CHAPTER 3

Vector and Matrix Concepts from a Geometric Viewpoint

3.1 INTRODUCTION

This chapter is, in part, designed to provide conceptual background for many of the vector and matrix operations described in Chapter 2. Here we are interested in "what goes on" when a scalar product, for example, is computed. Since geometry often provides a direct intuitive appeal to one's understanding, liberal use is made of diagrams and geometric reasoning.

To set the stage for the various geometric descriptions to come, we define a Euclidean space—the cornerstone of most multivariate procedures. This provides the setting in which point representations of vectors and such related concepts as vector length and angle are described. The operations of vector addition, subraction, multiplication by a scalar, and scalar product are then portrayed geometrically.

We next turn to a discussion of the meaning of linear independence and the dimensionality of a vector space. The concept of a *basis* of a vector space is described, and the process by which a basis can be changed is also illustrated geometrically. Special kinds of bases—orthogonal and orthonormal—are illustrated, as well as the Gram-Schmidt process of orthonormalizing an arbitrary basis. Some comments are also made regarding general (oblique) Cartesian coordinate systems.

Our discussion then turns to one of the most common types of transformations—orthogonal transformations (i.e., rotations) of axes. These transformations are portrayed in terms of simple geometric figures and also serve as illustrations of matrix multiplication in the context of Chapter 2.

We conclude the chapter with a geometric description of some commonly used association measures, such as covariance and correlation. Moreover, the idea of viewing a determinant of a matrix of association coefficients (e.g., a covariance or a correlation matrix) as a generalized scalar measure of dispersion is also described geometrically and tied in with counterpart material that has already been covered in Chapter 2. In brief, presentation of the material in this chapter covers some of the same ground discussed in Chapter 2. Here, however, our emphasis is on the *geometry* rather than the algebra of vectors.[1]

[1] In this chapter (and succeeding chapters as well) we shall typically present the material in terms of row vectors $\mathbf{a}'$, $\mathbf{b}'$, etc., particularly when *explicit* forms of the vectors are used, such as $\mathbf{a}' = (1, 2, 2)$. This is strictly to conserve on space. The reader should get used to moving back and forth between column vectors (as emphasized in Chapter 2) and row vectors as emphasized here.

3.2 EUCLIDEAN SPACE AND RECTANGULAR CARTESIAN COORDINATES

Before moving right into a discussion of the geometric aspects of vectors, it is useful to establish some preliminaries, even though they may be familiar to many readers. These preliminaries involve the construction of a coordinate system and a description of standard basis vectors.

3.2.1 Coordinate Systems

For illustrative purposes let us review a system that is familiar to most, namely, a three-dimensional coordinate system.[2] To do this, we need three things:

1. A point called the origin of the system, that will be identified by $\mathbf{0}'$, the zero vector.

2. Three lines, called the coordinate axes, that go through the origin. We shall assume for the time being that each line is perpendicular to the other two, and we shall call these lines rectangular Cartesian axes, denoted by x, y, and z

3. One point, other than the origin, on each of the three axes. We need these points to establish scale units and the notion of direction, positive or negative, relative to the origin. Here we assume that the unit of length on each axis is the same.

Figure 3.1 shows a simple illustration of the type of coordinate system that we can set up.

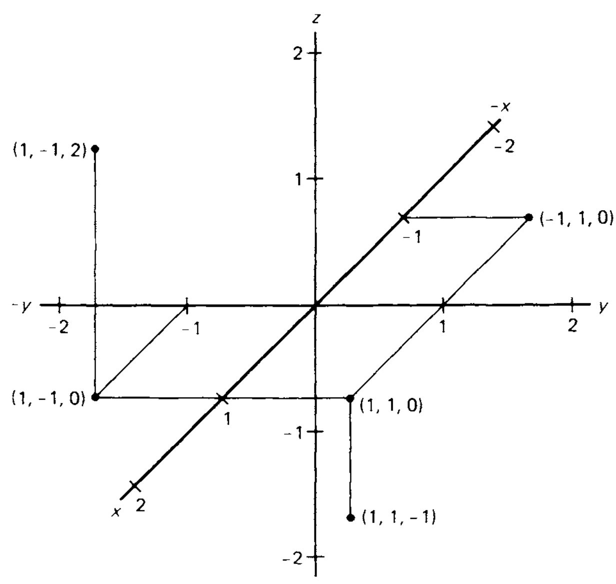

Fig. 3.1 A three-dimensional coordinate system with illustrative points.

[2] Later on we shall refer to two-dimensional as well as to three-dimensional systems. However, this particularization should cause no problems in interpretation.

By convention we have marked the positive directions of x, y, and z. Note further that three coordinate planes are also established:

1. The xy plane containing the x and y axes; this is the plane perpendicular to the z axis and containing the origin.
2. The xz plane containing the x and z axes; this is the plane perpendicular to the y axis and containing the origin.
3. The yz plane containing the y and z axes; this is the plane perpendicular to the x axis and containing the origin.

These planes cut the full space into eight octants. The first octant, for example, is the one above the xy plane in which all coordinates are positive.

Having established a coordinate system and the idea of signed distances along the axes, we can assign to each point in the space an ordered triple of real numbers:

$$\mathbf{a}' = (a_1, a_2, a_3)$$

where a_1 is the coordinate associated with the projection of $\mathbf{a}'$ onto the x axis, a_2 is the coordinate associated with the projection of $\mathbf{a}'$ onto the y axis, and a_3 is the coordinate associated with the projection of $\mathbf{a}'$ onto the z axis.

The (perpendicular) projection of a point onto a line is a vector on the line whose terminus or arrowhead is at the foot of the perpendicular dropped from the given point to the line. With the x, y, and z axes that have been set up in Fig. 3.1, the length of each projection is described on each axis by a single number, its coordinate. The coordinate is a *signed* distance from the origin; the sign is plus if the projection points in the positive direction and minus if the projection points in the negative direction. Figure 3.1 shows a few illustrative cases in different octants of the space.

In Chapter 2 we talked about a vector as a mathematical object having direction and magnitude. We need both characteristics since we can have an infinity of vectors, all having the same direction (but varying in length or magnitude), or all having the same length (but varying in direction). Furthermore, before we can talk meaningfully about direction, we need to fix a set of reference axes so that "direction" is considered relative to some standard.

In one sense vectors can originate and terminate anywhere in the space, as illustrated in Fig. 3.2. However, as also illustrated in Fig. 3.2, we can always move some arbitrary vector in a parallel direction so that the vector's tail starts at the origin. All vectors that start from the origin are called *position* vectors, and we essentially confine our attention to these. *Since we have not changed either the direction or the length of the arbitrary vector by this parallel displacement, any vector can be portrayed as a position vector.*

By concentrating our interest on position vectors, it turns out that any such vector can also be represented by a triple of numbers that we called components of a vector in Chapter 2. In the present context these components are also coordinates. By convention, the ith component of a vector is associated with the ith coordinate axis. This is illustrated in Fig. 3.3, by the projection of the terminus of $\mathbf{a}'$ onto x, y and z, the coordinate axes. Notice that each projection lies along the particular axis of interest. The (signed) length of each of these projections is, of course, described by a single number that is plus or minus, depending upon its direction along each axis relative to the origin.

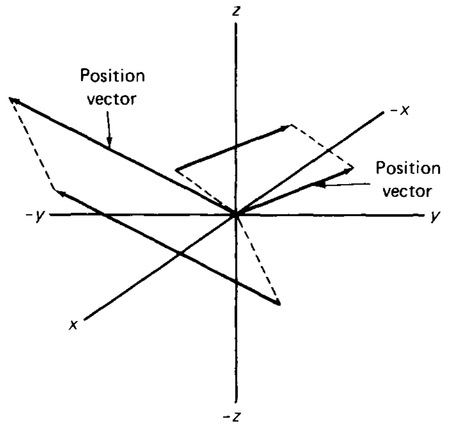

Fig. 3.2

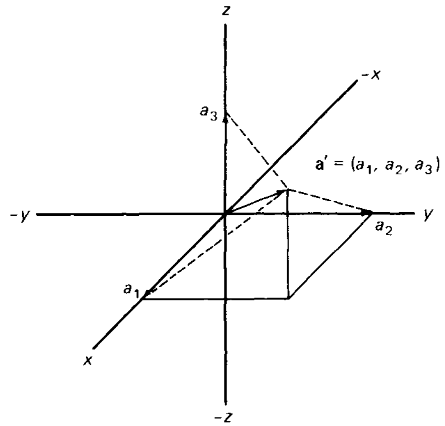

Fig. 3.3

Fig. 3.2 Parallel displacement ot arbitrary vectors.
Fig. 3.3 Vector components shown as signed lengths of projections.

Thus, given a fixed origin that is called $\mathbf{0}'$, we can always make a one-to-one correspondence between position vectors and points. For each point P we can find a corresponding position vector from the origin to P; for each position vector with its tail at $\mathbf{0}'$ we can locate a point P at the vector's terminus.

By restricting our attention to vectors emanating from the origin, any vector is both a geometric object, possessing length and direction, and an n-tuple of numbers (three numbers in this case). Since the vectors that we shall discuss *will* have their tails at the origin, *two vectors are equal if and only if they terminate at the same point.* If it were the case that two vectors had their tails at two different points, then they would be equal if and only if one of the vectors could be moved, without changing its direction or length, so that it coincided with the other.

In summary, then, by making sure that all of the vectors are position vectors (i.e., start at the origin of the coordinate system), we can pass freely back and forth between the *geometric* character of a vector (length and direction) and its *algebraic* character (an ordered n-tuple of scalars). The length of a vector's projection is given by the vector's coordinate on the x, y, and z axes, respectively, and the sign of its projection on x, y, and z depends on where the projection terminates, relative to the origin.

3.2.2 Standard Basis Vectors

Continuing on with the preliminaries, let us next consider Fig. 3.4. This figure shows a three-dimensional space with the vector $\mathbf{a}' = (1, 2, 2)$ appearing as a directed line segment.

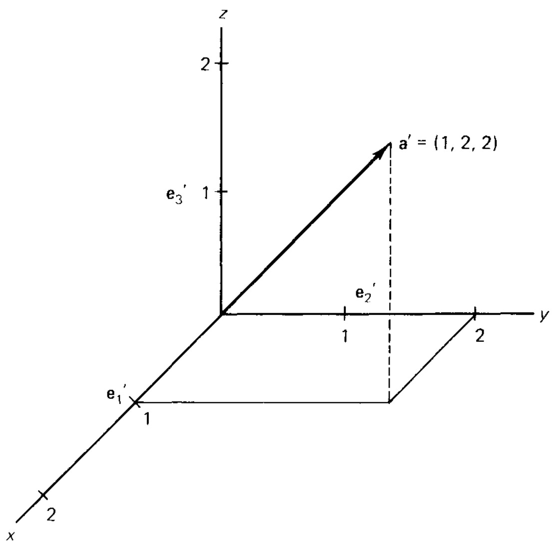

Fig. 3.4 Vector representation in three-dimensional space.

To set up this coordinate space, we define a special set of zero–one coordinate vectors, denoted $\mathbf{e}_i'$, as

1. vector $\mathbf{e}_1'$ of unit length in the positive (by convention) x direction:
$$\mathbf{e}_1' = (1, 0, 0)$$
2. vector $\mathbf{e}_2'$ of unit length in the positive y direction:
$$\mathbf{e}_2' = (0, 1, 0)$$
3. vector $\mathbf{e}_3'$ of unit length in the positive z direction:
$$\mathbf{e}_3' = (0, 0, 1)$$

We shall continue to let $\mathbf{0}'$, the zero vector, denote the origin of the space. As suggested in the discussion of vector addition and scalar multiplication of a vector in Chapter 2, we can now write the vector $\mathbf{a}' = (1, 2, 2)$ as a linear combination of the coordinate vectors:

$$\begin{aligned} 1\mathbf{e}_1' + 2\mathbf{e}_2' + 2\mathbf{e}_3' &= 1(1,0,0) + 2(0,1,0) + 2(0,0,1) \\ &= (1,0,0) + (0,2,0) + (0,0,2) \\ \mathbf{a}' &= (1,2,2) \end{aligned}$$

Note that what we have done is to perform scalar multiplication followed by vector addition, relative to the coordinate vectors $\mathbf{e}_i'$. We shall call the $\mathbf{e}_i'$ vectors a *standard basis* and comment later on the meaning of basis vectors, generally.

Note, further, that if we had the oppositely directed vector $-\mathbf{a}'$, this could also be represented in terms of the standard basis vectors as the linear combination:

$$-1\mathbf{e}_1' - 2\mathbf{e}_2' - 2\mathbf{e}_3' = (-1, 0, 0) + (0, -2, 0) + (0, 0, -2) = (-1, -2, -2)$$

In this case $-\mathbf{a}'$ would extend in the negative directions of x, y, and z.

What is shown above in particularized form can be generalized in accordance with the discussion of linear combinations of vectors in Chapter 2. As recalled:

Given p n-component vectors $\mathbf{b}_1', \mathbf{b}_2', \ldots, \mathbf{b}_p'$ *the n-component vector*

$$\mathbf{a}' = \sum_{i=1}^{p} k_i \mathbf{b}_i' = k_1 \mathbf{b}_1' + k_2 \mathbf{b}_2' + \cdots + k_p \mathbf{b}_p'$$

is a linear combination of p vectors, $\mathbf{b}_1', \mathbf{b}_2', \ldots, \mathbf{b}_p'$ *for any set of scalars* k_i $(i = 1, 2, \ldots, p)$.

In the illustration above we have $p = 3$ basis vectors, each containing $n = 3$ components. The components of the vector $\mathbf{a}' = (1, 2, 2)$ involve $p = 3$ scalars. The $\mathbf{b}_i'$ vectors in the more general expression above correspond to the specific $\mathbf{e}_i'$ vectors in the preceding numerical illustration.

The introduction of a set of standard basis vectors allows us to write *any* n-component vector $\mathbf{a}'$, relative to a standard basis of n-component $\mathbf{e}_i'$ vectors, as

$$\boxed{\mathbf{a}' = \sum_{i=1}^{n} a_i \mathbf{e}_i'}$$

where a_i denotes the ith component of $\mathbf{a}'$, and each of the n basis vectors has a 1 appearing in the ith position and zeros elsewhere. In this special case of a linear combination, the number of vectors p equals the number of components in $\mathbf{a}'$, namely, n.

Figure 3.5, incidentally, shows $\mathbf{a}'$ in terms of the triple of numbers (1, 2, 2). This point representation, as we now know, is equally acceptable for representing $\mathbf{a}'$ since the vector is already positioned with its tail at the origin.

The important point to remember is that $\mathbf{a}'$, itself, can be represented as a linear combination of other vectors—in this case, the standard basis vectors $\mathbf{e}_i'$. In a sense the $\mathbf{e}_i'$

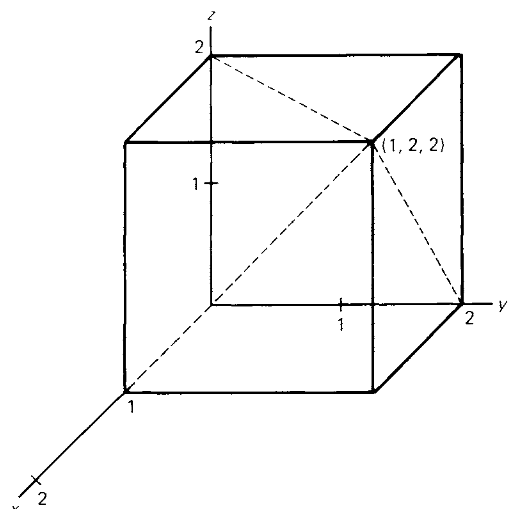

Fig. 3.5 Point representation in three-dimensional space.

represent a standard scale unit across the three axes. Any projection of $\mathbf{a}'$, then, can be considered as involving (signed) multiples of the appropriate $\mathbf{e}_i'$ vector.

With these preliminaries out of the way, we can now introduce the central concept of the chapter, namely, the Euclidean space and the associated idea of the distance between two points, or vector termini, in Euclidean space. This idea, in turn, leads to the concepts of angle and length.

3.2.3 Definition of Euclidean Space

A Euclidean space of n dimensions is the collection of all n-component vectors for which the operations of vector addition and multiplication by a scalar are permissible. Moreover, for any two vectors in the space, there is a nonnegative number, called the Euclidean distance between the two vectors.[3]

The function[4] that produces this nonnegative number is called a Euclidean distance function and is defined as

$$\boxed{\|\mathbf{a}' - \mathbf{b}'\| = [(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2]^{1/2}}$$

Alternatively, we can define $\|\mathbf{a}' - \mathbf{b}'\|$ in terms of a function of the now-familiar scalar product of $(\mathbf{a} - \mathbf{b})$ with itself:

$$\boxed{\|\mathbf{a}' - \mathbf{b}'\| = [(\mathbf{a} - \mathbf{b})'(\mathbf{a} - \mathbf{b})]^{1/2}}$$

where the vector $(\mathbf{a} - \mathbf{b})$ is a difference vector.

To get some geometric view of the Euclidean distance between two position vector termini (i.e., between two points), let us first examine Panel I of Fig. 3.6. Here in two dimensions are the two points

$$\mathbf{a}' = (1, 1); \qquad \mathbf{b}' = (1.5, 2)$$

Note that their straight-line distance can be represented by the square root of the hypotenuse of the right triangle, as sketched in the chart. In terms of the distance formula, we have

$$\|\mathbf{a}' - \mathbf{b}'\| = [(1 - 1.5)^2 + (1 - 2)^2]^{1/2} = \sqrt{1.25} = 1.12$$

Panel II of Fig. 3.6 merely extends the same idea to three dimensions for two new points:

$$\mathbf{a}' = (1, 1, -2); \qquad \mathbf{b}' = (2, 1, 2)$$

$$\|\mathbf{a}' - \mathbf{b}'\| = [(1 - 2)^2 + (1 - 1)^2 + (-2 - 2)^2]^{1/2} = \sqrt{17} = 4.12$$

[3] A more formal definition considers a Euclidean space as a finite-dimensional vector space on which a real-valued scalar or inner product is defined.

[4] The Euclidean metric is, itself, a special case of the Minkowski metric. The Minkowski metric also obeys the three distance axioms (positivity, symmetry, and triangle inequality). Since we have used the single bars $|\mathbf{A}|$ to denote the determinant of a matrix in Chapter 2, we use the double bars $\|\mathbf{a}' - \mathbf{b}'\|$ to denote the distance between two vectors, taken here to mean Euclidean distance.

