This book is in the
ADDISON-WESLEY SERIES
IN BEHAVIORAL SCIENCES: QUANTITATIVE METHODS

Consulting Editor

FREDERICK MOSTELLER

Elements of

CONTINUOUS MULTIVARIATE ANALYSIS

A. P. DEMPSTER
Department of Statistics, Harvard University

ADDISON-WESLEY PUBLISHING COMPANY
Reading, Massachusetts · Menlo Park, California · London · Don Mills, Ontario

Research work in connection with this book was supported by the grant GS-341 from the National Science Foundation to Harvard University. The numerical computations were also aided by the National Science Foundation through their grant GP-2723 to the Harvard Computing Center, and by the Faculty-Aide Program administered under the Faculty of Arts and Sciences by the Student Employment Office of Harvard University. The final touches were added to this book while I was a Fellow of the John Simon Guggenheim Memorial Foundation.

Cambridge, Massachusetts Arthur P. Dempster
July, 1968

CONTENTS

PART 1

ORIENTATION

CHAPTER 1

INTRODUCTION

1.1 WHAT THE BOOK IS ABOUT

The purpose of this book is to describe certain methods of analysis of statistical data arising from multivariate samples. A basic aim of such data analysis is to reduce large arrays of numbers to provide meaningful and reasonably complete summaries of whatever information resides in sample aggregates. Another aim is to draw inferences from sample aggregates to larger population aggregates from which the samples are drawn; that is, to understand how certain information about a sample provides uncertain information about a population. Usually statistical analyses are part of a larger endeavor. For example, a scientist may wish to test some scientific theory or hypothesis about a set of populations, or a decision-maker may require the output of statistical data analyses as part of the input for a cost-benefit study. It is a characteristic feature of this book, however, that the technical aspects of these larger endeavors are not explicitly treated.

Part 2 (Chapters 2 through 6) reviews the mathematical theory of vector spaces, emphasizing geometric and computational aspects. Vector space theory is fundamental both for purposes of description and illumination in the two statistical parts of the book. Part 3 (Chapters 7 through 11) presents the basic methods of data reduction while Part 4 (Chapters 12 through 15) is devoted to random sampling theory and its implications. Numerical examples are inserted throughout Part 3 and the same examples are reconsidered in the light of sampling theory in Chapter 14.

Underlying the organization of Parts 3 and 4 is an attitude towards statistical data analysis. While most books on statistical theory start out with sampling theory and attempt to make methods of data analysis follow, the attitude in this book is that the methods of data analysis are carried out largely because of the intrinsic appeal of the sample quantities computed. Such, at least, were the historical origins of the methods described here. Moreover, when viewed as producing descriptive or summary statistics, the methods have value even when assumptions like randomness of samples and normality of populations are quite unwarranted. Mathematical statistics can provide justifications for the methods used, and the main approaches to such justifications are sketched in Chapter 15. Mathematical statistics also suggests modifications to those raw sample statistics

which are especially misleading as estimates of their population analogues. Examples of such modifications are to be found in Chapter 14.

The next three sections of Chapter 1 relate to the third, second, and fourth parts of the book and are designed to convey in a rough way the dimensions of the subject matter. But the reader should recognize at the outset that the methods described are, while fundamental, a very restricted class. For one, variables are generally regarded here as producing numerical scores distributed along a line. It is tempting to apply the methods to classification variables taking the values 0 and 1 only, and some applications of this kind are discussed in Chapters 8 and 9. But more tailored methods and theories exist for analyzing data arising from crossclassifications. In addition, whole fields of data analysis methods which have grown up in particular disciplines are barely touched upon here. For example, there are several approaches to factor analysis and many variants of these methods mostly coming from attempts to structure sets of psychological variables. In this book, only one near relative—namely, principal component analysis—is discussed. Likewise, economists have created highly developed methods of solving sets of uncertain linear equations, whose only relatives in this book are multiple regression analyses. Engineers and psychologists talk a great deal about pattern analysis, which often turns out to be close to the multiple discriminant analysis of Chapter 10. But there are other techniques for clustering individuals which may be described as pattern analysis. Clearly there are much wider horizons to be considered. By studying a narrower class of proven methods in greater depth, this book aims to provide a framework to enable the reader to proceed more easily to these wider horizons.

1.2 INDIVIDUALS, VARIABLES, QUANTITIES, AND RELATED CONCEPTS

A *population* is the collection of all entities belonging to some recognizable and reasonably well-defined type or class, and a *sample* is a subset of a population. The choice of an interesting population is largely determined by the development of each individual scientific discipline. In this book, the entities which make up a given sample or its parent population will be called *individuals*. A sample individual is like a bird in hand, and it is not always clear what flock it came from nor how it was selected. Some sample surveys are designed with great care to be representative of a population under study, but samples of skulls such as those analyzed in Example 10.3 must be collected where they may be found and their representativeness is often a matter of hope. Thus, although the aim of an investigation is generally to study some populations, the data analyst's first emphasis must be on understanding the various angles from which to view a given sample. He may then be satisfied with a naive belief that what he learned from his sample will be roughly or at least possibly true of a population, or he may be willing to pay the price of assuming randomness of samples in order to have formal mathematical tools of statistical inference which provide precise probabilistic ways to think about the population.

The term *variable* will be reserved in this book for a quantitative real-valued attribute or characteristic which is possessed by every individual of a population. A variable is defined by some rule or procedure whereby its numerical *value* on any individual may be established. The rule should be reasonably precise and reasonably objective in the sense that different observers will generally feel that they agree on what is meant by a given variable determined over a specified population. A value of a variable may be directly observed, like the weight of an animal in grams or the response of a human subject to a questionnaire item on a five-point scale. Or a variable may be such that any value must be computed from directly observable variables, like a scale formed by weighting and summing the responses to several questionnaire items. In any particular study, a set of variables is usually dictated by the subject of investigation. The variables may be of the directly observable type or of the computed type. Usually there is no functional relation among them. A statistician may often suggest modifications of a given set of variables, replacing some by transforms or other computed variables. This is generally done for technical rather than fundamental reasons, however, and it is essentially assumed in this book that the set of variables entering an analysis is given.

The starting point of multivariate analysis is therefore a multivariate sample or several multivariate samples with given individuals and given variables. For each individual in each sample one has (ideally at least) all of the values of a specified set of variables.

The concept of variable needs to be distinguished from a related concept which will be called here *quantity*. Actually, there are four different technical terms which need to be distinguished and understood: namely, *variable*; *set of values of a variable*; *quantity*; and *value of a quantity*. The term *variable* connotes mathematically a function or mapping from the individuals of a population to a real line. The *set of values of the variable* is the set of real numbers associated with the individuals of the population by the variable. The term *quantity* as defined here also has the mathematical connotation of a function, but the real world context of a quantity differs sharply from that of a variable.

A simple and basic type of quantity will be defined first. Each pair consisting of an individual and a variable defines a real number which is the value of the variable on that individual . Such a pair defines a *quantity* whose associated real value is the *value of the quantity*. For example, the weight of a specific rat in grams is a quantity having a specific value. In its applied context a variable really does vary, for it takes different values over different individuals, while a quantity actually takes only one value. The set of actual values of the variable becomes the set of *possible* values of the quantity. The quantity itself should be viewed as a function which maps a generally hypothetical space of all possible states of the real world into the real line which includes the set of possible values of the quantity. Following a convention much used in probability theory, a single symbol, such as X, will be used ambiguously for a quantity or for the value of the quantity. One thinks of X as standing for the value of the quantity, and

rigorously speaking a function notation should be used for the quantity itself, but in practice such double notation is awkward and unnecessary.

A function of a variable or of a set of variables defined over a common population is again a variable over that same population. On the other hand, a function of a set of quantities is a meaningful concept provided only that each quantity may be regarded as defined over a common set of possible states of a real world. It is natural therefore to extend the notion of a quantity to include any function (or, in some theories, any measurable function) defined over the set of possible states of the real world. As before, the range or set of possible values of such a quantity is the real line while its actual value is a single number. Such quantities are the daily bread of statistics. Indeed, a function of a set of sample quantities is another quantity which is often called a sample *statistic*.

1.3 THE ROLE OF VECTOR SPACE THEORY

There is a strong emphasis in this book on geometric thinking as a means of visualizing and thereby improving an understanding of methods of data analysis and their associated normal sampling theory. At the same time, when precise mathematical reasoning is required it will be carried out in terms of the theory of finite dimensional vector spaces. This theory may be regarded as a precise mathematical framework underlying the heuristic patterns of geometric thought.

Consider a multivariate sample defined by n individuals and p variables. If this sample is observed, it may be represented by a set of np real numbers, whose typical member $X_j^{(i)}$ is the value of the quantity defined by the ith individual and the jth variable for $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, p$. These values are usually regarded in this book as forming a rectangular array or *data matrix* with n rows and p columns and with $X_j^{(i)}$ in the ith row and jth column. For each i, the row vector

$$[X_1^{(i)}, X_2^{(i)}, \ldots, X_p^{(i)}] \tag{1.3.1}$$

may be viewed as determining the coordinates of a point or a vector in a p-dimensional geometric space. This point carries the sample information about the ith individual, and the n points thus defined for $i = 1, 2, \ldots, n$ carry information equivalent to the data matrix. Any point in the p-dimensional space represents a potential sample individual and the space itself will be called the *individual-space* corresponding to the given p variables. On the other hand, for each j, the column vector

$$\begin{bmatrix} X_j^{(1)} \\ X_j^{(2)} \\ \cdot \\ \cdot \\ \cdot \\ X_j^{(n)} \end{bmatrix} \tag{1.3.2}$$

determines the coordinates of a point in an n-dimensional geometric space. The p points thus defined correspond to variables, and, like the n individuals in individual-space, these points also determine the whole data matrix. The n-dimensional space thus defined is special to the n sample individuals, but when $p \leq n$ it contains a p-dimensional subspace which is abstractly equivalent to a space called *variable-space* to be introduced shortly. Most methods of continuous multivariate analysis have simple descriptions in terms of either individual-space or variable-space, or both.

A restriction on this book is that almost all of the methods discussed are based on sample mean vectors and sample covariance matrices. Excepting some graphic techniques for plotting individuals, none of the information in each multivariate sample is retained except that contained in a special set of linear and quadratic statistics. Such a cutting operation should always be regarded with suspicion, even though mathematical statistics offers some good justifications in the case of random samples from multivariate normal populations. The attitude here is not meant to exclude more catholic approaches, but the range of methods in our restricted class is nontrivial and is unified by its close ties with vector space theory, and so these methods form a natural class for inclusion in a single volume.

The *sample mean vector* is defined to be

$$[\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_p], \tag{1.3.3}$$

where

$$\bar{X}_j = \frac{1}{n} \sum_{i=1}^{n} X_j^{(i)} \tag{1.3.4}$$

for $j = 1, 2, \ldots, p$. The definition of this mean vector should be regarded as illustrating vector space operations; specifically, the mean vector is found by taking the vector sum of the n vectors (1.3.1) and multiplying by the scalar $1/n$. Similarly, the *sample covariance* between the jth and kth variables is defined to be

$$\frac{1}{n-1} \sum_{i=1}^{n} (X_j^{(i)} - \bar{X}_j)(X_k^{(i)} - \bar{X}_k) = \frac{1}{n-1} \left[\sum_{i=1}^{n} X_j^{(i)} X_k^{(i)} - n \bar{X}_j \bar{X}_k \right]. \tag{1.3.5}$$

When $j = k$, this is called the *sample variance* of variable j. The square matrix whose element in row j and column k is the covariance (1.3.5) is called the *sample covariance matrix* of the given sample. The diagonal elements of the covariance matrix are variances and the matrix is symmetric in the sense that the covariances (1.3.5) are unchanged by interchanges of j and k. In Chapter 3 the sample covariance matrix will be identified with the vector space concept of an inner product defined over variable-space and the inner product concept will in turn be visualized geometrically in terms of an associated ellipsoid.

With such a choice of basic statistics it becomes natural to consider not only the p given variables but also all linear combinations of them. This is natural

because the means, variances, and covariances of the wider class of variables are determined by those of the original set. Suppose that $\alpha_1, \alpha_2, \ldots, \alpha_p$ denotes a given set of real numbers. Then

$$V = \alpha_1 V_1 + \alpha_2 V_2 + \cdots + \alpha_p V_p \tag{1.3.6}$$

defines a new variable over the same population. The rule for determining the value of V on a given individual asserts: first determine the values of each of $V_1, V_2, \ldots, V_p$ and then substitute these values for the variables in (1.3.6). For example, if a sample is defined as above by the values $X_j^{(i)}$ for $j = 1, 2, \ldots, p$ and $i = 1, 2, \ldots, n$, then the sample values of V in (1.3.6) are given by

$$X^{(i)} = \alpha_1 X_1^{(i)} + \alpha_2 X_2^{(i)} + \cdots + \alpha_p X_p^{(i)}, \tag{1.3.7}$$

for $i = 1, 2, \ldots, n$. The set of variables V thus defined when $\alpha_1, \alpha_2, \ldots, \alpha_p$ range over all sets of p real numbers will be called *variable-space* and will be formally identified in Chapter 2 with a p-dimensional vector space.

It is now natural to generalize (1.3.4) to define the sample mean of V, namely

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X^{(i)}. \tag{1.3.8}$$

It follows easily that

$$\bar{X} = \alpha_1 \bar{X}_1 + \alpha_2 \bar{X}_2 + \cdots + \alpha_p \bar{X}_p, \tag{1.3.9}$$

which shows, as mentioned above, that the sample means of the given p variables determine the sample means of the continuum of variables in variable-space. In a similar way, if $W = \beta_1 V_1 + \beta_2 V_2 + \cdots + \beta_p V_p$ is another variable with sample values $Y^{(1)}, Y^{(2)}, \ldots, Y^{(n)}$ and sample mean $\bar{Y}$, then (1.3.5) generalizes to

$$\frac{1}{n-1} \sum_{i=1}^{n} (X^{(i)} - \bar{X})(Y^{(i)} - \bar{Y}), \tag{1.3.10}$$

which will be called the *sample covariance* of V and W. The reader may easily show that if (1.3.5) is denoted by $\operatorname{cov}(V_j, V_k)$ and (1.3.10) is denoted by $\operatorname{cov}(V, W)$, then

$$\operatorname{cov}(V, W) = \sum_{j=1}^{p} \sum_{k=1}^{p} \alpha_j \beta_k \operatorname{cov}(V_j, W_k), \tag{1.3.11}$$

which shows how the sample covariance matrix determines the covariance between any pair of variables in variable-space.

Geometric or vector space reasoning avoiding unnecessary reference to particular sets of coordinate axes may often be carried out. Such reasoning, called *coordinate-free*, can simplify both mathematics and understanding by reducing matters to essentials. Concepts such as ellipsoids and linear projections which appear repeatedly may be visualized in a coordinate-free way.

Another important theme of vector space theory is *duality*. In Chapter 6 it will be seen that individual-space and variable-space stand in the formal mathematical relation of a pair of dual spaces. Any concept or chain of reasoning in terms of one space has a dual image in terms of the other, and one of these may sometimes appear simpler than the other. Alongside the mathematical concept of duality there is a corresponding pair of dual attitudes which may be assumed toward any variable. One common attitude to a variable is to regard it as an entity in itself; for example, one commonly feels that the concept of height, perhaps the concept of height of human beings in inches, is a very real thing closely tied to a world view which integrates many such concepts. On the other hand, a variable may be regarded as a device for producing a pattern of points on a line, these points being the values of the variable for a sample or population. More generally, a set of p variables may be viewed initially as entities in themselves, or they may be alternatively and dually regarded as a device for producing a pattern of points in a p-dimensional individual-space. It is part of the flavor of multivariate statistical analysis that it concentrates on the latter attitude by providing some ways of looking at linear and quadratic aspects of patterns in individual-space. This concentration is what makes statistics seem a dry subject, since it may obscure the more fascinating aspects of what variables mean, what motivated them, and how they relate to the world-picture of their associated scientific disciplines. Statistical data analysts must in practice try consciously to build better bridges between formal methods and their scientific contexts; the examples of Chapters 8 through 11 may help the reader in this direction, but they do not pretend to go very far.

1.4 THE ROLE OF SAMPLING THEORY

The final four chapters of this book are concerned with multivariate normal populations and with formal procedures for statistical inference which assume random samples from such normal populations. The mathematics of multivariate normal sampling theory is highly developed, and no attempt is made to review all that is known. Instead, a selection is made which appears to the author to provide those parts of the theory most relevant to data analysis. As always, stress is laid on geometric arguments to bring out the simplicity and elegance of the mathematics.

It is rarely possible to believe with certainty either that populations are normal or that samples are random. It is often possible to make data shed light on the failures of such assumptions or, less often, on the effect of these failures upon specific inferences. Such checking is not analyzed in a formal way in this book, but a few specific examples are given. Inevitably, one must make careful but informal judgments about how far any particular inference may be trusted.

Even accepting the assumptions of normality and randomness, the user of formal methods of statistical inference must decide between two sharply

contrasting approaches. To see the nature of the cleavage, suppose that a statistic has been decided upon to define an interesting characteristic of a sample or a population. The question which immediately arises is of the magnitude of the *sampling error* of the statistic, defined to be the difference between the sample value and the population value. Such a magnitude may be judged in one of two ways which will be called *postdictive* and *predictive*, and which determine a basic difference among schools of statistical inference. According to the first approach the population is regarded as fixed but unknown, and the randomness of the sample determines a probability distribution for the sample statistic and thence for the sampling error of the statistic. These *sampling distributions* provide the key to inference from sample to population. According to the second approach, the population is regarded as unknown, but it is the sample in hand which is viewed as fixed. A probability distribution is sought which may be used to provide judgments which are appropriate *after the sample is given*, and which concern the uncertainties of knowledge of the population. Such distributions are often called *posterior distributions*. Now the sampling distribution of a sampling error is not at all the same concept as the posterior distribution of a sampling error. The latter uses probability in the standard forward-looking mode (called here *predictive*) which regards probability as a measure of the uncertainty of some uncertain outcome or value. The sampling distribution on the other hand may be regarded as having a predictive interpretation only *before* the sample is observed, for it specifically reflects the randomness induced by the random sampling hypothesis. The question therefore is: what meaning can be attached to a sampling distribution *after* a sample is known? The answer is: for each hypothesized population which determines a sampling distribution for a particular sampling error, one knows which point along the sampling distribution was realized in the observed sample. To make this answer useful for statistical inference, one must add a principle asserting that an observed point in the far tail renders a sampling distribution implausible, and thus also renders implausible the hypothesized population which produced the sampling distribution. This form of reasoning, called here *postdictive* inference, includes significance testing and confidence region methods. Note that the word "confidence" may be misleading here because the argument proceeds in a double negative way, failing to reject certain hypotheses instead of positively supporting them.

Predictive inferences are undoubtedly more desirable in principle than are postdictive inferences. But the two main schools of predictive inference, namely Bayesian inference and fiducial inference, are not yet available in a form applicable in a routine way to multivariate data. Most statisticians feel unable to specify the prior distributions required by the Bayesian approach, especially in the presence of the large number of parameters required to specify a normal population of even modest dimension. At the same time, the fiducial methods are somewhat ambiguous. Thus, while predictive inference is analyzed further

in Chapter 15, the only applications of inference to data are those of Chapter 14 which rely on postdictive reasoning.

One of the standard terms of probability theory clashes with the basic terminology of this book. The offending terms are *random variable* and *variable*. Since the latter concept as defined in Section 1.2 is so basic to multivariate statistical analysis, the probabilist's term random variable will be altered here to *random quantity*. This use of the term quantity coincides with that introduced in Section 1.2. The adjective random signifies only that the quantity comes equipped with a probability measure or a family of probability measures. The value of a random quantity will be called a *random value* and denoted by a symbol like X. As already mentioned, X will be used in the familiar ambiguous way to denote either the quantity or its value.

1.5 RELATED WORKS: A VERY BRIEF SKETCH

Along with Harold Hotelling, R. A. Fisher, John Wishart, and M. S. Bartlett, one of the pioneers of multivariate statistical analysis among mathematical statisticians was S. S. Wilks. Wilks (1962) remains an excellent source for some of the basic work on multivariate normal sampling theory. A fuller exposition of a similar approach is given by Anderson (1958). Other distinguished theoretical statisticians have books which treat the subject, including Cramér (1946), Kendall (1957), Kendall and Stuart (1961, 1966), Rao (1952, 1965), and Roy (1957). The review paper by Bartlett (1947) is of historic importance and interest. Readers preferring mathematics of lower power may be helped by the books of Cooley and Lohnes (1962), Morrison (1967), or Seal (1964). Multiple regression analysis and related analysis of variance ideas are fundamental to multivariate analysis; a range of books on these topics is covered by the set Acton (1959), Deutsch (1965), Plackett (1960), Scheffé (1959), Draper and Smith (1966), and Williams (1959). Glimpses at a wide range of current work are to be found in Krishnaiah (1966). An extensive bibliography of statistical literature on multivariate statistical analysis will shortly be published by Anderson, Das Gupta, and Styan (1969). Other literature on multivariate analysis, especially original sources in research papers, will be cited throughout the text.

While the treatment of vector space theory in Chapters 2 to 6 is nearly self-contained and adequate for the purposes of this volume, many readers will benefit from deeper coverage of different approaches. Excellent books for the abstract mathematical content are Birkhoff and MacLane (1965) and Halmos (1958). For matrix theory, see MacDuffee (1943), Gantmacher (1959), and Rao (1965). Good sources for geometric orientation are Coxeter (1961) and Sommerville (1958). Important computational theory is given by Householder (1964), Varga (1962), and Wilkinson (1963, 1965). Books at a more down-to-earth level include Aitken (1958), Dwyer (1951), Faddeeva (1959), Horst (1963), Ralston and Wilf (1960), and Searle (1966).

PART 2

MATHEMATICAL FOUNDATIONS

CHAPTER 2

BASIC THEORY OF VECTOR SPACES

2.1 ABSTRACT VECTOR SPACES DEFINED

Most areas of applied mathematics lean heavily on the theory of vector spaces over the field of real numbers. The abstract mathematical content common to these applications is sketched below and briefly illustrated by the relation of the theory to multivariate analysis. More complete sources of the theory are listed in Section 1.5, while more detailed applications abound in the later parts of this book.

A *p-dimensional vector space over the field of real numbers* is a mathematical construct which specializes in the case $p = 1$ to the set of all real numbers with their familiar operations of addition and multiplication. In general, the elements of the vector space are called *vectors* and should be regarded for the moment as abstract entities, although in applications they will be identified with or at least labelled for concepts with a real world basis. Vectors will often be denoted by capital letters such as $U, V, W, \ldots$ The theory requires that any two vectors may be added to give a third vector, i.e., given any pair of vectors U and V there exists a unique vector W which will be written in the familiar way as

$$W = U + V. \tag{2.1.1}$$

This operation may be called *vector addition* to distinguish it from the specialized concept of *addition of real numbers*, but the term *addition* will be used for either and the meaning will be clear from the context. The theory requires also that any real number and any vector may be multiplied to give another vector, i.e., given any real number x and any vector U, the product of x and U is uniquely defined to be a vector Z which will be denoted by

$$Z = xU. \tag{2.1.2}$$

Real numbers are sometimes called *scalars* in vector space theory. The operation of passing from U to Z may be called *scaling by the factor* x and the multiplication operation itself is often called *scalar multiplication.*

The operations of addition and multiplication are required to obey the following rules:

a) Vectors form an additive abelian group, i.e.,

There exists a *zero* vector $\emptyset$ such that

$$\emptyset + V = V + \emptyset = V$$

for all V.

There exists a *negative vector* $-V$ for every V such that

$$V + (-V) = \emptyset.$$

Vector addition is commutative, i.e., (2.1.3)

$$V + U = U + V$$

for all U and V.

Vector addition is associative, i.e.,

$$(V + U) + W = V + (U + W)$$

for all U, V, W.

b) Multiplication of vectors by real numbers obeys:

$$x(U + V) = xU + xV, \qquad (x + y)V = xV + yV,$$
$$(xy)V = x(yV), \qquad \text{and} \qquad 1V = V \tag{2.1.4}$$

for all real numbers x and y and all vectors U and V.

These basic axioms together with some simple deductions from them make up a set of working rules which are used almost automatically by anyone familiar with vector spaces. Some of the simple deductions, not derived here, are as follows:

$$\begin{aligned} 0V &= \emptyset, && \text{for all } V, \\ x\emptyset &= \emptyset, && \text{for all } x, \text{ and} \\ (-1)V &= -V, && \text{for all } V. \end{aligned} \tag{2.1.5}$$

The notation $U - V$ abbreviates $U + (-V)$.

As indicated in Chapter 1, two applications of vector spaces will be used repeatedly in this book:

Example 2.1.1. Suppose that $V_1, V_2, \ldots, V_p$ represent p different observable or measurable variables. Here V_1 might be a symbol representing the concept of the height of a human being in inches, V_2 might similarly represent arm length in inches, etc. Then, as explained in Section 1.3, the notation

$$\sum_{t=1}^{p} \alpha_t V_t$$

is to be understood to represent a variable for which a value may be computed in the obvious way once $V_1, V_2, \ldots, V_p$ are assigned observed values. It is clear that the elements $\sum_1^p \alpha_t V_t$ can be regarded as the elements of a vector space satisfying the axioms. That is, sums are defined by

$$\left(\sum_{t=1}^{p} \alpha_t V_t\right) + \left(\sum_{t=1}^{p} \beta_t V_t\right) = \sum_{t=1}^{p} (\alpha_t + \beta_t) V_t, \tag{2.1.6}$$

and multiplication by a real number is defined by

$$x\left(\sum_{t=1}^{p} \alpha_t V_t\right) = \sum_{t=1}^{p} (x\alpha_t) V_t. \tag{2.1.7}$$

Also, $\sum_1^p \alpha_t V_t$ is defined to be the vector $\emptyset$ when $\alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$ and $\sum_1^p (-\beta_t) V_t$ is defined to be the negative vector of $\sum_1^p \beta_t V_t$. The reader may further check through the axioms. The space whose typical member is $\sum_1^p \alpha_t V_t$ will be referred to as *variable-space* throughout the book.

Variable-space suffers from the disadvantage that it does not allow changes of location of a variable although it does allow changes of scale. For example, if V represents a temperature variable in degrees Centigrade, then the same variable measured in degrees Fahrenheit, namely "$32 + 1.8V$," is not in the space of variables under consideration. This difficulty may be circumvented by a device of adding an artificial variable V_0 whose observed value is always unity, and replacing the space of variables $\sum_1^p \alpha_t V_t$ by the *augmented variable-space* $\sum_0^p \alpha_t V_t$. Then the variable loosely referred to above as "$32 + 1.8V$" can be written precisely as $32V_0 + 1.8V$ in the augmented space of variables. For the most part, it will not be necessary to bring in the artificial variable V_0 and attention may be focused on the original variable-space defined above.

Example 2.1.2. Suppose that the p-tuple of real numbers

$$[x_1, x_2, \ldots, x_p]$$

represents a set of measurements on the p variables $V_1, V_2, \ldots, V_p$. Again the reader may check that all such p-tuples form a vector space under the operations:

i) the sum of $[x_1, x_2, \ldots, x_p]$ and $[y_1, y_2, \ldots, y_p]$ is defined to be

$$[x_1 + y_1, x_2 + y_2, \ldots, x_p + y_p],$$

ii) the product of x and $[x_1, x_2, \ldots, x_p]$ is defined to be $[xx_1, xx_2, \ldots, xx_p]$.

Each point of this space may be thought of as a set of possible measurements on an individual, and the space will be referred to as *individual-space* throughout the book. In Chapter 6 the concept of *duality* which relates variable-space and individual-space will be discussed.

Again, it would be possible to consider an *augmented individual-space* consisting of $(p + 1)$-tuples

$$[x_0, x_1, x_2, \ldots, x_p],$$

where $x_1, x_2, \ldots, x_p$ are as above and x_0 denotes an observed value of the artificial variable V_0. However, in the application of this mathematical structure x_0 is always unity, and so it is clear that nothing essential can be gained by adding x_0 to the p-tuple $[x_1, x_2, \ldots, x_p]$.

2.2 SUBSPACES, HYPERPLANES, LINEAR DEPENDENCE, BASIS VECTORS, AND DIMENSION

Any subset of the set of vectors in a vector space is a *subspace* if it forms a vector space under the operations defined for the whole space. Clearly any linear combination of the vectors of a subspace also lies in the same subspace, where a *linear combination* of $U_1, U_2, \ldots, U_r$ means any vector of the form $\sum_1^r a_i U_i$ where $a_1, a_2, \ldots, a_r$ are any real numbers. Conversely, the set of all linear combinations of any subset $U_1, U_2, \ldots, U_r$ constitutes a subspace. This is called the *subspace spanned* by $U_1, U_2, \ldots, U_r$. The simplest examples of subspaces are the whole space and the subspace consisting only of $\emptyset$. The simplest nontrivial example of a subspace is the set of vectors αU for any given vector U as α ranges over all real numbers.

Given two subspaces $\mathscr{U}$ and $\mathscr{V}$ one may consider their intersection $\mathscr{U} \cap \mathscr{V}$ and union $\mathscr{U} \cup \mathscr{V}$. It may be checked that $\mathscr{U} \cap \mathscr{V}$ is itself a subspace but that $\mathscr{U} \cup \mathscr{V}$ cannot be a subspace unless $\mathscr{U}$ and $\mathscr{V}$ are identical. The elements of different subspaces $\mathscr{U}$ and $\mathscr{V}$ span a subspace which may be called the *direct sum* of $\mathscr{U}$ and $\mathscr{V}$ and written $\mathscr{U} \oplus \mathscr{V}$. Subspaces $\mathscr{U}$ and $\mathscr{V}$ are called *complementary* subspaces when their intersection is as small as possible, i.e., consists only of $\emptyset$, while their direct sum is the whole space.

Consider any subspace $\mathscr{U}$ of a vector space $\mathscr{E}$. For a given vector V the subset of $\mathscr{E}$ consisting of all vectors $V + U$ where U is in $\mathscr{U}$ will be called a *coset* of $\mathscr{U}$ and will be written $V + \mathscr{U}$. If W belongs to the coset $V + \mathscr{U}$, then the cosets $V + \mathscr{U}$ and $W + \mathscr{U}$ are identical, i.e., there are as many different ways of expressing the coset as there are vectors in the coset. Any vector V in $\mathscr{E}$ belongs to the coset $V + \mathscr{U}$ and to no other coset of $\mathscr{U}$; in other words, the cosets of $\mathscr{U}$ determine a partition of $\mathscr{E}$ into mutually exclusive subsets. Suppose that $\mathscr{V}$ is a subspace of $\mathscr{E}$ complementary to $\mathscr{U}$. Any such $\mathscr{V}$ defines a one-to-one correspondence between the vectors of $\mathscr{V}$ and the cosets of $\mathscr{U}$, defined by making V in $\mathscr{V}$ correspond to $V + \mathscr{U}$ in the class of cosets. If desired, this correspondence could be used to define vector space operations over the class of cosets, and it could be checked that these operations do not depend on the particular choice of $\mathscr{V}$ complementary to the given $\mathscr{U}$, i.e., there is a natural way to regard the cosets of $\mathscr{U}$ as forming a vector space. The group-theoretic term coset will be replaced later in this chapter by the more familiar geometric term *hyperplane*.

Vectors $U_1, U_2, \ldots, U_r$ are said to be *linearly independent* if the relation

$$c_1 U_1 + c_2 U_2 + \cdots + c_r U_r = \emptyset \tag{2.2.1}$$

implies that $c_1 = c_2 = \cdots = c_r = 0$. Otherwise they are *linearly dependent*, and at least one of them can be expressed as a linear combination of the rest. A *basis* of a vector space is a linearly independent set of vectors which spans the whole space. Any vector has a unique expression as a linear combination of the vectors of a given basis. A basis always exists but is never unique; in fact, any linearly independent set of vectors can be incorporated into a basis. Each basis of a given vector space has the same number of elements. If this number is finite it is called the *dimension* of the space, but otherwise the space is called *infinite-dimensional. There will be no discussion of infinite-dimensional vector spaces in this book.* At the other extreme, a vector space consisting only of $\emptyset$ is defined to have dimension 0. If $\mathscr{U}$ and $\mathscr{V}$ are subspaces, it can be shown that

$$d(\mathscr{U}) + d(\mathscr{V}) = d(\mathscr{U} \cap \mathscr{V}) + d(\mathscr{U} \oplus \mathscr{V}), \tag{2.2.2}$$

where $d(\ldots)$ denotes the dimension of a subspace. Thus, if $\mathscr{U}$ and $\mathscr{V}$ are complementary subspaces of a p-dimensional vector space,

$$d(\mathscr{U}) + d(\mathscr{V}) = p. \tag{2.2.3}$$

It may be shown that a set of r vectors is a linearly dependent set if and only if the subspace spanned by the set has dimension less than r.

It is clear that variable-space and individual-space in Examples 2.1.1 and 2.1.2 are both p-dimensional vector spaces. In Example 2.1.1, the vectors $V_1, V_2, \ldots, V_p$ constitute a basis, and $\alpha_1, \alpha_2, \ldots, \alpha_p$ represent the *coordinates* of $V = \sum_1^p \alpha_t V_t$ relative to this basis. In Example 2.1.2 the p-tuples

$$\begin{gathered}(1, 0, \ldots, 0)\\(0, 1, \ldots, 0)\\\vdots\\(0, 0, \ldots, 1)\end{gathered} \tag{2.2.4}$$

form a basis. On the other hand, it is clear that the augmented variable-space and the corresponding augmented individual-space are both $(p + 1)$-dimensional vector spaces.

2.3 LINEAR TRANSFORMATIONS

A mapping

$$V \to v \tag{2.3.1}$$

from the elements V of a vector space $\mathscr{E}$ to the elements v of a vector space $\mathscr{F}$ will be called a *linear transformation of $\mathscr{E}$ into $\mathscr{F}$* provided that

$$\alpha_1 V_1 + \alpha_2 V_2 \to \alpha_1 v_1 + \alpha_2 v_2, \tag{2.3.2}$$

for any real numbers α_1 and α_2, and any V_1 and V_2 in $\mathscr{E}$, where $V_1 \to v_1$ and $V_2 \to v_2$ under the mapping.

If $\emptyset$ denotes the zero vector in $\mathscr{E}$ and $\mathfrak{o}$ denotes the zero vector in $\mathscr{F}$, then clearly $\emptyset$ maps into $\mathfrak{o}$ under *any* linear transformation from $\mathscr{E}$ to $\mathscr{F}$. More generally, if $\mathscr{U}$ denotes the set of vectors in $\mathscr{E}$ which are carried into $\mathfrak{o}$ in $\mathscr{F}$, then it is easily checked that $\mathscr{U}$ must be a subspace of $\mathscr{E}$. $\mathscr{U}$ will, of course, be the subspace of dimension zero if only $\emptyset$ in $\mathscr{E}$ maps into $\mathfrak{o}$ in $\mathscr{F}$. A basic property of $\mathscr{U}$ is that if V in $\mathscr{E}$ maps into v in $\mathscr{F}$, then the whole coset $V + \mathscr{U}$ maps into v and conversely the coset $V + \mathscr{U}$ consists of all the vectors of $\mathscr{E}$ which map into v in $\mathscr{F}$. The range of the transformation, i.e., the set of vectors of $\mathscr{F}$ reached by mapping all V in $\mathscr{E}$ into $\mathscr{F}$, need not include all of $\mathscr{F}$ but must be a subspace $\mathscr{V}$ of $\mathscr{F}$. Thus, in general, a linear transformation is a many-one transformation from $\mathscr{E}$ to a subspace of $\mathscr{F}$. It becomes a one-one relationship between $\mathscr{E}$ and $\mathscr{F}$ if and only if $\mathscr{U}$ consists of $\emptyset$ and $\mathscr{V}$ consists of $\mathscr{F}$, and in this case there is a unique inverse mapping which is also a linear transformation. Such a one-one relationship may be called an *isomorphism.*

Suppose that $\mathscr{E}$, $\mathscr{U}$, $\mathscr{F}$, and $\mathscr{V}$ defined above have dimensions p, r, q, and s, respectively, where, since $\mathscr{U}$ is contained in $\mathscr{E}$ and $\mathscr{V}$ is contained in $\mathscr{F}$,

$$0 \leq r \leq p \qquad \text{and} \qquad 0 \leq s \leq q. \tag{2.3.3}$$

In particular, s is often called the *rank of the linear transformation* and obeys $s \leq p$ as well as $s \leq q$. This may be seen by noting that the transforms $v_1, v_2, \ldots,$ v_p in $\mathscr{F}$ of a set of basis vectors $V_1, V_2, \ldots, V_p$ in $\mathscr{E}$ must span $\mathscr{V}$, and no subspace of dimension s can be spanned by fewer than s vectors. Thus, the rank s of a linear transformation from a p-dimensional vector space $\mathscr{E}$ to a q-dimensional vector space $\mathscr{F}$ obeys

$$0 \leq s \leq \min(p, q). \tag{2.3.4}$$

An additional important relation is that

$$r + s = p. \tag{2.3.5}$$

To see this, consider a basis $V_1, V_2, \ldots, V_p$ of $\mathscr{E}$ such that $V_1, V_2, \ldots, V_r$ span $\mathscr{U}$. Then it may be checked that the transforms $v_1, v_2, \ldots, v_p$ of $V_1, V_2, \ldots, V_p$ satisfy the requirements that $v_1 = v_2 = \cdots = v_r = \mathfrak{o}$ and that $v_{r+1}, v_{r+2}, \ldots, v_p$ are linearly independent, the latter set therefore forming a basis of $\mathscr{V}$ with $s = p - r$ basis vectors.

Any given linear transformation may be completely specified by only the transforms $v_1, v_2, \ldots, v_p$ of an arbitrary basis $V_1, V_2, \ldots, V_p$ of $\mathscr{E}$. Conversely, an arbitrary linear transformation may be constructed by picking an arbitrary set of elements $v_1, v_2, \ldots, v_p$ in $\mathscr{F}$ and asserting that they are the transforms of a basis $V_1, V_2, \ldots, V_p$ of $\mathscr{E}$. By such constructions, it may be checked that any r and s satisfying the inequalities (2.3.3) and (2.3.4) are possible. In particular, if $p = q$ and one chooses a basis $v_1, v_2, \ldots, v_p$ of $\mathscr{F}$ to be the transforms of a basis $V_1, V_2, \ldots, V_p$ of $\mathscr{E}$, then $r = 0$ and $s = p$. Since this clearly defines an isomorphism between $\mathscr{E}$ and $\mathscr{F}$, it follows that an isomorphism may be defined between any pair of p-dimensional vector spaces, actually in infinitely many ways. Also, such an isomorphism has rank p, which is the largest rank allowed by (2.3.4).

It is sometimes useful to define vector space operations on linear transformations from $\mathscr{E}$ to $\mathscr{F}$. Thus, if A and B denote linear transformations from $\mathscr{E}$ to $\mathscr{F}$ carrying V in $\mathscr{E}$ into v_A and v_B in $\mathscr{F}$, then $\alpha\mathsf{A} + \beta\mathsf{B}$ may be defined as the mapping which carries V into $\alpha v_A + \beta v_B$ in $\mathscr{F}$.

A product notation for linear transformations is also useful. If C denotes a linear transformation from $\mathscr{E}$ to $\mathscr{F}$ and D denotes a linear transformation from $\mathscr{F}$ to $\mathscr{G}$, then CD will denote the transformation from $\mathscr{E}$ to $\mathscr{G}$ defined by first applying C to $\mathscr{E}$ and then applying D to $\mathscr{F}$. It is easily checked that CD is a linear transformation.

An important special case of the foregoing theory concerns linear transformations of $\mathscr{E}$ into itself, i.e., where the roles of $\mathscr{E}$ and $\mathscr{F}$ are both played by $\mathscr{E}$. Some additional concepts arise here. For example, one is led to consider the identity transformation I which carries each V in $\mathscr{E}$ into itself.

Clearly I satisfies

$$\mathsf{AI} = \mathsf{IA} = \mathsf{I} \tag{2.3.6}$$

for any A from $\mathscr{E}$ into itself.

A linear transformation A from $\mathscr{E}$ into itself which satisfies

$$\mathsf{AA} = \mathsf{A} \tag{2.3.7}$$

is often called an *idempotent transformation*, or, in geometrically more natural terms, a *projection*. Suppose that $\mathscr{U}$ and $\mathscr{V}$ denote, as before, the subset which maps into $\emptyset$ and the range space of A. It is easily shown that if A is idempotent, then $\mathscr{U}$ and $\mathscr{V}$ must be complementary and thence that every vector in $\mathscr{V}$ is left unchanged by A. Any vector W in $\mathscr{E}$ has a unique expression as

$$W = V + U \tag{2.3.8}$$

with V in $\mathscr{V}$ and U in $\mathscr{U}$, where A carries $V \to V$ and $U \to \emptyset$ and thence

$$W \to V. \tag{2.3.9}$$

Thus A is a projection in the sense that it removes the $\mathscr{U}$-component of W while leaving the $\mathscr{V}$-component unchanged. To construct an idempotent transformation of rank s, one need only specify a pair of complementary subspaces $\mathscr{U}$ and $\mathscr{V}$ of dimensions $p - s$ and s and use (2.3.9). Note that the identity I is a special projection—in fact the only projection with maximum rank p.

