Advanced Multivariate Statistics with Matrices

Mathematics and Its Applications

Managing Editor:

M. HAZEWINKEL
Centre for Mathematics and Computer Science, Amsterdam, The Netherlands

Volume 579

Advanced Multivariate Statistics with Matrices

by

Tõnu Kollo
University of Tartu,
Tartu, Estonia

and

Dietrich von Rosen
Swedish University of Agricultural Sciences,
Uppsala, Sweden

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-3418-0 (HB) Springer Dordrecht, Berlin, Heidelberg, New York
ISBN-10 1-4020-3419-9 (e-book) Springer Dordrecht, Berlin, Heidelberg, New York
ISBN-13 978-1-4020-3418-3 (HB) Springer Dordrecht, Berlin, Heidelberg, New York
ISBN-13 978-1-4020-3419-0 (e-book) Springer Dordrecht, Berlin, Heidelberg, New York

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

Printed on acid-free paper

Printed in the Netherlands.

To

Imbi, Kaarin, Ülle, Ardo
Tõnu

Tatjana, Philip, Sophie, Michael
Dietrich

TABLE OF CONTENTS

PREFACE

The book presents important tools and techniques for treating problems in modern multivariate statistics in a systematic way. The ambition is to indicate new directions as well as to present the classical part of multivariate statistical analysis in this framework. The book has been written for graduate students and statisticians who are not afraid of matrix formalism. The goal is to provide them with a powerful toolkit for their research and to give necessary background and deeper knowledge for further studies in different areas of multivariate statistics. It can also be useful for researchers in applied mathematics and for people working on data analysis and data mining who can find useful methods and ideas for solving their problems.

It has been designed as a textbook for a two semester graduate course on multivariate statistics. Such a course has been held at the Swedish Agricultural University in 2001/02. On the other hand, it can be used as material for series of shorter courses. In fact, Chapters 1 and 2 have been used for a graduate course "Matrices in Statistics" at University of Tartu for the last few years, and Chapters 2 and 3 formed the material for the graduate course "Multivariate Asymptotic Statistics" in spring 2002. An advanced course "Multivariate Linear Models" may be based on Chapter 4.

A lot of literature is available on multivariate statistical analysis written for different purposes and for people with different interests, background and knowledge. However, the authors feel that there is still space for a treatment like the one presented in this volume. Matrix algebra and theory of linear spaces are continuously developing fields, and it is interesting to observe how statistical applications benefit from algebraic achievements. Our main aim is to present tools and techniques whereas development of specific multivariate methods has been somewhat less important. Often alternative approaches are presented and we do not avoid complicated derivations.

Besides a systematic presentation of basic notions, throughout the book there are several topics which have not been touched or have only been briefly considered in other books on multivariate analysis. The internal logic and development of the material in this book is the following. In Chapter 1 necessary results on matrix algebra and linear spaces are presented. In particular, lattice theory is used. There are three closely related notions of matrix algebra which play a key role in the presentation of multivariate statistics: Kronecker product, vec-operator and the concept of matrix derivative. In Chapter 2 the presentation of distributions is heavily based on matrix algebra, what makes it possible to present complicated expressions of multivariate moments and cumulants in an elegant and compact way. The very basic classes of multivariate and matrix distributions, such as normal, elliptical and Wishart distributions, are studied and several relations and characteristics are presented of which some are new. The choice of the material in Chapter 2 has been made having in mind multivariate asymptotic distributions and multivariate expansions in Chapter 3. This Chapter presents general formal density expansions which are applied in normal and Wishart approximations. Finally, in Chapter 4 the results from multivariate distribution theory and approximations are used in presentation of general linear models with a special emphasis on the Growth Curve model.

The authors are thankful to the Royal Swedish Academy of Sciences and to the Swedish Institute for their financial support. Our sincere gratitude belongs also to the University of Tartu, Uppsala University and the Swedish Agricultural University for their support. Dietrich von Rosen gratefully acknowledges the support from the Swedish Natural Sciences Research Council, while Tõnu Kollo is indebted to the Estonian Science Foundation. Grateful thanks to Professors Heinz Neudecker, Kenneth Nordström and Muni Srivastava. Some results in the book stem from our earlier cooperation. Also discussions with Professors Kai-Tai Fang and Björn Holmquist have been useful for presentation of certain topics. Many thanks to our colleagues for support and stimulating atmosphere. Last but not least we are grateful to all students who helped improve the presentation of the material during the courses held on the material.

