Section 7.3 The Fundamental Spaces of a Matrix

THE FUNDAMENTAL SPACES OF A MATRIX

If A is an $m \times n$ matrix, then there are three important spaces associated with A:

1. The ***row space*** of A, denoted by $\text{row}(A)$, is the subspace of R^n that is spanned by the row vectors of A.
2. The ***column space*** of A, denoted by $\text{col}(A)$, is the subspace of R^m that is spanned by the column vectors of A.
3. The ***null space*** of A, denoted by $\text{null}(A)$, is the solution space of $A\mathbf{x} = \mathbf{0}$. This is a subspace of R^n.

→ row(A) = lin(A) ; col(A) = col(A) ; null(A) = nuc(A) .

If we consider A and A^T together, then there appear to be six such subspaces:

$$\text{row}(A), \quad \text{row}(A^T), \quad \text{col}(A), \quad \text{col}(A^T), \quad \text{null}(A), \quad \text{null}(A^T)$$

But transposing a matrix converts rows to columns, and columns to rows, so $\text{row}(A^T) = \text{col}(A)$ and $\text{col}(A^T) = \text{row}(A)$. Thus, of the six subspaces only the following four are distinct:

$\text{row}(A)$	$\text{null}(A)$
$\text{col}(A)$	$\text{null}(A^T)$

These are called the ***fundamental spaces*** of A. The dimensions of $\text{row}(A)$ and $\text{null}(A)$ are sufficiently important that there is some terminology associated with them.

Definition 7.3.1 The dimension of the row space of a matrix A is called the ***rank*** of A and is denoted by $\text{rank}(A)$; and the dimension of the null space of A is called the ***nullity*** of A and is denoted by $\text{nullity}(A)$.

→ rank(A) = pos(A) ; nullity(A) = nul(A) .

ORTHOGONAL COMPLEMENTS

One of the goals in this section is to develop some of the basic properties of the fundamental spaces. As a first step, we will need to establish some more results about orthogonality.

Recall from Section 3.5 that if $\mathbf{a}$ is a nonzero vector in R^n, then $\mathbf{a}^{\perp}$ is the set of all vectors in R^n that are orthogonal to $\mathbf{a}$. We call this set the *orthogonal complement* of $\mathbf{a}$ (or the *hyperplane through the origin with normal* $\mathbf{a}$). The following definition extends the idea of an orthogonal complement to sets with more than one vector.

Definition 7.3.2 If S is a nonempty set in R^n, then the ***orthogonal complement*** of S, denoted by $S^{\perp}$, is defined to be the set of all vectors in R^n that are orthogonal to every vector in S.

EXAMPLE 1 Orthogonal Complements of Subspaces of R^3

If L is a line through the origin of R^3, then $L^\perp$ is the plane through the origin that is perpendicular to L, and if W is a plane through the origin of R^3, then $W^\perp$ is the line through the origin that is perpendicular to W (Figure 7.3.1). ■

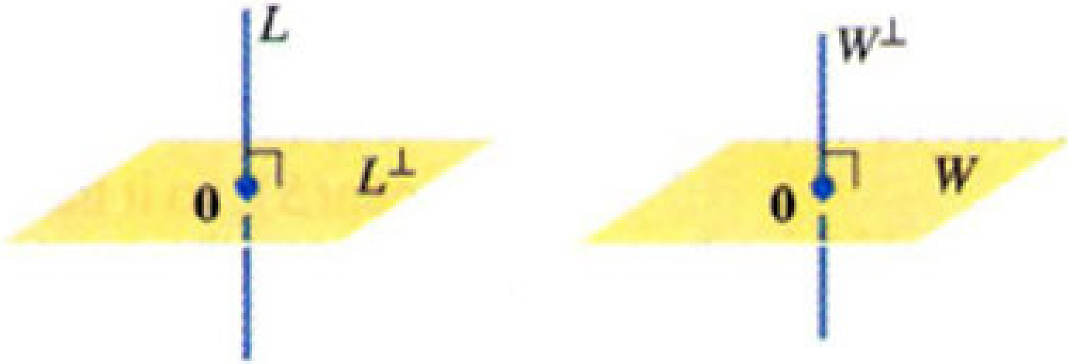

Figure 7.3.1

EXAMPLE 2 Orthogonal Complement of Row Vectors

If S is the set of row vectors of an $m \times n$ matrix A, then it follows from Theorem 3.5.6 that $S^\perp$ is the solution space of $A\mathbf{x} = \mathbf{0}$. ■

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 7.3.3 *If S is a nonempty set in R^n, then $S^\perp$ is a subspace of R^n.*

Proof The set $S^\perp$ contains the vector $\mathbf{0}$, so we can be assured that it is nonempty. We will show that it is closed under scalar multiplication and addition. For this purpose, let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $S^\perp$ and let c be a scalar.

