The hypothetical factors can be isolated mathematically using a method known as *principal component analysis*. The basic idea is to form a matrix X of deviations from the mean and then factor it into a product UW, where the columns of U correspond to the hypothetical factors. While in practice, the columns of X are positively correlated, the hypothetical factors should be uncorrelated. Thus, the column vectors of U should be mutually orthogonal (i.e., $\mathbf{u}_i^T\mathbf{u}_j = 0$ whenever $i \neq j$). The entries in each column of U measure how well the individual students exhibit the particular intellectual ability represented by that column. The matrix W measures to what extent each test depends on the hypothetical factors.

The construction of the principal component vectors relies on the covariance matrix $S = \frac{1}{n-1}X^TX$. Since it depends on the *eigenvalues* and *eigenvectors* of S, we will defer the details of the method until Chapter 6. In Section 5 of Chapter 6 we will revisit this application and learn an important factorization called the *singular value decomposition*, which is the main tool of principal component analysis.

References

1. Spearman, C., "'General Intelligence', Objectively Determined and Measured," *American Journal of Psychology*, **15**, 1904.
2. Hotelling, H., "Analysis of a Complex of Statistical Variables in Principal Components," *Journal of Educational Psychology*, **26**, 1933.
3. Maxwell, A. E., *Multivariate Analysis in Behavioral Research*, Chapman and Hall, London, 1977.

SECTION 5.1 EXERCISES

1. Find the angle between the vectors $\mathbf{v}$ and $\mathbf{w}$ in each of the following:

(a) $\mathbf{v} = (2, 1, 3)^T,\ \mathbf{w} = (6, 3, 9)^T$

(b) $\mathbf{v} = (2, -3)^T,\ \mathbf{w} = (3, 2)^T$

(c) $\mathbf{v} = (4, 1)^T,\ \mathbf{w} = (3, 2)^T$

(d) $\mathbf{v} = (-2, 3, 1)^T,\ \mathbf{w} = (1, 2, 4)^T$

2. For each pair of vectors in Exercise 1, find the scalar projection of $\mathbf{v}$ onto $\mathbf{w}$. Also find the vector projection of $\mathbf{v}$ onto $\mathbf{w}$.

3. For each of the following pairs of vectors $\mathbf{x}$ and $\mathbf{y}$, find the vector projection $\mathbf{p}$ of $\mathbf{x}$ onto $\mathbf{y}$ and verify that $\mathbf{p}$ and $\mathbf{x} - \mathbf{p}$ are orthogonal:

(a) $\mathbf{x} = (3, 4)^T,\ \mathbf{y} = (1, 0)^T$

(b) $\mathbf{x} = (3, 5)^T,\ \mathbf{y} = (1, 1)^T$

(c) $\mathbf{x} = (2, 4, 3)^T,\ \mathbf{y} = (1, 1, 1)^T$

(d) $\mathbf{x} = (2, -5, 4)^T,\ \mathbf{y} = (1, 2, -1)^T$

4. Let $\mathbf{x}$ and $\mathbf{y}$ be linearly independent vectors in $\mathbb{R}^2$. If $\|\mathbf{x}\| = 2$ and $\|\mathbf{y}\| = 3$, what, if anything, can we conclude about the possible values of $|\mathbf{x}^T\mathbf{y}|$?

5. Find the point on the line $y = 2x$ that is closest to the point $(5, 2)$.

6. Find the point on the line $y = 2x + 1$ that is closest to the point $(5, 2)$.

7. Find the distance from the point $(1, 2)$ to the line $4x - 3y = 0$.

8. In each of the following, find the equation of the plane normal to the given vector $\mathbf{N}$ and passing through the point P_0:

(a) $\mathbf{N} = (2, 4, 3)^T,\ P_0 = (0, 0, 0)$

(b) $\mathbf{N} = (-3, 6, 2)^T,\ P_0 = (4, 2, -5)$

(c) $\mathbf{N} = (0, 0, 1)^T,\ P_0 = (3, 2, 4)$

9. Find the equation of the plane that passes through the points

$$P_1 = (2, 3, 1),\quad P_2 = (5, 4, 3),\quad P_3 = (3, 4, 4)$$

10. Find the distance from the point $(1, 1, 1)$ to the plane $2x + 2y + z = 0$.

EXAMPLE 5 Let $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}$. $R(A^T)$ is spanned by $\mathbf{e}_1$ and $\mathbf{e}_2$, and $N(A)$ is spanned by $\mathbf{e}_3$. Any vector $\mathbf{x} \in \mathbb{R}^3$ can be written as a sum

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

where

$$\mathbf{y} = (x_1, x_2, 0)^T \in R(A^T) \qquad \text{and} \qquad \mathbf{z} = (0, 0, x_3)^T \in N(A)$$

If we restrict ourselves to vectors $\mathbf{y} \in R(A^T)$, then

$$\mathbf{y} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix} \rightarrow A\mathbf{y} = \begin{bmatrix} 2x_1 \\ 3x_2 \end{bmatrix}$$

In this case, $R(A) = \mathbb{R}^2$ and the inverse transformation from $R(A)$ to $R(A^T)$ is defined by

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \rightarrow \begin{bmatrix} \frac{1}{2}b_1 \\ \frac{1}{3}b_2 \\ 0 \end{bmatrix}$$

■

SECTION 5.2 EXERCISES

1. For each of the following matrices, determine a basis for each of the subspaces $R(A^T)$, $N(A)$, $R(A)$, and $N(A^T)$:

(a) $A = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$ **(b)** $A = \begin{bmatrix} 1 & 3 & 1 \\ 2 & 4 & 0 \end{bmatrix}$

(c) $A = \begin{bmatrix} 4 & -2 \\ 1 & 3 \\ 2 & 1 \\ 3 & 4 \end{bmatrix}$ **(d)** $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 2 & 2 \end{bmatrix}$

2. Let S be the subspace of $\mathbb{R}^3$ spanned by $\mathbf{x} = (1, -1, 1)^T$.

(a) Find a basis for $S^\perp$.

