

SHAYLE R. SEARLE

MATRIX ALGEBRA USEFUL FOR STATISTICS

**WILEY SERIES IN PROBABILITY
AND MATHEMATICAL STATISTICS**



Matrix Algebra Useful for Statistics

SHAYLE R. SEARLE

Professor of Biological Statistics

Biometrics Unit

New York State College of Agriculture and Life Sciences

Cornell University,

Ithaca, N.Y.



JOHN WILEY & SONS

New York • Chichester • Brisbane • Toronto • Singapore

Copyright © 1982 by John Wiley & Sons, Inc.

All rights reserved. Published simultaneously in Canada.

Reproduction or translation of any part of this work beyond that permitted by Section 107 or 108 of the 1976 United States Copyright Act without the permission of the copyright owner is unlawful. Requests for permission or further information should be addressed to the Permissions Department, John Wiley & Sons, Inc.

Library of Congress Cataloging in Publication Data:

Searle, S. R. (Shayle R.), 1928–

Matrix algebra useful for statistics.

(Wiley series in probability and mathematical statistics. Applied probability and statistics, ISSN 0271-6356)

Bibliography: p.

Includes index.

1. Matrices. 2. Statistics. I. Title. II. Series.

QA188.S4 512.9'434 82-4862

ISBN 0-471-86681-4 AACR2

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

To my wife Helen

Preface

Algebra is a mathematical shorthand for language, and matrices are a shorthand for algebra. Consequently, a special value of matrices is that they enable many mathematical operations, especially those arising in statistics and the quantitative sciences, to be expressed concisely and with clarity. The algebra of matrices is, of course, in no way new, but its presentation is often so surrounded by the trappings of mathematical generality that assimilation can be difficult for readers who have only limited ability or training in mathematics. Yet many such people nowadays find a knowledge of matrix algebra necessary for their work, especially where statistics and/or computers are involved. It is to these people that I address this book; and for them, I have attempted to keep the mathematical presentation as informal as possible.

The pursuit of knowledge frequently involves collecting data; and those responsible for the collecting must appreciate the need for analyzing their data to recover and interpret the information contained therein. Such people must therefore understand some of the mathematical tools necessary for this analysis, to an extent either that they can carry out their own analysis, or that they can converse with statisticians and mathematicians whose help will otherwise be needed. One of the necessary tools is matrix algebra. It is becoming as necessary to science today as elementary calculus has been for generations. Matrices originated in mathematics more than a century ago, but their broad adaptation to science is relatively recent, prompted by the widespread acceptance of statistical analysis of data, and of computers to do that analysis; both statistics and computing rely heavily on matrix algebra. The purpose of this book is therefore that of bringing to a broad spectrum of readers a knowledge of matrix algebra that is useful in the statistical analysis of data and in statistics generally.

The basic prerequisite for using the book is high school algebra. Differential calculus is used on only a few pages, which can easily be omitted;

nothing will be lost insofar as a general understanding of matrix algebra is concerned. Proofs and demonstrations of most of the theory are given, for without them the presentation would be lifeless. But in every chapter the theoretical development is profusely illustrated with elementary numerical examples and with illustrations taken from a variety of applied sciences. And the last three chapters are devoted solely to uses of matrix algebra in statistics, with Chapters 14 and 15 outlining two of the most widely used statistical techniques: regression and linear models.

The mainstream of the book is its first eleven chapters, beginning with one on introductory concepts that includes a discussion of subscript and summation notation. This is followed by four chapters dealing with basic arithmetic, special matrices, determinants and inverses. Chapters 6 and 7 are on rank and canonical forms, 8 and 9 deal with generalized inverses and solving linear equations, 10 is a collection of results on partitioned matrices and 11 describes eigenvalues and eigenvectors. Background theory for Chapter 11 is collected in an appendix, Chapter 11A, some summaries and miscellaneous topics make up Chapter 12, statistical illustrations constitute Chapter 13, and Chapters 14 and 15 describe regression and linear models. All chapters except the last two end with exercises.

Occasional sections and paragraphs can be omitted at a first reading, especially by those whose experience in mathematics is somewhat limited. These portions of the book are printed in small type and, generally speaking, contain material subsidiary to the main flow of the text—material that may be a little more advanced in mathematical presentation than the general level otherwise maintained.

