An Introduction to Applied Multivariate Statistics

M. S. Srivastava

E. M. Carter

An Introduction to Applied Multivariate Statistics

M. S. Srivastava
University of Toronto

and

E. M. Carter
University of Guelph

North-Holland
New York · Amsterdam · Oxford

Elsevier Science Publishing Co., Inc.
52 Vanderbilt Avenue, New York, NY 10017

Distributors outside the United States and Canada:

Elsevier Science Publishers B. V.
Post Office Box 211, 1000 AE Amsterdam, The Netherlands

Library of Congress Cataloging in Publication Data

Srivastava, M. S.
An introduction to applied multivariate statistics.

Bibliography: p.
Includes index.
1. Multivariate analysis. I. Carter, E. M. (Edward M.)
II. Title.
QA278.S687 519.5'35 82-2385
ISBN 0-444-00621-4 AACR2

Manufactured in the United States of America

Preface

This book is based on lectures given by the authors to first-year graduate students at the University of Toronto and the University of Guelph. Our primary aim is to teach those multivariate techniques applicable to the data available in such varied disciplines as forestry, biology, medicine, and education. Theoretical details have accordingly been kept to a bare minimum. A knowledge of matrix notation and manipulations will be helpful, and Chapter 1 should assist readers deficient in this area. The reader is, however, expected to have sufficient knowledge of elementary univariate theory.

Our emphasis is on methods in current use in multivariate statistics. For each new topic we present not only a description of the problem and its solution, but also several worked examples, chosen from many different fields. Each chapter ends with a discussion of the computer packages available and additional worked examples.

The likelihood ratio approach has been adopted for tests of the significance of a given hypothesis, and Roy's union–intersection principle and Bonferoni's inequalities for confidence intervals are introduced. For each test statistic, formulas for the observed significance levels are given. The percentage points needed to calculate confidence intervals appear as Appendixes.

Chapters 1 and 2 provide a review of the necessary matrix theory and statistical theory; Chapter 1 also contains a discussion of the SAS matrix procedures for calculating eigenvalues, eigenvectors, and the generalized inverse of a matrix. Chapters 3–7 are multivariate generalizations of univariate procedures for t-tests, analysis of variance, and multiple regression.

Chapter 12 gives tests of the assumptions required for these multivariate procedures to be valid, including tests for equality of covariance and independence. It is Chapters 8–11 that contain strictly multivariate procedures, beginning with discriminant analysis. Methods for finding functions that will discriminate among populations or groups are discussed, and procedures to reduce the number of characteristics necessary for discrimination are given in a section on stepwise discriminant analysis. Chapters 9–11 then introduce dimension-reducing procedures, including canonical correlation, which investigates the correlation between linear combinations of variables, and principal component analysis, which reduces the set of measured characteristics to fewer components. For example, 20 measurements on bulls might be reduced to two or three components for size and shape. Factor analysis performs the same reduction but assuming that the observations have an underlying structure; this method is often used to group responses to questionnaires and psychological tests.

Advanced topics are given toward the end of a chapter. For example, Chapter 3 covers the problem of incomplete data and tests for shift in the mean. The core material may be used in a one-semester course in applied multivariate statistics at the senior or first-year graduate level. Two semesters are required to cover the text in its entirety.

This book is dedicated to Jagdish Bahadur Srivastava and Ivy May Carter.

Contents

Chapter 1
Some Results on Matrices

1.1 Notation and Definitions

Suppose that $a_{11}, a_{12}, \ldots, a_{pq}$ is a collection of pq real numbers. The rectangular array of these elements

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1q} \\ a_{21} & a_{22} & \cdots & a_{2q} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pq} \end{pmatrix},$$

consisting of p rows and q columns, is called a $p \times q$ matrix. We write $A = (a_{ij})$: $p \times q$, where a_{ij} is the element in the ith row and jth column. For example, the matrix

$$A = \begin{pmatrix} 6 & 8 & 9 \\ 1 & 3 & 5 \end{pmatrix}$$

is a 2×3 matrix with $a_{11} = 6$, $a_{12} = 8$, $a_{13} = 9$, $a_{21} = 1$, $a_{22} = 3$, $a_{23} = 5$.

We shall now define some special matrices that will be used later.

