Preface

The purpose of this book is to present a version of multivariate statistical theory in which vector space and invariance methods replace, to a large extent, more traditional multivariate methods. The book is a text. Over the past ten years, various versions have been used for graduate multivariate courses at the University of Chicago, the University of Copenhagen, and the University of Minnesota. Designed for a one year lecture course or for independent study, the book contains a full complement of problems and problem solutions.

My interest in using vector space methods in multivariate analysis was aroused by William Kruskal's success with such methods in univariate linear model theory. In the late 1960s, I had the privilege of teaching from Kruskal's lecture notes where a coordinate free (vector space) approach to univariate analysis of variance was developed. (Unfortunately, Kruskal's notes have not been published.) This approach provided an elegant unification of linear model theory together with many useful geometric insights. In addition, I found the pedagogical advantages of the approach far outweighed the extra effort needed to develop the vector space machinery. Extending the vector space approach to multivariate situations became a goal, which is realized here. Basic material on vector spaces, random vectors, the normal distribution, and linear models take up most of the first half of this book.

Invariance (group theoretic) arguments have long been an important research tool in multivariate analysis as well as in other areas of statistics. In fact, invariance considerations shed light on most multivariate hypothesis testing, estimation, and distribution theory problems. When coupled with vector space methods, invariance provides an important complement to the traditional distribution theory-likelihood approach to multivariate analysis. Applications of invariance to multivariate problems occur throughout the second half of this book.

A brief summary of the contents and flavor of the ten chapters herein follows. In Chapter 1, the elements of vector space theory are presented. Since my approach to the subject is geometric rather than algebraic, there is an emphasis on inner product spaces where the notions of length, angle, and orthogonal projection make sense. Geometric topics of particular importance in multivariate analysis include singular value decompositions and angles between subspaces. Random vectors taking values in inner product spaces is the general topic of Chapter 2. Here, induced distributions, means, covariances, and independence are introduced in the inner product space setting. These results are then used to establish many traditional properties of the multivariate normal distribution in Chapter 3. In Chapter 4, a theory of linear models is given that applies directly to multivariate problems. This development, suggested by Kruskal's treatment of univariate linear models, contains results that identify all the linear models to which the Gauss–Markov Theorem applies.

Chapter 5 contains some standard matrix factorizations and some elementary Jacobians that are used in later chapters. In Chapter 6, the theory of invariant integrals (measures) is outlined. The many examples here were chosen to illustrate the theory and prepare the reader for the statistical applications to follow. A host of statistical applications of invariance, ranging from the invariance of likelihood methods to the use of invariance in deriving distributions and establishing inde-

pendence, are given in Chapter 7. Invariance arguments are used throughout the remainder of the book.

The last three chapters are devoted to a discussion of some traditional and not so traditional problems in multivariate analysis. Here, I have stressed the connections between classical likelihood methods, linear model considerations, and invariance arguments. In Chapter 8, the Wishart distribution is defined via its representation in terms of normal random vectors. This representation, rather than the form of the Wishart density, is used to derive properties of the Wishart distribution. Chapter 9 begins with a thorough discussion of the multivariate analysis of variance (MANOVA) model. Variations on the MANOVA model including multivariate linear models with structured covariances are the main topic of the rest of Chapter 9. An invariance argument that leads to the relationship between canonical correlations and angles between subspaces is the lead topic in Chapter 10. After a discussion of some distribution theory, the chapter closes with the connection between testing for independence and testing in multivariate regression models.

Throughout the book, I have assumed that the reader is familiar with the basic ideas of matrix and vector algebra in coordinate spaces and has some knowledge of measure and integration theory. As for statistical prerequisites, a solid first year graduate course in mathematical statistics should suffice. The book is probably best read and used as it was written—from front to back. However, I have taught short (one quarter) courses on topics in MANOVA using the material in Chapters 1, 2, 3, 4, 8, and 9 as a basis.

It is very difficult to compare this text with others on multivariate analysis. Although there may be a moderate amount of overlap with other texts, the approach here is sufficiently different to make a direct comparison inappropriate. Upon reflection, my attraction to vector space and invariance methods was, I think, motivated by a desire for a more complete understanding of multivariate statistical models and techniques. Over the years, I have found vector space ideas and invariance arguments have served me well in this regard. There are many multivariate topics not even mentioned here. These include discrimination and classification, factor analysis, Bayesian multivariate analysis, asymptotic results and decision theory results. Discussions of these topics can be found in one or more of the books listed in the Bibliography.

As multivariate analysis is a relatively old subject within statistics, a bibliography of the subject is very large. For example, the entries in *A Bibliography of Multivariate Analysis* by T. W. Anderson, S. Das Gupta, and G. H. P. Styan, published in 1972, number over 6000. The condensed bibliography here contains a few of the important early papers plus a sample of some recent work that reflects my bias. A more balanced view of the subject as a whole can be obtained by perusing the bibliographies of the multivariate texts listed in the Bibliography.

help I would like to acknowledge are Murray Clayton, Siu Chuen Ho, and Takeaki Kariya.

Most of the typing of the manuscript was done by Hanne Hansen. Her efforts are very much appreciated. For their typing of various corrections, addenda, changes, and so on, I would like to thank Melinda Hutson, Catherine Stepnes, and Victoria Wagner.

Morris L. Eaton
Minneapolis, Minnesota
May 1983