CHAPTER 8

Transforming matrices to diagonal form is important mathematically as well as in such applications as vibration analysis, face and fingerprint recognition, statistics, and data compression.

Diagonalization

source: Contemporary Linear Algebra, H. Anton / R.C. Busby

Section 8.1 Matrix Representations of Linear Transformations

MATRIX OF A LINEAR OPERATOR WITH RESPECT TO A BASIS

We know that every linear transformation $T: R^n \to R^m$ has an associated standard matrix

$$[T] = [T(\mathbf{e}_1) \mid T(\mathbf{e}_2) \mid \cdots \mid T(\mathbf{e}_n)]$$

with the property that

$$T(\mathbf{x}) = [T]\mathbf{x}$$

for every vector $\mathbf{x}$ in R^n. For the moment we will focus on the case where T is a linear operator on R^n, so the standard matrix $[T]$ is a square matrix of size $n \times n$.

Sometimes the form of the standard matrix fully reveals the geometric properties of a linear operator and sometimes it does not. For example, we can tell by inspection of the matrix

$$[T_1] = \begin{bmatrix} \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} & 0 \\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

that T_1 is a rotation through an angle of $\pi/4$ about the z-axis of an xyz-coordinate system.

In contrast, a casual inspection of the matrix

$$[T_2] = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{2}$$

provides only partial geometric information about the operator T_2; we can tell that T_2 is a rotation since the matrix $[T_2]$ is orthogonal and has determinant 1, but, unlike (1), this matrix does not explicitly reveal the axis and angle of rotation.

The difference between (1) and (2) has to do with the orientation of the standard basis. In the case of the operator T_1, the standard basis vector $\mathbf{e}_3$ aligns with the axis of rotation, and the basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ rotate in a plane perpendicular to $\mathbf{e}_3$, thereby making the axis and angle of rotation recognizable from the standard matrix.

However, in the case of the operator T_2, none of the standard basis vectors aligns with the axis of rotation (see Example 7 and Figure 6.2.9 of Section 6.2), so the operator T_2 does not transform $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ in a way that provides useful geometric information.

➡ Thus, although the standard basis is simple algebraically, it is not always the best basis from a geometric point of view.

Our primary goal in this section is to develop a way of using bases other than the standard basis to create matrices that describe the geometric behavior of a linear transformation better than the standard matrix. The key to doing this is to work with *coordinates* of vectors rather than with the vectors themselves, as we will now explain.

Suppose that

$$\mathbf{x} \xrightarrow{T} T(\mathbf{x})$$

is a linear operator on R^n and B is a basis for R^n. In the course of mapping $\mathbf{x}$ into $T(\mathbf{x})$ this operator creates a companion operator

$$[\mathbf{x}]_B \longrightarrow [T(\mathbf{x})]_B \tag{3}$$

that maps the coordinate matrix $[\mathbf{x}]_B$ into the coordinate matrix $[T(\mathbf{x})]_B$. In the exercises we will ask you to show that (3) is linear and hence must be a matrix transformation; that is, there must be a matrix A such that

$$A[\mathbf{x}]_B = [T(\mathbf{x})]_B$$

The following theorem shows how to find the matrix A.

Theorem 8.1.1 *Let $T: R^n \to R^n$ be a linear operator, let $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for R^n, and let*

$$A = \big[[T(\mathbf{v}_1)]_B \mid [T(\mathbf{v}_2)]_B \mid \cdots \mid [T(\mathbf{v}_n)]_B\big] \tag{4}$$

Then

$$[T(\mathbf{x})]_B = A[\mathbf{x}]_B \tag{5}$$

for every vector $\mathbf{x}$ in R^n. Moreover, the matrix A given by Formula (4) *is the only matrix with property* (5).

Proof Let $\mathbf{x}$ be any vector in R^n, and suppose that its coordinate matrix with respect to B is

$$[\mathbf{x}]_B = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

That is,

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

It now follows from the linearity of T that

$$T(\mathbf{x}) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + \cdots + c_nT(\mathbf{v}_n)$$

and from the linearity of coordinate maps that

$$[T(\mathbf{x})]_B = c_1[T(\mathbf{v}_1)]_B + c_2[T(\mathbf{v}_2)]_B + \cdots + c_n[T(\mathbf{v}_n)]_B$$

which we can write in matrix form as

$$[T(\mathbf{x})]_B = \big[[T(\mathbf{v}_1)]_B \mid [T(\mathbf{v}_2)]_B \mid \cdots \mid [T(\mathbf{v}_n)]_B\big] \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = A[\mathbf{x}]_B$$

This proves that the matrix A in (4) has property (5). Moreover, A is the only matrix with this property, for if there exists a matrix C such that

$$[T(\mathbf{x})]_B = A[\mathbf{x}]_B = C[\mathbf{x}]_B$$

for all $\mathbf{x}$ in R^n, then Theorem 7.11.6 implies that $A = C$. ■

The matrix A in (4) is called the ***matrix for T with respect to the basis B*** and is denoted by

$$[T]_B = \left[[T(\mathbf{v}_1)]_B \mid [T(\mathbf{v}_2)]_B \mid \cdots \mid [T(\mathbf{v}_n)]_B\right] \tag{6}$$

Using this notation we can write (5) as

$$[T(\mathbf{x})]_B = [T]_B[\mathbf{x}]_B \tag{7}$$

Recalling that the components of a vector in R^n are the same as its coordinates with respect to the standard basis S, it follows from (6) that

$$[T]_S = \left[[T(\mathbf{v}_1)]_S \mid [T(\mathbf{v}_2)]_S \mid \cdots \mid [T(\mathbf{v}_n)]_S\right] = [T(\mathbf{v}_1) \mid T(\mathbf{v}_2) \mid \cdots \mid T(\mathbf{v}_n)] = [T]$$

That is, *the matrix for a linear operator on R^n with respect to the standard basis is the same as the standard matrix for T.*

The basis consisting of $\{v_1, v_2, \ldots, v_n\}$ is then $\{e_1, e_2, \ldots, e_n\}$.