Null Matrix If all the elements of A are zero, then A is said to be a zero or null matrix, denoted 0_{pq} or, if there is no confusion, simply 0.

That is,

$$0_{32} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Square Matrix If $p = q$, then A is said to be a square matrix of order p.

Column Vector If $q=1$, then A is said to be a p-column vector, or simply a p-vector, and the vector will be written

$$\mathbf{a}=\begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix}.$$

For example, the 3-vector with entries 6, 8, and 9 is written

$$\mathbf{a}=\begin{pmatrix} 6 \\ 8 \\ 9 \end{pmatrix}.$$

Row Vector If $p=1$, then A is said to be a q-row vector. A will be written $\mathbf{a}'=(a_1,\ldots,a_q)$.

For example, the 4-row vector $\mathbf{a}'$ with entries 2, 4, 6, and 8 is written

$$\mathbf{a}'=(2,4,6,8).$$

Lower Triangular Matrix A square matrix with all elements above the main diagonal equal to zero is called a lower triangular matrix.

Examples are

$$\begin{pmatrix} 2 & 0 & 0 \\ 8 & 4 & 0 \\ 10 & 9 & 6 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 5 & 0 \end{pmatrix}.$$

Upper Triangular Matrix A square matrix with all elements below the main diagonal equal to zero is called an upper triangular matrix.

Diagonal Matrix A square matrix A with the off-diagonal elements equal to zero is called a diagonal matrix. If the diagonal entries are $a_1,\ldots,a_p$, then A is sometimes written D_a or $\mathrm{diag}(a_1,\ldots,a_p)$.

For example, if $a_1=1$, $a_2=2$, and $a_3=4$, then

$$A=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

Identity Matrix If A is a diagonal $p\times p$ matrix with all p diagonal entries equal to 1, then A is called an identity matrix, denoted I_p.

For example,

$$I_1 = [1], \qquad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Transpose Matrix If the rows and columns of a matrix are interchanged, the resulting matrix is called the transpose of A, denoted A'. Thus if $A = (a_{ij}) : p \times q$, then $A' = (a_{ji}) : q \times p$.

For example, if

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

then

$$A' = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Symmetric Matrix A square matrix is said to be symmetric if $A = A'$.

Examples of symmetric matrices are

$$\begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}.$$

Skew Symmetric Matrix A matrix A is said to be skew symmetric if $A = -A'$. If A is skew symmetric, then all the diagonal entries are zero.

Examples are

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 2 \\ -1 & -2 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & -1 & -2 \\ 1 & 0 & 3 \\ 2 & -3 & 0 \end{pmatrix}.$$

1.2 Matrix Operations

Sometimes, it is convenient to represent a matrix A as

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right),$$

where A_{ij} is an $m_i \times m_j$ submatrix of A $(i, j = 1, 2)$. A submatrix of A is a matrix obtained from A by deleting certain rows and columns. The above representation of A is called a partitioned matrix. If there are more than two partitions of rows (or columns), we can write $A = (A_{ij})$, where A_{ij} is the submatrix of A in the ith row and jth column partition. For example, if

$$A_{11}=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix},\qquad A_{12}=\begin{pmatrix}2\\ 3\end{pmatrix},\qquad A_{21}=(6,7),\qquad A_{22}=(8),$$

then

$$A=\left(\begin{array}{cc|c}1 & 2 & 2\\ 3 & 4 & 3\\ \hline 6 & 7 & 8\end{array}\right).$$

Another type of partitioned matrix can be written

$$A=\left(\begin{array}{c|c}A_{11} & \\ \cline{1-1} A_{21} & A_2\end{array}\right).$$

For example, if

$$A_{11}=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix},\qquad A_{21}=(6,7),\qquad A_2=\begin{pmatrix}2\\ 3\\ 8\end{pmatrix},$$

then

$$A=\left(\begin{array}{cc|c}1 & 2 & 2\\ 3 & 4 & 3\\ \cline{1-2} 6 & 7 & 8\end{array}\right).$$