(b) Give a geometrical description of S and $S^\perp$.

3. (a) Let S be the subspace of $\mathbb{R}^3$ spanned by the vectors $\mathbf{x} = (x_1, x_2, x_3)^T$ and $\mathbf{y} = (y_1, y_2, y_3)^T$. Let

$$A = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix}$$

Show that $S^\perp = N(A)$.

(b) Find the orthogonal complement of the subspace of $\mathbb{R}^3$ spanned by $(1, 2, 1)^T$ and $(1, -1, 2)^T$.

4. Let S be the subspace of $\mathbb{R}^4$ spanned by $\mathbf{x}_1 = (1, 0, -2, 1)^T$ and $\mathbf{x}_2 = (0, 1, 3, -2)^T$. Find a basis for $S^\perp$.

5. Let A be a 3×2 matrix with rank 2. Give geometric descriptions of $R(A)$ and $N(A^T)$, and describe geometrically how the subspaces are related.

6. Is it possible for a matrix to have the vector $(3, 1, 2)$ in its row space and $(2, 1, 1)^T$ in its null space? Explain.

7. Let $\mathbf{a}_j$ be a nonzero column vector of an $m \times n$ matrix A. Is it possible for $\mathbf{a}_j$ to be in $N(A^T)$? Explain.

8. Let S be the subspace of $\mathbb{R}^n$ spanned by the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$. Show that $\mathbf{y} \in S^\perp$ if and only if $\mathbf{y} \perp \mathbf{x}_i$ for $i = 1, \ldots, k$.

9. If A is an $m \times n$ matrix of rank r, what are the dimensions of $N(A)$ and $N(A^T)$? Explain.

10. Prove Corollary 5.2.5.

11. Prove: If A is an $m \times n$ matrix and $\mathbf{x} \in \mathbb{R}^n$, then either $A\mathbf{x} = \mathbf{0}$ or there exists $\mathbf{y} \in R(A^T)$ such that $\mathbf{x}^T\mathbf{y} \neq 0$. Draw a picture similar to Figure 5.2.2 to illustrate this result geometrically for the case where $N(A)$ is a two-dimensional subspace of $\mathbb{R}^3$.

12. Let A be an $m \times n$ matrix. Explain why the following are true.

(a) Any vector $\mathbf{x}$ in $\mathbb{R}^n$ can be uniquely written as a sum $\mathbf{y} + \mathbf{z}$, where $\mathbf{y} \in N(A)$ and $\mathbf{z} \in R(A^T)$.

(b) Any vector $\mathbf{b} \in \mathbb{R}^m$ can be uniquely written as a sum $\mathbf{u} + \mathbf{v}$, where $\mathbf{u} \in N(A^T)$ and $\mathbf{v} \in R(A)$.

13. Let A be an $m \times n$ matrix. Show that

(a) if $\mathbf{x} \in N(A^TA)$, then $A\mathbf{x}$ is in both $R(A)$ and $N(A^T)$.

(b) $N(A^TA) = N(A)$.

(c) A and A^TA have the same rank.

(d) if A has linearly independent columns, then A^TA is nonsingular.

14. Let A be an $m \times n$ matrix, B an $n \times r$ matrix, and $C = AB$. Show that

(a) $N(B)$ is a subspace of $N(C)$.

(b) $N(C)^\perp$ is a subspace of $N(B)^\perp$ and, consequently, $R(C^T)$ is a subspace of $R(B^T)$.

15. Let U and V be subspaces of a vector space W. Show that if $W = U \oplus V$, then $U \cap V = \{\mathbf{0}\}$.

16. Let A be an $m \times n$ matrix of rank r and let $\{\mathbf{x}_1, \ldots, \mathbf{x}_r\}$ be a basis for $R(A^T)$. Show that $\{A\mathbf{x}_1, \ldots, A\mathbf{x}_r\}$ is a basis for $R(A)$.

17. Let $\mathbf{x}$ and $\mathbf{y}$ be linearly independent vectors in $\mathbb{R}^n$ and let $S = \text{Span}(\mathbf{x}, \mathbf{y})$. We can use $\mathbf{x}$ and $\mathbf{y}$ to define a matrix A by setting

$$A = \mathbf{x}\mathbf{y}^T + \mathbf{y}\mathbf{x}^T$$

(a) Show that A is symmetric.

(b) Show that $N(A) = S^\perp$.

(c) Show that the rank of A must be 2.

5.3 Least Squares Problems

A standard technique in mathematical and statistical modeling is to find a *least squares* fit to a set of data points in the plane. The least squares curve is usually the graph of a standard type of function, such as a linear function, a polynomial, or a trigonometric polynomial. Since the data may include errors in measurement or experiment-related inaccuracies, we do not require the curve to pass through all the data points. Instead, we require the curve to provide an optimal approximation in the sense that the sum of squares of errors between the y values of the data points and the corresponding y values of the approximating curve are minimized.

The technique of least squares was developed independently by Adrien-Marie Legendre and Carl Friedrich Gauss. The first paper on the subject was published by Legendre in 1806, although there is clear evidence that Gauss had discovered it as a student nine years prior to Legendre's paper and had used the method to do astronomical calculations. Figure 5.3.1 is a portrait of Gauss.

Figure 5.3.1. Carl Friedrich Gauss