To show that $c\mathbf{v}$ and $\mathbf{u}+\mathbf{v}$ are vectors in $S^\perp$, we must show that $c\mathbf{v}\cdot\mathbf{x}=0$ and $(\mathbf{u}+\mathbf{v})\cdot\mathbf{x}=0$ for every vector $\mathbf{x}$ in S. But $\mathbf{u}$ and $\mathbf{v}$ are vectors in $S^\perp$, so $\mathbf{u}\cdot\mathbf{x}=0$ and $\mathbf{v}\cdot\mathbf{x}=0$. Thus, using properties of the dot product we obtain

$$c\mathbf{v}\cdot\mathbf{x}=c(\mathbf{v}\cdot\mathbf{x})=c(0)=0 \quad \text{and} \quad (\mathbf{u}+\mathbf{v})\cdot\mathbf{x}=(\mathbf{u}\cdot\mathbf{x})+(\mathbf{v}\cdot\mathbf{x})=0+0=0$$

which completes the proof. ■

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]**

EXAMPLE 3

Orthogonal Complement of Two Vectors

Find the orthogonal complement in an xyz-coordinate system of the set $S = \{\mathbf{v}_1, \mathbf{v}_2\}$, where

$$\mathbf{v}_1 = (1, -2, 1) \quad \text{and} \quad \mathbf{v}_2 = (3, -7, 5)$$

Solution It should be evident geometrically that $S^\perp$ is the line through the origin that is perpendicular to the plane determined by $\mathbf{v}_1$ and $\mathbf{v}_2$. One way to find this line is to use the result in Example 2 by letting A be the matrix with row vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ and solving the system $A\mathbf{x} = \mathbf{0}$. This system is

$$\begin{bmatrix} 1 & -2 & 1 \\ 3 & -7 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

'Corrigir no livro texto, pois lá está $[0\ 0\ 0]^T$.'

and a general solution is (verify)

$$x = 3t, \quad y = 2t, \quad z = t \tag{1}$$

Thus, $S^\perp$ is the line through the origin that is parallel to the vector $\mathbf{w} = (3, 2, 1)$.

EXAMPLE 3

Orthogonal Complement of Two Vectors

Alternative Solution A vector that is orthogonal to both $\mathbf{v}_1$ and $\mathbf{v}_2$ is

$$\mathbf{v}_1 \times \mathbf{v}_2 = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & -2 & 1 \\ 3 & -7 & 5 \end{vmatrix} = -3\mathbf{i} - 2\mathbf{j} - \mathbf{k} = (-3, -2, -1)$$

The vector $\mathbf{w} = -(\mathbf{v}_1 \times \mathbf{v}_2) = (3, 2, 1)$ is also orthogonal to $\mathbf{v}_1$ and $\mathbf{v}_2$ and is actually a simpler choice because it has fewer minus signs. This is the vector $\mathbf{w}$ that we obtained in our first solution, so again we find that the orthogonal complement is the line through the origin given parametrically by (1). ■

PROPERTIES OF ORTHOGONAL COMPLEMENTS

The following theorem lists three basic facts about orthogonal complements.

Theorem 7.3.4

(*a*) *If* W *is a subspace of* R^n, *then* $W^\perp \cap W = \{\mathbf{0}\}$.

(*b*) *If* S *is a nonempty subset of* R^n, *then* $S^\perp = \text{span}(S)^\perp$.

(*c*) *If* W *is a subspace of* R^n, *then* $(W^\perp)^\perp = W$.

We will prove parts (*a*) and (*b*); the proof of part (*c*) will be deferred until a later section in which we will have more mathematical tools to use.

Proof (a) The set $W^\perp \cap W$ contains at least the zero vector, since W and $W^\perp$ are subspaces of R^n. But this is the only vector in $W^\perp \cap W$, for if $\mathbf{v}$ is any vector in this intersection, then $\mathbf{v} \cdot \mathbf{v} = 0$, which implies that $\mathbf{v} = 0$ by part (*d*) of Theorem 1.2.6.

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]**

Proof (b) We must show that every vector in $\text{span}(S)^\perp$ is in $S^\perp$ and conversely. Accordingly, let $\mathbf{v}$ be any vector in $\text{span}(S)^\perp$. This vector is orthogonal to every vector in $\text{span}(S)$, so it has to be orthogonal to every vector in S, since S is contained in $\text{span}(S)$. Thus, $\mathbf{v}$ is in $S^\perp$.

Conversely, let $\mathbf{v}$ be any vector in $S^\perp$. To show that $\mathbf{v}$ is in $\text{span}(S)^\perp$, we must show that $\mathbf{v}$ is orthogonal to every linear combination of the vectors in S. Accordingly, let

$$\mathbf{w} = t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_k\mathbf{v}_k$$

be any such linear combination. Then using properties of the dot product we obtain

$$\mathbf{v} \cdot \mathbf{w} = \mathbf{v} \cdot (t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_k\mathbf{v}_k) = t_1(\mathbf{v} \cdot \mathbf{v}_1) + t_2(\mathbf{v} \cdot \mathbf{v}_2) + \cdots + t_k(\mathbf{v} \cdot \mathbf{v}_k)$$

But each dot product on the right side is zero, since $\mathbf{v}$ is orthogonal to every vector in S. Thus, $\mathbf{v} \cdot \mathbf{w} = 0$, which shows that $\mathbf{v}$ is orthogonal to every vector in $\text{span}(S)$. ■

In words, part (*b*) of Theorem 7.3.4 states that *the orthogonal complement of a nonempty set and the orthogonal complement of the subspace spanned by that set are the same*. Thus, for example, the orthogonal complement of the set of row vectors of a matrix is the same as the orthogonal complement of the row space of that matrix. Thus, we have established the following:

