Appendix A
Gentle Introduction to Vectors and Matrices

We present a very brief introduction to vectors and matrices, intended for readers with no prior exposure to matrix algebra. Specifically, we cover basic definitions and summarize some of the main properties of vectors and matrices. Vectors and matrices allow us to perform common mathematical operations (e.g., addition, subtraction, and multiplication) on a collection of numbers; they also facilitate the description of statistical methods for multivariate data. Our primary motivation for using them is the conciseness and compactness with which statistical techniques for analyzing longitudinal data can be presented when expressed in terms of vectors and matrices.

Mastery of the material presented in this section is a prerequisite for understanding the statistical methods for longitudinal data described in the book. Although we do not assume a profound understanding of matrix algebra, vectors and matrices are used extensively throughout the book to simplify notation and the reader is required to have some basic facility with the addition and multiplication of vectors and matrices.

Basic Concepts and Definitions

A *matrix* is a rectangular array of elements (e.g., numbers), arranged in rows and columns. For example,

$$A = \begin{pmatrix} 2 & 7 & 11 & 5 \\ 4 & 9 & 3 & 1 \\ 13 & 8 & 2 & 6 \end{pmatrix}$$

is a matrix with three rows and four columns. The *element* or *entry* in the i^{th} row and the j^{th} column of the matrix is referred to as the $(i, j)^{th}$ element of the matrix. For example, the entry in the 2^{nd} row and 3^{rd} column of A is 3. If we let a_{ij} denote the element in the i^{th} row and the j^{th} column of the matrix A, then

$$\begin{aligned} &a_{11} = 2, \quad &a_{12} = 7, \quad &a_{13} = 11, \quad &a_{14} = 5; \\ &a_{21} = 4, \quad &a_{22} = 9, \quad &a_{23} = 3, \quad &a_{24} = 1; \\ &a_{31} = 13, \quad &a_{32} = 8, \quad &a_{33} = 2, \quad &a_{34} = 6. \end{aligned}$$

The subscripts on the element a_{ij} denote its position in the i^{th} row and the j^{th} column of the matrix A.

The *dimension* of a matrix is the number of rows and columns in the matrix. By convention, the number of rows is listed first, and then the number of columns. Thus we refer to the matrix A above as being a 3×4, or a "3 by 4", matrix.

A *vector* is a special kind of matrix, having either a single row or a single column. For example,

$$V = \begin{pmatrix} 2 \\ 4 \\ 9 \\ 7 \\ 3 \end{pmatrix}$$

is a 5×1 (column) vector. Since the dimension of a vector corresponds to the number of elements in the vector, the dimension of a vector is often loosely referred to as its *length*.[1]

Finally, a *scalar* is a single element (e.g., a single number), and hence can be treated either as a single element vector or as a 1×1 matrix.

[1] In matrix algebra, vectors have a geometric meaning, denoting the coordinates of a point in Euclidean space. The geometric concept of the "length" (or magnitude) of a vector in Euclidean space has a very precise definition and technical meaning that is quite different from our informal use of the term here.

Transpose

The *transpose* is a function that interchanges the rows and columns of a matrix. That is, the first row becomes the first column, the second row becomes the second column, and so on. By convention, the transpose of a matrix A is denoted A' (or "A prime"). (Note that in some texts, a superscript T, instead of a prime, is used to denote the transpose of a matrix, e.g., A^T.)

For example, consider the 3×4 matrix A,

$$A = \begin{pmatrix} 2 & 7 & 11 & 5 \\ 4 & 9 & 3 & 1 \\ 13 & 8 & 2 & 6 \end{pmatrix}.$$

The transpose of A,

$$A' = \begin{pmatrix} 2 & 4 & 13 \\ 7 & 9 & 8 \\ 11 & 3 & 2 \\ 5 & 1 & 6 \end{pmatrix},$$

is the 4×3 matrix with rows and columns interchanged. Similarly, since a vector is a matrix with either a single row or column, if

$$V = \begin{pmatrix} 2 \\ 4 \\ 9 \\ 7 \\ 3 \end{pmatrix}, \quad \text{then} \quad V' = \begin{pmatrix} 2 & 4 & 9 & 7 & 3 \end{pmatrix}.$$

Examples of vectors and matrices that play key roles in the analysis of longitudinal data are the *response vector*, often denoted Y, and the *covariate matrix*, often denoted X. For example, consider the following data from a subject participating in a longitudinal clinical trial. In this trial, the subject was assigned to the placebo group (Group = 0 if assigned to placebo, Group = 1 if assigned to active treatment) and four repeated measures of blood lead levels were obtained at baseline (or week 0), week 1, week 4, and week 6:

Blood Lead	Treatment Group	Week
30.8	0	0
26.9	0	1
25.8	0	4
23.8	0	6

If we let Y denote the vector of repeated measurements of the response variable, then

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix} = \begin{pmatrix} 30.8 \\ 26.9 \\ 25.8 \\ 23.8 \end{pmatrix}.$$

Similarly we can let X denote a matrix of covariates associated with the vector of repeated measurements, with

$$X = \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \\ X_{41} & X_{42} & X_{43} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 4 \\ 1 & 0 & 6 \end{pmatrix}.$$

The first column of X contains only 1's, while the second column of X contains a variable denoting the treatment group assignment and the third column contains the times of the repeated measurements.