2.4 AFFINE GEOMETRY: VECTOR SPACE AS GEOMETRIC SPACE

It is often more natural to shed the formal abstract language of Sections 2.1, 2.2, and 2.3 and discuss vector spaces in geometric language. According to this

language, vectors are called *points*, and, in particular, the zero vector Ø is called the *origin*. A subspace $\mathscr{U}$ of dimension r is called a *hyperplane of dimension r through the origin*, and the cosets $V + \mathscr{U}$ are the r-dimensional *hyperplanes parallel* to $\mathscr{U}$. A *line* and a *plane* are hyperplanes of dimensions one and two respectively.

A line is uniquely determined by any two points on it. Specifically, the line through U_1 and U_2 may be written $U_1 + \mathscr{U}$ where $\mathscr{U}$ is the subspace spanned by $U_1 - U_2$. Alternatively, the points of this line may be written $\alpha_1 U_1 + \alpha_2 U_2$ where $\alpha_1 + \alpha_2 = 1$. The subset of these points satisfying $\alpha_1 \geq 0$ and $\alpha_2 \geq 0$ define the *line segment* U_1U_2 joining U_1 and U_2.

More generally, the smallest hyperplane containing a given set of points $U_1, U_2, \ldots, U_r$ is $U_1 + \mathscr{U}$ where $\mathscr{U}$ is the subspace spanned by the differences among the U_i. This hyperplane may also be characterized as the set of points

$$\sum_{i=1}^{r} \alpha_i U_i, \tag{2.4.1}$$

where

$$\sum_{i=1}^{r} \alpha_i = 1. \tag{2.4.2}$$

This hyperplane has dimension $r - 1$ if and only if no U_i lies in the smallest hyperplane containing the remaining U_j. In this case the set of points (2.4.1) obeying (2.4.2) and

$$\alpha_i \geq 0 \qquad \text{for} \qquad i = 1, 2, \ldots, r \tag{2.4.3}$$

defines the $(r - 1)$-dimensional *simplex* with *vertices* $U_1, U_2, \ldots, U_r$. A one-dimensional simplex is a line segment, a two-dimensional simplex a triangle, and a three-dimensional simplex a tetrahedron.

A different generalization of the one-dimensional concept of a line segment is that of a *parallelotope*, which specializes to a parallelogram in two dimensions and to a parallelepiped in three dimensions. For any linearly independent points $W_1, W_2, \ldots, W_r$ of $\mathscr{E}$ and any point U of $\mathscr{E}$, the set of points

$$D + \sum_{i=1}^{r} \beta_i W_i \tag{2.4.4}$$

such that

$$0 \leq \beta_i \leq 1 \qquad \text{for} \qquad i = 1, 2, \ldots, r \tag{2.4.5}$$

define the points of *an r-dimensional parallelotope*. The 2^r points such that each β_i is zero or unity are called the vertices of the parallelotope. The set of points such that $\beta_i = 0$ for a single specified i define an $(r - 1)$-dimensional parallelotope, as do those points such that $\beta_i = 1$. These subparallelotopes may be called the ith pair of parallel faces. A point on at least one face of a parallelotope lies on the surface of the parallelotope while the remaining points are interior points.

The point (2.4.4) such that

$$\beta_i = \tfrac{1}{2} \qquad \text{for} \qquad i = 1, 2, \ldots, r \tag{2.4.6}$$

defines the *center* of the parallelotope.

A basis $V_1, V_2, \ldots, V_p$ determines a *coordinate system* in the familiar sense of analytic geometry. The line segments $ØV_1, ØV_2, \ldots, ØV_p$ are the *coordinate axes* and the point $V = \sum_1^p \alpha_i V_i$ is said to have *coordinates* $\alpha_1, \alpha_2, \ldots, \alpha_p$.

The kind of geometry which can be carried on using only the operations allowed by vector space theory is called *affine geometry*. Affine geometry is characterized by the presence of concepts of parallelity of lines, planes, and hyperplanes, but the absence of concepts of length, angle, and perpendicularity. Note, however, that the ratio of the lengths of parallel line segments is an affine concept. Suppose that UV and U_1V_1 are parallel line segments and that $W = V - U$ and $W_1 = V_1 - U_1$. Then W and W_1 must lie in the same line through the origin, i.e., in the same subspace of dimension one, and hence $W_1 = \alpha W$ for some α. This α is the ratio of the length of the line segment U_1V_1 to that of UV.

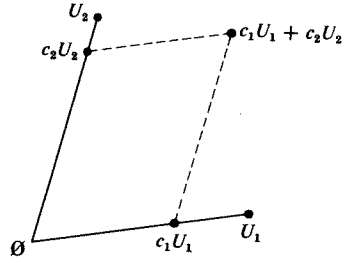

Fig. 2.4.1. Construction of the vector $c_1U_1 + c_2U_2$.

The reader may object that although geometric language has been introduced, nothing has been done to tie this language to his intuitive understanding of two- and three-dimensional geometric space. Consider therefore an "ordinary" geometric plane, or space of two dimensions, which may be thought of as a flat piece of paper extending to infinity in all directions. Suppose that a two dimensional vector space with basis vectors U_1 and U_2 is to be represented geometrically in this plane. The first step in the representation is to choose a point in the plane to be called the origin. Then any two line segments from the origin may be taken to represent the axes determined by $ØU_1$ and $ØU_2$. With these choices the correspondence between the vectors of the space and the points of the plane is defined. In order to construct the point corresponding to the vector $c_1U_1 + c_2U_2$, as illustrated in Fig. 2.4.1, one needs only the notions of (i) multiplying a line segment by a real number and (ii) moving a line segment parallel to itself. Note that (i) and (ii) are the geometric versions of the fundamental operations (2.1.1) and (2.1.2) on abstract vectors. This discussion could be repeated in three dimensions and embellished by explicitly tying together the

intuitive notion of a family of parallel lines or planes with the abstract notion of a subspace and its cosets. Note that, in thinking this way about vector spaces, care should be taken not to introduce concepts alien to affine geometry. It is impossible to draw geometric diagrams without some intuitive notions of length, angle, and perpendicularity, but if the space is regarded as affine these notions must not be used.

In four or more dimensions it is not possible to make concrete geometric representations of a vector space as illustrated in Fig. 2.4.1. However, it is possible and useful to develop a geometric intuition in p-dimensional space through trying to reason by analogy with two and three dimensional spaces, resorting when necessary to the rigorous abstract language of vector space theory.

In considering an affine plane or, in general, an affine p-dimensional space, it is not required to think of the origin as geometrically different from any other point. However, in order to regard affine geometric space as a vector space, a specific origin must be chosen, and as stated in Exercise 2.1.4 different vector operations result when different points are chosen as the origin. Thus, the concepts of an affine geometric space and a vector space are not identical, but the concepts of an affine geometric space with a specified origin and a vector space may be identified. Strictly speaking, *geometric affine properties* are those which continue to hold when defined in terms of a vector space arising from different choices of the origin $\mathcal{O}$.

When thinking in geometric terms with various possible choices for origin, it is convenient to widen the notion of a linear transformation of $\mathscr{E}$ into itself to include transformations which shift the origin as well as other points. Thus, if $\mathcal{O}$ is any point in an affine space $\mathscr{E}$ and the set $V_1, V_2, \ldots, V_p$ together with the origin $\mathcal{O}$ forms a basis of the resulting vector space, then a mapping of $\mathscr{E}$ into itself which carries

$$\sum_{i=1}^{p} \alpha_i V_i = \mathcal{O} + \sum_{i=1}^{p} \alpha_i (V_i - \mathcal{O}) \qquad \text{into} \qquad \mathcal{O}^* + \sum_{i=1}^{p} \alpha_i (V_i^* - \mathcal{O}^*)$$

for some set of $p + 1$ points $\mathcal{O}^*, V_1^*, \ldots, V_p^*$ will be called a *wide sense linear transformation*. Note that the wide sense linear transformation carries an origin $\mathcal{O}$ into another point $\mathcal{O}^*$ whereas the narrow sense definition always requires $\mathcal{O}$ to be invariant.

The simplest type of wide sense linear transformation is a *translation* $V \to V + \mathcal{O}^*$. In geometric terms, this translation simply shifts each point along a line parallel to $\mathcal{O}\mathcal{O}^*$ by an amount equal to the length of $\mathcal{O}\mathcal{O}^*$. It may be easily checked that the general transformation of the previous paragraph may be regarded as the net result of first carrying out the narrow sense linear transformation carrying $\mathcal{O}$ into itself and V_i into $V_i^* - \mathcal{O}^*$, for $i = 1, 2, \ldots, p$, and then carrying out the translation $V \to V + \mathcal{O}^*$. In this sense, translations

are the only new operations needed to define wide sense linear transformations from narrow sense linear transformations.

As suggested in Section 2.3, idempotent transformations have a natural geometric interpretation as projections. To define such a projection one needs a pair of hyperplanes $\mathscr{U}$ and $\mathscr{V}$ of dimensions r and s where $r + s = p$ and such that $\mathscr{U}$ and $\mathscr{V}$ intersect only in the origin. The operation of removing the component of W along $\mathscr{U}$ may also be regarded as the operation of replacing W by the intersection of the hyperplane $W + \mathscr{U}$ with the hyperplane $\mathscr{V}$. Thus the projection is often described as projection *along the family of hyperplanes parallel to $\mathscr{U}$ into the hyperplane $\mathscr{V}$*. The definition extends easily into a wide sense linear transformation with $\mathscr{V}$ replaced by a hyperplane $V + \mathscr{V}$ parallel to $\mathscr{V}$ but not through the origin, i.e., the transformation carries W into the intersection of $W + \mathscr{U}$ and $V + \mathscr{V}$.

2.5 MATRICES AND COORDINATE SYSTEMS: ANALYTIC VECTOR SPACE THEORY

Matrix algebra can be introduced without reference to vector space theory and has applications not related to vector space theory. Here, however, matrices are introduced primarily as mathematical and computational tools for describing and relating vectors. The elements of matrix theory are so widely known and available that they are only lightly treated here. In general, an $r \times s$ *matrix* is an array of r rows and s columns of real numbers and will be denoted by a bold-face letter. Typical notation would be

$$\mathbf{M} = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1s} \\ m_{21} & m_{22} & \cdots & m_{2s} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ m_{r1} & m_{r2} & \cdots & m_{rs} \end{bmatrix}, \tag{2.5.1}$$

where m_{ij} is often called the *element in position* (i, j). Occasionally the elements may be taken to be abstract vectors. If $r = 1$, a matrix may be called a *row vector* and if $s = 1$, a *column vector*. The matrix formed by interchanging the rows and columns of $\mathbf{M}$ is called the *transpose* of $\mathbf{M}$ and is denoted by $\mathbf{M}'$. When a matrix $\mathbf{M}$ is *square*, i.e., when $r = s$, one may ask whether $\mathbf{M}$ is *symmetric*, i.e., whether $\mathbf{M} = \mathbf{M}'$. The (i, j) elements of a matrix with $i = j$ will be called the *diagonal* or *main diagonal* elements and those with $i \neq j$ will be called the *off-diagonal* elements. The sum of the diagonal elements will be called the *trace* of the matrix and denoted by tr $\mathbf{M}$. The concepts of diagonal and trace are especially relevant to square symmetric matrices.

Matrices whose elements are all zero or all unity will be denoted by $\mathbf{0}$ and $\mathbf{1}$, respectively. The dimensions of any specific $\mathbf{0}$ or $\mathbf{1}$ must be defined in each

case or be clear from the context. Similarly, a square matrix of specific dimension whose diagonal elements are all unity and whose off-diagonal elements are all zero will be denoted by **I** and called an *identity matrix*.

Given an $r_1 \times s_1$ matrix $\mathbf{M}_1$ and an $r_2 \times s_2$ matrix $\mathbf{M}_2$ with $s_1 = r_2$, the product $\mathbf{M} = \mathbf{M}_1\mathbf{M}_2$ is defined to be the $r_1 \times s_2$ matrix whose (i, j) element is given by

$$m_{ij} = \sum_{t=1}^{s_1} m_{1it} m_{2tj}, \tag{2.5.2}$$

where m_{ij}, m_{1ij} and m_{2ij} denote the (i, j) elements of $\mathbf{M}, \mathbf{M}_1$, and $\mathbf{M}_2$, respectively. Two matrices of the same dimension may be added to give a third matrix of the same dimension, $\mathbf{N} = \mathbf{N}_1 + \mathbf{N}_2$ being defined by $n_{ij} = n_{1ij} + n_{2ij}$, where n_{ij}, n_{1ij}, and n_{2ij} denote the (i, j) elements of $\mathbf{N}, \mathbf{N}_1$, and $\mathbf{N}_2$, respectively. The product $x\mathbf{M}$ of a real number x and an $r \times s$ matrix $\mathbf{M}$ is defined to be an $r \times s$ matrix whose (i, j) element is xm_{ij} where m_{ij} is the (i, j) element of $\mathbf{M}$.

The first use of matrices is to represent the coordinates of individual vectors in a vector space relative to a given coordinate system (i.e., basis). Suppose that a vector space $\mathscr{E}$ has a basis $V_1, V_2, \ldots, V_p$. Then any vector $\sum_1^p \alpha_t V_t$ can be written $\boldsymbol{\alpha}\mathbf{V}$, where

$$\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_p] \quad \text{and} \quad \mathbf{V} = \begin{bmatrix} V_1 \\ V_2 \\ \cdot \\ \cdot \\ \cdot \\ V_p \end{bmatrix}. \tag{2.5.3}$$

It should be stressed that in this notation the elements of $\boldsymbol{\alpha}$ are real numbers but *the elements of* $\mathbf{V}$ *are abstract vectors.* It is easily checked that the operations defined for vectors and those defined for matrices agree in the sense that if $U_1 = \boldsymbol{\mu}_1\mathbf{V}$ and $U_2 = \boldsymbol{\mu}_2\mathbf{V}$, then $c_1U_1 + c_2U_2 = (c_1\boldsymbol{\mu}_1 + c_2\boldsymbol{\mu}_2)\mathbf{V}$. Thus one can dispense with the basis column vector $\mathbf{V}$ and simply identify the vectors of $\mathscr{E}$ with the p-tuples $\boldsymbol{\alpha}$ in the obvious way. This gives a matrix representation of the space $\mathscr{E}$.

This matrix representation is different for different choices of a basis in $\mathscr{E}$. A second use of matrices is to relate the representations corresponding to different bases. Suppose that the column vectors $\mathbf{U}$ and $\mathbf{V}$ represent different bases $U_1, U_2, \ldots, U_p$ and $V_1, V_2, \ldots, V_p$. Then any V_i has a unique expression as a linear combination of $U_1, U_2, \ldots, U_p$, i.e.,

$$V_i = \sum_{j=1}^{p} a_{ij} U_j, \; i = 1, 2, \ldots, p, \tag{2.5.4}$$

which may be written $\mathbf{V} = \mathbf{AU}$ in matrix notation where $\mathbf{A}$ has the (i, j) element

a_{ij}. Any V in $\mathscr{E}$ may be written $\sum_1^p \alpha_t V_t$ or $\sum_1^p \beta_t U_t$ where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are the coordinates relative to $\mathbf{V}$ and $\mathbf{U}$. Thus $\boldsymbol{\alpha}\mathbf{V} = \boldsymbol{\beta}\mathbf{U}$, or $\boldsymbol{\alpha}\mathbf{AU} = \boldsymbol{\beta}\mathbf{U}$, so that

$$\boldsymbol{\beta} = \boldsymbol{\alpha}\mathbf{A}. \tag{2.5.5}$$

Of course, one could equally well have expressed $\mathbf{U}$ in terms of $\mathbf{V}$ as $\mathbf{U} = \mathbf{BV}$ and have expressed

$$\boldsymbol{\alpha} = \boldsymbol{\beta}\mathbf{B}. \tag{2.5.6}$$

Note that $\mathbf{U} = \mathbf{BV} = \mathbf{BAU}$ and $\mathbf{V} = \mathbf{AU} = \mathbf{ABV}$ so that

$$\mathbf{BA} = \mathbf{AB} = \mathbf{I} \tag{2.5.7}$$

where $\mathbf{I}$ is an identity matrix. In (2.5.7) the result relating matrix multiplication and successive coordinate changes has been used implicitly. The general result, which may be easily checked, states that if the bases $\mathbf{U}$ and $\mathbf{V}$ are related by $\mathbf{U} = \mathbf{BV}$ and the bases $\mathbf{V}$ and $\mathbf{W}$ are related by $\mathbf{V} = \mathbf{CW}$, then the bases $\mathbf{U}$ and $\mathbf{W}$ are related by $\mathbf{U} = \mathbf{DW}$, where

$$\mathbf{D} = \mathbf{BC}. \tag{2.5.8}$$

A more general use of matrices is to represent linear transformations. Suppose that A is a linear transformation from $\mathscr{E}$ to $\mathscr{F}$ where $\mathscr{E}$ has the $p \times 1$ basis $\mathbf{V}$ and $\mathscr{F}$ has the $q \times 1$ basis $\mathbf{U}$. Then $\mathbf{V} \to \mathbf{AU}$ for some $p \times q$ matrix $\mathbf{A}$. Conversely, any $p \times q$ matrix $\mathbf{A}$ determines a linear transformation $\boldsymbol{\alpha}\mathbf{V} \to \boldsymbol{\alpha}\mathbf{AU}$ from $\mathscr{E}$ to $\mathscr{F}$ in terms of the bases $\mathbf{V}$ and $\mathbf{U}$. A further linear transformation B from $\mathscr{F}$ to $\mathscr{G}$ defines and is defined by an associated matrix $\mathbf{B}$ relative to bases $\mathbf{U}$ and $\mathbf{W}$ of $\mathscr{F}$ and $\mathscr{G}$, respectively. The product transformation AB from $\mathscr{E}$ to $\mathscr{G}$ is easily seen to be represented by the matrix product $\mathbf{AB}$ relative to the bases $\mathbf{V}$ and $\mathbf{W}$, for $\mathbf{V} \to \mathbf{AU}$ under A and $\mathbf{U} \to \mathbf{BW}$ under B so that $\mathbf{V} \to \mathbf{AU} \to \mathbf{ABW}$ under A followed by B.

The *rank* of any matrix $\mathbf{A}$ may be defined to be the rank of a linear transformation A which it represents. A $p \times p$ matrix $\mathbf{A}$ of rank p defines an isomorphism A between $\mathscr{E}$ and $\mathscr{F}$ which has a unique inverse B, in turn represented by the *inverse matrix* $\mathbf{B}$. Since $\mathsf{AB} = \mathsf{BA} = \mathsf{I}$, the inverse matrix $\mathbf{B}$ of $\mathbf{A}$ satisfies (2.5.7). A $p \times p$ matrix of rank p is called *nonsingular* or *of full rank*, whereas it is *singular* or *of less than full rank* if it has rank less than p. Every nonsingular matrix $\mathbf{A}$ has a unique inverse $\mathbf{B}$ satisfying (2.5.7), but no singular matrix has such an inverse.

Any $p \times p$ matrix $\mathbf{A}$ defines a transformation $\mathbf{V} \to \mathbf{AV}$ of $\mathscr{E}$ into itself. $\mathbf{A}$ is said to be *idempotent* if the transformation $\mathbf{V} \to \mathbf{AV}$ is idempotent, i.e., if

$$\mathbf{AA} = \mathbf{A}. \tag{2.5.9}$$

Mathematical theory can often be derived either by using purely vectorial or geometric reasoning, or by using analytic manipulations with matrices. The former reasoning often makes use of no basis and may then be called *coordinate-free*. The latter always must express vectors in terms of particular coordinate

systems and may be called *coordinate-dependent*. Simplifications in analytic arguments often result from the special choice of a basis or coordinate system.

2.6 PSEUDOINVERSION

From time to time it is useful to be equipped technically to handle matrices of less than full rank. These arise in various ways which can usually be related to linear transformations of less than full rank, and in geometric terms the situation is often easy to understand. In particular, consider the following generalization of the concept of inverse of a linear transformation.

Suppose that A denotes a linear transformation from a p-dimensional vector space $\mathscr{E}$ to a q-dimensional vector space $\mathscr{F}$. Suppose that $\mathscr{U}$, $\mathscr{V}$, r, and s are defined as in Section 2.3 where s is the rank of A. Suppose that $\mathscr{U}^*$ is any $(p-r)$-dimensional subspace of $\mathscr{E}$ complementary to $\mathscr{U}$ and $\mathscr{V}^*$ is any $(q-s)$-dimensional subspace of $\mathscr{F}$ complementary to $\mathscr{V}$. Define $\mathsf{I}(\mathscr{U}, \mathscr{U}^*)$ to be the projection of $\mathscr{E}$ into $\mathscr{U}^*$ along hyperplanes parallel to $\mathscr{U}$ and define $\mathsf{I}(\mathscr{V}^*, \mathscr{V})$ to be the projection of $\mathscr{F}$ into $\mathscr{V}$ along hyperplanes parallel to $\mathscr{V}^*$. Note that these projections both have rank s. Now the points U^* of $\mathscr{U}^*$ are in one-one correspondence with hyperplanes $U^* + \mathscr{U}$, and these hyperplanes are in one-one correspondence under A with the points of $\mathscr{V}$. In other words, A defines a one-one linear transformation from $\mathscr{U}^*$ to $\mathscr{V}$; call this A_1 and call its inverse B_1. Finally, define the linear transformation

$$\mathsf{B} = \mathsf{I}(\mathscr{V}^*, \mathscr{V})\mathsf{B}_1 \tag{2.6.1}$$

from $\mathscr{F}$ to $\mathscr{E}$ to be a *pseudoinverse* of A.

It is clear that given A there is a different pseudoinverse B for each different pair of chosen $\mathscr{U}^*$ and $\mathscr{V}^*$. Three basic and obvious properties of such a pseudoinverse are

i) B has the same rank s as A,

ii) $$\mathsf{AB} = \mathsf{I}(\mathscr{U}, \mathscr{U}^*), \tag{2.6.2}$$

iii) $$\mathsf{BA} = \mathsf{I}(\mathscr{V}^*, \mathscr{V}). \tag{2.6.3}$$

These properties say that B comes as close to being an inverse as is possible when A has rank s in the sense that a projection of rank s is as close to the identity as can be managed with a transformation of rank s. The particular definition is motivated by the following theorem.

Theorem 2.6.1. *Suppose that* A *is a given rank* s *linear transformation from* $\mathscr{E}$ *to* $\mathscr{F}$. *Suppose that* B *has rank* s *and* AB *and* BA *are both projections of rank* s. *Then* B *is a pseudoinverse of* A *as defined above.*

To prove this, define $\mathscr{V}^*$ to be the subspace of $\mathscr{F}$ which maps under B into $\emptyset$ in $\mathscr{E}$, and define $\mathscr{U}^*$ to be the range space of B in $\mathscr{E}$. The objective is to prove that B has the structure (2.6.1) with the $\mathscr{V}^*$ and $\mathscr{U}^*$ as defined. Since B has

rank s, $\mathscr{V}^*$ and $\mathscr{U}$ have dimensions $q-s$ and $s=p-r$. Also, AB carries any hyperplane parallel to $\mathscr{U}$ into a single point in the range space $\mathscr{U}^*$, and since AB is a projection of rank s it must be $\mathsf{I}(\mathscr{U}, \mathscr{U}^*)$. This proves, incidentally, that $\mathscr{U}$ and $\mathscr{U}^*$ must be complementary. Defining A_1 and B_1 as above, it follows that B is defined by B_1 over $\mathscr{V}$ and it remains only to show that $\mathscr{V}$ and $\mathscr{V}^*$ are complementary. If, to the contrary, the intersection of $\mathscr{V}$ and $\mathscr{V}^*$ were to contain more than the origin, then a larger subspace of $\mathscr{V}$ than simply the origin would map under B into the origin in $\mathscr{E}$. From this, $\mathscr{V}$ could not be in one-one correspondence with $\mathscr{U}^*$, yielding a contradiction which completes the proof.

In matrix terms, the analogue of A is a $p \times q$ matrix $\mathbf{A}$ of rank s. A *pseudoinverse of* $\mathbf{A}$ is naturally defined to be any $q \times p$ matrix $\mathbf{B}$ of rank s such that $\mathbf{AB}$ and $\mathbf{BA}$ are both idempotent of rank s. Clearly, $\mathbf{B}$ is a pseudoinverse of $\mathbf{A}$ if and only if B is a pseudoinverse of A where A and B are defined relative to bases $\mathbf{U}$ and $\mathbf{W}$ of $\mathscr{E}$ and $\mathscr{F}$ by

$$\mathbf{U} \to \mathbf{AW} \quad \text{and} \quad \mathbf{W} \to \mathbf{BU}. \tag{2.6.4}$$

The structure of $\mathbf{A}$ and its pseudoinverse $\mathbf{B}$ may be explored as follows. Consider a basis $\mathbf{U}^*$ of $\mathscr{E}$ such that the last $p-s$ elements $\mathbf{U}_2^*$ map into the origin under $\mathbf{U} \to \mathbf{AW}$, i.e., $\mathbf{U}_2^*$ spans the subspace $\mathscr{U}$ defined for the transformation A. Then the first s elements $\mathbf{U}_1^*$ span a complementary subspace $\mathscr{U}^*$. Similarly, suppose that $\mathbf{W}^*$ is a basis of $\mathscr{F}$ whose first s elements $\mathbf{W}_1^*$ span the range space $\mathscr{V}$ of A and whose last $q-s$ elements $\mathbf{W}_2^*$ span a complementary subspace $\mathscr{V}^*$. Since $\mathbf{U}^*$ and $\mathbf{W}^*$ determine $\mathscr{U}^*$ and $\mathscr{V}^*$, they determine a particular pseudoinverse B and its corresponding $\mathbf{B}$. Now A may be described as carrying

$$\mathbf{U}_1^* \to \mathbf{C}_{11}\mathbf{W}_1^* \quad \text{and} \quad \mathbf{U}_2^* \to \mathbf{o}, \tag{2.6.5}$$

where $\mathbf{C}_{11}$ is an $s \times s$ nonsingular matrix and $\mathbf{o}$ is the origin in $\mathscr{F}$. The corresponding pseudoinverse B carries

$$\mathbf{W}_1^* \to \mathbf{D}_{11}\mathbf{U}_1^* \quad \text{and} \quad \mathbf{W}_2^* \to \mathbf{O}, \tag{2.6.6}$$

where $\mathbf{D}_{11} = \mathbf{C}_{11}^{-1}$ and $\mathbf{O}$ is the origin in $\mathscr{E}$.

Now (2.6.5) and (2.6.6) may be written

$$\mathbf{U}^* \to \mathbf{C}_{11}^+\mathbf{W}^* \tag{2.6.7}$$

and

$$\mathbf{W}^* \to \mathbf{D}_{11}^+\mathbf{U}^* \tag{2.6.8}$$

where $\mathbf{C}_{11}^+$ is a $p \times q$ matrix with $\mathbf{C}_{11}$ in the first s rows and columns and zero elsewhere, and $\mathbf{D}_{11}^+$ is a $q \times p$ matrix with $\mathbf{D}_{11}$ in the first s rows and columns and zero elsewhere. Also there exist nonsingular matrices $\mathbf{G}$ and $\mathbf{H}$ such that

$$\mathbf{U}^* = \mathbf{GU} \quad \text{and} \quad \mathbf{W}^* = \mathbf{HW}. \tag{2.6.9}$$

From (2.6.7), (2.6.8), and (2.6.9), it follows that

$$\mathbf{A} = \mathbf{G}^{-1}\mathbf{C}_{11}^{+}\mathbf{H} \tag{2.6.10}$$

and

$$\mathbf{B} = \mathbf{H}^{-1}\mathbf{D}_{11}^{+}\mathbf{G}. \tag{2.6.11}$$

These formulas provide the following prescription for finding a pseudoinverse **B** for a given matrix **A**. *First find* **G** *and* $\mathbf{H}^{-1}$ *such that* $\mathbf{GAH}^{-1}$ *has the form of* $\mathbf{C}_{11}^{+}$. *Then find* $\mathbf{D}_{11} = \mathbf{C}_{11}^{-1}$ *and use* (2.6.11) *to find* **B**.

Finally, consider the special case where $s = \min(p, q)$. First suppose that $p < q$ while A has rank p. Then $\mathsf{AB} = \mathsf{I}$ since AB is a rank p projection of $\mathscr{E}$ into itself. Similarly, any pseudoinverse **B** of a $p \times q$ rank p matrix **A** satisfies $\mathbf{AB} = \mathbf{I}$. The situation with $q < p$ is similar. When $p = q$, B is simply the inverse of A, i.e., $\mathbf{B} = \mathbf{A}^{-1}$.

2.7 EXERCISES

2.1.1 Derive the relationships (2.1.5) from the given axioms.

2.1.2 From the axioms (2.1.3) and (2.1.4) prove the theorem that $xV = \emptyset$ for $x \neq 0$ implies $V = \emptyset$. Conversely, show that this theorem together with (2.1.3) and (2.1.4), excluding $1V = V$ from (2.1.4), is sufficient to imply that $1V = V$ for all V.

2.1.3 Show that variable-space and individual-space as defined in Examples 2.1.1 and 2.1.2 do in fact satisfy the axioms of vector space theory. What are the zero elements of these two spaces? How does one define the negative of a vector in these two spaces?

2.1.4 Suppose that $[a_1, a_2, \ldots, a_p]$ is a particular point in the vector space of p-tuples of real numbers $[x_1, x_2, \ldots, x_p]$. Show that this same space of p-tuples can be regarded as a vector space in a different way, where

i) the zero vector is taken to be $[a_1, a_2, \ldots, a_p]$,

ii) the sum of $[x_1, x_2, \ldots, x_p]$ and $[y_1, y_2, \ldots, y_p]$ is defined to be $[x_1 + y_1 - a_1, x_2 + y_2 - a_2, \ldots, x_p + y_p - a_p]$, and

iii) the product $\alpha[x_1, x_2, \ldots, x_p]$ is defined to be $[\alpha x_1 - (\alpha - 1)a_1, \alpha x_2 - (\alpha - 1)a_2, \ldots, \alpha x_p - (\alpha - 1)a_p]$.

What is the negative of $(x_1, x_2, \ldots, x_p)$ in this vector space?

2.2.1 Prove the unproved statements in the first two paragraphs of Section 2.2.

2.2.2 Show that any set of vectors including the zero vector is a linearly dependent set.

2.2.3 Show that vectors $V_1, V_2, \ldots, V_s$, all different from $\emptyset$, are linearly dependent if and only if some one of them may be expressed as a linear combination of the rest. Under what circumstances can $V_1, V_2, \ldots, V_s$ be linearly dependent while V_1 is not expressible as a linear combination of the rest?

2.2.4 Show that the cosets $V + \mathscr{V}$ and $W + \mathscr{V}$, where $\mathscr{V}$ is a subspace of a vector space $\mathscr{E}$, are identical if and only if $V - W$ belongs to $\mathscr{V}$.

2.3.1 Show that the single condition (2.3.2) is equivalent to the pair of conditions that

$$\alpha V \to \alpha v$$

for all real α and any V in $\mathscr{E}$, where $V \to v$, and that

$$V + U \to v + u$$

for any V and U in $\mathscr{E}$, where $V \to v$ and $U \to u$.

2.3.2 Show that the mapping $V \to V^* = -V$, for all V in a p-dimensional vector space $\mathscr{E}$, defines a linear transformation of rank p of $\mathscr{E}$ into itself.

2.3.3 Show that any linear transformation of a p-dimensional vector space $\mathscr{E}$ into itself has rank p if and only if it is an isomorphism.

2.3.4 Give an example of a specific linear transformation of rank 2 of the vector space spanned by the basis V_1, V_2, V_3, V_4 into itself. Specify $\mathscr{U}$ and $\mathscr{V}$ in the example.

2.3.5 Show that the set of points left invariant by a linear transformation of $\mathscr{E}$ into itself is a subspace $\mathscr{W}$ of $\mathscr{E}$. Show further that the dimension of $\mathscr{W}$ is less than or equal to the rank of the linear transformation, with equality only if the transformation is a projection.

2.3.6 Suppose that A is a linear transformation from $\mathscr{E}$ to $\mathscr{F}$, and that B is a linear transformation from $\mathscr{F}$ to $\mathscr{G}$. Show that the rank of AB is no greater than the smaller of the ranks of A and B.

2.3.7 Show that the vector space of linear transformations from $\mathscr{E}$ of dimension p to $\mathscr{F}$ of dimension q has dimension pq.

2.3.8 Suppose that $\mathscr{U}$ and $\mathscr{V}$ are complementary subspaces of the vector space $\mathscr{E}$. Suppose that I represents the identity mapping of $\mathscr{E}$ into itself, $\mathsf{I}(\mathscr{U}, \mathscr{V})$ represents the projection along $\mathscr{U}$ into $\mathscr{V}$, and $\mathsf{I}(\mathscr{V}, \mathscr{U})$ represents the projection along $\mathscr{V}$ into $\mathscr{U}$. Show that

$$\mathsf{I} - \mathsf{I}(\mathscr{U}, \mathscr{V}) = \mathsf{I}(\mathscr{V}, \mathscr{U}),$$

i.e., that the difference of I and a projection is another complementary projection. What is the transformation $\mathsf{I}(\mathscr{U}, \mathscr{V})\mathsf{I}(\mathscr{V}, \mathscr{U})$?

2.4.1 Interpret geometrically the following statements:

i) $\mathscr{U}$ and $\mathscr{V}$ are complementary subspaces of a three-dimensional vector space $\mathscr{E}$,

ii) $\mathscr{U}$ and $\mathscr{V}$ are subspaces of a three-dimensional vector space $\mathscr{E}$ such that $\mathscr{U}$ and $V + \mathscr{V}$ have no common point,

iii) U_1, U_2, U_3, and U_4 span a subspace of dimension 2 of a p-dimensional space $\mathscr{E}$, and

iv) U_1, U_2, U_3, and U_4 are linearly dependent in a p-dimensional space $\mathscr{E}$.

2.4.2 Give a definition of the statement that an r-dimensional hyperplane is parallel to an s-dimensional hyperplane in p-dimensional affine space.

2.4.3 Strictly speaking, a geometric property should not be called affine unless it is proved to hold regardless of the choice of an origin. Examples of such properties are (a) the parallelity or nonparallelity of two hyperplanes and (b) the ratios of lengths of parallel line segment ratios. Can you prove that (a) and (b) are affine properties?

2.4.4 Suppose that $U_1, U_2, \ldots, U_r$ are any r points in p-dimensional affine space. How would you define the center of gravity of these r points? Is this an affine concept?

2.4.5 Define in a natural way the set of 2^{r-s} parallel s-dimensional faces of a given r-dimensional parallelotope, including the vertices when $s = 0$ and the ordinary faces when $s = r - 1$. What is the condition that a face of dimension t be contained in a face of dimension s for $t < s$? Show that a face of dimension s is the intersection of all faces of dimension $r - 1$ which contain it.

2.4.6 Suppose that $\mathscr{U}$ and $\mathscr{V}$ are hyperplanes through the origin of dimensions r and s, respectively, in p-dimensional affine space $\mathscr{E}$. What is the possible range of values of dim $(\mathscr{U} \cap \mathscr{V})$? What is the condition on r and s such that $\mathscr{U} \cap \mathscr{V}$ must have dimension greater than zero?

2.4.7 Suppose that $U + \mathscr{U}$ and $V + \mathscr{V}$ are arbitrary hyperplanes of dimensions r and s, respectively, in p-dimensional affine space $\mathscr{E}$. Show that if $r + s \leq p$, then the hyperplanes do not necessarily have a common point, but that, if $r + s > p$, then the hyperplanes intersect in a hyperplane of dimension $\geq r + s - p$. (*Note. The results of Exercises* 2.4.6 *and* 2.4.7 *should be checked against three-dimensional geometric intuition.*)

2.4.8 Show that any parallelotope is invariant under the wide sense linear transformation of reflection in its center. What other wide sense linear transformations carry a parallelotope into itself?

2.5.1 Show that the matrix product $\mathbf{A}_1\mathbf{A}_2$ is not generally the same as $\mathbf{A}_2\mathbf{A}_1$ even when both products are defined and have the same dimensions.

2.5.2 Show that matrix multiplication is associative, i.e., that $(\mathbf{A}_1\mathbf{A}_2)\mathbf{A}_3 = \mathbf{A}_1(\mathbf{A}_2\mathbf{A}_3)$ provided that these products are all defined. Show how this property is used in deriving (2.5.5) and (2.5.8).

2.5.3 Show that any $m \times n$ matrix $\mathbf{A}$ has rank $r \leq \min(m, n)$.

2.5.4 Suppose that $\mathbf{M}$ is a $p \times q$ matrix of rank r and $\mathbf{N}$ is a $q \times q$ matrix of maximal rank q. Show that $\mathbf{MN}$ is, like $\mathbf{M}$, a $p \times q$ matrix of rank r.

2.5.5 If $\mathbf{A}_1$ and $\mathbf{A}_2$ are matrices of dimensions $p \times q$ and $q \times s$ and of ranks r_1 and r_2, respectively, what is known about the rank of $\mathbf{A}_1\mathbf{A}_2$?

2.5.6 Show that an $m \times n$ matrix $\mathbf{A}$ has rank $n - r$ if and only if there exist at most r linearly independent $n \times 1$ row vectors $\gamma_1, \gamma_2, \ldots, \gamma_r$ such that

$$\mathbf{A}\gamma_i = \mathbf{0} \qquad \text{for} \qquad i = 1, 2, \ldots, r.$$

2.5.7 Give a coordinate-free statement of the following result. The three equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= 0, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= 0, \text{ and} \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= 0 \end{aligned}$$

have a set of roots not all zero if and only if the matrix $\mathbf{A}$ of coefficients has rank less than 3. What is the nature of the set of all solutions if $\mathbf{A}$ has rank 2? Or if $\mathbf{A}$ has rank 1?

2.5.8 Under what conditions does the set of equations

$$\mathbf{AX} = \mathbf{B}$$

have a unique solution for $\mathbf{X}$, given that $\mathbf{A}$ and $\mathbf{B}$ have dimensions $n \times p$ and $p \times r$, respectively?

2.5.9 Show that the analytic form of a wide sense linear transformation of $\mathscr{E}$ into itself relative to a basis $\mathbf{V}$ is

$$\alpha \to \alpha^* = \mathbf{B} + \alpha\mathbf{A}.$$

2.5.10 Show that the rank of the sum of two $m \times n$ matrices is less than or equal to the sum of their ranks.

2.5.11 Show that, if $\mathbf{PQ}$, $\mathbf{P}^{-1}$, and $\mathbf{Q}^{-1}$ exist, then $(\mathbf{PQ})^{-1}$ exists and equals $\mathbf{Q}^{-1}\mathbf{P}^{-1}$.

2.6.1 Show that, if B is a pseudoinverse of A, then A is a pseudoinverse of B. What subspaces play the roles of $\mathscr{U}$, $\mathscr{U}^*$, $\mathscr{V}$, and $\mathscr{V}^*$ when A is regarded as a pseudoinverse of B?

2.6.2 Express the result of Exercise 2.6.1 in matrix terms.

2.6.3 Suppose that A is a $p \times q$ matrix of rank q where $p \geq q$. Suppose that a $p \times 1$ vector α is given and is known to have arisen as

$$\alpha = \mathsf{A}\beta$$

for some $q \times 1$ vector β. Show that β may be reconstructed from

$$\beta = \mathsf{B}\alpha,$$

where B is any pseudoinverse of A.

2.6.4 Suppose that B is a pseudoinverse of A. Show that B' is a pseudoinverse of A'.

2.6.5 Suppose that A is a $p \times q$ matrix with a $q \times p$ pseudoinverse matrix B. Suppose that $\mathbf{C}$ and $\mathbf{D}$ are nonsingular matrices of dimensions $p \times p$ and $q \times q$, respectively. Show that $\mathbf{D}^{-1}\mathsf{B}\mathbf{C}^{-1}$ is a pseudoinverse of $\mathbf{CAD}$.

CHAPTER 3

EUCLIDEAN VECTOR SPACES

3.1 ABSTRACT EUCLIDEAN VECTOR SPACES

Consider any p-dimensional vector space $\mathscr{E}$. It is proposed to enrich the structure of this space by giving it an *inner product*. An inner product is defined by a rule which assigns a real number (U, V) to each pair U and V of vectors in $\mathscr{E}$, i.e., an inner product is a function or mapping from the product space $\mathscr{E} \times \mathscr{E}$ to the real line. The term *inner product* will be used ambiguously either for the function or as an abbreviation for the *value* (U, V) *of the inner product* for given U and V.