Uppsala
November 2004
Tõnu Kollo
Dietrich von Rosen

INTRODUCTION

In 1958 the first edition of *An Introduction to Multivariate Statistical Analysis* by T. W. Anderson appeared and a year before S. N. Roy had published *Some Aspects of Multivariate Analysis*. Some years later, in 1965, *Linear Statistical Inference and Its Applications* by C. R. Rao came out. During the following years several books on multivariate analysis appeared: Dempster (1969), Morrison (1967), Press (1972), Kshirsagar (1972). The topic became very popular in the end of 1970s and the beginning of 1980s. During a short time several monographs were published: Giri (1977), Srivastava & Khatri (1979), Mardia, Kent & Bibby (1979), Muirhead (1982), Takeuchi, Yanai & Mukherjee (1982), Eaton (1983), Farrell (1985) and Siotani, Hayakawa, & Fujikoshi (1985). All these books made considerable contributions to the area though many of them focused on certain topics. In the last 20 years new results in multivariate analysis have been so numerous that it seems impossible to cover all the existing material in one book. One has to make a choice and different authors have made it in different directions. The first class of books presents introductory texts of first courses on undergraduate level (Srivastava & Carter, 1983; Flury, 1997; Srivastava, 2002) or are written for non-statisticians who have some data they want to analyze (Krzanowski, 1990, for example). In some books the presentation is computer oriented (Johnson, 1998; Rencher, 2002), for example). There are many books which present a thorough treatment on specific multivariate methods (Greenacre, 1984; Jolliffe, 1986; McLachlan, 1992; Lauritzen, 1996; Kshirsagar & Smith, 1995, for example) but very few are presenting foundations of the topic in the light of newer matrix algebra. We can refer to Fang & Zhang (1990) and Bilodeau & Brenner (1999), but there still seems to be space for a development. There exists a rapidly growing area of linear algebra related to mathematical statistics which has not been used in full range for a systematic presentation of multivariate statistics. The present book tries to fill this gap to some extent. Matrix theory, which is a cornerstone of multivariate analysis starting from T. W. Anderson, has been enriched during the last few years by several new volumes: Bhatia (1997), Harville (1997), Schott (1997b), Rao & Rao (1998), Zhang (1999). These books form a new basis for presentation of multivariate analysis.

The Kronecker product and vec-operator have been used systematically in Fang & Zhang (1990), but these authors do not tie the presentation to the concept of the matrix derivative which has become a powerful tool in multivariate analysis. Magnus & Neudecker (1999) has become a common reference book on matrix differentiation. In Chapter 2, as well as Chapter 3, we derive most of the results using matrix derivatives. When writing the book, our main aim was to answer the questions "Why?" and "In which way?", so typically the results are presented with proofs or sketches of proofs. However, in many situations the reader can also find an answer to the question "How?".

Before starting with the main text we shall give some general comments and

remarks about the notation and abbreviations.
Throughout the book we use boldface transcription for matrices and vectors. Matrices will be denoted by capital letters and vectors by ordinary small letters of Latin or Greek alphabets. Random variables will be denoted by capital letters from the end of the Latin alphabet. Notion appearing in the text for the first time is printed in *italics*. The end of proofs, definitions and examples is marked by ∎
To shorten proofs we use the following abbreviation

$$\underset{(1.3.2)}{=}$$

which should be red as "the equality is obtained by applying formula (1.3.2)". We have found it both easily understandable and space preserving. In numeration of Definitions, Theorems, Propositions and Lemmas we use a three position system. Theorem 1.2.10 is the tenth theorem of Chapter 1, Section 2. For Corollaries four integers are used: Corollary 1.2.3.1 is the first Corollary of Theorem 1.2.3. In a few cases when we have Corollaries of Lemmas, the capital L has been added to the last number, so Corollary 1.2.3.1L is the first corollary of Lemma 1.2.3. We end the Introduction with the List of Notation, where the page number indicates the first appearance or definition.

