Section 7.7 The Projection Theorem and Its Implications

ORTHOGONAL PROJECTIONS ONTO LINES IN R^2

In Sections 6.1 and 6.2 we discussed orthogonal projections onto lines through the origin of R^2 and onto the coordinate planes of a rectangular coordinate system in R^3. In this section we will be concerned with the more general problem of defining and calculating orthogonal projections onto subspaces of R^n.

To motivate the appropriate definition, we will revisit orthogonal projections onto lines through the origin of R^2 from another point of view.

Recall from Formula (21) of Section 6.1 that the standard matrix P_θ for the orthogonal projection of R^2 onto the line through the origin making an angle θ with the positive x-axis of a rectangular xy-coordinate system can be expressed as

$$P_\theta = \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix} \quad (1)$$

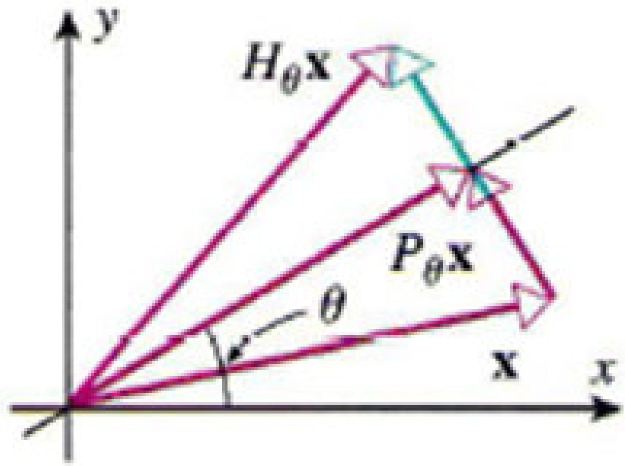


Figure 6.1.12

Now suppose that we are given a nonzero vector $\mathbf{a}$ in R^2, and let us consider how we might compute the orthogonal projection of a vector $\mathbf{x}$ onto the line $W = \text{span}\{\mathbf{a}\}$ without explicitly computing θ.

Figure 7.7.1 suggests that the vector $\mathbf{x}$ can be expressed as

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 \tag{2}$$

where $\mathbf{x}_1$ is the orthogonal projection of $\mathbf{x}$ onto W, and

$$\mathbf{x}_2 = \mathbf{x} - \mathbf{x}_1$$

is the orthogonal projection onto the line through the origin that is perpendicular to W.

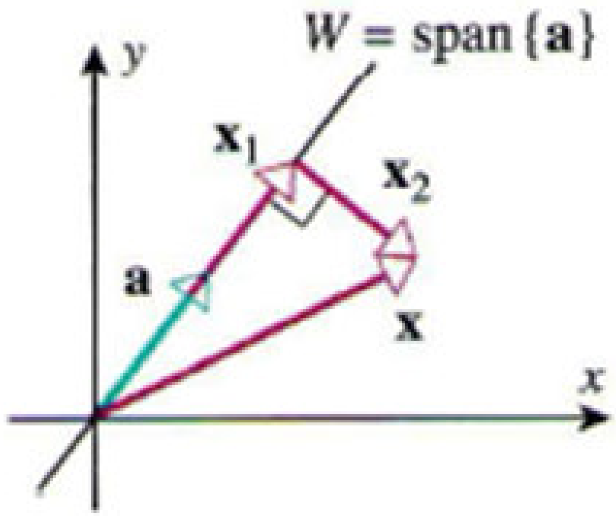


Figure 7.7.1

The vector $\mathbf{x}_1$ is some scalar multiple of $\mathbf{a}$, say

$$\mathbf{x}_1 = k\mathbf{a} \tag{3}$$

and the vector $\mathbf{x}_2 = \mathbf{x} - \mathbf{x}_1 = \mathbf{x} - k\mathbf{a}$ is orthogonal to $\mathbf{a}$, so we must have

$$(\mathbf{x} - k\mathbf{a}) \cdot \mathbf{a} = 0$$

which we can rewrite as

$$\mathbf{x} \cdot \mathbf{a} - k(\mathbf{a} \cdot \mathbf{a}) = 0$$

Solving for k and substituting in (3) yields

$$\mathbf{x}_1 = \frac{\mathbf{x} \cdot \mathbf{a}}{\mathbf{a} \cdot \mathbf{a}}\mathbf{a} = \frac{\mathbf{x} \cdot \mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} \tag{4}$$

which is a formula for the orthogonal projection of $\mathbf{x}$ onto the line span$\{\mathbf{a}\}$ in terms of $\mathbf{a}$ and $\mathbf{x}$. It is common to denote $\mathbf{x}_1$ by $\text{proj}_{\mathbf{a}}\mathbf{x}$ and to express Formula (4) as

$$\text{proj}_{\mathbf{a}}\mathbf{x} = \frac{\mathbf{x} \cdot \mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} \tag{5}$$

EXAMPLE 1 Orthogonal Projection onto a Line Through the Origin of R^2

Use Formula (5) to obtain the standard matrix P_θ for the orthogonal projection of R^2 onto the line W through the origin that makes an angle θ with the positive x-axis of a rectangular xy-coordinate system.

Solution The vector $\mathbf{u} = (\cos\theta, \sin\theta)$ is a unit vector along W (Figure 7.7.2), so if we use this vector as the $\mathbf{a}$ in Formula (5) and use the fact that $\|\mathbf{u}\| = 1$, then we obtain

$$\text{proj}_{\mathbf{u}}\mathbf{x} = (\mathbf{x} \cdot \mathbf{u})\mathbf{u}$$

In particular, the orthogonal projections of the standard unit vectors $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$ onto the line are

$$\text{proj}_{\mathbf{u}}\mathbf{e}_1 = (\mathbf{e}_1 \cdot \mathbf{u})\mathbf{u} = (\cos\theta)\mathbf{u} = (\cos^2\theta, \cos\theta\sin\theta) = (\cos^2\theta, \sin\theta\cos\theta)$$
$$\text{proj}_{\mathbf{u}}\mathbf{e}_2 = (\mathbf{e}_2 \cdot \mathbf{u})\mathbf{u} = (\sin\theta)\mathbf{u} = (\sin\theta\cos\theta, \sin^2\theta)$$

Expressing these vectors in column form yields the standard matrix

$$P_\theta = [\text{proj}_{\mathbf{u}}\mathbf{e}_1 \quad \text{proj}_{\mathbf{u}}\mathbf{e}_2] = \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix}$$

which is consistent with Formula (1). ■

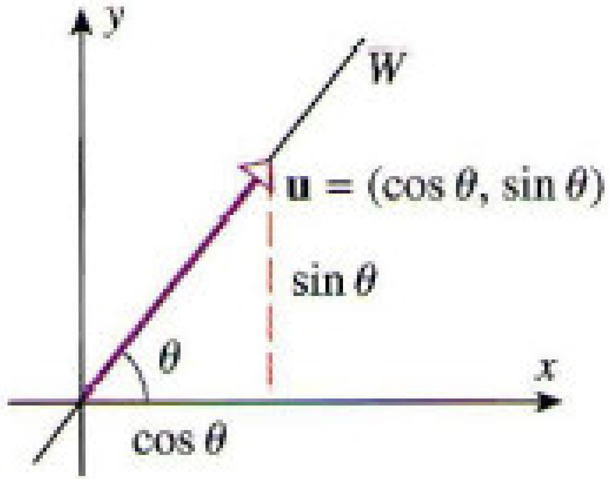


Figure 7.7.2

ORTHOGONAL PROJECTIONS ONTO LINES THROUGH THE ORIGIN OF R^n

The following theorem, which extends Formula (2) to R^n, is the foundation for defining orthogonal projections onto lines through the origin of R^n (Figure 7.7.3).