Chapters, and sections within chapters, are numbered with Arabic numerals 1, 2, 3, Within-chapter references to sections are by section number, but references across chapters use the decimal system; e.g., Section 1.3 is Section 3 of Chapter 1. These numbers are also shown in the running head of each page; e.g., [1.3] is found on page 4. Numbered equations are (1), (2), . . . , within each chapter. Those of one chapter are seldom referred to in another, but when they are, the chapter reference is explicit; otherwise “equation (3),” or more simply “(3),” means the equation numbered (3) in the chapter concerned. Exercises are in un-numbered sections and are referenced by their chapter number; e.g., Exercise 6.2 is Exercise 2 at the end of Chapter 6.

I am greatly indebted to George P. H. Styan for his exquisitely thorough readings of two drafts of the manuscript and his extensive and very helpful array of comments. Harold V. Henderson’s numerous suggestions for the final manuscript were equally as helpful. Readers of *Matrix Algebra for the*

Biological Sciences (Wiley, 1966), and students in fifteen years of my matrix algebra course at Cornell have also contributed many useful ideas. Particular thanks go to Mrs. Helen Seamon for her superb accuracy on the typewriter, her patience and fantastic attention to detail; such attributes are greatly appreciated.

Ithaca, New York
May 1982

SHAYLE R. SEARLE

Contents

1. Introduction	1
1. The scope of matrix algebra	1
2. General description of a matrix	3
3. Subscript notation	4
4. Summation notation	6
5. Dot notation	11
6. Definition of a matrix	12
7. Vectors and scalars	15
8. General notation	16
9. Illustrative examples	16
Exercises	17
2. Basic Operations	22
1. The transpose of a matrix	22
<i>a. A reflexive operation, 23</i>	
<i>b. Vectors, 24</i>	
2. Partitioned matrices	24
<i>a. Example, 24</i>	
<i>b. General specification, 26</i>	
<i>c. Transposing a partitioned matrix, 26</i>	
<i>d. Partitioning into vectors, 26</i>	
3. The trace of a matrix	27
4. Addition	28
5. Scalar multiplication	29
6. Subtraction	30
7. Equality and the null matrix	31

2. Basic Operations (Continued)

8. Multiplication	32
<i>a. The inner product of two vectors,</i>	<i>32</i>
<i>b. A matrix-vector product,</i>	<i>33</i>
<i>c. A product of two matrices,</i>	<i>36</i>
<i>d. Existence of matrix products,</i>	<i>38</i>
<i>e. Products with vectors,</i>	<i>39</i>
<i>f. Products with scalars,</i>	<i>42</i>
<i>g. Products with null matrices,</i>	<i>43</i>
<i>h. Products with diagonal matrices,</i>	<i>43</i>
<i>i. Identity matrices,</i>	<i>44</i>
<i>j. The transpose of a product,</i>	<i>44</i>
<i>k. The trace of a product,</i>	<i>45</i>
<i>l. Powers of a matrix,</i>	<i>46</i>
<i>m. Partitioned matrices,</i>	<i>48</i>
<i>n. Hadamard products,</i>	<i>49</i>
9. The Laws of algebra	50
<i>a. Associative laws,</i>	<i>50</i>
<i>b. The distributive law,</i>	<i>51</i>
<i>c. Commutative laws,</i>	<i>51</i>
10. Contrasts with scalar algebra	52
Exercises	53

3. Special Matrices

1. Symmetric matrices	60
<i>a. Products of symmetric matrices,</i>	<i>61</i>
<i>b. Properties of AA' and $A'A$,</i>	<i>61</i>
<i>c. Products of vectors,</i>	<i>63</i>
<i>d. Sums of outer products,</i>	<i>64</i>
<i>e. Elementary vectors,</i>	<i>65</i>
<i>f. Skew-symmetric matrices,</i>	<i>65</i>
2. Matrices having all elements equal	65
3. Idempotent matrices	68
4. Orthogonal matrices	69
<i>a. Definitions,</i>	<i>69</i>
<i>b. Special cases,</i>	<i>71</i>
(i) Helmert matrices,	71
(ii) Givens matrices,	72
(iii) Householder matrices,	72
5. Quadratic forms	73
6. Positive definite matrices	76
Exercises	78