The inner product must obey three simple rules:

i) *symmetry*, i.e.,

$$(U, V) = (V, U) \quad \text{for all } U \text{ and } V, \tag{3.1.1}$$

ii) *bilinearity*, i.e.,

$$(\mu_1 U_1 + \mu_2 U_2, V) = \mu_1(U_1, V) + \mu_2(U_2, V) \tag{3.1.2}$$

for all real numbers μ_1 and μ_2 and all vectors U_1, U_2, and V,

iii) *positiveness*, i.e.,

$$(U, U) > 0 \text{ for all } U \text{ different from } \emptyset. \tag{3.1.3}$$

The term bilinearity in (ii) is justified by the companion relation to (3.1.2) which follows from the application of (3.1.1) to each term in (3.1.2).

Some simple deductions from the axioms are

$$(\emptyset, V) = (V, \emptyset) = (\emptyset, \emptyset) = 0 \text{ for all } V, \text{ and} \tag{3.1.4}$$

$$\left(\sum_{i=1}^{r} \alpha_i U_i, \sum_{j=1}^{s} \beta_j V_j\right) = \sum_{i=1}^{r}\sum_{j=1}^{s} \alpha_i \beta_j (U_i, V_j), \tag{3.1.5}$$

for any real numbers $\alpha_1, \alpha_2, \ldots, \alpha_r$, $\beta_1, \ldots, \beta_s$ and vectors $U_1, \ldots, U_r$, $V_1, \ldots, V_s$. In particular,

$$\left(\sum_{i=1}^{p} \alpha_i V_i, \sum_{i=1}^{p} \beta_i V_i\right) = \sum_{i=1}^{p}\sum_{j=1}^{p} \alpha_i \beta_j (V_i, V_j), \tag{3.1.6}$$

where $V_1, \ldots, V_p$ denote any basis and $\sum_1^p \alpha_i V_i$ and $\sum_1^p \beta_i V_i$ are any two vectors in $\mathscr{E}$. Formula (3.1.6) implies that the inner product on $\mathscr{E}$ is defined by specifying it for all pairs of elements of any basis. In fact, if $\mathbf{Q}$ is the $p \times p$ matrix whose (i, j) element is (V_i, V_j), then the inner product of $\boldsymbol{\alpha}\mathbf{V}$ and $\boldsymbol{\beta}\mathbf{V}$ is given by

$$(\boldsymbol{\alpha}\mathbf{V}, \boldsymbol{\beta}\mathbf{V}) = \boldsymbol{\alpha}\mathbf{Q}\boldsymbol{\beta}', \tag{3.1.7}$$

a simple elegant formula. Such a matrix $\mathbf{Q}$ will be called the *inner product matrix* relative to the basis $\mathbf{V}$.

To show that the definition of an inner product is not empty, i.e., that an inner product exists, simply set

$$(\boldsymbol{\alpha}\mathbf{V}, \boldsymbol{\beta}\mathbf{V}) = \boldsymbol{\alpha}\boldsymbol{\beta}'. \tag{3.1.8}$$

This is equivalent to choosing $\mathbf{Q}$ in (3.1.7) to be $\mathbf{I}$. It is easily checked that the quantities defined by (3.1.8) satisfy the axiomatic requirements for an inner product. Using different bases in place of $\mathbf{V}$, one can define an inner product in different ways on the same space $\mathscr{E}$. Conversely, it will be seen in Chapter 4 that any inner product can be characterized as in (3.1.8) for some basis, in fact for an infinite family of bases.

A vector space with an inner product defined on it will be called a *Euclidean vector space*. In Euclidean vector space terms, the square root of the inner product (V, V) is called the *norm* of V, and a pair of vectors U and V is said to be *orthogonal* if $(U, V) = 0$. A set of vectors such that every pair of vectors in the set is orthogonal will be called a (*mutually*) *orthogonal* set, and, if the set constitutes a basis of $\mathscr{E}$, it will be called simply an *orthogonal basis*. A set of vectors is called *orthonormal* if it is orthogonal and each vector has unit norm. For example, $\mathbf{V}$ is an orthonormal set according to the inner product defined by (3.1.8). If an orthonormal set $\mathbf{V}$ constitutes a basis it is called an *orthonormal basis*. Two subspaces $\mathscr{U}$ and $\mathscr{V}$ are said to be *orthogonal* if $(U, V) = 0$ for every U in $\mathscr{U}$ and V in $\mathscr{V}$.

Example 3.1.1. Consider the variable-space of Example 2.1.1 where each variable or linear combination of variables is a point. Unfamiliar as it may be for a statistician to think of such variables as points, still it is very familiar to think of an inner product on this vector space. Simply substitute the familiar terms covariance and variance for the less familiar terms inner product and norm squared, i.e.,

$$(U, V) = \operatorname{cov}(U, V) \quad \text{and} \quad (U, U) = \operatorname{cov}(U, U) = \operatorname{var}(U). \tag{3.1.9}$$

It can be easily checked that the sample covariance function introduced in (1.3.5) satisfies the requirement for an inner product provided that a variable $\sum_1^p \alpha_t V_t$ has zero variance only if $\alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$. In fact, the covariance matrix whose (i, j) element is $\operatorname{cov}(V_i, V_j)$ plays the role of $\mathbf{Q}$ above and gives the standard form of the definition of an inner product in terms of its

values for a set of basis elements. The converse result that any inner product function is a possible covariance function is easy to prove. Thus the notions of covariance and inner product are abstractly identical.

Example 3.1.2. At this point, one might suspect that variances and covariances could be used to define a natural inner product for the individual-space of Example 2.1.2. This is true, but not trivial, and a discussion of this matter is postponed to Chapter 6.

3.2 EUCLIDEAN GEOMETRIC SPACE

Euclidean geometry bears the same relation to affine geometry as Euclidean vector spaces bear to ordinary vector spaces. The geometric terminology introduced in Section 2.4 remains valid, and additional terminology becomes possible when affine space is enriched to Euclidean space. For example, the norm of V may now be called the *distance* from $\emptyset$ to V or the *length* of the line segment $\emptyset V$. More generally, the distance from U to V or the length of the line segment UV is given by the norm of $U - V$, i.e.,

$$\sqrt{(U - V, U - V)}. \tag{3.2.1}$$

Similarly, the line segments UV and U_1V_1, or the lines determined by them, are said to make *angle* θ with one another, where

$$\cos\theta = \frac{(U - V, U_1 - V_1)}{(U - V, U - V)^{1/2}(U_1 - V_1, U_1 - V_1)^{1/2}}. \tag{3.2.2}$$

The line segments UV and U_1V_1 or the lines determined by them are said to be *perpendicular* if $\cos\theta = 0$, i.e., if

$$(U - V, U_1 - V_1) = 0. \tag{3.2.3}$$

Two hyperplanes $V + \mathscr{V}$ and $U + \mathscr{U}$ are said to be *perpendicular* if the subspaces $\mathscr{U}$ and $\mathscr{V}$ are orthogonal.

To check that these definitions of geometric terms agree with what one intuitively regards as ordinary Euclidean geometry in $p = 2$ dimensions, consider a Euclidean plane with an orthonormal basis W_1, W_2 corresponding to a pair of perpendicular line segments $\emptyset W_1$, $\emptyset W_2$ of unit length, which define perpendicular coordinate axes, as in Fig. 3.2.1. Given the coordinates of the end points of any pair of line segments UV and U_1V_1, the standard formulas of elementary analytic geometry for the length of UV, the cosine of the angle between UV and U_1V_1, and the perpendicularity of UV and U_1V_1 are seen to agree with (3.2.1), (3.2.2), and (3.2.3), respectively, when the inner product is defined as in (3.1.8) using the basis W_1, W_2. Similar checking could be done with three-dimensional Euclidean geometry. In the case of p dimensions, for $p > 3$ geometric intuition must be backed up by abstract theory.

The statistical importance of the present section is that the variable-space of Example 2.1.1, together with a definition of covariance between each pair of variables, may be regarded as an ordinary Euclidean p-dimensional geometric space with its associated well-known concepts and propositions. For example, variances may be regarded as squared lengths, zero covariance means perpendicular, and the *correlation coefficient*

$$\rho_{ij} = \text{cov}\,(V_i, V_j)/[\text{var}\,(V_i)\,\text{var}\,(V_j)]^{1/2} \tag{3.2.4}$$

between V_i and V_j is the cosine of the angle between $\emptyset V_i$ and $\emptyset V_j$.

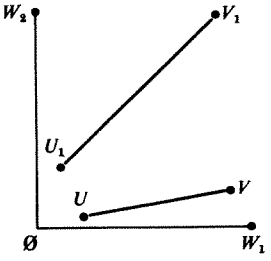

Fig. 3.2.1. The Euclidean plane with orthonormal coordinate axes $\emptyset W_1$ and $\emptyset W_2$ and arbitrary line segments UV and U_1V_1.

The concept of inner product is geometrically equivalent to the concept of an ellipsoid centered at the origin. In the present development an *ellipsoid centered at the origin* is defined to be the "solid figure" consisting of points V such that $(V, V) \leq 1$ according to some inner product. The points V such that $(V, V) = 1$ are said to be on the *surface* of the ellipsoid while the remaining points of the ellipsoid are interior points. Two different inner products cannot give rise to the same ellipsoid. To see this, note that the set of norms determined by an inner product uniquely define the whole inner product function, for

$$(U, V) = \tfrac{1}{4}[(U + V, U + V) - (U - V, U - V)]. \tag{3.2.5}$$

Consequently, if two inner products are different, there must exist a vector W with different norms under the two inner products. The vector αW for suitably chosen α has unit norm under one inner product and norm greater than unity under the other inner product, i.e., αW belongs to one ellipsoid but not to the other, and so the ellipsoids are different. The notation π will often be used either for an inner product or for its associated ellipsoid.

In terms of Euclidean geometry, the set of points at unit distance from the origin defines the unit sphere centered at the origin. However, even if an inner product is fixed in a space it is possible to consider a further range of inner products and the corresponding range of ellipsoids of which the unit sphere is just one. This idea is illustrated in the case $p = 2$ where a sphere is just a circle and an ellipsoid is an ellipse. Figure 3.2.2 is intended to be an affine plane where, because of an inability to draw such spaces without an implied Euclidean inner

product, vectors W_1 and W_2 appear to be orthonormal. If an inner product is defined by taking W_1, W_2 as an orthonormal basis, then the set of points at unit distance from Ø simply traces the unit circle. However, suppose that an arbitrary pair of vectors Z_1, Z_2 is chosen to be orthonormal for a *different* definition of inner product on the *same* space. Then it may be checked that the set of points at unit distance from the origin defines an ellipse as drawn. This ellipse has the property that the tangent line at Z_1 is parallel to $ØZ_2$ and the tangent line at Z_2 is parallel to $ØZ_1$, and no other ellipse has this property. These tangency properties will be derived in the p-dimensional case shortly. Note that the tangency properties appear obviously true in the case of the inner

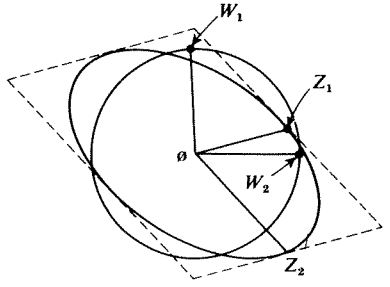

Fig. 3.2.2. Inner product ellipses associated with orthonormal bases W_1, W_2 and Z_1, Z_2. The latter is shown as conjugately contained in an associated parallelogram.

product represented by the unit circle, and, since any two bases are equivalent from an affine point of view, one might expect these properties to hold for any basis in a reasonable mathematical theory.

Any line through the center of an ellipsoid intersects the ellipsoid in a line segment which will be called an *axis* of the ellipsoid. The center divides such an axis into two *semi-axes* of equal length. In discussing a p-dimensional ellipsoid the word *conjugate* is often used in the same way that the word orthogonal is used in vector space language. Thus if two lines through Ø are orthogonal under an inner product, they determine *conjugate axes* of the associated ellipsoid. The line segments $ØW_1, ØW_2, \ldots, ØW_p$ defined by an orthonormal basis **W** are *a set of conjugate semi-axes* of the ellipsoid.

Suppose that $\mathscr{V}_0$ is the $(p-1)$-dimensional subspace of $\mathscr{E}$ orthogonal to some unit vector V_0 according to a given inner product. Then $V_0 + \mathscr{V}_0$ is defined to be the tangent hyperplane at V_0 to the ellipsoid of the inner product. This definition makes sense because (i) V_0 lies on both the hyperplane and the surface of the ellipsoid and (ii) every other point $V_0 + V$ lying on $V_0 + \mathscr{V}_0$ has

$$\begin{aligned}(V_0 + V, V_0 + V) &= (V_0, V_0) + 2(V_0, V) + (V, V)\\ &= (V_0, V_0) + (V, V)\\ &> (V_0, V_0),\end{aligned}$$

and so is outside of the ellipsoid. Now suppose that W_1 and W_2 are any pair

of orthonormal vectors. Then the tangent hyperplanes $W_1 + \mathscr{W}_1$ and $W_2 + \mathscr{W}_2$ clearly have the property that W_1 lies in $\mathscr{W}_2$ and W_2 lies in $\mathscr{W}_1$. More generally, if $W_1, W_2, \ldots, W_p$ constitute an orthonormal basis with tangent planes $W_i + \mathscr{W}_i$, then each W_i lies in $\mathscr{W}_j$ for $i \neq j$. In fact, $\mathscr{W}_j$ may be characterized as the subspace spanned by the W_i for $i \neq j$.

It is also natural to say that the ellipsoid is *conjugately contained* in the parallelotope with faces $\pm W_i + \mathscr{W}_i$, for the ellipsoid is tangent to each face of the parallelotope, and these tangent faces have the property of being conjugate to their points of contact. It is clear that only one ellipsoid can be conjugately inscribed in a given parallelotope, but it will be seen that many parallelotopes have this property relative to a given ellipsoid, i.e., a specified orthonormal basis determines an inner product, but a given inner product has many orthonormal bases.

In thinking of the ellipsoid of an inner product, it is instructive to note that the size of the ellipsoid varies inversely with the size of the inner product. To be more precise, changing the inner product from (U, V) to $\lambda(U, V)$ has the effect of multiplying the length of each axis by the factor $\lambda^{-1/2}$.

The foregoing discussion has defined only ellipsoids with center at Ø. To define an ellipsoid having an arbitrary center Ø*, simply translate an ellipsoid with center Ø, i.e., if V ranges over a given ellipsoid with center Ø, then Ø* + V ranges over the translated ellipsoid with center Ø*.

3.3 ORTHOGONAL LINEAR TRANSFORMATIONS AND ORTHOGONAL PROJECTIONS

An *orthogonal linear transformation* of a Euclidean space $\mathscr{E}$ into itself may be defined as a linear transformation $V \to V^*$ (narrow sense) which leaves the norm unchanged, i.e.,

$$(V, V) = (V^*, V^*) \tag{3.3.1}$$

for all V in $\mathscr{E}$. It is left to the reader to show that not only is the norm preserved, but every inner product is preserved, i.e.,

$$(V, W) = (V^*, W^*) \tag{3.3.2}$$

for every V and W in $\mathscr{E}$. Consequently, any configuration of lengths and angles determined by a set of vectors $U, V, W, \ldots$ is mirrored precisely by the set of transformed vectors $U^*, V^*, W^*, \ldots$ It is obvious that *every orthogonal linear transformation of the p-dimensional Euclidean space $\mathscr{E}$ into itself has rank p*, for if the rank were $r < p$, then a subspace $\mathscr{U}$ including vectors having nonzero norm would transform into the origin which has zero norm.

It is easily checked that any translation of a Euclidean space $\mathscr{E}$ leaves all lengths and angles invariant. Thus a translation followed by a narrow sense orthogonal linear transformation may naturally be called a *wide sense orthogonal linear transformation* because it leaves invariant all configurations of lengths

and angles concerning a set of points. If Euclidean space is regarded as physical space, then such a wide sense orthogonal linear transformation may be regarded as a rigid motion in the space, perhaps followed by reflections. An analytic approach to characterizing the class of all orthogonal linear transformations involves orthogonal matrices which are taken up in Section 3.4, and methods of constructing such matrices are presented in Chapter 4.

An *orthogonal projection* is a special type of linear projection where the family of parallel hyperplanes along which the projected points move is orthogonal to the hyperplane into which points are projected. The definition of an orthogonal projection requires that one have a pair of subspaces $\mathscr{U}$ and $\mathscr{V}$ which are both orthogonal and complementary. It will be proved in Section 4.1 that the set of vectors orthogonal to any subspace $\mathscr{U}$ defines a subspace $\mathscr{V}$ which is complementary to $\mathscr{U}$, and, since no other subspace can be both orthogonal and complementary to $\mathscr{U}$, the subspace $\mathscr{V}$ may be called *the orthogonal complement of* $\mathscr{U}$. If $\mathscr{V}$ is the orthogonal complement of $\mathscr{U}$, then $\mathscr{U}$ is the orthogonal complement of $\mathscr{V}$. If $\mathbf{W}$ is an orthonormal basis, then $W_1, W_2, \ldots, W_r$ and $W_{r+1}, W_{r+2}, \ldots, W_p$ span a pair of orthogonal complements. With this theory in hand, it is clear that the concept of *orthogonal projection into the hyperplane* $V + \mathscr{V}$ is uniquely determined by $V + \mathscr{V}$, for $\mathscr{V}$ determines the orthogonal complement $\mathscr{U}$ along which the projection takes place.

A different characterization of an orthogonal projection is the following. The orthogonal projection W^* of any point W into $V + \mathscr{V}$ is that point in $V + \mathscr{V}$ at minimum distance from W, i.e., that point W^* such that

$$(W - W^*, W - W^*) < (W - V^*, W - V^*) \tag{3.3.3}$$

for any V^* in $V + \mathscr{V}$ different from W^*. To prove this, note that the line segment WW^* lies in the hyperplane $W + \mathscr{U}$ where $\mathscr{U}$ is the orthogonal complement of $\mathscr{V}$, and hence WW^* is perpendicular to any line segment in $\mathscr{V}$. Thus,

$$(W - W^*, V^* - W^*) = 0, \tag{3.3.4}$$

so that

$$\begin{aligned}(W - V^*, W - V^*) &= ([W - W^*] - [V^* - W^*], [W - W^*] - [V^* - W^*]) \\ &= (W - W^*, W - W^*) - 2(W - W^*, V^* - W^*) + (V^* - W^*, V^* - W^*) \\ &= (W - W^*, W - W^*) + (V^* - W^*, V^* - W^*) \\ &> (W - W^*, W - W^*),\end{aligned} \tag{3.3.5}$$

whenever V^* is different from W^*, as required.

A third characterization of an orthogonal projection is as follows: if W^* in $V + \mathscr{V}$ has the property that WW^* is perpendicular to every line segment VV^* in $V + \mathscr{V}$, then W^* is the orthogonal projection of W into $V + \mathscr{V}$. The proofs of this theorem and its converse are left to the reader.

When $V + \mathscr{V}$ is simply $\mathscr{V}$, i.e., a hyperplane through the origin, the orthogonal projection W^* of W into $\mathscr{V}$ is often called the *component of* W *along* $\mathscr{V}$ and $W - W^*$ is called the *component of* W *orthogonal to* $\mathscr{V}$. Note that $W = W^* + [W - W^*]$ is the sum of these two components.

The reader should consider the intuitive meaning of the operation of orthogonal projection in ordinary Euclidean space of two and three dimensions.

3.4 POSITIVE DEFINITE SYMMETRIC MATRICES AND ORTHOGONAL MATRICES

A $p \times p$ matrix $\mathbf{Q}$ is said to be *positive definite* if $\boldsymbol{\alpha}\mathbf{Q}\boldsymbol{\alpha}' > 0$ for every $1 \times p$ vector $\boldsymbol{\alpha} \neq \mathbf{0}$. Given a Euclidean space $\mathscr{E}$ and a basis $\mathbf{V}$, the inner product matrix $\mathbf{Q}$ defined in Section 3.1 is clearly positive definite and symmetric. Conversely, any positive definite symmetric matrix may be used to define an inner product over a vector space $\mathscr{E}$ in terms of a basis $\mathbf{V}$ via (3.1.7). Thus the concept of positive definite symmetric matrix is the natural analytic coordinate-dependent counterpart of the concept of inner product. When the inner product is a covariance, the inner product matrix relative to $\mathbf{V}$ will be called the *covariance matrix* of $\mathbf{V}$. Thus the class of all possible covariance matrices is identical with the class of all positive definite symmetric matrices. Similarly, in geometric terms the class of all ellipsoids centered at the origin is in one-one correspondence with the class of all positive definite symmetric matrices. For given any basis $\mathbf{V}$, the ellipsoid of the inner product with matrix $\mathbf{Q}$ consists of points $\boldsymbol{\alpha}\mathbf{V}$ such that

$$\boldsymbol{\alpha}\mathbf{Q}\boldsymbol{\alpha}' \leq 1. \tag{3.4.1}$$

Since covariance inner products based on sample data must in general be represented in terms of a covariance matrix $\mathbf{Q}$, computations with such matrices are basic for multivariate analyses. Chapters 4 and 5 are largely concerned with operations on positive definite symmetric matrices.

A basic lemma which will be used over and over again, beginning in the next paragraph, is that *if* $\mathbf{Q}$ *and* $\mathbf{Q}^*$ *denote inner product matrices of* $\mathscr{E}$ *relative to a pair of bases* $\mathbf{V}$ *and* $\mathbf{V}^*$, *where* $\mathbf{V}^* = \mathbf{A}\mathbf{V}$, *then*

$$\mathbf{Q}^* = \mathbf{A}\mathbf{Q}\mathbf{A}'. \tag{3.4.2}$$

The proof is left to the reader.

A $p \times p$ matrix $\mathbf{G}$ is said to be *orthogonal* if

$$\mathbf{G}\mathbf{G}' = \mathbf{I}. \tag{3.4.3}$$

Equivalent definitions are obviously

$$\mathbf{G}' = \mathbf{G}^{-1} \tag{3.4.4}$$

or

$$\mathbf{G}'\mathbf{G} = \mathbf{I}. \tag{3.4.5}$$

Clearly $\mathbf{G}$ is orthogonal if and only if $\mathbf{G}'$ is orthogonal. Orthogonal matrices

arise in relating two orthonormal bases and in expressing an orthogonal linear transformation in terms of an orthonormal basis. Thus, if $\mathbf{U}$ is an orthonormal basis, then from (3.4.2) $\mathbf{U}^* = \mathbf{GU}$ is an orthonormal basis if and only if $\mathbf{G}$ is an orthogonal matrix. Note that the rows of $\mathbf{G}$ express $\mathbf{U}^*$ in terms of $\mathbf{U}$ while the columns of $\mathbf{G}$ express $\mathbf{U} = \mathbf{G}'\mathbf{U}^*$ in terms of $\mathbf{U}^*$. In a similar vein, it is clear that the linear transformation $\boldsymbol{\alpha}\mathbf{U} \rightarrow \boldsymbol{\alpha}\mathbf{GU}$ where $\mathbf{U}$ is an orthonormal basis of $\mathscr{E}$ is an orthogonal linear transformation if and only if $\mathbf{G}$ is an orthogonal matrix.

The simplest orthogonal matrix is no doubt $\mathbf{I}$. Perhaps the simplest nontrivial class of orthogonal matrices is defined by

$$\mathbf{G} = \mathbf{I} - 2\boldsymbol{\gamma}'\boldsymbol{\gamma}, \tag{3.4.6}$$

where $\boldsymbol{\gamma}$ is any $1 \times p$ vector satisfying $\boldsymbol{\gamma}\boldsymbol{\gamma}' = 1$. These matrices may be called *elementary orthogonal matrices* and are playing an increasingly important role in modern computational practice under the name *Householder transformations.* (See Exercise 4.3.3 and Section 5.4. See also Householder (1964), Wilkinson (1965), and papers referred to by these authors.) Regarded as a linear transformation in Euclidean space, $\mathbf{G}$ in (3.4.6) expresses the simple geometric notion of reflection. To see this, suppose that $\mathbf{U}$ is an orthonormal basis of $\mathscr{E}$. Then $\boldsymbol{\gamma}\mathbf{U}$ defines a unit vector and any $V = \boldsymbol{\alpha}\mathbf{U}$ may be decomposed into a component $(\boldsymbol{\alpha}\boldsymbol{\gamma}')\boldsymbol{\gamma}\mathbf{U}$ along $\boldsymbol{\alpha}\mathbf{U}$ and a component $V - (\boldsymbol{\alpha}\boldsymbol{\gamma}')\boldsymbol{\gamma}\mathbf{U}$ in the subspace orthogonal to $\boldsymbol{\gamma}\mathbf{U}$. The linear transformation of reflection in the subspace orthogonal to $\boldsymbol{\gamma}\mathbf{U}$ is defined by changing the sign of the component orthogonal to the subspace while leaving the component in the subspace unchanged. Thus $V = \boldsymbol{\alpha}\mathbf{U} = (\boldsymbol{\alpha}\boldsymbol{\gamma}')\boldsymbol{\gamma}\mathbf{U} + [\boldsymbol{\alpha} - (\boldsymbol{\alpha}\boldsymbol{\gamma}')\boldsymbol{\gamma}]\mathbf{U} \rightarrow -(\boldsymbol{\alpha}\boldsymbol{\gamma}')\boldsymbol{\gamma}\mathbf{U} + [\boldsymbol{\alpha} - (\boldsymbol{\alpha}\boldsymbol{\gamma}')\boldsymbol{\gamma}]\mathbf{U} = \boldsymbol{\alpha}[\mathbf{I} - 2\boldsymbol{\gamma}'\boldsymbol{\gamma}]\mathbf{U}$ in agreement with (3.4.6).

As a first illustration of the use of elementary orthogonal matrices, consider the problem of finding an orthogonal matrix with a given first row. Any $1 \times p$ matrix $\mathbf{g}$ satisfying $\mathbf{g}\mathbf{g}' = 1$ is a candidate for the first row of an orthogonal matrix. In terms of an orthonormal basis $\mathbf{U}$ of a space $\mathscr{E}$, what is needed is an orthogonal linear transformation $\mathbf{U} \rightarrow \mathbf{GU}$ carrying $U_1 \rightarrow \mathbf{gU}$. This is easily accomplished by reflection in the subspace which bisects the angle between the vectors U_1 and $\mathbf{gU}$. Thus $\boldsymbol{\gamma}\mathbf{U}$ lies along $U_1 - \mathbf{gU}$ or

$$\boldsymbol{\gamma} = \lambda(1 - g_1, -g_2, -g_3, \ldots, -g_p), \tag{3.4.7}$$

where

$$\lambda = \pm[2(1 - g_1)]^{-1/2} \tag{3.4.8}$$

and $\mathbf{G}$ is defined by (3.4.6).

The reader should check directly that $\mathbf{G}$ defined by (3.4.6) is orthogonal and that, if (3.4.7) is used, the first row of $\mathbf{G}$ is $\mathbf{g}$.

3.5 DETERMINANTS AND VOLUMES

The concept of determinant plays an incidental role in the theory of this book, usually in the context of describing a volume. A number called the *determinant*

of $\mathbf{A}$ and denoted by $\det \mathbf{A}$ may be associated with every square matrix $\mathbf{A}$. If $\mathbf{A}$ is a $p \times p$ matrix with (i, j) element a_{ij}, its determinant may be defined by the formula

$$\det \mathbf{A} = \sum \pm a_{1j(1)}\, a_{2j(2)} \cdots a_{pj(p)}, \tag{3.5.1}$$

where summation is over the $p!$ permutations $j(1), j(2), \ldots, j(p)$ of the integers $1, 2, \ldots, p$ and the signs $+$ and $-$ are used depending on whether an even or odd number of interchanges of pairs is required to modify the permutation $1, 2, \ldots, p$ into $j(1), j(2), \ldots, j(p)$.

A few of the basic properties of determinants will be given without proof. If $\mathbf{A}$ is a triangular matrix, i.e., either $a_{ij} = 0$ for all $i > j$ or $a_{ij} = 0$ for all $i < j$, then $\det \mathbf{A}$ is given by the product of the diagonal elements, i.e.,

$$\det \mathbf{A} = \prod_{i=1}^{p} a_{ii}. \tag{3.5.2}$$

If $\mathbf{A}$ and $\mathbf{B}$ are both $p \times p$ matrices, then

$$\det \mathbf{AB} = \det \mathbf{A} \times \det \mathbf{B}. \tag{3.5.3}$$

For any nonsingular matrix $\mathbf{A}$

$$\det \mathbf{A}^{-1} = 1/\det \mathbf{A}, \tag{3.5.4}$$

while $\mathbf{A}$ is nonsingular if and only if $\det \mathbf{A} \neq 0$. For any orthogonal matrix $\mathbf{G}$,

$$\det \mathbf{G} = \pm 1. \tag{3.5.5}$$

The concept of r-dimensional volume in p-dimensional Euclidean space generalizes the concept of length, which is volume in the case $r = 1$. Two parallel line segments have a ratio of lengths which is affinely determined, and they have actual lengths in Euclidean space whose ratio coincides with the affinely determined ratio of lengths. Similarly, it is possible to develop a theory of r-dimensional volumes whereby regions of parallel r-dimensional hyperplanes have affinely determined ratios of r-dimensional volumes and have consistent actual volumes when the space is regarded as Euclidean. To derive this theory would be to go too far afield into areas of measure and integration, so only a few facts will be given with some heuristic justification.

A set of r linearly independent points $W_1, W_2, \ldots, W_r$ determines the r-dimensional parallelotope consisting of the points $\sum_1^r c_i W_i$ where $0 \leq c_i \leq 1$ for $i = 1, 2, \ldots, r$. Suppose that $U_1, U_2, \ldots, U_r$ is an orthonormal set spanning the same subspace as $W_1, W_2, \ldots, W_r$ and that $\mathbf{W} = [W_1, W_2, \ldots, W_r]' = \mathbf{DU} = \mathbf{D}[U_1, U_2, \ldots, U_r]'$ for some $r \times r$ matrix $\mathbf{D}$. Then it can be shown (cf. Sommerville, 1958) that the volume of the parallelotope is given by $|\det \mathbf{D}|$. Thus the square of the volume is given by $\det \mathbf{Q}$ where $\mathbf{Q} = \mathbf{DD}'$ is the inner product matrix of $W_1, W_2, \ldots, W_r$. The role of this formula as a generalization of the familiar "area = base × height" rule for a parallelogram will be pointed out in Section 4.1.

An r-dimensional unit sphere consists of the set of points $\sum_1^r \alpha_i U_i$ such that

$$\sum_{i=1}^{r} \alpha_i^2 \leq 1 \tag{3.5.6}$$

where $U_1, U_2, \ldots, U_r$ is an orthonormal set. By ordinary multiple integration the volume of the region (3.5.6) is shown in many calculus books to be

$$K_r = \frac{\pi^{r/2}}{\Gamma(r/2 + 1)}, \tag{3.5.7}$$

where $\Gamma(\ldots)$ denotes the gamma function given by $\Gamma(m + 1) = m!$ and $\sqrt{\pi}(1/2)(3/2)\ldots([2m - 1]/2)$ for $m = 0, 1, 2, \ldots$. The integration for a general ellipsoid is not essentially more difficult and may be facilitated by a special choice of $U_1, U_2, \ldots, U_r$ which are orthonormal not only for the given inner product of the Euclidean space $\mathscr{E}$ but are also orthogonal for the inner product defined by the general ellipsoid. This amounts to locating the principal axes of the ellipsoid, as discussed in Section 5.2.

A more heuristic argument will be used to derive an expression for the volume of a general r-dimensional ellipsoid. Such an ellipsoid may be specified by a set of conjugate axes $\mathcal{O}W_1, \mathcal{O}W_2, \ldots, \mathcal{O}W_r$ where $W_1, W_2, \ldots, W_r$ are any set of linearly independent vectors. This ellipsoid is conjugately contained in the parallelotope of points $\sum_1^r c_i W_i$ where $-1 \leq c_i \leq 1$ for $i = 1, 2, \ldots, r$. This parallelotope is twice as long in every dimension as the parallelotope considered above and so has volume $2^r[\det \mathbf{Q}]^{1/2}$ where $\mathbf{Q}$ is the inner product matrix of $W_1, W_2, \ldots, W_r$. At the same time, the ratio of the volume of the ellipsoid to the volume of the parallelotope is known to be $K_r/2^r$, since the ratio is affinely invariant and therefore the same as in the special case where $W_1, W_2, \ldots, W_r$ form an orthonormal set. Thus the volume of the ellipsoid is

$$K_r[\det \mathbf{Q}]^{1/2}, \tag{3.5.8}$$

where K_r is defined in (3.5.7).

It may be helpful in interpreting (3.5.8) to express the ellipsoid in terms of an orthonormal set $\mathbf{U}$. The point $\boldsymbol{\alpha}\mathbf{U} = \boldsymbol{\beta}\mathbf{W}$ lies in the ellipsoid if $\boldsymbol{\beta}\boldsymbol{\beta}' \leq 1$. But $\mathbf{W} = \mathbf{D}\mathbf{U}$ where $\mathbf{D}\mathbf{D}' = \mathbf{Q}$ so that the desired definition of the ellipsoid in $\mathbf{U}$ coordinates is

$$\boldsymbol{\alpha}\mathbf{Q}^{-1}\boldsymbol{\alpha}' \leq 1. \tag{3.5.9}$$

3.6 SEMI-DEFINITE INNER PRODUCTS

The class of inner products over a vector space $\mathscr{E}$ may be widened by relaxing the axiom (3.1.3) to

$$(U, U) \geq 0 \qquad \text{for all } U \text{ in } \mathscr{E}. \tag{3.6.1}$$

The additional inner products permitted by this relaxation will be called *semi-definite* while the labels *proper* or *definite* will be used to distinguish the original class. The term *wide sense inner product* will be used to cover both types. The general theory of semi-definite inner products is easily derived once the theory of proper inner products is known.

A wide sense inner product is semi-definite if and only if there exists a $U \neq \mathcal{O}$ in $\mathscr{E}$ such that $(U, U) = 0$. From (3.1.1) and (3.1.2) it follows that the set of vectors such that $(U, U) = 0$ forms a subspace $\mathscr{U}$ of $\mathscr{E}$. A wide sense inner product is semi-definite if and only if its associated $\mathscr{U}$ has dimension greater than zero. Defining the *rank* of an inner product to be $p - \dim(\mathscr{U})$, it follows that a proper inner product has rank p while a semi-definite inner product has rank f, satisfying $0 \leq f \leq p - 1$.

Any vector U in $\mathscr{U}$ has the following properties:

$$(U, V) = 0 \qquad \text{for all } V \text{ in } \mathscr{E}, \tag{3.6.2}$$

and

$$(V + U, V + U) = (V, V) \qquad \text{for all } V \text{ in } \mathscr{E}. \tag{3.6.3}$$

Conversely, if (3.6.3) holds for a given U, then U is in $\mathscr{U}$. Thus, $\mathscr{U}$ may be characterized as that subspace whose elements do not affect the norm when added to any vector.

Suppose that $\mathscr{V}$ is any f-dimensional subspace complementary to $\mathscr{U}$. Clearly, if the inner product is considered only over $\mathscr{V}$, it is no longer semi-definite (although still of rank f). Furthermore, to specify a semi-definite inner product completely, one need only specify it on any $\mathscr{V}$ complementary to $\mathscr{U}$. For, if W_1 and W_2 are any vectors in $\mathscr{E}$ and $W_1 = V_1 + U_1$ and $W_2 = V_2 + U_2$ where V_1 and V_2 are in $\mathscr{V}$ and U_1 and U_2 are in $\mathscr{U}$, then from (3.6.2)

$$(W_1, W_2) = (V_1, V_2). \tag{3.6.4}$$

Thus the structure of semi-definite inner products as well as their means of construction has been made clear.

In geometric terms, a semi-definite inner product may be thought of in relation to the operation of projection along the family of $(p - f)$-dimensional hyperplanes parallel to $\mathscr{U}$ into the f-dimensional hyperplane $\mathscr{V}$. All those points which project into the same point in $\mathscr{V}$ are indistinguishable in relation to the semi-definite inner product, i.e., lie at distance zero from one another. The ellipsoid of a semi-definite inner product is an ellipsoidal cylinder based on the f-dimensional ellipsoid in the subspace $\mathscr{V}$ and including all of the hyperplanes $V + \mathscr{U}$ where $(V, V) = 1$.

A $p \times p$ symmetric matrix $\mathbf{Q}$ will be called *positive semi-definite symmetric* provided that

$$\boldsymbol{\alpha}\mathbf{Q}\boldsymbol{\alpha}' \geq 0 \tag{3.6.5}$$

for any $1 \times p$ vector $\boldsymbol{\alpha}$ with equality holding in (3.6.5) for some $\boldsymbol{\alpha} \neq \mathbf{0}$. If such a $\mathbf{Q}$ is used to define an inner product over a vector space $\mathscr{E}$ relative to a basis $\mathbf{V}$ of $\mathscr{E}$, then the resulting inner product is semi-definite, and conversely any

inner product matrix for a semi-definite inner product is a positive semi-definite symmetric matrix. It will now be shown that a semi-definite inner product and its inner product matrix relative to any basis must have the same rank. Suppose that a rank f inner product has the inner product matrix $\mathbf{Q}$ relative to a basis $\mathbf{V}$ of $\mathscr{E}$. Suppose that as above $\mathscr{U}$ is the $(p-f)$-dimensional subspace of vectors with zero norm and that $\mathscr{V}$ is any f-dimensional subspace complementary to $\mathscr{U}$. By the process of Chapter 4 there may be found an orthonormal basis $\mathbf{U}_1$ of $\mathscr{V}$. This basis $\mathbf{U}_1$ of $\mathscr{V}$ together with any basis $\mathbf{U}_2$ of $\mathscr{U}$ determines a basis

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{bmatrix} \tag{3.6.6}$$

of $\mathscr{E}$ whose inner product matrix $\mathbf{I}_f$ is a $p \times p$ matrix whose first f diagonal elements are unity and whose remaining elements are all zero. If $\mathbf{V} = \mathbf{DU}$, then

$$\mathbf{Q} = \mathbf{D}\mathbf{I}_f\mathbf{D}', \tag{3.6.7}$$

which, since $\mathbf{D}$ has rank p, clearly has rank f.

Incidentally, from (2.6.10) and (2.6.11) a pseudoinverse $\dot{\mathbf{Q}}$ of $\mathbf{Q}$ is given by

$$\dot{\mathbf{Q}} = \mathbf{C}'\mathbf{I}_f\mathbf{C}, \tag{3.6.8}$$

where $\mathbf{C} = \mathbf{D}^{-1}$. Note further that if $\mathbf{D}_1$ denotes the $p \times f$ matrix consisting of the first f columns of $\mathbf{D}$ and $\mathbf{C}_1$ is an $f \times p$ matrix consisting of the first f rows of $\mathbf{C}$, then (3.6.7) and (3.6.8) may be written

$$\mathbf{Q} = \mathbf{D}_1\mathbf{D}_1' \tag{3.6.9}$$

and

$$\dot{\mathbf{Q}} = \mathbf{C}_1'\mathbf{C}_1, \tag{3.6.10}$$

where $\mathbf{C}_1$ is a pseudoinverse of $\mathbf{D}_1$.

As an illustration of the concept of a wide sense inner product, consider a linear transformation A from a vector space $\mathscr{E}$ to a vector space $\mathscr{E}^*$. Curiously enough, A provides a very simple mechanism for carrying an inner product in the reverse direction from $\mathscr{E}^*$ to $\mathscr{E}$. For if π^* is a wide sense inner product defined over $\mathscr{E}^*$ with $(U^*, V^*)^*$ denoting the corresponding value for U^* and V^* in $\mathscr{E}^*$, then

$$(U, V) = (\mathsf{A}U, \mathsf{A}V)^* \tag{3.6.11}$$

for each U and V in $\mathscr{E}$ defines a wide sense inner product over π. The subspace $\mathscr{U}$ of $\mathscr{E}$ consisting of vectors with zero norm according to π may be characterized as that set of vectors in $\mathscr{E}$ which map under A into the subspace U^* of $\mathscr{E}^*$ whose vectors have zero norm according to π^*. In particular, if π^* is a proper inner product, then $\mathscr{U}$ is determined by A alone to be the subspace of $\mathscr{E}$ which maps into the origin in $\mathscr{E}^*$.

3.7 EXERCISES

3.1.1 Derive the formulas (3.1.4), (3.1.5), and (3.1.6).

3.1.2 Check that the inner product defined by (3.1.8) satisfies the axioms.

3.1.3 Show that the covariance function over variable-space defined in Section 1.3 satisfies the axioms for an inner product function.

3.1.4 Suppose that $U_1, U_2, \ldots, U_r$ are mutually orthogonal. Show that $\sum_1^r c_iU_i$ has norm $[\sum_1^r c_i^2(U_i, U_i)]^{1/2}$ and that the pair $\sum_1^r c_iU_i$ and $\sum_1^r d_iU_i$ has inner product $\sum_1^r c_id_i(U_i, U_i)$.

3.1.5 Construct a p-variate sample of size n which has the covariance matrix $\mathbf{I}$.

3.1.6 Show that the sum of two inner product functions is always an inner product function, but that the same is not always true of differences.

