GLOBAL EDITION

Linear Algebra with Applications

TENTH EDITION

Steven J. Leon • Lisette de Pillis

Linear Algebra with Applications

Tenth Edition
Global Edition

Steven J. Leon
University of Massachusetts Dartmouth

Lisette G. de Pillis
Harvey Mudd College

Pearson

Harlow, England • London • New York • Boston • San Francisco • Toronto • Sydney • Dubai • Singapore • Hong Kong
Tokyo • Seoul • Taipei • New Delhi • Cape Town • São Paulo • Mexico City • Madrid • Amsterdam • Munich • Paris • Milan

3.5 Change of Basis

Many applied problems can be simplified by changing from one coordinate system to another. Changing coordinate systems in a vector space is essentially the same as changing from one basis to another. For example, in describing the motion of a particle in the plane at a particular time, it is often convenient to use a basis for $\mathbb{R}^2$ consisting of a unit tangent vector $\mathbf{t}$ and a unit normal vector $\mathbf{n}$ instead of the standard basis $\{\mathbf{e}_1, \mathbf{e}_2\}$.

In this section, we discuss the problem of switching from one coordinate system to another. We will show that this can be accomplished by multiplying a given coordinate vector $\mathbf{x}$ by a nonsingular matrix S. The product $\mathbf{y} = S\mathbf{x}$ will be the coordinate vector for the new coordinate system.

Changing Coordinates in $\mathbb{R}^2$

The standard basis for $\mathbb{R}^2$ is $\{\mathbf{e}_1, \mathbf{e}_2\}$. Any vector $\mathbf{x}$ in $\mathbb{R}^2$ can be expressed as a linear combination:

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$$

The scalars x_1 and x_2 can be thought of as the *coordinates* of $\mathbf{x}$ with respect to the standard basis. Actually, for any basis $\{\mathbf{y}, \mathbf{z}\}$ for $\mathbb{R}^2$, it follows from Theorem 3.3.2 that a given vector $\mathbf{x}$ can be represented uniquely as a linear combination:

$$\mathbf{x} = \alpha\mathbf{y} + \beta\mathbf{z}$$

The scalars α and β are the coordinates of $\mathbf{x}$ with respect to the basis $\{\mathbf{y}, \mathbf{z}\}$. Let us order the basis elements so that $\mathbf{y}$ is considered the first basis vector and $\mathbf{z}$ is considered the second, and denote the ordered basis by $[\mathbf{y}, \mathbf{z}]$. We can then refer to the vector $(\alpha, \beta)^T$ as the *coordinate vector* of $\mathbf{x}$ with respect to $[\mathbf{y}, \mathbf{z}]$. Note that, if we reverse the order of the basis vectors and take $[\mathbf{z}, \mathbf{y}]$, then we must also reorder the coordinate vector. The coordinate vector of $\mathbf{x}$ with respect to $[\mathbf{z}, \mathbf{y}]$ will be $(\beta, \alpha)^T$. When we refer to a basis using subscripts, such as $\{\mathbf{u}_1, \mathbf{u}_2\}$, the subscripts assign an ordering to the basis vectors.

EXAMPLE 1 Let $\mathbf{y} = (2, 1)^T$ and $\mathbf{z} = (1, 4)^T$. The vectors $\mathbf{y}$ and $\mathbf{z}$ are linearly independent and hence form a basis for $\mathbb{R}^2$. The vector $\mathbf{x} = (7, 7)^T$ can be written as a linear combination:

$$\mathbf{x} = 3\mathbf{y} + \mathbf{z}$$

Thus, the coordinate vector of $\mathbf{x}$ with respect to $[\mathbf{y}, \mathbf{z}]$ is $(3, 1)^T$. Geometrically, the coordinate vector specifies how to get from the origin to the point $(7, 7)$ by moving first in the direction of $\mathbf{y}$ and then in the direction of $\mathbf{z}$. If, instead, we treat $\mathbf{z}$ as our first basis vector and $\mathbf{y}$ as the second basis vector, then

$$\mathbf{x} = \mathbf{z} + 3\mathbf{y}$$

The coordinate vector of $\mathbf{x}$ with respect to the ordered basis $[\mathbf{z}, \mathbf{y}]$ is $(1, 3)^T$. Geometrically, this vector tells us how to get from the origin to $(7, 7)$ by moving first in the direction of $\mathbf{z}$ and then in the direction of $\mathbf{y}$ (see Figure 3.5.1). ■