EXAMPLE 1 Matrix of a Linear Operator with Respect to a Basis B

Let $T : R^2 \rightarrow R^2$ be the linear operator whose standard matrix is

$$[T] = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix} \tag{8}$$

Find the matrix for T with respect to the basis $B = \{\mathbf{v}_1, \mathbf{v}_2\}$, where

$$\mathbf{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

Solution The images of the basis vectors under the operator T are

$$T(\mathbf{v}_1) = [T]\mathbf{v}_1 = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} = \mathbf{v}_1 = \mathbf{v}_1 + 0\mathbf{v}_2$$

$$T(\mathbf{v}_2) = [T]\mathbf{v}_2 = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{5}{\sqrt{2}} \\ \frac{5}{\sqrt{2}} \end{bmatrix} = 5\mathbf{v}_2 = 0\mathbf{v}_1 + 5\mathbf{v}_2$$

so the coordinate matrices of these vectors with respect to B are

$$[T(\mathbf{v}_1)]_B = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad [T(\mathbf{v}_2)]_B = \begin{bmatrix} 0 \\ 5 \end{bmatrix}$$

Thus, it follows from (6) that

$$[T]_B = \big[[T(\mathbf{v}_1)]_B \mid [T(\mathbf{v}_2)]_B\big] = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}$$

This matrix reveals geometric information about the operator T that was not evident from the standard matrix. It tells us that the effect of T is to stretch the $\mathbf{v}_2$-coordinate of a vector by a factor of 5 and to leave the $\mathbf{v}_1$-coordinate unchanged.

For example, Figure 8.1.1 shows the stretching effect that this operator has on a square of side 1 that is centered at the origin and whose sides align with $\mathbf{v}_1$ and $\mathbf{v}_2$. ■

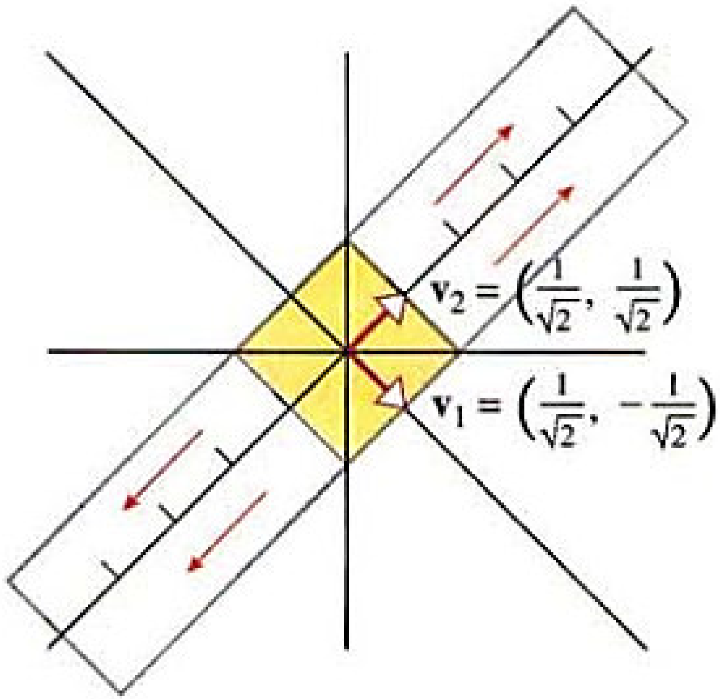


Figure 8.1.1

EXAMPLE 2 Uncovering Hidden Geometry

Let $T : R^3 \to R^3$ be the linear operator whose standard matrix is

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{9}$$

We showed in Example 7 of Section 6.2 that T is a rotation through an angle of $2\pi/3$ about an axis in the direction of the vector $\mathbf{n} = (1, 1, 1)$.

Let us now consider how the matrix for T would look with respect to an orthonormal basis $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ in which $\mathbf{v}_3 = \mathbf{v}_1 \times \mathbf{v}_2$ is a positive scalar multiple of $\mathbf{n}$ and $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthonormal basis for the plane W through the origin that is perpendicular to the axis of rotation (Figure 8.1.2). The rotation leaves the vector $\mathbf{v}_3$ fixed, so

$$T(\mathbf{v}_3) = \mathbf{v}_3 = 0\mathbf{v}_1 + 0\mathbf{v}_2 + 1\mathbf{v}_3$$