Let $A: m\times n$ and $B: p\times q$ be two matrices. The usual addition of two matrices is defined only if the matrices are of the same order. Thus if $A=(a_{ij}): p\times q$ and $B=(b_{ij}): p\times q$, then their *sum* is defined by

$$A+B=(a_{ij}+b_{ij}): p\times q.$$

The usual multiplication of A by B is defined only if the number of columns of A equal to the number of rows of B. Thus if $A=(a_{ij}): m\times n$ and $B=(b_{ij}): n\times q$, then the product of A and B (denoted AB) is defined as an $m\times q$ matrix $AB=C=(c_{ij})$, where

$$c_{ij}=\sum_{k=1}^{n} a_{ik}b_{kj},\qquad i=1,2,\ldots,m,\quad j=1,2,\ldots,q.$$

The following results can be obtained from the above definitions:

i. $(A+B)'=A'+B'$, $(A+B+C)'=A'+B'+C'$;
ii. $(AB)'=B'A'$, $(ABC)'=C'B'A'$;
iii. $A(B_1+B_2)=AB_1+AB_2$; and
iv. $\Sigma_{\alpha=1}^{k}AB_\alpha=A(\Sigma_{\alpha=1}^{k}B_\alpha)$.

Examples Let

$$A=\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{pmatrix},\qquad B=\begin{pmatrix}7 & 8 & 9\\ 10 & 11 & 12\end{pmatrix},\qquad C=\begin{pmatrix}3 & 1\\ 1 & 2\end{pmatrix}.$$

Then

1. $A' = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}, \; B' = \begin{pmatrix} 7 & 10 \\ 8 & 11 \\ 9 & 12 \end{pmatrix};$
2. $A + B = \begin{pmatrix} 8 & 10 & 12 \\ 14 & 16 & 18 \end{pmatrix};$
3. $(A+B)' = A' + B' = \begin{pmatrix} 8 & 14 \\ 10 & 16 \\ 12 & 18 \end{pmatrix};$
4. $CA = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 7 & 11 & 15 \\ 9 & 12 & 15 \end{pmatrix}.$

Note that AC is not defined.

We now define a few more matrices.

Semiorthogonal Matrix A matrix $A: p \times q$ is said to be semiorthogonal if $AA' = I_p$ $(q \geq p)$.

Examples are

$$(0.5, 0.5, 0.5, 0.5), \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Orthogonal Matrix A square matrix A is said to be orthogonal if $AA' = I_p$.

Idempotent Matrix A square matrix A is said to be idempotent if $A = A^2$.

Examples are

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Kronecker Product Let $A: m \times n$ and $B: p \times q$ be two matrices. Then the Kronecker product (or direct product) of A and B is defined as the $mp \times nq$ matrix $A \otimes B = (a_{ij}B)$, where $A = (a_{ij})$.

For example, for

$$A = \begin{pmatrix} 4 & 2 \\ 3 & 1 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

then

$$A \otimes B = \left(\begin{array}{cc|cc} 4 & 0 & 2 & 0 \\ 0 & 4 & 0 & 2 \\ \hline 3 & 0 & 1 & 0 \\ 0 & 3 & 0 & 1 \end{array}\right).$$

1.3 Determinants

The determinant of a square matrix $A=(a_{ij}): n\times n$ is defined as

$$|A| \equiv \sum_{\alpha}(-1)^{N(\alpha)} \prod_{j=1}^{n} a_{\alpha_j, j},$$

where Σ_{α} denotes the summation over the distinct permutations $\boldsymbol{\alpha}$ of the numbers $1,2,\ldots,n$ and $N(\boldsymbol{\alpha})$ is the total number of inversions of a permutation. An inversion of a permutation $\boldsymbol{\alpha}=(\alpha_1,\alpha_2,\ldots,\alpha_n)$ is an arrangement of two indices such that the larger index comes after the smaller index. For example, $N(2,1,4,3)=1+N(1,2,4,3)=2+N(1,2,3,4)=2$ because $N(1,2,3,4)=0$. Similarly, $N(4,3,1,2)=1+N(3,4,1,2)=3+N(3,1,2,4)=5$. The above expression of a determinant is denoted $|A|$ or $\det A$. If $|A|$ is real, then $|A|_{+}$ denotes the positive value of $|A|$. Note that

$$|A'| = \sum_{\alpha}(-1)^{N(\alpha)}\prod a_{j,\alpha_j} = \sum_{\alpha}(-1)^{N(\alpha)}\prod a_{\alpha_j, j} = |A|.$$