3.1.7 Show that any set of mutually orthogonal vectors, excluding $\emptyset$, is linearly independent.

3.1.8 Suppose that $\mathscr{U}$ is a subspace of a Euclidean space $\mathscr{E}$. Show that the set of all V orthogonal to every U in $\mathscr{U}$ forms a subspace $\mathscr{V}$ of $\mathscr{E}$. Show that $\emptyset$ is the only element common to both $\mathscr{U}$ and $\mathscr{V}$.

3.1.9 Suppose that $\mathscr{U}$ is a subspace of a Euclidean space $\mathscr{E}$. Show that if V in $\mathscr{E}$ can be written $V = V_1 + V_2$ with V_1 in $\mathscr{U}$ and V_2 orthogonal to $\mathscr{U}$, then this decomposition is unique.

3.2.1 Consider an arbitrary triangle in a two-dimensional vector space $\mathscr{E}$. Show that there is a unique inner product over $\mathscr{E}$ such that the triangle is equilateral with sides of unit length.

3.2.2 Derive the Pythagoras theorem in p-dimensional Euclidean space, i.e., show that the squared length of the hypotenuse of a right-angled triangle is equal to the sum of the squared lengths of the other two sides.

3.2.3 Suppose that $V_1, V_2, \ldots, V_p$ is a basis of an ordinary vector space, and suppose that A is a given point in this space. In Exercise 2.1.4 it was shown that the same set of elements may be regarded as a different vector space with origin A and basis $A + V_1$, $A + V_2, \ldots, A + V_p$. Either of these vector spaces may be made into Euclidean vector spaces by regarding the basis $V_1, V_2, \ldots, V_p$, in one case, or the basis $A + V_1$, $A + V_2, \ldots, A + V_p$, in the other case, as an orthonormal basis. Show that the concepts of length of a line segment and angle between two line segments are the same for these different Euclidean vector spaces, regardless of the choice of A. That is, these concepts may be regarded as belonging to Euclidean geometry proper, where no special origin is singled out.

3.2.4 Suppose that UV and U_1V_1 are parallel line segments in an affine space whose affine ratio of lengths is α. Show that this concept of length is consistent with the definition of length in Euclidean space, i.e., show that for any choice of an inner product over the affine space the ratio $(U - V, U - V)/(U_1 - V_1, U_1 - V_1)$ is α^2.

3.2.5 Give a definition of the angle between a line and a hyperplane of dimension r.

3.2.6 Show that the intersection of an ellipsoid with a hyperplane is either empty or is an ellipsoid in the hyperplane.

3.3.1 Show that the operation of orthogonal projection into a given hyperplane requires only the structure of Euclidean geometry, i.e., is independent of the choice of an origin.

3.3.2 Suppose that V is restricted to lie in a hyperplane $U + \mathscr{U}$ of a Euclidean space $\mathscr{E}$. How should V be chosen to minimize (V, V)?

3.4.1 Show that the sum of two positive definite symmetric matrices is itself positive definite symmetric.

3.4.2 Suppose that $\beta_1 V_1 + \beta_2 V_2 + \cdots + \beta_r V_r$ is the orthogonal projection of $\alpha_1 V_1 + \alpha_2 V_2 + \cdots + \alpha_p V_p$ into the subspace spanned by $V_1, V_2, \ldots, V_r$, where the inner product matrix is $\mathbf{Q}$ relative to $\mathbf{V}$. Write down a set of r linear equations which must be satisfied by $\beta_1, \beta_2, \ldots, \beta_r$.

3.4.3 Show that every $p \times p$ positive definite symmetric matrix has rank p.

3.4.4 Assuming that an orthonormal basis always exists, show that any positive definite symmetric matrix $\mathbf{Q}$ may be represented in the form

$$\mathbf{Q} = \mathbf{D}\mathbf{D}'$$

for some $p \times p$ matrix $\mathbf{D}$ of rank p. Deduce that any pair $\mathbf{Q}, \mathbf{Q}^*$ of positive definite symmetric matrices stand in the relation

$$\mathbf{Q}^* = \mathbf{A}\mathbf{Q}\mathbf{A}'$$

for some $p \times p$ matrix $\mathbf{A}$ of rank p.

3.4.5 Show that if $\mathbf{Q}$ and $\mathbf{A}$ are $p \times p$ rank p matrices where $\mathbf{Q}$ is positive definite symmetric, then $\mathbf{A}\mathbf{Q}\mathbf{A}'$ is positive definite symmetric. (In particular, $\mathbf{A}\mathbf{A}'$ is positive definite symmetric.)

3.4.6 Suppose that $\boldsymbol{\alpha}_0$ is a point on the surface of the ellipsoid defined by (3.4.1). Show that the set of points $\boldsymbol{\alpha}$ on the tangent plane to the ellipsoid at $\boldsymbol{\alpha}_0$ is characterized by the linear equation

$$\boldsymbol{\alpha}\mathbf{Q}\boldsymbol{\alpha}_0' = 1.$$

3.4.7 Suppose that $\mathbf{Q}$ is a positive definite symmetric matrix and $\mathbf{C}_1$ is any $p \times p$ matrix such that $\mathbf{C}_1\mathbf{Q}\mathbf{C}_1' = \mathbf{I}$. Show that $\mathbf{C}$ is a $p \times p$ matrix satisfying $\mathbf{C}\mathbf{Q}\mathbf{C}' = \mathbf{I}$ if and only if $\mathbf{C} = \mathbf{G}\mathbf{C}_1$ for some orthogonal matrix $\mathbf{G}$.

3.4.8 Suppose that $\mathbf{x}$ is a given $1 \times p$ vector and that $\mathbf{Q}$ is a given positive definite symmetric matrix. Show that $\boldsymbol{\alpha}\mathbf{Q}\boldsymbol{\alpha}'$ is minimized over choices of $\boldsymbol{\alpha}$ subject to $\boldsymbol{\alpha}\mathbf{x}' = 1$ by choosing

$$\boldsymbol{\alpha} = d^{-2}\mathbf{x}\mathbf{Q}^{-1},$$

where

$$d^2 = \mathbf{x}\mathbf{Q}^{-1}\mathbf{x}',$$

and that the resulting minimum is d^{-2}.

3.4.9 Suppose that $\mathbf{U}$ and $\mathbf{W}$ are orthonormal bases of Euclidean spaces $\mathscr{E}$ and $\mathscr{F}$, respectively. Suppose that a linear transformation $\mathbf{A}$ from $\mathscr{E}$ to $\mathscr{F}$ is defined by

$$\mathbf{U} \to \mathbf{A}\mathbf{W}$$

and suppose that a linear transformation $\mathbf{A}'$ from $\mathscr{F}$ to $\mathscr{E}$ is defined by

$$\mathbf{W} \to \mathbf{A}'\mathbf{U}.$$

Show that the definition of $\mathbf{A}'$ does not depend on the particular orthonormal bases chosen, and so is a coordinate-free Euclidean concept.

3.4.10 Suppose that $\mathscr{E}$ is a Euclidean space with an inner product denoted by π^*. Suppose that π denotes a new inner product defined over $\mathscr{E}$, and that $\mathbf{Q}$ denotes the inner product matrix of the new inner product relative to a basis $\mathbf{V}$ which is orthonormal relative to the original inner product. Show that the linear transformation

$$\mathbf{V} \to \mathbf{Q}\mathbf{V}$$

is coordinate-free, i.e., does not depend on the particular choice of a π^*-orthonormal $\mathbf{V}$.

3.4.11 The concept of a pseudoinverse $\mathbf{B}$ of a given linear transformation $\mathbf{A}$ from $\mathscr{E}$ to $\mathscr{F}$ was defined in Section 2.6. When $\mathscr{E}$ and $\mathscr{F}$ are Euclidean spaces it is natural to choose one such pseudoinverse and call it *the* pseudoinverse. Recall that $\mathbf{B}$ was defined in terms of subspaces $\mathscr{U}^*$ and $\mathscr{V}^*$ complementary to $\mathscr{U}$ and $\mathscr{V}$ in $\mathscr{E}$ and $\mathscr{F}$, respectively, where $\mathscr{U}$ and $\mathscr{V}$ are determined by $\mathbf{A}$. *The pseudoinverse* $\mathbf{B}$ *of* $\mathbf{A}$ may be defined as the special case where $\mathscr{U}^*$ and $\mathscr{V}^*$ are chosen to be *the* orthogonal complements of $\mathscr{U}$ and $\mathscr{V}$, respectively. Correspondingly, given any $p \times q$ matrix $\mathbf{A}$, the pseudoinverse $\mathbf{B}$ of $\mathbf{A}$ may be defined by asserting that

$$\mathbf{W} \to \mathbf{B}\mathbf{U}$$

is *the* pseudoinverse of

$$\mathbf{U} \to \mathbf{A}\mathbf{W},$$

where $\mathbf{U}$ and $\mathbf{W}$ are orthonormal bases of $\mathscr{E}$ and $\mathscr{F}$, respectively. Show that $\mathbf{B}$ is the pseudoinverse of $\mathbf{A}$ if and only if it can be represented in the form (2.6.10) and (2.6.11) with $\mathbf{G}$ and $\mathbf{H}$ both orthogonal matrices.

3.4.12 Show that $\mathbf{G} = \mathbf{G}' = \mathbf{G}^{-1}$ for an elementary orthogonal matrix $\mathbf{G}$. What is the geometric explanation for this?

3.5.1 Show that the volume of the simplex $\sum_1^r c_i W_i$ where $c_i \geq 0$ for $i = 1, 2, \ldots, r$ and $\sum_1^r c_i \leq 1$ is $1/r!$ times the volume of the parallelotope $\sum_1^r c_i W_i$ where $0 \leq c_i \leq 1$ for $i = 1, 2, \ldots, r$.

3.6.1 Derive the formulas (3.6.2) and (3.6.3).

3.6.2 Make a drawing of the ellipsoid of an inner product when $p = 3$ and $f = 2$.

3.6.3 Show that Equation (3.1.8) may be interpreted by saying that an inner product of rank p may be expressed as a sum of p inner products each of rank 1. Show that the sum of p arbitrary inner products may be an inner product of any rank from 1 to p.

3.6.4 Suppose that $\mathscr{U}$, $\mathscr{V}$, and $\mathbf{V}$ are as defined in the discussion preceding (3.6.7). Show that (3.6.8) defines *the* pseudoinverse of $\mathbf{Q}$ if $\mathscr{V}$ is the orthogonal complement of $\mathscr{U}$ according to the inner product which takes $\mathbf{V}$ to be orthonormal (cf. Exercises 3.4.10 and 3.4.11).

3.6.5 Suppose that $\mathscr{V}_1$ is any subspace of $\mathscr{E}$ such that all non-$\emptyset$ vectors in $\mathscr{V}_1$ have nonzero norm. Show that whatever the rank of the inner product, there is a unique decomposition of any vector V in $\mathscr{E}$ into a component in $\mathscr{V}_1$ and a component orthogonal to $\mathscr{V}_1$. (Assume, of course, the same result for the case of a full rank inner product.)

3.6.6 Show that if (3.6.9) is satisfied for any two different $p \times f$ matrices $\mathbf{D}_1$ and $\mathbf{D}_1^*$, then $\mathbf{D}_1 = \mathbf{D}_1^*\mathbf{G}_{11}$ for some $f \times f$ orthogonal matrix $\mathbf{G}_{11}$. [*Hint:* express the desired

relations first in terms of a basis **U** with inner product matrix $\mathbf{I}_f$ rather than in terms of **V** with inner product matrix **Q**.]

3.6.7 Suppose that **A** denotes a linear transformation of rank r from a p-dimensional vector space $\mathscr{E}$ to a q-dimensional vector space $\mathscr{E}^*$. Show that $(\mathbf{A}U, \mathbf{A}V)^*$ defines a rank r inner product over $\mathscr{E}$ if $(U^*, V^*)^*$ denotes a rank q inner product over $\mathscr{E}^*$. What is the rank of the induced inner product over $\mathscr{E}$ if the given inner product over $\mathscr{E}^*$ has rank $f < q$?

CHAPTER 4

SUCCESSIVE ORTHOGONALIZATION AND RELATED THEORY

4.1 THE PROCESS OF SUCCESSIVE ORTHOGONALIZATION

This chapter defines and explores a tool of great importance for working with Euclidean spaces, both for deriving mathematical theory and for carrying out computations with numerical data. Essentially, it is a process for using an ordered set of vectors $U_1, U_2, \ldots, U_s$ to produce a set of mutually orthogonal vectors.

Consider a set of vectors $U_1, U_2, \ldots, U_s$ in a Euclidean space $\mathscr{E}$ of dimension p. More general theory is outlined in Section 4.4, but here for simplicity it will be assumed that $U_1, U_2, \ldots, U_s$ are linearly independent and that the inner product over $\mathscr{E}$ has full rank p. Often $s = p$ in applications so that $U_1, U_2, \ldots, U_s$ is a basis of $\mathscr{E}$. From $U_1, U_2, \ldots, U_s$ *define in succession the set of vectors* $U_1^*, U_2^*, \ldots, U_s^*$ *as follows*:

$$
\begin{aligned}
U_1^* &= U_1 \\
U_2^* &= U_2 - \frac{(U_2, U_1^*)}{(U_1^*, U_1^*)} U_1^* \\
&\cdots \\
U_r^* &= U_r - \sum_{j=1}^{r-1} \frac{(U_r, U_j^*)}{(U_j^*, U_j^*)} U_j^* \\
&\cdots \\
U_s^* &= U_s - \sum_{j=1}^{s-1} \frac{(U_s, U_j^*)}{(U_j^*, U_j^*)} U_j^*.
\end{aligned}
\tag{4.1.1}
$$

This process of producing $U_1^*, U_2^*, \ldots, U_s^*$ from the ordered set $U_1, U_2, \ldots, U_s$ will be called the *process of successive orthogonalization.* To justify this terminology, it will be proved in the next paragraph that $U_1^*, U_2^*, \ldots, U_s^*$ are mutually

orthogonal. In passing, note that the process of successive orthogonalization of $U_1, U_2, \ldots, U_s$ includes the process of successive orthogonalization of $U_1, U_2, \ldots, U_t$ for $t = 1, 2, \ldots, s$ and that the U_j^* produced by these processes, for a given j, are identical.

An inductive proof that $U_1^*, U_2^*, \ldots, U_s^*$ are mutually orthogonal may be based on the following useful theorem: *if* $U_1^*, U_2^*, \ldots, U_{r-1}^*$ *form any set of* $r - 1$ *mutually orthogonal vectors,* V *is any vector, and*

$$\dot{V} = \sum_{j=1}^{r-1} \frac{(V, U_j^*)}{(U_j^*, U_j^*)} U_j^*, \tag{4.1.2}$$

then $V - \dot{V}$ *is orthogonal to each of* $U_1^*, U_2^*, \ldots, U_{r-1}^*$ *and therefore to the subspace spanned by* $U_1^*, U_2^*, \ldots, U_{r-1}^*$.

The theorem follows from the direct calculation that

$$\begin{aligned}(V - \dot{V}, U_i^*) &= (V, U_i^*) - (\dot{V}, U_i^*) \\ &= (V, U_i^*) - \sum_{j=1}^{r-1} \frac{(V, U_j^*)}{(U_j^*, U_j^*)} (U_j^*, U_i^*) \\ &= (V, U_i^*) - \frac{(V, U_i^*)}{(U_i^*, U_i^*)} (U_i^*, U_i^*) \\ &= 0,\end{aligned} \tag{4.1.3}$$

for $i = 1, 2, \ldots, r - 1$. The theorem may be applied $s - 1$ times in succession where V is chosen to be $U_2, U_3, \ldots, U_s$ and $r - 1$ is chosen to be $1, 2, \ldots, s - 1$, respectively. The first application of the theorem shows that U_2^* is orthogonal to U_1^*. The second application shows that, assuming U_2^* and U_1^* to be orthogonal, U_3^* is orthogonal to U_1^* and U_2^*. In general, the $(r - 1)$st application shows that, assuming $U_1^*, U_2^*, \ldots, U_{r-1}^*$ to be mutually orthogonal, $U_1^*, U_2^*, \ldots, U_r^*$ are mutually orthogonal. Thus, from $s - 1$ applications, it follows that $U_1^*, U_2^*, \ldots, U_s^*$ are mutually orthogonal, as required.

An important detail was neglected in the preceding. If $(U_r^*, U_r^*) = 0$ for some r, then the definition (4.1.1) breaks down for $t \geq r$. Actually, if this should occur, then the terms involving such U_r^* may simply be omitted from the right hand side of (4.1.1) to provide a modified and foolproof definition of $U_1^*, U_2^*, \ldots, U_s^*$. However, this difficulty does not arise when the inner product has full rank and $U_1, U_2, \ldots, U_s$ are linearly independent. For under these conditions $(U_r^*, U_r^*) = 0$ implies that $U_r^* = \emptyset$, and this together with the modified (4.1.1) would mean that $U_1, U_2, \ldots, U_s$ could be expressed in terms of fewer than s nonzero U_j^* which would imply linear dependence among $U_1, U_2, \ldots, U_s$.

The process of successive orthogonalization provides proofs for several assertions left unproved in Chapter 3. First, it shows that every Euclidean space possesses an orthogonal basis, for the $U_1^*, U_2^*, \ldots, U_p^*$ constructed as above

from any basis $U_1, U_2, \ldots, U_p$ constitute an orthogonal basis. Further, one may define

$$U_j^{**} = [(U_j^*, U_j^*)^{-1/2}] U_j^* \tag{4.1.4}$$

for $j = 1, 2, \ldots, p$ so that $U_1^{**}, U_2^{**}, \ldots, U_p^{**}$ constitute an orthonormal basis. Finally, in connection with the definition of an orthogonal projection in Section 3.3 it is required to show that the subspace $\mathscr{V}$ of vectors orthogonal to a given subspace $\mathscr{U}$ is complementary to $\mathscr{U}$, i.e., to show that the notion of orthogonal complement is properly defined. To prove this, suppose that $\mathscr{U}$ has dimension s, that $U_1, U_2, \ldots, U_s$ is a basis of $\mathscr{U}$, and that $U_1, U_2, \ldots, U_p$ is a basis of the whole space $\mathscr{E}$ which includes the basis $U_1, U_2, \ldots, U_s$ of $\mathscr{U}$. From these, construct the orthogonal basis $U_1^*, U_2^*, \ldots, U_p^*$ of $\mathscr{E}$ as in (4.1.1). It is clear from the construction that $U_1^*, U_2^*, \ldots, U_s^*$ is an orthogonal basis of $\mathscr{U}$. Also $U_{s+1}^*, U_{s+2}^*, \ldots, U_p^*$ clearly span a $(p - s)$-dimensional space both orthogonal to and complementary to $\mathscr{U}$. It remains only to show that any V orthogonal to $\mathscr{U}$ lies in the subspace spanned by $U_{s+1}^*, U_{s+2}^*, \ldots, U_p^*$. Certainly V may be written $V = \alpha_1^* U_1^* + \alpha_2^* U_2^* + \cdots + \alpha_p^* U_p^*$. But, taking inner products of each side with U_i^* yields $\alpha_i^* = 0$ for $i = 1, 2, \ldots, s$, as required.

The operation of orthogonal projection into a subspace $\mathscr{U}$ is especially easy to carry out when an orthogonal basis of $\mathscr{U}$ is available. For, suppose that $U_1^*, U_2^*, \ldots, U_s^*$ is an orthogonal basis of $\mathscr{U}$ and that $U_{s+1}^*, U_{s+2}^*, \ldots, U_p^*$ is an orthogonal basis of the orthogonal complement $\mathscr{V}$ of $\mathscr{U}$. Then any vector W may be written

$$W = \sum_{i=1}^{p} \frac{(W, U_i^*)}{(U_i^*, U_i^*)} U_i^* = \dot{W} + W^*, \tag{4.1.5}$$

where

$$\dot{W} = \sum_{i=1}^{s} \frac{(W, U_i^*)}{(U_i^*, U_i^*)} U_i^* \tag{4.1.6}$$

is the orthogonal projection of W into $\mathscr{U}$ or the component of W along $\mathscr{U}$, and

$$W^* = \sum_{i=s+1}^{p} \frac{(W, U_i^*)}{(U_i^*, U_i^*)} U_i^* \tag{4.1.7}$$

is the component of W orthogonal to $\mathscr{U}$. The proofs of these assertions are left to the reader. Note that the coordinates of any vector relative to an orthogonal basis are very simply expressed in terms of inner products, and that each coordinate multiplied by its associated basis vector simply gives the orthogonal projection of W into the one-dimensional subspace spanned by that basis vector.

The stepwise construction (4.1.1) of successive orthogonalization is a very important computational device, and is explored further from the computational viewpoint in Section 4.3. For pure mathematics, the ability to construct an orthonormal basis for an arbitrary inner product has led to the derivation of important theory such as the theorem: *given any inner product defined on a*

vector space $\mathscr{E}$ there exists a basis such that the inner product is defined by formula (3.1.8) *in terms of that basis, and furthermore such a basis can be constructed from an arbitrary basis in a finite number of steps.*

The theorem on the existence of an orthonormal basis corresponds to the geometric theorem that *every ellipsoid has a set of p conjugate axes.* In matrix algebra, this existence theorem corresponds to the following famous result: *if* $\mathbf{Q}$ *is any positive definite symmetric matrix, then there exists a nonsingular* $p \times p$ *matrix* $\mathbf{C}$ *such that*

$$\mathbf{CQC}' = \mathbf{I}. \tag{4.1.8}$$

To see this, suppose that $\mathbf{Q}$ is the inner product matrix of a basis $\mathbf{V}$. Then there exists a basis $\mathbf{U} = \mathbf{CV}$ with inner product matrix $\mathbf{I}$ and (4.1.8) is simply an application of (3.4.4).

It should be noted that the result of successive orthogonalization depends in general on the order in which the vectors are presented. Thus, with p vectors $U_1, U_2, \ldots, U_p$ there are in general $p!$ different orders of possible successive orthogonalization. Although the theory is always presented using a given order, the reader should keep in mind that in applications an order different from an arbitrary given order may often be appropriate.

4.2 SOME MATRIX THEORY RELATED TO SUCCESSIVE ORTHOGONALIZATION

Suppose that $\mathbf{Q}$ is a $p \times p$ positive definite symmetric matrix. By the device of regarding $\mathbf{Q}$ to be the inner product matrix relative to some basis of a vector space $\mathscr{E}$ and then applying successive orthogonalization to this basis, one is led to a wealth of identities all based on the elements of $\mathbf{Q}$. To begin, suppose that

$$\mathbf{U} = \begin{bmatrix} U_1 \\ U_2 \\ \cdot \\ \cdot \\ \cdot \\ U_p \end{bmatrix} \quad \text{and} \quad \mathbf{U}^* = \begin{bmatrix} U_1^* \\ U_2^* \\ \cdot \\ \cdot \\ \cdot \\ U_p^* \end{bmatrix} \tag{4.2.1}$$

denote the basis of $\mathscr{E}$ and its orthogonalization defined by (4.1.1) with $r = p$. In matrix notation (4.1.1) may be written

$$\mathbf{U} = \mathbf{BU}^*, \tag{4.2.2}$$

where $\mathbf{B}$ is a $p \times p$ triangular matrix with elements zero above the diagonal and unity along the diagonal. It is clear from (4.1.1) that $\mathbf{B}$ is uniquely determined by the inner product over $\mathscr{E}$ and therefore by $\mathbf{Q}$.

For simplicity of notation, the inverse matrix of $\mathbf{B}$ will be denoted by $\mathbf{A}$, so that

$$\mathbf{U}^* = \mathbf{AU}. \tag{4.2.3}$$

$\mathbf{A}$ may be found from $\mathbf{B}$ by a process of successive substitution to solve the equations (4.1.1) or (4.2.2) for $\mathbf{U}^*$ in terms of $\mathbf{U}$, i.e., the first row of (4.1.1) yields the first row of (4.2.3), the second row of (4.2.3) is found by substituting the first row of (4.2.3) into the second row of (4.1.1), the third row of (4.2.3) is found by substituting the already found first two rows of (4.2.3) into the third row of (4.1.1), and so on. It is clear from this process of construction that $\mathbf{A}$ like $\mathbf{B}$ is a triangular matrix with elements zero above the diagonal and unity along the diagonal.

A matrix $\mathbf{T}$ may be defined as the inner product matrix of $\mathbf{U}^*$, where, of course, $\mathbf{T}$ is a diagonal matrix with diagonal elements (U_s^*, U_s^*) in the (s, s) position for $s = 1, 2, \ldots, p$. It follows from (4.2.2) that

$$\mathbf{Q} = \mathbf{BTB}'. \tag{4.2.4}$$

The inverse of $\mathbf{Q}$ will be denoted by $\mathbf{P}$, and, by inverting both sides of (4.2.4) and replacing $\mathbf{B}^{-1}$ by $\mathbf{A}$, it follows that

$$\mathbf{P} = \mathbf{A}'\mathbf{T}^{-1}\mathbf{A}. \tag{4.2.5}$$

The above formulas also yield $\det \mathbf{Q}$ as a by-product. For, since $\mathbf{B}$ is triangular with diagonal elements unity, $\det \mathbf{B} = 1$, and hence

$$\begin{aligned} \det \mathbf{Q} &= \det \mathbf{BTB}' \\ &= \det \mathbf{B} \det \mathbf{T} \det \mathbf{B}' \\ &= \det \mathbf{T} \\ &= \prod_{s=1}^{p} (U_s^*, U_s^*). \end{aligned} \tag{4.2.6}$$

As discussed in Section 3.5, $\det \mathbf{Q}$ may be interpreted as the squared volume of the parallelotope generated by the basis $U_1, U_2, \ldots, U_p$. Formula (4.2.6) has special interest in this regard as the generalization to p dimensions of the formula "Area = base × height" for a parallelogram. In fact, the product of the first $p - 1$ terms of (4.2.6) simply gives the square of the $(p - 1)$-dimensional volume of the $(p - 1)$-dimensional parallelotope based on $U_1, U_2, \ldots, U_{p-1}$. This corresponds to the "base." The "height" is given by the component of U_p orthogonal to $U_1, U_2, \ldots, U_{p-1}$, and the squared length of this component is given by the last term of (4.2.6). Thus by repeated application of the "base × height" formula to the parallelotopes spanned by $U_1, U_2, \ldots, U_s$ for $s = 2, 3, \ldots, p$ one achieves an illuminating explanation of formula (4.2.6).

The remainder of this section is devoted to a set of relationships concerning the first s and last $p - s$ elements of $\mathbf{U}$ and $\mathbf{U}^*$. These will be of use in understanding the computational aspects of successive orthogonalization to be discussed in Section 4.3. For any $p \times 1$ column vector $\mathbf{G}$, denote the first s rows by $\mathbf{G}_1$ and the last $p - s$ rows by $\mathbf{G}_2$, so that

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \end{bmatrix}. \tag{4.2.7}$$

$p \times 1$

Similarly, for a $p \times p$ matrix $\mathbf{K}$, write

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_{11} & \mathbf{K}_{12} \\ \mathbf{K}_{21} & \mathbf{K}_{22} \end{bmatrix}, \tag{4.2.8}$$

where $\mathbf{K}_{11}$, $\mathbf{K}_{12}$, $\mathbf{K}_{21}$, and $\mathbf{K}_{22}$ are $s \times s$, $s \times (p-s)$, $(p-s) \times s$, and $(p-s) \times (p-s)$ matrices, respectively. This notation will be applied in particular to the $p \times 1$ vectors $\mathbf{U}$ and $\mathbf{U}^*$, and to the $p \times p$ matrices $\mathbf{B}$, $\mathbf{A}$, $\mathbf{Q}$, $\mathbf{P}$, and $\mathbf{T}$. Note that $\mathbf{A}_{12}$ and $\mathbf{B}_{12}$ consist entirely of zeros, and that $\mathbf{A}_{11}$, $\mathbf{A}_{22}$, $\mathbf{B}_{11}$, and $\mathbf{B}_{22}$ are triangular like $\mathbf{A}$ and $\mathbf{B}$. In addition, the notation $\mathbf{U}_{2.1}$ will be used for the vector of components of $\mathbf{U}_2$ orthogonal to $\mathbf{U}_1$, i.e.,

$$\mathbf{U}_{2.1} = \begin{bmatrix} U_{s+1.12\ldots s} \\ U_{s+2.12\ldots s} \\ \cdot \\ \cdot \\ \cdot \\ U_{p.12\ldots s} \end{bmatrix}, \tag{4.2.9}$$

where $U_{r.12\ldots s}$ denotes the component of U_r orthogonal to the subspace spanned by $U_1, U_2, \ldots, U_s$. Finally, the inner product matrix of $\mathbf{U}_{2.1}$ will be denoted by $\mathbf{Q}_{22.1}$. Observe that the notation of this paragraph assumes a fixed p and s.

One may now write (4.2.2) as

$$\mathbf{U}_1 = \mathbf{B}_{11}\mathbf{U}_1^* \quad \text{and} \quad \mathbf{U}_2 = \mathbf{B}_{21}\mathbf{U}_1^* + \mathbf{B}_{22}\mathbf{U}_2^*, \tag{4.2.10}$$

and (4.2.3) as

$$\mathbf{U}^* = \mathbf{A}_{11}\mathbf{U}_1 \quad \text{and} \quad \mathbf{U}_2^* = \mathbf{A}_{21}\mathbf{U}_1 + \mathbf{A}_{22}\mathbf{U}_2. \tag{4.2.11}$$

By substituting (4.2.10) into (4.2.11), one may express the submatrices of $\mathbf{A}$ in terms of the submatrices of $\mathbf{B}$ as

$$\begin{aligned} \mathbf{A}_{11} &= \mathbf{B}_{11}^{-1}, \\ \mathbf{A}_{21} &= -\mathbf{B}_{22}^{-1}\mathbf{B}_{21}\mathbf{B}_{11}^{-1}, \quad \text{and} \\ \mathbf{A}_{22} &= \mathbf{B}_{22}^{-1}. \end{aligned} \tag{4.2.12}$$

Similarly, one has

$$\begin{aligned} \mathbf{B}_{11} &= \mathbf{A}_{11}^{-1}, \\ \mathbf{B}_{21} &= -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1}, \quad \text{and} \\ \mathbf{B}_{22} &= \mathbf{A}_{22}^{-1}. \end{aligned} \tag{4.2.13}$$

From (4.2.12)

$$\mathbf{A}_{21}\mathbf{A}_{11}^{-1} = -\mathbf{B}_{22}^{-1}\mathbf{B}_{21} = \mathbf{A}_{21}\mathbf{B}_{11} = -\mathbf{A}_{22}\mathbf{B}_{21}, \tag{4.2.14}$$

and from (4.2.13)

$$\mathbf{B}_{21}\mathbf{B}_{11}^{-1} = -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} = \mathbf{B}_{21}\mathbf{A}_{11} = -\mathbf{B}_{22}\mathbf{A}_{21}. \tag{4.2.15}$$

It will be convenient in the sequel to use the special notation $\mathbf{H}_{21}$ *for the matrix expressed by* (4.2.15). Two special properties of $\mathbf{H}_{21}$ are:

i) *the first row of* $\mathbf{H}_{21}$ *is the negative of the first row of* $\mathbf{A}_{21}$

ii) *the last column of* $\mathbf{H}_{21}$ *is the last column of* $\mathbf{B}_{21}$.

These follow easily from the relations $\mathbf{H}_{21} = -\mathbf{B}_{22}\mathbf{A}_{21}$ and $\mathbf{H}_{21} = \mathbf{B}_{21}\mathbf{A}_{11}$, respectively. Property (i) may be generalized by noting from (4.2.10) that

$$\mathbf{U}_{2.1} = \mathbf{B}_{22}\mathbf{U}_2^*, \tag{4.2.16}$$

and thence from (4.2.16) and (4.2.11) that

$$\mathbf{U}_{2.1} = \mathbf{U}_2 - \mathbf{H}_{21}\mathbf{U}_1. \tag{4.2.17}$$

Thus the rows of $\mathbf{H}_{21}$ *determine the linear combinations of* $U_1, U_2, \ldots, U_s$ *which are the components of* $U_{s+1}, U_{s+2}, \ldots, U_p$ *along the subspace spanned by* $U_1, U_2, \ldots, U_s$.

It is of interest to express $\mathbf{Q}_{11}$, $\mathbf{Q}_{21}$, $\mathbf{Q}_{22}$, and $\mathbf{Q}_{22.1}$ in terms of $\mathbf{B}$ and $\mathbf{T}$. From (4.2.10) it follows that

$$\begin{aligned} \mathbf{Q}_{11} &= \mathbf{B}_{11}\mathbf{T}_{11}\mathbf{B}_{11}', \\ \mathbf{Q}_{21} &= \mathbf{B}_{21}\mathbf{T}_{11}\mathbf{B}_{11}', \quad \text{and} \\ \mathbf{Q}_{22} &= \mathbf{B}_{21}\mathbf{T}_{11}\mathbf{B}_{21}' + \mathbf{B}_{22}\mathbf{T}_{22}\mathbf{B}_{22}'. \end{aligned} \tag{4.2.18}$$

Also, from (4.2.16) it follows that the inner product matrix $\mathbf{Q}_{22.1}$ of $\mathbf{U}_{2.1}$ is given by

$$\mathbf{Q}_{22.1} = \mathbf{B}_{22}\mathbf{T}_{22}\mathbf{B}_{22}'. \tag{4.2.19}$$

From (4.2.19) it is clear that the first and second terms in the expression for $\mathbf{Q}_{22}$ in (4.2.18) represent the contributions to the inner product matrix of $\mathbf{U}_2$ from the components of $\mathbf{U}_2$ along the subspace spanned by $\mathbf{U}_1$, and from the components of $\mathbf{U}_2$ orthogonal to the subspace spanned by $\mathbf{U}_1$, respectively.

Matrix formulas are often seen which contain the products $\mathbf{Q}_{21}\mathbf{Q}_{11}^{-1}$ or $\mathbf{Q}_{21}\mathbf{Q}_{11}^{-1}\mathbf{Q}_{12}$. It will be useful therefore to note the following alternative expressions:

$$\mathbf{Q}_{21}\mathbf{Q}_{11}^{-1} = \mathbf{B}_{21}\mathbf{B}_{11}^{-1} = \mathbf{H}_{21}, \tag{4.2.20}$$

and

$$\mathbf{Q}_{21}\mathbf{Q}_{11}^{-1}\mathbf{Q}_{12} = \mathbf{H}_{21}\mathbf{Q}_{12} = \mathbf{B}_{21}\mathbf{T}_{11}\mathbf{B}_{21}' = \mathbf{Q}_{22} - \mathbf{Q}_{22.1}. \tag{4.2.21}$$

The reader may have noticed a duality between the formulas (4.2.4) and (4.2.5) and between the sets of formulas (4.2.12) and (4.2.13). The nature of this duality will be elaborated in Chapter 6, but it will be useful to give dual formulas here for (4.2.18) and (4.2.19). From (4.2.5), rewritten as

$$\begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}' \begin{bmatrix} \mathbf{T}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix},$$

it follows that

$$\mathbf{P}_{11} = \mathbf{A}_{21}'\mathbf{T}_{22}^{-1}\mathbf{A}_{21} + \mathbf{A}_{11}'\mathbf{T}_{11}^{-1}\mathbf{A}_{11},$$
$$\mathbf{P}_{12} = \mathbf{A}_{21}'\mathbf{T}_{22}^{-1}\mathbf{A}_{22}, \quad \text{and} \tag{4.2.22}$$
$$\mathbf{P}_{22} = \mathbf{A}_{22}'\mathbf{T}_{22}^{-1}\mathbf{A}_{22}.$$

The right side of the last line of (4.2.22) is now seen to be the inverse of the right side of (4.2.19), so that

$$\mathbf{Q}_{22.1} = \mathbf{P}_{22}^{-1}. \tag{4.2.23}$$

Dually, then, one might define

$$\mathbf{P}_{11.2} = \mathbf{Q}_{11}^{-1}, \tag{4.2.24}$$

and by inverting the first line of (4.2.18) find the dual formula of (4.2.19) to be

$$\mathbf{P}_{11.2} = \mathbf{A}_{11}'\mathbf{T}_{11}^{-1}\mathbf{A}_{11}. \tag{4.2.25}$$

Note that the duality here interchanges the roles of **B**, **Q**, and **T** with their transposed inverses, and also replaces the subscripts 1 and 2 with 2 and 1, respectively.

The formulas of this section all have analogues where the orthonormal basis **U**** is used in place of **U***. Here

$$\mathbf{U}^* = \mathbf{R}\mathbf{U}^{**}, \tag{4.2.26}$$

where **R** is the $p \times p$ diagonal matrix with diagonal elements $(U_s^*, U_s^*)^{1/2}$. Consequently

$$\mathbf{U} = \mathbf{D}\mathbf{U}^{**} \tag{4.2.27}$$

and

$$\mathbf{U}^{**} = \mathbf{C}\mathbf{U}, \tag{4.2.28}$$

where

$$\mathbf{C} = \mathbf{D}^{-1}, \quad \mathbf{D} = \mathbf{B}\mathbf{R}, \quad \text{and} \quad \mathbf{C} = \mathbf{R}^{-1}\mathbf{A}. \tag{4.2.29}$$

In words, this formula says that **D** is found by multiplying each column of **B** by the corresponding element of **R**, and **C** is found by multiplying each row of **A** by the corresponding element of $\mathbf{R}^{-1}$. Thus, **D** and **C** are triangular like **B** and **A**, but with the diagonal elements of **R** and $\mathbf{R}^{-1}$, respectively. In place of (4.2.4) and (4.2.5), one now has

$$\mathbf{Q} = \mathbf{D}\mathbf{D}' \tag{4.2.30}$$

and

$$\mathbf{P} = \mathbf{C}'\mathbf{C}, \tag{4.2.31}$$

and, in place of (4.2.10) and (4.2.11), one now has

$$\mathbf{U}_1 = \mathbf{D}_{11}\mathbf{U}_1^{**} \quad \text{and} \quad \mathbf{U}_2 = \mathbf{D}_{21}\mathbf{U}_1^{**} + \mathbf{D}_{22}\mathbf{U}_2^{**}, \tag{4.2.32}$$

and also

$$\mathbf{U}_1^{**} = \mathbf{C}_{11}\mathbf{U}_1 \quad \text{and} \quad \mathbf{U}_2^{**} = \mathbf{C}_{21}\mathbf{U}_1 + \mathbf{C}_{22}\mathbf{U}_2. \tag{4.2.33}$$

The analogues of formulas (4.2.12), (4.2.13), (4.2.14), (4.2.15), (4.2.18), (4.2.19), (4.2.22), and (4.2.25) are found by replacing **B** with **D**, **A** with **C**, and **T** with **I**. These analogues are left for the reader to write down and check.

4.3 COMPUTATIONAL METHODS RELATED TO SUCCESSIVE ORTHOGONALIZATION

Positive definite or semi-definite symmetric matrices may be computed from multivariate data in many ways, usually in the context of covariance matrices or their inverses. These ways will, of course, be discussed in detail in later chapters. Certain much-used further computations with such matrices are closely related to the process of successive orthogonalization. Beginning from a given **Q**, the computations proceed by stages to produce the matrices which appear in the various identities of Section 4.2. Many of these matrices have important statistical interpretations and it is therefore important to have efficient computing methods suitable for an electronic computer. Besides, these computing methods often have a simplicity and elegance which is not obvious from the definitions and identities of Section 4.2.

Q will denote throughout Section 4.3 a $p \times p$ symmetric positive definite matrix of full rank p.

4.3.1. Elimination procedures. Following the rows of (4.1.1) in order to find $U_2^*, U_3^*, \ldots, U_p^*$ corresponds to computing the rows of **B** in order. Such a direct approach is possible and is presented in Section 4.3.3, but a more convenient and widely used scheme computes the columns of **B** in order. The latter scheme and various natural extensions of it are the subject of Section 4.3.1.

The idea behind elimination procedures is a set of $p - 1$ stages where at stage s all of the terms in U_s^* on the right side of (4.1.1) are subtracted out at once, for $s = 1, 2, \ldots, p - 1$. After s such stages, one has in hand $U_1^*, U_2^*, \ldots, U_s^*$ in the first s rows of (4.1.1), $U_{s+1}^* = U_{s+1.12\ldots s}$ in row $s + 1$, and $U_{r.12\ldots s}$ in row r for $r = s + 2, s + 3, \ldots, p$. In geometric language, the procedure uses orthogonal projection to *eliminate* at stage s the components of $U_{s+1}, U_{s+2}, \ldots, U_p$ along U_s^* so that after s stages only the components of each of $U_{s+1}, U_{s+2}, \ldots, U_p$ perpendicular to all of $U_1, U_2, \ldots, U_s$ remain.

The computations will be described in three layers whose complexity increases as more quantities are carried along in the calculations. In the first layer, the main objective is simply to calculate **B** from **Q** by producing the columns of **B** one at a time. It turns out to be convenient to compute simultaneously the inner product matrices $\mathbf{Q}_{22.1}$ of $U_{s+1.12\ldots s}$ $(= U_{s+1}^*)$, $U_{s+2.12\ldots s}$, $\ldots, U_{p.12\ldots s}$, for $s = 1, 2, \ldots, p - 1$. Thus, after s stages of the first layer, one has in hand the first s columns of **B**, namely $\mathbf{B}_{11}$ and $\mathbf{B}_{21}$, together with $\mathbf{Q}_{22.1}$.