and hence

$$[T(\mathbf{v}_3)]_B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

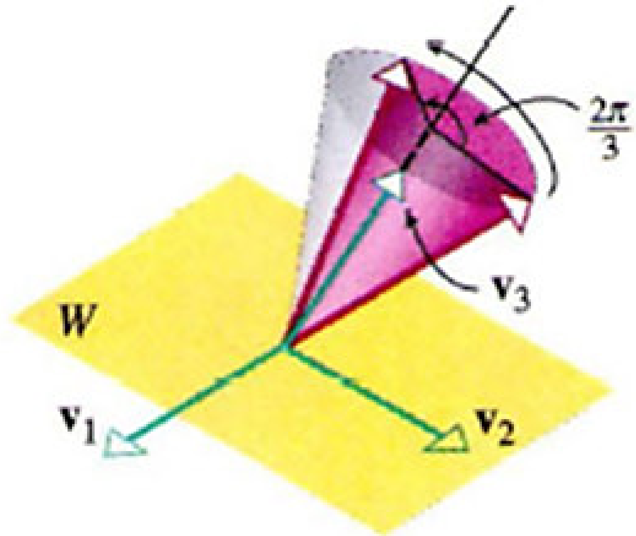


Figure 8.1.2

Also, $T(\mathbf{v}_1)$ and $T(\mathbf{v}_2)$ are linear combinations of $\mathbf{v}_1$ and $\mathbf{v}_2$, since these vectors lie in W. This implies that the third coordinate of both $[T(\mathbf{v}_1)]_B$ and $[T(\mathbf{v}_2)]_B$ must be zero, and the matrix for T with respect to the basis B must be of the form

$$[T]_B = \left[[T(\mathbf{v}_1)]_B \mid [T(\mathbf{v}_2)]_B \mid [T(\mathbf{v}_3)]_B\right] = \begin{bmatrix} \times & \times & 0 \\ \times & \times & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Since T behaves exactly like a rotation of R^2 in the plane W, the block of missing entries has the form of a rotation matrix in R^2. Thus,

$$[T]_B = \left[[T(\mathbf{v}_1)]_B \mid [T(\mathbf{v}_2)]_B \mid [T(\mathbf{v}_3)]_B\right] = \begin{bmatrix} \cos\frac{2\pi}{3} & -\sin\frac{2\pi}{3} & 0 \\ \sin\frac{2\pi}{3} & \cos\frac{2\pi}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

This matrix makes it clear that the angle of rotation is $2\pi/3$ and the axis of rotation is in the direction of $\mathbf{v}_3$, facts that are not directly evident from the standard matrix in (9). ■

CHANGING BASES

It is reasonable to conjecture that two matrices representing the same linear operator with respect to different bases must be related algebraically. To uncover that relationship, suppose that $T: R^n \to R^n$ is a linear operator and that $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $B' = \{\mathbf{v}'_1, \mathbf{v}'_2, \ldots, \mathbf{v}'_n\}$ are bases for R^n.

Also, let $P = P_{B \to B'}$ be the transition matrix from B to B' (so $P^{-1} = P_{B' \to B}$ is the transition matrix from B' to B). To find the relationship between $[T]_B$ and $[T]_{B'}$, consider the diagram in Figure 8.1.3, which links together the following four relationships schematically:

$$[T]_B[\mathbf{x}]_B = [T(\mathbf{x})]_B, \qquad [T]_{B'}[\mathbf{x}]_{B'} = [T(\mathbf{x})]_{B'}$$

$$P[T(\mathbf{x})]_B = P_{B \to B'}[T(\mathbf{x})]_B = [T(\mathbf{x})]_{B'}, \qquad P[\mathbf{x}]_B = P_{B \to B'}[\mathbf{x}]_B = [\mathbf{x}]_{B'}$$

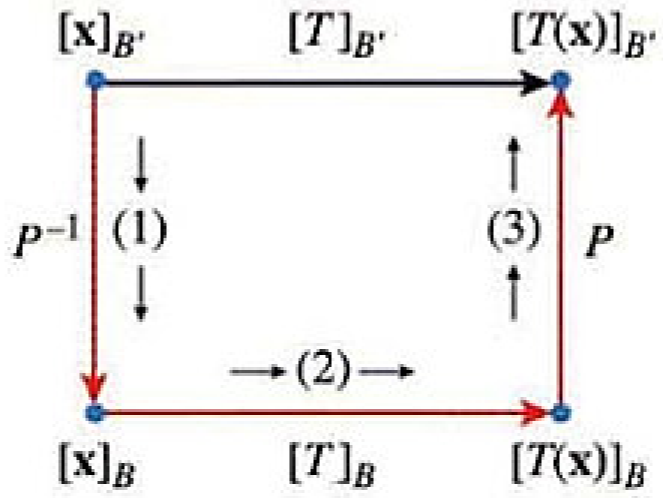


Figure 8.1.3

The diagram shows two different paths from $[\mathbf{x}]_{B'}$ to $[T(\mathbf{x})]_{B'}$, each of which corresponds to a different relationship between these vectors:

1. The direct path from $[\mathbf{x}]_{B'}$ to $[T(\mathbf{x})]_{B'}$ across the top of the diagram corresponds to the relationship

$$[T]_{B'}[\mathbf{x}]_{B'} = [T(\mathbf{x})]_{B'} \tag{10}$$

2. The path from $[\mathbf{x}]_{B'}$ to $[T(\mathbf{x})]_{B'}$ that goes down the left side, across the bottom, and up the right side corresponds to computing $[T(\mathbf{x})]_{B'}$ from $[\mathbf{x}]_{B'}$ by three successive matrix multiplications:

   (i) Multiply $[\mathbf{x}]_{B'}$ on the left by P^{-1} to obtain $P^{-1}[\mathbf{x}]_{B'} = [\mathbf{x}]_B$.
   (ii) Multiply $[\mathbf{x}]_B$ on the left by $[T]_B$ to obtain $[T]_B[\mathbf{x}]_B = [T(\mathbf{x})]_B$.
   (iii) Multiply $[T(\mathbf{x})]_B$ on the left by P to obtain $[T(\mathbf{x})]_{B'}$.

