

Cambridge University Press

978-0-521-86842-6 - The Coordinate-Free Approach to Linear Models

Michael J. Wichura

Frontmatter

[More information](#)

## *The Coordinate-Free Approach to Linear Models*

This book is about the coordinate-free, or geometric, approach to the theory of linear models, more precisely, Model I ANOVA and linear regression models with nonrandom predictors in a finite-dimensional setting. This approach is more insightful, more elegant, more direct, and simpler than the more common matrix approach to linear regression, analysis of variance, and analysis of covariance models in statistics. The book discusses the intuition behind and optimal properties of various methods of estimating and testing hypotheses about unknown parameters in the models.

Topics covered include inner product spaces, orthogonal projections, book orthogonal spaces, Tjur experimental designs, basic distribution theory, the geometric version of the Gauss-Markov theorem, optimal and nonoptimal properties of Gauss-Markov, Bayes and shrinkage estimators under the assumption of normality, the optimal properties of  $F$ -tests, and the analysis of covariance and missing observations.

Michael J. Wichura has 37 years of teaching experience in the Department of Statistics at the University of Chicago. He has served as an associate editor for the *Annals of Probability* and was the database editor for the *Current Index to Statistics* from 1995 to 2000. He is the author of the PiCTeX macros (for drawing pictures in TeX) and the PiCTeX manual and also of the TABLE macros and the TABLE manual.

Cambridge University Press

978-0-521-86842-6 - The Coordinate-Free Approach to Linear Models

Michael J. Wichura

Frontmatter

[More information](#)

## CAMBRIDGE SERIES IN STATISTICAL AND PROBABILISTIC MATHEMATICS

*Editorial Board:*

- R. Gill, *Department of Mathematics, Utrecht University*  
 B. D. Ripley, *Department of Statistics, University of Oxford*  
 S. Ross, *Epstein Department of Industrial & Systems Engineering, University of Southern California*  
 B. W. Silverman, *St. Peter's College, Oxford*  
 M. Stein, *Department of Statistics, University of Chicago*

This series of high-quality upper-division textbooks and expository monographs covers all aspects of stochastic applicable mathematics. The topics range from pure and applied statistics to probability theory, operations research, optimization, and mathematical programming. The books contain clear presentations of new developments in the field and also of the state of the art in classical methods. While emphasizing rigorous treatment of theoretical methods, the books also contain applications and discussions of new techniques made possible by advances in computational practice.

*Already Published*

1. *Bootstrap Methods and Their Application*, by A. C. Davison and D. V. Hinkley
2. *Markov Chains*, by J. Norris
3. *Asymptotic Statistics*, by A. W. van der Vaart
4. *Wavelet Methods for Time Series Analysis*, by Donald B. Percival and Andrew T. Walden
5. *Bayesian Methods*, by Thomas Leonard and John S. J. Hsu
6. *Empirical Processes in M-Estimation*, by Sara van de Geer
7. *Numerical Methods of Statistics*, by John F. Monahan
8. *A User's Guide to Measure Theoretic Probability*, by David Pollard
9. *The Estimation and Tracking of Frequency*, by B. G. Quinn and E. J. Hannan
10. *Data Analysis and Graphics Using R*, by John Maindonald and John Braun
11. *Statistical Models*, by A. C. Davison
12. *Semiparametric Regression*, by D. Ruppert, M. P. Wand, and R. J. Carroll
13. *Exercises in Probability*, by Loic Chaumont and Marc Yor
14. *Statistical Analysis of Stochastic Processes in Time*, by J. K. Lindsey
15. *Measure Theory and Filtering*, by Lakhdar Aggoun and Robert Elliott
16. *Essentials of Statistical Inference*, by G. A. Young and R. L. Smith
17. *Elements of Distribution Theory*, by Thomas A. Severini
18. *Statistical Mechanics of Disordered Systems*, by Anton Bovier

Cambridge University Press

978-0-521-86842-6 - The Coordinate-Free Approach to Linear Models

Michael J. Wichura

Frontmatter

[More information](#)