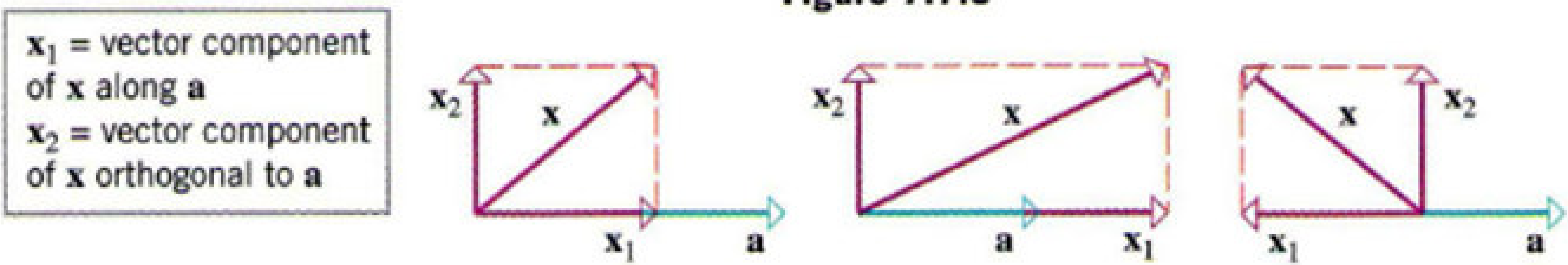
Theorem 7.7.1 *If* $\mathbf{a}$ *is a nonzero vector in* R^n*, then every vector* $\mathbf{x}$ *in* R^n *can be expressed in exactly one way as*

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 \tag{6}$$

where $\mathbf{x}_1$ *is a scalar multiple of* $\mathbf{a}$ *and* $\mathbf{x}_2$ *is orthogonal to* $\mathbf{a}$ *(and hence to* $\mathbf{x}_1$*). The vectors* $\mathbf{x}_1$ *and* $\mathbf{x}_2$ *are given by the formulas*

$$\mathbf{x}_1 = \frac{\mathbf{x}\cdot\mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} \quad \text{and} \quad \mathbf{x}_2 = \mathbf{x} - \mathbf{x}_1 = \mathbf{x} - \frac{\mathbf{x}\cdot\mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} \tag{7}$$

Figure 7.7.3



Proof There are two parts to the proof, an *existence* part and a *uniqueness* part.

The existence part is to show that there actually exist vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ that satisfy (6) such that $\mathbf{x}_1$ is a scalar multiple of $\mathbf{a}$ and $\mathbf{x}_2$ is orthogonal to $\mathbf{a}$;

and the uniqueness part is to show that if $\mathbf{x}_1'$ and $\mathbf{x}_2'$ is a second pair of vectors that satisfy these conditions, then $\mathbf{x}_1 = \mathbf{x}_1'$ and $\mathbf{x}_2 = \mathbf{x}_2'$.

For the existence part, we will show that the vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ in (7) satisfy the required conditions. It is obvious that $\mathbf{x}_1$ is a scalar multiple of $\mathbf{a}$ and that (6) holds, so let us focus on proving that $\mathbf{x}_2$ is orthogonal to $\mathbf{a}$, that is, $\mathbf{x}_2 \cdot \mathbf{a} = 0$. The computations are as follows:

$$\mathbf{x}_2 \cdot \mathbf{a} = (\mathbf{x} - \mathbf{x}_1) \cdot \mathbf{a} = \left(\mathbf{x} - \frac{\mathbf{x} \cdot \mathbf{a}}{\mathbf{a} \cdot \mathbf{a}}\mathbf{a}\right) \cdot \mathbf{a} = (\mathbf{x} \cdot \mathbf{a}) - \frac{\mathbf{x} \cdot \mathbf{a}}{\mathbf{a} \cdot \mathbf{a}}(\mathbf{a} \cdot \mathbf{a}) = 0$$

For the uniqueness part, suppose that $\mathbf{x}$ can also be written as $\mathbf{x} = \mathbf{x}_1' + \mathbf{x}_2'$, where $\mathbf{x}_1'$ is a scalar multiple of $\mathbf{a}$ and $\mathbf{x}_2'$ is orthogonal to $\mathbf{a}$. Then

$$\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{x}_1' + \mathbf{x}_2'$$

from which it follows that

$$\mathbf{x}_1 - \mathbf{x}_1' = \mathbf{x}_2' - \mathbf{x}_2 \tag{8}$$

Since $\mathbf{x}_1$ and $\mathbf{x}_1'$ are both scalar multiples of $\mathbf{a}$, so is their difference, and hence we can rewrite (8) in the form

$$k\mathbf{a} = \mathbf{x}_2' - \mathbf{x}_2 \tag{9}$$

Moreover, since $\mathbf{x}_2'$ and $\mathbf{x}_2$ are orthogonal to $\mathbf{a}$, so is their difference, and hence it follows from (9) that

$$k\mathbf{a} \cdot \mathbf{a} = k(\mathbf{a} \cdot \mathbf{a}) = 0 \tag{10}$$

From (9) to (10), take the inner product of both sides of (9) by **a**. As the right side is orthogonal to **a**, the result should be 0.

But $\mathbf{a} \cdot \mathbf{a} = \|\mathbf{a}\|^2 \neq 0$, since we assumed that $\mathbf{a} \neq \mathbf{0}$, so (10) implies that $k = 0$. Thus (9) implies that $\mathbf{x}_2 = \mathbf{x}_2'$ and (8), in turn, implies that $\mathbf{x}_1 = \mathbf{x}_1'$, which proves the uniqueness. ■

In the special case where $\mathbf{a}$ and $\mathbf{x}$ are vectors in R^2, the formula for $\mathbf{x}_1$ in (7) coincides with Formula (5) for the orthogonal projection of $\mathbf{x}$ onto span$\{\mathbf{a}\}$, and this suggests that we use Formula (7) to *define* orthogonal projections onto lines through the origin in R^n.