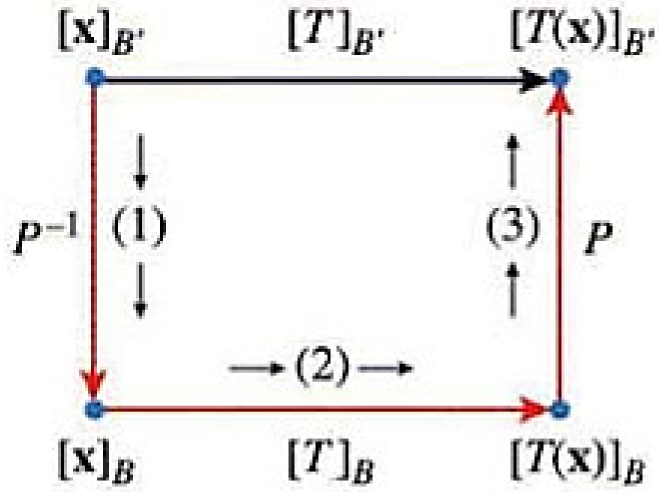


Figure 8.1.3

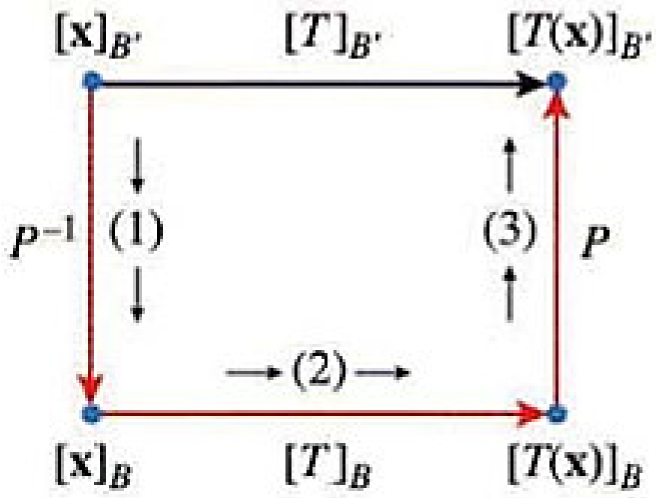


Figure 8.1.3

This process produces the relationship

$$(P[T]_B P^{-1})[\mathbf{x}]_{B'} = [T(\mathbf{x})]_{B'} \tag{11}$$

Thus, (10) and (11) together imply that

$$(P[T]_B P^{-1})[\mathbf{x}]_{B'} = [T]_{B'}[\mathbf{x}]_{B'}$$

Since this holds for all $\mathbf{x}$ in R^n, it follows from Theorem 7.11.6 that

$$P[T]_B P^{-1} = [T]_{B'}$$

Thus, we have established the following theorem that provides the relationship between the matrices for a fixed linear operator with respect to different bases.

$$[T]_{B'}[\mathbf{x}]_{B'} = [T(\mathbf{x})]_{B'} \tag{10}$$

Theorem 8.1.2 *If $T : R^n \to R^n$ is a linear operator, and if $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $B' = \{\mathbf{v}'_1, \mathbf{v}'_2, \ldots, \mathbf{v}'_n\}$ are bases for R^n, then $[T]_B$ and $[T]_{B'}$ are related by the equation*

$$[T]_{B'} = P[T]_B P^{-1} \tag{12}$$

in which

$$P = P_{B \to B'} = \left[[\mathbf{v}_1]_{B'} \mid [\mathbf{v}_2]_{B'} \mid \cdots \mid [\mathbf{v}_n]_{B'}\right] \tag{13}$$

is the transition matrix from B to B'. In the special case where B and B' are orthonormal bases the matrix P is orthogonal, so (12) *is of the form*

$$[T]_{B'} = P[T]_B P^T \tag{14}$$

When convenient, Formula (12) can be rewritten as

$$[T]_B = P^{-1}[T]_{B'} P \tag{15}$$

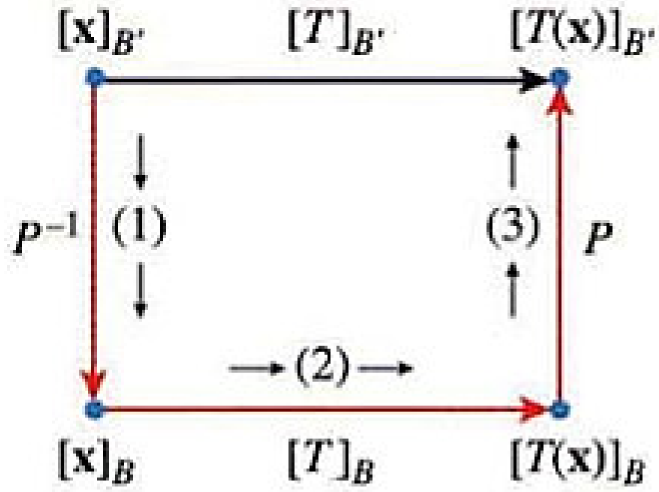
and in the case where the bases are orthonormal this equation can be expressed as

$$[T]_B = P^T[T]_{B'} P \tag{16}$$

REMARK When applying all of these formulas it is easy to lose track of whether P is the transition matrix from B to B', or vice versa, particularly if other notations are used for the bases.

A good way to keep everything straight is to draw Figure 8.1.3 with appropriate adjustments in notation. When creating the diagram you can choose either direction for the transition matrix P as long as you adhere to that direction when constructing the associated formula.