Consider now how to carry out the computations of stage $s + 1$ of the first layer while having in hand the output of stage s. Suppose that b_{ij} denotes the

(i,j) element of $\mathbf{B}$ and that the elements of $\mathbf{Q}_{22.1}$ are denoted by $q_{ij.12\ldots s}$ for i and $j = s+1, s+2, \ldots, p$. Suppose also that the elements of the analogue of $\mathbf{Q}_{22.1}$ with s raised to $s+1$ are denoted by $q_{ij.12\ldots\overline{s+1}}$ for i and $j = s+2, s+3, \ldots, p$.

From (4.1.1)

$$b_{i\,\overline{s+1}} = \frac{(U_i, U^*_{s+1})}{(U^*_{s+1}, U^*_{s+1})} = \frac{(U_{i.12\ldots s}, U^*_{s+1})}{(U^*_{s+1}, U^*_{s+1})} = \frac{q_{i\,\overline{s+1}.12\ldots s}}{q_{\overline{s+1}\,\overline{s+1}.12\ldots s}} \tag{4.3.1}$$

for $i = s+2, s+3, \ldots, p$. This provides the nontrivial elements of column $s+1$ of $\mathbf{B}$. The justification of the first step in (4.3.1) is that $U_i - U_{i.12\ldots s}$ consists of components along $U^*_1, U^*_2, \ldots, U^*_s$ which are orthogonal to U^*_{s+1} and therefore $(U_i - U_{i.12\ldots s}, U^*_{s+1}) = 0$ or $(U_i, U^*_{s+1}) = (U_{i.12\ldots s}, U^*_{s+1})$, as required. Similarly

$$\begin{aligned} q_{ij.12\ldots\overline{s+1}} &= (U_{i.12\ldots\overline{s+1}}, U_{j.12\ldots\overline{s+1}}) \\ &= (U_{i.12\ldots s} - b_{i\,\overline{s+1}}U^*_{s+1}, U_{j.12\ldots s} - b_{j\,\overline{s+1}}U^*_{s+1}) \\ &= q_{ij.12\ldots s} - b_{i\,\overline{s+1}}q_{j\,\overline{s+1}.12\ldots s} - b_{j\,\overline{s+1}}q_{i\,\overline{s+1}.12\ldots s} \\ &\qquad + b_{i\,\overline{s+1}}b_{j\,\overline{s+1}}q_{\overline{s+1}\,\overline{s+1}.12\ldots s}. \end{aligned} \tag{4.3.2}$$

From (4.3.1), the last three terms here are equal apart from sign, and thus

$$q_{ij.12\ldots\overline{s+1}} = q_{ij.12\ldots s} - \frac{q_{i\,\overline{s+1}.12\ldots s}q_{j\,\overline{s+1}.12\ldots s}}{q_{\overline{s+1}\,\overline{s+1}.12\ldots s}} \tag{4.3.3}$$

or

$$q_{ij.12\ldots\overline{s+1}} = q_{ij.12\ldots s} - b_{i\,\overline{s+1}}q_{j\,\overline{s+1}.12\ldots s}, \tag{4.3.4}$$

for i and $j = s+2, s+3, \ldots, p$. Formula (4.3.3) shows how to upstage $\mathbf{Q}_{22.1}$ from s to $s+1$, but (4.3.4) is computationally more convenient assuming that column $s+1$ of $\mathbf{B}$ has already been computed from (4.3.1). The calculation indicated by (4.3.3) is often called *pivotal condensation*. This completes the discussion of stage $s+1$ of the first layer of the elimination procedure.

The second layer has the basic objective of computing $\mathbf{A} = \mathbf{B}^{-1}$ along with $\mathbf{B}$. Recall that $\mathbf{A}$ may be found in stages by solving (4.2.2) to produce (4.2.3) where, at stage s, rows $1, 2, \ldots, s$ of (4.2.3) are substituted into row $s+1$ of (4.2.2) to produce row $s+1$ of $\mathbf{A}$. Under the elimination approach, it is more natural at stage s to substitute row s of (4.2.3) into rows $s+1, s+2, \ldots, p$ of (4.2.2), i.e., to *eliminate* U^*_s from the right side of (4.2.2). After s stages of this elimination approach one has produced the first s rows of (4.2.3) from the first s rows of (4.2.2), i.e., one has the first s rows $\mathbf{A}_{11}$ of $\mathbf{A}$. Also, one has modified the last $p-s$ rows of (4.2.2) into expressions for $U_{s+1}, U_{s+2}, \ldots, U_p$ in terms of $U_1, U_2, \ldots, U_s$ and $U^*_{s+1}, U^*_{s+2}, \ldots, U^*_p$, i.e., for $\mathbf{U}_2$ in terms of $\mathbf{U}_1$ and

$\mathbf{U}_{2.1}$. From (4.2.17) the coefficients of $\mathbf{U}_1$ here form $\mathbf{H}_{21}$. Thus, after s stages of the second layer of the elimination method, one may expect to have $\mathbf{A}_{11}$ and $\mathbf{H}_{21}$ together with the output $\mathbf{B}_{11}$, $\mathbf{B}_{21}$, and $\mathbf{Q}_{22.1}$ of the first layer.

Consider now how to upstage $\mathbf{A}_{11}$ and $\mathbf{H}_{21}$ from s to $s+1$. Finding row $s+1$ of $\mathbf{A}$ is trivial, for, as pointed out in Section 4.2 as property (i) of $\mathbf{H}_{21}$, the first s elements of row $s+1$ of $\mathbf{A}$ are simply the negatives of the first row of $\mathbf{H}_{21}$ while the remaining $p-s$ elements are simply a one followed by zeros. It remains therefore to consider only $\mathbf{H}_{21}$. Suppose that the elements of $\mathbf{H}_{21}$ are denoted by h_{ij} for $i = s+1, s+2, \ldots, p$ and $j = 1, 2, \ldots, s$, and the elements after upstaging s to $s+1$ are denoted by h^+_{ij} for $i = s+2, s+3, \ldots, p$ and $j = 1, 2, \ldots, s+1$. Now the stage $s+1$ elimination procedure alters the expression

$$U_i = \sum_{j=1}^{s} h_{ij}U_j + \sum_{t=s+1}^{i} b_{it}U^*_t \tag{4.3.5}$$

into

$$U_i = \sum_{j=1}^{s+1} h^+_{ij}U_j + \sum_{t=s+2}^{i} b_{it}U^*_t \tag{4.3.6}$$

for $i = s+2, s+3, \ldots, p$, by substituting

$$U^*_{s+1} = U_{s+1} + \sum_{j=1}^{s} a_{\overline{s+1}\,j}U_j = U_{\overline{s+1}} - \sum_{j=1}^{s} h_{\overline{s+1}\,j}U_j. \tag{4.3.7}$$

Carrying out the substitution and comparing coefficients with (4.3.6) one finds

$$h^+_{ij} = h_{ij} - b_{i\,\overline{s+1}}h_{\overline{s+1}\,j} \tag{4.3.8}$$

for $i = s+2, s+3, \ldots, p$ and $j = 1, 2, \ldots, s$. An alternative to (4.3.8) derived from (4.3.1) is

$$h^+_{ij} = h_{ij} - \frac{q_{i\,\overline{s+1}.12\ldots s}h_{\overline{s+1}\,j}}{q_{\overline{s+1}\,\overline{s+1}.12\ldots s}}. \tag{4.3.9}$$

Formulas (4.3.8) and (4.3.9) are analogous to (4.3.4) and (4.3.3), respectively; the first is computationally more direct, assuming that $b_{i\,\overline{s+1}}$ has already been computed. This completes the discussion of stage $s+1$ of the second layer of the elimination procedure.

The third layer when added to the first two brings $\mathbf{Q}^{-1}$ $(= \mathbf{P})$ into the system by finding after s stages $\mathbf{Q}^{-1}_{11}$ $(= \mathbf{P}_{11.2})$ and upstaging it from s to $s+1$. Denote the elements of $\mathbf{Q}^{-1}_{11}$ by $p_{ij.\overline{s+1}\,\overline{s+2}\ldots p}$ for i and $j = 1, 2, \ldots, s$ and the upstaged elements by $p_{ij.\overline{s+2}\,\overline{s+3}\ldots p}$ for i and $j = 1, 2, \ldots, s+1$. Now $\mathbf{Q}^{-1}_{11}$ could have been calculated from (4.2.25) and the upstaged $\mathbf{Q}^{-1}_{11}$ could be calculated from the upstaged (4.2.25). Note, however, that most of the labor in calculating the upstaged (4.2.25) was involved in calculating the original (4.2.25). Indeed, by

careful comparison of these two versions of (4.2.25) the reader may check that

$$p_{ij.\overline{s+2}\,\overline{s+3}\ldots p} = p_{ij.\overline{s+1}\,\overline{s+2}\ldots p} + \frac{a_{\overline{s+1}\,i}a_{\overline{s+1}\,j}}{q_{\overline{s+1}\,\overline{s+1.12}\ldots s}} = p_{ij.\overline{s+1}\,\overline{s+2}\ldots p} + \frac{h_{\overline{s+1}\,i}h_{\overline{s+1}\,j}}{q_{\overline{s+1}\,\overline{s+1.12}\ldots s}} \tag{4.3.10}$$

for i and $j = 1, 2, \ldots, s$, and

$$p_{ij.\overline{s+2}\,\overline{s+3}\ldots p} = \frac{a_{\overline{s+1}\,i}a_{\overline{s+1}\,j}}{q_{\overline{s+1}\,\overline{s+1.\,2}\ldots s}} \tag{4.3.11}$$

when either i or j or both are $s + 1$. Note that (4.3.11) may be explicitly written

$$p_{i\,\overline{s+1.s+2}\,\overline{s+3}\ldots p} = -\frac{h_{\overline{s+1}\,i}}{q_{\overline{s+1}\,\overline{s+1.12}\ldots s}} = p_{\overline{s+1}\,i.\overline{s+2}\,\overline{s+3}\ldots p}$$

for $i = 1, 2, \ldots, s$, and (4.3.12)

$$p_{\overline{s+1}\,\overline{s+1.s+2}\,\overline{s+3}\ldots p} = \frac{1}{q_{\overline{s+1}\,\overline{s+1.12}\ldots s}}.$$

Formulas (4.3.10) and (4.3.12) express the upstaged $\mathbf{Q}_{11}^{-1}$ in terms of the output of layers 1, 2, and 3 of the first s stages of the elimination procedures. This process of upstaging $\mathbf{Q}_{11}^{-1}$ is sometimes called *bordering*.

The foregoing formulas provide computational routines for essentially all of the matrices discussed in Section 4.2. Note that the first diagonal element of $\mathbf{Q}_{22.1}$ is (U_{s+1}^*, U_{s+1}^*) which is the $(s + 1, s + 1)$ element of $\mathbf{T}$, so that the elements of $\mathbf{T}$ may be picked off from the successive $\mathbf{Q}_{22.1}$ as s increases. Knowing $\mathbf{T}$, $\det \mathbf{Q}$ may easily be computed from (4.2.6). From $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{T}$ one has $\mathbf{R}$, and thence $\mathbf{C}$ and $\mathbf{D}$ from (4.2.29).

Before illustrating these calculations it will be helpful to provide a very simple means of describing them.

4.3.2. The sweep operator. A $p \times p$ matrix $\mathbf{M}$ will be said to have been *swept on row k and column k* if $\mathbf{M}$ is replaced by another $p \times p$ matrix $\mathbf{N}$ whose (i, j) elements n_{ij} are related to the (i, j) elements m_{ij} of $\mathbf{M}$ as follows:

$$\begin{aligned} n_{kk} &= -1/m_{kk} \\ n_{ik} &= m_{ik}/m_{kk} \\ n_{kj} &= m_{kj}/m_{kk} \\ n_{ij} &= m_{ij} - m_{ik} \cdot m_{kj}/m_{kk}, \end{aligned} \tag{4.3.13}$$

for $i \neq k$ and $j \neq k$. For brevity $\mathbf{N}$ will be denoted by $\mathrm{SWP}[k]\mathbf{M}$ and the result of successively applying the operations $\mathrm{SWP}[k_1]$, $\mathrm{SWP}[k_2]$, $\ldots$, $\mathrm{SWP}[k_t]$ to $\mathbf{M}$ will be denoted by $\mathrm{SWP}[k_1, k_2, \ldots, k_t]$. The terminology here is borrowed from Beaton (1964).

With an electronic computer the operations (4.3.13) would usually be carried out as follows, beginning from a set of p^2 registers containing the elements of $\mathbf{M}$. First, m_{kk} is replaced by $n_{kk} = -1/m_{kk}$. Second, the remaining elements m_{ik} of column k are replaced by $n_{ik} = -m_{ik}n_{kk}$. Third, the elements m_{ij} in neither row k nor column k are replaced by $n_{ij} = m_{ij} - n_{ik}m_{kj}$. Finally, the remaining elements m_{kj} in row k are replaced by $n_{kj} = -m_{kj}n_{kk}$. Each stage here is designed to use only the output of the previous stages, so that no waste motion of saving numbers in special registers is needed. Note, however, that in the primary applications of this book $\mathbf{M}$ is a symmetric matrix, whence (4.3.13) shows that $\mathbf{N}$ is also a symmetric matrix. It follows that only $p(p + 1)/2$ registers are required to store $\mathbf{M}$, and the foregoing routine may be altered slightly to fit this circumstance. In any case, it is clear that a computer subroutine can be readily programmed to carry out the sweep operations one at a time and thence successively, i.e., $\mathrm{SWP}[k_1, k_2, \ldots, k_t]\mathbf{M}$ is quite amenable to electronic computation.

The calculations of each stage of the elimination procedure of Section 4.3.1 may be viewed simply as a sweep operation. To see this, set

$$\mathbf{Q}_{(012\ldots s)} = \begin{bmatrix} -\mathbf{Q}_{11}^{-1} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{Q}_{22.1} \end{bmatrix} \tag{4.3.14}$$

where $\mathbf{H}_{12} = \mathbf{H}_{21}'$. (The remaining notation was introduced in Section 4.2.) The notation (4.3.14) should be construed to include the limiting cases

$$\mathbf{Q}_{(0)} = \mathbf{Q} \tag{4.3.15}$$

and

$$\mathbf{Q}_{(012\ldots p)} = -\mathbf{Q}^{-1}. \tag{4.3.16}$$

The reason for the notation (4.3.14) is the following simple but powerful theorem which is an immediate consequence of (4.3.1), (4.3.3), (4.3.9), (4.3.10), and (4.3.12) together with the definition (4.3.13).

Theorem 4.3.1.

$$\mathbf{Q}_{(012\ldots\overline{s+1})} = \mathrm{SWP}[s + 1]\mathbf{Q}_{(012\ldots s)} \tag{4.3.17}$$

for $s = 0, 1, 2, \ldots, p - 1$ *and consequently*

$$\mathbf{Q}_{(012\ldots s)} = \mathrm{SWP}[1, 2, \ldots, s]\mathbf{Q} \tag{4.3.18}$$

for $s = 1, 2, \ldots, p$

The remarkable feature of Theorem 4.3.1 is that it shows how the same computing operation is involved at each stage s. By simply carrying out the operations $\mathrm{SWP}[1]$, $\mathrm{SWP}[2]$, $\ldots$, $\mathrm{SWP}[p]$, one can pick up $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{T}$ along the way and finally arrive at $\mathrm{SWP}[1, 2, \ldots, p]\mathbf{Q} = -\mathbf{Q}^{-1}$. (Recall that the first diagonal element of $\mathbf{Q}_{22.1}$ is (U_{s+1}^*, U_{s+1}^*), that column s of $\mathbf{H}_{21}$ is that part of column s of $\mathbf{B}$ below the diagonal, and that the first row of $\mathbf{H}_{21}$ is the negative of that part of row $s + 1$ of $\mathbf{A}$ to the left of the diagonal.)

Example 4.3 The following is a 4 × 4 sample covariance matrix:

$$\mathbf{Q} = \begin{bmatrix} 19.1434 & 9.0356 & 9.7634 & 3.2394 \\ 9.0356 & 11.8658 & 4.6232 & 2.4746 \\ 9.7634 & 4.6232 & 12.2978 & 3.8794 \\ 3.2394 & 2.4746 & 3.8794 & 2.4604 \end{bmatrix}.$$

The calculations of the elimination procedure to find **B**, **T**, **A**, and $\mathbf{Q}^{-1}$ will be illustrated on this **Q**. The reader should attempt to reproduce the numbers in the subsequent matrices.

The first stage is essentially to compute

$$\text{SWP}[1]\mathbf{Q} = \begin{bmatrix} -0.0522373 & 0.471995 & 0.510014 & 0.169218 \\ 0.471995 & 7.60104 & 0.01492 & 0.94562 \\ 0.510014 & 0.01492 & 7.31833 & 2.22726 \\ 0.169218 & 0.94562 & 2.22726 & 1.91224 \end{bmatrix}.$$

At this stage one knows that

$$\mathbf{T} = \begin{bmatrix} 19.1434 & 0 & 0 & 0 \\ 0 & 7.60104 & 0 & 0 \\ 0 & 0 & ? & 0 \\ 0 & 0 & 0 & ? \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.471995 & 1 & 0 & 0 \\ 0.510014 & ? & 1 & 0 \\ 0.169218 & ? & ? & 1 \end{bmatrix},$$

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -0.471995 & 1 & 0 & 0 \\ ? & ? & 1 & 0 \\ ? & ? & ? & 1 \end{bmatrix},$$

and $19.1434^{-1} = 0.0522373$.

The second stage is to compute

$$\text{SWP}[1, 2]\mathbf{Q} = \text{SWP}[2]\text{SWP}[1]\mathbf{Q}$$

$$= \begin{bmatrix} -0.0815463 & 0.0620961 & 0.509088 & 0.110499 \\ 0.0620961 & -0.131561 & 0.0019629 & 0.1294067 \\ 0.509088 & 0.0019629 & 7.318301 & 2.22574 \\ 0.110499 & 0.1244067 & 2.22574 & 1.79460 \end{bmatrix}$$

from which

$$\mathbf{T} = \begin{bmatrix} 19.1434 & 0 & 0 & 0 \\ 0 & 7.60104 & 0 & 0 \\ 0 & 0 & 7.318301 & 0 \\ 0 & 0 & 0 & ? \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.471995 & 1 & 0 & 0 \\ 0.510014 & 0.0019629 & 1 & 0 \\ 0.169218 & 0.1244067 & ? & 0 \end{bmatrix},$$

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -0.471995 & 1 & 0 & 0 \\ -0.509088 & -0.0019629 & 1 & 0 \\ ? & ? & ? & 1 \end{bmatrix},$$

and

$$\begin{bmatrix} 19.1434 & 9.0356 \\ 9.0356 & 11.8658 \end{bmatrix}^{-1} = \begin{bmatrix} 0.0815463 & -0.0620961 \\ -0.0620961 & 0.131561 \end{bmatrix}.$$

The third stage produces

$$\text{SWP}[1, 2, 3]\mathbf{Q} = \text{SWP}[3]\text{SWP}[1, 2]\mathbf{Q}$$

$$= \begin{bmatrix} -0.1169601 & 0.0619596 & 0.0695633 & -0.044331 \\ 0.0619596 & -0.131562 & 0.0002682 & 0.123810 \\ 0.0695633 & 0.0002682 & -0.136643 & 0.304132 \\ -0.044331 & 0.123810 & 0.304132 & 1.11768 \end{bmatrix},$$

from which

$$\mathbf{T} = \begin{bmatrix} 19.1434 & 0 & 0 & 0 \\ 0 & 7.60104 & 0 & 0 \\ 0 & 0 & 7.318301 & 0 \\ 0 & 0 & 0 & 1.11768 \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.471995 & 1 & 0 & 0 \\ 0.510014 & 0.0019629 & 1 & 0 \\ 0.169218 & 0.1244067 & 0.304132 & 1 \end{bmatrix},$$

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -0.471995 & 1 & 0 & 0 \\ -0.509088 & -0.0019629 & 1 & 0 \\ 0.044331 & -0.123810 & -0.304132 & 1 \end{bmatrix},$$

and

$$\begin{bmatrix} 19.1434 & 9.0356 & 9.7634 \\ 9.0356 & 11.8658 & 4.6232 \\ 9.7634 & 4.6232 & 12.2978 \end{bmatrix}^{-1} = \begin{bmatrix} 0.1169601 & -0.0619596 & -0.0695633 \\ -0.0619596 & 0.131562 & -0.0002682 \\ -0.0695633 & -0.0002682 & 0.136643 \end{bmatrix}.$$

If desired, a final stage may be added to yield

$$\mathbf{Q}^{-1} = -\text{SWP}[4]\text{SWP}[1, 2, 3]\mathbf{Q}$$

$$= \begin{bmatrix} 0.118718 & -0.066871 & -0.081626 & 0.039663 \\ -0.066871 & 0.145277 & 0.033422 & -0.110774 \\ -0.081626 & 0.033422 & 0.219400 & -0.272110 \\ 0.039663 & 0.110774 & -0.272110 & 0.894710 \end{bmatrix}.$$

Other quantities of interest might be $\det \mathbf{Q} = \det \mathbf{T} = 1190.04$ and the square roots of $\mathbf{T}$, namely,

$$\mathbf{R} = \begin{bmatrix} 4.3753 & 0 & 0 & 0 \\ 0 & 2.7570 & 0 & 0 \\ 0 & 0 & 2.7052 & 0 \\ 0 & 0 & 0 & 1.0572 \end{bmatrix}.$$

From (4.2.9), $\mathbf{D}$ is found from $\mathbf{B}$ by multiplying each column of $\mathbf{B}$ by the corresponding element of $\mathbf{R}$, giving

$$\mathbf{D} = \begin{bmatrix} 4.3753 & 0 & 0 & 0 \\ 2.0651 & 2.7570 & 0 & 0 \\ 2.2315 & 0.0054 & 2.7052 & 0 \\ 0.5712 & 0.3430 & 0.8277 & 1.0572 \end{bmatrix}.$$

Similarly, $\mathbf{C}$ is found from $\mathbf{A}$ by dividing each row of $\mathbf{A}$ by the corresponding element of $\mathbf{R}$, giving

$$\mathbf{C} = \begin{bmatrix} 0.22855 & 0 & 0 & 0 \\ -0.17120 & 0.36271 & 0 & 0 \\ -0.18819 & -0.00073 & 0.36966 & 0 \\ 0.04193 & -0.11711 & -0.28768 & 0.94589 \end{bmatrix}.$$

These calculations were performed on a desk calculator. Variations on the calculations—to be exhibited later in Section 4.3—were performed independently on a desk calculator and show minor deviations from those given above due to rounding error.

Two important properties of the sweep operator will now be demonstrated. The first of these is commutativity, i.e.,

$$\text{SWP}[i, j]\mathbf{M} = \text{SWP}[j, i]\mathbf{M}, \tag{4.3.19}$$

from which it follows that

$$\text{SWP}[i_1, i_2, \ldots, i_t]\mathbf{M} = \text{SWP}[j_1, j_2, \ldots, j_t]\mathbf{M}, \tag{4.3.20}$$

where $i_1, i_2, \ldots, i_t$ and $j_1, j_2, \ldots, j_t$ are permutations of the same set of integers. Formula (4.3.19) may be deduced directly from (4.3.13), but it is interesting to note also that it can be deduced from (4.3.18). The point is that while all of the parts $-\mathbf{Q}_{11}^{-1}$, $\mathbf{H}_{21}$, and $\mathbf{Q}_{22.1}$ of $\mathbf{Q}_{(012\ldots s)}$ depend on the parts $\mathbf{U}_1$ and $\mathbf{U}_2$ of the basis $\mathbf{U}$ they do not depend on the order of the basis elements of $\mathbf{U}_1$, at least not in any meaningful sense. Thus $\mathbf{Q}_{11}^{-1}$ is the inverse of the inner product matrix of $\mathbf{U}_1$, and if $\mathbf{Q}_{11}^{-1}$ is computed using the basis elements in one order it can equally well be computed using the basis elements in any other order. Similar remarks apply to $\mathbf{H}_{21}$ via (4.2.17) and to $\mathbf{Q}_{22.1}$ which is the inner product matrix of the

components of $\mathbf{U}_2$ orthogonal to $\mathbf{U}_1$. These computations using the basis elements in a different order are, however, nothing more than the application of the sweep operators SWP[1], SWP[2], ..., SWP[s] in a different order. Consequently, permuting the order of these operations cannot affect the outcome.

The second property of the sweep operator is that it is very easily undone. Indeed, the equations (4.3.13) may be solved to yield

$$\begin{aligned} m_{kk} &= -1/n_{kk} \\ m_{ik} &= -n_{ik}/n_{kk} \\ m_{kj} &= -n_{kj}/n_{kk} \\ m_{ij} &= n_{ij} - n_{ik}n_{kj}/n_{kk} \end{aligned} \tag{4.3.21}$$

for $j \neq i$ and $k \neq i$. The operator defined by (4.3.21) may be denoted by

$$\mathbf{M} = \text{RSW}[k]\mathbf{N} \tag{4.3.22}$$

and may be called the *reverse sweep operator on row and column k*. RSW operators commute with each other, and as with SWP operators, RSW[$k_1, k_2, \ldots, k_t$] will denote the result of successively applying RSW[k_1], RSW[k_2], ..., RSW[k_t] in any order.

The formulas

$$\begin{aligned} \text{SWP}[1, 2, \ldots, s]\begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} &= \begin{bmatrix} -\mathbf{Q}_{11}^{-1} & \mathbf{Q}_{11}^{-1}\mathbf{Q}_{12} \\ \mathbf{Q}_{21}\mathbf{Q}_{11}^{-1} & \mathbf{Q}_{22} - \mathbf{Q}_{21}\mathbf{Q}_{11}^{-1}\mathbf{Q}_{12} \end{bmatrix} \text{ and} \\ \text{RSW}[1, 2, \ldots, s]\begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} &= \begin{bmatrix} -\mathbf{Q}_{11}^{-1} & -\mathbf{Q}_{11}^{-1}\mathbf{Q}_{12} \\ -\mathbf{Q}_{21}\mathbf{Q}_{11}^{-1} & \mathbf{Q}_{22} - \mathbf{Q}_{21}\mathbf{Q}_{11}^{-1}\mathbf{Q}_{12} \end{bmatrix} \end{aligned} \tag{4.3.23}$$

are useful mathematical characterizations of the general SWP and RSW operators.

4.3.3. The assimilation operator. In place of the elimination procedure described above, it is possible to proceed in stages which produce after stage s the successive orthogonalization of $U_1, U_2, \ldots, U_s$ without having touched $U_{s+1}, U_{s+2}, \ldots, U_p$. At stage s, U_{s+1} is *assimilated* into the picture, for $s = 1, 2, \ldots, p-1$. First, however, a more general assimilation problem is considered.

Suppose that computations on $\mathbf{Q}$ have proceeded to the stage SWP[1, 2, ..., s]$\mathbf{Q}$ when the p-dimensional space $\mathscr{E}$ with basis $U_1, U_2, \ldots, U_p$ is enlarged to a $(p+r)$-dimensional space $\mathscr{E}^*$ by the addition of new basis variables $U_{p+1}, U_{p+2}, \ldots, U_{p+r}$. At the same time the $p \times p$ inner product matrix

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} \tag{4.3.24}$$

is enlarged to the $(p+r) \times (p+r)$ inner product matrix

$$\mathbf{Q}^* = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} & \mathbf{Q}_{13} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} & \mathbf{Q}_{23} \\ \mathbf{Q}_{31} & \mathbf{Q}_{32} & \mathbf{Q}_{33} \end{bmatrix}. \tag{4.3.25}$$

The basic computing task considered here is to produce SWP[1, 2, . . . , s]$\mathbf{Q}^*$ from SWP[1, 2, . . . , s]$\mathbf{Q}$ and the last r rows and columns of $\mathbf{Q}^*$, i.e., to pass from

$$\begin{bmatrix} -\mathbf{Q}_{11}^{-1} & \mathbf{H}_{12} & \mathbf{Q}_{13} \\ \mathbf{H}_{21} & \mathbf{Q}_{22.1} & \mathbf{Q}_{23} \\ \mathbf{Q}_{31} & \mathbf{Q}_{32} & \mathbf{Q}_{33} \end{bmatrix} \tag{4.3.26}$$

to

$$\left[\begin{array}{cc|c} -\mathbf{Q}_{11}^{-1} & \mathbf{H}_{12} & \mathbf{H}_{13} \\ \mathbf{H}_{21} & \mathbf{Q}_{22.1} & \mathbf{Q}_{23.1} \\ \hline \mathbf{H}_{31} & \mathbf{Q}_{32.1} & \mathbf{Q}_{33.1} \end{array}\right], \tag{4.3.27}$$

where, from (4.3.18),

$$\mathbf{H}_{13} = \mathbf{H}'_{31} = \mathbf{Q}_{11}^{-1}\mathbf{Q}_{13}, \tag{4.3.28}$$

$$\mathbf{Q}_{23.1} = \mathbf{Q}'_{32.1} = \mathbf{Q}_{23} - \mathbf{H}_{21}\mathbf{Q}_{13}, \tag{4.3.29}$$

and

$$\mathbf{Q}_{33.1} = \mathbf{Q}_{33} - \mathbf{H}_{31}\mathbf{Q}_{13}. \tag{4.3.30}$$

It is convenient to introduce the operator notation ASM[$p + 1, p + 2$, . . . , $p + r$; 1, 2, . . . , s] for the passage from (4.3.26) to (4.3.27) defined by (4.3.28), (4.3.29), and (4.3.30). In general terms, the operator may be written

$$\begin{aligned} \text{ASM}[p+1, p+2, \ldots, p+r; 1, 2, \ldots, s] &\begin{bmatrix} \mathbf{K}_{11} & \mathbf{K}_{12} & \mathbf{K}_{13} \\ \mathbf{K}_{21} & \mathbf{K}_{22} & \mathbf{K}_{23} \\ \mathbf{K}_{31} & \mathbf{K}_{32} & \mathbf{K}_{33} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{K}_{11} & \mathbf{K}_{12} & -\mathbf{K}_{11}\mathbf{K}_{13} \\ \mathbf{K}_{21} & \mathbf{K}_{22} & \mathbf{K}_{23} - \mathbf{K}_{21}\mathbf{K}_{13} \\ -\mathbf{K}_{31}\mathbf{K}_{11} & \mathbf{K}_{32} - \mathbf{K}_{31}\mathbf{K}_{12} & \mathbf{K}_{33} + \mathbf{K}_{31}\mathbf{K}_{11}\mathbf{K}_{13} \end{bmatrix}. \end{aligned} \tag{4.3.31}$$

The notation ASM stands for *assimilate*; more fully, the assimilation operator ASM[$p + 1, p + 2, \ldots, p + r$; 1, 2, . . . , s] assimilates $U_{p+1}, U_{p+2}, \ldots, U_{p+r}$ into a situation where $U_1, U_2, \ldots, U_s$ have already been swept out.

Two alternative characterizations of the ASM operator follow. First, it is clear that the assimilation of $U_{p+1}, U_{p+2}, \ldots, U_{p+r}$ may itself be carried out in r steps. Thus ASM[$p + 1, p + 2, \ldots, p + r$; 1, 2, . . . , s] may be described as the result of applying in order ASM[$p + 1$; 1, 2, . . . , s] to the first $p + 1$ rows and columns, ASM[$p + 2$; 1, 2, . . . , s] to the first $p + 2$ rows and columns, and so on to ASM[$p + r$; 1, 2, . . . , s] applied to the whole matrix. This is the direct way to program the ASM operator for an electronic computer. The second characterization of the ASM[$p + 1, p + 2, \ldots, p + r$; 1, 2, . . . , s] operator is that it is the result of applying RSW[1, 2, . . . , s] to the first p rows and columns followed by SWP[1, 2, . . . , s] applied to the whole matrix. The first step here returns the first p rows and columns of (4.3.25) to $\mathbf{Q}$, and the second step then finds SWP[1, 2, . . . , s]$\mathbf{Q}^*$ directly from $\mathbf{Q}^*$. This second characterization of the ASM operator is given for its mathematical interest; it is clearly an inefficient way to carry out the calculations.

As an example of the ASM operator consider the following operations applied in sequence to $\mathbf{Q}$:

SWP[1] applied to the upper left 1 × 1 submatrix,
ASM[2; 1] applied to the upper left 2 × 2 submatrix,
SWP[2] applied to the upper left 2 × 2 submatrix,
.
.
.
ASM[s; 1, 2, . . . , s − 1] applied to the upper left s × s submatrix,
SWP[s] applied to the upper left s × s submatrix,
.
.
.
ASM[p; 1, 2, . . . , p − 1] applied to the whole matrix,
SWP[p] applied to the whole matrix.

Just before the SWP[s] operation, row s provides the negative of row s of $\mathbf{A}$. Just after the SWP[s] operation, the upper left s × s submatrix provides $-\mathbf{Q}_{11}^{-1}$. The final result of the sequence of operations is $-\mathbf{Q}^{-1}$.

Example 4.3 (*First continuation*) The computing sequence just described is applied to the 4 × 4 matrix $\mathbf{Q}$ used to illustrate the SWP operator:

$$-0.052237,$$

$$\begin{bmatrix} -0.052237 & 0.471993 \\ 0.471993 & 7.601060 \end{bmatrix},$$

$$\begin{bmatrix} -0.081546 & 0.062096 \\ 0.062096 & -0.131561 \end{bmatrix},$$

$$\begin{bmatrix} -0.081546 & 0.062096 & 0.509084 \\ 0.062096 & -0.131561 & 0.001965 \\ 0.509084 & 0.001965 & 7.318324 \end{bmatrix},$$

$$\begin{bmatrix} -0.116959 & 0.061959 & 0.069563 \\ 0.061959 & -0.131562 & 0.000269 \\ 0.069563 & 0.000269 & -0.136643 \end{bmatrix},$$

$$\begin{bmatrix} -0.116959 & 0.061959 & 0.069563 & -0.044310 \\ 0.061959 & -0.131562 & 0.000269 & 0.123809 \\ 0.069563 & 0.000269 & -0.136643 & 0.304085 \\ -0.04431 & 0.123809 & 0.304085 & 1.117893 \end{bmatrix},$$

$$\begin{bmatrix} -0.118715 & 0.066866 & 0.081616 & -0.039637 \\ 0.066866 & -0.145274 & -0.033409 & 0.110752 \\ 0.081616 & -0.033409 & -0.219359 & 0.272016 \\ -0.039637 & 0.110752 & 0.272016 & -0.894540 \end{bmatrix}.$$

The final result here is $-\mathbf{Q}^{-1}$. The reader should also locate in this sequence the elements of **A** and **T**.

4.3.4. The multistandardize operator (cf. Beaton, 1964). For later computations involving eigenvalues and eigenvectors (to be discussed in Chapter 5) it is often desirable to have in hand an inner product matrix where a subset of the vectors **U** has been replaced by an orthonormal set spanning the same subspace as the original subset. Usually any orthonormal set of linear combinations of the specified subset will suffice, and the process of successive orthogonalization is a convenient means to such an end.

Suppose that **Q** is the inner product matrix for a basis **U** and, for simplicity, that the subset $\mathbf{U}_1$ is to be replaced by the orthonormal subset $\mathbf{U}_1^{**}$ found by orthogonalizing $\mathbf{U}_1$ in the given order. The objectives are to find two $p \times p$ matrices $\mathbf{Q}_{[012\ldots s]}$ and $\mathbf{C}_{[012\ldots s]}$ where $\mathbf{Q}_{[012\ldots s]}$ is the inner product matrix of the new basis $[\mathbf{U}_1^{**}, \mathbf{U}_2]'$ and $\mathbf{C}_{[012\ldots s]}$ expresses the basis $[\mathbf{U}_1^{**}, \mathbf{U}_2]'$ in terms of **U**, i.e.,

$$\begin{bmatrix}\mathbf{U}_1^{**}\\ \mathbf{U}_2\end{bmatrix} = \mathbf{C}_{[012\ldots s]}\mathbf{U}. \tag{4.3.32}$$

The limiting cases $\mathbf{Q}_{[0]}$ and $\mathbf{C}_{[0]}$ will be taken to be **Q** and **I**.

The *multistandardize operator* MST$[1, 2, \ldots, s]$ is a computation which modifies the pair **[Q, I]** in a set of $2p^2$ registers into the pair $[\mathbf{Q}_{[012\ldots s]}, \mathbf{C}_{[012\ldots s]}]$. It does this in s stages where stage $r+1$ for $r = 0, 1, \ldots, s-1$ modifies $[\mathbf{Q}_{[012\ldots r]}, \mathbf{C}_{[012\ldots r]}]$ into $[\mathbf{Q}_{[012\ldots \overline{r+1}]}, \mathbf{C}_{[012\ldots \overline{r+1}]}]$. The following four paragraphs will describe how to carry out stage $r+1$ in such a multistandardize operation.

First note that, when $\mathbf{Q}_{[012\ldots s]}$ and $\mathbf{C}_{[012\ldots s]}$ are partitioned in the usual way, only one part of each is unknown at the outset. Thus

$$\mathbf{Q}_{[012\ldots s]} = \begin{bmatrix}\mathbf{I} & \mathbf{Q}_{12}^{**}\\ \mathbf{Q}_{21}^{**} & \mathbf{Q}_{22}\end{bmatrix}, \tag{4.3.33}$$

where **I** and $\mathbf{Q}_{22}$ are the known inner product matrices of $\mathbf{U}_1^{**}$ and $\mathbf{U}_2$ and where $\mathbf{Q}_{21}^{**}$ or its transpose $\mathbf{Q}_{12}^{**}$ contain the unknown inner products between $\mathbf{U}_1^{**}$ and $\mathbf{U}_2$. Similarly

$$\mathbf{C}_{[012\ldots s]} = \begin{bmatrix}\mathbf{C}_{11} & \mathbf{0}'\\ \mathbf{0} & \mathbf{I}\end{bmatrix}, \tag{4.3.34}$$

where $\mathbf{C}_{11}$ is the part of **C** which expresses $\mathbf{U}_1^{**}$ in terms of $\mathbf{U}_1$ as in (4.2.33), and where **0** and its transpose consist of zeros and **I** is the $(p-s) \times (p-s)$ identity which expresses $\mathbf{U}_2$ in terms of $\mathbf{U}_2$. These partitions may also be applied with s replaced by r for $r = 0, 1, 2, \ldots, s-1$.

Suppose that (U_i, U_j^{**}) is denoted by q_{ij}^{**} for $i \geq j$. The formulas underlying the computations at stage $r+1$ are

$$U_{r+1}^{*} = U_{r+1} - q_{r+1\,1}^{**}U_1^{**} - q_{r+1\,2}^{**}U_2^{**} - \cdots - q_{r+1\,r}^{**}U_r^{**} \tag{4.3.35}$$

and

$$U_{r+1}^{**} = (U_{r+1}^{*}, U_{r+1}^{*})^{-1/2}U_{r+1}^{*}. \tag{4.3.36}$$

From (4.3.35),

$$(U_{r+1}^{*}, U_{r+1}^{*}) = q_{\overline{r+1}\,\overline{r+1}} - (q_{r+1\,1}^{**})^2 - (q_{r+1\,2}^{**})^2 - \cdots - (q_{r+1\,r}^{**})^2, \tag{4.3.37}$$

and from (4.3.36)

$$q_{i\,r+1}^{**} = (q_{i\,\overline{r+1}} - q_{r+1\,1}^{**}q_{i1}^{**} - q_{r+1\,2}^{**}q_{i2}^{**} - \cdots - q_{r+1\,r}^{**}q_{ir}^{**})/(U_{r+1}^{*}, U_{r+1}^{*})^{1/2} \tag{4.3.38}$$

for $i = r+2, r+3, \ldots, p$.

Passage from $\mathbf{Q}_{[012\ldots r]}$ to $\mathbf{Q}_{[012\ldots \overline{r+1}]}$ requires only modification of row and column $r+1$. The last $(p-r-1)$ elements of row and column $r+1$ are modified by replacing $q_{i\,\overline{r+1}} = q_{\overline{r+1}\,i}$ by $q_{i\,\overline{r+1}}^{**}$ calculated from (4.3.38) with $i = r+2, r+3, \ldots, p$. Note that (4.3.38) requires preliminary computation of $(U_{r+1}^{*}, U_{r+1}^{*})$ via (4.3.37). The remaining elements of row and column $r+1$ are modified to zeros except for the diagonal element which becomes unity.

From (4.3.32) it is seen that linear combinations of the vectors $\mathbf{U}_1^{**}$ and $\mathbf{U}_2$ correspond to linear combinations of rows of $\mathbf{C}_{[012\ldots s]}$. It is therefore clear from (4.3.35) and (4.3.36) that $\mathbf{C}_{[012\ldots \overline{r+1}]}$ is produced from $\mathbf{C}_{[012\ldots r]}$ by altering row $r+1$ of $\mathbf{C}_{[012\ldots r]}$ as follows: first replace row $r+1$ of $\mathbf{C}_{[012\ldots r]}$ by

$$[\text{row } r+1 \text{ of } \mathbf{C}_{[012\ldots r]}] - q_{r+1\,1}^{**}[\text{row 1 of } \mathbf{C}_{[012\ldots r1]} - q_{r+1\,2}^{**}[\text{row 2 of } \mathbf{C}_{[012\ldots r]}] - \cdots - q_{r+1\,r}^{**}[\text{row } r \text{ of } \mathbf{C}_{[012\ldots r]}], \tag{4.3.39}$$

and then divide through the resulting row $r+1$ by $(U_{r+1}^{*}, U_{r+1}^{*})^{1/2}$.

