Stevens, J. P. (2009). Applied multivariate statistics for the social sciences (5th Edition). New York: Routledge.

r" in some studies.
effect" that I men-

riables. If a univari-
e, then how many
d on overall alpha?

ependent variables.
variable, then how
bound on overall

gnificance: A sur-
ide toward educa-
ventually respond.
. A higher positive
ble responses. The
group sizes, along

l, which is signifi-

red on 4 variables.
d yet be an outlier

ject (checking for
at the .001 level?
control on?
f interest to occur.
interest. What are

2

Matrix Algebra

2.1 Introduction

A matrix is simply a rectangular array of elements. The following are examples of matrices:

$$\underset{2 \times 4}{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 5 & 6 & 9 \end{bmatrix}} \qquad \underset{4 \times 3}{\begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 5 \\ 5 & 6 & 8 \\ 1 & 4 & 10 \end{bmatrix}} \qquad \underset{2 \times 2}{\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}}$$

The numbers underneath each matrix are the dimensions of the matrix, and indicate the size of the matrix. The first number is the number of rows and the second number the number of columns. Thus, the first matrix is a 2×4 since it has 2 rows and 4 columns.

A familiar matrix in educational research is the score matrix. For example, suppose we had measured six subjects on three variables. We could represent all the scores as a matrix:

$$\begin{array}{cc} & \textit{Variables} \\ & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \text{Subjects} \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} & \begin{bmatrix} 10 & 4 & 18 \\ 12 & 6 & 21 \\ 13 & 2 & 20 \\ 16 & 8 & 16 \\ 12 & 3 & 14 \\ 15 & 9 & 13 \end{bmatrix} \end{array}$$

$$\begin{array}{cc} & \textit{Variables} \\ & \begin{array}{ccccc} 1 & 2 & 3 & & p \end{array} \\ \text{Subjects} \begin{array}{c} 1 \\ 2 \\ \\ N \end{array} & \begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1p} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & & \vdots \\ x_{N1} & x_{N2} & x_{N3} & \cdots & x_{Np} \end{bmatrix} \end{array}$$

The first subscript indicates the row and the second subscript the column. Thus, x_{12} represents the score of subject 1 on variable 2 and x_{2p} represents the score of subject 2 on variable p.

The *transpose* **A′** of a matrix **A** is simply the matrix obtained by interchanging rows and columns.

Example 2.1

$$\mathbf{A} = \begin{bmatrix} 2 & 3 & 6 \\ 5 & 4 & 8 \end{bmatrix} \Rightarrow \mathbf{A}' = \begin{bmatrix} 2 & 5 \\ 3 & 4 \\ 6 & 8 \end{bmatrix}$$

The first row of **A** has become the first column of **A′** and the second row of **A** has become the second column of **A′**.

$$\mathbf{B} = \begin{bmatrix} 3 & 4 & 2 \\ 5 & 6 & 5 \\ 1 & 3 & 8 \end{bmatrix} \rightarrow \mathbf{B}' = \begin{bmatrix} 3 & 5 & 1 \\ 4 & 6 & 3 \\ 2 & 5 & 8 \end{bmatrix}$$

In general, if a matrix **A** has dimensions $r \times s$, then the dimensions of the transpose are $s \times r$.

A matrix with a single row is called a row vector, and a matrix with a single column is called a column vector. Vectors are always indicated by small letters and a row vector by a transpose, for example, **x′**, **y′**, and so on. Throughout this text a matrix or vector is denoted by a boldface letter.

Example 2.2

$$\mathbf{x}' = (1,2,3) \qquad \mathbf{y} = \begin{bmatrix} 4 \\ 6 \\ 8 \\ 7 \end{bmatrix} 4 \times 1 \text{ column vector}$$

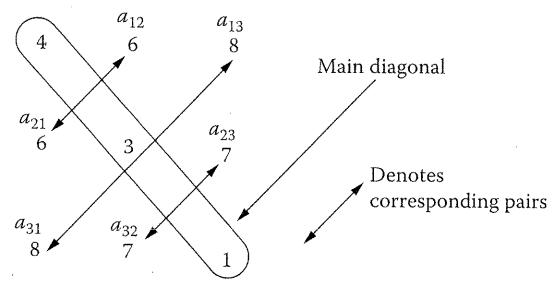
1×3 row vector

A row vector that is of particular interest to us later is the vector of means for a group of subjects on several variables. For example, suppose we have measured 100 subjects on the

California Psychological Inventory and have obtained their average scores on five of the subscales. We could represent their five means as a column vector, and the transpose of this column vector is a row vector **x**′.

$$\mathbf{x} = \begin{bmatrix} 24 \\ 31 \\ 22 \\ 27 \\ 30 \end{bmatrix} \rightarrow \quad \mathbf{x}' = (24, 31, 22, 27, 30)$$

The elements on the diagonal running from upper left to lower right are said to be on the main diagonal of a matrix. A matrix **A** is said to be *symmetric* if the elements below the main diagonal are a mirror reflection of the corresponding elements above the main diagonal. This is saying $a_{12} = a_{21}$, $a_{13} = a_{31}$, and $a_{23} = a_{32}$ for a 3×3 matrix, since these are the corresponding pairs. This is illustrated by:



In general, a matrix **A** is symmetric if $a_{ij} = a_{ji}$, $i \neq j$, i.e., if all corresponding pairs of elements above and below the main diagonal are equal.