Theorem 7.3.5 *If A is an $m \times n$ matrix, then the row space of A and the null space of A are orthogonal complements.*

Theorem 7.3.6 *If A is an $m \times n$ matrix, then the column space of A and the null space of A^T are orthogonal complements.*

The results in these two theorems are captured by the formulas

$$\begin{aligned} \text{row}(A)^\perp &= \text{null}(A), \qquad & \text{null}(A)^\perp &= \text{row}(A) \\ \text{col}(A)^\perp &= \text{null}(A^T), \qquad & \text{null}(A^T)^\perp &= \text{col}(A) \end{aligned} \tag{2}$$

Theorem 7.3.4

(*a*) *If W is a subspace of R^n, then $W^\perp \cap W = \{\mathbf{0}\}$.*

→ (*b*) *If S is a nonempty subset of R^n, then $S^\perp = \text{span}(S)^\perp$.*

(*c*) *If W is a subspace of R^n, then $(W^\perp)^\perp = W$.*

The following theorem provides an important computational tool for studying relationships between the fundamental spaces of a matrix. The first two statements in the theorem go to the heart of Gauss–Jordan elimination and Gaussian elimination and were simply accepted to be true when we developed those methods.

Theorem 7.3.7

(*a*) *Elementary row operations do not change the row space of a matrix.*

(*b*) *Elementary row operations do not change the null space of a matrix.*

(*c*) *The nonzero row vectors in any row echelon form of a matrix form a basis for the row space of the matrix.*

We will prove parts (*a*) and (*b*) and give an informal argument for part (*c*).

Proof (a) Observe first that when you multiply a row of a matrix A by a nonzero scalar or when you add a scalar multiple of one row to another, you are computing a linear combination of row vectors of A. Thus, if B is obtained from A by a succession of elementary row operations, then every vector in $\text{row}(B)$ must be in $\text{row}(A)$.

However, if B is obtained from A by elementary row operations, then A can be obtained from B by performing the inverse operations in reverse order. Thus, every vector in $\text{row}(A)$ must also be in $\text{row}(B)$, from which we conclude that $\text{row}(A) = \text{row}(B)$.

Proof (b) By part (*a*), performing an elementary row operation on a matrix does not change the row space of the matrix and hence does not change the orthogonal complement of the row space. But the orthogonal complement of the row space of A is the null space of A (Theorem 7.3.5), so performing an elementary row operation on a matrix A does not change the null space of A.

Proof (c) The nonzero row vectors in a row echelon form of a matrix A form a basis for the row space of A because they span the row space by part (*a*) of this theorem, and they are linearly independent by Example 4 of Section 7.1. ■

Theorem 7.3.8 *If A and B are matrices with the same number of columns, then the following statements are equivalent.*

(a) A and B have the same row space.

(b) A and B have the same null space.

(c) The row vectors of A are linear combinations of the row vectors of B, and conversely.

We will prove the equivalence of parts (a) and (b) and leave the proof of equivalence $(a) \Leftrightarrow (c)$ as an exercise. The equivalence $(b) \Leftrightarrow (c)$ will then follow as a logical consequence.

***Proof* (a) ⇔ (b)** The row space and null space of a matrix are orthogonal complements of one another. Thus, if A and B have the same row space, then they must have the same null space; and conversely. ■

FINDING BASES BY ROW REDUCTION

We now turn to the problem of finding a basis for the subspace W of R^n that is spanned by a given set of vectors

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s\}$$

There are two variations of this problem, each requiring different methods:

1. If *any* basis for W will suffice for the problem at hand, then we can start by forming a matrix A that has $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s$ as row vectors. This makes W the row space of A, so a basis can be found by reducing A to row echelon form and extracting the nonzero rows.
2. If the basis must consist of vectors from the *original set* S, then the preceding method is not appropriate because elementary row operations usually alter row vectors. A method for solving this kind of basis problem will be discussed later.

EXAMPLE 4 Finding a Basis by Row Reduction

(a) Find a basis for the subspace W of R^5 that is spanned by the vectors

$$\mathbf{v}_1 = (1, 0, 0, 0, 2), \qquad \mathbf{v}_2 = (-2, 1, -3, -2, -4)$$
$$\mathbf{v}_3 = (0, 5, -14, -9, 0), \quad \mathbf{v}_4 = (2, 10, -28, -18, 4)$$

(b) Find a basis for $W^{\perp}$.

Solution (a) The subspace spanned by the given vectors is the row space of the matrix

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

Reducing this matrix to row echelon form yields

$$U = \begin{bmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & -3 & -2 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{4}$$

Extracting the nonzero rows yields the basis vectors

$$\mathbf{w}_1 = (1, 0, 0, 0, 2), \quad \mathbf{w}_2 = (0, 1, -3, -2, 0), \quad \mathbf{w}_3 = (0, 0, 1, 1, 0)$$

which we have expressed in comma-delimited notation for consistency with the form of the original vectors. Since there are three basis vectors, we have shown that $\dim(W) = 3$.

Alternatively, we can take the matrix A all the way to the reduced row echelon form

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{5}$$

which yields the basis vectors

$$\mathbf{w}'_1 = (1, 0, 0, 0, 2), \quad \mathbf{w}'_2 = (0, 1, 0, 1, 0), \quad \mathbf{w}'_3 = (0, 0, 1, 1, 0) \tag{6}$$

Although it is extra work to obtain the reduced row echelon form, the resulting basis vectors are usually simpler in that they have more zeros. Whether it justifies the extra work depends on the purpose for which the basis will be used.