4. Determinants	84
1. Expansion by minors	84
<i>a. First- and second-order determinants, 85</i>	
<i>b. Third-order determinants, 86</i>	
<i>c. n-order determinants, 89</i>	
2. Formal definition	90
3. Basic properties	92
<i>a. Determinant of a transpose, 92</i>	
<i>b. Two rows the same, 93</i>	
<i>c. Cofactors, 93</i>	
<i>d. Adding multiples of a row (column) to a row (column), 95</i>	
<i>e. Products, 96</i>	
(i) Reduction to triangular form, 96	
(ii) Two useful lemmas, 97	
(iii) Determinant of a product, 98	
4. Elementary row operations	99
<i>a. Factorization, 101</i>	
<i>b. A row (column) of zeros, 102</i>	
<i>c. Interchanging rows (columns), 102</i>	
<i>d. Adding a row to a multiple of a row, 102</i>	
5. Examples	103
6. Diagonal expansion	106
7. The Laplace expansion	109
8. Sums and differences of determinants	111
Exercises	112
5. Inverse Matrices	119
1. Introduction: solving equations	119
2. Products equal to I	123
3. Cofactors of a determinant	125
4. Derivation of the inverse	125
5. Conditions for existence of the inverse	129
6. Properties of the inverse	130
7. Some simple special cases	131
<i>a. Inverses of order 2, 131</i>	
<i>b. Diagonal matrices, 132</i>	
<i>c. I and J matrices, 132</i>	
<i>d. Orthogonal matrices, 132</i>	
<i>e. Idempotent matrices, 133</i>	
8. Equations and algebra	133
<i>a. Solving linear equations, 133</i>	
(i) Age distributions in wild populations, 133	
(ii) Input-output analysis in economics, 134	
(iii) Least squares equations, 136	
<i>b. Algebraic simplifications, 136</i>	

5. Inverse Matrices (Continued)

- 9. Computers and inverses 139
 - a. *The arithmetic of linear equations*, 140
 - b. *Rounding error*, 142
 - (i) Addition, 143
 - (ii) Inverting a matrix, 144
 - (iii) Solving linear equations, 145
- 10. Left and right inverses 147
 - Exercises 148

6. Rank 155

- 1. Linear combinations of vectors 155
- 2. Linear transformations 157
- 3. Linear dependence and independence 159
 - a. *Definitions*, 159
 - b. *General characteristics*, 161
 - (i) Sets of vectors, 161
 - (ii) Some a 's zero, 161
 - (iii) Existence and non-uniqueness of nonzero a 's, 161
 - (iv) Null vectors, 161
- 4. Linearly dependent vectors 162
 - a. *At least two a 's are nonzero*, 162
 - b. *Vectors are linear combinations of others*, 162
 - c. *Partitioning matrices*, 163
 - d. *Zero determinants*, 164
 - e. *Inverse matrices*, 164
 - f. *Testing for dependence (simple cases)*, 164
- 5. Linearly independent (LIN) vectors 166
 - a. *Nonzero determinants and inverse matrices*, 166
 - b. *Linear combinations of LIN vectors*, 167
 - c. *A maximum number of LIN vectors*, 167
- 6. The number of LIN rows and columns in a matrix 169
- 7. The rank of a matrix 171
- 8. Rank and inverse matrices 172
- 9. Permutation matrices 173
- 10. Full-rank factorization 175
 - a. *Basic development*, 175
 - b. *The general case*, 177
 - c. *Matrices of full row (column) rank*, 177
- 11. Vector spaces 177
 - a. *Euclidean space*, 178
 - b. *Vector spaces*, 178
 - c. *Spanning sets and bases*, 179

6. Rank (*Continued*)

- d. Many spaces of order n , 179*
- e. Subspaces, 180*
- f. The range and null space of a matrix, 180*
- Exercises 181

7. Canonical Forms 184

- 1. Elementary operators 184
 - a. Row operations, 184*
 - b. Transposes, 185*
 - c. Column operations, 185*
 - d. Inverses, 185*
- 2. Rank and the elementary operators 186
 - a. Rank, 186*
 - b. Products of elementary operators, 186*
 - c. Equivalence, 187*
- 3. Finding the rank of a matrix 187
 - a. Some special LIN vectors, 187*
 - b. Calculating rank, 188*
 - c. A general procedure, 189*
- 4. Reduction to equivalent canonical form 190
 - a. Row operations, 190*
 - b. Column operations, 191*
 - c. The equivalent canonical form, 192*
 - d. Non-uniqueness of \mathbf{P} and \mathbf{Q} , 194*
 - e. Existence is assured, 194*
 - f. Full-rank factorization, 194*
- 5. Rank of a product matrix 196
- 6. Symmetric matrices 199
 - a. Row and column operations, 199*
 - b. The diagonal form, 200*
 - c. The canonical form under congruence, 201*
 - d. Two special provisions, 202*
 - (i) Null rows, 202
 - (ii) Negative elements in the diagonal form, 203
 - e. Full-rank factorization, 205*
- 7. Non-negative definite matrices 205
 - a. Diagonal elements and principal minors, 205*
 - b. Congruent canonical form, 206*
 - c. Full-rank factorization, 206*
 - d. Quadratic forms as sums of squares, 206*
 - e. Full row (column) rank matrices, 208*
- Exercises 209