This completes the definition of the multistandardize operator MST$[1, 2, \ldots, s]$ with the property

$$\text{MST}[1, 2, \ldots, s][\mathbf{Q}, \mathbf{I}] = [\mathbf{Q}_{[012\ldots s]}, \mathbf{C}_{[012\ldots s]}]. \tag{4.3.40}$$

Example 4.3 (*Second continuation*) Beginning from

$$[\mathbf{Q}, \mathbf{I}] = \begin{bmatrix}19.1434 & 9.0356 & 9.7634 & 3.2394 & 1 & 0 & 0 & 0\\ 9.0356 & 11.8658 & 4.6232 & 2.4746 & 0 & 1 & 0 & 0\\ 9.7634 & 4.6232 & 12.2978 & 3.8794 & 0 & 0 & 1 & 0\\ 3.2394 & 2.4746 & 3.8794 & 2.4604 & 0 & 0 & 0 & 1\end{bmatrix},$$

the multistandardize operations will be applied to yield in succession $[\mathbf{Q}_{[01]}, \mathbf{C}_{[01]}]$, $[\mathbf{Q}_{[012]}, \mathbf{C}_{[012]}]$, $[\mathbf{Q}_{[0123]}, \mathbf{C}_{[0123]}]$, and $[\mathbf{Q}_{[01234]}, \mathbf{C}_{[01234]}] = [\mathbf{I}, \mathbf{C}]$.

DEUPTEST

The first stage simply replaces U_1 by $U_1^{**} = 0.22855U_1$ whence

$$\mathbf{Q}_{[01]} = \begin{bmatrix} 1 & 2.0651 & 2.2314 & 0.7404 \\ 2.0651 & 11.8658 & 4.6232 & 2.4746 \\ 2.2314 & 4.6232 & 12.2978 & 3.8794 \\ 0.7404 & 2.4746 & 3.8794 & 2.4604 \end{bmatrix}$$

and

$$\mathbf{C}_{[01]} = \begin{bmatrix} 0.22855 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The second stage requires computing from (4.3.37):

$$(U_2^*, U_2^*) = 11.8658 - 4.26460$$
$$= 7.6018, \text{ whence}$$
$$(U_2^*, U_2^*)^{1/2} = 2.7571.$$

Next, following (4.3.33) and (4.3.38),

$$\mathbf{Q}_{[012]} = \begin{bmatrix} 1 & 0 & 2.23143 & 0.74036 \\ 0 & 1 & 0.00548 & 0.34300 \\ 2.23143 & 0.00548 & 12.2978 & 2.8794 \\ 0.74036 & 0.34300 & 3.8794 & 2.4604 \end{bmatrix}.$$

Then, following (4.3.39),

$$\mathbf{C}_{[012]} = \begin{bmatrix} 0.22855 & 0 & 0 & 0 \\ -0.17119 & 0.36270 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The third stage follows the same general pattern as the second except with increasing complexity. Thus

$$(U_3^*, U_3^*) = 12.2978 - 4.97928 - 0.00003$$
$$= 7.3185, \text{ whence}$$
$$(U_3^*, U_3^*)^{1/2} = 2.7053.$$

Similarly,

$$\mathbf{Q}_{[0123]} = \begin{bmatrix} 1 & 0 & 0 & 0.74036 \\ 0 & 1 & 0 & 0.34300 \\ 0 & 0 & 1 & 0.82262 \\ 0.74036 & 0.34300 & 0.82262 & 2.4604 \end{bmatrix}$$

and

$$\mathbf{C}_{[0123]} = \begin{bmatrix} 0.22855 & 0 & 0 & 0 \\ -0.17119 & 0.36270 & 0 & 0 \\ -0.18817 & -0.00074 & 0.36964 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The final stage differs in that $\mathbf{Q}_{[01234]}$ is known to be simply $\mathbf{I}$. Thus, one needs

$$(U_4^*, U_4^*) = 1.11791,$$
$$(U_4^*, U_4^*)^{1/2} = 1.0573,$$

and

$$\mathbf{C} = \mathbf{C}_{[01234]} = \begin{bmatrix} 0.22855 & 0 & 0 & 0 \\ -0.17119 & 0.36270 & 0 & 0 \\ -0.18817 & -0.00074 & 0.36964 & 0 \\ 0.04190 & -0.11709 & -0.28759 & 0.94581 \end{bmatrix}.$$

It is clear that the rows and columns $1, 2, \ldots, s$ were arbitrarily chosen and that the definition may be extended to define $\text{MST}[i_1, i_2, \ldots, i_s]$ where $i_1, i_2, \ldots, i_s$ are any subset of the integers $1, 2, \ldots, p$. The vectorial definitions all proceed as before except that $U_1, U_2, \ldots, U_s$ are replaced by $U_{i_1}, U_{i_2}, \ldots, U_{i_s}$ and the desired orthonormal basis is that arising from the orthogonalization of $U_{i_1}, U_{i_2}, \ldots, U_{i_s}$ in this order. Precise details are left to the reader, but clearly (4.3.39) easily generalizes to

$$\text{MST}[i_1, i_2, \ldots, i_s][\mathbf{Q}, \mathbf{I}] = [\mathbf{Q}_{[0\,i_1 i_2 \ldots i_s]}, \mathbf{C}_{[0\,i_1 i_2 \ldots i_s]}]. \tag{4.3.41}$$

The definitions were originally given in the special case only because the partitions (4.3.33) and (4.3.34) are more easily displayed in that case.

A further observation is that the definition of the operator does not require the initial pair to be $[\mathbf{Q}, \mathbf{I}]$; rather, any $p \times p$ matrix $\mathbf{K}$ may be substituted for $\mathbf{I}$. It follows that (4.3.41) generalizes to

$$\text{MST}[i_1, i_2, \ldots, i_s][\mathbf{Q}, \mathbf{K}] = [\mathbf{Q}_{[0\,i_1 i_2 \ldots i_s]}, \mathbf{C}_{[0\,i_1 i_2 \ldots i_s]}\mathbf{K}]. \tag{4.3.42}$$

The idea behind definition (4.3.42) is that the basis $\mathbf{U}$ may have been derived from an initial basis $\mathbf{V}$ where $\mathbf{U} = \mathbf{K}\mathbf{V}$. Then (4.3.42) extends to

$$\begin{bmatrix} \mathbf{U}_1^{**} \\ \mathbf{U}_2 \end{bmatrix} = \mathbf{C}_{[012\ldots s]}\mathbf{K}\mathbf{V} \tag{4.3.43}$$

so that $\mathbf{C}_{[012\ldots s]}\mathbf{K}$ is an important matrix for relating back to the original basis $\mathbf{V}$. If (4.3.43) is used in place of (4.3.32) throughout the foregoing discussion, one is naturally led to the generalization (4.3.42) rather than (4.3.41).

The term *multistandardize* was suggested by Beaton (1964) who recognized the usefulness of the MST operator. The term arises as follows: if a single vector U_i in a basis $\mathbf{U}$ is replaced by $(U_i, U_i)^{-1/2}U_i$ having a unit norm, then the inner

product matrix $\mathbf{Q}$ of $\mathbf{U}$ is replaced by dividing row and column i of $\mathbf{Q}$ by $(U_i, U_i)^{-1/2}$. This is the same as the operation MST[i] applied to $\mathbf{Q}$. In the statistical context where U_i is a variable, the transformation $U_i \to (U_i, U_i)^{-1/2}U_i$ reduces U_i to a rescaled variable with unit norm or standard deviation which is often called a *standardized* variable, and so MST[i] is closely related to the *standardization* of variable i. The term *multistandardize* is a generalization referring to standardizing not a single variable but a block of variables, producing not a single variable with unit norm but a block of orthonormal variables.

4.4 SITUATIONS WITH LESS THAN FULL RANK

In Section 4.1 it was pointed out that the definition (4.1.1) for $U_1^*, U_2^*, \ldots, U_r^*$ may be made more general by agreeing to drop from the right side of (4.1.1) any term which is a multiple of U_s^* for any s such that $(U_s^*, U_s^*) = 0$. By similarly dropping terms in U_j^* where $(U_j^*, U_j^*) = 0$ on the right sides of (4.1.2) and (4.1.3) it is easily proved that $U_1^*, U_2^*, \ldots, U_r^*$ are mutually orthogonal in the general case. Recall, of course, that any U_s^* such that $(U_s^*, U_s^*) = 0$ is orthogonal to every V in $\mathscr{E}$.

The condition $(U_s^*, U_s^*) = 0$ means either that $U_s^* = \emptyset$, or that the inner product is of less than full rank, or both.

The possibility $U_s^* = \emptyset$ implies linear dependence among $U_1, U_2, \ldots, U_s$. More precisely, *if the inner product over $\mathscr{E}$ has full rank, if the subspace spanned by $U_1, U_2, \ldots, U_r$ has dimension $s \leq r$, and if the number of U_j^* different from $\emptyset$ is q, then $s = q$ and the subset of q non-$\emptyset$ U_j^* constitute an orthogonal basis of the subspace spanned by $U_1, U_2, \ldots, U_r$.* To prove this it will be shown that the subspaces spanned by $U_1, U_2, \ldots, U_r$ and $U_1^*, U_2^*, \ldots, U_r^*$ are identical. Clearly the latter is contained in the former because linear combinations of $U_1, U_2, \ldots, U_r$ are used to define the latter. Also (4.1.1) expresses each U_i in terms of the $U_1^*, U_2^*, \ldots, U_r^*$, so that the former subspace is contained in the latter, as required. To prove that the dimension of this subspace is q, one need only show that the s non-$\emptyset$ U_j^* which span it are linearly independent. (This was left to the reader as Exercise 3.1.7.)

The theorem just proved shows how the process of successive orthogonalization may, in principle, be used to check on the linear dependence of a set of vectors and to find the dimension of the subspace spanned by these vectors. Since the successive orthogonalization of $U_1, U_2, \ldots, U_r$ contains the successive orthogonalization of $U_1, U_2, \ldots, U_t$ for every $t \leq r$, it produces as a by-product the dimension of each of the subspaces spanned by $U_1, U_2, \ldots, U_t$ for $t = 1, 2, \ldots, r$. It is of interest to note that while the dimensions of these subspaces are affine properties, the structure of Euclidean vector spaces was used in finding them. It follows that if the inner product were changed, resulting in general in a different set of $U_1^*, U_2^*, \ldots, U_r^*$, the property of whether a particular $U_j^* = \emptyset$ or not would not be changed.

Turning now to the possibility that the inner product over $\mathscr{E}$ has less than full rank, suppose that $U_1, U_2, \ldots, U_p$ is a basis of $\mathscr{E}$ and that the inner product has rank f so that the inner product matrix $\mathbf{Q}$ of the basis $\mathbf{V}$ is positive semi-definite symmetric of rank f. Since, as before, the $U_1^*, U_2^*, \ldots, U_p^*$ defined by the general version of (4.1.1) span the same space as $U_1, U_2, \ldots, U_p$, it must be that $U_1^*, U_2^*, \ldots, U_p^*$ form a basis of $\mathscr{E}$, and $U_s^* = \emptyset$ is impossible for any s. Furthermore, it is clear that the set of U_s^* such that $(U_s^*, U_s^*) = 0$ must span the $(p - f)$-dimensional subspace $\mathscr{U}$ of vectors with zero norm and that the remaining U_s^* span a complementary f-dimensional subspace $\mathscr{V}$. In particular, it follows that the number of U_s^* with positive norm must be f. This result shows how in principle to compute the rank of a given positive semi-definite symmetric matrix $\mathbf{Q}$.

The general definition of $\mathbf{U}^*$ discussed above requires that $\mathbf{U} = \mathbf{B}\mathbf{U}^*$ where the elements b_{is} of $\mathbf{B}$ for $i > s$ are chosen to be zero when $(U_s^*, U_s^*) = 0$. However, if such zero elements were to be replaced by arbitrary numbers, the $\mathbf{U}_j^*$ would be altered only by the addition of vectors of zero norm. In particular, the diagonal inner product matrix of $\mathbf{T}$ of $\mathbf{U}^*$ would be the same for each such choice of $\mathbf{B}$, and therefore (4.2.4) would continue to hold. Also, any such $\mathbf{B}$ is triangular with elements unity along the main diagonal and zero above, so that $\mathbf{B}$ has a unique inverse $\mathbf{A}$. Defining $\dot{\mathbf{T}}$ to be the matrix formed by inverting the nonzero elements of $\mathbf{T}$, it follows from (2.6.10) and (2.6.11) that

$$\dot{\mathbf{Q}} = \mathbf{A}'\dot{\mathbf{T}}\mathbf{A} \tag{4.4.1}$$

is a pseudoinverse of $\mathbf{Q}$ for any of the choices of $\mathbf{B}$ and thence of $\mathbf{A}$. Formula (4.4.1) is a generalization of (4.2.5).

The theory of this section has practical implications of several kinds. It is sometimes convenient in statistics to consider sets of variables including some which are known to be linear combinations of others; in effect, this means that linear dependence is built in. In such situations care must be taken not to program an electronic computer to carry out (4.1.1) in its original form, since the attempt to divide by zero will either produce wild results due to rounding error or will stop the machine.

Another possible complication arises when the input variables are not linearly dependent and small values of (U_s^*, U_s^*) are encountered due to empirical relationships among variables. In such situations, blind following of (4.1.1) may again lead to trouble because the value (U_s^*, U_s^*) used at stage s may consist largely of rounding error. In situations where the order of orthogonalization has some latitude, programs may be written which determine the order as the calculation proceeds in a way which controls the problem of small (U_s^*, U_s^*). But if the problem becomes too severe, the computation may need to be abandoned or its precision increased. For the most part, such questions of accuracy of computed values are beyond the scope of this book.

4.5 EXERCISES

4.1.1 Show that the successive orthogonalization of U_1, U_2 in general produces different vectors from the successive orthogonalization of U_2, U_1. What is the exceptional case? Illustrate your answer with a plane diagram.

4.1.2 Show that the subspace spanned by $U_1^*, U_2^*, \ldots, U_s^*$ defined by (4.1.1) is the same as the subspace spanned by $U_1, U_2, \ldots, U_s$.

4.1.3 For any $t \leq s$, show that

$$U_s - \sum_{j=1}^{t-1} \frac{(U_s, U_j^*)}{(U_j^*, U_j^*)} U_j^* \quad \text{and} \quad \sum_{j=1}^{t-1} \frac{(U_s, U_j^*)}{(U_j^*, U_j^*)} U_j^*$$

denote the components of U_s along and orthogonal to the subspace spanned by U_1, $U_2, \ldots, U_{t-1}$. In particular, deduce that the mapping $U_s \to U_s^*$ is the orthogonal projection of U_s into the subspace orthogonal to $U_1, U_2, \ldots, U_{s-1}$.

4.1.4 Show that, if $\mathscr{V}$ is the orthogonal complement of $\mathscr{U}$, then $\mathscr{U}$ is the orthogonal complement of $\mathscr{V}$.

4.2.1 Show that, given any basis $\mathbf{U}$ of a p-dimensional Euclidean space $\mathscr{E}$, there exists one and only one *orthogonal* basis $\mathbf{U}^* = \mathbf{AU}$ such that $\mathbf{A}$ is triangular with elements unity along the diagonal and zero above the diagonal. Show also that there exists one and only one *orthonormal* basis $\mathbf{U}^{**} = \mathbf{CU}$ such that $\mathbf{C}$ is triangular with elements zero above the diagonal.

4.2.2 In the notation of Section 4.2, show that

$$\mathbf{T} = \mathbf{AQA}' \quad \text{and} \quad \mathbf{T}^{-1} = \mathbf{B}'\mathbf{PB}.$$

4.2.3 Write down the analogues of (4.2.12), (4.2.13), (4.2.14), (4.2.15), (4.2.18), (4.2.19), (4.2.22), and (4.2.25) in terms of $\mathbf{C}$, $\mathbf{D}$, and $\mathbf{I}$ in place of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{T}$. Show that these may be derived from (4.2.32) and (4.2.33) in the same way that their analogues were derived from (4.2.10) and (4.2.11). Show also that they may be derived alternatively by the direct substitution of $\mathbf{A} = \mathbf{RC}$, $\mathbf{B} = \mathbf{DR}^{-1}$, and $\mathbf{T} = \mathbf{RR}$ into the analogous formulas.

4.2.4 Show that $\mathbf{Q}_{21}\mathbf{Q}_{11}^{-1} = \mathbf{B}_{21}\mathbf{B}_{11}^{-1} = -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} = \mathbf{D}_{21}\mathbf{D}_{11}^{-1} = -\mathbf{C}_{22}^{-1}\mathbf{C}_{21}$ and dually that $\mathbf{P}_{12}\mathbf{P}_{22}^{-1} = \mathbf{A}_{21}'\mathbf{A}_{22}'^{-1} = -\mathbf{B}_{11}'^{-1}\mathbf{B}_{21}' = \mathbf{C}_{21}'\mathbf{C}_{22}'^{-1} = \mathbf{D}_{11}'^{-1}\mathbf{D}_{21}'$.

4.2.5 Show directly that $\mathbf{U}_2 - \mathbf{Q}_{21}\mathbf{Q}_{11}^{-1}\mathbf{U}_1$ is orthogonal to $\mathbf{U}_1$, and hence must be $\mathbf{U}_{2.1}$. Then show that the inner product matrix of $\mathbf{U}_2 - \mathbf{Q}_{21}\mathbf{Q}_{11}^{-1}\mathbf{U}_1$ is $\mathbf{Q}_{21}\mathbf{Q}_{11}^{-1}\mathbf{Q}_{12}$ and deduce that $\mathbf{Q}_{22.1} = \mathbf{Q}_{22} - \mathbf{Q}_{21}\mathbf{Q}_{11}^{-1}\mathbf{Q}_{12}$.

4.2.6 Show that $\mathbf{U}_2 = \mathbf{B}_{21}\mathbf{B}_{11}^{-1}\mathbf{U}_1 + \mathbf{B}_{22}\mathbf{U}_2^*$, and by looking at the expression for U_{s+1} given by the first row of this equality, show that the first row of $\mathbf{B}_{21}\mathbf{B}_{11}^{-1}$ consists of

$$[-a_{\overline{s+1}\,1}, -a_{\overline{s+1}\,2}, \ldots, -a_{\overline{s+1}\,s}],$$

i.e., of the negatives of the first s elements of row $s+1$ of $\mathbf{A}$. Note that the remaining elements are known, for $a_{\overline{s+1}\,\overline{s+1}} = 1$ and $a_{\overline{s+1}\,t} = 0$ for $t > s+1$.

4.2.7 By writing $\mathbf{B}_{21}\mathbf{B}_{11}^{-1} = \mathbf{B}_{21}\mathbf{A}_{11}$, show directly that the last column of $\mathbf{B}_{21}\mathbf{B}_{11}^{-1}$ consists of $[b_{\overline{s+1}\,s}, b_{\overline{s+2}\,s}, \ldots, b_{p\,s}]$, i.e., of the elements below the diagonal element of the $(s+1)$st column of $\mathbf{B}$.

4.2.8 From Exercises 4.2.6 and 4.2.7 deduce that

$$a_{\overline{s+1}\,s} = -b_{\overline{s+1}\,s} \quad \text{for} \quad s = 1, 2, \ldots, p-1.$$

4.2.9 What are the dual results of Exercises 4.2.6 and 4.2.7?

4.2.10 Show that $\mathbf{Q}_{22.1} = \mathbf{Q}_{22} - \mathbf{H}_{21}\mathbf{Q}_{11}\mathbf{H}_{21}'$ and dually that $\mathbf{Q}_{11}^{-1} = \mathbf{P}_{11} - \mathbf{H}_{21}'\mathbf{P}_{22}\mathbf{H}_{21}$.

4.3.1 How many additions or subtractions and how many multiplications or divisions are required to compute $\mathbf{B}$ and $\mathbf{T}$ by the elimination procedure of Section 4.3.1? How many are required to compute $\mathbf{B}$, $\mathbf{T}$, and $\mathbf{A}$? How many are required to compute $\mathbf{B}$, $\mathbf{T}$, $\mathbf{A}$, and $\mathbf{Q}^{-1}$? Compare the number of multiplications and divisions in each of these three categories with the number of multiplications required to multiply two $p \times p$ matrices, especially for p large.

4.3.2 The elimination method of Section 4.3.1 may be modified into the *square root method.* This approach bypasses $\mathbf{B}$ and $\mathbf{A}$ and finds $\mathbf{D}$ and $\mathbf{C}$ directly. Finding $\mathbf{D}$ by this method may be described as finding the coefficients in the set of equations

$$\begin{aligned}
d_{11}U_1^{**} &= U_1^* = U_1 \\
d_{22}U_2^{**} &= U_2^* = U_2 - d_{21}U_1^{**} \\
&\cdots \\
d_{pp}U_p^{**} &= U_p^* = U_p - d_{p1}U_1^{**} - d_{p2}U_2^{**} \cdots - d_{p\,\overline{p-1}}U_{p-1}^{**}.
\end{aligned}$$

Show that the column of coefficients of U_1^{**} may be computed directly from $\mathbf{Q}$ as $(U_s, U_1)/(U_1, U_1)^{1/2}$ for $s = 2, 3, \ldots, p$. Show that this column multiplied by its transpose may be subtracted from the last $p-1$ rows and columns of $\mathbf{Q}$ to yield $\mathbf{Q}_{22.1}$ with $s = 1$, which completes the first stage of the square root version of the elimination method of finding $\mathbf{D}$. Describe the remaining steps required to find $\mathbf{D}$, $\mathbf{C}$, and $\mathbf{Q}^{-1}$ by this method, illustrating your answer with the matrix $\mathbf{Q}$ of Example 4.3.

4.3.3 *Triangularization.* The process of successive orthogonalization leads to a somewhat different set of computational procedures when the computations are based directly on the coordinates of the set $\mathbf{U}$ relative to an orthonormal basis $\mathbf{W}$ of $\mathscr{E}$ rather than on the inner product matrix $\mathbf{Q}$ of the set $\mathbf{U}$. As in the text, there is a choice between expressions involving $\mathbf{U}^*$, and hence $\mathbf{A}$ and $\mathbf{B}$, or expressions involving $\mathbf{U}^{**}$, and hence $\mathbf{C}$ and $\mathbf{D}$. Because of the convenience of having orthogonal matrices represent relations between the orthonormal sets $\mathbf{W}$ and $\mathbf{U}^{**}$, the latter representation is used here.

Suppose that $\mathbf{E}$ is any $s \times p$ matrix of maximum rank s for some s on $1 \leq s \leq p$. Then $\mathbf{E}$ may viewed as defining a set of linearly independent vectors $\mathbf{U} = \mathbf{EW}$ where $\mathbf{W}$ is an orthonormal basis of $\mathscr{E}$. The orthonormal set $\mathbf{U}^{**}$ produced by successive orthogonalization of $\mathbf{U}$ may be expressed as $\mathbf{U}^{**} = \mathbf{GW}$ where $\mathbf{G}$ is an $s \times p$ matrix obeying $\mathbf{GG}' = \mathbf{I}$. Since $\mathbf{U} = \mathbf{D}_{11}\mathbf{U}^{**}$ where $\mathbf{D}_{11}$ is lower triangular,

$$\mathbf{E} = \mathbf{D}_{11}\mathbf{G} \tag{4.5.1}$$

which may be called the *triangularization* of $\mathbf{E}$.

There are various ways to compute $\mathbf{G}$ and $\mathbf{D}_{11}$ of which three will be sketched here. The reader is asked to supply missing details as an exercise.

a) *Triangularization by assimilation.* Suppose that the rows of $\mathbf{E}$ are denoted by $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_s$, that the rows of $\mathbf{G}$ are denoted by $\mathbf{G}_1, \mathbf{G}_2, \ldots, \mathbf{G}_s$, and that the (i,j) element of $\mathbf{D}_{11}$ is denoted by d_{ij} with $d_{ij} = 0$ for $j > i$. From (4.5.1), $\mathbf{E}_1 = d_{11}\mathbf{G}_1$ and, since $\mathbf{G}_1\mathbf{G}_1' = 1$, $d_{11} = \pm(\mathbf{E}_1\mathbf{E}_1')^{1/2}$, thus determining $\mathbf{G}_1$. Now suppose that the first s rows of both $\mathbf{D}_{11}$ and $\mathbf{G}$ are determined. Then $\mathbf{G}_{r+1}\mathbf{W}$ is found by removing from $\mathbf{E}_{r+1}\mathbf{W}$ components along $\mathbf{G}_1\mathbf{W}, \mathbf{G}_2\mathbf{W}, \ldots, \mathbf{G}_r\mathbf{W}$ and scaling for unit length. Thus

$$k\mathbf{G}_{r+1} = \mathbf{E}_{r+1} - (\mathbf{E}_{r+1}\mathbf{G}_1')\mathbf{G}_1 - (\mathbf{E}_{r+1}\mathbf{G}_2')\mathbf{G}_2 \cdots - (\mathbf{E}_{r+1}\mathbf{G}_r')\mathbf{G}_r, \qquad (4.5.2)$$

and after computing the right side of (4.5.2) one finds k from $k^2 = (\mathbf{G}_{r+1}\mathbf{G}_{r+1}')$. Note that row $r+1$ of $\mathbf{D}_{11}$ is given by $[\mathbf{E}_{r+1}\mathbf{G}_1', \mathbf{E}_{r+1}\mathbf{G}_2', \ldots, \mathbf{E}_{r+1}\mathbf{G}_r', k]$.

b) *Triangularization by elimination.* After finding $\mathbf{G}_1$, the components of each of $\mathbf{E}_2, \mathbf{E}_3, \ldots, \mathbf{E}_s$ along $\mathbf{G}_1$ may be removed yielding, say, $\mathbf{E}_{2.1}, \mathbf{E}_{3.1}, \ldots, \mathbf{E}_{s.1}$. After scaling, $\mathbf{E}_{2.1}$ yields $\mathbf{G}_2$. Components of $\mathbf{E}_{3.1}, \mathbf{E}_{4.1}, \ldots, \mathbf{E}_{s.1}$ along $\mathbf{G}_2$ are then removed, and so on. Further details are left to the reader.

c) *Triangularization by Householder transformations* (Householder, 1958). This method produces an $s \times p$ matrix $\mathbf{D}^*$ and a $p \times p$ orthogonal matrix $\mathbf{G}^*$ such that $\mathbf{E} = \mathbf{D}^*\mathbf{G}^*$ and the (i,j) elements of $\mathbf{D}^*$ are zero for $j > i$. Dropping the last $p - s$ columns of $\mathbf{D}^*$ yields $\mathbf{D}_{11}$ and dropping the last $p - s$ rows of $\mathbf{G}^*$ yields $\mathbf{G}$ where (4.5.1) holds. Because the last $p - s$ columns of $\mathbf{D}^*$ are all $\mathbf{0}$, dropping them along with the last $p - s$ rows of $\mathbf{G}^*$ does not affect the product $\mathbf{D}^*\mathbf{G}^*$, but the truncated product is then in the form required of $\mathbf{D}_{11}\mathbf{G}$ and so must be $\mathbf{D}_{11}\mathbf{G}$. $\mathbf{D}^*$ and $\mathbf{G}^*$ are found using a sequence of $p \times p$ orthogonal matrices $\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_s$ and successively writing $\mathbf{E} = (\mathbf{E})(\mathbf{I}) = (\mathbf{E}\mathbf{H}_1')(\mathbf{H}_1) = (\mathbf{E}\mathbf{H}_1'\mathbf{H}_2')(\mathbf{H}_2\mathbf{H}_1) = \cdots = (\mathbf{E}\mathbf{H}_1'\mathbf{H}_2' \cdots \mathbf{H}_s')(\mathbf{H}_s\mathbf{H}_{s-1} \cdots \mathbf{H}_1)$ where $\mathbf{D}^* = \mathbf{E}\mathbf{H}_1'\mathbf{H}_2' \cdots \mathbf{H}_s'$ and $\mathbf{G}^* = \mathbf{H}_s\mathbf{H}_{s-1} \cdots \mathbf{H}_1$. $\mathbf{H}_1$ is chosen to be an elementary orthogonal matrix such that $\mathbf{E}\mathbf{H}_1'$ has a first row whose only nonzero element is its first. In other words, the linear transformation $\mathbf{X} \to \mathbf{X}\mathbf{H}_1'$ carries $\mathbf{E}_1 \to [k, 0, 0, \ldots, 0]$, or, since such transformations are self-inverting, carries $[k, 0, 0, \ldots, 0] \to \mathbf{E}_1$. The construction of such an $\mathbf{H}_1$ was described in Section 3.4. More generally, $\mathbf{E}\mathbf{H}_1'\mathbf{H}_2' \cdots \mathbf{H}_r'$ has the property that its (i,j) elements are zero for $j > i$ and $i = 1, 2, \ldots, r$. Then $\mathbf{H}_{r+1}$ is chosen to be a matrix whose first r rows and columns are like an identity matrix but whose remaining part is an $(s - r) \times (s - r)$ elementary orthogonal matrix such that the superdiagonal elements of row $r + 1$ of $\mathbf{E}\mathbf{H}_1'\mathbf{H}_2' \cdots \mathbf{H}_r'$ are eliminated when multiplied by $\mathbf{H}_{r+1}'$. The definition of $\mathbf{H}_{r+1}$ for $r = 1, 2, \ldots, s - 1$ is essentially the same as that of $\mathbf{H}_1$ in a different context. The reader is invited to supply further details.

4.3.4 In addition to the sweep operator SWP[i], define the following operators on $p \times p$ matrices:

IOP: the identity operation.
NOP: the operator which changes all signs.
MOP[i]: the operator which changes the signs of the off-diagonal elements in row and column i.
POP[i]: the operator which changes the signs of the (i,i) element and the elements (j,k) with $j \neq i$ and $k \neq i$.

Suppose that the operator defined by the successive application of the operators $A, B, \ldots, C$ is denoted by $C \cdots BA$. Show that

$$\begin{aligned} \text{SWP}[i]\text{SWP}[i] &= \text{MOP}[i] \\ \text{MOP}[i]\text{SWP}[i] &= \text{SWP}[i]\text{MOP}[i] \\ \text{SWP}[i]\text{POP}[i] &= \text{NOPSWP}[i] \\ \text{SWP}[i]\text{NOP} &= \text{POP}[i]\text{SWP}[i] \\ \text{NOP} &= \text{SWP}[i]\text{NOPSWP}[i]. \end{aligned}$$

What are the analogous identities involving RSW[i] instead of SWP[i]? Show that the SWP[i] and RSW[j] commute except for confusion of signs. Describe this sign confusion in detail.

4.3.5 *Inversion of nonsymmetric matrices.* Suppose that $\mathbf{G}$ is a $p \times p$ nonsingular matrix and is regarded as carrying the basis $\mathbf{V}$ of a vector space $\mathscr{E}$ into a basis $\mathbf{W} = \mathbf{G}\mathbf{V}$ of $\mathscr{E}$. Suppose that $\mathbf{U}^{(s)}$ is the set of vectors $W_1, W_2, \ldots, W_s, V_{s+1}, \ldots, V_p$, and suppose that $\mathbf{U}^{(s)}$ is a basis of $\mathscr{E}$ for $s = 1, 2, \ldots, p-1$. The basis $\mathbf{V}$ may be modified by stages into $\mathbf{W}$ via

$$\mathbf{V} = \mathbf{U}^{(0)} \to \mathbf{U}^{(1)} \to \mathbf{U}^{(2)} \to \cdots \to \mathbf{U}^{(p)} = \mathbf{W}.$$

Define $\mathbf{H}^{(s)}$ to be a matrix with (i,j) element $h_{ij}^{(s)}$ whose row i for $i = 1, 2, \ldots, s$ expresses $-V_i$ in terms of $\mathbf{U}^{(s)}$ and for $i = s+1, s+2, \ldots, p$ expresses W_i in terms of $\mathbf{U}^{(s)}$, i.e.,

$$-V_i = \sum_{j=1}^{s} h_{ij}^{(s)} W_j + \sum_{j=s+1}^{p} h_{ij}^{(s)} V_j$$

for $i = 1, 2, \ldots, p$, and

$$W_i = \sum_{j=1}^{s} h_{ij}^{(s)} W_j + \sum_{j=s+1}^{p} h_{ij}^{(s)} V_j$$

for $i = s+1, s+2, \ldots, p$. Show that $\mathbf{H}^{(s+1)} = \text{SWP}[s+1]\mathbf{H}^{(s)}$, and deduce that $\mathbf{G}^{-1} = -\text{SWP}[1, 2, \ldots, p]\mathbf{G}$.

4.3.6 *Solution of linear equations.* The formulation of Exercise 4.3.5 will be continued here. Suppose that $\mathbf{Y}$ is a $q \times p$ matrix; consider finding a $q \times p$ matrix $\mathbf{X}$ such that

$$\mathbf{X}\mathbf{G} = \mathbf{Y}.$$

The problem is thus to solve a collection of q sets of linear equations each having p unknowns and each having the same matrix of coefficients $\mathbf{G}$. In vector terms, $\mathbf{Y}\mathbf{V}$ determines q points in $\mathscr{E}$, and the problem is to express these q points as $\mathbf{X}\mathbf{W}$ in terms of the basis $\mathbf{W} = \mathbf{G}\mathbf{V}$. Suppose these q points are $\mathbf{Z}^{(s)}\mathbf{U}^{(s)}$ in terms of the basis $\mathbf{U}^{(s)}$. Then passage from $\mathbf{Y}$ to $\mathbf{X}$ may be achieved via

$$\mathbf{Y} = \mathbf{Z}^{(0)} \to \mathbf{Z}^{(1)} \to \mathbf{Z}^{(2)} \to \cdots \to \mathbf{Z}^{(p)} = \mathbf{X}.$$

Show that

$$\begin{bmatrix} \mathbf{H}^{(s+1)} \\ \mathbf{Z}^{(s+1)} \end{bmatrix} = \text{SWP}[s+1] \begin{bmatrix} \mathbf{H}^{(s)} \\ \mathbf{Z}^{(s)} \end{bmatrix},$$

so that

$$\begin{bmatrix} -\mathbf{G}^{-1} \\ \mathbf{X} \end{bmatrix} = \text{SWP}[1, 2, \ldots, p] \begin{bmatrix} \mathbf{G} \\ \mathbf{Y} \end{bmatrix}.$$

Note that although the sweep operator was defined in Section 4.3.2 for square matrices, the same definition is used here for a $(p+q) \times p$ matrix.

4.3.7 The computational methods in Exercises 4.3.5 and 4.3.6 may not be applied with arbitrary $\mathbf{G}$ because the sets $\mathbf{U}^{(s)}$ are not necessarily linearly independent. Show that the method may be applied for some rearrangement of the rows of $\mathbf{G}$, i.e., that for some $W_{\pi(1)}$ the set $W_{\pi(1)}, V_2, \ldots, V_p$ must be a basis such that given such a $W_{\pi(1)}$ there must be a $W_{\pi(2)}$ such that the set $W_{\pi(1)}, W_{\pi(2)}, V_3, \ldots, V_p$ is a basis, and so on. Show that this leads to performance of a sweeping operation at stage s where the row $\pi(s)$ and the column s selected for special treatment are in general different, i.e., the definition of sweep operations must be slightly generalized. It may be further shown that $\pi(s)$ may be chosen to be any i such that $h_{is}^{(s-1)} \neq 0$, where $h_{is}^{(s-1)}$ refers to the generalized process with rows taken in the order $\pi(1), \pi(2), \ldots, \pi(s-1)$. It should now be clear how to carry out matrix inversion and solution of linear equations for arbitrary $\mathbf{G}$.

4.3.8 Describe in detail the set of calculations required to carry out the ASM$[p+1, p+2, \ldots, p+r; 1, 2, \ldots, s]$ operator using the scheme of r successive stages.

4.3.9 Suppose that RSM$[p+1, p+2, \ldots, p+r; 1, 2, \ldots, s]$ denotes the operator inverse of ASM$[p+1, p+2, \ldots, p+r; 1, 2, \ldots, s]$. Describe the three characterizations of the RSM operator analogous to the three characterizations of the ASM operator given in Section 4.3.3.

4.3.10 What are the special features of the output of

$$\text{MST}[1, 2, \ldots, s]\text{MST}[s+1, s+2, \ldots, p][\mathbf{Q}, \mathbf{I}]$$

and

$$\text{MST}[1]\text{MST}[2] \ldots \text{MST}[p][\mathbf{Q}, \mathbf{I}]?$$

4.4.1 Suppose that the vectors U_1, U_2, U_3, U_4, U_5 are all nonzero and span a two-dimensional subspace of a three-dimensional Euclidean space. How many and which of the $U_1^*, U_2^*, \ldots, U_5^*$ produced by successive orthogonalization are $\emptyset$? Suppose that $\mathbf{Q}$ denotes the 5×5 inner product matrix of $U_1, U_2, \ldots, U_5$. Show that $\mathbf{Q}$ is a positive semi-definite symmetric matrix of rank 2.

4.4.2 Describe a computing procedure to determine the rank of a symmetric positive semi-definite matrix $\mathbf{Q}$.

4.4.3 Suppose that a positive semi-definite symmetric matrix

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix}$$

of rank f is regarded as an inner product matrix relative to a basis

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{bmatrix} \text{ of a vector space } \mathscr{E},$$

where the partitions refer as usual to s and $p-s$ rows and columns. Suppose that $\mathscr{U}_1$ denotes the intersection of the subspace $\mathscr{U}$ of $\mathscr{E}$ of vectors with zero norm with the subspace spanned by $\mathbf{U}_1$. Suppose that $\mathscr{V}_1$ is any complementary subspace to $\mathscr{U}_1$ in the subspace spanned by $\mathbf{U}_1$. Show that $\mathscr{U}_1$ consists of vectors with zero norm and $\mathscr{V}_1$ consists of vectors with nonzero norm. Show that the dimension f_1 of $\mathscr{V}_1$ is the same as the rank of $\mathbf{Q}_{11}$ and that $f_1 \leq s$ and $f_1 \leq f$. Show that any V in $\mathscr{E}$ decomposes

uniquely into a component in $\mathscr{V}_1$ and a component orthogonal to $\mathscr{V}_1$ (cf. Exercise 3.6.5), and define the latter set of components of $\mathbf{U}_2$ to be $\mathbf{U}_{2.1}$. Show that the definitions of $\mathbf{U}_{2.1}$ differ for different $\mathscr{V}_1$ but only by vectors of zero norm. Show that the inner product matrix $\mathbf{Q}_{22.1}$ of $\mathbf{U}_{2.1}$ is consequently well-defined in the sense of not depending on the choice of $\mathscr{V}_1$.

4.4.4 Following the notation of Exercise 4.4.3, suppose that $\dot{\mathbf{Q}}_{11}$ is a pseudoinverse of $\mathbf{Q}_{11}$ defined as in (3.6.8) with the roles of $\mathbf{Q}$, $\mathscr{U}$, and $\mathscr{V}$ played here by $\mathbf{Q}_{11}$, $\mathscr{U}_1$, and $\mathscr{V}_1$. Show that

$$\mathbf{Q}_{22.1} = \mathbf{Q}_{22} - \mathbf{Q}_{21}\dot{\mathbf{Q}}_{11}\mathbf{Q}_{12}.$$

In particular, $\dot{\mathbf{Q}}_{11}$ may be taken as (4.4.1) with any choice of the null columns of $\mathbf{B}$, in which case

$$\mathbf{Q}_{22.1} = \mathbf{B}_{22}\mathbf{T}_{22}\mathbf{B}_{22}'.$$

In another particular case $\dot{\mathbf{Q}}_{11}$ may be *the* pseudoinverse of $\mathbf{Q}_{11}$ (cf. Exercise 3.6.4).

4.4.5 Suppose that the first s rows and columns of $\mathbf{Q}$ in Exercise 4.4.3 are further partitioned into $s = f_1 + (s - f_1)$ so that

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{111} & \mathbf{Q}_{112} & \mathbf{Q}_{121} \\ \mathbf{Q}_{121} & \mathbf{Q}_{122} & \mathbf{Q}_{122} \\ \mathbf{Q}_{211} & \mathbf{Q}_{212} & \mathbf{Q}_{22} \end{bmatrix}$$

where

$$\mathbf{Q}_{11} = \begin{bmatrix} \mathbf{Q}_{111} & \mathbf{Q}_{112} \\ \mathbf{Q}_{121} & \mathbf{Q}_{122} \end{bmatrix}$$

and

$$\mathbf{Q}_{21} = [\mathbf{Q}_{211}, \mathbf{Q}_{212}] = \begin{bmatrix} \mathbf{Q}_{121} \\ \mathbf{Q}_{122} \end{bmatrix}' = \mathbf{Q}_{12}'.$$

Suppose also that the first s rows and columns are arranged such that $\mathbf{Q}_{111}$ has full rank f_1.