Definition 7.7.2 If $\mathbf{a}$ is a nonzero vector in R^n, and if $\mathbf{x}$ is any vector in R^n, then the ***orthogonal projection of*** $\mathbf{x}$ ***onto*** span$\{\mathbf{a}\}$ is denoted by $\text{proj}_{\mathbf{a}}\mathbf{x}$ and is defined as

$$\text{proj}_{\mathbf{a}}\mathbf{x} = \frac{\mathbf{x} \cdot \mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} \tag{11}$$

The vector $\text{proj}_{\mathbf{a}}\mathbf{x}$ is also called the ***vector component of*** $\mathbf{x}$ ***along*** $\mathbf{a}$, and $\mathbf{x} - \text{proj}_{\mathbf{a}}\mathbf{x}$ is called the ***vector component of*** $\mathbf{x}$ ***orthogonal to*** $\mathbf{a}$.

EXAMPLE 2 Calculating Vector Components

Let $\mathbf{x} = (2, -1, 3)$ and $\mathbf{a} = (4, -1, 2)$. Find the vector components of $\mathbf{x}$ along $\mathbf{a}$ and orthogonal to $\mathbf{a}$.

Solution Since

$$\mathbf{x} \cdot \mathbf{a} = (2)(4) + (-1)(-1) + (3)(2) = 15 \tag{12}$$

and

$$\|\mathbf{a}\|^2 = \mathbf{a} \cdot \mathbf{a} = 4^2 + (-1)^2 + 2^2 = 21 \tag{13}$$

it follows from (7) and (11) that the vector component of $\mathbf{x}$ along $\mathbf{a}$ is

$$\mathbf{x}_1 = \text{proj}_{\mathbf{a}}\mathbf{x} = \frac{\mathbf{x} \cdot \mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} = \tfrac{15}{21}(4, -1, 2) = \left(\tfrac{20}{7}, -\tfrac{5}{7}, \tfrac{10}{7}\right)$$

and that the vector component of $\mathbf{x}$ orthogonal to $\mathbf{a}$ is

$$\mathbf{x}_2 = \mathbf{x} - \text{proj}_{\mathbf{a}}\mathbf{x} = (2, -1, 3) - \left(\tfrac{20}{7}, -\tfrac{5}{7}, \tfrac{10}{7}\right) = \left(-\tfrac{6}{7}, -\tfrac{2}{7}, \tfrac{11}{7}\right)$$

As a check, you may want to confirm that $\mathbf{x}_1$ and $\mathbf{x}_2$ are orthogonal and that $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$. ■

Sometimes we will be interested in finding the *length* of $\text{proj}_{\mathbf{a}}\mathbf{x}$ but will not need the projection itself. A formula for this length can be derived from (11) by writing

$$\|\text{proj}_{\mathbf{a}}\mathbf{x}\| = \left\|\frac{\mathbf{x}\cdot\mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a}\right\| = \frac{|\mathbf{x}\cdot\mathbf{a}|}{\|\mathbf{a}\|^2}\|\mathbf{a}\|$$

which simplifies to

$$\|\text{proj}_{\mathbf{a}}\mathbf{x}\| = \frac{|\mathbf{x}\cdot\mathbf{a}|}{\|\mathbf{a}\|} \tag{14}$$

EXAMPLE 3 Computing the Length of an Orthogonal Projection

Use Formula (14) to compute the length of $\text{proj}_{\mathbf{a}}\mathbf{x}$ for the vectors **a** and **x** in Example 2.

Solution Using the results in (12) and (13), we obtain

$$\|\text{proj}_{\mathbf{a}}\mathbf{x}\| = \frac{|\mathbf{x}\cdot\mathbf{a}|}{\|\mathbf{a}\|} = \frac{|15|}{\sqrt{21}} = \frac{15}{\sqrt{21}}$$

We leave it for you to check this result by calculating directly from the vector $\text{proj}_{\mathbf{a}}\mathbf{x}$ obtained in Example 2. ■

PROJECTION OPERATORS ON R^n

Since the vector **x** in Definition 7.7.2 is arbitrary, we can use Formula (11) to define an operator $T : R^n \to R^n$ by

$$T(\mathbf{x}) = \text{proj}_{\mathbf{a}}\mathbf{x} = \frac{\mathbf{x} \cdot \mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} \tag{15}$$

This is called the ***orthogonal projection of*** R^n ***onto*** **span{a}**. We leave it as an exercise for you to show that this is a linear operator. The following theorem provides a formula for the standard matrix for T.

Theorem 7.7.3 *If* **a** *is a nonzero vector in* R^n, *and if* **a** *is expressed in column form, then the standard matrix for the linear operator* $T(\mathbf{x}) = \text{proj}_{\mathbf{a}}\mathbf{x}$ *is*

$$P = \frac{1}{\mathbf{a}^T\mathbf{a}}\mathbf{a}\mathbf{a}^T \tag{16}$$

This matrix is symmetric and has rank 1.

Proof The column vectors of the standard matrix for a linear transformation are the images of the standard basis vectors under the transformation. Thus, if we denote the jth entry of $\mathbf{a}$ by a_j, then the jth column of P is given by

$$T(\mathbf{e}_j) = \text{proj}_{\mathbf{a}}\mathbf{e}_j = \frac{\mathbf{e}_j \cdot \mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} = \frac{a_j}{\|\mathbf{a}\|^2}\mathbf{a}$$

Accordingly, partitioning P into column vectors yields

$$P = \left[\frac{a_1}{\|\mathbf{a}\|^2}\mathbf{a} \;\middle|\; \frac{a_2}{\|\mathbf{a}\|^2}\mathbf{a} \;\middle|\; \cdots \;\middle|\; \frac{a_n}{\|\mathbf{a}\|^2}\mathbf{a}\right] = \frac{1}{\|\mathbf{a}\|^2}[a_1\mathbf{a} \mid a_2\mathbf{a} \mid \cdots \mid a_n\mathbf{a}]$$

$$= \frac{1}{\|\mathbf{a}\|^2}\mathbf{a}[a_1 \mid a_2 \mid \cdots \mid a_n] = \frac{1}{\mathbf{a}^T\mathbf{a}}\mathbf{a}\mathbf{a}^T$$

which proves (16). Finally, the matrix $\mathbf{a}\mathbf{a}^T$ is symmetric and has rank 1 (Theorem 7.4.8), so P, being a nonzero scalar multiple of $\mathbf{a}\mathbf{a}^T$, must also be symmetric and have rank 1. ■

We leave it as an exercise for you to show that the matrix P in Formula (16) does not change if $\mathbf{a}$ is replaced by any nonzero scalar multiple of $\mathbf{a}$. This means that P is determined by the *line* onto which it projects and not by the particular basis vector $\mathbf{a}$ that is used to span the line.

In particular, we can use a unit vector $\mathbf{u}$ along the line, in which case $\mathbf{u}^T\mathbf{u} = \|\mathbf{u}\|^2 = 1$, and the formula for P simplifies to

$$P = \mathbf{u}\mathbf{u}^T \tag{17}$$

Thus, we have shown that *the standard matrix for the orthogonal projection of R^n onto a line through the origin can be obtained by finding a unit vector along the line and forming the outer product of that vector with itself.*

EXAMPLE 4

Example 1
Revisited

Use Formula (17) to obtain the standard matrix P_θ for the orthogonal projection of R^2 onto the line W through the origin that makes an angle θ with the positive x-axis of a rectangular xy-coordinate system.

