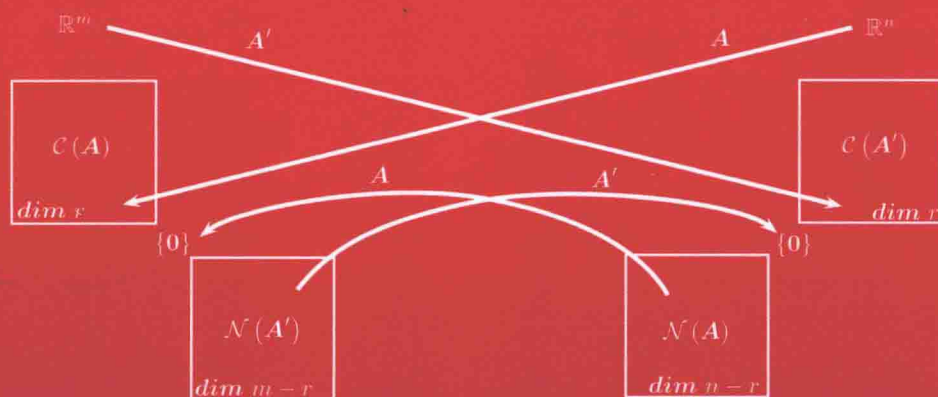


Texts in Statistical Science

Linear Algebra and Matrix Analysis for Statistics



Sudipto Banerjee
Anindya Roy



CRC Press
Taylor & Francis Group

A CHAPMAN & HALL BOOK

Texts in Statistical Science

Linear Algebra and Matrix Analysis for Statistics

Sudipto Banerjee

Professor of Biostatistics
School of Public Health
University of Minnesota, U.S.A.

Anindya Roy

Professor of Statistics

Department of Mathematics and Statistics
University of Maryland, Baltimore County, U.S.A.



CRC Press

Taylor & Francis Group

Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group an **informa** business

A CHAPMAN & HALL BOOK

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2014 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works

Printed on acid-free paper
Version Date: 20140407

International Standard Book Number-13: 978-1-4200-9538-8 (Hardback)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, micro-filming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging-in-Publication Data

Banerjee, Sudipto, author.

Linear algebra and matrix analysis for statistics / Sudipto Banerjee, professor of biostatistics, School of Public Health, University of Minnesota, U.S.A., Anindya Roy, professor of statistics, Department of Mathematics and Statistics, University of Maryland, Baltimore County, U.S.A.
pages cm. -- (Chapman & Hall/CRC texts in statistical science)

Summary: "Linear algebra and the study of matrix algorithms have become fundamental to the development of statistical models. Using a vector-space approach, this book provides an understanding of the major concepts that underlie linear algebra and matrix analysis. Each chapter introduces a key topic, such as infinite-dimensional spaces, and provides illustrative examples. The authors examine recent developments in diverse fields such as spatial statistics, machine learning, data mining, and social network analysis. Complete in its coverage and accessible to students without prior knowledge of linear algebra, the text also includes results that are useful for traditional statistical applications."-- Provided by publisher.

Includes bibliographical references and index.

ISBN 978-1-4200-9538-8 (hardback)

1. Algebras, Linear. 2. Matrices. 3. Mathematical statistics. I. Roy, Anindya, 1970- author. II. Title.

QA184.2.B36 2014
512'.5--dc23

2014004884

Visit the Taylor & Francis Web site at
<http://www.taylorandfrancis.com>

and the CRC Press Web site at
<http://www.crcpress.com>

Linear Algebra and Matrix Analysis for Statistics

CHAPMAN & HALL/CRC

Texts in Statistical Science Series

Series Editors

Francesca Dominici, *Harvard School of Public Health, USA*

Julian J. Faraway, *University of Bath, UK*

Martin Tanner, *Northwestern University, USA*

Jim Zidek, *University of British Columbia, Canada*

Analysis of Failure and Survival Data

P.J. Smith

The Analysis of Time Series: An Introduction, Sixth Edition

C. Chatfield

Applied Bayesian Forecasting and Time Series Analysis

A. Pole, M. West, and J. Harrison

Applied Categorical and Count Data Analysis

W. Tang, H. He, and X.M. Tu

Applied Nonparametric Statistical Methods, Fourth Edition

P. Sprent and N.C. Smeeton

Applied Statistics: Handbook of GENSTAT Analyses

E.J. Snell and H. Simpson

Applied Statistics: Principles and Examples

D.R. Cox and E.J. Snell

Applied Stochastic Modelling, Second Edition

B.J.T. Morgan

Bayesian Data Analysis, Third Edition

A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, and D.B. Rubin

Bayesian Ideas and Data Analysis: An Introduction for Scientists and Statisticians

R. Christensen, W. Johnson, A. Branscum, and T.E. Hanson

Bayesian Methods for Data Analysis, Third Edition

B.P. Carlin and T.A. Louis

Beyond ANOVA: Basics of Applied Statistics

R.G. Miller, Jr.