Show that SWP$[1, 2, \ldots, f_1]\mathbf{Q}$ has the form

$$\begin{bmatrix} -\mathbf{Q}_{111}^{-1} & \mathbf{H}_{112} & \mathbf{H}_{121} \\ \mathbf{H}_{121} & \mathbf{0} & \mathbf{0} \\ \mathbf{H}_{211} & \mathbf{0} & \mathbf{Q}_{22.1} \end{bmatrix}$$

where $[\mathbf{H}_{211}, \mathbf{0}] = \begin{bmatrix} \mathbf{H}_{121} \\ \mathbf{0} \end{bmatrix}'$ is one choice for $\mathbf{H}_{21}$ in Exercise 4.4.3 and $\mathbf{Q}_{22.1}$ is the uniquely defined inner product matrix of $\mathbf{U}_{2.1}$ arising from any choice of $\mathbf{H}_{21}$. Note that in computational practice a subset of f_1 of the first s rows and columns having a full rank inner product matrix may not be known. In this case one need only set out to perform SWP[1], SWP[2], ..., SWP[s] in order, omitting any operation which is undefined because of a zero element in the pivotal diagonal position. As a result one will carry out just f_1 sweeping operations and arrive at a matrix like SWP$[1, 2, \ldots, f_1]\mathbf{Q}$ above, except that generally a different subset of f_1 of the first s rows and columns will have been swept.

4.4.6 Supposing that $\mathbf{A}$ is any $p \times q$ matrix of rank $q < p$, show how the idea of triangularization may be used to find the pseudoinverse of $\mathbf{A}$ defined in Exercise 3.4.11.

CHAPTER 5

THE RELATION BETWEEN TWO INNER PRODUCTS

5.1 BASIC THEORY

The statistical techniques of canonical correlation analysis, multiple linear discriminant analysis, and principal component analysis are the most sophisticated methods of multivariate statistical analysis considered in this book. Each of these techniques may be naturally viewed as relating two different inner products on the same vector space $\mathscr{E}$. Consideration of such relationships leads to a mathematical theory of *eigenvalues* and *eigenvectors* (sometimes called *characteristic values* and *characteristic vectors*, *proper values* and *proper vectors*, or *latent roots* and *latent vectors*) in Euclidean spaces.

Suppose that π_1 and π_2 are symbols for two different inner product functions on a p-dimensional vector space $\mathscr{E}$, and suppose that the inner product of U and V according to π_i is denoted by $(U, V)_i$ for $i = 1, 2$. One might ask: for what vectors V in $\mathscr{E}$ is the ratio

$$\lambda = (V, V)_1/(V, V)_2 \tag{5.1.1}$$

maximized or minimized? The answer to this question and many related questions comes from the following theory which indicates that in terms of a specially chosen basis $\mathbf{W}$ of $\mathscr{E}$ the relationship between π_1 and π_2 is simply and clearly displayed. Until stated otherwise, π_1 and π_2 will be assumed to have full rank p.

The theory depends on the following simple lemmas.

Lemma 5.1.1. *The supremum of the ratio* (5.1.1) *is achieved for at least one vector V in $\mathscr{E}$.*

In proving this lemma, one need only consider vectors V such that $(V, V)_2 = 1$, for the ratio (5.1.1) is invariant under multiplication of V by a scalar; consequently any value of (5.1.1) taken on by some U in $\mathscr{E}$ is also taken on by $V = (U, U)_2^{-1/2}U$, satisfying $(V, V)_2 = 1$. In analytic terms, with reference to an orthonormal basis of π_2, the problem is to find $\boldsymbol{\alpha}$ which maximizes $\boldsymbol{\alpha}\mathbf{Q}_1\boldsymbol{\alpha}'$ subject

to the restriction $\boldsymbol{\alpha}\boldsymbol{\alpha}' = 1$. The existence of such an $\boldsymbol{\alpha}$ follows because the function $\boldsymbol{\alpha}\mathbf{Q}_1\boldsymbol{\alpha}'$ is continuous and finite over the closed domain $\boldsymbol{\alpha}\boldsymbol{\alpha}' = 1$.

Lemma 5.1.2. *Suppose that W_1 is a vector which maximizes* (5.1.1) *over all V in $\mathscr{E}$. Then any vector orthogonal to W_1 according to π_1 is also orthogonal to W_1 according to π_2 and vice versa, i.e., the $(p-1)$-dimensional subspace $\mathscr{V}_1$ of $\mathscr{E}$ orthogonal to W_1 is the same for π_1 and π_2.*

The lemma will be proved by contradiction. Suppose that $\mathscr{V}_{1i}$ is the subspace of $\mathscr{E}$ orthogonal to W_1 according to π_i, for $i = 1, 2$, and suppose that $\mathscr{V}_{11}$ and $\mathscr{V}_{12}$ are different. This would imply that a vector W_1^* orthogonal to $\mathscr{V}_{11}$ according to π_2 is not a multiple of W_1. Thus W_1 may be decomposed into

$$W_1 = W_{11} + W_{12}, \tag{5.1.2}$$

where W_{11} lies along W_2^*, W_{12} lies in $\mathscr{V}_{11}$, and $(W_{11}, W_{12})_2 = 0$. Also, since $\mathscr{V}_{11}$ and $\mathscr{V}_{12}$ are different, W_{12} is different from $\emptyset$. Thus

$$(W_1, W_1)_2 = (W_{11}, W_{11})_2 + (W_{12}, W_{12})_2,$$

whence

$$(W_{11}, W_{11})_2 < (W_1, W_1)_2. \tag{5.1.3}$$

On the other hand, according to π_1, W_1 and W_{12} are orthogonal, so that $(W_{11}, W_{11})_1 = (W_1, W_1)_1 + (W_{12}, W_{12})_1$, whence

$$(W_{11}, W_{11})_1 > (W_1, W_1)_1. \tag{5.1.4}$$

From (5.1.3) and (5.1.4) it follows that

$$\frac{(W_{11}, W_{11})_1}{(W_{11}, W_{11})_2} > \frac{(W_1, W_1)_1}{(W_1, W_1)_2}, \tag{5.1.5}$$

which is impossible since W_1 maximizes (5.1.1). Thus the lemma is proved by contradiction.

Lemmas 5.1.1 and 5.1.2 are used to prove:

Theorem 5.1.1. *There exists a basis $\mathbf{W}$ of $\mathscr{E}$ which is orthogonal according to both π_1 and π_2.*

The existence of such a basis is demonstrated by constructing it. For W_1 take any vector which maximizes (5.1.1). Suppose that $\mathscr{V}_1$ is the subspace orthogonal to W_1 according to both π_1 and π_2. Now consider π_1 and π_2 as inner products over $\mathscr{V}_1$. Choose W_2 to be any vector in $\mathscr{V}_1$ which maximizes (5.1.1) over V in $\mathscr{V}_1$. Suppose that $\mathscr{V}_2$ is the subspace of $\mathscr{V}_1$ orthogonal to W_2 according to both π_1 and π_2. Choose W_3 to be any vector in $\mathscr{V}_2$ which maximizes (5.1.1) over V in $\mathscr{V}_2$. Continue thus until $\mathbf{W}$ with the required property is constructed.

The vectors $W_1, W_2, \ldots, W_p$ of a basis $\mathbf{W}$ which is orthogonal relative to both π_1 and π_2 will be called *eigenvectors of π_1 relative to π_2*. The corresponding

$$\lambda_i = (W_i, W_i)_1/(W_i, W_i)_2 \tag{5.1.6}$$

will be called *eigenvalues of π_1 relative to π_2*. The constructive proof of Theorem 5.1.1 leaves some latitude in the choice of eigenvectors. Theorem 5.1.2 will describe in detail the range of possible bases $\mathbf{W}$ allowed by this constructive proof and will show that the construction always produces the same set of ordered eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_p$. Theorem 5.1.3 will show that, apart from order, any basis of eigenvectors must belong to the class of such bases specified in Theorem 5.1.2, and consequently that the set of eigenvalues of π_1 relative to π_2, again apart from order, is uniquely determined.

If $[W_1, W_2, \ldots, W_p]$ is a basis orthogonal relative to both π_1 and π_2 and if $[c_1, c_2, \ldots, c_p]$ is any set of nonzero coefficients, then $[c_1W_1, c_2W_2, \ldots, c_pW_p]$ is also a basis orthogonal relative to both π_1 and π_2. The latter basis will be called a *scaled* version of the former. It follows that a basis of eigenvectors may be unique only up to arbitrary changes of scale. In other words, it is possible only that the set of p one-dimensional subspaces spanned by $W_1, W_2, \ldots, W_p$ can be uniquely determined. Necessary and sufficient conditions that such uniqueness obtains are implied by the following theorem.

Theorem 5.1.2. *Suppose that $\lambda_1, \lambda_2, \ldots, \lambda_p$ and $W_1, W_2, \ldots, W_p$ denote a particular set of eigenvalues and corresponding eigenvectors of π_1 relative to π_2 found by the construction procedure used in the proof of Theorem* 5.1.1. *Then any other realization of the construction procedure (resulting from an alternative choice of W_i at any stage) yields the same set of $\lambda_1, \lambda_2, \ldots, \lambda_p$. Suppose that $s_1 < s_2 < \cdots < s_r$ denote the set of indices such that*

$$\begin{aligned} \lambda_1 &= \cdots = \lambda_{s_1} \\ &> \lambda_{s_1+1} = \cdots = \lambda_{s_2} \\ &> \lambda_{s_2+1} = \cdots = \lambda_{s_3} \\ &\cdots \\ &> \lambda_{s_r+1} = \cdots = \lambda_p. \end{aligned} \tag{5.1.7}$$

Then the subspaces $\mathscr{W}_1, \mathscr{W}_2, \ldots, \mathscr{W}_{r+1}$ spanned by $[W_1, \ldots, W_{s_1}]$, $[W_{s_1+1}, \ldots, W_{s_2}], \ldots, [W_{s_r+1}, \ldots, W_p]$ are uniquely determined under any realization of the construction procedure. The inner product π_1 is a multiple of π_2 over each of $\mathscr{W}_1, \mathscr{W}_2, \ldots, \mathscr{W}_{r+1}$. The sets $[W_1, \ldots, W_{s_1}]$, $[W_{s_1+1}, \ldots, W_{s_2}], \ldots, [W_{s_r+1}, \ldots, W_p]$ are any orthogonal bases of their corresponding subspaces. In particular, if

$$\lambda_1 > \lambda_2 > \cdots > \lambda_p, \tag{5.1.8}$$

then $W_1, W_2, \ldots, W_p$ are uniquely determined up to a set of scale factors.

The proof of Theorem 5.1.2 is omitted, but requires only a careful following-out of the constructive proof of Theorem 5.1.1. The difference is that one must now check into the uniqueness of W_s at each stage and show that the range of possibilities is as specified by the theorem in terms of subspaces $\mathscr{W}_1, \mathscr{W}_2, \ldots, \mathscr{W}_{r+1}$ over which the ratio of norms is constant.

Theorem 5.1.3. *The class of bases $\mathbf{W}$ defined as in Theorem* 5.1.2 *may be characterized as the only bases which are orthogonal relative to both π_1 and π_2. Consequently, the eigenvalues of π_1 relative to π_2 are a well-defined set of numbers.*

The proof will show that any contemplated basis element $\boldsymbol{\delta}\mathbf{W}$ must be a member of a basis in the permitted class, where $\boldsymbol{\delta} = [\delta_1, \delta_2, \ldots, \delta_p]$ and $\mathbf{W}$ is a basis in the permitted class which is orthonormal relative to π_2. The subspace of vectors orthogonal to $\boldsymbol{\delta}\mathbf{W}$ according to π_2 consists of those $\boldsymbol{\alpha}\mathbf{W}$ satisfying

$$\sum_{i=1}^{p} \delta_i\alpha_i = 0. \tag{5.1.9}$$

Similarly

$$\sum_{i=1}^{p} \lambda_i\delta_i\alpha_i = 0 \tag{5.1.10}$$

is the condition for $\boldsymbol{\delta}\mathbf{W}$ and $\boldsymbol{\alpha}\mathbf{W}$ to be orthogonal relative to π_1. Since $\boldsymbol{\delta}\mathbf{W}$ is a member of a doubly orthogonal basis, the vectors $\boldsymbol{\alpha}\mathbf{W}$ satisfying (5.1.9) must coincide with those satisfying (5.1.10). When the λ_i are distinct, (5.1.9) and (5.1.10) define the same subspaces only when all the δ_i are zero except one, i.e., $\boldsymbol{\delta}\mathbf{W}$ is simply an element of $\mathbf{W}$ rescaled. If equality holds for a set of λ_i, then the corresponding δ_i for this set may be nonzero while the remaining δ_i are zero. Again $\boldsymbol{\delta}\mathbf{W}$ is an element of a permitted basis. Thus Theorem 5.1.3 is proved.

It is often convenient to fix the arbitrary scaling of a basis $\mathbf{W}$ of eigenvectors by making it orthonormal according to π_1 or π_2. For example, if $\mathbf{W}$ is scaled so that $(W_i, W_i)_2 = 1$ for $i = 1, 2, \ldots, p$, then $(W_i, W_i)_1 = \lambda_i$ for $i = 1, 2, \ldots, p$, i.e., $\mathbf{W}$ has inner product matrix $\mathbf{L}$ according to π_1 where $\mathbf{L}$ is a diagonal matrix with diagonal elements $\lambda_1, \lambda_2, \ldots, \lambda_p$. In this case, if $U = \boldsymbol{\alpha}\mathbf{W}$ and $V = \boldsymbol{\beta}\mathbf{W}$, then

$$(U, V)_1 = \sum_{i=1}^{p} \lambda_i\alpha_i\beta_i, \tag{5.1.11}$$

whereas

$$(U, V)_2 = \sum_{i=1}^{p} \alpha_i\beta_i. \tag{5.1.12}$$

Formula (5.1.11) shows that π_1 and π_2 are differently weighted sums of p semi-definite inner products each of rank 1, i.e., defining

$$(U, V)_{(i)} = \alpha_i\beta_i \quad \text{for} \quad i = 1, 2, \ldots, p, \tag{5.1.13}$$

it follows that

$$(U, V)_1 = \sum_{i=1}^{p} \lambda_i (U, V)_{(i)} \tag{5.1.14}$$

and

$$(U, V)_2 = \sum_{i=1}^{p} (U, V)_{(i)}.$$

The requirement that π_1 and π_2 both have full rank will now be relaxed. In the full rank case, each λ_i must be finite and nonzero. If π_1 or π_2 or both are semi-definite, and if no non-Ø V has zero norm relative to both π_1 and π_2, then it may be checked that the full rank theory goes through unchanged except that zero and infinite values of the λ_i are allowed. Finally, suppose that $\mathscr{U}_i$ is the subspace with zero norm according to π_i for $i = 1, 2$, and suppose that $\mathscr{U}_1 \cap \mathscr{U}_2$ is larger than Ø. Then it may be checked that the theory goes through in any subspace $\mathscr{V}$ complementary to $\mathscr{U}_1 \cap \mathscr{U}_2$. Moreover, the eigenvalues do not depend on the choice of $\mathscr{V}$ and the eigenvectors for different choices of $\mathscr{V}$ differ by arbitrary zero-norm vectors in $\mathscr{U}_1 \cap \mathscr{U}_2$.

The theory of this section treats π_1 and π_2 asymmetrically only in that the ratios of norms had π_1 in the numerator and π_2 in the denominator. It is clear that reversing the roles of π_1 and π_2 results in reciprocal values for eigenvalues but does not affect the eigenvectors.

5.2 RELATIONS BETWEEN TWO ELLIPSOIDS

Given an affine space $\mathscr{E}$, two inner products π_1 and π_2 on $\mathscr{E}$ may be defined by specifying their respective ellipsoids in $\mathscr{E}$. The symbols π_1 and π_2 may be used to denote the ellipsoids as well as the inner products. In ellipsoid language, Theorem 5.1.1 may be expressed as:

Theorem 5.2.1. *Given any two p-dimensional ellipsoids π_1 and π_2 with a common center, there exists a set of p lines through the center which define sets of conjugate axes for both π_1 and π_2.*

The directions of these p lines are determined by the line segments $ØW_1$, $ØW_2, \ldots, ØW_p$ where $W_1, W_2, \ldots, W_p$ is a basis of eigenvectors of π_1 relative to π_2. The eigenvalue λ_i is the ratio of the squared length of the semi-axis of π_2 in the direction of W_i to the squared length of the semi-axis of π_1 in the same direction for $i = 1, 2, \ldots, p$. This situation is pictured in Fig. 5.2.1 in two dimensions. It is left to the reader to rewrite similarly Theorems 5.1.2 and 5.1.3 in ellipsoid terms.

Consider the geometric picture of the ellipsoids π_1 and π_2 when $\mathscr{E}$ is regarded as a Euclidean space with inner product π_2. In this Euclidean space, π_2 becomes the unit sphere, and the special conjugate axes of π_1 are called *principal axes*. The eigenvalues λ_i are the inverses of the squared lengths of the *principal semi-axes*. These concepts are familiar in two dimensions, as illustrated in Fig. 5.2.2.

Note that, if equality among eigenvalues occurs so that the norms are in a constant ratio over a subspace $\mathscr{W}$ as in Theorem 5.1.2, then the intersections of the ellipsoids π_1 and π_2 with $\mathscr{W}$ are ellipsoids which differ only by a scale factor.

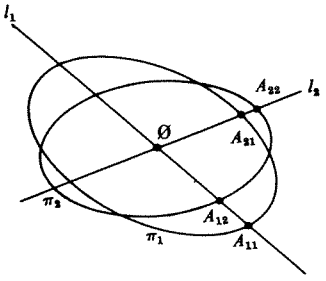

Fig. 5.2.1. Ellipses π_1 and π_2 centered at Ø in the plane. Lines l_1 and l_2 define a pair of common conjugate axes. The eigenvalues of π_1 relative to π_2 are $\lambda_i = [ØA_{i2}/ØA_{i1}]^2$ for $i = 1, 2$.

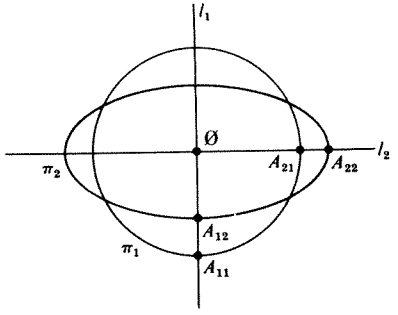

Fig. 5.2.2. The same picture as Fig. 5.2.1 drawn with $ØA_{21}$ and $ØA_{11}$ perpendicular with the same (unit) length, so that π_1 is a (unit) circle.

5.3 RELATED MATRIX THEORY

The basic theory of Section 5.1 can be used to derive corresponding and equivalent theory relating positive definite symmetric matrices. Suppose that $\mathbf{V}$ is any basis of a vector space $\mathscr{E}$, and that inner products π_1 and π_2 are defined in $\mathscr{E}$ using positive definite matrices $\mathbf{Q}_1$ and $\mathbf{Q}_2$ as inner product matrices relative to $\mathbf{V}$. For a basis $\mathbf{W} = \mathbf{AV}$ the corresponding inner product matrices are $\mathbf{AQ}_1\mathbf{A}'$ and $\mathbf{AQ}_2\mathbf{A}'$. Theorem 5.1.1 then yields the purely analytic theorem:

Theorem 5.3.1. *Given any pair $\mathbf{Q}_1$ and $\mathbf{Q}_2$ of positive definite symmetric matrices, there exists a nonsingular matrix $\mathbf{A}$ such that both $\mathbf{AQ}_1\mathbf{A}'$ and $\mathbf{AQ}_2\mathbf{A}'$ are diagonal matrices with positive diagonal elements.*

Alternatively, the two inner products π_1 and π_2 could be defined relative to a basis $\mathbf{U}$ which is orthonormal relative to π_2, so that only the inner product

matrix $\mathbf{Q}$ of π_1 relative to $\mathbf{U}$ is needed to specify both π_1 and π_2. If Theorem 5.1.1 is applied in this case, and if the basis $\mathbf{W}$ is scaled to be orthonormal according to π_2, then $\mathbf{W} = \mathbf{CU}$ for some orthogonal matrix $\mathbf{C}$. Thus:

Theorem 5.3.2. *Given any positive definite symmetric matrix* $\mathbf{Q}$, *there exists an orthogonal matrix* $\mathbf{C}$ *such that* $\mathbf{CQC}'$ *is a diagonal matrix with positive diagonal elements.*

The reader may check that Theorems 5.3.1 and 5.3.2 can be deduced from each other and so are equivalent, and that either Theorem 5.3.1 or 5.3.2 implies Theorem 5.1.1.

Suppose that the diagonal elements of $\mathbf{AQ}_1\mathbf{A}'$ and $\mathbf{AQ}_2\mathbf{A}'$ as found in Theorem 5.3.1 are denoted by $\mu_{11}, \mu_{12}, \ldots, \mu_{1p}$ and $\mu_{21}, \mu_{22}, \ldots, \mu_{2p}$. Then one may define

$$\lambda_i = \frac{\mu_{1i}}{\mu_{2i}} \tag{5.3.1}$$

for $i = 1, 2, \ldots, p$ and, if desired, the rows of $\mathbf{A}$ may be arranged so that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$. Note that the individual rows $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_p$ of $\mathbf{A}$ determine the individual μ_{1i}, μ_{2i}, and λ_i because $\boldsymbol{\alpha}_i\mathbf{Q}_1\boldsymbol{\alpha}_i' = \mu_{1i}$ and $\boldsymbol{\alpha}_i\mathbf{Q}_2\boldsymbol{\alpha}_i' = \mu_{2i}$, so that

$$\lambda_i = \frac{\boldsymbol{\alpha}_i\mathbf{Q}_1\boldsymbol{\alpha}_i'}{\boldsymbol{\alpha}_i\mathbf{Q}_2\boldsymbol{\alpha}_i'}. \tag{5.3.2}$$

If $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are regarded as the inner product matrices of two inner products π_1 and π_2 relative to a basis $\mathbf{V}$ of a vector space $\mathscr{E}$, then it is clear that $\boldsymbol{\alpha}_1\mathbf{V}$, $\boldsymbol{\alpha}_2\mathbf{V}, \ldots, \boldsymbol{\alpha}_p\mathbf{V}$ form an orthogonal basis relative to π_1 and π_2 and

$$\lambda_i = \frac{\boldsymbol{\alpha}_i\mathbf{Q}_1\boldsymbol{\alpha}_i'}{\boldsymbol{\alpha}_i\mathbf{Q}_2\boldsymbol{\alpha}_i'} = \frac{(\boldsymbol{\alpha}_i\mathbf{V}, \boldsymbol{\alpha}_i\mathbf{V})_1}{(\boldsymbol{\alpha}_i\mathbf{V}, \boldsymbol{\alpha}_i\mathbf{V})_2}. \tag{5.3.3}$$

Thus the basis $\mathbf{W} = \mathbf{AV}$ consists of the eigenvectors of π_1 relative to π_2 and $\lambda_1, \lambda_2, \ldots, \lambda_p$ are the corresponding eigenvalues. Since they relate to the corresponding vector space quantities, it is natural to say that $\lambda_1, \lambda_2, \ldots, \lambda_p$ defined by (5.3.1) are the *eigenvalues of* $\mathbf{Q}_1$ *relative to* $\mathbf{Q}_2$ and that the rows $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_p$ of $\mathbf{A}$ are the *associated eigenvectors of* $\mathbf{Q}_1$ *relative to* $\mathbf{Q}_2$.

If $\mathbf{Q}_2 = \mathbf{I}$ the eigenvalues and eigenvectors of $\mathbf{Q}_1$ relative to $\mathbf{I}$ are referred to simply as the *eigenvalues* and *eigenvectors* of $\mathbf{Q}_1$ (with no reference to $\mathbf{Q}_2$). Thus it may be checked that the rows of $\mathbf{C}$ in Theorem 5.3.2 are a set of eigenvectors of $\mathbf{Q}$, and the associated eigenvalues are given by the diagonal elements of $\mathbf{CQC}'$.

If the eigenvalues of a positive definite symmetric matrix $\mathbf{Q}$ are denoted by $\lambda_1, \lambda_2, \ldots, \lambda_p$, and if corresponding orthonormal eigenvectors are denoted by $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_p$, then

$$\mathbf{CQC}' = \mathbf{L} \tag{5.3.4}$$

where $\mathbf{C}$ is an orthogonal matrix with rows $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_p$ and $\mathbf{L}$ is a diagonal matrix with diagonal elements $\lambda_1, \lambda_2, \ldots, \lambda_p$. From (5.3.4),

$$\mathbf{C}'\mathbf{CQC}'\mathbf{C} = \mathbf{C}'\mathbf{LC}, \quad \text{or}$$

$$\mathbf{Q} = \sum_{i=1}^{p} \lambda_i\boldsymbol{\gamma}_i'\boldsymbol{\gamma}_i. \tag{5.3.5}$$

This formula is equivalent to (5.1.11) or (5.1.14).

5.4 COMPUTATIONAL METHODS

5.4.1. A brief description of some basic approaches. The eigenvalues and eigenvectors of a given $\mathbf{Q}$ may *not* be computed in a finite number of steps based on the arithmetic operations of adding, multiplying, dividing, and extracting roots, except when $p \leq 4$ and the approach of Section 5.5 is followed. This is an important difference from the theory of Chapter 4. As a result, the practical determination of eigenvalues and eigenvectors must ultimately involve some iterative procedure of successive approximations, and this is rarely feasible except with an electronic computer. There are many competing approaches which have different advantages in different situations. The subject is highly developed and highly technical; consequently, this section will attempt only to convey some understanding without complete details of two methods, namely the Jacobi and QR methods, which have seen considerable use in recent years. Further theory, methods, and details may be found in Wilkinson (1965).

The method of Jacobi (1846) is easily described from first principles. It applies to any real symmetric matrix $\mathbf{Q}$. The aim is to find an orthogonal matrix $\mathbf{C}$ such that $\mathbf{L} = \mathbf{CQC}'$ is a diagonal matrix. Or in geometric terms, the aim is to pass from an original orthonormal basis $\mathbf{U}$ of a Euclidean space $\mathscr{E}$ to a new orthonormal basis $\mathbf{W} = \mathbf{CU}$ such that the inner product defined by the inner product matrix $\mathbf{Q}$ relative to $\mathbf{U}$ has a diagonal inner product matrix $\mathbf{CQC}'$ relative to $\mathbf{W}$. In these terms, the approach is to find a sequence of orthonormal bases $\mathbf{U} = \mathbf{U}^{(0)} \to \mathbf{U}^{(1)} \to \mathbf{U}^{(2)} \to \cdots$ such that $\lim_{s\to\infty} \mathbf{U}^{(s)} = \mathbf{W}$. This means that the corresponding sequence of orthogonal $\mathbf{C}^{(s)}$ defined by $\mathbf{U}^{(s)} = \mathbf{C}^{(s)}\mathbf{U}$ converges to $\mathbf{C}$ and the sequence $\mathbf{Q}^{(s)} = \mathbf{C}^{(s)}\mathbf{QC}^{(s)\prime}$ converges to $\mathbf{CQC}' = \mathbf{L}$.

This basic idea is to make each transformation $\mathbf{U}^{(s)} \to \mathbf{U}^{(s+1)}$ a plane rotation affecting only a pair of elements of $\mathbf{U}^{(s)}$. Thus at stage s, one chooses indices r and t and defines

$$\begin{aligned} U_r^{(s+1)} &= \cos\theta U_r^{(s)} + \sin\theta U_t^{(s)}, \\ U_t^{(s+1)} &= -\sin\theta U_r^{(s)} + \cos\theta U_t^{(s)}, \quad \text{and} \\ U_j^{(s+1)} &= U_j^{(s)} \quad \text{for} \quad j \neq r, j \neq t. \end{aligned} \tag{5.4.1}$$

The angle θ is chosen to carry out the two-dimensional eigenvalue computation in the space spanned by $U_r^{(s)}$ and $U_t^{(s)}$, i.e., θ is determined by the requirement that $\mathbf{Q}^{(s+1)}$ shall have zeros in the symmetric (r, t) and (t, r) positions. (See Exercise 5.4.1.) A rule must be established for determining r and t at each stage s, but the choice of a rule is not generally critical for convergence and many such rules will be apparent once the proof of convergence is understood.

Passage from $\mathbf{Q}^{(s)}$ to $\mathbf{Q}^{(s+1)}$ annihilates one pair of off-diagonal elements (r, t) and (t, r), but a later stage involving one of r and t will in general bring back a nonzero value in the (r, t) and (t, r) positions. This explains why a finite number of stages cannot be used to annihilate one by one all the symmetric pairs of off-diagonal elements. Nevertheless there is a very simple sense in which each stage brings one measurably closer to the limiting diagonal form. The theory depends on a lemma which asserts:

> ***Lemma 5.4.1.*** *The sum of squares of the p^2 elements of $\mathbf{Q}^{(s+1)}$ is identical to the sum of squares of the p^2 elements of $\mathbf{Q}^{(s)}$, so that this sum of squares is constant through all the stages. Moreover, at stage s, the invariant sum may be broken into three sums each of which is invariant, these parts consisting of the sum of squares of the 4 elements in both rows and columns r or t, the sum of squares of the $(p-2)^2$ elements in neither rows nor columns r or t, and the sum of squares of the remaining $2(p-2)$ elements in rows r or t but not columns r or t or in columns r or t but not rows r or t.*

From the lemma it may be easily seen that one effect of the passage from $\mathbf{Q}^{(s)}$ to $\mathbf{Q}^{(s+1)}$ is to reduce the sum of squares of the off-diagonal elements by $2q_{rt}^{(s)2}$ while increasing the sum of squares of the diagonal elements by the same amount, where $q_{rt}^{(s)}$ denotes the (r, t) element of $\mathbf{Q}^{(s)}$.

It is natural, in view of the above lemma to measure the distance from $\mathbf{U}^{(s)}$ to $\mathbf{W}$ by the sum of squares of the off-diagonal elements of $\mathbf{Q}^{(s)}$, for, if this sum could be reduced to zero, then the required diagonalization would be precisely achieved. Also, it is clear that many rules of choosing r and t at stage s will result in this distance tending to zero as $s \to \infty$. The most obvious rule is to choose r and t $(r \neq t)$ to maximize $|q_{rt}^{(s)}|$. However, for large p, an electronic computer may find it wasteful of time to locate this maximum for each s. Consequently, a rule taking any (r, t) pair such that $|q_{rt}^{(s)}|$ exceeds some assigned threshold is often used. When no values of $|q_{rt}^{(s)}|$ exceed the threshold, then the threshold may be lowered, and so on, until the distance from $\mathbf{U}^{(s)}$ to $\mathbf{W}$ is arbitrarily small.

In the practical application of the method, one starts with $\mathbf{Q}$ and $\mathbf{I}$ in a set of memory locations in a computer. The contents of these registers are then altered by stages so that $[\mathbf{Q}, \mathbf{I}] \to [\mathbf{Q}^{(1)}, \mathbf{C}^{(1)}] \to [\mathbf{Q}^{(2)}, \mathbf{C}^{(2)}] \to \cdots$, finally stopping at a stage for which $\mathbf{Q}^{(s)}$ is acceptably close to diagonal. For further details see Exercise 5.4.1.

Example 5.4. The eigenvalues and eigenvectors of

$$\mathbf{Q} = \begin{bmatrix} 19.1434 & 9.0356 & 9.7634 & 3.2394 \\ 9.0356 & 11.8658 & 4.6232 & 2.4746 \\ 9.7634 & 4.6232 & 12.2978 & 3.8794 \\ 3.2394 & 2.4746 & 3.8794 & 2.4604 \end{bmatrix}$$

were computed as below by the Jacobi method:

$$\mathbf{Q}^{(1)} = \begin{bmatrix} 25.245387 & 0 & 10.678569 & 4.0694975 \\ 0 & 5.7638128 & -1.6328245 & 0.2377977 \\ 10.678569 & -1.6328245 & 12.297800 & 3.879400 \\ 4.0694975 & 0.2377977 & 3.879400 & 2.460400 \end{bmatrix},$$

$$\mathbf{C}^{(1)} = \begin{bmatrix} 0.8287229 & 0.5596592 & 0 & 0 \\ -0.5596592 & 0.8287229 & 0 & 0 \\ 0 & 0 & 1.0 & 0 \\ 0 & 0 & 0 & 1.0 \end{bmatrix},$$

$$\mathbf{Q}^{(2)} = \begin{bmatrix} 31.259260 & -0.8012371 & 0 & 5.4495008 \\ -0.8012371 & 5.7638128 & -1.4227210 & 0.2377977 \\ 0 & -1.4227210 & 6.2839258 & 1.3832912 \\ 5.4495008 & 0.2377977 & 1.3832912 & 2.460400 \end{bmatrix},$$

$$\mathbf{C}^{(2)} = \begin{bmatrix} 0.7220871 & 0.4876451 & 0.4907062 & 0 \\ -0.5596592 & 0.8287229 & 0 & 0 \\ -0.4066594 & -0.2746282 & 0.8713251 & 0 \\ 0 & 0 & 0 & 1.0 \end{bmatrix},$$

$$\mathbf{Q}^{(3)} = \begin{bmatrix} 32.255954 & -0.7453803 & 0.2488708 & 0 \\ -0.7453803 & 5.7638128 & -1.4227210 & 0.3780697 \\ 0.2488708 & -1.4227210 & 6.2839258 & 1.3607196 \\ 0 & 0.3780697 & 1.3607196 & 1.4637057 \end{bmatrix},$$

$$\mathbf{C}^{(3)} = \begin{bmatrix} 0.7103045 & 0.4796880 & 0.4826991 & 0.179912 \\ -0.5596592 & 0.8287229 & 0 & 0 \\ -0.4066594 & -0.2746282 & 0.8713251 & 0 \\ -0.1299122 & -0.08773323 & -0.08828395 & 0.9836827 \end{bmatrix},$$

and so on to

$$\mathbf{Q}^{(16)} = \begin{bmatrix} 32.280122 & -0.0000016 & 0 & 0 \\ -0.0000016 & 4.9317448 & 0 & -0.0000005 \\ 0 & 0 & 7.5673957 & 0 \\ 0 & -0.0000005 & 0 & 0.9881297 \end{bmatrix},$$

$$\mathbf{C}^{(16)} = \begin{bmatrix} 0.7214751 & 0.4526087 & 0.4921702 & 0.1799647 \\ -0.6868510 & 0.4923919 & 0.4466156 & 0.2938031 \\ -0.03142772 & -0.7305346 & 0.6688321 & 0.1341453 \\ 0.08198438 & -0.1378935 & -0.3331163 & 0.9291380 \end{bmatrix},$$

$$\mathbf{Q}^{(17)} = \begin{bmatrix} 32.280121 & 0 & 0 & 0 \\ 0 & 4.9317446 & 0 & -0.0000005 \\ 0 & 0 & 7.5673957 & 0 \\ 0 & -0.0000005 & 0 & 0.9881297 \end{bmatrix},$$

$$\mathbf{C}^{(17)} = \begin{bmatrix} 0.7214751 & 0.4526087 & -0.4921702 & 0.1799647 \\ -0.6868510 & 0.4923919 & 0.4466156 & 0.2938031 \\ -0.03142772 & -0.7305346 & 0.6688321 & 0.1341453 \\ -0.08198438 & -0.1378935 & -0.3331163 & 0.9291380 \end{bmatrix}.$$

At this point the iterations were stopped and $\mathbf{Q}^{(17)}$ declared diagonal to the order of accuracy carried in these calculations. The diagonal elements of $\mathbf{Q}^{(17)}$ are the computed (approximate) eigenvalues, and $\mathbf{C}^{(17)}$ is the computed (approximate) matrix of eigenvectors.

The QR method of Francis (1961, 1962) is a modern refinement of a very old concept called *powering*. The idea is best understood in terms of the linear transformation $\mathbf{U} \to \mathbf{QU}$ (cf. Exercise 3.4.10) which sends the eigenvectors $\mathbf{W} = \mathbf{CU}$ into $\mathbf{CQU} = \mathbf{CQC'W} = \mathbf{LW}$, i.e., this is the transformation which simply stretches the component along each eigenvector by a factor which is the corresponding eigenvalue. In other terms, the transformation carries

$$\boldsymbol{\alpha}\mathbf{U} \to \beta_1\lambda_1 W_1 + \beta_2\lambda_2 W_2 + \cdots + \beta_p\lambda_p W_p \tag{5.4.2}$$

where $\boldsymbol{\beta} = \boldsymbol{\alpha}\mathbf{C}'$ expresses the $\mathbf{W}$ coordinates of $\boldsymbol{\alpha}\mathbf{U}$. Moreover, the result of successively applying this transformation s times, namely $\mathbf{U} \to \mathbf{Q}^s\mathbf{U}$, carries

$$\boldsymbol{\alpha}\mathbf{U} \to \beta_1\lambda_1^s W_1 + \beta_2\lambda_2^s W_2 + \cdots + \beta_p\lambda_p^s W_p. \tag{5.4.3}$$

Assuming for simplicity the strict ordering $\lambda_1 > \lambda_2 > \cdots > \lambda_p$, it is clear that, provided $\boldsymbol{\alpha}$ is chosen so that $\beta_1 \neq 0$, the first coefficient $\beta_1\lambda_1^s$ in (5.4.3) comes to dominate all the others as $s \to \infty$. Thus, the sequence of vectors $\boldsymbol{\alpha}\mathbf{Q}, \boldsymbol{\alpha}\mathbf{Q}^2, \ldots, \boldsymbol{\alpha}\mathbf{Q}^s, \ldots$ must tend, after scaling for unit norm, to the first eigenvector, which is the first row of $\mathbf{C}$. At the same time the ratio of the norms of $\boldsymbol{\alpha}\mathbf{Q}^{s+1}$ and $\boldsymbol{\alpha}\mathbf{Q}^s$ tends to λ_1. This is an old method of computing the largest eigenvalue and its associated eigenvector. It may be applied again to the right side of (5.3.5) with the first term removed to yield the second eigenvalue and eigenvector, and so on.

Instead of starting from a single vector $\boldsymbol{\alpha}\mathbf{U}$, the QR method starts from a basis $\mathbf{AU}$ and determines all the eigenvectors in one sequence of iterations. Formula (5.4.3) generalizes to

$$\mathbf{AU} \to \begin{bmatrix} b_{11}\lambda_1^s W_1 + b_{12}\lambda_2^s W_2 + \cdots + b_{1p}\lambda_p^s W_p \\ b_{21}\lambda_1^s W_1 + b_{22}\lambda_2^s W_2 + \cdots + b_{2p}\lambda_p^s W_p \\ \vdots \\ b_{p1}\lambda_1^s W_1 + b_{p2}\lambda_2^s W_2 + \cdots + b_{pp}\lambda_p^s W_p^s \end{bmatrix}, \tag{5.4.4}$$

where $\mathbf{B} = \mathbf{AC}'$. Assuming for $r = 1, 2, \ldots, p$ that the first r rows and columns of $\mathbf{B}$ define a nonsingular $r \times r$ matrix, because $\lambda_1 > \lambda_2 > \cdots > \lambda_p$ it is clear that for $r = 1, 2, \ldots, p$ the subspace spanned by the first r elements in the right side of (5.4.4) tends as $s \to \infty$ to the subspace spanned by $W_1, W_2, \ldots, W_r$. It follows that, if the basis on the right side of (5.4.4) is orthogonalized to yield an orthonormal basis, the sequence of orthonormal bases thus defined for $s = 0, 1, 2, \ldots$ tends to the basis $\mathbf{W}$ of eigenvectors as $s \to \infty$. This sequence of bases will be denoted by $\mathbf{U}^{(s)} = \mathbf{C}^{(s)}\mathbf{U}$ for $s = 0, 1, 2, \ldots$ where the coordinate matrices $\mathbf{C}^{(s)}$ tend to $\mathbf{C}$ and the inner product matrices $\mathbf{Q}^{(s)} = \mathbf{C}^{(s)}\mathbf{Q}\mathbf{C}^{(s)\prime}$ tend to $\mathbf{L}$ as $s \to \infty$.