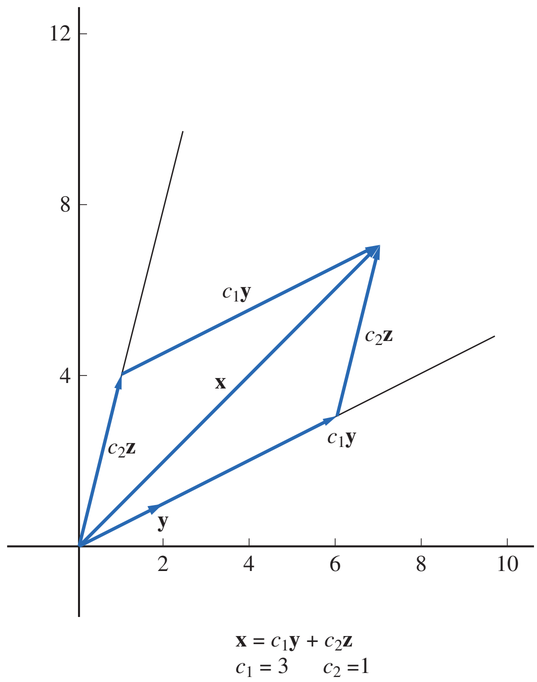

Figure 3.5.1.

As an example of a problem for which it is helpful to change coordinates, consider the following application.

APPLICATION 1 Population Migration

Suppose that the total population of a large metropolitan area remains relatively fixed; however, each year 6 percent of the people living in the city move to the suburbs and 2 percent of the people living in the suburbs move to the city. If, initially, 30 percent of the population lives in the city and 70 percent lives in the suburbs, what will these percentages be in 10 years? 30 years? 50 years? What are the long-term implications?

The changes in population can be determined by matrix multiplications. If we set

$$A = \begin{Bmatrix} 0.94 & 0.02 \\ 0.06 & 0.98 \end{Bmatrix} \quad \text{and} \quad \mathbf{x}_0 = \begin{Bmatrix} 0.30 \\ 0.70 \end{Bmatrix}$$

then the percentages of people living in the city and suburbs after one year can be calculated by setting $\mathbf{x}_1 = A\mathbf{x}_0$. The percentages after two years can be calculated by setting $\mathbf{x}_2 = A\mathbf{x}_1 = A^2\mathbf{x}_0$. In general, the percentages after n years will be given by $\mathbf{x}_n = A^n\mathbf{x}_0$. If we calculate these percentages for $n = 10$, 30, and 50 years and round to the nearest percent, we get

$$\mathbf{x}_{10} = \begin{Bmatrix} 0.27 \\ 0.73 \end{Bmatrix} \qquad \mathbf{x}_{30} = \begin{Bmatrix} 0.25 \\ 0.75 \end{Bmatrix} \qquad \mathbf{x}_{50} = \begin{Bmatrix} 0.25 \\ 0.75 \end{Bmatrix}$$

In fact, as n increases, the sequence of vectors $\mathbf{x}_n = A^n\mathbf{x}_0$ converges to a limit $\mathbf{x} = (0.25, 0.75)^T$. The limit vector $\mathbf{x}$ is called a *steady-state vector* for the process.

To understand why the process approaches a steady state, it is helpful to switch to a different coordinate system. For the new coordinate system, we will pick vectors $\mathbf{u}_1$ and $\mathbf{u}_2$, for which it is easy to see the effect of multiplication by the matrix A. In particular, if we pick $\mathbf{u}_1$ to be any multiple of the steady-state vector $\mathbf{x}$, then $A\mathbf{u}_1$ will equal $\mathbf{u}_1$. Let us choose $\mathbf{u}_1 = (1 \;\; 3)^T$ and $\mathbf{u}_2 = (-1 \;\; 1)^T$. The second vector was chosen because the effect of multiplying by A is just to scale the vector by a factor of 0.92. Thus, our new basis vectors satisfy

$$A\mathbf{u}_1 = \begin{Bmatrix} 0.94 & 0.02 \\ 0.06 & 0.98 \end{Bmatrix} \begin{Bmatrix} 1 \\ 3 \end{Bmatrix} = \begin{Bmatrix} 1 \\ 3 \end{Bmatrix} = \mathbf{u}_1$$

$$A\mathbf{u}_2 = \begin{Bmatrix} 0.94 & 0.02 \\ 0.06 & 0.98 \end{Bmatrix} \begin{Bmatrix} -1 \\ 1 \end{Bmatrix} = \begin{Bmatrix} -0.92 \\ 0.92 \end{Bmatrix} = 0.92\mathbf{u}_2$$