# *The Coordinate-Free Approach to Linear Models*

---

MICHAEL J. WICHURA



**CAMBRIDGE**  
UNIVERSITY PRESS

Cambridge University Press

978-0-521-86842-6 - The Coordinate-Free Approach to Linear Models

Michael J. Wichura

Frontmatter

[More information](#)

CAMBRIDGE UNIVERSITY PRESS

Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press

32 Avenue of the Americas, New York, NY 10013-2473, USA

[www.cambridge.org](http://www.cambridge.org)

Information on this title: [www.cambridge.org/9780521868426](http://www.cambridge.org/9780521868426)

© Michael J. Wichura 2006

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2006

Printed in the United States of America

*A catalog record for this publication is available from the British Library.*

*Library of Congress Cataloging in Publication Data*

Wichura, Michael J. (Michael John)

The coordinate-free approach to linear models / Michael J. Wichura.

p. cm. – (Cambridge series on statistical and probabilistic mathematics)

Includes bibliographical references and index.

ISBN-13: 978-0-521-86842-6 (hardback)

ISBN-10: 0-521-86842-4 (hardback)

1. Linear models (Statistics) 2. Analysis of variance. 3. Regression analysis.

4. Analysis of covariance. I. Title. II. Series.

QA279.W53 2006

519.5–dc22 2006004133

ISBN-13 978-0-521-86842-6 hardback

ISBN-10 0-521-86842-4 hardback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party Internet Web sites referred to in this publication and does not guarantee that any content on such Web sites is, or will remain, accurate or appropriate.

Cambridge University Press

978-0-521-86842-6 - The Coordinate-Free Approach to Linear Models

Michael J. Wichura

Frontmatter

[More information](#)

---

In memoriam  
William H. Kruskal  
1919–2005

## CONTENTS

<b>Preface</b> . . . . .	xi
<b>1. Introduction</b> . . . . .	1
1. Orientation . . . . .	1
2. An illustrative example . . . . .	2
3. Notational conventions . . . . .	4
<b>2. Topics in Linear Algebra</b> . . . . .	6
1. Orthogonal projections . . . . .	6
2. Properties of orthogonal projections . . . . .	12
2A. Characterization of orthogonal projections . . . . .	12
2B. Differences of orthogonal projections . . . . .	13
2C. Sums of orthogonal projections . . . . .	16
2D. Products of orthogonal projections . . . . .	17
2E. An algebraic form of Cochran's theorem . . . . .	19
3. Tjur's theorem . . . . .	21
4. Self-adjoint transformations and the spectral theorem . . . . .	32
5. Representation of linear and bilinear functionals . . . . .	36
6. Problem set: Cleveland's identity . . . . .	40
7. Appendix: Rudiments . . . . .	41
7A. Vector spaces . . . . .	42
7B. Subspaces . . . . .	43
7C. Linear functionals . . . . .	43
7D. Linear transformations . . . . .	43
<b>3. Random Vectors</b> . . . . .	45
1. Random vectors taking values in an inner product space . . . . .	45
2. Expected values . . . . .	46
3. Covariance operators . . . . .	47
4. Dispersion operators . . . . .	49
5. Weak sphericity . . . . .	51
6. Getting to weak sphericity . . . . .	52
7. Normality . . . . .	52
8. The main result . . . . .	54
9. Problem set: Distribution of quadratic forms . . . . .	57