Solution Since $\mathbf{u} = (\cos\theta, \sin\theta)$ is a unit vector along the line, we write this vector in column form and apply Formula (17) to obtain

$$P_\theta = \mathbf{u}\mathbf{u}^T = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} [\cos\theta \quad \sin\theta] = \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix} \tag{18}$$

This agrees with the result in Example 1. ■

EXAMPLE 5 The Standard Matrix for an Orthogonal Projection

(a) Find the standard matrix P for the orthogonal projection of R^3 onto the line spanned by the vector $\mathbf{a} = (1, -4, 2)$.

(b) Use the matrix to find the orthogonal projection of the vector $\mathbf{x} = (2, -1, 3)$ onto the line spanned by $\mathbf{a}$.

(c) Show that P has rank 1, and interpret this result geometrically.

Solution (a) Expressing $\mathbf{a}$ in column form we obtain

$$\mathbf{a}^T\mathbf{a} = [1 \quad -4 \quad 2]\begin{bmatrix} 1 \\ -4 \\ 2 \end{bmatrix} = 21 \quad \text{and} \quad \mathbf{a}\mathbf{a}^T = \begin{bmatrix} 1 \\ -4 \\ 2 \end{bmatrix}[1 \quad -4 \quad 2] = \begin{bmatrix} 1 & -4 & 2 \\ -4 & 16 & -8 \\ 2 & -8 & 4 \end{bmatrix}$$

Thus, it follows from (16) that the standard matrix P for the orthogonal projection is

$$P = \frac{1}{21}\begin{bmatrix} 1 & -4 & 2 \\ -4 & 16 & -8 \\ 2 & -8 & 4 \end{bmatrix} = \begin{bmatrix} \frac{1}{21} & -\frac{4}{21} & \frac{2}{21} \\ -\frac{4}{21} & \frac{16}{21} & -\frac{8}{21} \\ \frac{2}{21} & -\frac{8}{21} & \frac{4}{21} \end{bmatrix} \tag{19}$$

[We could also have obtained this result by normalizing $\mathbf{a}$ and applying Formula (17) using the normalized vector $\mathbf{u}$.] Note that P is symmetric, as expected.

$$P = \frac{1}{21}\begin{bmatrix} 1 & -4 & 2 \\ -4 & 16 & -8 \\ 2 & -8 & 4 \end{bmatrix} = \begin{bmatrix} \frac{1}{21} & -\frac{4}{21} & \frac{2}{21} \\ -\frac{4}{21} & \frac{16}{21} & -\frac{8}{21} \\ \frac{2}{21} & -\frac{8}{21} & \frac{4}{21} \end{bmatrix} \tag{19}$$

Solution (b) The orthogonal projection of $\mathbf{x}$ onto the line spanned by $\mathbf{a}$ is the product $P\mathbf{x}$ with $\mathbf{x}$ expressed in column form. Thus,

$$\text{proj}_{\mathbf{a}}\mathbf{x} = P\mathbf{x} = \begin{bmatrix} \frac{1}{21} & -\frac{4}{21} & \frac{2}{21} \\ -\frac{4}{21} & \frac{16}{21} & -\frac{8}{21} \\ \frac{2}{21} & -\frac{8}{21} & \frac{4}{21} \end{bmatrix}\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} \frac{12}{21} \\ -\frac{48}{21} \\ \frac{24}{21} \end{bmatrix} = \begin{bmatrix} \frac{4}{7} \\ -\frac{16}{7} \\ \frac{8}{7} \end{bmatrix}$$

Solution (c) The matrix P in (19) has rank 1 since the second and third columns are scalar multiples of the first. This tells us that the column space of P is one-dimensional, which makes sense because the column space is the range of the linear operator represented by P, and we know that this is a line through the origin. ■

ORTHOGONAL PROJECTIONS ONTO GENERAL SUBSPACES

Theorem 7.7.4 ***(Projection Theorem for Subspaces)*** *If W is a subspace of R^n, then every vector $\mathbf{x}$ in R^n can be expressed in exactly one way as*

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 \tag{20}$$

where $\mathbf{x}_1$ is in W and $\mathbf{x}_2$ is in $W^{\perp}$.

Proof We will leave the case where $W = \{\mathbf{0}\}$ as an exercise, so we may assume that $W \neq \{\mathbf{0}\}$ and hence has a basis. Let $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ be a basis for W, and form the matrix M that has these basis vectors as successive columns. This makes W the column space of M and $W^{\perp}$ the null space of M^T. Thus, the proof will be complete if we can show that every vector $\mathbf{x}$ in R^n can be expressed in exactly one way as

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$$

where $\mathbf{x}_1$ is in the column space of M and $M^T\mathbf{x}_2 = \mathbf{0}$.

However, to say that $\mathbf{x}_1$ is in the column space of M is equivalent to saying that $\mathbf{x}_1 = M\mathbf{v}$ for some vector $\mathbf{v}$ in R^k, and to say that $M^T\mathbf{x}_2 = \mathbf{0}$ is equivalent to saying that $M^T(\mathbf{x} - \mathbf{x}_1) = \mathbf{0}$.

Thus, if we can show that the equation

$$M^T(\mathbf{x} - M\mathbf{v}) = \mathbf{0} \tag{21}$$

has a unique solution for $\mathbf{v}$, then $\mathbf{x}_1 = M\mathbf{v}$ and $\mathbf{x}_2 = \mathbf{x} - \mathbf{x}_1$ will be uniquely determined vectors with the required properties.

To do this, let us rewrite (21) as

$$M^TM\mathbf{v} = M^T\mathbf{x} \tag{22}$$

The matrix M in this equation has full column rank, since its column vectors are linearly independent.

Thus, it follows from Theorem 7.5.10 that M^TM is invertible, so (22) has the unique solution

$$\mathbf{v} = (M^TM)^{-1}M^T\mathbf{x} \tag{23}$$

■

In the special case where W is a line through the origin of R^n, the vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ in this theorem are those given in Theorem 7.7.1; this suggests that we *define* the vector $\mathbf{x}_1$ in (20) to be the ***orthogonal projection of* x *on W***.