8. Generalized Inverses	212
1. The Moore–Penrose inverse	212
2. Generalized inverses	212
<i>a. Derivation from row operations, 214</i>	
<i>b. Derivation from the diagonal form, 215</i>	
3. Other names and symbols	216
4. An algorithm	217
<i>a. An easy form, 217</i>	
<i>b. A general form, 217</i>	
5. Arbitrariness in a generalized inverse	219
6. Symmetric matrices	220
<i>a. Non-negative definite matrices, 220</i>	
<i>b. A general algorithm, 221</i>	
<i>c. The matrix $\mathbf{X}'\mathbf{X}$, 221</i>	
Exercises	222
 9. Solving Linear Equations	 227
1. Equations having many solutions	227
2. Consistent equations	228
<i>a. Definition, 228</i>	
<i>b. Existence of solutions, 229</i>	
<i>c. Tests for consistency, 232</i>	
3. Equations having one solution	233
4. Deriving solutions using generalized inverses	235
<i>a. Obtaining a solution, 235</i>	
<i>b. Obtaining many solutions, 236</i>	
<i>c. All possible solutions, 237</i>	
<i>d. Combinations of solutions, 239</i>	
5. Linearly independent solutions	239
6. An invariance property	242
7. Equations $\mathbf{Ax} = \mathbf{0}$	245
<i>a. General properties, 245</i>	
<i>b. Orthogonal solutions, 246</i>	
<i>c. Orthogonal vector spaces, 248</i>	
8. A complete example	249
9. Least squares equations	251
Exercises	252

10. Partitioned Matrices	257
1. Orthogonal matrices	257
2. Determinants	258
3. Inverses	260
4. Schur complements	261
5. Generalized inverses	261
6. Direct sums	263
7. Direct products	265
Exercises	267
 11. Eigenvalues and Eigenvectors	 272
1. Introduction: age distribution vectors	272
2. Derivation of eigenvalues	274
3. Elementary properties of eigenvalues	276
<i>a. Eigenvalues of powers of a matrix, 276</i>	
<i>b. Eigenvalues of a scalar-by-matrix product, 277</i>	
<i>c. Eigenvalues of polynomials, 277</i>	
<i>d. The sum and product of eigenvalues, 278</i>	
4. Calculating eigenvectors	279
<i>a. A general method, 279</i>	
<i>b. Simple roots, 280</i>	
<i>c. Multiple roots, 281</i>	
5. The similar canonical form	282
<i>a. Derivation, 282</i>	
<i>b. Uses, 285</i>	
6. Symmetric matrices	290
<i>a. Eigenvalues all real, 290</i>	
<i>b. Symmetric matrices are diagonalizable, 290</i>	
<i>c. Eigenvectors are orthogonal, 290</i>	
(i) Different eigenvalues, 291	
(ii) Multiple eigenvalues, 291	
(iii) The canonical form under orthogonal similarity, 291	
<i>d. Rank equals number of nonzero eigenvalues, 292</i>	
7. Dominant eigenvalues	293
8. Factoring the characteristic equation	298
Exercises	299

