2.5 LEARNING SUMMARY

- A *vector* is an ordered set of real numbers. A vector of order ($n \times 1$) is called a *column vector* and is written as follows:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$$

  A vector of order ($1 \times m$) is called a *row vector*. The scalars xa_i are referred to as *components* or *elements* of the vector.

- A matrix is a rectangular array of real numbers arranged into n rows and m columns:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

  - If $m = n$, then the matrix is called a *square matrix*.
  - If $x_{ij} = x_{ji}$ for all elements of a square matrix, then the matrix is called a *symmetric matrix*.
  - If $x_{ij} = 0$ for all off-diagonal elements of a square matrix, then the matrix is called a *diagonal matrix*.
  - If $x_{ii} = 1$ for all the on-diagonal elements of a diagonal matrix, then the matrix is called an *identity matrix* and is usually denoted **I.**

- We often find it useful to organize multivariate data into a matrix. The rows of the matrix correspond to objects (i.e., the entities under study) and the columns correspond to variables (i.e., measured characteristics of the objects).
- The data in a matrix can be displayed in a Cartesian coordinate system. Each row of the matrix is plotted as a point in the coordinate space (where the axes of the space correspond to the columns of the matrix).
- Vector and matrix operations have an underlying geometric interpretation:
  - *Scalar multiplication* has the effect of uniformly stretching or shrinking the data configuration.
  - *Vector multiplication* has the effect of projecting the data points onto a single line defined by the vector.

TABLE 2.1 Heights and weights of 20 women: Raw data, mean-centered data, and standardized data

	X_1	X_2	X_{d1}	X_{d2}	X_{s1}	X_{s2}
1	57	93	−5.85	−30.6	−1.77427	−1.96516
2	58	110	−4.85	−13.6	−1.47098	−0.87341
3	60	99	−2.85	−24.6	−0.86439	−1.57984
4	59	111	−3.85	−12.6	−1.16768	−0.80918
5	61	115	−1.85	−8.6	−0.56109	−0.55230
6	60	122	−2.85	−1.6	−0.86439	−0.10275
7	62	110	−0.85	−13.6	−0.25780	−0.87341
8	61	116	−1.85	−7.6	−0.56109	−0.48808
9	62	122	−0.85	−1.6	−0.25780	−0.10275
10	63	128	0.15	4.4	0.04549	0.28257
11	62	134	−0.85	10.4	−0.25780	0.66790
12	64	117	1.15	−6.6	0.34879	−0.42386
13	63	123	0.15	−0.6	0.04549	−0.03853
14	65	129	2.15	5.4	0.65208	0.34679
15	64	135	1.15	11.4	0.34879	0.73212
16	66	128	3.15	4.4	0.95538	0.28257
17	67	135	4.15	11.4	1.25867	0.73212
18	66	148	3.15	24.4	0.95538	1.56699
19	68	142	5.15	18.4	1.56197	1.18167
20	69	155	6.15	31.4	1.86526	2.01654
Mean	62.85	123.6	0.0	0.0	0.0	0.0
Stdev	3.3	15.5	3.3	15.5	1.0	1.0

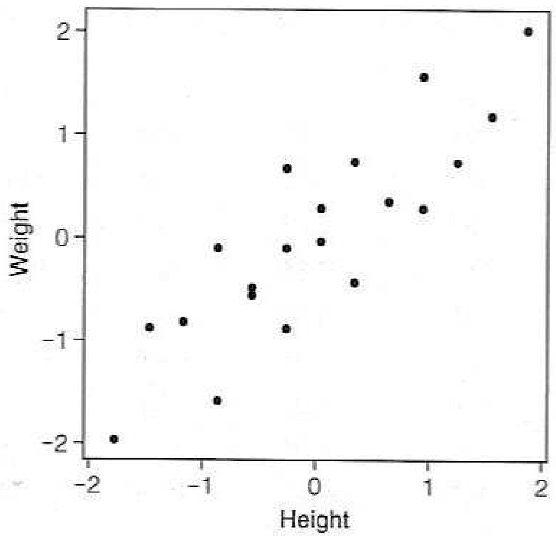

FIGURE 2.3 Scatter plot of weight versus height for 20 women

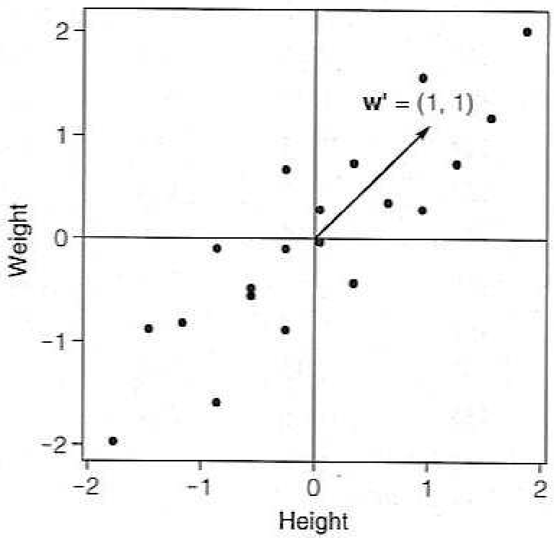

FIGURE 2.4 Vector **w**' = (1, 1) superimposed on scatter plot of weight versus height for 20 women