Figure 8.1.3



Since many linear operators are defined by their standard matrices, it is important to consider the special case of Theorem 8.1.2 in which $B' = S$ is the standard basis for R^n. In this case $[T]_{B'} = [T]_S = [T]$, and the transition matrix P from B to B' has the simplified form

$$P = P_{B\to B'} = P_{B\to S} = \big[[\mathbf{v}_1]_S \mid [\mathbf{v}_2]_S \mid \cdots \mid [\mathbf{v}_n]_S\big] = [\mathbf{v}_1 \mid \mathbf{v}_2 \mid \cdots \mid \mathbf{v}_n]$$

Thus, we have the following result.

Theorem 8.1.3 *If $T : R^n \to R^n$ is a linear operator, and if $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for R^n, then $[T]$ and $[T]_B$ are related by the equation*

$$[T] = P[T]_B P^{-1} \tag{17}$$

in which

$$P = [\mathbf{v}_1 \mid \mathbf{v}_2 \mid \cdots \mid \mathbf{v}_n] \tag{18}$$

is the transition matrix from B to the standard basis. In the special case where B is an orthonormal basis the matrix P is orthogonal, so (17) *is of the form*

$$[T] = P[T]_B P^T \tag{19}$$

When convenient, Formula (17) can be rewritten as

$$[T]_B = P^{-1}[T]P \tag{20}$$

and in the case where B is an orthonormal basis this equation can be expressed as

$$[T]_B = P^T[T]P \tag{21}$$

Formula (17) [or (19) in the orthogonal case] tells us that the process of changing from the standard basis for R^n to a basis B produces a factorization of the standard matrix for T as

$$[T] = P[T]_B P^{-1} \tag{22}$$

in which P is the transition matrix from the basis B to the standard basis S. To understand the geometric significance of this factorization, let us use it to compute $T(\mathbf{x})$ by writing

$$T(\mathbf{x}) = [T]\mathbf{x} = (P[T]_B P^{-1})\mathbf{x} = P[T]_B(P^{-1}\mathbf{x})$$

Reading from right to left, this equation tells us that $T(\mathbf{x})$ can be obtained by first mapping the standard coordinates of $\mathbf{x}$ to B-coordinates using the matrix P^{-1}, then performing the operation on the B-coordinates using the matrix $[T]_B$, and then using the matrix P to map the resulting vector back to standard coordinates.

EXAMPLE 3 Example 1 Revisited from the Viewpoint of Factorization

In Example 1 we considered the linear operator $T: R^2 \to R^2$ whose standard matrix is

$$A = [T] = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$$

and we showed that

$$[T]_B = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}$$

with respect to the orthonormal basis $B = \{\mathbf{v}_1, \mathbf{v}_2\}$ that is formed from the vectors

$$\mathbf{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

In this case the transition matrix from B to S is

$$P = [\mathbf{v}_1 \mid \mathbf{v}_2] = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

so it follows from (17) that $[T]$ can be factored as

$$\underset{[T]}{\begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}} = \underset{P}{\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}} \underset{[T]_B}{\begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}} \underset{P^{-1}}{\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}}$$

Reading from right to left, this equation tells us that $T(\mathbf{x})$ can be computed by first transforming standard coordinates to B-coordinates, then stretching the $\mathbf{v}_2$-coordinate by a factor of 5 while leaving the $\mathbf{v}_1$-coordinate fixed, and then transforming B-coordinates back to standard coordinates. ■

EXAMPLE 4 Example 2 Revisited from the Viewpoint of Factorization

In Example 2 we considered the rotation $T : R^3 \to R^3$ whose standard matrix is

$$A = [T] = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

and we showed that

$$[T]_B = \begin{bmatrix} \cos \frac{2\pi}{3} & -\sin \frac{2\pi}{3} & 0 \\ \sin \frac{2\pi}{3} & \cos \frac{2\pi}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

with respect to any orthonormal basis $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ in which $\mathbf{v}_3 = \mathbf{v}_1 \times \mathbf{v}_2$ is a positive multiple of the vector $\mathbf{n} = (1, 1, 1)$ along the axis of rotation and $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthonormal basis for the plane W that passes through the origin and is perpendicular to the axis of rotation.

To find a specific basis of this form, recall from Example 7 of Section 6.2 that the equation of the plane W is

$$x + y + z = 0$$

and recall from Example 10 of Section 7.9 that the vectors

$$\mathbf{v}_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}$$

form an orthonormal basis for W. Since

$$\mathbf{v}_3 = \mathbf{v}_1 \times \mathbf{v}_2 = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{vmatrix} = \frac{1}{\sqrt{3}}\mathbf{i} + \frac{1}{\sqrt{3}}\mathbf{j} + \frac{1}{\sqrt{3}}\mathbf{k}$$

the transition matrix from $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ to the standard basis is

$$P = [\mathbf{v}_1 \mid \mathbf{v}_2 \mid \mathbf{v}_3] = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

why?

Since this matrix is orthogonal, it follows from (19) that $[T]$ can be factored as

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} \cos\frac{2\pi}{3} & -\sin\frac{2\pi}{3} & 0 \\ \sin\frac{2\pi}{3} & \cos\frac{2\pi}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

$$[T] \quad = \quad P \quad [T]_B \quad P^T$$

Reading from right to left, this tells us that $T(\mathbf{x})$ can be computed by first transforming standard coordinates to B-coordinates, then rotating through an angle of $2\pi/3$ about an axis in the direction of $\mathbf{v}_3$, and then transforming B-coordinates back to standard coordinates. ■