The BUGS Book: A Practical Introduction to Bayesian Analysis

D. Lunn, C. Jackson, N. Best, A. Thomas, and D. Spiegelhalter

A Course in Categorical Data Analysis

T. Leonard

A Course in Large Sample Theory

T.S. Ferguson

Data Driven Statistical Methods

P. Sprent

Decision Analysis: A Bayesian Approach

J.Q. Smith

Design and Analysis of Experiments with SAS

J. Lawson

Elementary Applications of Probability Theory, Second Edition

H.C. Tuckwell

Elements of Simulation

B.J.T. Morgan

Epidemiology: Study Design and Data Analysis, Third Edition

M. Woodward

Essential Statistics, Fourth Edition

D.A.G. Rees

Exercises and Solutions in Statistical Theory

L.L. Kupper, B.H. Neelon, and S.M. O'Brien

Exercises and Solutions in Biostatistical Theory

L.L. Kupper, B.H. Neelon, and S.M. O'Brien

Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models

J.J. Faraway

A First Course in Linear Model Theory

N. Ravishanker and D.K. Dey

Generalized Additive Models: An Introduction with R

S. Wood

Generalized Linear Mixed Models: Modern Concepts, Methods and Applications

W. W. Stroup

Graphics for Statistics and Data Analysis with R

K.J. Keen

Interpreting Data: A First Course in Statistics

A.J.B. Anderson

Introduction to General and Generalized Linear Models

H. Madsen and P. Thyregod

An Introduction to Generalized Linear Models, Third Edition

A.J. Dobson and A.G. Barnett

Introduction to Multivariate Analysis

C. Chatfield and A.J. Collins

Introduction to Optimization Methods and Their Applications in Statistics

B.S. Everitt

Introduction to Probability with R

K. Baclawski

Introduction to Randomized Controlled Clinical Trials, Second Edition

J.N.S. Matthews

Introduction to Statistical Inference and Its Applications with R

M.W. Trosset

Introduction to Statistical Limit Theory

A.M. Polansky

Introduction to Statistical Methods for Clinical Trials

T.D. Cook and D.L. DeMets

Introduction to Statistical Process Control

P. Qiu

Introduction to the Theory of Statistical Inference

H. Liero and S. Zwanzig

Large Sample Methods in Statistics

P.K. Sen and J. da Motta Singer

Linear Algebra and Matrix Analysis for Statistics

S. Banerjee and A. Roy

Logistic Regression Models

J.M. Hilbe

Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference, Second Edition

D. Gamerman and H.F. Lopes

Mathematical Statistics

K. Knight

Modeling and Analysis of Stochastic Systems, Second Edition

V.G. Kulkarni

Modelling Binary Data, Second Edition

D. Collett

Modelling Survival Data in Medical Research, Second Edition

D. Collett

Multivariate Analysis of Variance and Repeated Measures: A Practical Approach for Behavioural Scientists