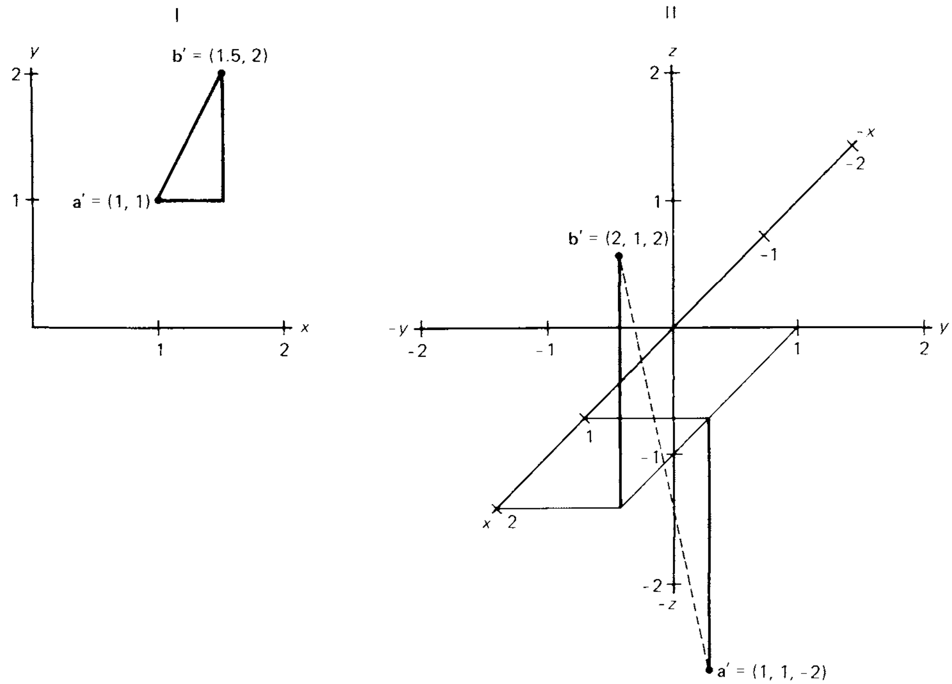

Fig. 3.6 Illustrations of Euclidean distances between pairs of points. Key: I, two dimensions; II, three dimensions.

Euclidean distance, then, entails adding up the squared differences in projections of each point on each axis in turn, and then taking the square root of the sum.

As might be surmised from either panel in Fig. 3.6, the Euclidean distance function possesses the following properties:

$\|\|\mathbf{a}' - \mathbf{b}'\|\| > 0$ unless $\mathbf{a}' - \mathbf{b}' = \mathbf{0}'$;		positivity
$\|\|\mathbf{a}' - \mathbf{b}'\|\| = \|\|\mathbf{b}' - \mathbf{a}'\|\|$;	symmetry	
$\|\|\mathbf{a}' - \mathbf{b}'\|\| + \|\|\mathbf{b}' - \mathbf{c}'\|\| \geqslant \|\|\mathbf{a}' - \mathbf{c}'\|\|$;		triangle inequality

The first of the above properties, positivity, precludes the possibility of negative distances. Symmetry, the second property, means that the distance from $\mathbf{a}'$ to $\mathbf{b}'$ is the same as the distance from $\mathbf{b}'$ to $\mathbf{a}'$. The third property, triangle inequality, states that the sum of the distances between $\mathbf{a}'$ and $\mathbf{b}'$ and between $\mathbf{b}'$ and some third point $\mathbf{c}'$ is no less than the direct distance between $\mathbf{a}'$ and $\mathbf{c}'$. If $\mathbf{b}'$ lies on the line connecting $\mathbf{a}'$ and $\mathbf{c}'$, then the sum of the distances of $\mathbf{a}'$ to $\mathbf{b}'$ and $\mathbf{b}'$ to $\mathbf{c}'$ equals the direct distance from $\mathbf{a}'$ to $\mathbf{c}'$.

We next define the concept of vector length or magnitude. The length of a vector $\mathbf{a}' = (a_1, a_2, \ldots, a_n)$ is defined as

$$\boxed{\|\mathbf{a}'\| = \left[\sum_{i=1}^{n} a_i^2\right]^{1/2}}$$

Note that this is a special case of the Euclidean distance function in which the second vector is the origin of the space, or the $\mathbf{0}'$ vector. That is,

$$||\mathbf{a}'|| = ||\mathbf{a}' - \mathbf{0}'|| = [(a_1-0)^2 + (a_2-0)^2 + \cdots + (a_n-0)^2]^{1/2}$$

Furthermore, we can also observe that the *squared* vector length equals the scalar product of **a** with itself:

$$\boxed{||\mathbf{a}'||^2 = \mathbf{a}'\mathbf{a}}$$

Thus, in the case of $\mathbf{a}' = (1, 2, 2)$, we see that

$$\begin{aligned} ||\mathbf{a}'||^2 &= \sum_{i=1}^{3} a_i^2 = (1)^2 + (2)^2 + (2)^2 \\ &= [(a_1-0)^2 + (a_2-0)^2 + (a_3-0)^2] = (1)^2 + (2)^2 + (2)^2 \\ &= \mathbf{a}'\mathbf{a} = (1,2,2)'(1,2,2) = (1)^2 + (2)^2 + (2)^2 = 9 \end{aligned}$$

are all equivalent ways of finding the squared length of $\mathbf{a}'$. The square root of $||\mathbf{a}'||^2$, that is, $\sqrt{9} = 3$, is, of course, the Euclidean distance or vector length of the vector terminus as measured from the origin $\mathbf{0}'$.

We now discuss some of these notions in more detail. Since it will be intuitively easier to present the concepts in terms of the standard basis vectors $\mathbf{e}_i'$—that is, where rectangular (mutually perpendicular) Cartesian coordinates are used—we discuss this case first and later briefly discuss more general coordinate systems in which the axes are *not* necessarily mutually perpendicular, although the space is still assumed to be Euclidean.

3.3 GEOMETRIC REPRESENTATION OF VECTORS

We have already commented on the fact that a vector can be equally well represented by the directed line segment, starting from the origin (Fig. 3.4), or the triple of point coordinates (Fig. 3.5). In both representations, the coordinate on each axis is the foot of the perpendicular dropped from $\mathbf{a}'$ to each axis.

Our interest now is in expanding some of these geometric notions so as to come up with graphical counterparts to the various algebraic operations on vectors that were described in Chapter 2.

3.3.1 Length and Direction Angles of a Single Vector

Let us again examine vector $\mathbf{a}' = (1, 2, 2)$, represented as the directed line segment in Fig. 3.4. As shown earlier, the length or Euclidean distance, denoted $||\mathbf{a}'||$, of $\mathbf{a}'$ from the origin $\mathbf{0}'$ is

$$\begin{aligned} ||\mathbf{a}'|| &= [(a_1-0)^2 + (a_2-0)^2 + (a_3-0)^2]^{1/2} \\ &= [(1)^2 + (2)^2 + (2)^2]^{1/2} = 3 \end{aligned}$$

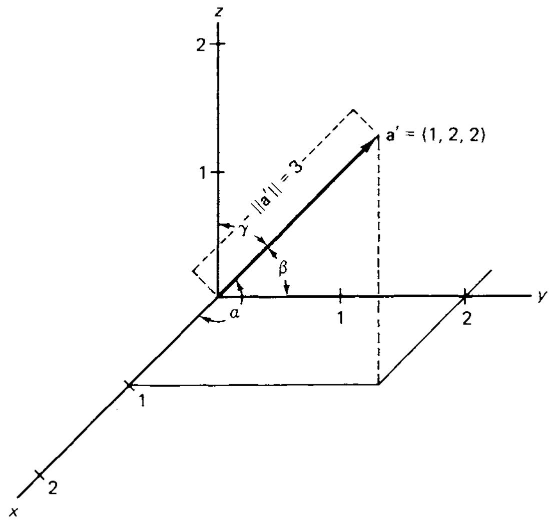

Fig. 3.7 Vector length and direction cosines.

Our interest now focuses on various aspects of length and *direction angles.* First, observe from Fig. 3.7 that $\mathbf{a}'$ makes angles of α, β, and γ with the three reference axes x, y, and z. The cosines of these angles are called *direction cosines* relative to the reference axes and are computed as *ratios of each vector component to the vector's length.*

Since the length of $\mathbf{a}'$ is 3 and the components of $\mathbf{a}'$ are 1, 2, and 2, the cosines of α, β, and γ, respectively, are

$$\cos\alpha = \tfrac{1}{3}; \qquad \cos\beta = \tfrac{2}{3}; \qquad \cos\gamma = \tfrac{2}{3}$$

Notice that these can be written out in full as

$$\cos\alpha = \frac{a_1}{[a_1^2 + a_2^2 + a_3^2]^{1/2}} = \frac{1}{3}$$

$$\cos\beta = \frac{a_2}{[a_1^2 + a_2^2 + a_3^2]^{1/2}} = \frac{2}{3}$$

$$\cos\gamma = \frac{a_3}{[a_1^2 + a_2^2 + a_3^2]^{1/2}} = \frac{2}{3}$$

The angles corresponding to these cosines are

$$\alpha \cong 71°; \qquad \beta \cong 48°; \qquad \gamma \cong 48°$$

Notice that our use of the cosine is in accordance with basic trigonometry. For example, the cosine of the angle α, which the vector $\mathbf{a}'$ makes with the x axis, is equal to the length of the adjacent side of the right triangle, formed by projection of $\mathbf{a}'$ onto the x axis, divided by the hypotenuse of that right triangle. The adjacent side has length 1, or unit distance from the origin, and the hypotenuse is of length 3. Hence, the cosine of α is $\frac{1}{3}$. By similar reasoning the cosine of β is $\frac{2}{3}$ with respect to the y axis, and that of γ is $\frac{2}{3}$ with respect to the z axis.

We can also discuss some of these notions in somewhat more general terms. Once a coordinate system is chosen, any position vector that emanates from the origin can be represented by

1. the angles α, β, γ, made by the line with respect to the x, y, and z axes, where $0 \leqslant \alpha, \beta, \gamma \leqslant 180°$, and
2. the vector's length or magnitude.

We have already discussed the case of vectors that emanate from locations other than the origin of the space. Therefore, by appropriate parallel displacement to a position vector, any vector in the space can be represented by its direction angles and length.

If we had a vector $-\mathbf{a}' = (-1, -2, -2)$ that was oppositely directed from $\mathbf{a}'$, this would cause no problems since the direction cosines and angles would then be

$$\cos\alpha = -\tfrac{1}{3}; \qquad \alpha = 109° = 180° - 71°$$

$$\cos\beta = -\tfrac{2}{3}; \qquad \beta = 132° = 180° - 48°$$

$$\cos\gamma = -\tfrac{2}{3}; \qquad \gamma = 132° = 180° - 48°$$

It is also useful to examine the sum of the squared cosines of α, β, and γ. Since $a_1{}^2 + a_2{}^2 + a_3{}^2 = \|\mathbf{a}'\|^2$, we have

$$\boxed{\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1}$$

We can state the above result in words: The sum of the squares of the direction cosines of some vector $\mathbf{a}'$*, originating at the origin, is equal to 1.* This fact holds true in any dimensionality, not just three dimensions.

Furthermore, it is a simple matter to work backward to find the components of a vector if we know its direction angles and length. Continuing with the illustrative vector,

$$\mathbf{a}' = (1, 2, 2)$$

with direction angles and cosines,

$$\alpha = 71°, \quad \cos\alpha = \tfrac{1}{3}; \qquad \beta = 48°, \quad \cos\beta = \tfrac{2}{3}; \qquad \gamma = 48°, \quad \cos\gamma = \tfrac{2}{3}$$

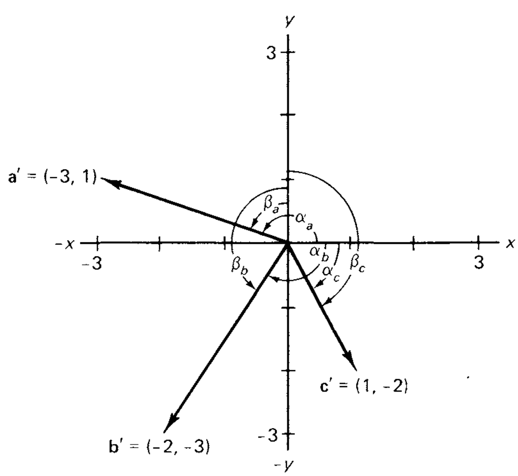

Fig. 3.8 Direction angles and lengths of illustrative vectors.

and length $\|\mathbf{a}'\| = 3$, we have, by simple algebra, the vector components:

$$a_1 = \tfrac{1}{3}(3) = 1; \qquad a_2 = \tfrac{2}{3}(3) = 2; \qquad a_3 = \tfrac{2}{3}(3) = 2$$

In the case of negative coordinates, the cosines, of course, would be negative for those axes involving negative projections.

Working with negative cosines is most easily shown in two dimensions, as illustrated in Fig. 3.8. Here are portrayed three different vectors, terminating in three different quadrants. We first note the smaller angle ($\leqslant 180°$) made by each vector with each axis. Then, by means of the formulas shown earlier, the direction cosines of each vector are computed as follows:

$$\boxed{\mathbf{a}'} \quad \cos\alpha_a = \frac{-3}{[(-3)^2 + (1)^2]^{1/2}} = -0.95; \qquad \cos\beta_a = \frac{1}{[(-3)^2 + (1)^2]^{1/2}} = 0.32$$

$$\boxed{\mathbf{b}'} \quad \cos\alpha_b = \frac{-2}{[(-2)^2 + (-3)^2]^{1/2}} = -0.55; \qquad \cos\beta_b = \frac{-3}{[(-2)^2 + (-3)^2]^{1/2}} = -0.83$$

$$\boxed{\mathbf{c}'} \quad \cos\alpha_c = \frac{1}{[(1)^2 + (-2)^2]^{1/2}} = 0.45; \qquad \cos\beta_c = \frac{-2}{[(1)^2 + (-2)^2]^{1/2}} = -0.89$$

with correspondent direction angles:

$$\begin{aligned} \alpha_a &\cong 161°; & \beta_a &\cong 71° \\ \alpha_b &\cong 123°; & \beta_b &\cong 146° \\ \alpha_c &\cong 63°; & \beta_c &\cong 153° \end{aligned}$$

as shown in Fig. 3.8.

Notice, in particular, that as any angle becomes obtuse, the formulas for finding direction cosines still hold since changes in the sign of the cosine are taken care of by corresponding changes in the appropriate vector components.

In summary, then, any position vector is uniquely determined by knowledge of its magnitude and direction. In turn, its direction is given by the angles it makes with the reference axes. These angles are obtained from the cosines that are computed by the expression

$$\boxed{\cos\Psi_i = \frac{a_i}{\|\mathbf{a}'\|}}$$

where Ψ_i denotes the angle between the vector and the ith reference axis ($0° \leqslant \Psi_i \leqslant 180°$), a_i denotes the ith component of $\mathbf{a}'$, and $\|\mathbf{a}'\|$ denotes its length.

Note, in particular, that if $\cos\Psi_i = 0$, *then the angle is* 90° *and the vector is said to be orthogonal or perpendicular to the ith reference axis.*

3.3.2 Geometric Aspects of Vector Addition and Multiplication by a Scalar

While we have earlier discussed in Chapter 2 the rules of vector addition and subtraction and multiplication of a vector by a scalar, it is useful now to show these

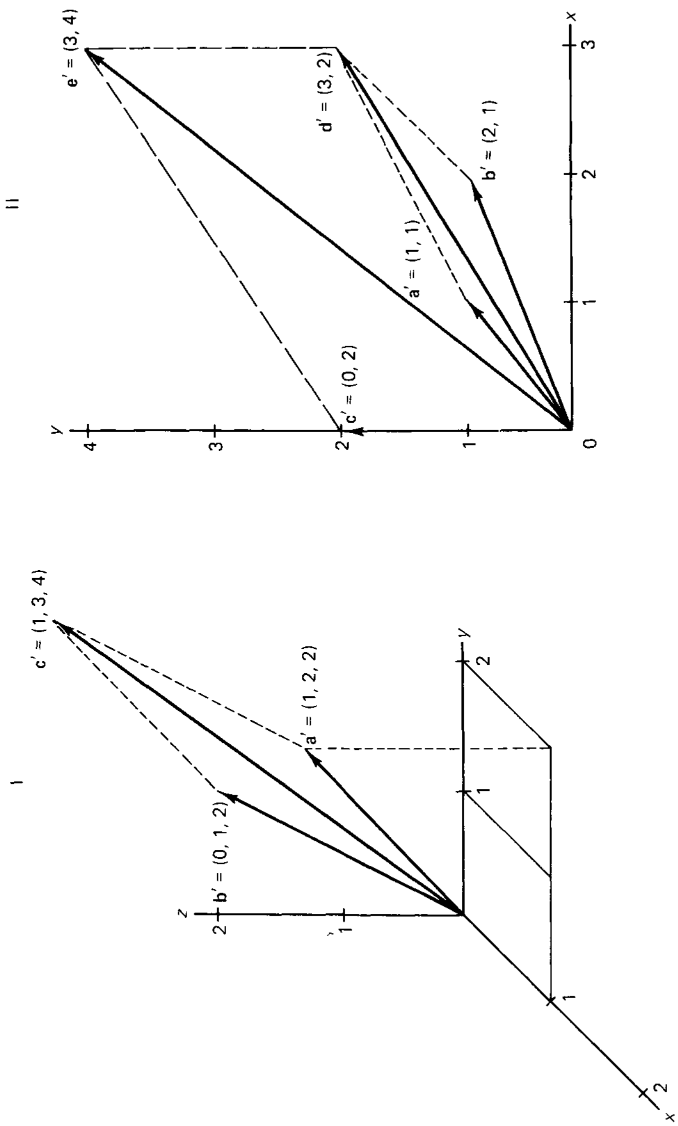

Fig. 3.9 Illustrations of vector addition.

operations geometrically. First consider the two vectors $\mathbf{a}' = (1, 2, 2)$ and $\mathbf{b}' = (0, 1, 2)$ shown in Panel I of Fig. 3.9. As already known from Chapter 2, their vector sum is

$$(1+0, 2+1, 2+2) = (1,3,4)$$

We can formalize this by saying that if $\mathbf{a}'$ *and* $\mathbf{b}'$ *are* $1 \times n$ *vectors, their vector sum is defined by*

$$\boxed{\mathbf{a}' + \mathbf{b}' = (a_1 + b_1, a_2 + b_2, \ldots, a_i + b_i, \ldots, a_n + b_n)}$$

As noted from Panel I of the figure, vector addition proceeds on a component-by-component basis. Geometrically, $\mathbf{c}' = \mathbf{a}' + \mathbf{b}'$ is represented by the diagonal of a parallelogram determined by $\mathbf{a}'$ and $\mathbf{b}'$.