EXAMPLE 5 Factoring the Standard Matrix for a Reflection

Recall from Formula (2) of Section 6.2 that the standard matrix for the reflection T of R^2 about the line L through the origin making an angle θ with the positive x-axis of a rectangular xy-coordinate system is

$$[T] = H_\theta = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}$$

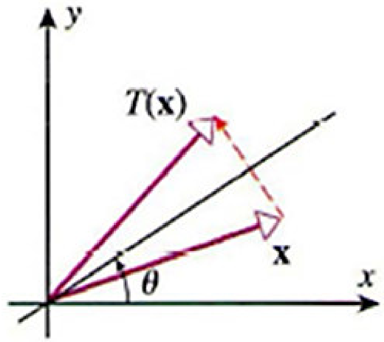


Figure 6.1.9

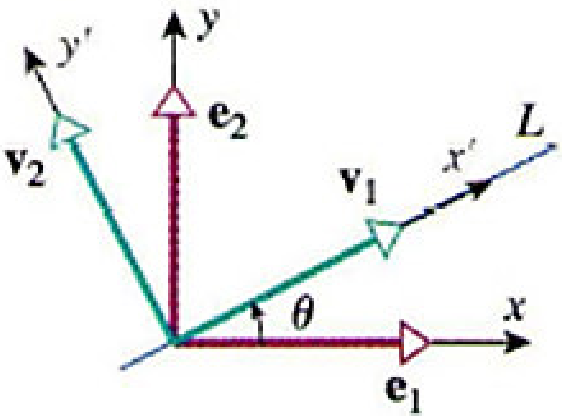
The fact that this matrix represents a reflection is not immediately evident because the standard unit vectors along the positive x- and y-axes have no special relationship to the line L. Suppose, however, that we rotate the coordinate axes through the angle θ to align the new x'-axis with L, and we let $\mathbf{v}_1$ and $\mathbf{v}_2$ be unit vectors along the x'- and y'-axes, respectively (Figure 8.1.4). Since

$$T(\mathbf{v}_1) = \mathbf{v}_1 = \mathbf{v}_1 + 0\mathbf{v}_2 \quad \text{and} \quad T(\mathbf{v}_2) = -\mathbf{v}_2 = 0\mathbf{v}_1 + (-1)\mathbf{v}_2$$

it follows that the matrix for T with respect to the basis $B = \{\mathbf{v}_1, \mathbf{v}_2\}$ is

$$[T]_B = [[T(\mathbf{v}_1)]_B \mid [T(\mathbf{v}_2)]_B] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Figure 8.1.4



Also, it follows from Example 8 of Section 7.11 that the transition matrices between the standard basis S and the basis B are

$$P = P_{B\to S} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad \text{and} \quad P^T = P_{S\to B} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

Thus, Formula (19) implies that

$$\underset{[T]}{\begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}} = \underset{P}{\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}} \underset{[T]_B}{\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}} \underset{P^T}{\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}}$$

Reading from right to left, this equation tells us that $T(\mathbf{x})$ can be computed by first rotating the xy-axes through the angle θ to convert standard coordinates to B-coordinates, then reflecting about the x'-axis, and then rotating through the angle $-\theta$ to convert back to standard coordinates.

■

MATRIX OF A LINEAR TRANSFORMATION WITH RESPECT TO A PAIR OF BASES

Up to now we have focused on matrix representations of linear operators. We will now consider the corresponding idea for linear transformations. Recall that every linear transformation $T: R^n \to R^m$ has an associated $m \times n$ standard matrix

$$[T] = [T(\mathbf{e}_1) \mid T(\mathbf{e}_2) \mid \cdots \mid T(\mathbf{e}_n)]$$

with the property that

$$T(\mathbf{x}) = [T]\mathbf{x}$$

If B and B' are bases for R^n and R^m, respectively, then the transformation

$$\mathbf{x} \xrightarrow{T} T(\mathbf{x})$$

creates an associated transformation

$$[\mathbf{x}]_B \to [T(\mathbf{x})]_{B'}$$

that maps the coordinate matrix $[\mathbf{x}]_B$ into the coordinate matrix $[T(\mathbf{x})]_{B'}$. As in the operator case, this associated transformation is linear and hence must be a matrix transformation; that is, there must be a matrix A such that

$$A[\mathbf{x}]_B = [T(\mathbf{x})]_{B'}$$

The following generalization of Theorem 8.1.1 shows how to find A.

Theorem 8.1.4 *Let $T: R^n \to R^m$ be a linear transformation, let $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $B' = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ be bases for R^n and R^m, respectively, and let*

$$A = \left[[T(\mathbf{v}_1)]_{B'} \mid [T(\mathbf{v}_2)]_{B'} \mid \cdots \mid [T(\mathbf{v}_n)]_{B'}\right] \tag{23}$$

Then

$$[T(\mathbf{x})]_{B'} = A[\mathbf{x}]_B \tag{24}$$

for every vector $\mathbf{x}$ in R^n. Moreover, the matrix A given by Formula (23) *is the only matrix with property* (24).

The matrix A in (23) is called the ***matrix for T with respect to the bases B and B′*** and is denoted by the symbol $[T]_{B',B}$. With this notation Formulas (23) and (24) can be expressed as

$$[T]_{B',B} = \left[[T(\mathbf{v}_1)]_{B'} \mid [T(\mathbf{v}_2)]_{B'} \mid \cdots \mid [T(\mathbf{v}_n)]_{B'}\right] \tag{25}$$

and

$$[T(\mathbf{x})]_{B'} = [T]_{B',B}[\mathbf{x}]_B \tag{26}$$

REMARK Observe the order of the subscripts in the notation $[T]_{B',B}$—the right subscript denotes the basis for the domain and the left subscript denotes the basis for the codomain (Figure 8.1.5). Also, note how the basis for the domain seems to "cancel" in Formula (26) (Figure 8.1.6).