The initial vector $\mathbf{x}_0$ can be written as a linear combination of the new basis vectors:

$$\mathbf{x}_0 = \begin{Bmatrix} 0.30 \\ 0.70 \end{Bmatrix} = 0.25 \begin{Bmatrix} 1 \\ 3 \end{Bmatrix} - 0.05 \begin{Bmatrix} -1 \\ 1 \end{Bmatrix} = 0.25\mathbf{u}_1 - 0.05\mathbf{u}_2$$

It follows that

$$\mathbf{x}_n = A^n\mathbf{x}_0 = 0.25\mathbf{u}_1 - 0.05(0.92)^n\mathbf{u}_2$$

The entries of the second component approach 0 as n gets large. In fact, for $n > 27$, the entries will be small enough so that the rounded values of $\mathbf{x}_n$ are all equal to

$$0.25\mathbf{u}_1 = \begin{Bmatrix} 0.25 \\ 0.75 \end{Bmatrix}$$

This application is an example of a type of mathematical model called a *Markov process*. The sequence of vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots$ is called a *Markov chain*. The matrix A has a special structure in that its entries are nonnegative and its columns all add up to 1. Such matrices are called *stochastic matrices*. More precise definitions will be given later when we study these types of applications in Chapter 6. What we want to stress here is that the key to understanding such processes is to switch to a basis for which the effect of the matrix is quite simple. In particular, if A is $n \times n$, then we will want to choose basis vectors so that the effect of the matrix A on each basis vector $\mathbf{u}_j$ is simply to scale it by some factor λ_j, that is,

$$A\mathbf{u}_j = \lambda_j\mathbf{u}_j \quad j = 1, 2, \ldots, n \tag{1}$$

In many applied problems involving an $n \times n$ matrix A, the key to solving the problem often is to find basis vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$ and scalars $\lambda_1, \ldots, \lambda_n$ such that (1) is satisfied. The new basis vectors can be thought of as a natural coordinate system to use with the matrix A, and the scalars can be thought of as natural frequencies for the basis vectors. We will study these types of applications in more detail in Chapter 6.

Changing Coordinates

Once we have decided to work with a new basis, we have the problem of finding the coordinates with respect to that basis. Suppose, for example, that instead of using the standard basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ for $\mathbb{R}^2$, we wish to use a different basis, say,

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Indeed, we may want to switch back and forth between the two coordinate systems. Let us consider the following two problems:

I. Given a vector $\mathbf{x} = (x_1, x_2)^T$, find its coordinates with respect to $\mathbf{u}_1$ and $\mathbf{u}_2$.

II. Given a vector $c_1\mathbf{u}_1 + c_2\mathbf{u}_2$, find its coordinates with respect to $\mathbf{e}_1$ and $\mathbf{e}_2$.

We will solve **II** first, since it turns out to be the easier problem. To switch bases from $\{\mathbf{u}_1, \mathbf{u}_2\}$ to $\{\mathbf{e}_1, \mathbf{e}_2\}$, we must express the old basis elements $\mathbf{u}_1$ and $\mathbf{u}_2$ in terms of the new basis elements $\mathbf{e}_1$ and $\mathbf{e}_2$.

$$\begin{aligned} \mathbf{u}_1 &= 3\mathbf{e}_1 + 2\mathbf{e}_2 \\ \mathbf{u}_2 &= \mathbf{e}_1 + \mathbf{e}_2 \end{aligned}$$

It follows then that

$$\begin{aligned} c_1\mathbf{u}_1 + c_2\mathbf{u}_2 &= (3c_1\mathbf{e}_1 + 2c_1\mathbf{e}_2) + (c_2\mathbf{e}_1 + c_2\mathbf{e}_2) \\ &= (3c_1 + c_2)\mathbf{e}_1 + (2c_1 + c_2)\mathbf{e}_2 \end{aligned}$$