2.5 LEARNING SUMMARY

- A *vector* is an ordered set of real numbers. A vector of order ($n \times 1$) is called a *column vector* and is written as follows:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$$

  A vector of order ($1 \times m$) is called a *row vector*. The scalars xa_i are referred to as *components* or *elements* of the vector.

- A matrix is a rectangular array of real numbers arranged into n rows and m columns:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

  - If $m = n$, then the matrix is called a *square matrix*.
  - If $x_{ij} = x_{ji}$ for all elements of a square matrix, then the matrix is called a *symmetric matrix*.
  - If $x_{ij} = 0$ for all off-diagonal elements of a square matrix, then the matrix is called a *diagonal matrix*.
  - If $x_{ii} = 1$ for all the on-diagonal elements of a diagonal matrix, then the matrix is called an *identity matrix* and is usually denoted **I**.

- We often find it useful to organize multivariate data into a matrix. The rows of the matrix correspond to objects (i.e., the entities under study) and the columns correspond to variables (i.e., measured characteristics of the objects).
- The data in a matrix can be displayed in a Cartesian coordinate system. Each row of the matrix is plotted as a point in the coordinate space (where the axes of the space correspond to the columns of the matrix).
- Vector and matrix operations have an underlying geometric interpretation:
  - *Scalar multiplication* has the effect of uniformly stretching or shrinking the data configuration.
  - *Vector multiplication* has the effect of projecting the data points onto a single line defined by the vector.

$$\mathbf{z}_1 = \mathbf{X}_s\mathbf{w}$$

where $\mathbf{X}_s$ is a matrix of order (20 × 2), $\mathbf{w}$ is a column vector of
is a column vector of order (20 × 1). The vector $\mathbf{z}_1$ contains th
each of the 20 women in the sample, which are effectively the p
the points onto the directed line segment given by the vector w
2.8. We can now create a new variable, Z_1, that takes on the va

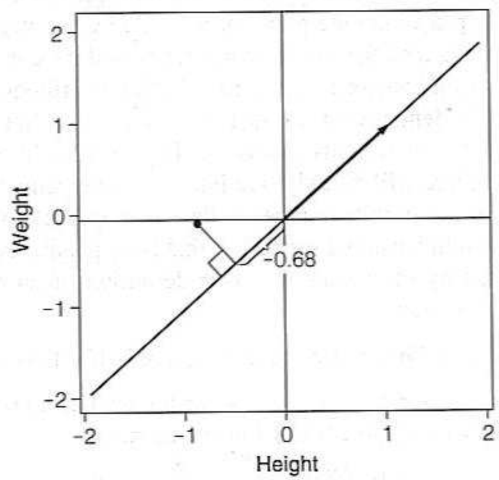

FIGURE 2.7
Projection of point (−.86, −.10) onto axis defined by vector $\mathbf{w}' = (.707, .707)$

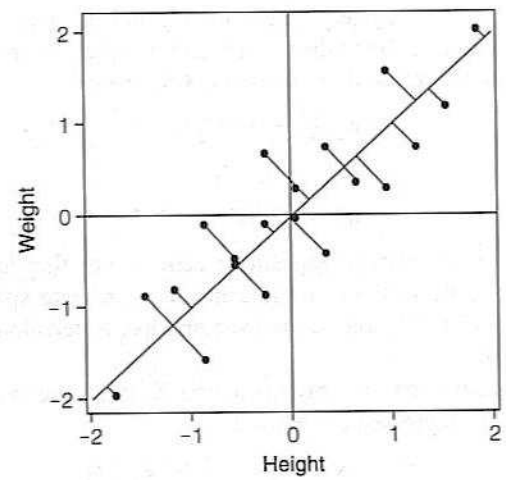

FIGURE 2.8
New variable $\mathbf{z}_1 = \mathbf{X}_s\,\mathbf{w}$ (points projected onto axis defined by $\mathbf{w}$)