<b>4. Gauss-Markov Estimation</b>	60
1. Linear functionals of $\mu$	60
2. Estimation of linear functionals of $\mu$	62
3. Estimation of $\mu$ itself	67
4. Estimation of $\sigma^2$	70
5. Using the wrong inner product	72
6. Invariance of GMEs under linear transformations	74
7. Some additional optimality properties of GMEs	75
8. Estimable parametric functionals	78
9. Problem set: Quantifying the Gauss-Markov theorem	85
<b>5. Normal Theory: Estimation</b>	89
1. Maximum likelihood estimation	89
2. Minimum variance unbiased estimation	90
3. Minimality of $P_M Y$	92
4. James-Stein estimation	97
5. Problem set: Admissible minimax estimation of $\mu$	104
<b>6. Normal Theory: Testing</b>	110
1. The likelihood ratio test	111
2. The $F$ -test	112
3. Monotonicity of the power of the $F$ -test	117
4. An optimal property of the $F$ -test	121
5. Confidence intervals for linear functionals of $\mu$	127
6. Problem set: Wald's theorem	136
<b>7. Analysis of Covariance</b>	141
1. Preliminaries on nonorthogonal projections	141
1A. Characterization of projections	142
1B. The adjoint of a projection	142
1C. An isomorphism between $J$ and $I^\perp$	143
1D. A formula for $P_{J;I}$ when $J$ is given by a basis	143
1E. A formula for $P'_{J;I}$ when $J$ is given by a basis	145
2. The analysis of covariance framework	146
3. Gauss-Markov estimation	147
4. Variances and covariances of GMEs	150
5. Estimation of $\sigma^2$	152
6. Scheffé intervals for functionals of $\mu_M$	153
7. $F$ -testing	155
8. Problem set: The Latin square design	159
<b>8. Missing Observations</b>	164
1. Framework and Gauss-Markov estimation	164
2. Obtaining $\hat{\mu}$	169
2A. The consistency equation method	169
2B. The quadratic function method	172
2C. The analysis of covariance method	173

Cambridge University Press  
978-0-521-86842-6 - The Coordinate-Free Approach to Linear Models  
Michael J. Wichura  
Frontmatter  
[More information](#)

---

## CONTENTS

ix

3. Estimation of $\sigma^2$ . . . . .	176
4. $F$ -testing . . . . .	177
5. Estimation of linear functionals . . . . .	181
6. Problem set: Extra observations . . . . .	184
<b>References</b> . . . . .	188
<b>Index</b> . . . . .	190



## PREFACE

When I was a graduate student in the mid-1960s, the mathematical theory underlying analysis of variance and regression became clear to me after I read a draft of William Kruskal's monograph on the so-called coordinate-free, or geometric, approach to these subjects. Alas, with Kruskal's demise, this excellent treatise will never be published.

From time to time during the 1970s, 80s, and early 90s, I had the good fortune to teach the coordinate-free approach to linear models, more precisely, to Model I analysis of variance and linear regression with nonrandom predictors. While doing so, I evolved my own set of lecture notes, presented here. With regard to inspiration and content, my debt to Kruskal is clear. However, my notes differ from Kruskal's in many ways. To mention just a few, my notes are intended for a one- rather than three-quarter course. The notes are aimed at statistics graduate students who are already familiar with the basic concepts of linear algebra, such as linear subspaces and linear transformations, and who have already had some exposure to the matrix formulation of the GLM, perhaps through a methodology course, and who are interested in the underlying theory. I have also included Tjur experimental designs and some of the highlights of the optimality theory for estimation and testing in linear models under the assumption of normality, feeling that the elegant setting provided by the coordinate-free approach is a natural one in which to place these jewels of mathematical statistics. As he alluded to in his conversation with Zabell (1994), Kruskal always wished that he could have brought "his book, his potential book, his unborn book" to life. Out of deference to Kruskal, who was my colleague here at the University of Chicago, I have not until now made my notes public.

For motivation, Chapter 1 presents an example illustrating Kruskal's claim in his 1961 Berkeley Symposium paper that "the coordinate-free approach . . . permits a simpler, more general, more elegant, and more direct treatment of the general theory . . . than do its notational counterparts, the matrix and scalar approaches." I hope that as s/he works through the book, the reader will be more and more convinced that this is indeed the case.

The last section of Chapter 2 reviews the "elementary" concepts from linear algebra which the reader is assumed to know already. The first five sections of that chapter develop the "nonelementary" tools we need, such as (finite-dimensional, real) inner product spaces, orthogonal projections, the spectral theorem for self-adjoint linear transformations, and the representation of linear and bilinear functionals. Sec-

tion 2.3 uses the notions of book orthogonal subspaces and orthogonal projections, along with the inclusion and intersection of subspaces, to discuss Tjur experimental designs, thereby giving a unified treatment of the algebra underlying the models one usually encounters in a first course in analysis of variance, and much more.