Examples

1. $\begin{vmatrix} 2 & 1 \\ 1 & 5 \end{vmatrix} = 2\times 5 - 1\times 1 = 9,$
2. $\begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 4-1=3.$

The following are immediate consequences of the definition of $|A|$.

i. If the ith row (or column) is multiplied by a constant c, the value of the determinant is multiplied by c. Hence $|cA|=c^n|A|$ if A is an $n\times n$ matrix.

ii. If any two rows (or columns) of a matrix are interchanged, the sign of the determinant is changed. Hence if two rows (or columns) of a matrix are identical, the value of the determinant is zero.

iii. The value of the determinant is unchanged if in the ith row (or column) a cth multiple of the jth row (or column) is added. Hence the value of the determinant is zero if a row (or column) is a linear combination of other rows (or columns).

iv. $|I_n|=1$, $|D_a|=\prod a_i$.

v. $|AB|=|A||B|$ if $A: p\times p$ and $B: p\times p$,

vi. $|AA'|\geq 0$.

vii.

$$\begin{vmatrix} I & 0 \\ C & A \end{vmatrix} = |A|, \qquad \begin{vmatrix} A & C \\ 0 & I \end{vmatrix} = |A|, \qquad \begin{vmatrix} A & C \\ 0 & B \end{vmatrix} = \begin{vmatrix} A & 0 \\ C & B \end{vmatrix} = |A||B|,$$

where A and B are square matrices.

viii. $|I_p + AB| = |I_q + BA|$, where A and B are $p\times q$ and $q\times p$ matrices.

1.3.1 Cofactors of a Square Matrix

Let A^i_j be the submatrix of $A: n \times n$ obtained by deleting the ith row and the jth column of A. Let $m_{ij} = |A^i_j|$. Then $c_{ij}(A) = (-1)^{i+j}|A^i_j| = (-1)^{i+j} m_{ij}$ is known as the *cofactor* of a_{ij}. Note that

$$|A| = \sum_{j=1}^{n} c_{ij}(A) a_{ij} = \sum_{i=1}^{n} c_{ij}(A) a_{ij},$$

$$0 = \sum_{j=1}^{n} c_{ij} a_{kj}, \quad k \neq i, \qquad 0 = \sum_{i=1}^{n} c_{ij} a_{ik}, \quad j \neq k.$$

Example Let

$$A = \begin{pmatrix} 1 & 6 & 5 & 7 \\ 6 & 9 & 10 & 12 \\ 3 & 7 & 8 & 10 \\ 2 & 5 & 9 & 11 \end{pmatrix},$$

and let $i = 2$ and $j = 3$. Then

$$A^2_3 = \begin{pmatrix} 1 & 6 & 7 \\ 3 & 7 & 10 \\ 2 & 5 & 11 \end{pmatrix},$$

and $c_{23}(A) = (-1)^{2+3}|A^2_3|$.

1.3.2 Minor, Principal Minor, and Trace of a Matrix

Let $A^{(i_1,\ldots,i_t)}_{(j_1,\ldots,j_t)}$ be the submatrix of $A: m \times n$ obtained by taking the $i_1, i_2, \ldots, i_t$ rows and $j_1, j_2, \ldots, j_t$ columns; note that it is a square submatrix of A. Then $|A^{(i_1,\ldots,i_t)}_{(j_1,\ldots,j_t)}|$ is know as a *minor* of order t. If $i_1 = j_1, i_2 = j_2, \ldots, i_t = j_t$, then it is known as a *principal minor* of order t.

For any square matrix, the sum of all principal minors of order t is called the tth trace of A. Symbolically, $\operatorname{tr}_t(A) = \Sigma_\alpha |A^{(i_1,\ldots,i_t)}_{(i_1,\ldots,i_t)}|$. Thus, when $t = 1$, we have $\operatorname{tr}_1 A \equiv \operatorname{tr} A = \Sigma_{i=1}^{n} a_{ii}$.