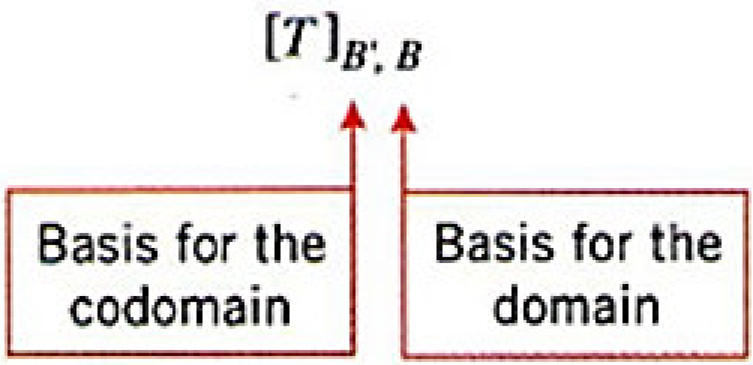


Figure 8.1.5

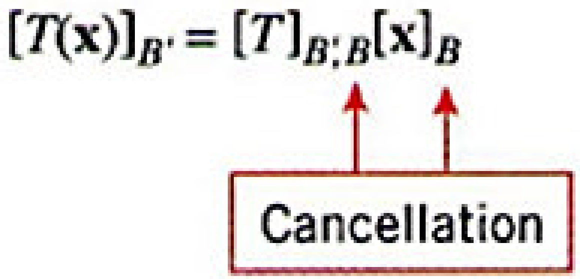


Figure 8.1.6

Recalling that the components of a vector in R^n or R^m are the same as its coordinates with respect to the standard basis for that space, it follows from (25) that if $S = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is the standard basis for R^n and S' is the standard basis for R^m, then

$$[T]_{S',S} = \big[[T(\mathbf{e}_1)]_{S'} \mid [T(\mathbf{e}_2)]_{S'} \mid \cdots \mid [T(\mathbf{e}_n)]_{S'}\big] = [T(\mathbf{e}_1) \mid T(\mathbf{e}_2) \mid \cdots \mid T(\mathbf{e}_n)] = [T]$$

That is, the matrix for a linear transformation from R^n to R^m with respect to the standard bases for those spaces is the same as the standard matrix for T.

EXAMPLE 6 Matrix of a Linear Transformation

Let $T : R^2 \to R^3$ be the linear transformation defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_2 \\ -5x_1 + 13x_2 \\ -7x_1 + 16x_2 \end{bmatrix}$$

Find the matrix for T with respect to the bases $B = \{\mathbf{v}_1, \mathbf{v}_2\}$ for R^2 and $B' = \{\mathbf{v}'_1, \mathbf{v}'_2, \mathbf{v}'_3\}$ for R^3, where

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 5 \\ 2 \end{bmatrix}; \quad \mathbf{v}'_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \mathbf{v}'_2 = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}, \quad \mathbf{v}'_3 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Solution Using the given formula for T we obtain

$$T(\mathbf{v}_1) = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}, \quad T(\mathbf{v}_2) = \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}$$

by solving linear systems

(verify), and expressing these vectors as linear combinations of $\mathbf{v}'_1$, $\mathbf{v}'_2$, and $\mathbf{v}'_3$ we obtain (verify)

$$T(\mathbf{v}_1) = -\mathbf{v}'_2 - \tfrac{3}{2}\mathbf{v}'_3 \quad \text{and} \quad T(\mathbf{v}_2) = \tfrac{5}{2}\mathbf{v}'_1 + \tfrac{1}{2}\mathbf{v}'_2 - \tfrac{3}{4}\mathbf{v}'_3$$

Thus,

$$[T]_{B',B} = \big[[T(\mathbf{v}_1)]_{B'} \mid [T(\mathbf{v}_2)]_{B'}\big] = \begin{bmatrix} 0 & \frac{5}{2} \\ -1 & \frac{1}{2} \\ -\frac{3}{2} & -\frac{3}{4} \end{bmatrix}$$

■

EFFECT OF CHANGING BASES ON MATRICES OF LINEAR TRANSFORMATIONS

Theorems 8.1.2 and 8.1.3 and the related factorizations all have analogs for linear transformations. For example, suppose that B_1 and B_2 are bases for R^n, that B_1' and B_2' are bases for R^m, that U is the transition matrix from B_2 to B_1, and that V is the transition matrix from B_2' to B_1'. Then the diagram in Figure 8.1.7 suggests that

$$[T]_{B_1',B_1} = V[T]_{B_2',B_2}U^{-1} \tag{27}$$

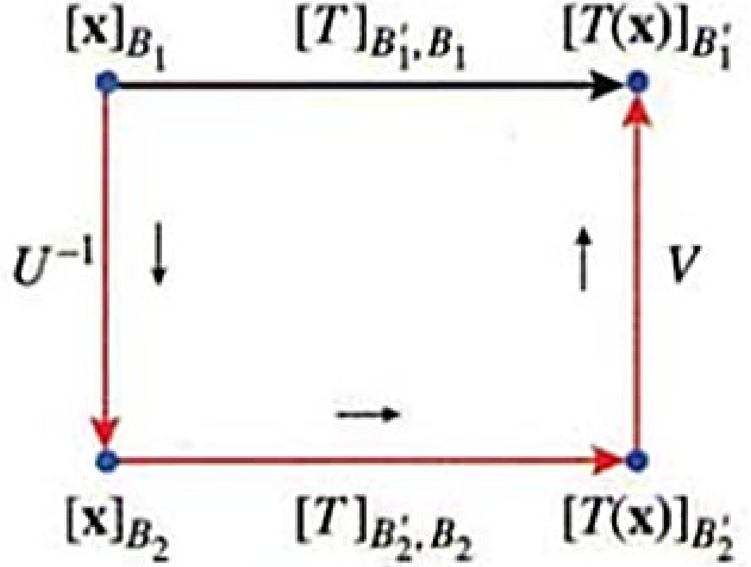


Figure 8.1.7

In particular, if B_1 and B_1' are the standard bases for R^n and R^m, respectively, and if B and B' are any bases for R^n and R^m, respectively, then it follows from (27) that

$$[T] = V[T]_{B',B}U^{-1} \tag{28}$$

where U is the transition matrix from B to the standard basis for R^n and V is the transition matrix from B' to the standard basis for R^m.