Solution (b) It follows from Theorem 7.3.5 that $W^{\perp}$ is the null space of A, so our problem reduces to finding a basis for the solution space of the homogeneous system $A\mathbf{x} = \mathbf{0}$. We will use the canonical basis produced by Gauss–Jordan elimination. Most of the work has already been done, since R in (5) is the reduced row echelon form of A. We leave it for you to use R to obtain the general solution

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = s \begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 0 \\ -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} \tag{7}$$

Thus, the vectors

$$\mathbf{u}_1 = (-2, 0, 0, 0, 1) \quad \text{and} \quad \mathbf{u}_2 = (0, -1, -1, 1, 0)$$

form a basis for $W^{\perp}$. As a check, we leave it for you to confirm that $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthogonal to all of the basis vectors for W that were obtained in part (a). ■

To be in ***reduced row echelon form*** a matrix must have the following properties:

1. If a row does not consist entirely of zeros, then the first nonzero number in the row is a 1. We call this a ***leading* 1**.
2. If there are any rows that consist entirely of zeros, then they are grouped together at the bottom of the matrix.
3. In any two successive rows that do not consist entirely of zeros, the leading 1 in the lower row occurs farther to the right than the leading 1 in the higher row.
4. Each column that contains a leading 1 has zeros everywhere else.

A matrix that has the first three properties is said to be in ***row echelon form***. (Thus, a matrix in reduced row echelon form is of necessity in row echelon form, but not conversely.)

DETERMINING WHETHER A VECTOR IS IN A GIVEN SUBSPACE

Consider the following three problems:

Problem 1. Given a set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ in R^m, find conditions on the numbers $b_1, b_2, \ldots, b_m$ under which $\mathbf{b} = (b_1, b_2, \ldots, b_m)$ will lie in span(S).

Problem 2. Given an $m \times n$ matrix A, find conditions on the numbers $b_1, b_2, \ldots, b_m$ under which $\mathbf{b} = (b_1, b_2, \ldots, b_m)$ will lie in col(A).

Problem 3. Given a linear transformation $T : R^n \to R^m$, find conditions on the numbers $b_1, b_2, \ldots, b_m$ under which $\mathbf{b} = (b_1, b_2, \ldots, b_m)$ will lie in ran(T).

Although these problems look different at the surface, they are just different formulations of the same problem (why?). The following example illustrates three ways of attacking the first formulation of the problem.

→ From (1) to (2), the vectors in S should be the columns of A;

from (2) to (3), A represents the linear transformation T.

EXAMPLE 6 Conditions for a Vector to Lie in a Given Subspace

What conditions must a vector $\mathbf{b} = (b_1, b_2, b_3, b_4, b_5)$ satisfy to lie in the subspace of R^5 spanned by the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{v}_4$ in Example 4?

Solution 1 The most direct way to solve this problem is to look for conditions under which the vector equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 + x_4\mathbf{v}_4 = \mathbf{b} \tag{8}$$

has a solution for x_1, x_2, x_3, and x_4. This is the vector form of the linear system $C\mathbf{x} = \mathbf{b}$ in which $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{v}_4$ are the successive column vectors of C. The augmented matrix for this system is

$$\left[\begin{array}{rrrr|r} 1 & -2 & 0 & 2 & b_1 \\ 0 & 1 & 5 & 10 & b_2 \\ 0 & -3 & -14 & -28 & b_3 \\ 0 & -2 & -9 & -18 & b_4 \\ 2 & -4 & 0 & 4 & b_5 \end{array}\right] \tag{9}$$

and our goal is to determine conditions on the b's under which this system is consistent. This is what we referred to as a "consistency problem" (see 3.3.10). As in Example 8 of that section, we reduce the left side of (9) to row echelon form, which yields (verify)

$$\left[\begin{array}{rrrr|c} 1 & -2 & 0 & 2 & b_1 \\ 0 & 1 & 5 & 10 & b_2 \\ 0 & 0 & 1 & 2 & b_3 + 3b_2 \\ 0 & 0 & 0 & 0 & b_4 - b_3 - b_2 \\ 0 & 0 & 0 & 0 & b_5 - 2b_1 \end{array}\right]$$

Thus, for the system to be consistent the components of **b** must satisfy the two conditions

$$\begin{aligned} b_4 - \quad b_3 - b_2 &= 0 \\ b_5 - 2b_1 \qquad\quad &= 0 \end{aligned} \quad \text{or} \quad \begin{aligned} b_4 &= b_2 + b_3 \\ b_5 &= \quad 2b_1 \end{aligned}$$

For example, the vector $(7, -2, 5, 3, 14)$ lies in W, but $(7, -2, 5, 3, 6)$ and $(0, -1, 3, -2, 0)$ do not (verify).