The computations are carried out by following the steps $[\mathbf{A}, \mathbf{Q}] \to [\mathbf{C}^{(0)}, \mathbf{Q}^{(0)}] \to [\mathbf{C}^{(1)}, \mathbf{Q}^{(1)}] \to \cdots$, and these computations are conveniently described using the concept of triangularization discussed in Exercise 4.3.3. Thus $\mathbf{C}^{(0)}$ is defined by

$$\mathbf{A} = \mathbf{DC}^{(0)}, \tag{5.4.5}$$

where $\mathbf{D}$ is triangular with zeros above the diagonal, and

$$\mathbf{Q}^{(0)} = \mathbf{C}^{(0)}\mathbf{Q}\mathbf{C}^{(0)\prime}. \tag{5.4.6}$$

For $s = 0, 1, 2, \ldots$, the orthogonal matrix $\mathbf{K}^{(s)}$ such that $\mathbf{U}^{(s+1)} = \mathbf{K}^{(s)}\mathbf{U}^{(s)}$ may be defined by

$$\mathbf{Q}^{(s)} = \mathbf{D}^{(s)}\mathbf{K}^{(s)}, \tag{5.4.7}$$

where again $\mathbf{D}^{(s)}$ is triangular with zeros above the diagonal. Having $\mathbf{K}^{(s)}$, one finds

$$\mathbf{C}^{(s+1)} = \mathbf{K}^{(s)}\mathbf{C}^{(s)} \tag{5.4.8}$$

and

$$\mathbf{Q}^{(s+1)} = \mathbf{K}^{(s)}\mathbf{Q}^{(s)}\mathbf{K}^{(s)\prime} = \mathbf{K}^{(s)}\mathbf{D}^{(s)}, \tag{5.4.9}$$

and is ready for the next stage. The justification of (5.4.7) is provided by showing that it leads to the correct $\mathbf{C}^{(s)}$, i.e., that

$$\mathbf{AQ}^s = \mathbf{FC}^{(s)}, \tag{5.4.10}$$

where $\mathbf{F}$ is a triangular matrix with zeros above the diagonal. To demonstrate (5.4.10), note that

$$\begin{aligned} \mathbf{Q}^{(s)} &= \mathbf{K}^{(s-1)}\mathbf{Q}^{(s-1)}\mathbf{K}^{(s-1)\prime} \\ &= \mathbf{K}^{(s-1)}\mathbf{K}^{(s-2)}\mathbf{Q}^{(s-2)}\mathbf{K}^{(s-2)\prime}\mathbf{K}^{(s-1)\prime} \\ &\;\;\vdots \\ &= \mathbf{K}^{(s-1)} \cdots \mathbf{K}^{(1)}\mathbf{C}^{(0)}\mathbf{Q}\mathbf{C}^{(0)\prime}\mathbf{K}^{(1)\prime} \cdots \mathbf{K}^{(s-1)\prime} \\ &= \mathbf{C}^{(s)}\mathbf{Q}\mathbf{C}^{(s)\prime} \end{aligned} \tag{5.4.11}$$

from which

$$\mathbf{Q}^{(s)}\mathbf{C}^{(s)} = \mathbf{C}^{(s)}\mathbf{Q}. \tag{5.4.12}$$

Applying (5.4.8), (5.4.7), and (5.4.12) in order yields

$$\mathbf{D}\mathbf{D}^{(0)}\mathbf{D}^{(1)} \cdots \mathbf{D}^{(s)}\mathbf{C}^{(s+1)} = [\mathbf{D}\mathbf{D}^{(0)}\mathbf{D}^{(1)} \ldots \mathbf{D}^{(s-1)}\mathbf{C}^{(s+1)}]\mathbf{Q} \tag{5.4.13}$$

and applying (5.4.13) for $s-1, s-2, \ldots, 0$ followed by (5.4.5) yields

$$\mathbf{D}\mathbf{D}^{(0)}\mathbf{D}^{(1)} \ldots \mathbf{D}^{(s-1)}\mathbf{C}^{(s)} = \mathbf{A}\mathbf{Q}^{s}, \tag{5.4.14}$$

which is in the form (5.4.10) with

$$\mathbf{F} = \mathbf{D}\mathbf{D}^{(0)}\mathbf{D}^{(1)} \ldots \mathbf{D}^{(s-1)}.$$

Obviously, the initial choice of $\mathbf{A}$ can greatly affect the rate of convergence of the QR method. For example, if by a lucky guess $\mathbf{A}$ were exactly $\mathbf{C}$ then one iteration would produce the exact correct answer. Likewise, if $\mathbf{A}$ were near $\mathbf{C}$ in some sense then one would expect fewer iterations to be needed. The usefulness of the QR method is greatly enhanced by a trick choice of $\mathbf{A}$ which *tridiagonalizes* $\mathbf{Q}$.

The $p \times p$ symmetric matrix $\mathbf{Q}$ is said to be *tridiagonalized* by the orthogonal matrix $\mathbf{G}$ provided that

$$\mathbf{T} = \mathbf{G}\mathbf{Q}\mathbf{G}' \tag{5.4.15}$$

has (i,j) elements $t_{ij} = 0$ if $|j-i| > 1$.

A happy combination of circumstances makes this idea important for eigenvalue calculations. First, it may be carried out by a succession of $p-1$ elementary orthogonal transformations and is therefore computationally highly tractable. Second, it obviously yields an inner product matrix $\mathbf{T}$ which will in general be much closer to the derived $\mathbf{L}$ than is the original $\mathbf{Q}$, so that it is attractive as a first step in several approaches to eigenvalue computation. Third, the QR method has the property that, if $\mathbf{Q}^{(0)}$ is tridiagonal, then the whole sequence $\mathbf{Q}^{(s)}$ for $s = 1, 2, \ldots$ is tridiagonal, so that the choice of $\mathbf{A}$ to be the tridiagonalizing $\mathbf{G}$ of (5.4.15) means that the whole QR iteration procedure involves only tridiagonal inner product matrices. This may be especially important for large p because it means that $\mathbf{Q}^{(s)}$ may be stored in $2p-1$ memory registers of a computer rather than the $p(p+1)/2$ registers required in general. The procedure for tridiagonalizing $\mathbf{Q}$ by Householder transformations will be described in the following paragraph. For a proof that the QR method preserves the tridiagonal form the reader is referred to Wilkinson (1965).

$\mathbf{T}$ in (5.4.15) is produced by a sequence of steps $\mathbf{T} = \mathbf{T}^{(0)} \to \mathbf{T}^{(1)} \to \cdots \to \mathbf{T}^{(p-1)} = \mathbf{T}$, where $\mathbf{T}^{(s-1)}$ has zeros in positions (i,j) and (j,i) for $j > i+1$ and $i = 1, 2, \ldots, s-1$. The next stage uses an orthogonal matrix $\mathbf{J}^{(s)}$ to create $\mathbf{T}^{(s)} = \mathbf{J}^{(s)}\mathbf{T}^{(s-1)}\mathbf{J}^{(s)\prime}$ which has the required zeros in the sth row and column. The end result after $p-1$ stages is then (5.4.15) with

$$\mathbf{G} = \mathbf{J}^{(p-1)}\mathbf{J}^{(p-2)} \ldots \mathbf{J}^{(1)}.$$

$\mathbf{J}^{(s)}$ has the form

$$\mathbf{J}^{(s)} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{L}^{(s)} \end{bmatrix}, \tag{5.4.16}$$

where $\mathbf{I}$ has dimensions $s \times s$ and $\mathbf{L}^{(s)}$ is a $(p-s) \times (p-s)$ elementary orthogonal matrix. $\mathbf{J}^{(s)}$ carries the basis $\mathbf{V}^{(s-1)}$ with inner product matrix $\mathbf{T}^{(s-1)}$ into the basis $\mathbf{V}^{(s)} = \mathbf{J}^{(s)}\mathbf{V}^{(s-1)}$ with inner product matrix $\mathbf{T}^{(s)}$ but operates only on the last $p-s$ elements of $\mathbf{V}^{(s-1)}$. Consider any orthogonal transformation leaving $V_1^{(s-1)}, \ldots, V_s^{(s-1)}$ fixed and sending $V_{s+1}^{(s-1)}$ into $V = \sum_{s+1}^{p} t_{si}^{(s-1)} V_i^{(s-1)}$, where $t_{ij}^{(s-1)}$ denotes the (i,j) element of $\mathbf{T}^{(s-1)}$. Any vector $W = \sum_{s+1}^{p} c_{s+i} V_{s+i}^{(s-1)}$ is orthogonal to V according to the original inner product if and only if $\sum_{s+1}^{p} c_{s+i} t_{si}^{(s-1)} = 0$ and is orthogonal to $V_s^{(s-1)}$ if and only if $[0, 0, \ldots, 0, c_{s+i}, \ldots, c_p]\mathbf{T}^{(s-1)}[0, 0, \ldots, 1, 0, \ldots, 0]' = 0$. But these two conditions are identical, so that any transformation of this type is a candidate for $\mathbf{J}^{(s)}$. The suggested $\mathbf{J}^{(s)}$ employs for $\mathbf{L}^{(s)}$ in (5.4.16) the $(p-s) \times (p-s)$ elementary orthogonal matrix which reflects in the bisector of the angle between $V_{s+1}^{(s-1)}$ and $\sum_{s+1}^{p} t_{si}^{(s-1)} V_i^{(s-1)}$.

The QR method may be conveniently and elegantly carried out using a long sequence of Householder transformations. The first $p-1$ of these reduce $\mathbf{Q}$ to tridiagonal form. Then each iteration of the QR method proper requires a triangularization which can be carried out by $p-1$ Householder transformations as described in part (*c*) of Exercise 4.3.3. The QR method now appears to be favored by numerical analysts, since it generally performs well, regarding both speed and round-off error. Because of its dependence on powering, however, the rate of convergence of the QR method is sensitive to the ratios among $\lambda_1, \lambda_2, \ldots, \lambda_p$, which are of course initially unknown. For example, if $\lambda_1/\lambda_2 \gg 1$ and $\lambda_2/\lambda_3 \gg 1$ but $\lambda_3, \lambda_4, \ldots, \lambda_p$ have ratios close to unity, then the QR method will quickly find the first two eigenvalues and eigenvectors, but may be slow to converge to the rest.

The Jacobi method was used in the examples of Chapter 9, 10, and 11.

5.4.2. The SDG operator. The computing operation of finding $\mathbf{L}$ and $\mathbf{C}$ from $\mathbf{Q}$ will be written in this book as an operator SDG applied to the pair $[\mathbf{Q}, \mathbf{I}]$ yielding the pair $[\mathbf{L}, \mathbf{C}]$. SDG abbreviates step-diagonalize as in Beaton (1964). More precisely, the notation

$$[\mathbf{L}, \mathbf{C}] = \mathrm{SDG}[1, 2, \ldots, p][\mathbf{Q}, \mathbf{I}] \tag{5.4.17}$$

will be used. The bracket $[1, 2, \ldots, p]$ indicates that all p rows and columns of $\mathbf{Q}$ were used. In many statistical applications, the eigenvalue analysis is applied only to a subset of the rows and columns of $\mathbf{Q}$. For example, selecting rows and columns $1, 2, \ldots, s$, one would write

$$[\dot{\mathbf{L}}, \dot{\mathbf{C}}] = \mathrm{SDG}[1, 2, \ldots, s][\mathbf{Q}, \mathbf{I}], \tag{5.4.18}$$

where

$$\dot{\mathbf{C}} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \tag{5.4.19}$$

with $\mathbf{C}_{11}$ the matrix of eigenvectors of $\mathbf{Q}_{11}$ and

$$\dot{\mathbf{L}} = \dot{\mathbf{C}}\mathbf{Q}\dot{\mathbf{C}}', \tag{5.4.20}$$

the partitions referring in the obvious way to the first s and last $p - s$ rows and columns. The extension of the SDG operator to any subset of rows and columns should now be clear.

One further extension of the operator notion is often used. For any $p \times p$ matrix $\mathbf{K}$, define

$$[\dot{\mathbf{L}}, \dot{\mathbf{K}}] = \text{SDG}[1, 2, \ldots, s][\mathbf{Q}, \mathbf{K}] \tag{5.4.21}$$

to be a generalization of (5.4.16) where $\dot{\mathbf{K}} = \dot{\mathbf{C}}\mathbf{K}$. Thus, if the eigenvalue analysis starts with a basis $\mathbf{U} = \mathbf{KV}$, then the SDG operator (5.4.21) produces the bases $\dot{\mathbf{W}} = \dot{\mathbf{C}}\mathbf{U} = \dot{\mathbf{C}}\mathbf{KV}$ directly in terms of $\mathbf{V}$. For example, if $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are inner product matrices relative to a basis $\mathbf{V}$ and if the eigenvalues and eigenvectors of $\mathbf{Q}_1$ relative to $\mathbf{Q}_2$ are required, then one might first compute $\text{MST}[1, 2, \ldots, p][\mathbf{Q}_2, \mathbf{I}] = [\mathbf{I}, \mathbf{K}]$. In terms of the basis $\mathbf{U} = \mathbf{KV}$, one then needs the eigenvalues of $\mathbf{KQ}_1\mathbf{K}'$ relative to $\mathbf{I}$, so the operator $\text{SDG}[1, 2, \ldots, p]$ $[\mathbf{KQ}_1\mathbf{K}', \mathbf{K}]$ produces $[\mathbf{L}, \mathbf{CK}]$ where $\mathbf{CK}$ expresses the eigenvectors in terms of the original basis $\mathbf{V}$.

5.5 AN ANALYTIC DEFINITION OF EIGENVALUES AND EIGENVECTORS

Eigenvalues are often defined as roots of a certain determinantal equation, and the corresponding eigenvectors are defined by a corresponding set of linear equations. Since different definitions have been given already, these analytic definitions may be derived as a theorem.

Theorem 5.5.1. *Suppose that $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are positive definite symmetric matrices, that $\lambda_1, \lambda_2, \ldots, \lambda_p$ are the eigenvalues of $\mathbf{Q}_1$ relative to $\mathbf{Q}_2$, and that $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_p$ are corresponding eigenvectors. Then $\lambda_1, \lambda_2, \ldots, \lambda_p$ are the roots of the equation*

$$\det(\mathbf{Q}_1 - \lambda\mathbf{Q}_2) = 0, \tag{5.5.1}$$

and $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_p$ satisfy the equations

$$\boldsymbol{\alpha}_i(\mathbf{Q}_1 - \lambda_i\mathbf{Q}_2) = \mathbf{0} \tag{5.5.2}$$

for $i = 1, 2, \ldots, p$.

To prove this theorem, think of $\mathbf{Q}_1$ and $\mathbf{Q}_2$ as inner product matrices of π_1 and π_2 relative to a basis $\mathbf{V}$ of $\mathscr{E}$. Then consider changing to the basis $\mathbf{W} = \mathbf{AV}$

where the rows of $\mathbf{A}$ are $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_p$, i.e., $\mathbf{W}$ is the basis of eigenvectors $\boldsymbol{\alpha}_1\mathbf{V}, \boldsymbol{\alpha}_2\mathbf{V}, \ldots, \boldsymbol{\alpha}_p\mathbf{V}$. The inner product matrices of π_1 and π_2 relative to $\mathbf{W}$ are $\mathbf{M}_1 = \mathbf{AQ}_1\mathbf{A}'$ and $\mathbf{M}_2 = \mathbf{AQ}_2\mathbf{A}'$, where $\mathbf{M}_1$ and $\mathbf{M}_2$ are diagonal matrices with diagonal elements $\mu_{11}, \mu_{12}, \ldots, \mu_{1p}$ and $\mu_{21}, \mu_{22}, \ldots, \mu_{2p}$ such that $\lambda_i = \mu_{1i}/\mu_{2i}$ for $i = 1, 2, \ldots, p$. Thus

$$\begin{aligned} \det(\mathbf{Q}_1 - \lambda\mathbf{Q}_2) &= \det(\mathbf{A}^{-1}\mathbf{A}(\mathbf{Q}_1 - \lambda\mathbf{Q}_2)\mathbf{A}'\mathbf{A}'^{-1}) \\ &= [\det \mathbf{A}]^{-2} \det(\mathbf{AQ}_1\mathbf{A}' - \lambda\mathbf{AQ}_2\mathbf{A}') \\ &= [\det \mathbf{A}]^{-2} \det(\mathbf{M}_1 - \lambda\mathbf{M}_2) \\ &= [\det \mathbf{A}]^{-2} \prod_{i=1}^{p} (\mu_{1i} - \lambda\mu_{2i}) \\ &= (-1)^p [\det \mathbf{A}]^{-2} \prod_{i=1}^{p} \mu_{2i} \prod_{i=1}^{p} (\lambda - \lambda_i), \end{aligned} \tag{5.5.3}$$

so that the roots of (5.5.1) are $\lambda_1, \lambda_2, \ldots, \lambda_p$. Similarly

$$\begin{aligned} \boldsymbol{\alpha}_i(\mathbf{Q}_1 - \lambda_i\mathbf{Q}_2) &= \boldsymbol{\alpha}_i\mathbf{A}^{-1}\mathbf{A}(\mathbf{Q}_1 - \lambda_i\mathbf{Q}_2)\mathbf{A}'\mathbf{A}'^{-1} \\ &= \mu_i\mathbf{1}_i(\mathbf{M}_1 - \lambda_i\mathbf{M}_2)\mathbf{A}'^{-1} \\ &= \mathbf{0}\mathbf{A}'^{-1} = \mathbf{0}, \end{aligned} \tag{5.5.4}$$

where $\mathbf{1}_i$ is the ith row of $\mathbf{I}$. The only unexplained step in (5.5.4) is the step $\boldsymbol{\alpha}_i\mathbf{A}^{-1} = \mu_i\mathbf{1}_i$. If $\boldsymbol{\alpha}$ denotes coordinates relative to $\mathbf{V}$ and $\boldsymbol{\beta}$ denotes coordinates relative to $\mathbf{W} = \mathbf{AV}$, then $\boldsymbol{\beta} = \boldsymbol{\alpha}\mathbf{A}^{-1}$. But relative to $\mathbf{W}$, the ith eigenvector has coordinates $\mu_i\mathbf{1}_i$, and hence $\boldsymbol{\alpha}_i\mathbf{A}^{-1} = \mu_i\mathbf{1}_i$ for some μ_i. This completes the proof of the theorem. Theorem 5.5.1 is often given in the special case where $\mathbf{Q}_2 = \mathbf{I}$.

It is not necessary that π_1 should be of full rank, i.e., $\mathbf{Q}_1$ need only be *positive semi-definite* of rank r. In this case there are r nonzero eigenvalues, i.e., the equation (5.5.1) has a zero root of multiplicity $p - r$.

The equations (5.5.1) provide an alternative to the iterative method of computation. Given $\mathbf{Q}_1$ and $\mathbf{Q}_2$, however, it is not easy to compute the left side of (5.5.1). As with the iterative computational methods, difficulties may be eased by changing to a basis $\mathbf{U} = \mathbf{KV}$ which is orthonormal according to π_2. This requires finding $\mathbf{K}$ such that $\mathbf{KQ}_2\mathbf{K}' = \mathbf{I}$ and then finding $\mathbf{Q} = \mathbf{KQ}_1\mathbf{K}'$. Applying (5.5.1) and (5.5.2) to the simplified situation leads to

$$\det(\mathbf{Q} - \lambda\mathbf{I}) = 0 \tag{5.5.5}$$

and

$$\boldsymbol{\gamma}_i(\mathbf{Q} - \lambda_i\mathbf{I}) = \mathbf{0}, \tag{5.5.6}$$

where the roots of (5.5.5) are the same as those of (5.5.1), and the $\boldsymbol{\gamma}_i$ found from (5.5.6) are related to the $\boldsymbol{\alpha}_i$ in (5.5.1) via $\mathbf{U} = \mathbf{KV}$, i.e.,

$$\boldsymbol{\alpha}_i = \boldsymbol{\gamma}_i\mathbf{K}. \tag{5.5.7}$$

Even in this form it is not easy to compute the polynomial $\det(\mathbf{Q} - \lambda\mathbf{I})$. A further simplification is produced by the tridiagonalization device of (5.4.14).

Thus (5.5.5) may be transformed further into

$$\det(\mathbf{T} - \lambda\mathbf{I}) = 0, \tag{5.5.8}$$

which, being the determinant of a tridiagonal matrix, is more easily computed. The eigenvectors are then determined by

$$\boldsymbol{\delta}_i(\mathbf{T} - \lambda_i\mathbf{I}) = \mathbf{0}, \tag{5.5.9}$$

where

$$\boldsymbol{\alpha}_i = \boldsymbol{\delta}_i\mathbf{GK}. \tag{5.5.10}$$

According to (5.5.5) or (5.5.8) the eigenvalues may be found as the roots of a polynomial of degree p in λ. Having these roots, each corresponding eigenvector may be deduced from a set of linear equations as in (5.5.6) or (5.5.9). Note that only $p - 1$ of the equations (5.5.6) or (5.5.9) for a given i are linearly independent and that, because of the arbitrary scale factor, any element of $\boldsymbol{\gamma}_i$ or $\boldsymbol{\delta}_i$ may be arbitrarily fixed, yielding $p - 1$ equations in $p - 1$ unknowns.

This approach to computing eigenvalues and eigenvectors is not pursued in detail because for most statistical purposes one requires both eigenvalues and eigenvectors, and the methods of Section 5.4.1 are generally faster.

5.6 AN APPLICATION TO ANGLES BETWEEN SUBSPACES OF A EUCLIDEAN SPACE

Consider a p-dimensional Euclidean space $\mathscr{E}$ with two specified subspaces $\mathscr{U}$ and $\mathscr{V}$ of dimensions m and n, respectively. The theory of eigenvectors and eigenvalues provides an easy means to an understanding of the system of angles between the hyperplanes $\mathscr{U}$ and $\mathscr{V}$. One may ask, for example, what is the smallest angle between $\mathscr{U}$ and $\mathscr{V}$? In other words, how does one find a vector U in $\mathscr{U}$ and a vector V in $\mathscr{V}$ such that the angle between U and V is minimum, and what is the angle?

Given any U in $\mathscr{U}$, a vector in $\mathscr{V}$ making the smallest angle with U is given by the orthogonal projection of U into $\mathscr{V}$, which may be denoted by $\dot{U}$. Moreover, if θ is the angle from U to $\dot{U}$

$$\cos^2\theta = \frac{(\dot{U}, \dot{U})}{(U, U)}, \tag{5.6.1}$$

and the problem is to choose U to maximize $\cos^2\theta$. To place the problem in the same form as that considered at the start of this chapter, one need only define a pair of inner products π_1 and π_2 over $\mathscr{U}$ where, for any U and W in $\mathscr{U}$,

$$(U, W)_1 = (\dot{U}, \dot{W}) \tag{5.6.2}$$

and

$$(U, W)_2 = (U, W). \tag{5.6.3}$$

Here, the brackets subscripted by i refer to π_i for $i = 1, 2$ and the brackets without subscripts refer to the inner product implied by the assertion that $\mathscr{E}$ is

Euclidean. As before, $\dot{U}$ and $\dot{W}$ refer to the orthogonal projection of U and W into $\mathscr{V}$. The reader should check that π_2 as defined in (5.6.3) is an inner product, albeit possibly semi-definite. It follows that the largest eigenvalue of π_1 relative to π_2 maximizes $\cos^2\theta$ in (5.6.1). The general picture is provided by the following theorem.

Theorem 5.6.1. *Suppose that $\mathscr{U}$ and $\mathscr{V}$ are subspaces of dimensions m and n of a p-dimensional Euclidean space $\mathscr{E}$ where $m \leq n$. Then there exist an orthonormal basis $U_1, U_2, \ldots, U_m$ of $\mathscr{U}$ and an orthonormal basis $V_1, V_2, \ldots, V_n$ of $\mathscr{V}$ with the properties that*

$$(U_i, V_j) = 0 \quad \text{for} \quad i \neq j. \tag{5.6.4}$$

If the angles θ_i between U_i and V_i for $i = 1, 2, \ldots, m$ are ordered so that

$$\theta_1 \leq \theta_2 \leq \cdots \leq \theta_m, \tag{5.6.5}$$

then θ_1 is the smallest angle between $\mathscr{U}$ and $\mathscr{V}$, θ_2 is the smallest angle between the subspace of $\mathscr{U}$ orthogonal to U_1 and the subspace of $\mathscr{V}$ orthogonal to V_1, and so on. If the inequalities in (5.6.5) *are strict inequalities, then the orthonormal sets $U_1, U_2, \ldots, U_m$ and $V_1, V_2, \ldots, V_m$ satisfying* (5.6.4) *are unique. On the other hand, if for some s and t,*

$$\theta_{s-1} < \theta_s = \theta_{s+1} = \cdots = \theta_{s+t-1} < \theta_{s+t}, \tag{5.6.6}$$

then only the subspaces spanned by $U_s, U_{s+1}, \ldots, U_{s+t-1}$ and $V_s, V_{s+1}, \ldots, V_{s+t-1}$ are uniquely determined, and any orthonormal basis of one such t-dimensional subspace corresponds to an orthonormal basis of the other such t-dimensional subspace satisfying (5.6.4) *and* (5.6.5). *In either case, any orthogonal transformation of the set $V_{m+1}, V_{m+2}, \ldots, V_n$ leads to another possible choice of $V_{m+1}, V_{m+2}, \ldots, V_n$, whose only requirement from* (5.6.4) *is that they shall be orthogonal to $\mathscr{U}$.*

As already indicated, the basic idea leading to this theorem is that $U_1, U_2, \ldots, U_m$ shall be taken as a basis of eigenvectors of π_1 relative to π_2. For some s on $0 \leq s \leq m$, the orthogonal projections $\dot{U}_1, \dot{U}_2, \ldots, \dot{U}_s$ will have positive length, and $V_1, V_2, \ldots, V_s$ may be taken to be that orthonormal set found by rescaling $\dot{U}_1, \dot{U}_2, \ldots, \dot{U}_s$ to unit length. It follows that the subspace of $\mathscr{V}$ orthogonal to $V_1, V_2, \ldots, V_s$ is orthogonal to $\mathscr{U}$, and any orthonormal sets $U_{s+1}, U_{s+2}, \ldots, U_m$ and $V_{s+1}, V_{s+2}, \ldots, V_n$ may be chosen to complete the required bases of $\mathscr{U}$ and $\mathscr{V}$. The uniqueness properties claimed for the bases $U_1, U_2, \ldots, U_m$ and $V_1, V_2, \ldots, V_n$ in the theorem follow directly from Theorems 5.1.1, 5.1.2, and 5.1.3. The extremal properties of $\theta_1, \theta_2, \ldots$ follow from the constructive proof of Theorem 5.1.1 (cf. Exercise 5.1.3).

In this geometric framework, the powering procedures of Section 5.4.1 have a simple, elegant interpretation. Because the bases $U_1, U_2, \ldots, U_m$ and

$V_1, V_2, \ldots, V_n$ are orthonormal, the orthogonality relations imply that $\dot{U}_i = \cos\theta_i V_i$ for $i = 1, 2, \ldots, m$. Thus, if A is any vector in $\mathscr{U}$ and $A = \sum_1^m a_i U_i$, then the orthogonal projection $\dot{A}$ of A into $\mathscr{V}$ is given by

$$\dot{A} = \sum_{i=1}^{m} a_i \dot{U}_i = \sum_{i=1}^{m} [a_i \cos\theta_i] V_i. \tag{5.6.7}$$

Similarly, if $B = \sum_1^n b_i V_i$ is any vector in $\mathscr{V}$ and the orthogonal projection of B into $\mathscr{U}$ is denoted by $\ddot{B}$, then one may deduce that $\ddot{V}_i = \cos\theta_i U_i$ for $i = 1, 2, \ldots, m$ and $\ddot{V}_i = \emptyset$ for $i = m + 1, \ldots, n$. Thence

$$\ddot{B} = \sum_{i=1}^{n} b_i \ddot{V}_i = \sum_{i=1}^{m} [b_i \cos\theta_i] U_i. \tag{5.6.8}$$

If one performs the transformations $A \to \dot{A}$ and $B \to \ddot{B}$ in succession one has

$$\sum_{i=1}^{m} a_i U_i \to \sum_{i=1}^{m} [a_i \cos\theta_i] V_i \to \sum_{i=1}^{m} [a_i \cos^2\theta_i] U_i, \tag{5.6.9}$$

which has the form of the transformation $\mathbf{U} \to \mathbf{QU}$ considered in connection with the powering methods. It follows in particular that if a vector A_0 is projected orthogonally into $\mathscr{V}$, then back into $\mathscr{U}$ and so on back and forth, the pair of vectors in $\mathscr{U}$ and $\mathscr{V}$ converges to a pair of vectors which make the smallest angle between $\mathscr{V}$ and $\mathscr{U}$.

In the sequel it will be convenient to call the angles $\theta_1, \theta_2, \ldots, \theta_m$ in (5.6.5) the *canonical angles* between $\mathscr{U}$ and $\mathscr{V}$. Corresponding bases $\mathbf{U}$ and $\mathbf{V}$ as in (5.6.4) will be called *canonical bases*. The angles $\theta_{m+1}, \theta_{m+2}, \ldots, \theta_n$ between $V_{m+1}, V_{m+2}, \ldots, V_n$ and $\mathscr{U}$ are all $\pi/2$, but may also be called canonical angles.

5.7 EXERCISES

5.1.1 Suppose that λ_1 and λ_p are the largest and smallest eigenvalues, respectively, of π_1 relative to π_2. Show that any value of $\lambda = (V, V)_1/(V, V)_2$ between λ_1 and λ_p occurs for some V in $\mathscr{E}$.

5.1.2 Suppose that $\lambda_1, \lambda_2, \ldots, \lambda_p$ denote the eigenvalues of π_1 relative to π_2. For any U and V in $\mathscr{E}$, define π_3 from

$$(U, V)_3 = c_1(U, V)_1 + c_2(U, V)_2.$$

What is the condition on $c_1, c_2, \lambda_1, \lambda_2, \ldots, \lambda_p$ such that $(V, V)_3 > 0$ for all V in $\mathscr{E}$ and $(U, V)_3$ defines an inner product? Show that any basis orthogonal relative to both π_1 and π_2 is orthogonal relative to π_3, whether or not π_3 defines a proper inner product. What are the eigenvalues of π_3 relative to π_2?

5.1.3 Suppose that $W_{(1)}, W_{(2)}, \ldots, W_{(r)}$ denote a subset of a set of eigenvectors of π_1 relative to π_2, where the corresponding eigenvalues $\lambda_{(1)}, \lambda_{(2)}, \ldots, \lambda_{(p)}$ are taken in non-increasing order. Suppose that $\mathscr{U}$ denotes the subspace of $\mathscr{E}$ spanned by $W_{(1)}, W_{(2)}, \ldots,$

$W_{(r)}$. Show that $(V, V)_1/(V, V)_2$ has maximum value $\lambda_{(1)}$ and minimum value $\lambda_{(r)}$ for V in $\mathscr{U}$.

5.1.4 Prove Theorem 5.1.2.

5.1.5 Characterize the set of all inner products π_2 which have a given set of eigenvalues relative to a given inner product π_1.

5.2.1 Suppose that the ellipsoid π_1 has eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_p$ relative to the ellipsoid π_2 in affine space $\mathscr{E}$. Show that the affine ratio of the volume of π_1 to the volume of π_2 is given by

$$\left[\prod_{i=1}^{p} \lambda_i\right]^{-1/2}.$$

Consequently, deduce that, if $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are inner product matrices of π_1 and π_2 relative to a basis $\mathbf{U}$, then the product $\prod_{i=1}^{p} \lambda_i$ may be expressed as $\det \mathbf{Q}_1/\det \mathbf{Q}_2$.

5.2.2 Consider a general ellipsoid π in a Euclidean space $\mathscr{E}$. Show that any axis of the ellipsoid which is perpendicular to the tangent plane at its intersection with the ellipsoid is a principal axis of the ellipsoid.

5.3.1 Suppose that $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are positive definite symmetric matrices. Given the theory proving the existence of an orthogonal basis of a Euclidean vector space, show that the problem of finding a $1 \times p$ matrix $\boldsymbol{\alpha}$ to maximize $\boldsymbol{\alpha}\mathbf{Q}_1\boldsymbol{\alpha}'/\boldsymbol{\alpha}\mathbf{Q}_2\boldsymbol{\alpha}'$ can be reduced to the problem of finding a $1 \times p$ matrix $\boldsymbol{\beta}$ to maximize $\boldsymbol{\beta}\mathbf{Q}\boldsymbol{\beta}'$ subject to the condition $\boldsymbol{\beta}\boldsymbol{\beta}' = 1$, where $\mathbf{Q}$ is a certain positive definite symmetric matrix.

5.3.2 Show that a positive definite symmetric matrix $\mathbf{Q}$ has the same eigenvalues as $\mathbf{DQD}'$ if and only if $\mathbf{D}$ is orthogonal. Show also that if the eigenvectors of $\mathbf{Q}$ are denoted by the rows of a $p \times p$ matrix $\mathbf{C}$, then the eigenvectors of $\mathbf{DQD}'$ for orthogonal $\mathbf{C}$ are given by the rows of $\mathbf{CD}'$.

5.3.3 State and prove the extension of Theorem 5.3.2 to cover positive semi-definite symmetric matrices.

5.3.4 Suppose that $\mathbf{P}$ is any symmetric $p \times p$ matrix, and that

$$\mu = \inf [\boldsymbol{\beta}\mathbf{P}\boldsymbol{\beta}'],$$

where $\boldsymbol{\beta}$ ranges over all row vectors such that $\boldsymbol{\beta}\boldsymbol{\beta}' = 1$. Set

$$\mathbf{Q} = \mathbf{P} + \nu\mathbf{I},$$

where ν is any real number greater than μ. Show that $\mathbf{Q}$ is positive definite symmetric. Deduce also that there exists an orthogonal matrix $\mathbf{C}$ such that

$$\mathbf{CPC}' = \mathbf{L},$$

where $\mathbf{L}$ is a diagonal matrix whose diagonal elements $\nu_1, \nu_2, \ldots, \nu_p$ are not necessarily positive.

5.3.5 Suppose that the positive definite symmetric matrix $\mathbf{Q}$ is operated on by the orthogonal matrix $\mathbf{C}$ to produce $\mathbf{CQC}'$. Show that $\mathbf{Q}$ and $\mathbf{CQC}'$ have the same trace, i.e., the same sum of diagonal elements. Deduce that the sum of the eigenvalues of $\mathbf{Q}$ is given by the trace of $\mathbf{Q}$. Can this result be extended to any symmetric matrix $\mathbf{P}$ (cf. Exercise 5.3.4)?

5.3.6 Suppose that a positive definite symmetric $p \times p$ matrix $\mathbf{Q}$ is the sum $\mathbf{Q}_1 + \mathbf{Q}_2 + \cdots + \mathbf{Q}_p$ of p positive semi-definite matrices each of rank 1. Show that the $\mathbf{Q}_i$ are necessarily the terms in the sum (5.3.5) corresponding to an eigenvalue analysis of $\mathbf{Q}$ (Cochran's theorem).

5.4.1 Show that the angle θ in (5.4.1) is determined by the equation

$$q_{rt}^{(s)} \cos^2 \theta + [q_{tt}^{(s)} - q_{rr}^{(s)}] \cos \theta \sin \theta - q_{rt}^{(s)} \sin^2 \theta = 0,$$

and that the following relations describe how to compute $\mathbf{Q}^{(s+1)}$ from $\mathbf{Q}^{(s)}$ and θ:

$$\begin{aligned}
q_{rr}^{(s+1)} &= q_{rr}^{(s)} \cos^2 \theta + 2q_{rt}^{(s)} \cos \theta \sin \theta + q_{tt}^{(s)} \sin^2 \theta, \\
q_{tt}^{(s+1)} &= q_{rt}^{(s)} \sin^2 \theta - 2q_{rt}^{(s+1)} \cos \theta \sin \theta + q_{tt}^{(s)} \cos^2 \theta, \\
q_{rt}^{(s+1)} &= q_{tr}^{(s+1)} = 0, \\
q_{ir}^{(s+1)} &= q_{ri}^{(s+1)} = q_{ir}^{(s)} \cos \theta + q_{it}^{(s)} \sin \theta \qquad \text{for} \qquad i \neq r, t, \\
q_{it}^{(s+1)} &= q_{ti}^{(s+1)} = -q_{ir}^{(s)} \sin \theta + q_{it}^{(s)} \cos \theta \qquad \text{for} \qquad i \neq r, t, \\
q_{ij}^{(s+1)} &= q_{ij}^{(s)} \qquad \text{for} \qquad i \neq r, t, \qquad \text{and} \qquad j \neq r, t.
\end{aligned}$$

Check also that the rows of $\mathbf{C}^{(s+1)}$ are computed from the rows of $\mathbf{C}^{(s)}$ in the same way that the rows of $\mathbf{U}^{(s+1)}$ are computed from the rows of $\mathbf{U}^{(s)}$, as in (5.4.1).

5.4.2 Suppose that $\mathbf{D}$ is any $q \times q$ orthogonal matrix and $\mathbf{P}$ is any $q \times q$ symmetric matrix. Show that the sum of squares of the elements of $\mathbf{P}$ and of $\mathbf{DPD}'$ are the same. Show how this result can be applied to prove Lemma 5.4.1.

5.4.3 Check the first stage of computation in Example 5.4.

5.4.4 Carry out the tridiagonalization of $\mathbf{Q}$ in Example 5.4 using a pair of elementary orthogonal transformations.

5.4.5 Starting from the tridiagonalized version of $\mathbf{Q}$ computed in Exercise 5.4.4, carry out one stage of the QR method.

5.5.1 Show how to compute directly from determinants of submatrices of $\mathbf{Q}$ the elementary symmetric functions of the eigenvalues of $\mathbf{Q}$, namely,

$$\begin{gathered}
\sum_{i=1}^{p} \lambda_i \\
\sum_{i=1}^{p} \sum_{j=i}^{p} \lambda_i \lambda_j \\
\cdot \\
\cdot \\
\cdot \\
\lambda_1 \lambda_2 \cdots \lambda_p.
\end{gathered}$$

5.5.2 Use the calculus method of Lagrange multipliers to show that the maximum of the quadratic form $\boldsymbol{\alpha}\mathbf{Q}\boldsymbol{\alpha}'$, subject to the condition $\boldsymbol{\alpha}\boldsymbol{\alpha}' = 1$, must satisfy

$$\boldsymbol{\alpha}\mathbf{Q} = \lambda\boldsymbol{\alpha}, \tag{5.7.1}$$

where λ satisfies

$$\det (\mathbf{Q} - \lambda\mathbf{I}) = 0. \tag{5.7.2}$$

Show also that the resulting maximum is λ.

5.5.3 Show that eigenvectors $\boldsymbol{\gamma}_i$ and $\boldsymbol{\gamma}_j$ satisfying (5.5.6) for eigenvalues λ_i and λ_j satisfying $\lambda_i \neq \lambda_j$ must also satisfy

$$\boldsymbol{\gamma}_i \mathbf{Q} \boldsymbol{\gamma}_j = 0. \tag{5.7.3}$$

5.5.4 Suppose (5.7.1) and (5.7.2) are used to define eigenvectors and eigenvalues for an arbitrary matrix $\mathbf{Q}$. Show that, in general, the eigenvectors $\boldsymbol{\gamma}_1^*, \boldsymbol{\gamma}_2^*, \ldots, \boldsymbol{\gamma}_p^*$ of $\mathbf{Q}'$ are different from the eigenvectors $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_p$ of $\mathbf{Q}$, but that

$$\boldsymbol{\gamma}_i \mathbf{Q} \boldsymbol{\gamma}_j^{*\prime} = 0 \qquad \text{if} \qquad \lambda_i \neq \lambda_j \tag{5.7.4}$$

is the appropriate generalization of (5.7.3).

5.5.5 Show, as an alternative to (5.5.5) and (5.5.6), that formulas (5.5.1) and (5.5.2) may be written

$$\det (\mathbf{Q}_1 \mathbf{Q}_2^{-1} - \lambda\mathbf{I}) = 0 \tag{5.7.5}$$

and

$$\boldsymbol{\gamma}_i (\mathbf{Q}_1 \mathbf{Q}_2^{-1} - \lambda_i \mathbf{I}) = \mathbf{0}. \tag{5.7.6}$$

Show that $\mathbf{Q}_1\mathbf{Q}_2^{-1}$ is not in general a symmetric matrix, so that its eigenvalues as defined by $\mathbf{Q}_1\mathbf{Q}_2^{-1}$ may not be found by iterative methods designed for symmetric matrices.

5.6.1 What becomes of the theorem of Section 5.6 when $m = 1$?

5.6.2 Show that (5.6.4) defines a definite or semi-definite inner product over $\mathscr{U}$.

5.6.3 Show that the inner products defined in (5.6.4) and (5.6.5) satisfy $(W, W)_1 / (W, W)_2 \leq 1$ and that such ratios may be interpreted as $\cos^2 \theta$ for some θ.

5.6.4 Show that the angles $\theta_1, \theta_2, \ldots, \theta_m$ defined in Section 5.6 may all be zero, may all be $\pi/2$, or more generally may all be equal to any θ on $0 \leq \theta \leq \pi/2$. Relate the number of θ_i which are zero to the dimension of the intersection of $\mathscr{U}$ and $\mathscr{V}$.

5.6.5 Suppose that $\mathbf{U}$ denotes a basis of $\mathscr{U}$ and $\mathbf{V}$ a basis of $\mathscr{V}$. Suppose that the orthogonal projections of $\mathbf{U}$ into $\mathscr{V}$ into $\mathscr{U}$ are denoted by $\mathbf{SV}$ and $\mathbf{RU}$, respectively, where $\mathbf{U}$, $\mathbf{V}$, $\mathbf{R}$, and $\mathbf{S}$ have dimensions $m \times 1$, $n \times 1$, $n \times m$, and $m \times n$. Show that the iteration procedure is equivalent to computing a product of the form $\mathbf{SRSR} \ldots \mathbf{SR}$. What would be an efficient way to compute such a product?