Panel II of Fig. 3.9 shows a case for three vectors in two dimensions, $\mathbf{a}'$, $\mathbf{b}'$ and $\mathbf{c}'$. When $\mathbf{a}'$ and $\mathbf{b}'$ are added, their sum is represented by $\mathbf{d}'$, the diagonal of a parallelogram. The parallelogram rule also applies as $\mathbf{d}'$ is added to $\mathbf{c}'$, resulting in their vector sum, shown by $\mathbf{e}'$.

Vector subtraction presents no major additional complications. Suppose, for example, that we wish to show the difference

$$\mathbf{d}' = \mathbf{a}' - \mathbf{b}' = (1-0, 2-1, 2-2) = (1,1,0)$$

geometrically. Figure 3.10 shows the *difference* vector, denoted by $\mathbf{d}'$, as a vector emanating from the origin with the *same length and direction* as that indicated by the line connecting the arrowheads of $\mathbf{a}'$ and $\mathbf{b}'$. Notice, then, that we maintain the concept of position vector by making a parallel displacement of the difference between $\mathbf{a}'$ and $\mathbf{b}'$ so that $\mathbf{d}'$ starts from the origin.

If we had the vector $\mathbf{a}'$ and another vector $-\mathbf{a}'$, it would, of course, be the case that

$$\mathbf{a}' + (-\mathbf{a}') = \mathbf{0}'$$

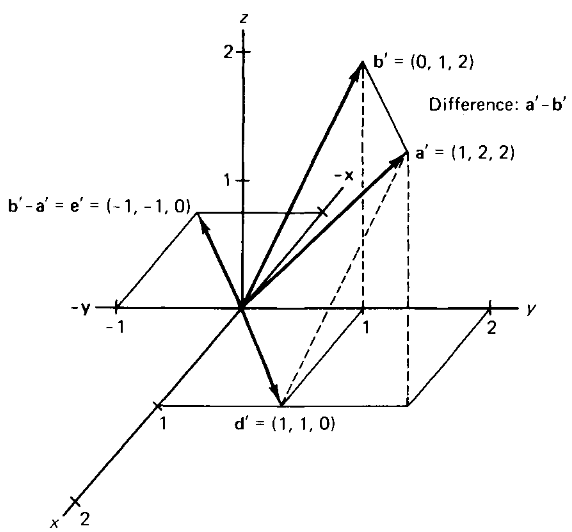

Fig. 3.10 Illustrations of vector subtraction.

Note also that the vector subtraction

$$\mathbf{e}' = \mathbf{b}' - \mathbf{a}' = (0-1, 1-2, 2-2) = (-1, -1, 0)$$

is handled analogously and, furthermore, that $-1\mathbf{e}' = \mathbf{d}'$, as should be the case. We find that $\mathbf{d}'$ and $\mathbf{e}'$ are merely oppositely directed vectors of equal length.[5]

Multiplication of a vector **a** *by a scalar k is formally defined as*

$$\boxed{k(\mathbf{a}') = (ka_1, ka_2, \ldots, ka_i, \ldots, ka_n)}$$

and is also illustrated in Fig. 3.10 for the special case in which $k = -1$. That is,

$$\mathbf{e}' = -1(\mathbf{d}') = -1(1, 1, 0) = (-1, -1, 0)$$

As a more general example, Fig. 3.11 shows the case of multiplying the vector

$$\mathbf{a}' = (2, 3, 2)$$

by $k_1 = -1$, $k_2 = \frac{1}{2}$, $k_3 = 2$. We note that the sign of k determines the direction of $k\mathbf{a}'$ while the magnitude of k determines how far $k\mathbf{a}'$ extends in the appropriate direction from the origin, relative to $\|\mathbf{a}'\|$, the length of $\mathbf{a}'$ when $k = 1$.

As a concluding example we combine the operations of addition and scalar multiplication of a vector by considering the case of a linear combination:

$$\tfrac{1}{2}\mathbf{a}' + 2\mathbf{b}' = \tfrac{1}{2}(1, 2, 2) + 2(0, 1, 2) = (\tfrac{1}{2}, 1, 1) + (0, 2, 4) = (\tfrac{1}{2}, 3, 5)$$

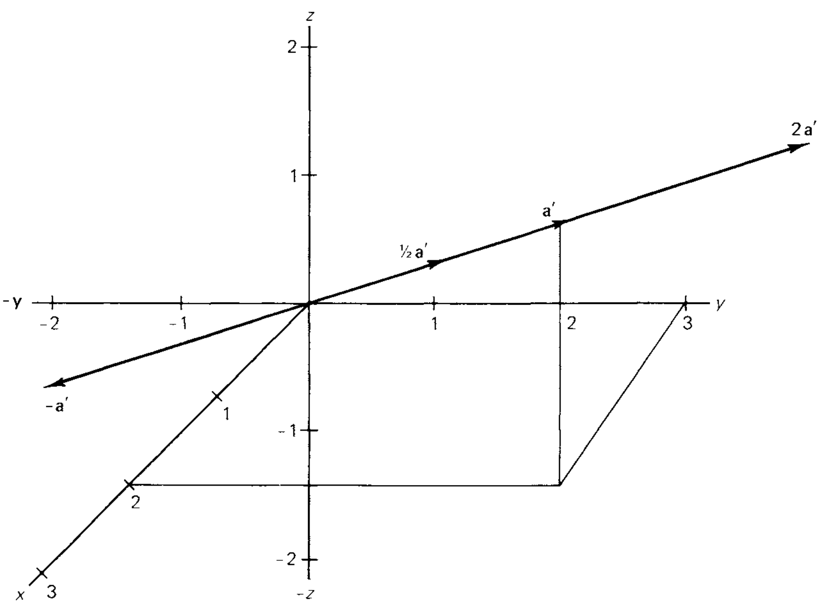

Fig. 3.11 Illustrations of vector multiplication by a scalar.

[5] The vector $\mathbf{e}'$ is used here as an arbitrary vector and is not to be confused with the standard basis vectors $\mathbf{e}_i'$, introduced earlier in the chapter.

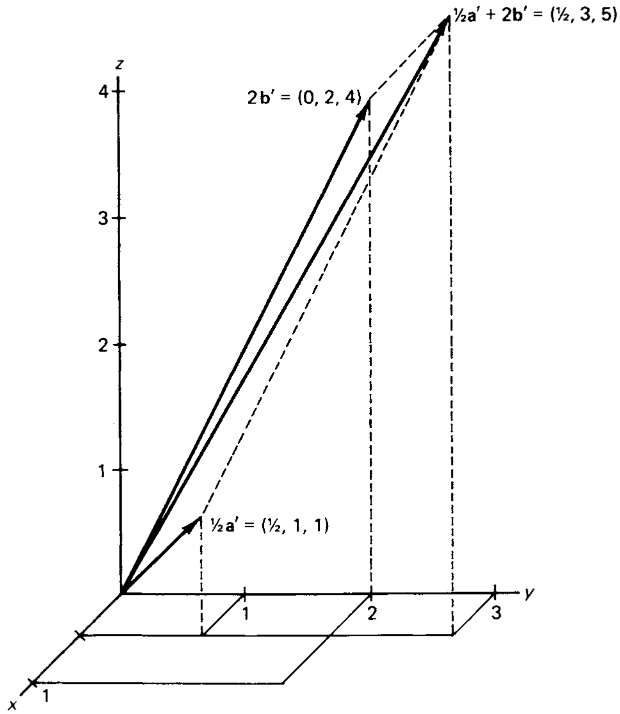

Fig. 3.12 An illustration of the combined operations of scalar multiplication and addition.

The result of these operations appears in Fig. 3.12. This same idea can, of course, be extended to more than two vectors. For example, in three dimensions the sum of three three-component vectors would be represented by the diagonal of a parallelepiped formed from the three contributing vectors. As long as we confine the number of components of each vector to at most three, it becomes quite straightforward to picture the operations of addition, subtraction, scalar multiplication of a vector, and their generalization, a linear combination of vectors.

The properties of addition, subtraction, and multiplication of a vector by a scalar were listed in Chapter 2. These properties, of course, apply here since our current purpose is simply to portray the same vector relations geometrically rather than algebraically.

3.3.3 Distance and Angle between Two Vectors

In Section 3.3.1 we considered the special case of the angle between two vectors when one of those vectors was a coordinate axis. We can now discuss the general situation of the angle between any pair of position vectors in Euclidean space. Suppose we continue to consider the case of the two vectors $\mathbf{a}' = (1, 2, 2)$ and $\mathbf{b}' = (0, 1, 2)$. As shown earlier, vector $\mathbf{a}'$ has length 3. Vector $\mathbf{b}'$ has length

$$\|\mathbf{b}'\| = [(0)^2 + (1)^2 + (2)^2]^{1/2} = \sqrt{5}$$

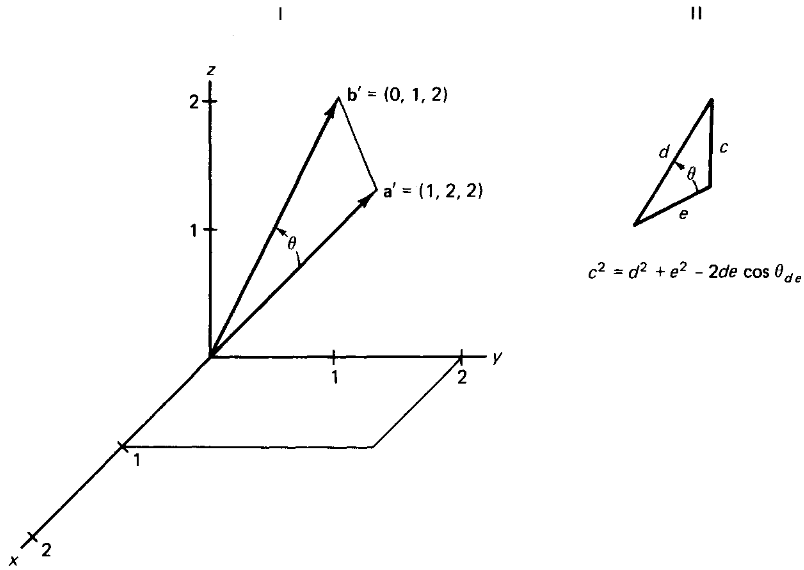

Fig. 3.13 Finding the angle between the two vectors. Key: I, position vectors; II, arbitrary triangle.

with direction cosines and angles, with respect to x, y, and z, of

$$\cos\alpha^* = 0/\sqrt{5} = 0; \qquad \alpha^* = 90^\circ$$
$$\cos\beta^* = 1/\sqrt{5} = 0.45; \qquad \beta^* \cong 63^\circ$$
$$\cos\gamma^* = 2/\sqrt{5} = 0.89; \qquad \gamma^* \cong 27^\circ$$

So far, nothing new. Now we ask: What is the Euclidean distance between $\mathbf{a}'$ and $\mathbf{b}'$?

Again, as we know, the distance between $\mathbf{a}'$ and $\mathbf{b}'$ can be computed as

$$\|\mathbf{a}' - \mathbf{b}'\| = [(1-0)^2 + (2-1)^2 + (2-2)^2]^{1/2} = \sqrt{2} = 1.41$$

That is, we find the difference between the two vectors on a component-by-component basis, square each of these differences, sum the squared differences, and then take the square root of the result. Notice that this is similar to finding a vector's length in which the origin, or zero vector, plays the role of the second vector.

Again, nothing new. However, at this point we can note from Panel I of Fig. 3.13 that $\mathbf{a}'$ and $\mathbf{b}'$ make some angle θ with each other. The problem, now, is to determine what this angle is. That is, analogous to the case of finding the angle that a single vector makes with each of the reference axes, we now wish to find the angle between two different vectors referred to the same set of coordinate axes. To do so, we make use of the *cosine law of trigonometry*.

As the reader may recall from basic trigonometry, the law of cosines states:

For any triangle with sides c, d, and e, the square of any side is equal to the sum of the squares of the other two sides minus twice the product of the other two sides and the *cosine of their included angle* θ. Or, to illustrate (see arbitrary triangle in Panel II of Fig. 3.13),

$$\boxed{c^2 = d^2 + e^2 - 2de \cos \theta_{de}}$$

Similarly, we could find d^2 or e^2, as the case may be.

Returning to our specific example in Panel I of Fig. 3.13, by simple algebra we can first express the law of cosines in terms of the cosine of

$$\boxed{\cos \theta_{\mathbf{a}'\mathbf{b}'} = \frac{\|\mathbf{a}'\|^2 + \|\mathbf{b}'\|^2 - \|\mathbf{a}' - \mathbf{b}'\|^2}{2\|\mathbf{a}'\| \cdot \|\mathbf{b}'\|}}$$

where the above formula represents the particularized version of

$$\boxed{\cos \theta_{de} = \frac{d^2 + e^2 - c^2}{2de}}$$

as applying to any triangle of interest.

In terms of our specific problem, the cosine of the angle θ between the vectors $\mathbf{a}'$ and $\mathbf{b}'$ is expressed as a ratio in which the numerator is the squared length of $\mathbf{a}'$ plus the squared length of $\mathbf{b}'$ minus the squared length of the difference vector $\mathbf{a}' - \mathbf{b}'$; the denominator of the ratio is simply 2 times the product of the lengths of $\mathbf{a}'$ and $\mathbf{b}'$.

If we then substitute the appropriate numerical quantities, we have

$$\cos \theta_{\mathbf{a}'\mathbf{b}'} = \frac{\|\mathbf{a}'\|^2 + \|\mathbf{b}'\|^2 - \|\mathbf{a}' - \mathbf{b}'\|^2}{2\|\mathbf{a}'\| \cdot \|\mathbf{b}'\|} = \frac{9 + 5 - 2}{2(3)(\sqrt{5})} = \frac{12}{13.416} = 0.894$$

with the correspondent angle

$$\theta_{\mathbf{a}'\mathbf{b}'} \cong 27^\circ$$

Notice further that we can turn this procedure around. If we know the ***angle*** that two vectors make with each other and their lengths, another way of finding the squared distance between them makes use of a rearrangement of the above formula to

$$\boxed{\|\mathbf{a}' - \mathbf{b}'\|^2 = \|\mathbf{a}'\|^2 + \|\mathbf{b}'\|^2 - 2 \cos \theta_{\mathbf{a}'\mathbf{b}'} \|\mathbf{a}'\| \cdot \|\mathbf{b}'\|}$$

The concepts illustrated here for two dimensions also hold true in higher dimensions since two noncollinear vectors will entail a (plane) triangle embedded in higher dimensionality.[6] The vector lengths, of course, will be based on projections on all axes of the higher-dimensional space.

A few other observations are of interest. First, if the angle θ between two vectors is 90°, then $\cos \theta = 0$, and one has the familiar Pythagorean theorem for a right triangle in which the square of the hypotenuse is equal to the sum of the squares of the sides. In the case where $\cos \theta = 0$, the two vectors are said to be *orthogonal* (as mentioned earlier). If $\cos \theta_{\mathbf{a}'\mathbf{b}'} = 1$, then $\mathbf{a}'$ and $\mathbf{b}'$ are collinear in the same direction, and the sum of the

[6] By noncollinear is meant that the vectors are not superimposed so that all points of one vector fall on the other vector.

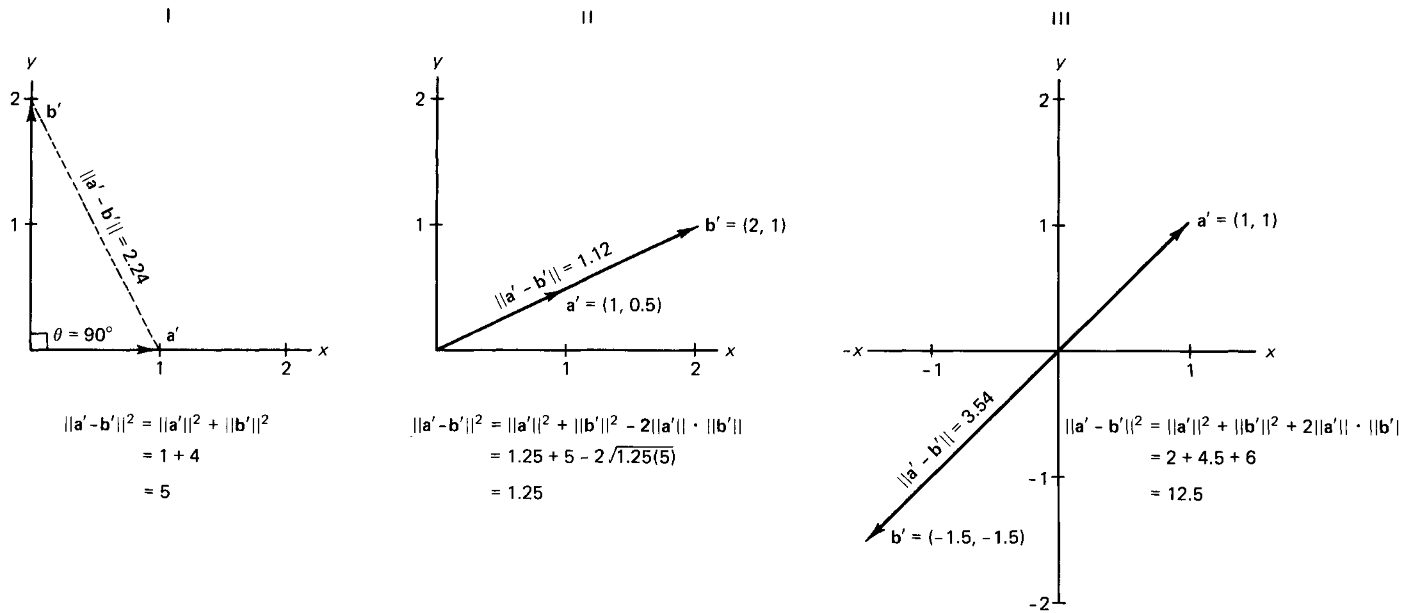

Fig. 3.14 Some special cases of the squared distance between two vectors.

squared lengths of $\mathbf{a}'$ and $\mathbf{b}'$ is appropriately reduced by $2\|\mathbf{a}'\| \cdot \|\mathbf{b}'\|$. If $\cos\theta_{\mathbf{a}'\mathbf{b}'} = -1$, then $\mathbf{a}'$ and $\mathbf{b}'$ are oppositely directed, and the sum of the squared lengths of $\mathbf{a}'$ and $\mathbf{b}'$ is increased by $2\|\mathbf{a}'\| \cdot \|\mathbf{b}'\|$.