Solution 2 Here is a way to attack the same problem by focusing on rows rather than columns. It follows from Theorem 7.2.5 that $\mathbf{b}$ will lie in $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ if and only if this space has the same dimension as $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{b}\}$, that is, if and only if the matrix A with row vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{v}_4$ has the same rank as the matrix that results when $\mathbf{b}$ is adjoined to A as an additional row vector. The matrix A is given in (3), so that adjoining $\mathbf{b}$ as an additional row vector yields

$$\left[\begin{array}{rrrrr} 1 & 0 & 0 & 0 & 2 \\ -2 & 1 & -3 & -2 & -4 \\ 0 & 5 & -14 & -9 & 0 \\ 2 & 10 & -28 & -18 & 4 \\ \hdashline b_1 & b_2 & b_3 & b_4 & b_5 \end{array}\right] \tag{10}$$

To determine conditions under which (3) and (10) have the same rank, we will begin by reducing the A portion of (10) to reduced row echelon form to reveal a basis for the row space of A (a row echelon form will also work). This yields

$$\left[\begin{array}{ccccc} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hdashline b_1 & b_2 & b_3 & b_4 & b_5 \end{array}\right] \tag{11}$$

For this matrix to have the same rank as A (rank 3) it would have to be possible to make the last row zero using elementary row operations. Accordingly, we will now start "zeroing out" those entries in the last row that lie in the pivot columns of A by adding suitable multiples of the pivot rows of A to the bottom row. We leave it for you to verify that this yields

$$\left[\begin{array}{ccccc} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & b_4 - b_3 - b_2 & b_5 - 2b_1 \end{array}\right] \tag{12}$$

Thus, for (11) to have rank 3 we must have $b_4 - b_3 - b_2 = 0$ and $b_5 - 2b_1 = 0$. These are the same conditions that we obtained by the first method.

Solution 3 Here is a third way to attack the same problem. To say that $\mathbf{b} = (b_1, b_2, b_3, b_4, b_5)$ lies in the subspace W spanned by the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{v}_4$ is the same as saying that $\mathbf{b}$ is orthogonal to every vector in $W^{\perp}$. But we showed in part (b) of Example 4 that the vectors

$$\mathbf{u}_1 = (-2, 0, 0, 0, 1) \quad \text{and} \quad \mathbf{u}_2 = (0, -1, -1, 1, 0)$$

form a basis for $W^{\perp}$. Thus, $\mathbf{b}$ will be orthogonal to every vector in $W^{\perp}$ if and only if it is orthogonal to $\mathbf{u}_1$ and $\mathbf{u}_2$. If we write the orthogonality conditions $\mathbf{u}_1 \cdot \mathbf{b} = 0$ and $\mathbf{u}_2 \cdot \mathbf{b} = 0$ in component form, we obtain

$$-2b_1 + b_5 = 0 \quad \text{and} \quad -b_2 - b_3 + b_4 = 0$$

which are the same conditions we obtained by the first two methods. ■

EXAMPLE 7 A Useful Algorithm

Determine which of the three vectors $\mathbf{b}_1 = (7, -2, 5, 3, 14)$, $\mathbf{b}_2 = (7, -2, 5, 3, 6)$, and $\mathbf{b}_3 = (0, -1, 3, -2, 0)$, if any, lie in the subspace of R^5 spanned by the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{v}_4$ in Example 4.

Solution One way to solve this problem is to consider the matrix C that has $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{v}_4$ as its successive column vectors, and determine which of the $\mathbf{b}$'s, if any, lie in the column space of C by investigating whether the systems $C\mathbf{x} = \mathbf{b}_1$, $C\mathbf{x} = \mathbf{b}_2$, and $C\mathbf{x} = \mathbf{b}_3$ are consistent. An efficient way of doing this was presented in Example 7 of Section 3.3. As in that example, we adjoin the column vectors $\mathbf{b}_1$, $\mathbf{b}_2$, or $\mathbf{b}_3$ to C and consider the partitioned matrix

$$\left[\begin{array}{cccc|c|c|c} 1 & -2 & 0 & 2 & 7 & 7 & 0 \\ 0 & 1 & 5 & 10 & -2 & -2 & -1 \\ 0 & -3 & -14 & -28 & 5 & 5 & 3 \\ 0 & -2 & -9 & -18 & 3 & 3 & -2 \\ 2 & -4 & 0 & 4 & 14 & 6 & 0 \end{array}\right]$$

If we now apply row operations to this until the submatrix C is in row echelon form, we obtain (verify)

$$\left[\begin{array}{cccc|c|c|c} 1 & -2 & 0 & 2 & 7 & 7 & 0 \\ 0 & 1 & 5 & 10 & -2 & -2 & -1 \\ 0 & 0 & 1 & 2 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 & 0 & -8 & 0 \end{array}\right]$$

At this point we can see that the system $C\mathbf{x} = \mathbf{b}_1$ is consistent but the systems $C\mathbf{x} = \mathbf{b}_2$ and $C\mathbf{x} = \mathbf{b}_3$ are not. Thus, vector $\mathbf{b}_1 = (7, -2, 5, 3, 14)$ lies in span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ but vectors $\mathbf{b}_2 = (7, -2, 5, 3, 6)$ and $\mathbf{b}_3 = (0, -1, 3, -2, 0)$ do not. This is consistent with the observation made at the end of the first solution in Example 6. ■

Section 7.4 The Dimension Theorem and Its Implications

THE DIMENSION THEOREM FOR MATRICES

Recall from Theorem 2.2.2 that if $A\mathbf{x} = \mathbf{0}$ is a homogeneous linear system with n unknowns, and if the reduced row echelon form of the augmented matrix has r nonzero rows, then the system has $n - r$ free variables; we called this the *dimension theorem for homogeneous linear systems*.