Example Let A be the 4×4 matrix given in Section 1.3.1. Let $i_1 = 2$, $i_2 = 3$, $j_1 = 2$, and $j_2 = 3$. Then

$$A^{(2,3)}_{(2,3)} = \begin{pmatrix} 9 & 10 \\ 7 & 8 \end{pmatrix}, \qquad A^{(1)}_{(1)} = 1, \quad A^{(2)}_{(2)} = 9$$

and $\begin{vmatrix} 9 & 10 \\ 7 & 8 \end{vmatrix}$ is a principal minor of order 2. The sum of all principal minors of order 2,

$$\begin{vmatrix} 9 & 10 \\ 7 & 8 \end{vmatrix} + \begin{vmatrix} 1 & 6 \\ 6 & 9 \end{vmatrix} + \begin{vmatrix} 1 & 5 \\ 3 & 8 \end{vmatrix} + \begin{vmatrix} 9 & 12 \\ 5 & 11 \end{vmatrix} + \begin{vmatrix} 1 & 7 \\ 2 & 11 \end{vmatrix} + \begin{vmatrix} 8 & 10 \\ 9 & 11 \end{vmatrix} = 4 = \operatorname{tr}_2(A),$$

and $\operatorname{tr}_1 A = 1 + 9 + 8 + 11$.

From the definition of a determinant, we get the following property of a determinant.

Let A be a square matrix of order n. Then the determinant $|A+\lambda I_n|$ is a polynomial of degree n in λ and can be written

$$|A+\lambda I_n| = \sum_{i=1}^{n} \lambda^i \operatorname{tr}_{n-i} A, \quad \text{where} \quad \operatorname{tr}_0 A = 1.$$

If $\lambda_1, \lambda_2, \ldots, \lambda_n$ denote the roots of this polynomial, then

$$\begin{aligned} \operatorname{tr}_1(A) &= \sum_{i=1}^{n} \lambda_i, \\ \operatorname{tr}_2(A) &= \lambda_1\lambda_2 + \lambda_1\lambda_3 + \cdots + \lambda_{n-1}\lambda_n, \\ &\vdots \\ \operatorname{tr}_n(A) &= \lambda_1\lambda_2\cdots\lambda_n = |A|. \end{aligned}$$

We shall write $\operatorname{tr} A$ for $\operatorname{tr}_1 A$. It can easily be verified that

i. $\operatorname{tr} AB = \operatorname{tr} BA$,
ii. $\operatorname{tr} ABC = \operatorname{tr} BCA = \operatorname{tr} CAB$,
iii. $\operatorname{tr} A = \operatorname{tr} A'$,
iv. $\operatorname{tr}(A+B) = \operatorname{tr} A + \operatorname{tr} B$,
v. $\operatorname{tr}(A+B+C) = \operatorname{tr} A + \operatorname{tr} B + \operatorname{tr} C$,
vi. $\operatorname{tr}(\sum_{\alpha=1}^{k} A_\alpha) = \sum_{\alpha=1}^{k}(\operatorname{tr} A_\alpha)$, and
vii. $\operatorname{tr} c = c$, where c is a scalar.

1.4 Rank of a Matrix

An $m \times n$ matrix A is said to be of rank r, denoted $\rho(A) = r$, if and only if (iff) there is at least one nonzero minor of order r from A and all of the minors of order $r+1$ are zero. It is easy to establish the following:

i. $\rho(A) = 0$ iff $A = 0$.
ii. If A is an $m \times n$ matrix and $A \neq 0$, then $1 \leq \rho(A) \leq \min(m, n)$. Further, $\rho(A) = \rho(A')$.
iii. $\max(\rho(A), \rho(B)) \leq \rho(A \vdots B) \leq \min(n, \rho(A) + \rho(B))$, where n is the number of rows in A.
iv.
$$\rho\begin{pmatrix} 0 & Q \\ R & 0 \end{pmatrix} = \rho\begin{pmatrix} Q & 0 \\ 0 & R \end{pmatrix} = \rho(R) + \rho(Q).$$
v. $\rho(AB) \leq \min(\rho(A), \rho(B))$.
vi. $\rho(AB) = \rho(A)$ if $\rho(B) = p$, where $B: p \times q$ and $p \leq q$.

Note that $\rho(A)$ need not be equal to $\rho(A^2)$. For example, if

$$A = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix},$$