These latter two relationships are easily seen by recalling that for scalars we have the identities

$$\boxed{(x-y)^2 \equiv x^2 + y^2 - 2xy}$$

and

$$\boxed{[x-(-y)]^2 = (x+y)^2 \equiv x^2 + y^2 + 2xy}$$

Figure 3.14 shows geometrical examples of all three of the preceding cases.

Later on, when we discuss some of the more common measures of statistical association, we shall find that the above relationships are useful in portraying various statistical measures from a geometric standpoint. At this point, however, we proceed to a geometric description of still another concept of vector algebra, namely, the *scalar product* of two vectors and its relationship to Euclidean distance.

3.3.4 The Scalar Product of Two Vectors

In Chapter 2 we defined the scalar (or inner or dot) product of two vectors $\mathbf{a}$ and $\mathbf{b}$ (of conformable order) as

$$\boxed{\mathbf{a}'\mathbf{b}}$$

in which, if $\mathbf{a}' = (a_1, a_2, \ldots, a_k, \ldots, a_n)$ and $\mathbf{b}' = (b_1, b_2, \ldots, b_k, \ldots, b_n)$, then

$$\boxed{\mathbf{a}'\mathbf{b} = \sum_{k=1}^{n} a_k b_k}$$

and the result was a single number, or scalar.

A geometrically motivated (and more general) definition of scalar product, which takes into consideration the *angle* θ made between the two vectors and their respective lengths, can now be presented. This definition of scalar product is given by the expression

$$\boxed{\mathbf{a}'\mathbf{b} = \|\mathbf{a}\| \cdot \|\mathbf{b}\| \cos\theta_{\mathbf{ab}}}$$

In the above example in which $\mathbf{a}' = (1, 2, 2)$ and $\mathbf{b}' = (0, 1, 2)$, we have

$$\mathbf{a}'\mathbf{b} = 3(\sqrt{5})(0.894) = 6$$

We also recall that $\mathbf{a}'\mathbf{b} = \mathbf{b}'\mathbf{a}$. The geometric counterpart of this is

$$\boxed{\mathbf{b}'\mathbf{a} = \|\mathbf{b}\| \cdot \|\mathbf{a}\| \cos\theta_{\mathbf{ab}}}$$

Moreover, the counterpart to $\mathbf{a}'\mathbf{a}$, the scalar product of a vector with itself, is simply

$$\boxed{\|\mathbf{a}\| \cdot \|\mathbf{a}\| \cos \theta_{\mathbf{aa}} = \|\mathbf{a}\|^2 \cdot 1}$$

This is the vector's squared length, inasmuch as the angle θ that a vector makes with itself is, of course, zero; hence $\cos \theta_{\mathbf{aa}} = 1$.

3.3.5 Vector Projections and Scalar Products

Still another way of looking at the scalar product of two vectors is in terms of the signed length of a projection of one vector along another. At the beginning of this chapter, we talked informally about the projection of a vector onto the coordinate axes x, y, and z. Its projection on some axis, say x, was referred to as the signed distance from the origin, along x, to the foot of a perpendicular dropped from the vector onto x. Similar interpretations pertained to the vector's projections on axes y and z.

However, suppose we have two arbitrary position vectors in the space. Clearly, we could consider the projection of one vector onto the other, in a fashion analogous to coordinate projections. This concept is most simply described in two dimensions. Accordingly, let us select two arbitrary vectors

$$\mathbf{a}' = (1, 2) \qquad \text{and} \qquad \mathbf{b}' = (0, 2)$$

These vectors are shown in Panel I of Fig. 3.15. We now project $\mathbf{b}'$ onto $\mathbf{a}'$ by dropping a perpendicular from $\mathbf{b}'$'s terminus to $\mathbf{a}'$. *The number*

$$\boxed{\|\mathbf{b}_p'\| = \left|\|\mathbf{b}'\| \cos \theta_{\mathbf{a}'\mathbf{b}'}\right| = \frac{\mathbf{a}'\mathbf{b}}{\|\mathbf{a}'\|}}$$

is defined as the length[7] *of the projection of the vector* $\mathbf{b}'$ *along the vector* $\mathbf{a}'$. The length of the projection is also frequently called the *component* of $\mathbf{b}'$ along $\mathbf{a}'$.

This concept is most easily understood by first recalling that the cosine can be viewed in terms of the length of the projection of a *unit length* vector, in this case one in the direction of $\mathbf{b}'$, onto the adjacent side (vector $\mathbf{a}'$) of a right triangle. Here, the unit length is multiplied by $\|\mathbf{b}'\|$. In this example $\|\mathbf{b}'\| = [(0)^2 + (2)^2]^{1/2} = 2$.

The cosine $\theta_{\mathbf{a}'\mathbf{b}'}$ is next found from the cosine law:

$$\cos \theta_{\mathbf{a}'\mathbf{b}'} = \frac{\|\mathbf{a}'\|^2 + \|\mathbf{b}'\|^2 - \|\mathbf{a}' - \mathbf{b}'\|^2}{2\|\mathbf{a}'\| \cdot \|\mathbf{b}'\|} = \frac{5 + 4 - 1}{2(\sqrt{5})(\sqrt{4})} = 0.89$$

$$\theta_{\mathbf{a}'\mathbf{b}'} \cong 27^\circ$$

[7] Note that in defining $\|\mathbf{b}_p'\|$ we use the *absolute value* of the expression $\|\mathbf{b}'\| \cos \theta_{\mathbf{a}'\mathbf{b}'}$ since lengths are taken to be nonnegative. However, the *signed* distance is in the direction of $\mathbf{a}'$ if $\cos_{\mathbf{a}'\mathbf{b}'}$ is positive (i.e., the angle $\theta_{\mathbf{a}'\mathbf{b}'}$ is acute) and in the direction of $-\mathbf{a}'$ if $\cos \theta_{\mathbf{a}'\mathbf{b}'}$ is negative (i.e., the angle $\theta_{\mathbf{a}'\mathbf{b}'}$ is obtuse).

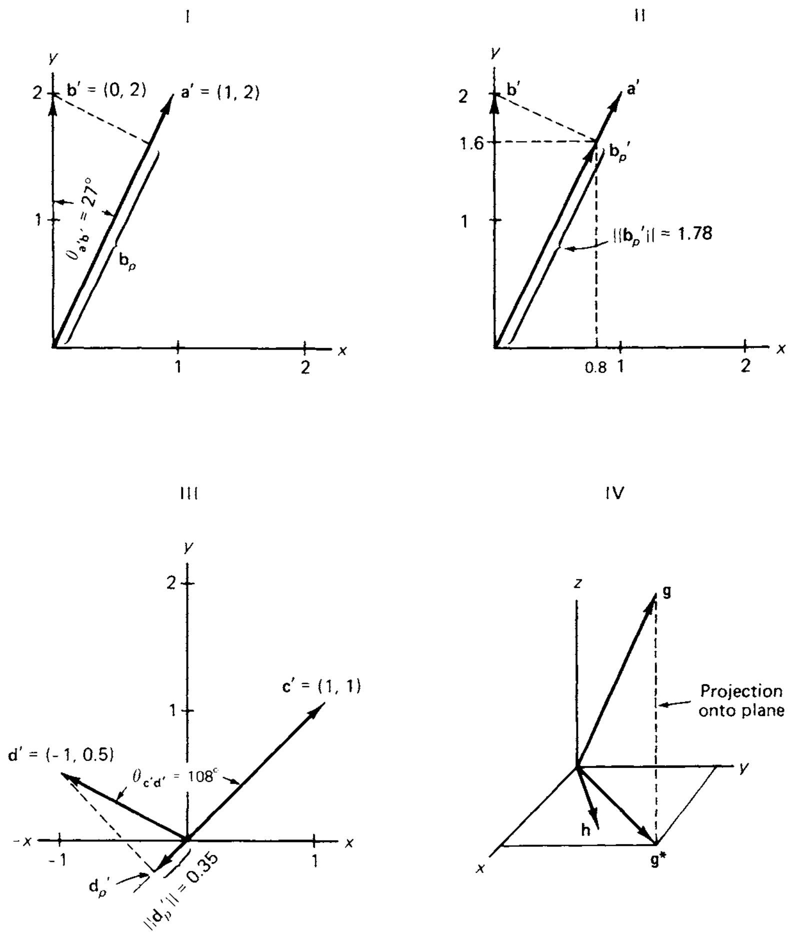

Fig. 3.15 Geometric interpretation of vector projection.

Panel I of Fig. 3.15 shows the projection vector $\mathbf{b}_p'$ along the direction of $\mathbf{a}'$. We note that $\mathbf{b}'$ makes an angle of 27° with $\mathbf{a}'$. First, let us consider the projection *vector* $\mathbf{b}_p'$, and then let us consider its length. The projection vector is found by the formula

$$\boxed{\mathbf{b}_p' = \left[\frac{\|\mathbf{b}'\| \cos \theta_{\mathbf{a}'\mathbf{b}'}}{\|\mathbf{a}'\|}\right]\mathbf{a}'}$$

In terms of the problem, we have

$$\mathbf{b}_p' = \left[\frac{2(0.89)}{\sqrt{5}}\right](1, 2) = (0.8, 1.6)$$

Panel II of Fig. 3.15 shows the coordinates of $\mathbf{b}_p' = (0.8, 1.6)$. Since the angle $\theta_{\mathbf{a}'\mathbf{b}'} = 27°$ is acute, $\mathbf{b}_p'$ is in the *same* direction as $\mathbf{a}'$.

The length of $\mathbf{b}_p'$ is given by

$$\|\mathbf{b}_p'\| = \left| \|\mathbf{b}'\| \cos \theta_{\mathbf{a}'\mathbf{b}'} \right| = 2(0.89) = 1.78$$

and this also appears in Panel II of Fig. 3.15.

Should we desire the length of the projection of a along b, this is obtained analogously as

$$\|\mathbf{a}_p'\| = \left| \|\mathbf{a}'\| \cos \theta_{\mathbf{a}'\mathbf{b}'} \right| = \sqrt{5}(0.89) = 2$$

Notice that if $\mathbf{a}'$ and $\mathbf{b}'$ are each of unit length, the length of the projection of $\mathbf{b}'$ along $\mathbf{a}'$ (or $\mathbf{a}'$ along $\mathbf{b}'$) is simply $|\cos \theta_{\mathbf{a}'\mathbf{b}'}|$.

If the two vectors should make an obtuse angle with each other, the procedure remains the same, but the direction of the projection vector is opposite to that of the reference vector. Panel III of Fig. 3.15 shows a case in which $\mathbf{c}' = (1, 1)$ and $\mathbf{d}' = (-1, 0.5)$ make an angle of 108° with each other. The cosine of this angle is -0.316, and we have

$$\|\mathbf{d}_p'\| = \left| \|\mathbf{d}'\| \cos \theta_{\mathbf{c}'\mathbf{d}'} \right| = \left| 1.12(-0.316) \right| = 0.35$$

as shown in Panel III. However, since $\cos \theta_{\mathbf{c}'\mathbf{d}'}$ is negative, the direction of $\mathbf{d}_p'$ is opposite to that of $\mathbf{c}'$.

The idea of (orthogonal) projection can, of course, be extended to the projection of a vector in three dimensions into a subspace, such as the xy plane in Panel IV of Fig. 3.15. For example, the vector g can be projected into the xy plane by dropping a perpendicular from the terminus of g to the xy plane. The distance between the foot of the projection (represented by the terminus of $\mathbf{g}^*$) and $\mathbf{g}$ must be the minimum distance between g and the xy plane. Any other vector in the xy plane, such as $\mathbf{h}$, must have a terminus that is farther away from g since the hypotenuse of a right triangle must be longer than either side. Subspace projections are discussed later (in Section 4.6.4).

All of this discussion can be straightforwardly related to the geometric aspects of a scalar product. The scalar product $\mathbf{a}'\mathbf{b}$ was earlier defined in general terms as

$$\boxed{\mathbf{a}'\mathbf{b} = \cos \theta_{\mathbf{ab}} \|\mathbf{a}\| \cdot \|\mathbf{b}\|}$$

which can now be expressed in absolute-value terms as the product of two scalars:

$$|\mathbf{a}'\mathbf{b}| = \|\mathbf{b}_p'\| \cdot \|\mathbf{a}'\| = 1.78(\sqrt{5}) = 4$$

Furthermore, the preceding definition of projection length is fully consistent with an informal description presented at the beginning of the chapter. For example, if we have the vector $\mathbf{a}' = (1, 2)$, its projection lengths onto the standard basis vectors $\mathbf{e}_1'$ and $\mathbf{e}_2'$ are found as follows:

$$\mathbf{a}_{p1}' = \left[\frac{\mathbf{a}'\mathbf{e}_1}{\|\mathbf{e}_1'\| \cdot \|\mathbf{e}_1'\|} \right] \mathbf{e}_1' = \left[\frac{\|\mathbf{a}\| \cos \theta_{\mathbf{a}'\mathbf{e}_1'}}{\|\mathbf{e}_1'\|} \right] \mathbf{e}_1' = \left[\frac{(1,2)'(1,0)}{1} \right] (1,0) = (1,0)$$

It follows that

$$\|\mathbf{a}'_{p1}\| = 1$$

$$\mathbf{a}'_{p2} = \left[\frac{\mathbf{a}'\mathbf{e}_2}{\|\mathbf{e}'_2\| \cdot \|\mathbf{e}'_2\|}\right]\mathbf{e}_2{}' = \left[\frac{(1,2)'(0,1)}{1}\right](0,1) = (0,2)$$

and

$$\|\mathbf{a}'_{p2}\| = 2$$

Incidentally, we shall always take $\theta_{\mathbf{a}'\mathbf{b}'}$ to be the smaller angle between $\mathbf{a}'$ and $\mathbf{b}'$. If the vectors are oppositely directed, the direction of the projection will be the negative of the reference vector's direction since $\cos\theta_{\mathbf{a}'\mathbf{b}'}$ will be negative, as illustrated in Panel III of Fig. 3.15.

3.3.6 Recapitulation

At this point we have provided geometric interpretations of all the various algebraic operations on vectors that were illustrated in Chapter 2. In particular, the addition of two vectors followed a parallelogram rule, as illustrated in Fig. 3.9. Subtraction of two vectors also involved a parallelogram rule, in which the difference vector was displaced so as to start at the origin; this is shown in Fig. 3.10.

Multiplication of a vector by a scalar k involves stretching the vector if $k > 1$ and compressing it if $0 < k < 1$. These cases are illustrated in Fig. 3.11. If $k = 1$, the vector remains unchanged. If $k = 0$ the vector becomes $\mathbf{0}$, the zero vector. If k is negative, the vector is stretched and oppositely directed if $|k| > 1$ and compressed and oppositely directed if $0 < |k| < 1$, as shown in Fig. 3.11.

The operations of sum and difference between two (or more) vectors and multiplication of a vector by a scalar are summarized in terms of the concept of linear combination, as illustrated in Fig. 3.12.

The definition of a Euclidean space enabled us to consider the distance and angle between two vectors. By means of the cosine law, illustrated in Fig. 3.13, the cosine of the angle formed by two vectors and their lengths were related to the (squared) Euclidean distance between them. This concept, in turn, led to the geometric portrayal of the projection of one vector onto another, as illustrated in Fig. 3.15. From here it was a short step toward portraying the scalar product of two vectors as a *signed* distance involving the product of the component (projection length) of one vector along some reference vector and the reference vector's length. In short, all of the algebraic operations of Chapter 2 involving vectors were given geometric interpretations here.

We can summarize the various formulas involving aspects of the scalar product as follows:

1. $\|\mathbf{a}-\mathbf{b}\|^2 = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - 2\|\mathbf{a}\| \cdot \|\mathbf{b}\| \cos\theta_{\mathbf{ab}} = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - 2[\mathbf{a}'\mathbf{b}]$

2. $\cos\theta_{\mathbf{ab}} = \dfrac{\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - \|\mathbf{a}-\mathbf{b}\|^2}{2\|\mathbf{a}\| \cdot \|\mathbf{b}\|} = \dfrac{\mathbf{a}'\mathbf{b}}{\|\mathbf{a}\| \cdot \|\mathbf{b}\|}$

3. $$\|\mathbf{b}_p\| = \left|\frac{\mathbf{a}'\mathbf{b}}{\|\mathbf{a}\|}\right| = \left|\,\|\mathbf{b}\|\cos\theta_{\mathbf{ab}}\right|$$

$$\|\mathbf{a}_p\| = \left|\frac{\mathbf{a}'\mathbf{b}}{\|\mathbf{b}\|}\right| = \left|\,\|\mathbf{a}\|\cos\theta_{\mathbf{ab}}\right|$$

It is worth noting that the scalar product plays a central role in all of these formulas.

3.4 LINEAR DEPENDENCE OF VECTORS

In the beginning of the chapter we chose a set of reference vectors $\mathbf{e}_i'$, called standard basis vectors, that in three dimensions were defined as follows:

$$\mathbf{e}_1' = (1,0,0); \qquad \mathbf{e}_2' = (0,1,0); \qquad \mathbf{e}_3' = (0,0,1)$$

As we shall see in a moment this set of vectors is linearly independent. The concept of linear independence plays a major role in vector algebra and multivariate analysis. As we know from elementary geometry, a line is one-dimensional, an area is two-dimensional, and a volume is three-dimensional. By analogy, a space of n dimensions entails "hypervolume."

Loosely speaking, linear independence of vectors has to do with the minimum number of vectors in terms of which any given vector in the space can be expressed and, in effect, is related to the "volume" of the space spanned by the vectors. Linearly dependent vectors display a kind of redundancy or superfluity in the sense that at least one vector of a linearly dependent set can be written as a linear combination of the other vectors.

Somewhat more formally, if $\mathbf{a}_1', \mathbf{a}_2', \ldots, \mathbf{a}_p'$ denote a set of p vectors and $k_1, k_2, \ldots, k_p$ denote a set of p scalars, it may be the case that the following linear equation is satisfied:

$$\boxed{k_1\mathbf{a}_1' + k_2\mathbf{a}_2' + \cdots + k_p\mathbf{a}_p' = \mathbf{0}'}$$

where $\mathbf{0}'$ is the zero vector.