We will see later that the vector $\mathbf{x}_2$ is the ***orthogonal projection of* x *on*** $W^{\perp}$. We will denote these vectors by $\mathbf{x}_1 = \text{proj}_W\mathbf{x}$ and $\mathbf{x}_2 = \text{proj}_{W^{\perp}}\mathbf{x}$, respectively (Figure 7.7.4). Thus, Formula (20) can be expressed as

$$\mathbf{x} = \text{proj}_W\mathbf{x} + \text{proj}_{W^{\perp}}\mathbf{x} \qquad (24)$$

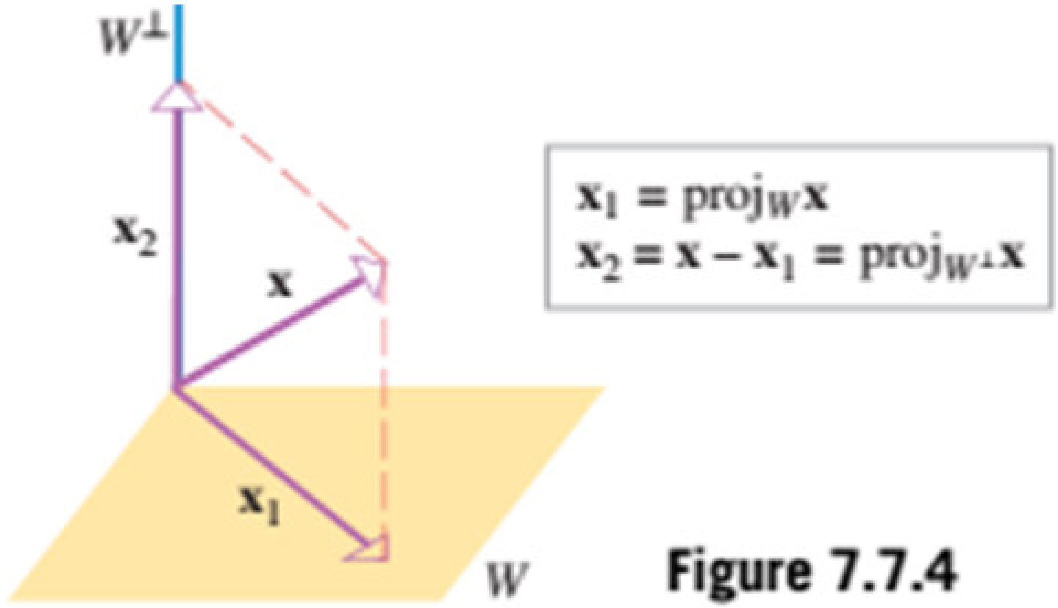


Figure 7.7.4

The proof of the following result follows from Formula (23) in the proof of Theorem 7.7.4 and the relationship $\mathbf{x}_1 = \text{proj}_W \mathbf{x} = M\mathbf{v}$ that was established in that proof.

Theorem 7.7.5 *If W is a nonzero subspace of R^n, and if M is any matrix whose column vectors form a basis for W, then*

$$\text{proj}_W \mathbf{x} = M(M^T M)^{-1} M^T \mathbf{x} \tag{25}$$

for every column vector $\mathbf{x}$ in R^n.

Formula (25) can be used to define the linear operator

$$T(\mathbf{x}) = \text{proj}_W(\mathbf{x}) = M(M^T M)^{-1} M^T \mathbf{x} \tag{26}$$

on R^n whose standard matrix P is

$$P = M(M^T M)^{-1} M^T \tag{27}$$

We call this operator the ***orthogonal projection of R^n onto W***.

REMARK When working with Formulas (26) and (27), it is important to keep in mind that the matrix M is not unique, since its column vectors can be any basis vectors for W; that is, no matter what basis vectors you use to construct M, you will obtain the same operator T and the same matrix P.

EXAMPLE 6 Orthogonal Projection of R^3 onto a Plane Through the Origin

(a) Find the standard matrix P for the orthogonal projection of R^3 onto the plane $x - 4y + 2z = 0$.

(b) Use the matrix P to find the orthogonal projection of the vector $\mathbf{x} = (1, 0, 4)$ onto the plane.

Solution (a) Our strategy will be to find a basis for the plane, create a matrix M with the basis vectors as columns, and then use (27) to obtain P. To find a basis for the plane, we will view the equation $x - 4y + 2z = 0$ as a linear system of one equation in three unknowns and find a basis for the solution space. Solving the equation for its leading variable x and assigning arbitrary values to the free variables yields the general solution

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4t_1 - 2t_2 \\ t_1 \\ t_2 \end{bmatrix} = t_1 \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

The two column vectors on the right side form a basis for the solution space, so we take the matrix M to be

$$M = \begin{bmatrix} 4 & -2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

As stressed, matrix M is not unique!

Thus,

$$M^TM = \begin{bmatrix} 4 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & -2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 17 & -8 \\ -8 & 5 \end{bmatrix} \quad \text{and} \quad (M^TM)^{-1} = \begin{bmatrix} \frac{5}{21} & \frac{8}{21} \\ \frac{8}{21} & \frac{17}{21} \end{bmatrix}$$

and therefore

$$P = M(M^TM)^{-1}M^T = \begin{bmatrix} 4 & -2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{5}{21} & \frac{8}{21} \\ \frac{8}{21} & \frac{17}{21} \end{bmatrix} \begin{bmatrix} 4 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{20}{21} & \frac{4}{21} & -\frac{2}{21} \\ \frac{4}{21} & \frac{5}{21} & \frac{8}{21} \\ -\frac{2}{21} & \frac{8}{21} & \frac{17}{21} \end{bmatrix} \tag{28}$$

Solution (b) The orthogonal projection of $\mathbf{x}$ onto the plane $x - 4y + 2z = 0$ is $P\mathbf{x}$ with $\mathbf{x}$ expressed in column form. Thus,

$$P\mathbf{x} = \begin{bmatrix} \frac{20}{21} & \frac{4}{21} & -\frac{2}{21} \\ \frac{4}{21} & \frac{5}{21} & \frac{8}{21} \\ -\frac{2}{21} & \frac{8}{21} & \frac{17}{21} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} \frac{12}{21} \\ \frac{36}{21} \\ \frac{66}{21} \end{bmatrix} = \begin{bmatrix} \frac{4}{7} \\ \frac{12}{7} \\ \frac{22}{7} \end{bmatrix}$$

If preferred, this can be written in the comma-delimited form $P\mathbf{x} = \left(\frac{4}{7}, \frac{12}{7}, \frac{22}{7}\right)$ for consistency with the comma-delimited notation that was originally used for $\mathbf{x}$. As a check, you may want to confirm that $P\mathbf{x}$ is in the plane $x - 4y + 2z = 0$ and $\mathbf{x} - P\mathbf{x}$ is orthogonal to $P\mathbf{x}$. ■

WHEN DOES A MATRIX REPRESENT AN ORTHOGONAL PROJECTION?

We now turn to the problem of determining what properties an $n \times n$ matrix P must have in order to represent an orthogonal projection onto a k-dimensional subspace W of R^n. Some of the properties are clear. For example, since W is k-dimensional, the column space of P must be k-dimensional, and P must have rank k.