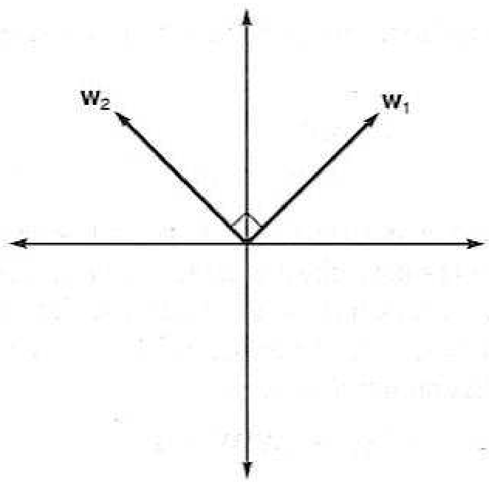

FIGURE 2.9 Orthogonal vectors $\mathbf{w}_1' = (.707, .707)$ and $\mathbf{w}_2' = (-.707, .707)$

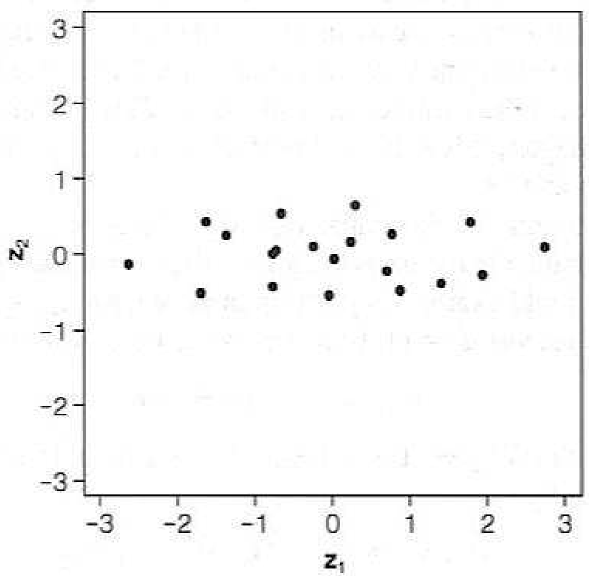

FIGURE 2.10 Plot of $\mathbf{z}_1$ and $\mathbf{z}_2$ (from orthogonal rotation of $\mathbf{x}_1$ and $\mathbf{x}_2$)

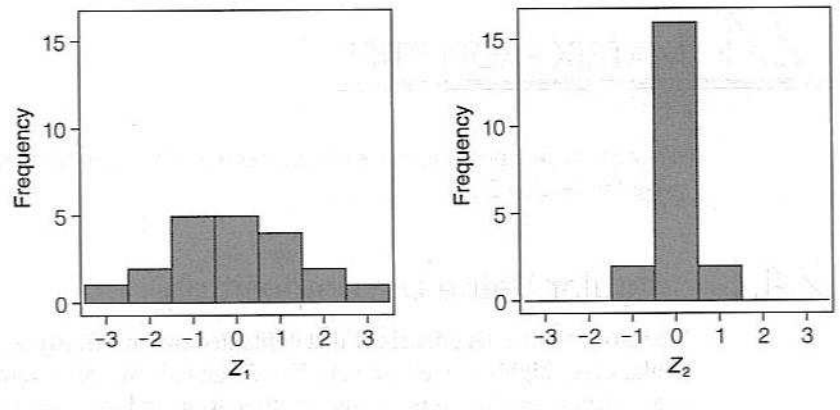

FIGURE 2.11
Histograms of Z_1 and Z_2 (from orthogonal rotation of X_1 and X_2)

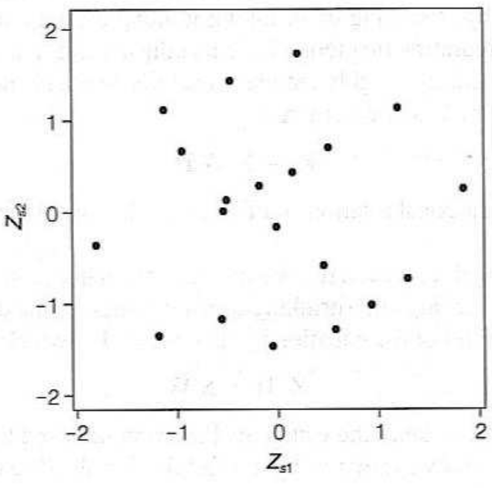

FIGURE 2.12
Scatter plot of standardized variables Z_{s1} and Z_{s2}

2.4.1 Singular Value Decomposition

We started with a data matrix **X** that contained two arbitrarily scaled and (in this particular case) highly correlated variables X_1 and X_2. We discovered that we could use one matrix operation to rotate the configuration and another to change the shape of the configuration by stretching or shrinking it along different axes. As a result, we began with a configuration that looked like an ellipse tilted at a 45 degree angle, and ended up with something roughly circular in nature. In matrix notation, the transformations we performed can be written as

$$\mathbf{Z}_s = \mathbf{X}\,\mathbf{W}\,\mathbf{D}^{-1} \qquad \textbf{(2.19)}$$

where **W** is our orthogonal rotation and $\mathbf{D}^{-1}$ is our diagonal stretching and shrinking transformation.