11A. Appendix to Chapter 11	305
1. Proving the diagonability theorem	305
<i>a. The number of nonzero eigenvalues never exceeds rank, 305</i>	
<i>b. A lower bound on $r(\mathbf{A} - \lambda_k \mathbf{I})$, 306</i>	
<i>c. Proof of the diagonability theorem, 307</i>	
<i>d. All symmetric matrices are diagonable, 307</i>	
2. Other results for symmetric matrices	308
<i>a. Spectral decomposition, 308</i>	
<i>b. Non-negative definite (n.n.d.) matrices, 309</i>	
<i>c. Simultaneous diagonalization of two symmetric matrices, 312</i>	
3. The Cayley–Hamilton theorem	314
4. The singular-value decomposition	316
Exercises	318
 12. Miscellanea	 320
1. Orthogonal matrices—a summary	320
2. Idempotent matrices—a summary	320
3. The matrix $a\mathbf{I} + b\mathbf{J}$ —a summary	322
4. Non-negative definite matrices—a summary	322
5. Canonical forms and other decompositions—a summary	323
6. Matrix Functions	325
<i>a. Functions of matrices, 325</i>	
<i>b. Matrices of functions, 325</i>	
7. Iterative solution of nonlinear equations	326
8. Vectors of differential operators	327
<i>a. Scalars, 327</i>	
<i>b. Vectors, 328</i>	
<i>c. Quadratic forms, 329</i>	
9. Vec and vech operators	332
<i>a. Definitions, 332</i>	
<i>b. Properties of vec, 333</i>	
<i>c. Vec-permutation matrices, 334</i>	
<i>d. Relationships between vec and vech, 334</i>	
10. Other calculus results	334
<i>a. Differentiating inverses, 335</i>	
<i>b. Differentiating traces, 335</i>	
<i>c. Differentiating determinants, 336</i>	
<i>d. Jacobians, 338</i>	
<i>e. Aitken's integral, 340</i>	
<i>f. Hessians, 341</i>	
11. Matrices with elements that are complex numbers	341
Exercises	342

13. Applications in Statistics	346
1. Variance–covariance matrices	347
2. Correlation matrices	348
3. Matrices of sums of squares and cross-products	349
<i>a. Data matrices, 349</i>	
<i>b. Uncorrected sums of squares and products, 349</i>	
<i>c. Means, and the centering matrix, 350</i>	
<i>d. Corrected sums of squares and products, 351</i>	
4. The multivariate normal distribution	352
5. Quadratic forms and χ^2 -distributions	355
6. Least squares equations	357
7. Contrasts among means	358
Exercises	360
14. The Matrix Algebra of Regression Analysis	363
1. General description	363
<i>a. Linear models, 364</i>	
<i>b. Observations, 364</i>	
<i>c. Nonlinear models, 365</i>	
2. Estimation	366
3. Several regressor variables	368
4. Deviations from means	369
5. The statistical model	371
6. Unbiasedness and variances	372
7. Predicted y -values	373
8. Estimating the error variance	374
9. Partitioning the total sum of squares	376
10. Multiple correlation	378
11. The no-intercept model	379
12. Analysis of variance	380
13. Testing linear hypotheses	382
<i>a. Stating a hypothesis, 382</i>	
<i>b. The F-statistic, 383</i>	
<i>c. Equivalent statements of a hypothesis, 385</i>	
<i>d. Special cases, 386</i>	
(i) $H: \mathbf{b} = \mathbf{0}$, 386	
(ii) $H: \mathbf{b} = \mathbf{b}_0$, 386	
(iii) $H: \mathbf{\lambda}'\mathbf{b} = m$, 386	
14. Confidence intervals	386
15. Fitting subsets of the x -variables	387
16. Reductions in sums of squares: the $R(\cdot \cdot)$ notation	389

15. An Introduction to Linear Statistical Models	392
1. General description	392
2. The normal equations	395
a. <i>A general form, 395</i>	
b. <i>Many solutions, 396</i>	
3. Solving the normal equations	397
a. <i>Generalized inverses of $\mathbf{X}'\mathbf{X}$, 397</i>	
b. <i>Solutions, 397</i>	
4. Expected values and variances	399
5. Predicted y -values	400
6. Estimating the error variance	401
a. <i>Error sum of squares, 401</i>	
b. <i>Expected value, 402</i>	
c. <i>Estimation, 402</i>	
7. Partitioning the total sum of squares	403
8. Coefficient of determination	404
9. Analysis of variance	405
10. The $R(\cdot \cdot)$ notation	407
11. Estimable functions	408
12. Testing linear hypotheses	411
13. Confidence intervals	415
14. Some particular models	416
a. <i>The one-way classification, 416</i>	
b. <i>Two-way classification, no interactions, balanced data, 418</i>	
c. <i>Two-way classification, no interactions, unbalanced data, 422</i>	
15. The $R(\cdot \cdot)$ notation (<i>Continued</i>)	424
References	429
Index	433