We also know from (27) that if M is any $n \times k$ matrix whose column vectors form a basis for W, then

$$P^T = \left(M(M^TM)^{-1}M^T\right)^T = M(M^TM)^{-1}M^T = P$$

so P must be symmetric. Moreover,

$$\begin{aligned} P^2 &= \left(M(M^TM)^{-1}M^T\right)\left(M(M^TM)^{-1}M^T\right) \\ &= M(M^TM)^{-1}(M^TM)(M^TM)^{-1}M^T = M(M^TM)^{-1}M^T = P \end{aligned}$$

so P must be the same as its square. This makes sense intuitively, since the orthogonal projection of R^n onto W leaves vectors in W unchanged.

In particular, it leaves $P\mathbf{x}$ unchanged for each $\mathbf{x}$ in R^n, so

$$P^2\mathbf{x} = P(P\mathbf{x}) = P\mathbf{x}$$

and this implies that $P^2 = P$.

A matrix that is the same as its square is said to be ***idempotent***. Thus, we have shown that the standard matrix for an orthogonal projection of R^n onto a k-dimensional subspace has rank k, is symmetric, and is idempotent. In the exercises we will ask you to prove that the converse is also true, thereby establishing the following theorem.

Theorem 7.7.6 *An $n \times n$ matrix P is the standard matrix for an orthogonal projection of R^n onto a k-dimensional subspace of R^n if and only if P is symmetric, idempotent, and has rank k. The subspace W is the column space of P.*

EXAMPLE 7

Properties of Orthogonal Projections

In Example 5 we showed that the standard matrix P in (19) for the orthogonal projection of R^3 onto the line spanned by the vector $\mathbf{a} = (1, -4, 2)$ has rank 1, which is consistent with the fact that the line is one-dimensional. In accordance with Theorem 7.7.6, the matrix P is symmetric (verify) and is idempotent, since

$$P^2 = \begin{bmatrix} \frac{1}{21} & -\frac{4}{21} & \frac{2}{21} \\ -\frac{4}{21} & \frac{16}{21} & -\frac{8}{21} \\ \frac{2}{21} & -\frac{8}{21} & \frac{4}{21} \end{bmatrix} \begin{bmatrix} \frac{1}{21} & -\frac{4}{21} & \frac{2}{21} \\ -\frac{4}{21} & \frac{16}{21} & -\frac{8}{21} \\ \frac{2}{21} & -\frac{8}{21} & \frac{4}{21} \end{bmatrix} = \begin{bmatrix} \frac{1}{21} & -\frac{4}{21} & \frac{2}{21} \\ -\frac{4}{21} & \frac{16}{21} & -\frac{8}{21} \\ \frac{2}{21} & -\frac{8}{21} & \frac{4}{21} \end{bmatrix} = P \qquad \blacksquare$$

EXAMPLE 8 Identifying Orthogonal Projections

Show that

$$A = \begin{bmatrix} \frac{1}{9} & \frac{2}{9} & \frac{2}{9} \\ \frac{2}{9} & \frac{4}{9} & \frac{4}{9} \\ \frac{2}{9} & \frac{4}{9} & \frac{4}{9} \end{bmatrix}$$

is the standard matrix for an orthogonal projection of R^3 onto a line through the origin, and find the line.

Solution We leave it for you to confirm that A is symmetric, idempotent, and has rank 1. Thus, it follows from Theorem 7.7.6 that A represents an orthogonal projection of R^3 onto a line through the origin. That line is the column space of A, and since the second and third column vectors are scalar multiples of the first, we can take the first column vector of A as a basis for the line.

Moreover, since any scalar multiple of this column vector is also a basis for the line, we might as well multiply by 9 to simplify the components. Thus, the line can be expressed in comma-delimited form as the span of the vector $\mathbf{a} = (1, 2, 2)$, or it can be expressed parametrically in xyz-coordinates as

$$x = t, \quad y = 2t, \quad z = 2t$$

■

STRANG DIAGRAMS

Formula (24) is useful for studying systems of linear equations. To see why this is so, suppose that A is an $m \times n$ matrix, so that $A\mathbf{x} = \mathbf{b}$ is a linear system of m equations in n unknowns. Since $\mathbf{x}$ is a vector in R^n, we can apply Formula (24) with $W = \text{row}(A)$ and $W^\perp = \text{null}(A)$ to express $\mathbf{x}$ as a sum of two orthogonal terms

$$\mathbf{x} = \mathbf{x}_{\text{row}(A)} + \mathbf{x}_{\text{null}(A)} \tag{29}$$

where $\mathbf{x}_{\text{row}(A)}$ is the orthogonal projection of $\mathbf{x}$ onto the row space of A, and $\mathbf{x}_{\text{null}(A)}$ is the orthogonal projection of $\mathbf{x}$ onto the null space of A;

similarly, since $\mathbf{b}$ is a vector in R^m, we can apply Formula (24) to $\mathbf{b}$ with $W = \text{col}(A)$ and $W^\perp = \text{null}(A^T)$ to express $\mathbf{b}$ as a sum of two orthogonal terms

$$\mathbf{b} = \mathbf{b}_{\text{col}(A)} + \mathbf{b}_{\text{null}(A^T)} \tag{30}$$

where $\mathbf{b}_{\text{col}(A)}$ is the orthogonal projection of $\mathbf{b}$ onto the column space of A, and $\mathbf{b}_{\text{null}(A^T)}$ is the orthogonal projection of $\mathbf{b}$ onto the null space of A^T.

$$\mathbf{x} = \text{proj}_W \mathbf{x} + \text{proj}_{W^\perp} \mathbf{x} \tag{24}$$

$$\mathbf{x} = \mathbf{x}_{\text{row}(A)} + \mathbf{x}_{\text{null}(A)} \qquad (29) \qquad\qquad \mathbf{b} = \mathbf{b}_{\text{col}(A)} + \mathbf{b}_{\text{null}(A^T)} \qquad (30)$$

The decompositions in (29) and (30) can be pictured as in Figure 7.7.5 in which we have represented the fundamental spaces of A as perpendicular lines. We will call this a ***Strang diagram***.*