However, for a homogeneous system, the augmented matrix and the coefficient matrix have the same number of zero rows in their reduced row echelon forms (the number being the rank of A), so we can restate the dimension theorem for homogeneous linear systems as

$$\text{number of free variables} = n - \text{rank}(A)$$

nonzero

or, alternatively, as

$$\text{rank}(A) + \text{number of free variables} = \text{number of columns of } A \tag{1}$$

Theorem 2.2.2 ***(Dimension Theorem for Homogeneous Systems)*** *If a homogeneous linear system has n unknowns, and if the reduced row echelon form of its augmented matrix has r nonzero rows, then the system has $n - r$ free variables.*

But each free variable produces a parameter in a general solution of the system $A\mathbf{x} = \mathbf{0}$, so the number of free variables is the same as the dimension of the solution space of the system (which is the nullity of A). Thus, we can rewrite (1) as

$$\text{rank}(A) + \text{nullity}(A) = \text{number of columns of } A$$

and hence we have established the following matrix version of the dimension theorem.

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

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

EXAMPLE 1

The Dimension Theorem for Matrices

In Example 4 of the last section we saw that the reduced row echelon form of the matrix

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

has three nonzero rows [Formula (5) of Section 7.3], and we saw that the null space of A has dimension 2 [Formula (7) of Section 7.3]. Thus,

$$\text{rank}(A) + \text{nullity}(A) = 3 + 2 = 5$$

which is consistent with Formula (2) and the fact that A has five columns. ■

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (5) \qquad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = s\begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} + t\begin{bmatrix} 0 \\ -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} \quad (7)$$

EXTENDING A LINEARLY INDEPENDENT SET TO A BASIS

It follows from part (*b*) of Theorem 7.2.2 that every linearly independent set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ in R^n can be enlarged to a basis for R^n by adding appropriate linearly independent vectors to it. One way to find such vectors is to form the matrix A that has $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ as row vectors, thereby making the subspace spanned by these vectors into the row space of A.

By solving the homogeneous linear system $A\mathbf{x} = \mathbf{0}$, we can find a basis for the null space of A. This basis has $n - k$ vectors, say $\mathbf{w}_{k+1}, \ldots, \mathbf{w}_n$, by the dimension theorem for matrices, and each of the $\mathbf{w}$'s is orthogonal to all of the $\mathbf{v}$'s, since null(A) and row(A) are orthogonal.

This orthogonality implies that the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{w}_{k+1}, \ldots, \mathbf{w}_n\}$ is linearly independent (Exercise P4) and hence forms a basis for R^n.

EXAMPLE 2

Extending a Linearly Independent Set to a Basis

The vectors

$$\mathbf{v}_1 = (1, 3, -1, 1) \quad \text{and} \quad \mathbf{v}_2 = (0, 1, 1, 6)$$

are linearly independent, since neither vector is a scalar multiple of the other. Enlarge the set $\{\mathbf{v}_1, \mathbf{v}_2\}$ to a basis for R^4.

Solution We will find a basis for the null space of the matrix

$$A = \begin{bmatrix} 1 & 3 & -1 & 1 \\ 0 & 1 & 1 & 6 \end{bmatrix}$$

by solving the linear system $A\mathbf{x} = \mathbf{0}$. The reduced row echelon form of the augmented matrix for the system is

$$\left[\begin{array}{cccc|c} 1 & 0 & -4 & -17 & 0 \\ 0 & 1 & 1 & 6 & 0 \end{array}\right]$$

The reduced row echelon form of the augmented matrix for the system is

$$\left[\begin{array}{cccc|c} 1 & 0 & -4 & -17 & 0 \\ 0 & 1 & 1 & 6 & 0 \end{array}\right]$$

so a general solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4s + 17t \\ -s - 6t \\ s \\ t \end{bmatrix} = s \begin{bmatrix} 4 \\ -1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 17 \\ -6 \\ 0 \\ 1 \end{bmatrix}$$

Thus, the vectors

$$\mathbf{v}_1 = (1, 3, -1, 1), \quad \mathbf{v}_2 = (0, 1, 1, 6), \quad \mathbf{w}_3 = (4, -1, 1, 0), \quad \mathbf{w}_4 = (17, -6, 0, 1)$$

form a basis for R^4. ■

SOME CONSEQUENCES OF THE DIMENSION THEOREM FOR MATRICES

Theorem 7.4.2 *If an $m \times n$ matrix A has rank k, then:*

(*a*) *A has nullity $n - k$.*

(*b*) *Every row echelon form of A has k nonzero rows.*

(*c*) *Every row echelon form of A has $m - k$ zero rows.*

(*d*) *The homogeneous system $A\mathbf{x} = \mathbf{0}$ has k pivot variables (leading variables) and $n - k$ free variables.*

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 3 Consequences of the Dimension Theorem for Matrices

State some facts about a 5×7 matrix A with nullity 3.

