Journal of Statistical Software
July 2016, Volume 71, Book Review 1. doi: 10.18637/jss.v071.b01

Reviewer: Abdolvahab Khademi
University of Massachusetts

Basics of Matrix Algebra for Statistics with R

> Nick Fieller
> Chapman & Hall/CRC, Boca Raton, 2016.
> ISBN 9781498712361. xviii+226 pp. USD 59.95 (H).
> https://www.crcpress.com/9781498712361

Native array programming languages such as MATLAB and Mathematica have extensively been used for mathematical and scientific computation where the data are mainly in matrix form. Although originally conceived as a statistical programming language, R is also a native matrix language which could be exploited for mathematical applications.

Matrix algebra is widely used in statistical computation and manipulations, especially in high-dimensional data analysis. Knowledge of matrix algebra is, therefore, essential to learning, understanding, or using areas of statistics based on matrix data. Although there is a plethora of books on matrix and linear algebra that are written for a general audience who need and use matrix algebra in their fields, a few matrix algebra books have been written specifically for use in statistics. The advantage of such books over more general matrix algebra texts is that the contents are selected to address the needs for statisticians and the exercises and examples are also geared towards applications in statistics. There is another category of books that attempt to integrate a programming language such as R in presenting matrix algebra concepts and solving problems using that language instead of hand calculation. *Basics of Matrix Algebra for Statistics with R* belongs to the latter category, which attempts to present matrix mathematics concepts most related and applicable to statistics using R.

The book is organized concisely in nine chapters, each with exercises at the end in the form of calculation, proofs, and computation. Solutions to all exercises are provided with different degrees of detail depending on the type of the problem.

Chapter 1 (*Introduction*) outlines the objectives of the book and introduces the R programming language to the beginner. Topics include obtaining and installing R, the structure of the language, and inputting data. Basic examples of the implementation of matrix operations in R are also given, such as transposing, inverting, addition, and finding the trace and the determinant of a matrix.

Chapter 2 (*Vectors and matrices*) presents the fundamental concepts of vectors and matrices. General topics such as matrix arithmetic, trace, special matrices, normal matrices, manipulation of partitioned matrices, and linear and quadratic forms are introduced with corresponding

R code. Chapter 3 (*Rank of matrices*) presents the concept of matrix rank and the combination of matrices, rank factorization, and the significance of a matrix rank in statistical analysis. This very brief chapter ends with exercises on calculation and proof. Chapter 4 (*Determinants*) introduces determinants and their properties. In Chapter 5 (*Inverses*) the author presents matrix inversion, its application in linear and multivariate statistics, and its implementation in R.

The more complex concepts are introduced in the subsequent chapters. *Eigenanalysis of real symmetric matrices* in Chapter 6 presents the fundamental concepts of eigenvalues widely used in multivariate statistical methods, such as principal components analysis, linear discriminant analysis, and partial least squares. Properties of eigenanalysis, matrix decomposition, and eigenanalysis of special matrix structures frequently used in statistics are the prominent concepts presented in this chapter, ending with exercises mostly focused on proofs. The calculus of vectors and matrices is taken up in Chapter 7 (*Vector and matrix calculus*) where differentiation on scalars, vectors, and matrices is presented, with a side note on the use of differentiation in constrained optimization using eigenanalysis. Computation is minimal in this chapter and its exercises.

Chapter 8 (*Further topics*) presents three extended topics: matrix decompositions, generalized inverses, and special products. The decomposition section includes topics on QR, LU, LDU, Cholesky, and Schur decompositions with corresponding R code and applications in statistics. The second section introduces the Moore-Penrose inverse and generalized inverses. The section on product of matrices includes Hadamard products and Kronecker products with their decomposition and eigenanalysis. R code is provided for implementation.

Key applications to statistics in Chapter 9 introduces the use of matrix algebra in statical concepts and formulations such as the multivariate normal distribution, maximum likelihood estimation, and the likelihood ratio test. Statistical methods such as multi-sample tests, principal component analysis, linear discriminant analysis, canonical correlation analysis, classical scaling, linear models, and crossover clinical trials are presented.

As stated at the beginning, *Basics of Matrix Algebra for Statistics with R* belongs to the category of mathematics books that integrate a programming language with substantive content. On the substantive side, the author has meticulously selected matrix algebra topics that are fundamental to learning, using, and understanding statistics. In this manner, the reader is saved time by focusing on matrix mathematics which is of most relevance to statistics. In addition, an instructor also benefits from the concise introduction to matrix algebra related to statistics. Therefore, this book can easily be adopted as a matrix algebra supplemental book in a syllabus on statistics. The exercises are short but rigorous, with detailed solutions provided at the end of the book.

As for the programming aspect, the book fares well in integrating basic R code for basic matrix algebra concepts, especially in the first six chapters. However, as the materials become more mathematical in nature, coding receives secondary attention (and minimal in some chapters, e.g., Chapters 3 and 7). However, in Chapter 9, the reader is again given the chance to see the application of coding in implementation of several statistical methods.

Overall, the book fares better in introducing matrix algebra for statistics than integrating R coding. What could make this book more valuable for those trying to do matrix algebra through R would be the integration of more advanced code and more challenging exercises requiring R coding. Nevertheless, as a traditional text to teach practical matrix algebra to students taking multivariate and more advanced statistics courses, this book can be of good use.

Reviewer:

Abdolvahab Khademi
University of Massachusetts
College of Education &
Department of Mathematics and Statistics
Amherst MA 01002, United States of America
E-mail: khademi@math.umass.edu

Journal of Statistical Software http://www.jstatsoft.org/
published by the Foundation for Open Access Statistics http://www.foastat.org/
July 2016, Volume 71, Book Review 1 *Published:* 2016-07-04
doi:10.18637/jss.v071.b01