For example, if $k_1 = k_2 = \ldots = k_p = 0$, *any* set of p vectors trivially satisfies the above equation. If, however, the equation can be satisfied *without* all k_i being equal to zero, the solution is called "nontrivial."

If a nontrivial solution can be found, then we say that the set of vectors is linearly dependent. If only the trivial solution is satisfied, the set of vectors is said to be linearly independent.

To illustrate the case of nontrivial satisfaction of the above equation, let us assume three four-component vectors:

$$\mathbf{a}_1' = (1,2,0,4); \qquad \mathbf{a}_2' = (-1,0,5,1); \qquad \mathbf{a}_3' = (1,6,10,14)$$

and let

$$k_1 = 3; \qquad k_2 = 2; \qquad k_3 = -1$$

Since

$$\begin{aligned} &3(1,2,0,4) + 2(-1,0,5,1) - 1(1,6,10,14) \\ &\qquad = (3,6,0,12) + (-2,0,10,2) + (-1,-6,-10,-14) = (0,0,0,0) \end{aligned}$$

it is seen that $\mathbf{a}_1'$, $\mathbf{a}_2'$, and $\mathbf{a}_3'$ are linearly dependent and at least one of the vectors is a linear combination of the remaining $p - 1$ vectors. To see that this is so, we note that

$$\begin{aligned} \mathbf{a}_1' &= -\frac{k_2}{k_1}(\mathbf{a}_2') - \frac{k_3}{k_1}(\mathbf{a}_3') \\ &= -\tfrac{2}{3}(-1,0,5,1) + \tfrac{1}{3}(1,6,10,14) \\ &= (\tfrac{2}{3}, 0, -\tfrac{10}{3}, -\tfrac{2}{3}) + (\tfrac{1}{3}, 2, \tfrac{10}{3}, \tfrac{14}{3}) \\ \mathbf{a}_1' &= (1,2,0,4) \end{aligned}$$

and $\mathbf{a}_1'$ is, indeed, a linear combination of $\mathbf{a}_2'$ and $\mathbf{a}_3'$. It is also pertinent to note that any set of p vectors is *always* linearly dependent if $p > n$, where n is the number of vector components in an n by 1 column vector or a 1 by n row vector, as the case may be.

While no proof of this assertion is given, the statement relates to the fact that if one wished to solve n equations for n unknowns, one could take the first n vectors, assuming they are linearly independent, and solve for any of the other vectors as linear combinations of these n linearly independent vectors.

The concept of linear independence is of particular importance to multivariate analysis. A set of linearly independent vectors is said to *span* some Euclidean space of interest. Ultimately the idea of linear independence relates to the *dimensionality* of the space in which the researcher is working. And, as we shall see, once a set of such vectors is found, all other vectors can be expressed as linear combinations of these.

In brief, then, two ideas are involved in the study of linear independence. First, we wish to find a set of *nonredundant* vectors. Second, we wish to make sure that we have *enough* linearly independent vectors to span some space of interest or, as indicated earlier, to contain some hypervolume of interest.

3.4.1 Dimensionality of a Vector Space and the Concept of Basis

In line with our earlier discussions involving geometric analogy, we can now examine the dimensionality of a vector space. *The dimensionality of a vector space is equal to the maximum number of linearly independent vectors in that space.* To illustrate for the case of three dimensions, we return to the $\mathbf{e}_i'$ standard coordinate vectors:

$$\mathbf{e}_1' = (1,0,0); \qquad \mathbf{e}_2' = (0,1,0); \qquad \mathbf{e}_3' = (0,0,1)$$

If we set up the equation

$$k_1 \mathbf{e}_1' + k_2 \mathbf{e}_2' + k_3 \mathbf{e}_3' = 0$$

we find that the above equation is satisfied only if $k_1 = k_2 = k_3 = 0$. Hence, $\mathbf{e}_1'$, $\mathbf{e}_2'$, and $\mathbf{e}_3'$ are linearly independent, and the dimensionality of the space is three. *In general, if* $\mathbf{a}_1', \mathbf{a}_2', \ldots, \mathbf{a}_n'$ *denote a set of n linearly independent n-component vectors, then any other vector of that n-space can be written as*

$$\boxed{\mathbf{b}' = k_1 \mathbf{a}_1' + k_2 \mathbf{a}_2' + \cdots + k_n \mathbf{a}_n'}$$

and the $\mathbf{a}_1', \mathbf{a}_2', \ldots, \mathbf{a}_n'$ *vectors are said to constitute a basis of the n-space.* In a space of n dimensions, *any* set of n linearly independent vectors can constitute a basis of the space. Thus the basis vectors $\mathbf{e}_i'$ above represent only one type of basis, one that we have called the standard basis.

The $\mathbf{e}_i'$ standard basis vectors, however, are particularly convenient. *Indeed, unless stated otherwise we shall assume that the particular basis being chosen is the standard basis.* Still, we should indicate that any other set of n linearly independent vectors could qualify as the reference set. Accordingly, we spend some time on the process by which one can change one set of basis vectors to some other set, for example, to a set of standard basis vectors.

3.4.2 Change of Basis Vectors

Up to this point we have emphasized rectangular Cartesian coordinates, where it is natural to view the coordinate vectors $\mathbf{e}_i'$ as both (a) mutually orthogonal (i.e., exhibiting pairwise scalar products of zero) and (b) of unit length. *This type of basis is called orthonormal.* In this intuitively simple case, the vector $\mathbf{a}' = (a_1, a_2, \ldots, a_n)$ can be easily written as

$$\boxed{\mathbf{a}' = a_1 \mathbf{e}_1' + a_2 \mathbf{e}_2' + \cdots + a_n \mathbf{e}_n'}$$

where $\mathbf{e}_1' = (1, 0, \ldots, 0)$, $\mathbf{e}_2' = (0, 1, 0, \ldots, 0)$, and $\mathbf{e}_n' = (0, 0, \ldots, 1)$. Hence $(a_1, a_2, \ldots, a_n)$ are the coordinates of $\mathbf{a}'$ relative to the orthonormal basis $\mathbf{e}_1', \mathbf{e}_2', \ldots, \mathbf{e}_n'$.

An equally satisfactory way of showing this concept is to represent the standard coordinate vectors in columnar form. A given vector $\mathbf{a}$ can then be written as

$$\mathbf{a} = a_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \cdots + a_n \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix}$$

and we have an illustration of a linear combination of standard basis vectors in which the components of $\mathbf{a}$ (i.e., a_1, a_2, etc.) are the scalars of interest.

Orthonormal bases are easy to work with, and we shall usually assume that this type of basis, more specifically, the standard basis vectors $\mathbf{e}_i$, underlies the coordinate representation of interest. However, as indicated above, any set of linearly independent vectors, unit length or not, orthogonal or not, can be used to define a basis. Hence, it is pertinent to point out how one can move from one basis of a space to some other basis of that space.

Accordingly, let us now illustrate the idea of *general* coordinate systems whose basis vectors need not be mutually orthogonal or of unit length. Suppose we start with two sets of basis vectors—first, the more familiar $\mathbf{e}_i'$ standard basis vectors, $\mathbf{e}_1' = (1, 0)$ and $\mathbf{e}_2' = (0, 1)$, and second, another set of basis vectors $\mathbf{f}_1' = c_1\mathbf{e}_1' + c_2\mathbf{e}_2'$ and $\mathbf{f}_2' = d_1\mathbf{e}_1' + d_2\mathbf{e}_2'$.

To be specific, we let $c_1 = 0.707, c_2 = 0.707, d_1 = 0.940$, and $d_2 = 0.342$. Then

$$\mathbf{f}_1' = 0.707\mathbf{e}_1' + 0.707\mathbf{e}_2' = (0.707, 0.707)$$

$$\mathbf{f}_2' = 0.940\mathbf{e}_1' + 0.342\mathbf{e}_2' = (0.940, 0.342)$$

Note that $\mathbf{f}_1'$ and $\mathbf{f}_2'$ are each of unit length but are not orthogonal; that is, $\mathbf{f}_1'\mathbf{f}_2 \neq 0$. Note further that we can write the preceding equations in columnar form as

$$\mathbf{f}_1 = \begin{bmatrix} f_{11} \\ f_{21} \end{bmatrix} = 0.707 \overset{\mathbf{e}_1}{\begin{bmatrix} 1 \\ 0 \end{bmatrix}} + 0.707 \overset{\mathbf{e}_2}{\begin{bmatrix} 0 \\ 1 \end{bmatrix}} = \begin{bmatrix} 0.707 \\ 0.707 \end{bmatrix}$$

$$\mathbf{f}_2 = \begin{bmatrix} f_{12} \\ f_{22} \end{bmatrix} = 0.940 \overset{\mathbf{e}_1}{\begin{bmatrix} 1 \\ 0 \end{bmatrix}} + 0.342 \overset{\mathbf{e}_2}{\begin{bmatrix} 0 \\ 1 \end{bmatrix}} = \begin{bmatrix} 0.940 \\ 0.342 \end{bmatrix}$$

Figure 3.16 shows a plot of $\mathbf{f}_1$ and $\mathbf{f}_2$ relative to the standard basis $\mathbf{e}_1$ and $\mathbf{e}_2$. By finding their projections on $\mathbf{e}_1$ and $\mathbf{e}_2$, we can note that their coordinates are given by the preceding equations.

Now let us select a new vector $\mathbf{a} = a_1\mathbf{f}_1 + a_2\mathbf{f}_2$. That is, we shall assume that $\mathbf{a}$ is referred to the new (and nonorthogonal) basis, $\mathbf{f}_1$ and $\mathbf{f}_2$. To be specific, we assume that the coordinates of $\mathbf{a}$ relative to $\mathbf{f}_1$ and $\mathbf{f}_2$ are

$$a_1 = 0.5; \qquad a_2 = 0.5$$

We can find these coordinates by extending lines parallel to $\mathbf{f}_1$ and $\mathbf{f}_2$ and noting the coordinates of OQ and OR, respectively, on $\mathbf{f}_1$ and $\mathbf{f}_2$. The basis vectors $\mathbf{f}_1$ and $\mathbf{f}_2$ are often called *oblique* Cartesian axes since the angle that they make with each other is *not* equal to 90°. Notice, however, that $\mathbf{a}$ is still given by the (parallelogram) law for vector addition:

$$\mathbf{a} = a_1\mathbf{f}_1 + a_2\mathbf{f}_2$$

and that $a_1\mathbf{f}_1$ and $a_2\mathbf{f}_2$ are scalar multiples of $\mathbf{f}_1$ and $\mathbf{f}_2$, respectively. Since we have chosen $\mathbf{f}_1$ and $\mathbf{f}_2$ to be of unit length, the coordinates a_1 and a_2 are merely the lengths OQ and OR. Had $\mathbf{f}_1$ and $\mathbf{f}_2$ not been of unit length, a_1 and a_2 would still be regarded as

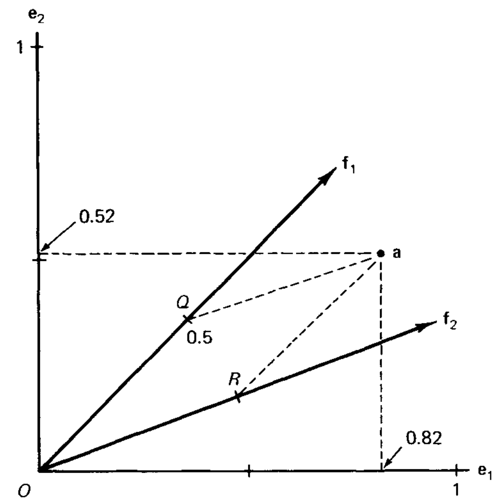

Fig. 3.16 An illustration of generalized coordinates and change of basis.

coordinates, but they would correspond to lengths of $OQ/\|\mathbf{f}_1\|$ and $OR/\|\mathbf{f}_2\|$, respectively.

We now seek a set of coordinates for the vector a, referred to the oblique basis $\mathbf{f}_1$ and $\mathbf{f}_2$, in terms of the *original* (and standard) basis $\mathbf{e}_1$ and $\mathbf{e}_2$. This can be done by the following substitution:

$$\mathbf{a} = a_1\mathbf{f}_1 + a_2\mathbf{f}_2$$

But, since $\mathbf{f}_1$ and $\mathbf{f}_2$ have been defined in terms of $\mathbf{e}_1$ and $\mathbf{e}_2$, we can write

$$\mathbf{a} = a_1(c_1\mathbf{e}_1 + c_2\mathbf{e}_2) + a_2(d_1\mathbf{e}_1 + d_2\mathbf{e}_2)$$
$$= (c_1a_1 + d_1a_2)\mathbf{e}_1 + (c_2a_1 + d_2a_2)\mathbf{e}_2$$

However, since $\mathbf{e}_1$ and $\mathbf{e}_2$ denote a basis, we can also represent a in terms of $\mathbf{e}_1$ and $\mathbf{e}_2$ as

$$\mathbf{a} = a_1{}^*\mathbf{e}_1 + a_2{}^*\mathbf{e}_2$$

Hence, through substitution of $c_1a_1 + d_1a_2$ for $a_1{}^*$, and $c_2a_2 + d_2a_2$ for $a_2{}^*$, we find

$$a_1{}^* = c_1a_1 + d_1a_2 = 0.707(0.5) + 0.940(0.5) = 0.82$$
$$a_2{}^* = c_2a_1 + d_2a_2 = 0.707(0.5) + 0.342(0.5) = 0.52$$

As can be observed from Fig. 3.16, the length of the projection of a on $\mathbf{e}_1$ is, indeed, 0.82, and its projection length on $\mathbf{e}_2$ is 0.52.

Thus, one can work "backward" to relate a vector described in terms of one set of basis vectors to a description of that same vector in terms of another set of basis vectors, assuming we know how the basis vectors themselves are connected. And, as a matter of fact, one can *always* find an orthonormal set of axes (mutually orthogonal and of unit length) by which a set of arbitrary basis vectors can be represented, even though the original axes might be oblique and not of unit length. The next section illustrates one procedure for finding an orthonormal basis from an initial set of nonorthonormal basis vectors.

3.4.3 Finding an Orthonormal Basis

As indicated earlier, a special kind of basis in a vector space—one of particular value in multivariate analysis—is an orthonormal basis. This basis is characterized by the facts that (a) the scalar product of any pair of basis vectors is zero and (b) each basis vector is of unit length. As we know, the standard basis vectors $\mathbf{e}_i$ represent one such orthonormal basis.

In multivariate data analysis, it is usually the case that multiple measurements on a set of objects will be associated; for example, weight will be correlated with height. Sometimes we may want to transform the original (and correlated) variables to a set of uncorrelated variables. As will be shown later, this process can be viewed as transforming a set of n nonorthogonal vectors into a set of n orthogonal vectors. In the process we may also want to make all of these vectors unit length; this is the "norming" aspect of the process.

We have already observed that the scalar product is a central concept in vector algebra and is a function that assigns a real number to each pair of vectors in the Euclidean space of interest. *In particular, the concepts of vector length, distance, and cosine can all be expressed in terms of the single idea of a scalar product:*

$$\boxed{\begin{aligned} \|\mathbf{a}\| &= [\mathbf{a}'\mathbf{a}]^{1/2} \\ \|\mathbf{a}-\mathbf{b}\| &= [\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - 2(\mathbf{a}'\mathbf{b})]^{1/2} \\ \cos\theta_{\mathbf{ab}} &= \frac{\mathbf{a}'\mathbf{b}}{\|\mathbf{a}\|\cdot\|\mathbf{b}\|} \end{aligned}}$$

As we shall see in a moment, the scalar product also provides a simple representation of vectors that are mutually orthogonal (perpendicular):

$$\boxed{\begin{gathered} \mathbf{a} \text{ and } \mathbf{b} \text{ are orthogonal} \\ \text{if and only if} \\ \mathbf{a}'\mathbf{b} = 0 \end{gathered}}$$

We can now proceed to construct an orthonormal basis, one whose vectors are mutually orthogonal *and* of unit length.

Any arbitrary basis can be transformed to an orthonormal basis by a procedure known as *Gram–Schmidt orthonormalization.* To illustrate the process, consider the three arbitrary row vectors:

$$\mathbf{a}_1' = (2, 1, 2); \qquad \mathbf{a}_2' = (3, -1, 5); \qquad \mathbf{a}_3' = (0, 1, -1)$$

The Gram–Schmidt process starts out by selecting (arbitrarily) one of the vectors, say $\mathbf{a}_1'$, as the first reference vector.[8] The idea here is to keep this vector fixed and then find other vectors, two other vectors in this case, so that the resultant sets are mutually orthogonal. As a final step each of the orthogonal vectors is normalized to unit length. To start off the process we first set

$$\mathbf{b}_1' = \mathbf{a}_1'$$

and then find

$$\mathbf{b}_2' = \mathbf{a}_2' - \left[\frac{\mathbf{a}_2'\mathbf{b}_1}{\mathbf{b}_1'\mathbf{b}_1}\right]\mathbf{b}_1' = \mathbf{a}_2' - \left[\frac{(3 \times 2) + (-1 \times 1) + (5 \times 2)}{2^2 + 1^2 + 2^2}\right]\mathbf{b}_1'$$

$$= (3, -1, 5) - (15/9)(2, 1, 2) = (-1/3, -8/3, 5/3)$$

$$\mathbf{b}_2' = (-0.33, -2.67, 1.67)$$

Let us now examine the expression

$$\left[\frac{\mathbf{a}_2'\mathbf{b}_1}{\mathbf{b}_1'\mathbf{b}_1}\right]\mathbf{b}_1' = (15/9)(2, 1, 2) = (10/3, 5/3, 10/3)$$

This expression is the orthogonal projection of $\mathbf{a}_2'$ onto $\mathbf{b}_1'$ (as discussed in Section 3.3.5).