D.J. Hand and C.C. Taylor

Multivariate Statistics: A Practical Approach

B. Flury and H. Riedwyl

Multivariate Survival Analysis and Competing Risks

M. Crowder

Nonparametric Methods in Statistics with SAS Applications

O. Korosteleva

Pólya Urn Models

H. Mahmoud

Practical Data Analysis for Designed Experiments

B.S. Yandell

Practical Longitudinal Data Analysis

D.J. Hand and M. Crowder

Practical Multivariate Analysis, Fifth Edition

A. Afifi, S. May, and V.A. Clark

Practical Statistics for Medical Research

D.G. Altman

A Primer on Linear Models

J.F. Monahan

Principles of Uncertainty

J.B. Kadane

Probability: Methods and Measurement

A. O'Hagan

Problem Solving: A Statistician's Guide, Second Edition

C. Chatfield

Randomization, Bootstrap and Monte Carlo Methods in Biology, Third Edition

B.F.J. Manly

Readings in Decision Analysis

S. French

Sampling Methodologies with Applications

P.S.R.S. Rao

Stationary Stochastic Processes: Theory and Applications

G. Lindgren

Statistical Analysis of Reliability Data

M.J. Crowder, A.C. Kimber, T.J. Sweeting, and R.L. Smith

Statistical Methods for Spatial Data Analysis

O. Schabenberger and C.A. Gotway

Statistical Methods for SPC and TQM

D. Bissell

Statistical Methods in Agriculture and Experimental Biology, Second Edition

R. Mead, R.N. Curnow, and A.M. Hasted

Statistical Process Control: Theory and Practice, Third Edition

G.B. Wetherill and D.W. Brown

Statistical Theory: A Concise Introduction

F. Abramovich and Y. Ritov

Statistical Theory, Fourth Edition

B.W. Lindgren

Statistics for Accountants

S. Letchford

Statistics for Epidemiology

N.P. Jewell

Statistics for Technology: A Course in Applied Statistics, Third Edition

C. Chatfield

Statistics in Engineering: A Practical Approach

A.V. Metcalfe

Statistics in Research and Development, Second Edition

R. Caulcutt

Stochastic Processes: An Introduction, Second Edition

P.W. Jones and P. Smith

Survival Analysis Using S: Analysis of Time-to-Event Data

M. Tableman and J.S. Kim

The Theory of Linear Models

B. Jørgensen

Time Series Analysis

H. Madsen

Time Series: Modeling, Computation, and Inference

R. Prado and M. West

Understanding Advanced Statistical Methods

P.H. Westfall and K.S.S. Henning

To my parents, Shyamali and Sunit, my wife, Sharbani, and my son, Shubho.

—Sudipto Banerjee

To my wife, Ishani.

—Anindya Roy

Preface

Linear algebra constitutes one of the core mathematical components in any modern curriculum involving statistics. Usually students studying statistics are expected to have seen at least one semester of linear algebra (or applied linear algebra) at the undergraduate level. In particular, students pursuing graduate studies in statistics or biostatistics are expected to have a sound conceptual grasp of vector spaces and subspaces associated with matrices, orthogonality, projections, quadratic forms and so on.

As the relevance and attraction of statistics as a discipline for graduate studies continues to increase for students with more diverse academic preparations, the need to accommodate their mathematical needs also keeps growing. In particular, many students find their undergraduate preparation in linear algebra rather different from what is required in graduate school. There are several excellent texts on the subject that provide as comprehensive a coverage of the subject as possible at the undergraduate level. However, some of these texts cater to a broader audience (e.g., scientists and engineers) and several formal concepts that are important in theoretical statistics are not emphasized.

There are several excellent texts on linear algebra. For example, there are classics by Halmos (1974), Hoffman and Kunze (1984) and Axler (1997) that make heavy use of vector spaces and linear transformations to provide a coordinate-free approach. A remarkable feature of the latter is that it develops the subject without using determinants at all. Then, there are the books by Strang (2005, 2009) and Meyer (2001) that make heavy use of echelon forms and canonical forms to reveal the properties of subspaces associated with a matrix. This approach is tangible, but may not turn out to be the most convenient to derive and prove results often encountered in statistical modeling. Among texts geared toward statistics, Rao and Bhimsankaram (2000), Searle (1982) and Graybill (2001) have stood the test of time. The book by Harville (1997) stands out in its breadth of coverage and is already considered a modern classic. Several other excellent texts exist for statisticians including Healy (2000), Abadir and Magnus (2005), Schott (2005) and Gentle (2010). The concise text by Bapat (2012) is a delightful blend of linear algebra and statistical linear models.

While the above texts offer excellent coverage, some expect substantial mathematical maturity from the reader. Our attempt here has been to offer a more gradual exposition to linear algebra without really dumbing down the subject. The book tries to be as self-contained as possible and does not assume any prior knowledge of linear

algebra. However, those who have seen some elementary linear algebra will be able to move more quickly through the early chapters. We have attempted to present both the vector space approach as well as the canonical forms in matrix theory.