Solution Here are some possibilities:

- $\text{rank}(A) = 7 - 3 = 4$.
- Every row echelon form of A has $5 - 4 = 1$ zero row.
- The homogeneous system $A\mathbf{x} = \mathbf{0}$ has 4 pivot variables and $7 - 4 = 3$ free variables. ■

EXAMPLE 4 Restrictions Imposed by the Dimension Theorem for Matrices

Can a 5×7 matrix A have a one-dimensional null space?

Solution No. Otherwise, the rank of A would be

$$\text{rank}(A) = 7 - \text{nullity}(A) = 7 - 1 = 6$$

which is impossible, since the five row vectors of A cannot span a six-dimensional space. ■

THE DIMENSION THEOREM FOR SUBSPACES

Theorem 7.4.3 ***(The Dimension Theorem for Subspaces)*** *If W is a subspace of R^n, then*

$$\dim(W) + \dim(W^\perp) = n \tag{3}$$

Proof If $W = \{\mathbf{0}\}$, then $W^\perp = R^n$, in which case $\dim(W) + \dim(W^\perp) = 0 + n = n$. If $W \neq \{\mathbf{0}\}$, then choose a basis for W and form a matrix A that has these basis vectors as row vectors. The matrix A has n columns since its row vectors come from R^n. Moreover, the row space of A is W and the null space of A is $W^\perp$, so it follows from Theorem 7.4.1 that

$$\dim(W) + \dim(W^\perp) = \text{rank}(A) + \text{nullity}(A) = n$$ ■

EXAMPLE 5 The Dimension Theorem for Subspaces

In Example 4 of the last section we considered a subspace W spanned by four vectors, $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{v}_4$ in R^5. In part (a) of that example we found a basis for W with three vectors, and in part (b) we found a basis for $W^\perp$ with two vectors. Thus,

$$\dim(W) + \dim(W^\perp) = 3 + 2 = 5$$

which is consistent with Formula (3) and the fact that R^5 is five-dimensional. ■

Theorem 7.4.5 *If W is a subspace of R^n with dimension $n-1$, then there is a nonzero vector $\mathbf{a}$ for which $W = \mathbf{a}^\perp$; that is, W is a hyperplane through the origin of R^n.*

Proof Let W be a subspace of R^n with dimension $n-1$. It follows from the dimension theorem for subspaces that $\dim(W^\perp) = 1$; and this implies that $W^\perp$ is the span of some nonzero vector in R^n, say $W^\perp = \text{span}\{\mathbf{a}\}$. Thus, it follows from parts (*b*) and (*c*) of Theorem 7.3.4 that

$$W = (W^\perp)^\perp = (\text{span}\{\mathbf{a}\})^\perp = \mathbf{a}^\perp$$

which shows that W is the hyperplane through the origin of R^n with normal $\mathbf{a}$. ■

Theorem 7.4.3 ***(The Dimension Theorem for Subspaces)*** *If W is a subspace of R^n, then*

$$\dim(W) + \dim(W^\perp) = n \quad (3)$$

Since hyperplanes through the origin of R^n are the subspaces of dimension $n-1$, their orthogonal complements are the subspaces of dimension 1, which are the lines through the origin of R^n. Thus, we have the following geometric result.

Theorem 7.4.6 *The orthogonal complement of a hyperplane through the origin of R^n is a line through the origin of R^n, and the orthogonal complement of a line through the origin of R^n is a hyperplane through the origin of R^n. Specifically, if $\mathbf{a}$ is a nonzero vector in R^n, then the line $\text{span}\{\mathbf{a}\}$ and the hyperplane $\mathbf{a}^{\perp}$ are orthogonal complements of one another.*

RANK 1 MATRICES

Matrices of rank 1 will play an important role in our later work, so we will conclude this section by discussing some basic results about them. Here are some facts about an $m \times n$ matrix A that follow immediately from our previous work:

- If $\text{rank}(A) = 1$, then $\text{nullity}(A) = n - 1$, so the row space of A is a line through the origin of R^n and the null space is a hyperplane through the origin of R^n. Conversely, if the row space of A is a line through the origin of R^n, or, equivalently, if the null space of A is a hyperplane through the origin of R^n, then A has rank 1.

- If $\text{rank}(A) = 1$, then the row space of A is spanned by some nonzero vector $\mathbf{a}$, so all row vectors of A are scalar multiples of $\mathbf{a}$ and the null space of A is $\mathbf{a}^\perp$. Conversely, if the row vectors of A are all scalar multiples of some nonzero vector $\mathbf{a}$, then A has rank 1 and the null space of A is the hyperplane $\mathbf{a}^\perp$.

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

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

EXAMPLE 6

Some Rank 1 Matrices

The following matrices have rank 1 because in each case the row vectors are expressible as scalar multiples of any nonzero row vector:

$$\begin{bmatrix} 2 & -4 & -6 & 0 \\ -3 & 6 & 9 & 0 \end{bmatrix}, \quad \begin{bmatrix} 3 & -1 \\ -6 & 2 \\ -3 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 \\ -2 & -2 & -2 \\ 0 & 0 & 0 \end{bmatrix}$$

Observe that in each case the column vectors are also scalar multiples of a nonzero vector. This is because the row space and column space of a matrix always have the same dimension—a result that will be proved in the next section. ■

Rank 1 matrices arise when outer products of nonzero column vectors are computed. To see why this is so, suppose that

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

are nonzero, and recall from Definition 3.1.11 that the outer product of **u** with **v** is

$$\mathbf{u}\mathbf{v}^T = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix} [v_1 \quad v_2 \quad \cdots \quad v_n] = \begin{bmatrix} u_1 v_1 & u_1 v_2 & \cdots & u_1 v_n \\ u_2 v_1 & u_2 v_2 & \cdots & u_2 v_n \\ \vdots & \vdots & & \vdots \\ u_m v_1 & u_m v_2 & \cdots & u_m v_n \end{bmatrix} \tag{4}$$

This matrix has rank 1 since all row vectors are scalar multiples of the nonzero vector $\mathbf{v}^T$ and at least one of the components of **u** is nonzero.