The "residual" is then equal to the difference

$$\mathbf{b}_2' = \mathbf{a}_2' - \left[\frac{\mathbf{a}_2'\mathbf{b}_1}{\mathbf{b}_1'\mathbf{b}_1}\right]\mathbf{b}_1' = (3, -1, 5) - (10/3, 5/3, 10/3)$$

$$= (-1/3, -8/3, 5/3)$$

and should be orthogonal to $\mathbf{b}_1'$, as is shown illustratively in *two* dimensions, in Fig. 3.17. That is,

$$\mathbf{b}_2'\mathbf{b}_1 = (-1/3, -8/3, 5/3)\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} = 0$$

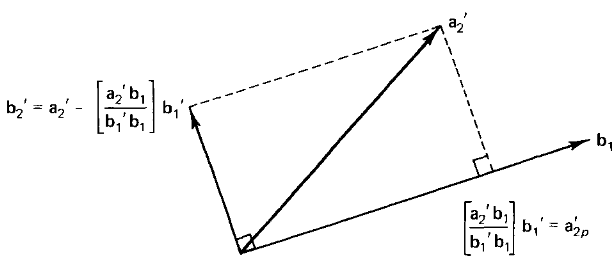

Fig. 3.17 Finding the orthogonal projection of $\mathbf{a}_2'$ onto $\mathbf{b}_1'$ (illustrated in two dimensions).

[8] It should be mentioned that the specific results of the Gram–Schmidt process depend on the order on which the vectors are selected; however, in any case the resulting set will be orthogonal and of unit length.

We encountered orthogonal projections, both in the beginning of the chapter and in Section 3.3.5. Accordingly, there is nothing new here, except for the fact that we are now interested in the *orthogonal complement* or that part of the $\mathbf{a}_2'$ vector that does *not* lie along the reference vector. In this case it is the vector

$$\mathbf{b}_2' = (-0.33, -2.67, 1.67)$$

As we noted above, the scalar product of $\mathbf{b}_2'\mathbf{b}_1$ is indeed zero. Thus, $\mathbf{b}_2'$ is now orthogonal to $\mathbf{b}_1' = \mathbf{a}_1'$. We now have to orthogonalize $\mathbf{a}_3'$ with regard to the two, already orthogonal, vectors $\mathbf{b}_2'$ and $\mathbf{b}_1'$:

$$\begin{aligned}
\mathbf{b}_3' &= \mathbf{a}_3' - \left[\frac{\mathbf{a}_3'\mathbf{b}_2}{\mathbf{b}_2'\mathbf{b}_2}\right]\mathbf{b}_2' - \left[\frac{\mathbf{a}_3'\mathbf{b}_1}{\mathbf{b}_1'\mathbf{b}_1}\right]\mathbf{b}_1' \\
&= (0, 1, -1) + \frac{13}{30}(-1/3, -8/3, 5/3) + \frac{1}{9}(2, 1, 2) \\
&= (0, 1, -1) + (-13/90, -104/90, 65/90) + (2/9, 1/9, 2/9) \\
&= (7/90, -4/90, -5/90) \\
\mathbf{b}_3' &= (0.08, -0.04, -0.06)
\end{aligned}$$

And, in general for r vectors, we would have

$$\mathbf{b}_r' = \mathbf{a}_r' - \left[\frac{\mathbf{a}_r'\mathbf{b}_{r-1}}{\mathbf{b}_{r-1}'\mathbf{b}_{r-1}}\right]\mathbf{b}_{r-1}' - \cdots - \left[\frac{\mathbf{a}_r'\mathbf{b}_1}{\mathbf{b}_1'\mathbf{b}_1}\right]\mathbf{b}_1'$$

After the $\mathbf{b}'$'s are obtained, we would find that they are mutually orthogonal. Each set is then normalized by its respective divisor $\|\mathbf{b}_i'\|$. That is, we find the length of each of the $\mathbf{b}'$'s and divide each vector component by the length of that vector. In the above example, the lengths of $\mathbf{b}_1'$, $\mathbf{b}_2'$, and $\mathbf{b}_3'$, respectively, are 3, 3.17, and 0.108. The normalized vectors then become

$$\begin{aligned}
\mathbf{b}_1^{*\prime} &= (1/3)(2, 1, 2) = (0.67, 0.33, 0.67) \\
\mathbf{b}_2^{*\prime} &= (1/3.17)(-0.33, -2.67, 1.67) = (-0.10, -0.84, 0.53) \\
\mathbf{b}_3^{*\prime} &= (1/0.108)(0.08, -0.04, -0.06) = (0.74, -0.37, -0.56)
\end{aligned}$$

Within rounding error, we first note that all three vectors have unit length. If we then find the scalar product of each pair of vectors, we observe, again within rounding error, that all three scalar products equal zero.

We conclude by saying that the vectors $\mathbf{b}_1^{*\prime}$, $\mathbf{b}_2^{*\prime}$, $\mathbf{b}_3^{*\prime}$ form a three-dimensional *orthonormal* basis—one whose axes are mutually orthogonal and of unit length.

It is rather difficult to show the Gram–Schmidt procedure for the specific vectors utilized in our example. This being the case, Fig. 3.18 shows a more stylized conceptualization of the procedure. The pictures first show orthonormalization of the first two vectors in two dimensions and then orthonormalization of all three in three dimensions. (In the figure the orthonormalized vectors are expressed as column vectors.)

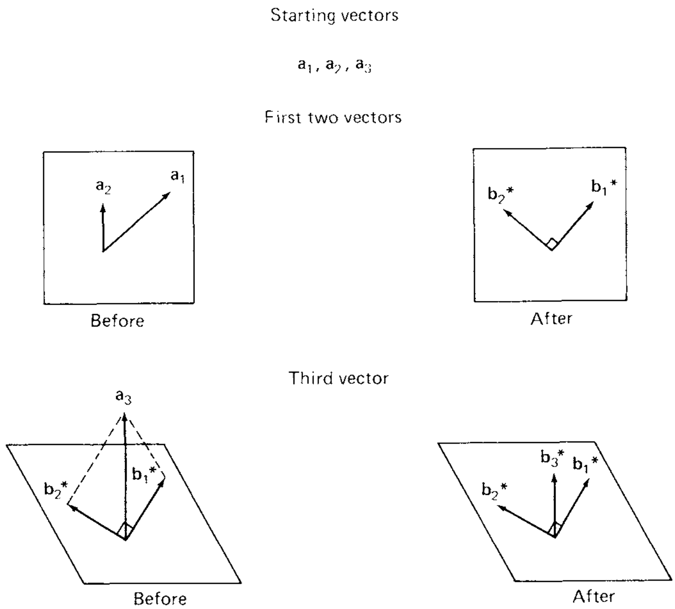

Fig. 3.18 Conceptualization of Gram–Schmidt orthonormalization procedure.

3.4.4 Scalar Products in Oblique Coordinate Systems

Now that we have discussed both oblique and orthonormal bases, it is of interest to point out that vector addition and subtraction as well as multiplication of a vector by a scalar are carried out the same way under either oblique or orthonormal basis conditions. However, this correspondence does *not* hold in the case of the scalar product.

The reason why the usual scalar product formula does not work in the nonorthogonal case is most easily seen by observing the two basis vectors **a** and **b** in the diagram of Fig. 3.19. Note that the angle they make with each other is 45° rather than 90°. If we

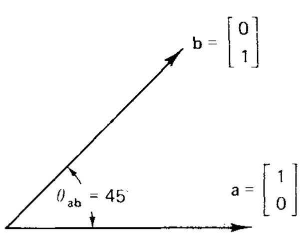

Fig. 3.19 Scalar product in oblique coordinate system.

(incorrectly) assumed that these basis vectors were orthogonal, their scalar product would be zero by application of the special case of a scalar product that entails $\mathbf{a}'\mathbf{b}$.[9]

However, using the formula described in Section 3.3.4, we find the following:

$$\mathbf{a}'\mathbf{b} = \|\mathbf{a}\| \cdot \|\mathbf{b}\| \cos \theta_{\mathbf{ab}} = 1(1)(0.707) = 0.707$$

As noted above, this form of computing the scalar product is *not* dependent on referring vectors to an orthonormal basis. Still, as can be observed from this section, orthonormal bases simplify calculations quite a lot.

In more general terms, if we have two column vectors $\mathbf{a}$ and $\mathbf{b}$ referred to oblique unit length basis vectors $\mathbf{f}_i$, this situation can be represented as

$$\mathbf{a} = a_1\mathbf{f}_1 + a_2\mathbf{f}_2 + \cdots + a_n\mathbf{f}_n$$

and

$$\mathbf{b} = b_1\mathbf{f}_1 + b_2\mathbf{f}_2 + \cdots + b_n\mathbf{f}_n$$

In this case the scalar product between $\mathbf{a}$ and $\mathbf{b}$ is given by

$$\boxed{\mathbf{a}'\mathbf{b} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i b_j \cos \theta_{ij}}$$

where θ_{ij} is the *angle* between the pair of basis vectors $\mathbf{f}_i$ and $\mathbf{f}_j$ for $i, j = 1, 2, \ldots, n$. However, if $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ are also orthogonal, $\cos \theta_{ij} = 0$ for all pairs of basis vectors in which $j \neq i$ and, hence, we have the special case, discussed earlier, of

$$\mathbf{a}'\mathbf{b} = \sum_{i=1}^{n} a_i b_i$$

What if the oblique $\mathbf{f}_i$ and $\mathbf{f}_j$ are not of unit length? If this is the case, then the scalar product becomes

$$\boxed{\mathbf{a}'\mathbf{b} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i b_j \cos \theta_{ij} \|\mathbf{f}_i\| \cdot \|\mathbf{f}_j\|}$$

Thus, if we need to account for basis vectors whose lengths are not equal to unity, the more general expression above is applicable.[10]

[9] It is important to note that the definition of scalar product as $\mathbf{a}'\mathbf{b} = \sum_{i=1}^{n} a_i b_i$ in Chapter 2 has implicitly assumed that both vectors are referred to standard basis vectors $\mathbf{e}_i$. In general, this will indeed be the case; however, in oblique coordinate systems the geometrically oriented definition $\mathbf{a}'\mathbf{b} = \|\mathbf{a}\| \cdot \|\mathbf{b}\| \cos \theta_{\mathbf{ab}}$ should be used.

[10] The computation of scalar products in an oblique coordinate system entails the concept of a (positive definite) quadratic form, a topic that is discussed in Chapter 5. Throughout the book, however, we shall emphasize the simpler case in which the standard basis vectors $\mathbf{e}_i$ are assumed to be applicable.

3.5 ORTHOGONAL TRANSFORMATIONS

Up to this point we have presented geometric interpretations of all of the principal algebraic operations that were performed on vectors in Chapter 2, such as addition, subtraction, the scalar product of two vectors, and so on. So far, matrices have been ignored for the most part, except in our discussion of basis vectors.

It is now appropriate to discuss some preliminary aspects of a matrix transformation of a vector or set of vectors. In this chapter we limit our discussion to a special but quite important case, namely, orthogonal transformations or rotations.

Most readers probably have an intuitive idea about what is meant by a rigid rotation of a set of points. Often in multivariate analysis we wish to perform a transformation on a set of points that will preserve their angles, lengths, and interpoint distances, while at the same time referring them to a new, perhaps simpler, coordinate system. Since rotations, as a special type of linear transformation, play such a key role in understanding more general kinds of matrix transformations, this concept is introduced here and related to earlier discussions of distance and angle.

The definition of an orthogonal matrix as used in multivariate analysis differs somewhat from researcher to researcher. We shall use the term "orthogonal" to refer to a square matrix $\mathbf{A}$ that exhibits the property

$$\boxed{\mathbf{A}'\mathbf{A} = \mathbf{A}\mathbf{A}' = \mathbf{I}}$$

That is, any two column vectors or any two row vectors in the matrix $\mathbf{A}$ are mutually orthogonal and, furthermore, each vector is of unit length. Some authors call this type of matrix "square, orthonormal," but we shall use the more common term of orthogonal matrix.

3.5.1 Axis and Point Rotations

To motivate the discussion let us consider the column vector $\mathbf{a} = \begin{bmatrix}2\\1\end{bmatrix}$ in the diagram of Fig. 3.20. We adopt a set of standard basis vectors $\mathbf{e}_i$ for the space in order to simplify

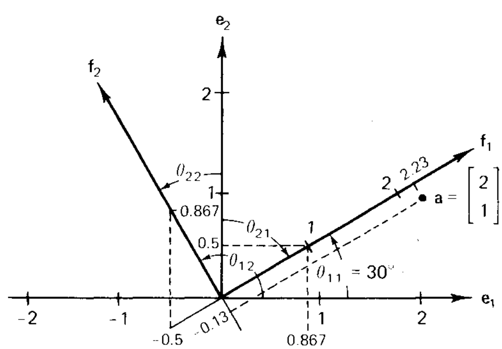

Fig. 3.20 Rotation of reference axes.

subsequent discussion. Now suppose we wish to apply an orthogonal transformation to the vector a. Geometrically, this can mean one of two things:

1. we can rigidly rotate the *axes*, either counterclockwise or clockwise, from their original $\mathbf{e}_i$ orientation, while leaving the point fixed, or
2. we can leave the original $\mathbf{e}_i$ axes fixed and rigidly rotate the *vector* $\mathbf{a} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ to a new location.

Let us consider the first case. Suppose we wish to rotate the original axes $\mathbf{e}_1$ and $\mathbf{e}_2$ counterclockwise through an angle of 30°, as shown in Fig. 3.20. To do this we shall need a set of direction cosines for each angle made by the (new) $\mathbf{f}_1$ and $\mathbf{f}_2$ axes with the (original) $\mathbf{e}_1$ and $\mathbf{e}_2$ axes. Let us first find the cosines that we shall need. From basic trigonometry we have

$$\cos 30^\circ = \sqrt{3}/2 = 0.867; \qquad \cos 60^\circ = \tfrac{1}{2} = 0.5; \qquad \cos 120^\circ = -\tfrac{1}{2} = -0.5$$

Next, we shall use the symbol θ_{ij} to denote angles between pairs of axes, where i denotes the original axis and j denotes the new axis. If we examine the four angles $\theta_{11}, \theta_{12}, \theta_{21}, \theta_{22}$, in which the first subscript refers to the old axis and the second to the new axis in Fig. 3.20, we see that

1. θ_{11} involves a 30° counterclockwise rotation with $\cos 30^\circ = 0.867$.
2. θ_{12} involves a 120° counterclockwise rotation with $\cos 120^\circ = -0.5$.
3. θ_{21} involves a 60° clockwise rotation with $\cos 60^\circ = 0.5$.
4. θ_{22} involves a 30° counterclockwise rotation with $\cos 30^\circ = 0.867$.

The angle of 30° that $\mathbf{f}_1$ makes with $\mathbf{e}_1$ involves a cosine that is equal to 0.867. And, since $\mathbf{f}_1$ makes an angle of 60° (with a cosine of 0.5) with $\mathbf{e}_2$, we have the linear combination

$$\mathbf{f}_1 = \cos\theta_{11}\mathbf{e}_1 + \cos\theta_{21}\mathbf{e}_2 = 0.867\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0.5\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.867 \\ 0.5 \end{bmatrix}$$

as the coordinates of $\mathbf{f}_1$.

Similarly, we can compute the coordinates of $\mathbf{f}_2$ as follows:

$$\mathbf{f}_2 = \cos\theta_{12}\mathbf{e}_1 + \cos\theta_{22}\mathbf{e}_2 = -0.5\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0.867\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.5 \\ 0.867 \end{bmatrix}$$

As can be seen from Fig. 3.20, $\mathbf{f}_1$ and $\mathbf{f}_2$ display the coordinates indicated above. We also note that the sum of squares of each set of direction cosines is unity. That is,

$$(0.867)^2 + (0.5)^2 = 1; \qquad (-0.5)^2 + (0.867)^2 = 1$$

At this point we have expressed $\mathbf{f}_1$ and $\mathbf{f}_2$ in terms of $\mathbf{e}_1$ and $\mathbf{e}_2$. We also know that the (assumed fixed) point a is expressed in terms of $\mathbf{e}_1$ and $\mathbf{e}_2$, the *original* basis vectors, as

$$\mathbf{a} = 2\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Our problem, now, is to find the coordinates of that same point—which we can call a*—in terms of the *new* basis vectors $\mathbf{f}_i$.

We can express this transformation in the form of a matrix postmultiplied by a vector. That is, we can let $\mathbf{a}^* = \left[\begin{smallmatrix} a_1^* \\ a_2^* \end{smallmatrix}\right]$ denote the *new* coordinates of the point a by the following substitution:

$$\begin{bmatrix} a_1^* \\ a_2^* \end{bmatrix} = \begin{bmatrix} \cos\theta_{11} & \cos\theta_{21} \\ \cos\theta_{12} & \cos\theta_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

$$\begin{bmatrix} 2.23 \\ -0.13 \end{bmatrix} = \begin{bmatrix} 0.867 & 0.5 \\ -0.5 & 0.867 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

These are the coordinates of the point with respect to the new basis $\mathbf{f}_i$ in Fig. 3.20.

Let us examine the transformation somewhat more closely in Fig. 3.20. First, as noted above, we see that the unit length portion of $\mathbf{f}_1$ has coordinates of $\mathbf{f}_1 = \left[\begin{smallmatrix} 0.867 \\ 0.5 \end{smallmatrix}\right]$ with respect to $\mathbf{e}_1$ and $\mathbf{e}_2$, respectively. Similarly, the unit length portion of $\mathbf{f}_2$ has coordinates of $\mathbf{f}_2 = \left[\begin{smallmatrix} -0.5 \\ 0.867 \end{smallmatrix}\right]$ with respect to $\mathbf{e}_1$ and $\mathbf{e}_2$, respectively.

However, we can turn the coin over and look at the coordinates of $\mathbf{e}_i$ in terms of the new axes $\mathbf{f}_i$. If we project $\mathbf{e}_1$ and $\mathbf{e}_2$ onto $\mathbf{f}_1$ and $\mathbf{f}_2$, we have, from Fig. 3.20,

$$\mathbf{g}_1 = \begin{bmatrix} 0.867 \\ -0.5 \end{bmatrix}; \qquad \mathbf{g}_2 = \begin{bmatrix} 0.5 \\ 0.867 \end{bmatrix}$$

where we use $\mathbf{g}_1$ and $\mathbf{g}_2$ to denote the fact that the reference vectors are now the $\mathbf{f}_i$. Since a has been defined originally in terms of $\mathbf{e}_i$, and the $\mathbf{e}_i$ have now been represented in terms of $\mathbf{f}_i$, we have

$$\mathbf{a} = 2\mathbf{g}_1 + 1\mathbf{g}_2$$

$$\mathbf{a} = 2\begin{bmatrix} 0.867 \\ -0.5 \end{bmatrix} + 1\begin{bmatrix} 0.5 \\ 0.867 \end{bmatrix} = \begin{bmatrix} 2.23 \\ -0.13 \end{bmatrix}$$

But, as already shown, this can also be written as

$$\begin{bmatrix} a_1^* \\ a_2^* \end{bmatrix} = \begin{bmatrix} \cos\theta_{11} & \cos\theta_{21} \\ \cos\theta_{12} & \cos\theta_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

$$\begin{bmatrix} 2.23 \\ -0.13 \end{bmatrix} = \begin{bmatrix} 0.867 & 0.5 \\ -0.5 & 0.867 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Thus, while the point remains fixed, its coordinates are determined by the particular basis by which they are expressed.