Although we adopt the vector space approach for much of the later development, the book does not begin with vector spaces. Instead, it addresses the rudimentary mechanics of linear systems using Gaussian elimination and the resultant decompositions (Chapters 1–3). Chapter 4 introduces Euclidean vector spaces using less abstract concepts and makes connections to systems of linear equations wherever possible. Chapter 5 is on the rank of a matrix. Why devote an entire chapter to rank? We believe that the concept of rank is *that* important for a thorough understanding. In several cases we show how the same result may be derived using multiple techniques, which, we hope, will offer insight into the subject and ensure a better conceptual grasp of the material. Chapter 6 introduces complementary subspaces and oblique projectors. Chapter 7 introduces orthogonality and orthogonal projections and leads us to the Fundamental Theorem of Linear Algebra, which connects the four fundamental subspaces associated with a matrix. Chapter 8 builds upon the previous chapter and focuses on orthogonal projectors, which is fundamental to linear statistical models, and also introduces several computational techniques for orthogonal reduction. Chapter 9 revisits linear equations from a more mature perspective and shows how the theoretical concepts developed thus far can be handy in analyzing solutions for linear systems. The reader, at this point, will have realized that there is much more to linear equations than Gaussian elimination and echelon forms. Chapter 10 discusses determinants. Unlike some classical texts, we introduce determinants a bit late in the game and present it as a useful tool for characterizing and obtaining certain useful results. Chapter 11 introduces eigenvalues and eigenvectors and is the first time complex numbers make an appearance. Results on general real matrices are followed by those of real symmetric matrices. The popular algorithms for eigenvalues and eigenvectors are outlined both for symmetric and unsymmetric matrices. Chapter 12 derives the Singular value decomposition and the Jordan Canonical Form and presents an accessible proof of the latter. Chapter 13 is devoted to quadratic forms, another topic of fundamental importance to statistical theory and methods. Chapter 14 presents Kronecker and Hadamard products and other related materials that have become conspicuous in multivariate statistics and econometrics. Chapters 15 and 16 provide a taste of some more advanced topics but, hopefully, in a more accessible manner than more advanced texts. The former presents some aspects of linear iterative systems and convergence of matrices, while the latter introduces more general vector spaces, linear transformations and Hilbert spaces.

We remark that this is not a book on matrix computations, although we describe several numerical procedures in some detail. We have refrained from undertaking a thorough exploration of the most numerically stable algorithms as they would require a lot more theory and be too much of a digression. However, readers who grasp the material provided here should find it easier to study more specialized texts on matrix computations (e.g., Golub and Van Loan, 2013; Trefthen and Bau III, 1997). Also, while we have included many exercises that can be solved using languages such as

MATLAB and R, we decided not to marry the text to a specific language or platform. We have also not included statistical theory and applications here. This decision was taken neither in haste nor without deliberation. There are plenty of excellent texts on the theory of linear models, regression and modeling that make abundant use of linear algebra. Our hope is that readers of this text will find it easier to grasp the material in such texts. In fact, we believe that this book can be used as a companion text in the more theoretical courses on linear regression or, perhaps, stand alone as a one-semester course devoted to linear algebra for statistics and econometrics.

Finally, we have plenty of people to thank for this book. We have been greatly influenced by our teachers at the Indian Statistical Institute, Kolkata. The book by Rao and Bhimsankaram (2000), written by two of our former teachers whose lectures and notes are still vivid in our minds, certainly shaped our preparation. Sudipto Banerjee would also like to acknowledge Professor Alan Gelfand of Duke University with whom he has had several discussions regarding the role of linear algebra in Bayesian hierarchical models and spatial statistics. The first author also thanks Dr. Govindan Rangarajan of the Indian Institute of Science, Bangalore, India, Dr. Anjana Narayan of California Polytechnic State University, Pomona, and Dr. Mohan Delampady of Indian Statistical Institute, Bangalore, for allowing the author to work on this manuscript as a visitor in their respective institutes. We thank the Division of Biostatistics at the University of Minnesota, Twin Cities, and the Department of Statistics at the University of Maryland, Baltimore, for providing us with an ambience most conducive to this project. Special mention must be made of Dr. Rajarshi Guhaniyogi, Dr. Joao Monteiro and Dr. Qian Ren, former graduate students at the University of Minnesota, who have painstakingly helped with proof-reading and typesetting parts of the text. This book would also not have happened without the incredible patience and cooperation of Rob Calver, Rachel Holt, Sarah Gelson, Kate Gallo, Charlotte Byrnes and Shashi Kumar at CRC Press/Chapman and Hall. Finally, we thank our families, whose ongoing love and support made all of this possible.

SUDIPTO BANERJEE
ANINDYA ROY

Minneapolis, Minnesota
Baltimore, Maryland
August, 2013.