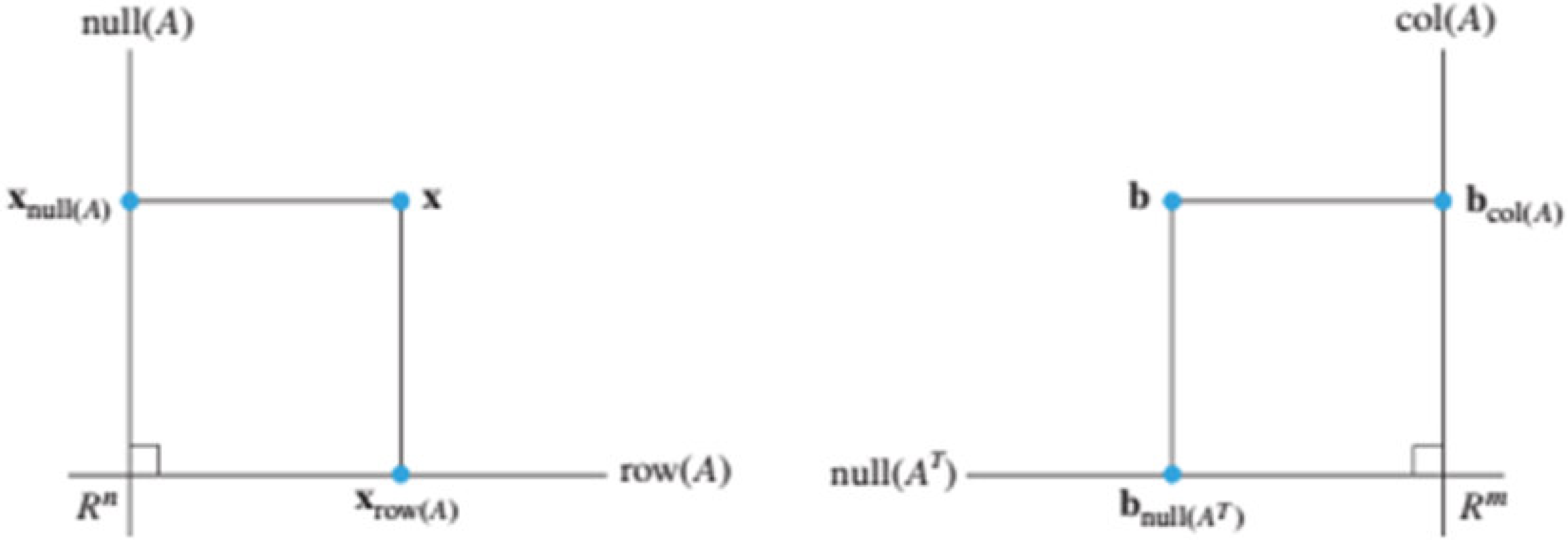


Figure 7.7.5

The fact that the fundamental spaces are represented by lines in a Strang diagram is pictorial only and is not intended to suggest that those spaces are one-dimensional.

What we do know for sure is that

$$\dim(\text{row}(A)) + \dim(\text{null}(A)) = n \tag{31}$$

$$\dim(\text{col}(A)) + \dim(\text{null}(A^T)) = m \tag{32}$$

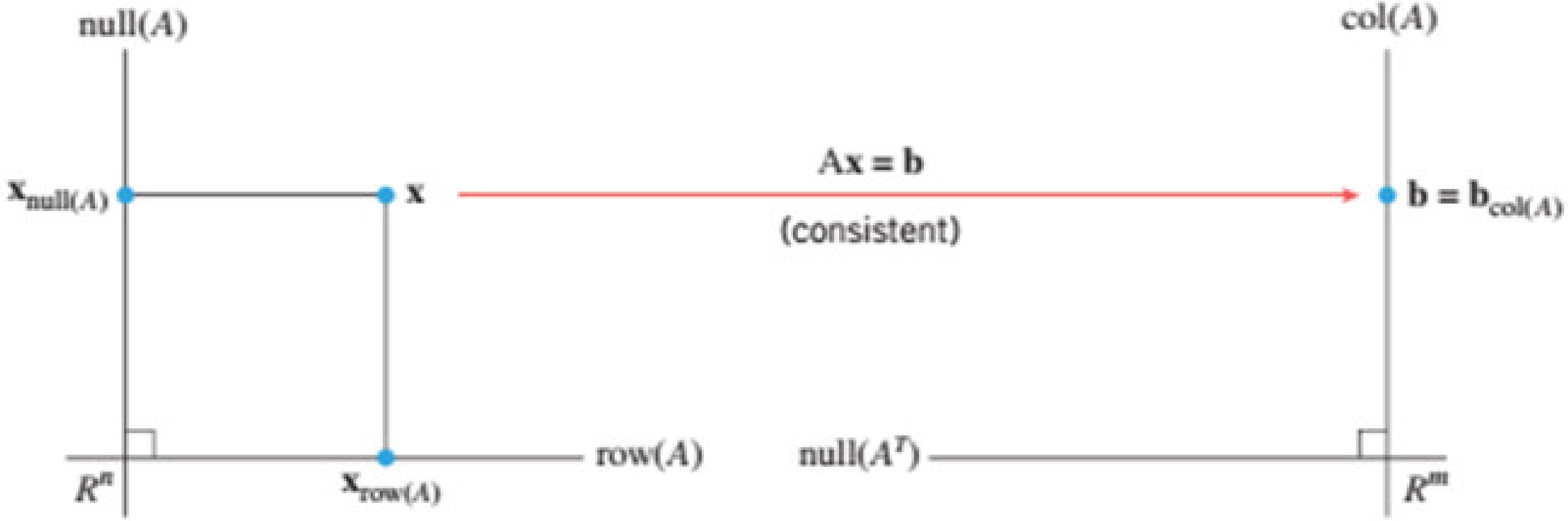
[see (5) in Section 7.5].

Also, we know from Theorem 3.5.5 that the system $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is in the column space of A, that is, if and only if $\mathbf{b}_{\text{null}(A^T)} = \mathbf{0}$ in (30).

$$\mathbf{b} = \mathbf{b}_{\text{col}(A)} + \mathbf{b}_{\text{null}(A^T)} \tag{30}$$

This is illustrated by the Strang diagrams in Figure 7.7.6.

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



Strang diagram for a consistent linear system

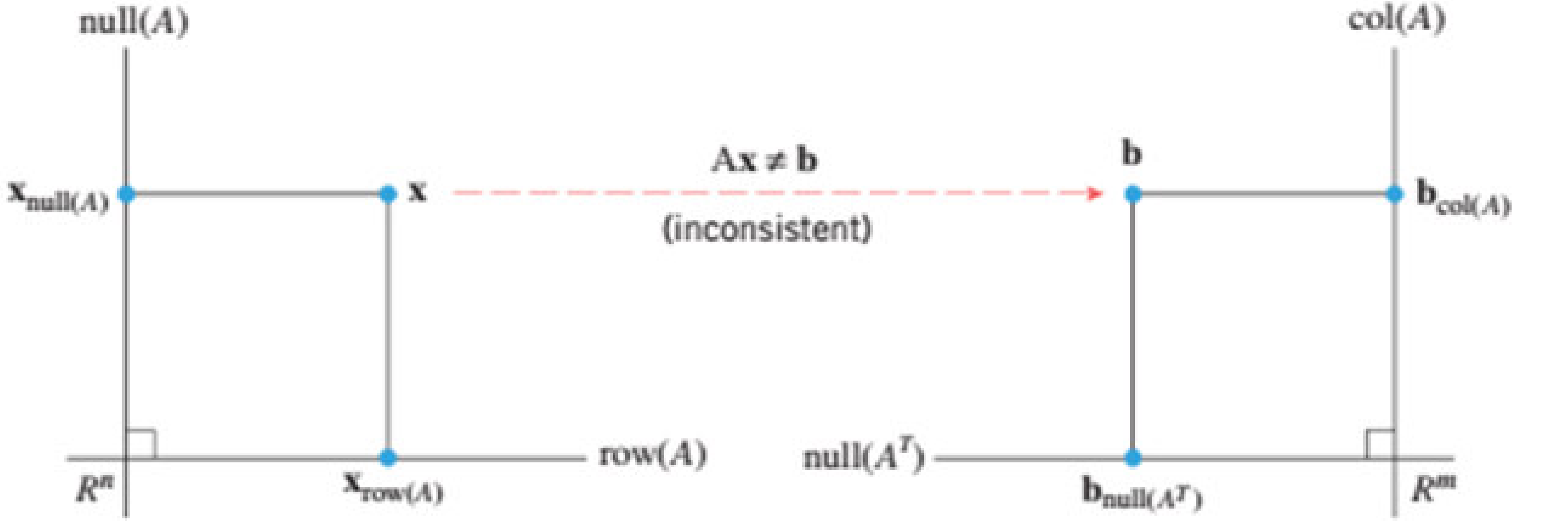


Figure 7.7.6 Strang diagram for an inconsistent linear system

“Teorema da Perpendicular Dupla”