In summary, we see that the new coordinates of the original point $\mathbf{a} = \left[\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right]$ are now $a_1^* = 2.23$ and $a_2^* = -0.13$ in terms of the $\mathbf{f}_1$ and $\mathbf{f}_2$ axes. However, there is a second way of looking at this transformation. That is, we can make the original $\mathbf{e}_1$, $\mathbf{e}_2$ plane *remain the same* and assume that it is the point $\left[\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right]$ which moves from its old location to the position $\left[\begin{smallmatrix} 2.23 \\ -0.13 \end{smallmatrix}\right]$. This second way of interpreting things is shown in Fig. 3.21. Notice in

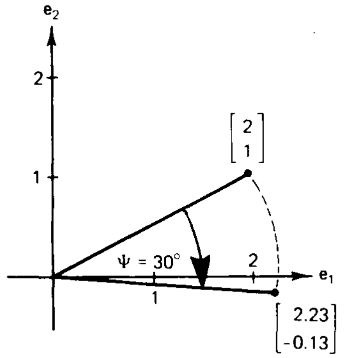

Fig. 3.21 Rotation of point with basis vectors fixed.

this case that the point $[\begin{smallmatrix}2\\1\end{smallmatrix}]$ is rotated *clockwise* 30°.[11] Either interpretation is equally suitable. The one that is selected will depend on the researcher's purpose since it is only *relative* motion that is indicated.

In Chapter 4 we shall explore basis vector and coordinate transformations much more thoroughly. At this point, however, we wish merely to show that there are two compatible ways of looking at things:

1. One can rotate the basis vectors and refer the unchanged point to the new reference axes.
2. One can rotate the point and refer its new location to the original reference axes.

In each case, under rigid rotations we should note that angles and distances are preserved. Finally, we could simplify the angular representation of the preceding rotation—in the special case of two dimensions—by means of a *single* angle of rotation.

If we let $\Psi = \theta_{11}$, we can note the following:

$$\begin{bmatrix} \cos\theta_{11} & \cos\theta_{21} \\ \cos\theta_{12} & \cos\theta_{22} \end{bmatrix} = \begin{bmatrix} \cos\Psi & \sin\Psi \\ -\sin\Psi & \cos\Psi \end{bmatrix}$$

It is instructive to see how the rotation of a set of basis vectors through the single angle Ψ (in the case of two dimensions) leads to the matrix above.

3.5.2 The Trigonometry of Rotation

The trigonometry of rotation can be shown fairly straightforwardly. Panel I of Fig. 3.22 shows a point A with original coordinates a_1 and a_2 in the $\mathbf{e}_1$, $\mathbf{e}_2$ basis. If the

[11] As will be discussed in more detail in Chapter 4, the *clockwise* rotation of $\mathbf{a} = [\begin{smallmatrix}2\\1\end{smallmatrix}]$ can also be represented by the product

$$\begin{bmatrix} 2.23 \\ -0.13 \end{bmatrix} = \begin{bmatrix} 0.867 & 0.5 \\ -0.5 & 0.867 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

where $[\begin{smallmatrix}2\\1\end{smallmatrix}]$ and $[\begin{smallmatrix}2.23\\-0.13\end{smallmatrix}]$ are both expressed in terms of the $\mathbf{e}_i$ basis vectors.

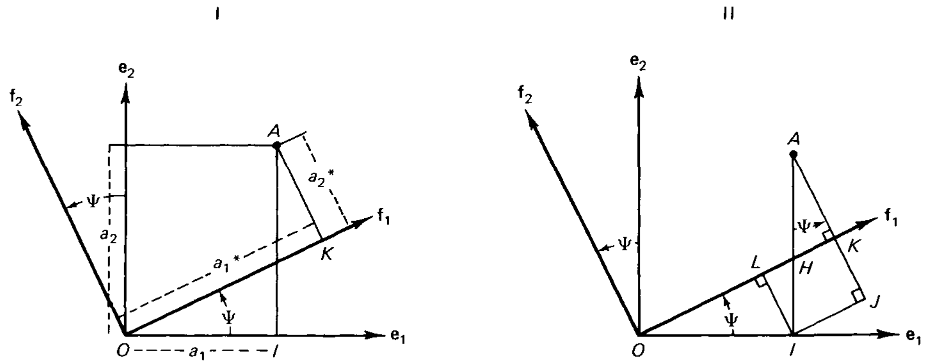

Fig. 3.22 A trigonometric demonstration of basis vector rotation.

original axes are rotated counterclockwise through the angle Ψ, we obtain $a_1{}^*$ and $a_2{}^*$ as the coordinates of A in the $\mathbf{f}_1$, $\mathbf{f}_2$ basis. We now ask: How can $a_1{}^*$ and $a_2{}^*$ be expressed in terms of the old coordinates a_1 and a_2?

The trigonometric argument is simple to describe. Panel II shows the construction of the rectangle $IJKL$. Angle OHI is the complement of the angle Ψ and, in turn, equals angle AHK. Hence, angle HAK is equal to Ψ, the angle of rotation. Given these facts, we can now say

$$a_1{}^* = OK = OH + HK = OL + IJ = OI \cos\Psi + AI \sin\Psi = a_1 \cos\Psi + a_2 \sin\Psi$$

$$a_2{}^* = AK = AJ - JK = AJ - IL = AI \cos\Psi - OI \sin\Psi = a_2 \cos\Psi - a_1 \sin\Psi$$

The coordinates of A in terms of the new basis vectors $\mathbf{f}_1$, $\mathbf{f}_2$ are then given by

$$a_1{}^* = a_1 \cos\Psi + a_2 \sin\Psi; \qquad a_2{}^* = -a_1 \sin\Psi + a_2 \cos\Psi$$

or, in matrix form,

$$\begin{bmatrix} a_1{}^* \\ a_2{}^* \end{bmatrix} = \begin{bmatrix} \cos\Psi & \sin\Psi \\ -\sin\Psi & \cos\Psi \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

as desired.

It should be remembered, however, that expressing a basis vector rotation in terms of a single angle Ψ is restricted to *two* dimensions. On the other hand, the more cumbersome notation involving four angles

$$\theta_{11}; \quad \theta_{12}; \quad \theta_{21}; \quad \theta_{22}$$

is more general, since the concept of direction cosines generalizes to three or more dimensions. Thus, any time that we work with rotations involving three or more dimensions we shall assume that direction cosines are involved throughout.

3.5.3 Higher-Dimensional Rotations

What has been illustrated above for the case of two-dimensional rotations can be extended to three or more dimensions by using the appropriate matrix of direction

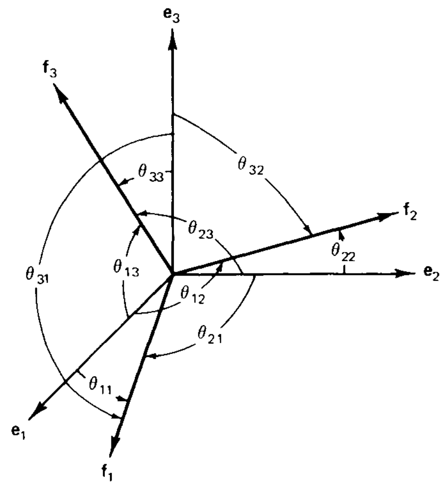

Fig. 3.23 Rotation in three dimensions.

cosines. Figure 3.23 portrays, in general form, a counterclockwise rotation from the $\mathbf{e}_i$ basis to an $\mathbf{f}_i$ basis. The coordinates of some point $\mathbf{a}$ in the original basis can be expressed as $\mathbf{a}^*$ in the new basis by

$$\begin{bmatrix} a_1^* \\ a_2^* \\ a_3^* \end{bmatrix} = \begin{bmatrix} \cos\theta_{11} & \cos\theta_{21} & \cos\theta_{31} \\ \cos\theta_{12} & \cos\theta_{22} & \cos\theta_{32} \\ \cos\theta_{13} & \cos\theta_{23} & \cos\theta_{33} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

Fig. 3.23 shows the angles that are considered in this more complex case. However, no new principles are involved.

Orthogonal matrices play a central role in various multivariate procedures, and their special properties should be noted; these are taken up next.

3.5.4 Properties of an Orthogonal Matrix

Now that we have illustrated what goes on when a point, or points, are subjected to a rotation, let us examine some of the properties of the transformation matrix used in Section 3.5.1. We have

$$\mathbf{A} = \begin{bmatrix} \cos\theta_{11} & \cos\theta_{21} \\ \cos\theta_{12} & \cos\theta_{22} \end{bmatrix} = \begin{bmatrix} 0.867 & 0.5 \\ -0.5 & 0.867 \end{bmatrix}$$

First, let us check on the following:

$$\mathbf{A}'\mathbf{A} = \begin{bmatrix} 0.867 & -0.5 \\ 0.5 & 0.867 \end{bmatrix} \begin{bmatrix} 0.867 & 0.5 \\ -0.5 & 0.867 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{A}\mathbf{A}' = \begin{bmatrix} 0.867 & 0.5 \\ -0.5 & 0.867 \end{bmatrix} \begin{bmatrix} 0.867 & -0.5 \\ 0.5 & 0.867 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

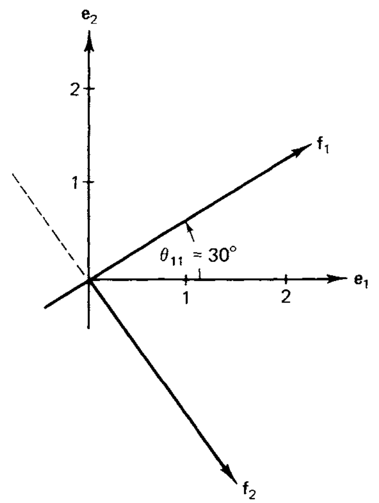

Fig. 3.24 Improper rotation.

As can be seen, within rounding error, we obtain an identity matrix in each case. Hence **A** is observed to be an orthogonal matrix.

However, before concluding that *all* matrices that satisfy the above conditions are rigid rotations, let us consider the following modification of **A**.

$$\mathbf{B} = \begin{bmatrix} \cos\theta_{11} & \cos\theta_{21} \\ -\cos\theta_{12} & -\cos\theta_{22} \end{bmatrix}$$

Note that **B** differs from **A** only in the fact that the second-row entries of **A** have been each multiplied by -1. If we examine the properties of **B**, we see that

$$\mathbf{B}'\mathbf{B} = \mathbf{B}\mathbf{B}' = \mathbf{I}$$

That is, the *same* conditions are met with the **B** matrix as were met with the **A** matrix.[12]

However, what is happening here is something that is a bit different from a rigid rotation. Figure 3.24 illustrates what is going on. In this latter case we have a rigid rotation that leads to a new axis $\mathbf{f}_1$ which is in the same orientation as $\mathbf{f}_1$ in Fig. 3.20 but an axis $\mathbf{f}_2$ which is the *negative* of $\mathbf{f}_2$ in Fig. 3.20.

This new situation represents a case of rotation followed by a *reflection* of the $\mathbf{f}_2$ axis. Alternatively, we could have affixed minus signs to the first row of **A** and, in this case, it would be the $\mathbf{f}_1$ axis that was reflected. However, if all entries of **A** are multiplied by -1, a rigid rotation of $\theta_{11} + 180°$ would result. It is only when an odd number of rows receive minus signs that we have what is known as an "improper" rotation, that is, a rotation followed by reflection.

[12] The reader may verify this numerically or, in more general terms, write out the implied trigonometric relationships.

How do we know before hand whether a proper versus improper rotation is involved? It turns out that this distinction is revealed by examining the determinant of **A**.

1. If the determinant of **A** equals 1, then a proper rotation is involved.
2. If the determinant of **A** equals -1, then an improper rotation is involved.

And, it turns out that any orthogonal matrix will have a determinant that is *either* 1 or -1.

To sum up, if $\mathbf{A}'\mathbf{A} = \mathbf{A}\mathbf{A}' = \mathbf{I}$, we say the matrix is orthogonal. If $|\mathbf{A}| = -1$, it represents a rotation followed by an odd number of reflections, for example, one axis in the 2×2 case, one or three in the 3×3 case, one or three in the 4×4 case, and so on. If $|\mathbf{A}| = 1$, then we are dealing with a proper rotation.[13]

3.6 GEOMETRIC ASPECTS OF CROSS-PRODUCT MATRICES AND DETERMINANTS

In Chapter 2 we defined a determinant as a scalar function of a square matrix. Evaluation of a determinant in terms of both cofactor expansion and the pivotal method was also described and illustrated numerically. At this point attention focuses on the geometric aspects of a determinant and, in particular, its role in portraying *generalized variance* among a set of statistical variables. In the course of describing this relationship, we shall also point out geometric analogies to a number of common statistical concepts.

By way of introduction, we first illustrate some geometric aspects of a determinant at a simple two-dimensional level. We then discuss how determinants can be linked with various statistical measures of interest to multivariate analysis.

3.6.1 The Geometric Interpretation of a Determinant

Certain aspects of the determinant of a matrix can be expressed in geometric format. To illustrate, let us consider the unit square *OIJK*, as shown in Fig. 3.25. The coordinates representing the vertices of the square are

$$O = (0,0); \qquad I = (1,0); \qquad J = (1,1); \qquad K = (0,1)$$

Next, suppose we were to multiply these coordinates by the matrix

$$\mathbf{T} = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}$$

as follows:

$$\mathbf{U} = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \begin{array}{c} \begin{array}{cccc} O & I & J & K \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \end{array}; \qquad \mathbf{U} = \begin{array}{c} \begin{array}{cccc} O & I^* & J^* & K^* \end{array} \\ \begin{bmatrix} 0 & 2 & 3 & 1 \\ 0 & 3 & 7 & 4 \end{bmatrix} \end{array}$$

[13] It should be mentioned that any reflection of two or more axes can itself be represented by a proper rotation followed by just *one* reflection. For example, in the 3×3 case, only one axis need be reflected after an appropriate rotation is made.

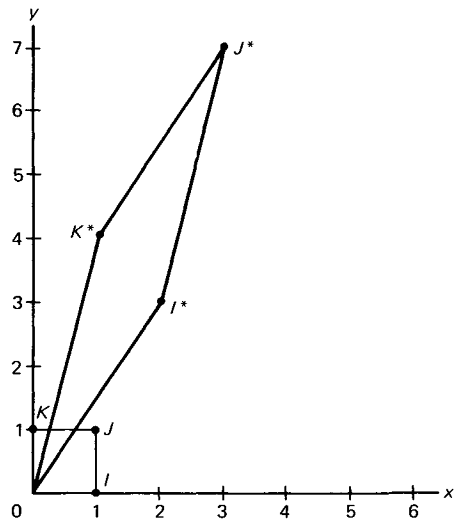

Fig. 3.25 Geometric aspects of a determinant.

These transformed points also appear in Fig. 3.25 as the quadrilateral $OI^*J^*K^*$.

The key aspect of this transformation has to do with the ratio of the area of the quadrilateral to the area of the original unit square. *The ratio of the two areas equals the determinant of the transformation matrix* **T**. That is,

$$|\mathbf{T}| = (2 \times 4) - (3 \times 1) = 5$$

Thus, if one were to measure the area of $OI^*J^*K^*$ and compare it to the area of $OIJK$, one would find that it is exactly five times the latter area. And this would be true for any starting figure that is transformed by **T**.

This concept generalizes to determinants of matrices of order 3 x 3 and higher. In the 3 x 3 case, the determinant measures the ratio of volumes between the original and transformed figures. In the 4 x 4 and higher-order cases, the determinant measures the ratio of hypervolumes between original and transformed figures.

Finally, if the sign of the determinant should be negative, this does not affect the ratio between hypervolumes of original and transformed figures. Rather, the presence of a negative determinant has to do with the orientation of the transformed figure in the space of interest.[14] Hence, it is the *absolute value* of the determinant that indicates the ratio of hypervolumes. Moreover, if that absolute value is less than unity, then the transformed figure's hypervolume is a fraction of that of the original figure.

3.6.2 The Geometry of Statistical Measures

In Chapter 1 we introduced a small and illustrative data bank (Table 1.2), involving only twelve cases and three variables. In Chapter 2 we used this miniature data bank

[14] To illustrate this, the reader should work out the case of $\mathbf{T} = \begin{bmatrix} -2 & -1 \\ 3 & 4 \end{bmatrix}$ with $|\mathbf{T}| = -5$. This entails a reflection of the quadrilateral in Fig. 3.25 across the y axis. However, the ratio of areas is still 5 : 1.

(Table 2.2) to illustrate the application of matrix operations in the computation of various cross-product matrices, such as the SSCP, covariance, and correlation matrices.

We continue to refer to this sample data bank. However, in line with the focus of Chapter 3, the data of the sample problem are now discussed from a geometric viewpoint.

In the course of analyzing multivariate data, it is useful to make various scatter plots for showing relationships among variables. Figures 1.2 and 1.3 are illustrations of the more usual type of plot in which variables are treated as axes, and cases (employees in this example) are treated as points. This more conventional way of portraying data is often called a response surface or point model, since with one criterion variable and two predictors X_1 and X_2, one could visualize the fitting of a response surface to Y, the criterion variable.

Alternatively, however, we could imagine that each of the twelve employees, or cases, represents a dimension, and each of the three variables in the sample problem represents a *vector* embedded in a twelve-dimensional space. (Actually, if the three variables are linearly independent, they will lie in only a three-dimensional subspace of the original twelve dimensions, as is discussed in more detail in later chapters.)

For the moment, let us simplify things even further and consider only two of the variables of the sample problem, namely, Y and X_1. If so, a vector representation of these two variables could be portrayed in only two dimensions, embedded in the full, twelve-dimensional space.