Contents

Preface	xv
1 Matrices, Vectors and Their Operations	1
1.1 Basic definitions and notations	2
1.2 Matrix addition and scalar-matrix multiplication	5
1.3 Matrix multiplication	7
1.4 Partitioned matrices	14
1.4.1 2×2 partitioned matrices	14
1.4.2 General partitioned matrices	16
1.5 The “trace” of a square matrix	18
1.6 Some special matrices	20
1.6.1 Permutation matrices	20
1.6.2 Triangular matrices	22
1.6.3 Hessenberg matrices	24
1.6.4 Sparse matrices	26
1.6.5 Banded matrices	27
1.7 Exercises	29
2 Systems of Linear Equations	33
2.1 Introduction	33
2.2 Gaussian elimination	34
2.3 Gauss-Jordan elimination	42
2.4 Elementary matrices	44
2.5 Homogeneous linear systems	48
2.6 The inverse of a matrix	51
2.7 Exercises	61

3	More on Linear Equations	63
3.1	The LU decomposition	63
3.2	Crout's Algorithm	71
3.3	LU decomposition with row interchanges	72
3.4	The LDU and Cholesky factorizations	77
3.5	Inverse of partitioned matrices	79
3.6	The LDU decomposition for partitioned matrices	81
3.7	The Sherman-Woodbury-Morrison formula	81
3.8	Exercises	83
4	Euclidean Spaces	87
4.1	Introduction	87
4.2	Vector addition and scalar multiplication	87
4.3	Linear spaces and subspaces	89
4.4	Intersection and sum of subspaces	92
4.5	Linear combinations and spans	94
4.6	Four fundamental subspaces	97
4.7	Linear independence	103
4.8	Basis and dimension	116
4.9	Change of basis and similar matrices	123
4.10	Exercises	124
5	The Rank of a Matrix	129
5.1	Rank and nullity of a matrix	129
5.2	Bases for the four fundamental subspaces	138
5.3	Rank and inverse	141
5.4	Rank factorization	144
5.5	The rank-normal form	148
5.6	Rank of a partitioned matrix	150
5.7	Bases for the fundamental subspaces using the rank normal form	151
5.8	Exercises	152

6	Complementary Subspaces	155
6.1	Sum of subspaces	155
6.2	The dimension of the sum of subspaces	156
6.3	Direct sums and complements	159
6.4	Projectors	163
6.5	The column space-null space decomposition	170
6.6	Invariant subspaces and the Core-Nilpotent decomposition	172
6.7	Exercises	176
7	Orthogonality, Orthogonal Subspaces and Projections	179
7.1	Inner product, norms and orthogonality	179
7.2	Row rank = column rank: A proof using orthogonality	184
7.3	Orthogonal projections	185
7.4	Gram-Schmidt orthogonalization	191
7.5	Orthocomplementary subspaces	197
7.6	The Fundamental Theorem of Linear Algebra	200
7.7	Exercises	203
8	More on Orthogonality	207
8.1	Orthogonal matrices	207
8.2	The QR decomposition	211
8.3	Orthogonal projection and projector	215
8.4	Orthogonal projector: Alternative derivations	220
8.5	Sum of orthogonal projectors	222
8.6	Orthogonal triangularization	225
8.6.1	The modified Gram-Schmidt process	226
8.6.2	Reflectors	227
8.6.3	Rotations	232
8.6.4	The rectangular QR decomposition	239
8.6.5	Computational effort	240
8.7	Orthogonal similarity reduction to Hessenberg forms	243
8.8	Orthogonal reduction to bidiagonal forms	248
8.9	Some further reading on statistical linear models	252
8.10	Exercises	253

9	Revisiting Linear Equations	257
9.1	Introduction	257
9.2	Null spaces and the general solution of linear systems	257
9.3	Rank and linear systems	259
9.4	Generalized inverse of a matrix	263
9.5	Generalized inverses and linear systems	268
9.6	The Moore-Penrose inverse	270
9.7	Exercises	274
10	Determinants	277
10.1	Introduction	277
10.2	Some basic properties of determinants	280
10.3	Determinant of products	288
10.4	Computing determinants	290
10.5	The determinant of the transpose of a matrix—revisited	291
10.6	Determinants of partitioned matrices	292
10.7	Cofactors and expansion theorems	297
10.8	The minor and the rank of a matrix	300
10.9	The Cauchy-Binet formula	303
10.10	The Laplace expansion	306
10.11	Exercises	308
11	Eigenvalues and Eigenvectors	311
11.1	The Eigenvalue equation	313
11.2	Characteristic polynomial and its roots	316
11.3	Eigenspaces and multiplicities	320
11.4	Diagonalizable matrices	325
11.5	Similarity with triangular matrices	334
11.6	Matrix polynomials and the Cayley-Hamilton Theorem	338
11.7	Spectral decomposition of real symmetric matrices	348
11.8	Computation of eigenvalues	352
11.9	Exercises	368