Definition 3.1.11 If **u** and **v** are column vectors with the same size, then the product $\mathbf{u}^T\mathbf{v}$ is called the ***matrix inner product*** of **u** with **v**; and if **u** and **v** are column vectors of any size, then the product $\mathbf{u}\mathbf{v}^T$ is called the ***matrix outer product*** of **u** with **v**.

EXAMPLE 7 A Rank 1 Matrix Arising from a Product $\mathbf{u}\mathbf{v}^T$

Let

$$\mathbf{u} = \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} 1 \\ 3 \\ -2 \\ -1 \end{bmatrix}$$

Then

$$\mathbf{u}\mathbf{v}^T = \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix} [1 \quad 3 \quad -2 \quad -1] = \begin{bmatrix} -2 & -6 & 4 & 2 \\ 1 & 3 & -2 & -1 \\ 4 & 12 & -8 & -4 \end{bmatrix}$$

which is a matrix of rank 1. ■

We saw in (4) that the outer product of nonzero column vectors has rank 1. The following theorem shows that all rank 1 matrices arise from outer products.

Theorem 7.4.7 *If* $\mathbf{u}$ *is a nonzero* $m \times 1$ *matrix and* $\mathbf{v}$ *is a nonzero* $n \times 1$ *matrix, then the outer product*

$$A = \mathbf{u}\mathbf{v}^T$$

has rank 1. Conversely, if A is an $m \times n$ *matrix with rank 1, then A can be factored into a product of the above form.*

Proof Only the converse remains to be proved. Accordingly, let A be any $m \times n$ matrix of rank 1. The row vectors of A are all scalar multiples of some nonzero row vector $\mathbf{v}^T$, so we can express A in the form

$$A = \begin{bmatrix} u_1\mathbf{v}^T \\ u_2\mathbf{v}^T \\ \vdots \\ u_m\mathbf{v}^T \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix} \mathbf{v}^T = \mathbf{u}\mathbf{v}^T$$

where $\mathbf{u}$ is the column vector with components $u_1, u_2, \ldots, u_m$. These components cannot all be zero, otherwise A would have rank 0. ■

The proof of theorem 7.4.7 suggests a method for factoring a rank 1 matrix into a product $\mathbf{u}\mathbf{v}^T$ of a column vector times a row vector—take $\mathbf{v}^T$ to be any nonzero row of A and take the entries of the column vector $\mathbf{u}$ to be the scalars that produce the successive rows of A from $\mathbf{v}^T$. Here is an example.

EXAMPLE 8

Factoring a Rank 1 Matrix into the Form $\mathbf{u}\mathbf{v}^T$

We will factor the first matrix in Example 6, taking $\mathbf{v}^T$ to be the first row. This yields

$$\begin{bmatrix} 2 & -4 & -6 & 0 \\ -3 & 6 & 9 & 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -\frac{3}{2} \end{bmatrix} [2 \quad -4 \quad -6 \quad 0] = \mathbf{u}\mathbf{v}^T$$

■

SYMMETRIC RANK 1 MATRICES

If $\mathbf{u}$ is a nonzero column vector, then

$$\mathbf{u}\mathbf{u}^T = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} [u_1 \quad u_2 \quad \cdots \quad u_n] = \begin{bmatrix} u_1^2 & u_1u_2 & \cdots & u_1u_n \\ u_2u_1 & u_2^2 & \cdots & u_2u_n \\ \vdots & \vdots & & \vdots \\ u_nu_1 & u_nu_2 & \cdots & u_n^2 \end{bmatrix} \tag{5}$$

which, in addition to having rank 1, is symmetric. This is part of the following theorem whose proof is outlined in the exercises.

Theorem 7.4.8 *If $\mathbf{u}$ is a nonzero $n \times 1$ column vector, then the outer product $\mathbf{u}\mathbf{u}^T$ is a symmetric matrix of rank 1. Conversely, if A is a symmetric $n \times n$ matrix of rank 1, then it can be factored as $A = \mathbf{u}\mathbf{u}^T$ or else as $A = -\mathbf{u}\mathbf{u}^T$ for some nonzero $n \times 1$ column vector $\mathbf{u}$.*

EXAMPLE 9

A Symmetric Matrix of Rank One Arising from $\mathbf{u}\mathbf{u}^T$

If

$$\mathbf{u} = \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix}$$

then

$$\mathbf{u}\mathbf{u}^T = \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix} [-2 \quad 1 \quad 3] = \begin{bmatrix} 4 & -2 & -6 \\ -2 & 1 & 3 \\ -6 & 3 & 9 \end{bmatrix}$$

which we see directly is symmetric and has rank 1. ■