Square and Symmetric Matrices

A matrix is said to be *square* if it has the same number of rows and columns. A square matrix is *symmetric* if it equals its transpose. For example,

$$S = \begin{pmatrix} 2 & 3 & 7 & 11 \\ 3 & 9 & 1 & 2 \\ 7 & 1 & 5 & 8 \\ 11 & 2 & 8 & 4 \end{pmatrix}$$

is a symmetric matrix since it equals its transpose

$$S' = \begin{pmatrix} 2 & 3 & 7 & 11 \\ 3 & 9 & 1 & 2 \\ 7 & 1 & 5 & 8 \\ 11 & 2 & 8 & 4 \end{pmatrix}.$$

Examples of symmetric matrices that play an important role in the analysis of longitudinal data are the covariance and correlation matrices for the repeated measures on the same individuals.

Finally, a *diagonal* matrix is a special case of a symmetric square matrix that has non-zero elements only in the main diagonal positions, and zeros elsewhere. The main diagonal elements are those in the same row and column, from the upper left to the lower right corners of the matrix. For example,

$$D = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 9 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

is a diagonal matrix. The diagonal matrix having all ones along the main diagonal is known as the *identity* matrix and is often denoted by I or I_n, where the subscript n denotes the dimension of the identity matrix. Thus

$$I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Arithmetic Operations

Addition and subtraction of matrices are defined only for matrices of the same dimension. That is, the matrices must share the same number of rows and the same number of columns. The sum of two matrices is obtained by adding their corresponding elements. For example,

$$\begin{aligned} \begin{pmatrix} 2 & 7 & 11 \\ 4 & 9 & 3 \\ 13 & 8 & 2 \end{pmatrix} + \begin{pmatrix} 3 & 2 & 14 \\ 7 & 8 & 4 \\ 6 & 5 & 9 \end{pmatrix} &= \begin{pmatrix} 2+3 & 7+2 & 11+14 \\ 4+7 & 9+8 & 3+4 \\ 13+6 & 8+5 & 2+9 \end{pmatrix} \\ &= \begin{pmatrix} 5 & 9 & 25 \\ 11 & 17 & 7 \\ 19 & 13 & 11 \end{pmatrix}. \end{aligned}$$

Subtraction of matrices is defined in a similar way. For example,

$$\begin{aligned} \begin{pmatrix} 2 & 7 & 11 \\ 4 & 9 & 3 \\ 13 & 8 & 2 \end{pmatrix} - \begin{pmatrix} 3 & 2 & 14 \\ 7 & 8 & 4 \\ 6 & 5 & 9 \end{pmatrix} &= \begin{pmatrix} 2-3 & 7-2 & 11-14 \\ 4-7 & 9-8 & 3-4 \\ 13-6 & 8-5 & 2-9 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 5 & -3 \\ -3 & 1 & -1 \\ 7 & 3 & -7 \end{pmatrix}. \end{aligned}$$

Scalar Multiplication of a Matrix

A scalar is a single number, as opposed to a vector or matrix of numbers. The scalar multiple of a matrix is formed by multiplying each element of the matrix by the scalar. For example, if

$$A = \begin{pmatrix} 2 & 7 & 11 & 5 \\ 4 & 9 & 3 & 1 \\ 13 & 8 & 2 & 6 \end{pmatrix}, \quad \text{then} \quad 2A = \begin{pmatrix} 4 & 14 & 22 & 10 \\ 8 & 18 & 6 & 2 \\ 26 & 16 & 4 & 12 \end{pmatrix}.$$

Multiplication of Matrices

The multiplication of two matrices is somewhat more involved. The multiplication of two matrices A and B, denoted AB, is defined only if the number of columns of A is equal to the number of rows of B. For example, if A is a $p \times q$ matrix and B is a $q \times r$ matrix, then the product of the two matrices AB is a $p \times r$ matrix. Letting C be the product of A and B,

$$C = AB,$$

the $(i, j)^{th}$ element of C is the sum of the products of the corresponding elements in the i^{th} row of A and the j^{th} column of B. Specifically, if c_{ij} is the element in the i^{th} row and the j^{th} column of the matrix $C = AB$, then

$$c_{ij} = \sum_{k=1}^{q} a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{iq} b_{qj}, \quad i = 1, ..., p; \quad j = 1, ..., r;$$

where q is the number of columns in A or the number of rows in B. Matrix multiplication is best understood by considering a simple example. Suppose

$$A = \begin{pmatrix} 2 & 7 & 11 \\ 4 & 9 & 3 \\ 13 & 8 & 2 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 2 \\ 3 & 1 \\ 2 & 4 \end{pmatrix}$$

then

$$\begin{aligned} AB &= \begin{pmatrix} (2 \times 1) + (7 \times 3) + (11 \times 2) & (2 \times 2) + (7 \times 1) + (11 \times 4) \\ (4 \times 1) + (9 \times 3) + (3 \times 2) & (4 \times 2) + (9 \times 1) + (3 \times 4) \\ (13 \times 1) + (8 \times 3) + (2 \times 2) & (13 \times 2) + (8 \times 1) + (2 \times 4) \end{pmatrix} \\ &= \begin{pmatrix} 45 & 55 \\ 37 & 29 \\ 41 & 42 \end{pmatrix}. \end{aligned}$$

Note that the order of multiplication is very important. For example, if A and B are both square matrices of the same dimension, then AB is usually not equal to BA.