THE DOUBLE PERP THEOREM

In Theorem 7.3.4 we stated without proof that if W is a subspace of R^n, then $(W^\perp)^\perp = W$. Although this result may seem somewhat technical, it is important because it establishes that orthogonal complements occur in “companion pairs” in the sense that each of the spaces W and $W^\perp$ is the orthogonal complement of the other.

Thus, for example, knowing that the null space of a matrix is the orthogonal complement of the row space automatically implies that the row space is the orthogonal complement of the null space.

We now have all of the mathematical machinery to prove this result.

Theorem 7.7.8 ***(Double Perp Theorem)*** *If W is a subspace of R^n, then $(W^\perp)^\perp = W$.*

REMARK If W is a subspace of R^n and $W^\perp$ is its orthogonal complement, then the equation $(W^\perp)^\perp = W$ in part (*c*) of theorem 7.3.4 states that W is the orthogonal complement of $W^\perp$. This establishes a symmetry that allows us to say that *W and $W^\perp$ are orthogonal complements of one another*. Note, however, that it is required that W be a subspace of R^n (not just a subset) for this to be true.

Proof Our strategy will be to show that every vector in W is in $(W^{\perp})^{\perp}$, and conversely that every vector in $(W^{\perp})^{\perp}$ is in W. Stated using set notation, we will be showing that $W \subset (W^{\perp})^{\perp}$ and $(W^{\perp})^{\perp} \subset W$, thereby proving that $W = (W^{\perp})^{\perp}$.

Let $\mathbf{w}$ be any vector in W. By definition, $W^{\perp}$ consists of all vectors that are orthogonal to every vector in W. Thus, every vector in $W^{\perp}$ is orthogonal to $\mathbf{w}$, and this implies that $\mathbf{w}$ is in $(W^{\perp})^{\perp}$.

Conversely, let $\mathbf{w}$ be any vector in $(W^{\perp})^{\perp}$. It follows from the projection theorem (Theorem 7.7.4) that $\mathbf{w}$ can be expressed uniquely as

$$\mathbf{w} = \mathbf{w}_1 + \mathbf{w}_2$$

where $\mathbf{w}_1$ is a vector in W and $\mathbf{w}_2$ is a vector in $W^{\perp}$.

To show that $\mathbf{w}$ is a vector in W, we will show that $\mathbf{w}_2 = \mathbf{0}$ (which then makes $\mathbf{w} = \mathbf{w}_1$, which we know to be in W). Toward this end, observe that $\mathbf{w}_2$ is orthogonal to $\mathbf{w}$, so

$$\mathbf{w}_2 \cdot \mathbf{w} = 0$$

as **w** is in $(W^\mathsf{T})^\mathsf{T}$ (hypothesis)

which implies that

$$(\mathbf{w}_2 \cdot \mathbf{w}_1) + (\mathbf{w}_2 \cdot \mathbf{w}_2) = 0$$

However, $\mathbf{w}_2$ is orthogonal to $\mathbf{w}_1$ (why?), so this equation simplifies to $\mathbf{w}_2 \cdot \mathbf{w}_2 = 0$, which implies that $\mathbf{w}_2 = \mathbf{0}$. ■

ORTHOGONAL PROJECTIONS ONTO $W^{\perp}$

Given a subspace W of R^n, the standard matrix for the orthogonal projection $\text{proj}_{W^\perp}\mathbf{x}$ can be computed in one of two ways—you can either apply Formula (26) with the column vectors of M taken to be basis vectors for $W^{\perp}$, or you can use the fact that $\mathbf{x} = \text{proj}_W\mathbf{x} + \text{proj}_{W^\perp}\mathbf{x}$ to write $\text{proj}_{W^\perp}\mathbf{x}$ in terms of $\text{proj}_W\mathbf{x}$ as

$$\text{proj}_{W^\perp}\mathbf{x} = \mathbf{x} - \text{proj}_W\mathbf{x} = I\mathbf{x} - P\mathbf{x} = (I - P)\mathbf{x} \tag{35}$$

It now follows from (27) and (35) that the standard matrix for $\text{proj}_{W^\perp}\mathbf{x}$ can be expressed in terms of the standard matrix P for $\text{proj}_W\mathbf{x}$ as

$$I - P = I - M(M^TM)^{-1}M^T \tag{36}$$

where the column vectors of M form a basis for W.

$$\text{proj}_W(\mathbf{x}) = M(M^TM)^{-1}M^T\mathbf{x} \tag{26}$$

$$P = M(M^TM)^{-1}M^T \tag{27}$$

EXAMPLE 9 Orthogonal Projection onto an Orthogonal Complement

In Example 6 we showed that the standard matrix for the orthogonal projection of R^3 onto the plane $x - 4y + 2z = 0$ is

$$P = \begin{bmatrix} \frac{20}{21} & \frac{4}{21} & -\frac{2}{21} \\ \frac{4}{21} & \frac{5}{21} & \frac{8}{21} \\ -\frac{2}{21} & \frac{8}{21} & \frac{17}{21} \end{bmatrix}$$

In this case the orthogonal complement of the plane is the line through the origin that is perpendicular to the given plane, that is, the line spanned by the vector $\mathbf{a} = (1, -4, 2)$. It follows from (36) that the orthogonal projection of R^3 onto this line is

$$I - P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} \frac{20}{21} & \frac{4}{21} & -\frac{2}{21} \\ \frac{4}{21} & \frac{5}{21} & \frac{8}{21} \\ -\frac{2}{21} & \frac{8}{21} & \frac{17}{21} \end{bmatrix} = \begin{bmatrix} \frac{1}{21} & -\frac{4}{21} & \frac{2}{21} \\ -\frac{4}{21} & \frac{16}{21} & -\frac{8}{21} \\ \frac{2}{21} & -\frac{8}{21} & \frac{4}{21} \end{bmatrix}$$

Note that this is consistent with the result that we obtained in Example 5 using Theorem 7.7.3. ■