From Table 2.3 we note that the covariance and correlation matrices for only the Y, X_1 pair of variables are

$$\mathbf{C} = \begin{array}{c} \\ Y \\ X_1 \end{array}\!\!\begin{array}{c} \begin{array}{cc} Y & X_1 \end{array} \\ \begin{bmatrix} 29.52 & 19.44 \\ 19.44 & 14.19 \end{bmatrix} \end{array}; \qquad \mathbf{R} = \begin{array}{c} \\ Y \\ X_1 \end{array}\!\!\begin{array}{c} \begin{array}{cc} Y & X_1 \end{array} \\ \begin{bmatrix} 1.0 & 0.95 \\ 0.95 & 1.0 \end{bmatrix} \end{array}$$

As recalled, both the $\mathbf{C}$ and $\mathbf{R}$ matrices are based on mean-corrected variables; as such, the origin of the space will be taken at the centroid of the variables, represented by the $\mathbf{0}$ vector.[15]

As in Chapter 2, we can define the variance of a variable X_1 as

$$\boxed{s_{x_1}^2 = \frac{\Sigma x_{i1}^2}{m} \qquad \text{where } x_{i1} = X_{i1} - \bar{X}_1 \text{ (i.e., each } x \text{ is expressed as a deviation about the mean, and } m \text{ denotes the number of cases}^{16}\text{)}}$$

[15] By centroid is meant a vector whose components are the arithmetic means of Y and X_1, respectively. Then, if we allow the centroid to represent the origin or 0 vector, the individual vectors are position vectors whose termini are expressed as deviations from the mean of Y and X_1, respectively.

[16] One could use $m - 1$ in the denominator if one wished to have an unbiased estimate of the population variance. (Such adjustment does not mean that the sample standard deviation is an unbiased estimate of the population standard deviation, however.) Here, for purpose of simplification, we omit the adjustment and use m in the denominator.

Similarly, the correlation of a pair of variables is defined as

$$r_{yx_1} = \frac{\Sigma y_i x_{i1}}{\sqrt{\Sigma y_i^2}\sqrt{\Sigma x_{i1}^2}}$$ where y_i and x_{i1} are each expressed as deviations about their respective means

When Y and X_1 are each expressed in mean-corrected form, their correlation is related to their scalar product as follows:

$$r_{yx_1} = \frac{\mathbf{y}'\mathbf{x}_1}{\|\mathbf{y}\| \cdot \|\mathbf{x}_1\|} = \cos \theta_{\mathbf{yx}_1}$$

which, we see, is just the cosine of $\theta_{\mathbf{yx}_1}$, their angle of separation.

In terms of the sample problem, the correlation is

$$r_{yx_1} = \cos \theta_{\mathbf{yx}_1} = \frac{233.28}{\sqrt{354.24} \cdot \sqrt{170.28}} = 0.95$$

where the scalar product and squared vector lengths are computed from Y_d and X_{d1} in Table 1.2. The covariance of a pair of variables is defined as

$$\text{cov}_{yx_1} = r_{yx_1} s_y s_{x_1} = \frac{\Sigma y_i x_{i1}}{m}$$ where s_y and s_{x_1} are standard deviations of Y and X_1, respectively

Our current objective is to tie in these statistical notions with concepts from vector algebra and, in particular, to show how the *determinant* of a covariance matrix can be used as a generalized scalar measure of dispersion.

To do this, we first imagine a geometric space in which the axes are the m cases (e.g., the employees). The variable Y can then be thought of as a "test" vector $\mathbf{y}$ in a space of m persons. Similarly, the variable X_1 can be thought of as another vector $\mathbf{x}_1$ in the person space. The components of $\mathbf{y}$ and the components of $\mathbf{x}_1$ each sum to zero since each variable is expressed in terms of deviations from its own mean.

Once this has been done, we can see that the length of the vector $\mathbf{y}$ turns out to be proportional to the standard deviation of the variable Y. That is,

$$\|\mathbf{y}\| = \sqrt{\sum_{i=1}^{m} y_i^2} = \sqrt{m}\, s_y$$

Similarly, for the vector $\mathbf{x}_1$ we have

$$\|\mathbf{x}_1\| = \sqrt{\sum_{i=1}^{m} x_{i1}^2} = \sqrt{m}\, s_{x_1}$$

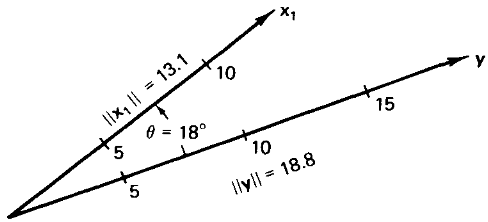

Fig. 3.26 A vector representation of the covariance between Y and X_1 (sample problem).

The correlation r_{yx_1} between the two test vectors, measured as deviations from each variable's mean but based on vector lengths that are each normalized to unity, is equal to the cosine of the angle separating them. Furthermore, the scalar product of these two vectors is proportional to their covariance. That is,

$$\mathbf{y}'\mathbf{x}_1 = \cos\theta_{\mathbf{yx}_1} \|\mathbf{y}\| \cdot \|\mathbf{x}_1\| = m[\mathrm{cov}_{yx_1}] = m[r_{yx_1} s_y s_{x_1}]$$

Insofar as the sample problem is concerned, Fig. 3.26 shows a plot of the vectors $\mathbf{y}$ and $\mathbf{x}_1$. The angle corresponding to a cosine of 0.95 (denoting their correlation) is 18°. Their respective lengths are

$$\|\mathbf{y}\| = \sqrt{12} \cdot \sqrt{29.52} = 18.82; \qquad \|\mathbf{x}_1\| = \sqrt{12} \cdot \sqrt{14.19} = 13.05$$

If one is dealing with standardized scores, then vector lengths would, of course, each be equal to $\sqrt{m}$ since s_y and s_{x_1} would each be equal to unity.

In brief, with mean-corrected variables, all three cross-product matrices—the SSCP, covariance, and correlation matrices—can be portrayed in geometric terms. The key concept involves the scalar product between two variables. In all three cases, we have

$$\cos\theta_{\mathbf{yx}_1} = r_{yx_1}$$

The vector lengths in each case are as follows:

SSCP matrix:	$\sqrt{\sum_{i=1}^{m} y_i^2}$; $\sqrt{\sum_{i=1}^{m} x_{i1}^2}$
Covariance matrix:	$\sqrt{m}\, s_y$; $\sqrt{m}\, s_{x_1}$
Correlation matrix:	$\sqrt{m}$; $\sqrt{m}$

This complementary view of association between variables will serve us in good stead in the interpretation of various aspects of multivariate analysis in later chapters.

3.6.3 A Generalized Variance Measure

Having established the various correspondences shown above, we now wish to illustrate how the determinant of the covariance matrix represents a scalar measure of generalized variance.[17] Given, illustratively, two variables Y and X, the covariance matrix can be written as

$$\mathbf{C} = \begin{bmatrix} s_y{}^2 & r_{yx} s_y s_x \\ r_{yx} s_y s_x & s_x{}^2 \end{bmatrix}$$

where the main diagonal components are variances, and the off-diagonal component is their covariance. If we compute the determinant of this matrix, we get

$$|\mathbf{C}| = s_y{}^2 s_x{}^2 - r_{yx}^2 s_y{}^2 s_x{}^2 = s_y{}^2 s_x{}^2 (1 - r_{yx}^2) = s_y{}^2 s_x{}^2 (1 - \cos^2 \theta_{\mathrm{yx}})$$

and, from basic trigonometry, in which we have the identity $\sin^2 \theta + \cos^2 \theta = 1$, we can write

$$|\mathbf{C}| = s_y{}^2 s_x{}^2 \sin^2 \theta_{\mathrm{yx}} = (s_y s_x \sin \theta_{\mathrm{yx}})^2$$

where θ_{yx} is the angle between the (deviation-score) vectors y and x.

As pointed out earlier, the standard deviation of a variable is $1/\sqrt{m}$ times the length of its corresponding vector. Hence, we have

$$\boxed{s_y s_x \sin \theta_{\mathrm{yx}} = \frac{\|\mathbf{y}\|}{\sqrt{m}} \cdot \frac{\|\mathbf{x}\|}{\sqrt{m}} \sin \theta_{\mathrm{yx}}}$$

and we can conveniently set up the equivalence

$$\boxed{h = \frac{\|\mathbf{y}\|}{\sqrt{m}} \sin \theta}$$

as the height of a parallelogram with base given by

$$\frac{\|\mathbf{x}\|}{\sqrt{m}}$$

as shown in Fig. 3.27. So, if the vector lengths are each scaled by $1/\sqrt{m}$, we see that the area of the resulting (scaled) parallelogram equals $s_y s_x \sin \theta_{\mathrm{yx}}$. The *square* of this area

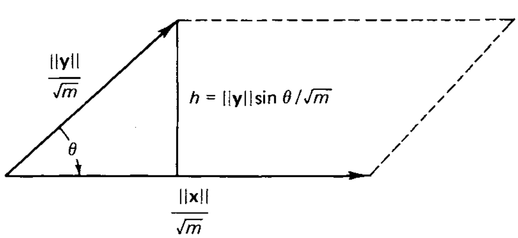

Fig. 3.27 Representing the determinant of a 2 x 2 covariance matrix as the area of a parallelogram.

[17] For example, even in a 2 x 2 covariance matrix we have four dispersionlike entries. Our interest here is on developing a *single* number that represents the four entries in certain multivariate statistical applications.

equals the generalized variance (i.e., determinant of the dispersion matrix). If n variables are involved, the generalized variance equals the square of the volume formed by n such (rescaled) vectors.

Figure 3.27 shows, in general form, the nature of this parallelogram in the two-variable case. We see that each vector appears in scaled (by $1/\sqrt{m}$) form, and the parallelogram is completed as shown.

For a numerical illustration of the correspondence, let us again refer to the sample problem of Table 2.3. We illustrate the equivalence for only the first two variables Y and X_1.

First, from the covariance matrix involving variables Y and X_1,

$$\mathbf{C} = \begin{array}{c} \\ Y \\ X_1 \end{array} \begin{array}{c} \begin{array}{cc} Y & X_1 \end{array} \\ \begin{bmatrix} 29.52 & 19.44 \\ 19.44 & 14.19 \end{bmatrix} \end{array}$$

we obtain

$$\|\mathbf{y}\| = \sqrt{12} \cdot \sqrt{29.52} = 18.82; \qquad \|\mathbf{x}_1\| = \sqrt{12} \cdot \sqrt{14.19} = 13.05$$

$$\cos \theta_{\mathbf{yx}_1} = 0.95; \qquad \sin \theta_{\mathbf{yx}_1} = 0.31$$

Hence the area of the parallelogram formed by $\mathbf{y}$ and $\mathbf{x}_1$ is

$$\frac{\|\mathbf{y}\|}{\sqrt{m}} \cdot \frac{\|\mathbf{x}_1\|}{\sqrt{m}} \cdot \sin \theta_{\mathbf{yx}_1} = \frac{18.82}{\sqrt{12}} \cdot \frac{13.05}{\sqrt{12}} (0.31) = 6.34$$

The square of 6.34 is equal to 40.22. This value, within rounding error, equals the determinant of $\mathbf{C}$, the covariance matrix. Thus we have shown geometrically and numerically how the determinant of $\mathbf{C}$ is equal to the square of the area of the parallelogram in Fig. 3.27.

The concept of generalized variance is quite important in multivariate analysis since it enables us to portray a matrix of variances and covariances *in terms of a single number,* namely, the determinant of the covariance matrix. Just as importantly, we also see that the statistical measures of standard deviation, covariance, and correlation can be portrayed in terms of length and/or angle of test vectors in person and, more generally, object space.

3.7 SUMMARY

The purpose of this chapter has been to describe a number of the vector and matrix operations outlined in Chapter 2 from a geometric standpoint. After setting up a rectangular Cartesian coordinate system and defining the concept of a Euclidean space, we discussed such topics as vector length and angle, vector addition and subtraction, scalar multiplication of a vector, and the scalar product of two vectors from a geometric point of view.

We then described the notion of linear independence. We also illustrated how the Gram-Schmidt process could be employed to find an orthonormal basis starting from any given (arbitrary) basis. Following this we briefly discussed the idea of generalized (nonorthogonal) coordinate systems.

Matrix times vector multiplication was introduced from a geometric viewpoint for the special case of orthogonal (i.e., proper or improper rotation) matrices. The properties of this class of matrices were discussed, and their application was illustrated numerically. We concluded the chapter with a geometric representation of various statistical measures, including the central concept of generalized variance, as applicable to multivariate statistical tests to be considered in later chapters.

REVIEW QUESTIONS

1. Sketch a three-dimensional coordinate system.

a. Plot points with coordinates $(2, 1, 0)$, $(1, -1, -1)$, $(\sqrt{3}, \pi, -2)$. What is the length of each?

b. What is the set of points whose x and y coordinates sum to 1?

c. What is the graph of $z = x^2$?

d. What is the graph of the inequality $x^2 + y^2 + z^2 \leqslant 1$?

2. Let P be the point $(4, 3, -1)$, Q be the point $(1, 0, 2)$, and R be the midpoint of the segment joining P and Q.

a. What are the coordinates of R?

b. Sketch the vectors PR, OR, and PQ, where O denotes the origin.

c. Verify that $PR = PQ/2$ by computing the distance from P to R, R to Q, and P to Q; then show that the first two distances are each half of the last distance.

3. In the context of linear combinations,

a. find a scalar k such that

$$(1, 0, 2) + k(2, 1, 1) = (-1, -1, 1)$$

b. find scalars k_1, k_2, and k_3 such that

$$k_1(5\mathbf{e}_1 + \mathbf{e}_2) + k_2(\mathbf{e}_2 + \mathbf{e}_3) + k_3(\mathbf{e}_3) = 5\mathbf{e}_1 + 3\mathbf{e}_2 + \mathbf{e}_3$$

c. find k_1 and k_2 such that

$$k_1(5\mathbf{e}_1 + \mathbf{e}_2) + k_2(\mathbf{e}_1 - \mathbf{e}_2) = \mathbf{0}$$

4. Let $\mathbf{a}$ and $\mathbf{b}$ be vectors with given lengths and angle θ. Compute their scalar product under the conditions

a. $\|\mathbf{a}\| = 0.5$; $\|\mathbf{b}\| = 4$; $\theta = 45°$

b. $\|\mathbf{a}\| = 4$; $\|\mathbf{b}\| = 1$; $\theta = 90°$

c. $\|\mathbf{a}\| = 1$; $\|\mathbf{b}\| = 1$; $\theta = 120°$

d. What is the possible range of values for the scalar product $\mathbf{a}'\mathbf{b}$ if

$$\|\mathbf{a}\| = 2 \quad \text{and} \quad \|\mathbf{b}\| = 3?$$

5. Let a, b, and c be vectors. Let $\|\mathbf{a}_p\|$ be the component of a along c and let $\|\mathbf{b}_p\|$ be the component of b along c. What is the component of a + b along c? Sketch the relationship in two-dimensional space.

6. Let $\mathbf{a}' = (2, -1, 7)$ and $\mathbf{b}' = (-3, 6, 1)$. Find direction cosines for

a. $\mathbf{a}' + \mathbf{b}'$ b. $\mathbf{a}' - \mathbf{b}'$ c. $5\mathbf{a}' + 10\mathbf{b}'$ d. $\frac{1}{2}(\mathbf{a}' - \mathbf{b}')$

Next, find the cosine of the angle between the two vectors obtained in parts c and d.

7. Apply the Gram–Schmidt orthonormalization procedure to the following sets of vectors:

a. $\mathbf{a}' = (1, 2, 3)$; $\mathbf{b}' = (3, 0, 2)$; $\mathbf{c}' = (3, 1, 1)$

b. $\mathbf{a}' = (2, 1)$; $\mathbf{b}' = (1, 2)$; $\mathbf{c}' = (1, 1)$

c. What do you notice about the vectors obtained in part b?

8. Find coordinates of the vector $\mathbf{a}' = (2, 3)$ relative to the basis vectors $\mathbf{f}_1' = (1, -1)$ and $\mathbf{f}_2' = (3, 5)$.

9. Show that the vectors $\mathbf{a}' = (1, 4, -2)$ and $\mathbf{b}' = (2, 1, 3)$ are orthogonal and find a third vector that is orthogonal to both.

10. Find the equations for the ellipse $4x^2 + y^2 = 4$ and the circle $x^2 + y^2 = 1$ after the xy axes have been rotated counterclockwise through angles of

a. $45°$ b. $60°$ c. $120°$

11. Find x and y so that the vectors $(4, -2, 1, 7)$ and $(2, -3, x, y)$ are linearly dependent.

12. Express the standard basis vectors $\mathbf{e}_1' = (1, 0, 0)$, $\mathbf{e}_2' = (0, 1, 0)$, and $\mathbf{e}_3' = (0, 0, 1)$ as linear combinations of $\mathbf{f}_1' = (1, 2, 4)$, $\mathbf{f}_2' = (-2, 1, 5)$, and $\mathbf{f}_3' = (-1, -1, 2)$.

13. Rotate the vector $\mathbf{a}' = (1, 2)$ counterclockwise through an angle of $45°$ while keeping the basis vectors fixed. Rotate $\mathbf{b}' = (3, 2)$ clockwise through an angle of $60°$.

a. What is the scalar product $\mathbf{a}'\mathbf{b}$ before and after the two rotations?
b. What are the vector lengths of $\mathbf{a}'$ and $\mathbf{b}'$ after the rotations?
c. Show each of the above steps geometrically.

14. Assume that we have the expression

$$\begin{bmatrix} a_1^* \\ a_2^* \end{bmatrix} = \begin{bmatrix} \cos 45° & -\sin 45° \\ \sin 45° & \cos 45° \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

and OP is the line joining the origin $\mathbf{0}$ to the point $P = (2, 3)$. Show in diagram form the position of OP^*, the rotated point.

15. Apply the transformation

$$\begin{bmatrix} a_1^* \\ a_2^* \end{bmatrix} = \begin{bmatrix} 3 & 5 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

to a square with vertices of (1, 1), (3, 1), (3, 3), (1, 3) and show geometrically that the ratio of the area of the new figure to the area of the original is 2 to 1.

16. In the sample problem of Table 2.3, consider the full 3 x 3 covariance matrix.

a. Plot the mean-corrected $\mathbf{y}$, $\mathbf{x}_1$, and $\mathbf{x}_2$ in a three-dimensional space.
b. Plot the standardized form of $\mathbf{y}$, $\mathbf{x}_1$, and $\mathbf{x}_2$ in a three-dimensional space.
c. Show how the correlation between $\mathbf{y}$ and $\mathbf{x}_2$ is related to

$$\cos \theta_{\mathbf{y}\mathbf{x}_2} = \frac{\mathbf{y}'\mathbf{x}_2}{\|\mathbf{y}\| \cdot \|\mathbf{x}_2\|}$$