Regarding Figure 8.1.7,

the above idea is

$B_1 \rightarrow S$ and $B_1' \rightarrow S'$

$B_2 \rightarrow B$ and $B_2' \rightarrow B'$

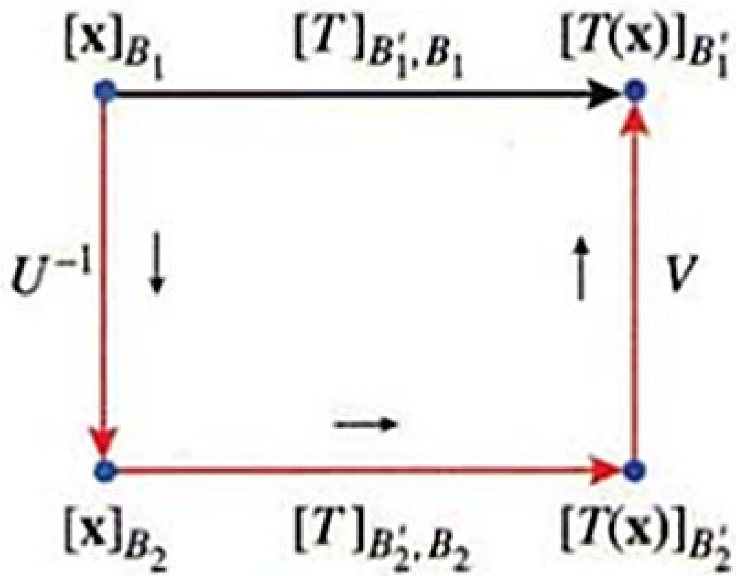


Figure 8.1.7

REPRESENTING LINEAR OPERATORS WITH TWO BASES

A linear operator $T: R^n \rightarrow R^n$ can be viewed as a linear transformation in which the domain and codomain are the same. Thus, instead of choosing a single basis B and representing T by the matrix $[T]_B$, we can choose two different bases for R^n, say B and B', and represent T by the matrix $[T]_{B',B}$. Indeed, we will ask you to show in the exercises that

$$[T]_B = [T]_{B,B}$$

That is, the single-basis representation of T with respect to B can be viewed as the two-basis representation in which both bases are B.

EXAMPLE 7 Matrices of Identity Operators

Recall from Example 5 of Section 6.1 that the operator $T_I(\mathbf{x}) = \mathbf{x}$ that maps each vector in R^n into itself is called the *identity operator* on R^n.

(a) Find the standard matrix for T_I.

(b) Find the matrix for T_I with respect to an arbitrary basis B.

(c) Find the matrix for T_I with respect to a pair of arbitrary bases B and B'.

Solution (a) The standard matrix for T_I is the $n \times n$ identity matrix, since

$$[T_I] = [T_I(\mathbf{e}_1) \mid T_I(\mathbf{e}_2) \mid \cdots \mid T_I(\mathbf{e}_n)] = [\mathbf{e}_1 \mid \mathbf{e}_2 \mid \cdots \mid \mathbf{e}_n] = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

In solution (a), vector $\mathbf{e}_j$ written in terms of the standard basis S is the same $\mathbf{e}_j$.

Solution (b) If $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is any basis for R^n, then

$$[T_I]_B = \big[[T_I(\mathbf{v}_1)]_B \mid [T_I(\mathbf{v}_2)]_B \mid \cdots \mid [T_I(\mathbf{v}_n)]_B\big] = \big[[\mathbf{v}_1]_B \mid [\mathbf{v}_2]_B \mid \cdots \mid [\mathbf{v}_n]_B\big]$$

But for each of these column vectors we have $[\mathbf{v}_i]_B = \mathbf{e}_i$ (why?), so

$$[T_I]_B = [\mathbf{e}_1 \mid \mathbf{e}_2 \mid \cdots \mid \mathbf{e}_n]$$

That is, $[T_I]_B$ is the $n \times n$ identity matrix.

In solution (b), vector $\mathbf{v}_j$ written in terms of basis B is the same $\mathbf{v}_j$.

Solution (c) If $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $B' = \{\mathbf{v}'_1, \mathbf{v}'_2, \ldots, \mathbf{v}'_n\}$ are any bases for R^n, then

$$[T_I]_{B',B} = \big[[T_I(\mathbf{v}_1)]_{B'} \mid [T_I(\mathbf{v}_2)]_{B'} \mid \cdots \mid [T_I(\mathbf{v}_n)]_{B'}\big] = \big[[\mathbf{v}_1]_{B'} \mid [\mathbf{v}_2]_{B'} \mid \cdots \mid [\mathbf{v}_n]_{B'}\big]$$

which is the transition matrix $P_{B \to B'}$ [see Formula (11) of Section 7.11]. ■