The multiplication of a vector by a matrix is a particularly important operation that plays a key role in longitudinal analysis. Let B be a $p \times 1$ vector and X be a $n \times p$ matrix. Then the product,

$$C = XB,$$

is a $n \times 1$ vector with

$$c_i = \sum_{k=1}^{p} x_{ik} b_k, \quad i = 1, ..., n;$$

where x_{ij} is the element in the i^{th} row and the j^{th} column of the matrix X and b_j is the element in the j^{th} row of the vector B. That is,

$$c_1 = x_{11}b_1 + x_{12}b_2 + \cdots + x_{1p}b_p,$$

$$c_2 = x_{21}b_1 + x_{22}b_2 + \cdots + x_{2p}b_p,$$

$$c_3 = x_{31}b_1 + x_{32}b_2 + \cdots + x_{3p}b_p,$$

and so on.

Let us return to the example introduced earlier, with repeated measures of blood lead levels obtained on four occasions. Letting Y denote the vector of repeated measurements of the response variable,

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix},$$

and X a matrix of covariates associated with the vector of repeated measurements,

$$X = \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \\ X_{41} & X_{42} & X_{43} \end{pmatrix},$$

a linear regression model for the mean of each response can be expressed in vector and matrix notation as

$$E(Y) = X\,\beta,$$

where $E(Y)$ denotes the expected value or mean of Y (see *Properties of Expectations and Variances* in Appendix B) and β is a 3×1 vector of regression coefficients,

$$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}.$$

Specifically, the product

$$E(Y) = X\,\beta,$$

is a 4×1 vector

$$\begin{pmatrix} E(Y_1) \\ E(Y_2) \\ E(Y_3) \\ E(Y_4) \end{pmatrix} = \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \\ X_{41} & X_{42} & X_{43} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} \beta_1 X_{11} + \beta_2 X_{12} + \beta_3 X_{13} \\ \beta_1 X_{21} + \beta_2 X_{22} + \beta_3 X_{23} \\ \beta_1 X_{31} + \beta_2 X_{32} + \beta_3 X_{33} \\ \beta_1 X_{41} + \beta_2 X_{42} + \beta_3 X_{43} \end{pmatrix}.$$

That is,

$$E(Y) = X\,\beta,$$

is simply a shorthand representation for the following series of linear regression equations

$$E(Y_1) = \beta_1 X_{11} + \beta_2 X_{12} + \beta_3 X_{13},$$

$$E(Y_2) = \beta_1 X_{21} + \beta_2 X_{22} + \beta_3 X_{23},$$

$$E(Y_3) = \beta_1 X_{31} + \beta_2 X_{32} + \beta_3 X_{33},$$

$$E(Y_4) = \beta_1 X_{41} + \beta_2 X_{42} + \beta_3 X_{43}.$$

Inverse

The *inverse* of a square matrix A, denoted A^{-1}, is defined as a square matrix whose elements are such that

$$AA^{-1} = A^{-1}A = I,$$

where I is the identity matrix, a diagonal matrix having all ones along the main diagonal. That is, the product of A by its inverse is equal to the identity matrix. The inverse of a square matrix does not always exist. The inverse of a matrix only exists if the matrix is *non-singular*.

In matrix algebra the inverse plays the role of the reciprocal, and thus multiplication by an inverse, A^{-1}, can loosely be thought of as "division" by the matrix A. Methods for calculating the inverse of a matrix will not be discussed here. In practice, the inverse of a matrix is usually obtained with the aid of a computer.

Finally, the *determinant* of a square matrix is a unique scalar (or single number) function of its elements and is denoted by $|A|$. For example, if A is a 2×2 matrix,

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

then the determinant of A is the following function of its elements

$$|A| = a_{11}\, a_{22} - a_{12}\, a_{21}.$$

The corresponding expression for the determinant of a 3×3 square matrix, and of matrices of higher dimensions, is more involved and the details are not important. A useful property of the determinant is that it provides a test of whether the inverse of a matrix exists. In particular, if $|A| \neq 0$, then the inverse of A exists; if $|A| = 0$, then the matrix is said to be *singular* and the inverse of A does not exist.

The determinant also plays a role in the definition of the multivariate normal distribution (see Section 3.2). The multivariate normal density includes a term involving the determinant of the covariance matrix. The determinant of the covariance matrix is often referred to as the *generalized variance* and characterizes the salient features of the variation expressed by the covariance matrix in a single number summary.