Thus, the coordinate vector of $c_1\mathbf{u}_1 + c_2\mathbf{u}_2$ with respect to $\{\mathbf{e}_1, \mathbf{e}_2\}$ is

$$\mathbf{x} = \begin{bmatrix} 3c_1 + c_2 \\ 2c_1 + c_2 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

If we set

$$U = (\mathbf{u}_1, \mathbf{u}_2) = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}$$

then, given any coordinate vector $\mathbf{c}$ with respect to $\{\mathbf{u}_1, \mathbf{u}_2\}$, to find the corresponding coordinate vector $\mathbf{x}$ with respect to $\{\mathbf{e}_1, \mathbf{e}_2\}$, we simply multiply U times $\mathbf{c}$:

$$\mathbf{x} = U\mathbf{c} \tag{2}$$

The matrix U is called the *transition matrix* from the ordered basis $\{\mathbf{u}_1, \mathbf{u}_2\}$ to the standard basis $\{\mathbf{e}_1, \mathbf{e}_2\}$.

To solve problem **I**, we must find the transition matrix from $\{\mathbf{e}_1, \mathbf{e}_2\}$ to $\{\mathbf{u}_1, \mathbf{u}_2\}$. The matrix U in (2) is nonsingular, since its column vectors, $\mathbf{u}_1$ and $\mathbf{u}_2$, are linearly independent. It follows from (2) that

$$\mathbf{c} = U^{-1}\mathbf{x}$$

Thus, given a vector

$$\mathbf{x} = (x_1, x_2)^T = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$$

we need only multiply by U^{-1} to find its coordinate vector with respect to $\{\mathbf{u}_1, \mathbf{u}_2\}$. U^{-1} is the transition matrix from $\{\mathbf{e}_1, \mathbf{e}_2\}$ to $\{\mathbf{u}_1, \mathbf{u}_2\}$.

EXAMPLE 2 Let $\mathbf{u}_1 = (3,2)^T$, $\mathbf{u}_2 = (1,1)^T$, and $\mathbf{x} = (7,4)^T$. Find the coordinates of $\mathbf{x}$ with respect to $\mathbf{u}_1$ and $\mathbf{u}_2$.

Solution

By the preceding discussion, the transition matrix from $\{\mathbf{e}_1, \mathbf{e}_2\}$ to $\{\mathbf{u}_1, \mathbf{u}_2\}$ is the inverse of

$$U = (\mathbf{u}_1, \mathbf{u}_2) = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$$

Thus,

$$\mathbf{c} = U^{-1}\mathbf{x} = \begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 7 \\ 4 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$$

is the desired coordinate vector and

$$\mathbf{x} = 3\mathbf{u}_1 - 2\mathbf{u}_2$$ ■

EXAMPLE 3 Let $\mathbf{b}_1 = (1,-1)^T$ and $\mathbf{b}_2 = (-2,3)^T$. Find the transition matrix from $\{\mathbf{e}_1, \mathbf{e}_2\}$ to $\{\mathbf{b}_1, \mathbf{b}_2\}$ and determine the coordinates of $\mathbf{x} = (1,2)^T$ with respect to $\{\mathbf{b}_1, \mathbf{b}_2\}$.

Solution

The transition matrix from $\{\mathbf{b}_1, \mathbf{b}_2\}$ to $\{\mathbf{e}_1, \mathbf{e}_2\}$ is

$$B = (\mathbf{b}_1, \mathbf{b}_2) = \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix}$$

and hence the transition matrix from $\{\mathbf{e}_1, \mathbf{e}_2\}$ to $\{\mathbf{b}_1, \mathbf{b}_2\}$ is

$$B^{-1} = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}$$

The coordinate vector of $\mathbf{x}$ with respect to $\{\mathbf{b}_1, \mathbf{b}_2\}$ is

$$\mathbf{c} = B^{-1}\mathbf{x} = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \end{pmatrix}$$

and hence

$$\mathbf{x} = 7\mathbf{b}_1 + 3\mathbf{b}_2$$ ■

Now let us consider the general problem of changing from one ordered basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ of $\mathbb{R}^2$ to another ordered basis $\{\mathbf{u}_1, \mathbf{u}_2\}$. In this case, we assume that, for a given vector $\mathbf{x}$, its coordinates with respect to $\{\mathbf{v}_1, \mathbf{v}_2\}$ are known:

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

Now we wish to represent $\mathbf{x}$ as a sum $d_1\mathbf{u}_1 + d_2\mathbf{u}_2$. Thus, we must find scalars d_1 and d_2 so that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = d_1\mathbf{u}_1 + d_2\mathbf{u}_2 \tag{3}$$

If we set $V = (\mathbf{v}_1, \mathbf{v}_2)$ and $U = (\mathbf{u}_1, \mathbf{u}_2)$, then equation (3) can be written in matrix form:

$$V\mathbf{c} = U\mathbf{d}$$

It follows that

$$\mathbf{d} = U^{-1}V\mathbf{c}$$

Thus, given a vector $\mathbf{x}$ in $\mathbb{R}^2$ and its coordinate vector $\mathbf{c}$ with respect to the ordered basis $\{\mathbf{v}_1, \mathbf{v}_2\}$, to find the coordinate vector of $\mathbf{x}$ with respect to the new basis $\{\mathbf{u}_1, \mathbf{u}_2\}$, we simply multiply $\mathbf{c}$ by the transition matrix $S = U^{-1}V$.

EXAMPLE 4 Find the transition matrix corresponding to the change of basis from $\{\mathbf{v}_1, \mathbf{v}_2\}$ to $\{\mathbf{u}_1, \mathbf{u}_2\}$, where

$$\mathbf{v}_1 = \begin{pmatrix} 5 \\ 2 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 7 \\ 3 \end{pmatrix} \qquad \text{and} \qquad \mathbf{u}_1 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Solution

The transition matrix from $\{\mathbf{v}_1, \mathbf{v}_2\}$ to $\{\mathbf{u}_1, \mathbf{u}_2\}$ is given by

$$S = U^{-1}V = \begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 5 & 7 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ -4 & -5 \end{pmatrix}$$ ■

The change of basis from $\{\mathbf{v}_1, \mathbf{v}_2\}$ to $\{\mathbf{u}_1, \mathbf{u}_2\}$ can also be viewed as a two-step process. First we change from $\{\mathbf{v}_1, \mathbf{v}_2\}$ to the standard basis, $\{\mathbf{e}_1, \mathbf{e}_2\}$, and then we change from the standard basis to $\{\mathbf{u}_1, \mathbf{u}_2\}$. Given a vector $\mathbf{x}$ in $\mathbb{R}^2$, if $\mathbf{c}$ is the coordinate vector of $\mathbf{x}$ with respect to $\{\mathbf{v}_1, \mathbf{v}_2\}$ and $\mathbf{d}$ is the coordinate vector of $\mathbf{x}$ with respect to $\{\mathbf{u}_1, \mathbf{u}_2\}$, then

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 = d_1\mathbf{u}_1 + d_2\mathbf{u}_2$$

Since V is the transition matrix from $\{\mathbf{v}_1, \mathbf{v}_2\}$ to $\{\mathbf{e}_1, \mathbf{e}_2\}$ and U^{-1} is the transition matrix from $\{\mathbf{e}_1, \mathbf{e}_2\}$ to $\{\mathbf{u}_1, \mathbf{u}_2\}$, it follows that

$$V\mathbf{c} = \mathbf{x} \qquad \text{and} \qquad U^{-1}\mathbf{x} = \mathbf{d}$$

and hence

$$U^{-1}V\mathbf{c} = U^{-1}\mathbf{x} = \mathbf{d}$$

As before, we see that the transition matrix from $\{\mathbf{v}_1, \mathbf{v}_2\}$ to $\{\mathbf{u}_1, \mathbf{u}_2\}$ is $U^{-1}V$ (see Figure 3.5.2).

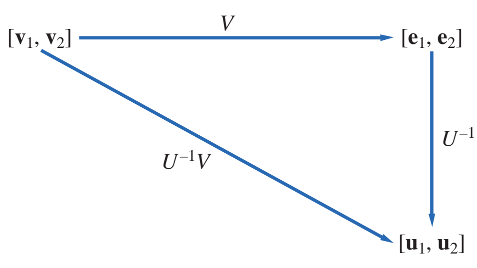

Figure 3.5.2.