LIST OF NOTATION

$\circ$ – elementwise or Hadamard product, p. 3
$\otimes$ – Kronecker or direct product, tensor product, p. 81, 41
$\oplus$ – direct sum, p. 27
$\boxplus$ – orthogonal sum, p. 27
$\mathbf{A}$ – matrix, p. 2
$\mathbf{a}$ – vector, p. 2
c – scalar, p. 3
$\mathbf{A}'$ – transposed matrix, p. 4
$\mathbf{I}_p$ – identity matrix, p. 4
$\mathbf{A}_d$ – diagonalized matrix $\mathbf{A}$, p. 6
$\mathbf{a}_d$ – diagonal matrix, $\mathbf{a}$ as diagonal, p. 6
diag$\mathbf{A}$ – vector of diagonal elements of $\mathbf{A}$, p. 6
$|\mathbf{A}|$ – determinant of $\mathbf{A}$, p. 7
$r(\mathbf{A})$ – rank of $\mathbf{A}$, p. 9
p.d. – positive definite, p. 12
$\mathbf{A}^-$ – generalized inverse, g-inverse, p. 15
$\mathbf{A}^+$ – Moore-Penrose inverse, p. 17
$\mathbf{A}(K)$ – patterned matrix (pattern K), p. 97
$\mathbf{K}_{p,q}$ – commutation matrix, p. 79
vec – vec-operator, p. 89
$\mathbf{A}^{\otimes k}$ – k-th Kroneckerian power, p. 84
$V^k(\mathbf{A})$ – vectorization operator, p. 115
$R^k(\mathbf{A})$ – product vectorization operator, p. 115

$\frac{d\mathbf{Y}}{d\mathbf{X}}$ – matrix derivative, p. 127
m.i.v. – mathematically independent and variable, p. 126
$\mathbf{J}(\mathbf{Y} \to \mathbf{X})$ – Jacobian matrix, p. 156
$|\mathbf{J}(\mathbf{Y} \to \mathbf{X})|_+$ – Jacobian, p. 156
$\mathbb{A}^\perp$ – orthocomplement, p. 27
$\mathbb{B} \perp \mathbb{A}$ – perpendicular subspace, p. 27
$\mathbb{B}|\mathbb{A}$ – commutative subspace, p. 31
$\mathcal{R}(\mathbf{A})$ – range space, p. 34
$\mathcal{N}(\mathbf{A})$ – null space, p. 35
$\mathcal{C}(\mathbf{C})$ – column space, p. 48
$\mathbf{X}$ – random matrix, p. 171
$\mathbf{x}$ – random vector, p. 171
X – random variable, p. 171
$f_\mathbf{x}(\mathbf{x})$ – density function, p. 174
$F_\mathbf{x}(\mathbf{x})$ – distribution function, p. 174
$\varphi_\mathbf{x}(\mathbf{t})$ – characteristic function, p. 174
$E[\mathbf{x}]$ – expectation, p. 172
$D[\mathbf{x}]$ – dispersion matrix, p. 173
$c_k[\mathbf{x}]$ – $k-$th cumulant, p. 181
$m_k[\mathbf{x}]$ – $k-$th moment, p. 175
$\overline{m}_k[\mathbf{x}]$ – $k-$th central moment, p. 175
$mc_k[\mathbf{x}]$ – $k-$th minimal cumulant, p. 185
$mm_k[\mathbf{x}]$ – $k-$th minimal moment, p. 185
$m\overline{m}_k[\mathbf{x}]$ – $k-$th minimal central moment, p. 185
$\mathbf{S}$ – sample dispersion matrix, p. 284
$\mathbf{R}$ – sample correlation matrix, p. 289
Ω – theoretical correlation matrix, p. 289
$N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ – multivariate normal distribution, p. 192
$N_{p,n}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$ – matrix normal distribution, p. 192
$E_p(\boldsymbol{\mu}, \mathbf{V})$ – elliptical distribution, p. 224
$W_p(\boldsymbol{\Sigma}, n)$ – central Wishart distribution, p. 237
$W_p(\boldsymbol{\Sigma}, n, \boldsymbol{\Delta})$ – noncentral Wishart distribution, p. 237
$M\beta_I(p, m, n)$ – multivariate beta distribution, type I, p. 249
$M\beta_{II}(p, m, n)$ – multivariate beta distribution, type II, p. 250
$\xrightarrow{\mathcal{D}}$ – convergence in distribution, week convergence, p. 277
$\xrightarrow{\mathcal{P}}$ – convergence in probability, p. 278
$O_P(\cdot)$ – p. 278
$o_P(\cdot)$ – p. 278
$(\mathbf{X})()'$ – $(\mathbf{X})(\mathbf{X})'$, p. 355