Chapter 3 develops basic distribution theory for random vectors taking values in inner product spaces — the first- and second-moment structures of such vectors, and the key fact that if one splits a spherical normal random vector up into its components in mutually orthogonal subspaces, then those components are independent and have themselves spherical normal distributions within their respective subspaces.

The geometric version of the Gauss-Markov theorem is discussed in Chapter 4, from the point of view of both estimating linear functionals of the unknown mean vector and estimating the mean vector itself. These results are based just on assumptions about the first and second moment structures of the data vector. For an especially nice example of how the geometric viewpoint is more insightful than the matricial one, be sure to work the “four-penny problem” in Exercise 4.2.17.

Estimation under the assumption of normality is taken up in Chapter 5. In some respects, Gauss-Markov estimators are optimal; for example, they are minimum variance unbiased. However, in other respects they are not optimal. Indeed, Bayesian considerations lead naturally to the James-Stein shrinkage type estimators, which can significantly outperform GMEs in terms of mean square error.

Again under the assumption of normality,  $F$ -testing of null hypotheses and the related issue of interval estimation are taken up in Chapter 6. Chapters 7 and 8 deal with the analysis of covariance and missing observations, respectively.

The book is written at the level of Halmos’s *Finite-Dimensional Vector Spaces* (but Halmos is not a prerequisite). Thus the reader will on the one hand need to be comfortable with the yoga of definitions, theorems, and proofs, but on the other hand be comforted by knowing that the abstract ideas will be illustrated by concrete examples and presented with (what I hope are) some insightful comments. To get a feeling for the coordinate-free approach before embarking on a serious study of this book, you might find it helpful to first read one or more of the brief elementary nontechnical expositions of the subject that have appeared in *The American Statistician*, for example, Herr (1980), Bryant (1984), or Saville and Wood (1986). From the perspective of mathematical statistics, there are some very elegant results, and some notable surprises, connected with the optimality theory for Gauss-Markov estimation and  $F$ -testing under the assumption of normality. In the sections that deal with these matters — in particular Sections 5.4, 5.5, 6.3, 6.4, and 6.6 — the mathematics is somewhat harder than elsewhere, corresponding to the greater depth of the theory.

Each of the chapters following the Introduction contains numerous exercises, along with a problem set that develops some topic complementing the material in that chapter. Altogether there are about 200 exercises. Most of them are easy but, I hope, instructive. I typically devote most of the class time to having students present solutions to the exercises. Some exercises foreshadow what is to come, by

Cambridge University Press

978-0-521-86842-6 - The Coordinate-Free Approach to Linear Models

Michael J. Wichura

Frontmatter

[More information](#)

## PREFACE

xiii

covering a special case of material that will be presented in full generality later. Moreover, the assertions of some exercises are appealed to later in the text. If you are working through the book on your own, you should at least read over each exercise, even if you do not work things out. The problem sets, several of which are based on journal articles, are harder than the exercises and require a sustained effort for their completion.

The students in my courses have typically worked through the whole book in one quarter, but that is admittedly a brisk pace. One semester would be less demanding. For a short course, you could concentrate on the parts of the book that flesh out the outline of the coordinate-free viewpoint that Kruskal set out in Section 2 of his aforementioned Berkeley Symposium paper. That would involve: Chapter 1, for motivation; Sections 2.1, 2.2, 2.5, and 2.7 for notation and basic results from linear algebra; Sections 3.1–3.8 for distribution theory; Sections 4.1–4.7 for the properties of Gauss-Markov estimators; Sections 5.1 and 5.2 for estimation under the assumption of normality; and Sections 6.1, 6.2, and 6.5 for hypothesis testing and interval estimation under normality.

Various graduate students, in particular Neal Thomas and Nathaniel Schenker, have made many comments that have greatly improved this book. My thanks go to all of them, and also to David Van Dyke and Peter Meyer for suggesting how easy/hard each of the exercises is. Thanks are also due to Mitzi Nakatsuka for her help in converting the notes to  $\text{\TeX}$ , and to Persi Diaconis for his advice and encouragement.

*Michael J. Wichura*  
University of Chicago