With a little algebra, we can rewrite the expression in equation (2.19) in terms of $\mathbf{Z}_s$. To undo the stretching and shrinking transformation of the diagonal matrix $\mathbf{D}^{-1}$, we multiply both sides of the equation by its inverse (**D**), which yields

$$\mathbf{Z}_s\,\mathbf{D} = \mathbf{X}\,\mathbf{W} \qquad \textbf{(2.20)}$$

Recall that we can undo the orthogonal rotation imposed by **W** by rotating the configuration the other way; that is, by multiplying by **W′**. The result leaves us with

$$\mathbf{X} = \mathbf{Z}_s\,\mathbf{D}\,\mathbf{W}' \qquad \textbf{(2.21)}$$

What we have shown does not constitute a formal proof, but the intuition is important. What equation (2.21) says is that any data matrix **X** can be decomposed into three component parts: a matrix of variables ($\mathbf{Z}_s$) that are uncorrelated and that have unit variance, a stretching and shrinking transformation (**D**), and an orthogonal rotation (**W′**). The process of finding these three components is called *singular value decomposition* (or SVD).

$$\mathrm{cov}(X_i, X_j) = \frac{1}{(n-1)} \sum_{k=1}^{n} (x_{ki} - \bar{x}_{.i})(x_{kj} - \bar{x}_{.j}) \qquad \textbf{(2.23)}$$

When the data are standardized, $\mathrm{var}(X_i) = 1$ for all variables X_i, and the covariances are interpretable as correlations (constrained to vary between -1 and 1).

If we create the matrix $\mathbf{X}_d$ to contain mean-centered data (i.e., the variables X_{d1} and X_{d2} from the third and fourth columns of Table 2.1), then we can represent the co-variance matrix using the following matrix notation:

$$\mathbf{S} = 1/(n-1)\, \mathbf{X}_d' \mathbf{X}_d \qquad \textbf{(2.24)}$$

The expression in equation (2.24) is sometimes called a *matrix cross-product*. If we were to calculate the matrix cross-product for our sample data (for nonstandardized values), the results would be:

$$\mathbf{S} = \begin{bmatrix} 10.87 & 44.51 \\ 44.51 & 242.42 \end{bmatrix}$$

The correlation between the variables X_{d1} and X_{d2} implied by these numbers is $44.51/\sqrt{10.87 \times 242.42} = 0.867$, which is quite high.

2.5 LEARNING SUMMARY

- A *vector* is an ordered set of real numbers. A vector of order $(n \times 1)$ is called a *column vector* and is written as follows:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$$

  A vector of order $(1 \times m)$ is called a *row vector.* The scalars xa_i are referred to as *components* or *elements* of the vector.

- A matrix is a rectangular array of real numbers arranged into n rows and m columns:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix}$$

  - If $m = n$, then the matrix is called a *square matrix.*
  - If $x_{ij} = x_{ji}$ for all elements of a square matrix, then the matrix is called a *symmetric matrix.*
  - If $x_{ij} = 0$ for all off-diagonal elements of a square matrix, then the matrix is called a *diagonal matrix.*
  - If $x_{ii} = 1$ for all the on-diagonal elements of a diagonal matrix, then the matrix is called an *identity matrix* and is usually denoted **I.**
- We often find it useful to organize multivariate data into a matrix. The rows of the matrix correspond to objects (i.e., the entities under study) and the columns correspond to variables (i.e., measured characteristics of the objects).
- The data in a matrix can be displayed in a Cartesian coordinate system. Each row of the matrix is plotted as a point in the coordinate space (where the axes of the space correspond to the columns of the matrix).
- Vector and matrix operations have an underlying geometric interpretation:
  - *Scalar multiplication* has the effect of uniformly stretching or shrinking the data configuration.
  - *Vector multiplication* has the effect of projecting the data points onto a single line defined by the vector.

- *Matrix multiplication* has the effect of rotating the configuration onto a new set of axes. In some special cases (where the columns of the matrix are orthogonal), the rotation preserves the overall shape of the configuration.
- Any data matrix can be represented as the product of three component parts: a matrix of variables that are uncorrelated with the same variance $\mathbf{Z}_s$, a stretching and shrinking transformation matrix $\mathbf{D}$, and an orthogonal rotation matrix $\mathbf{W}'$. This result, called a *singular value decomposition* (SVD), is summarized as follows:

$$\mathbf{X} = \mathbf{Z}_s \, \mathbf{D} \, \mathbf{W}'$$

- The variances and covariances in a set of multivariate data can be summarized by the following cross-product matrix:

$$1/(n-1)\mathbf{X}_d' \, \mathbf{X}_d$$

where $\mathbf{X}_d$ is a matrix containing the mean-differenced data (i.e., each column has zero mean).

- The *determinant* of the cross-product matrix is a measure of the generalized variance in the data.