An example of a symmetric matrix that is frequently encountered in statistical work is that of a correlation matrix. For example, here is the matrix of intercorrelations for four subtests of the Differential Aptitude Test for boys:

$$\begin{array}{ll} & \begin{array}{cccc} \quad VR & \quad NA & \quad Cler. & \quad Mech. \end{array} \\ \begin{array}{l} \text{Verbal Reas.} \\ \text{Numerical Abil.} \\ \text{Clerical Speed} \\ \text{Mechan. Reas.} \end{array} & \begin{bmatrix} 1.00 & .70 & .19 & .55 \\ .70 & 1.00 & .36 & .50 \\ .19 & .36 & 1.00 & .16 \\ .55 & .50 & .16 & 1.00 \end{bmatrix} \end{array}$$

2.2 Addition, Subtraction, and Multiplication of a Matrix by a Scalar

Two matrices **A** and **B** are added by adding corresponding elements.

Example 2.3

$$\mathbf{A}=\begin{bmatrix}2 & 3\\ 3 & 4\end{bmatrix}\quad \mathbf{B}=\begin{bmatrix}6 & 2\\ 2 & 5\end{bmatrix}$$

$$\mathbf{A}+\mathbf{B}=\begin{bmatrix}2+6 & 3+2\\ 3+2 & 4+5\end{bmatrix}=\begin{bmatrix}8 & 5\\ 5 & 9\end{bmatrix}$$

Notice the elements in the (1, 1) positions, that is, 2 and 6, have been added, and so on.

Only matrices of the same dimensions can be added. Thus addition would not be defined for these matrices:

$$\begin{bmatrix}2 & 3 & 1\\ 1 & 4 & 6\end{bmatrix}+\begin{bmatrix}1 & 4\\ 5 & 6\end{bmatrix}\text{ not defined}$$

Two matrices of the same dimensions are subtracted by subtracting corresponding elements.

$$\underset{\mathbf{A}}{\begin{bmatrix}2 & 1 & 5\\ 3 & 2 & 6\end{bmatrix}}-\underset{\mathbf{B}}{\begin{bmatrix}1 & 4 & 2\\ 1 & 2 & 5\end{bmatrix}}=\underset{\mathbf{A}-\mathbf{B}}{\begin{bmatrix}1 & -3 & 3\\ 2 & 0 & 1\end{bmatrix}}$$

Multiplication of a matrix or a vector by a scalar (number) is accomplished by multiplying each element of the matrix or vector by the scalar.

Example 2.4

$$2(3,1,4)=(6,2,8)\qquad 1/3\begin{bmatrix}4\\ 3\end{bmatrix}=\begin{bmatrix}4/3\\ 1\end{bmatrix}$$

$$4\begin{bmatrix}2 & 1\\ 1 & 5\end{bmatrix}=\begin{bmatrix}8 & 4\\ 4 & 20\end{bmatrix}$$

2.2.1 Multiplication of Matrices

There is a restriction as to when two matrices can be multiplied. Consider the product **AB**. Then *the number of columns in* **A** *must equal the number of rows in* **B**. For example, if **A** is 2×3, then **B** must have 3 rows, although **B** could have any number of columns. If two matrices can be multiplied they are said to be *conformable*. The dimensions of the product matrix, call it **C**, are simply the number of rows of **A** by number of columns of **B**. In the above example, if **B** were 3×4, then **C** would be a 2×4 matrix. In general then, if **A** is an $r \times s$ matrix and **B** is an $s \times t$ matrix, then the dimensions of the product **AB** are $r \times t$.

Example 2.5

$$\underset{2\times 3}{\overset{\mathbf{A}}{\begin{bmatrix} 2 & 1 & 3 \\ 4 & 5 & 6 \end{bmatrix}}} \quad \underset{3\times 2}{\overset{\mathbf{B}}{\begin{bmatrix} 1 & 0 \\ 2 & 4 \\ -1 & 5 \end{bmatrix}}} = \underset{2\times 2}{\overset{\mathbf{C}}{\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}}}$$

Notice first that **A** and **B** can be multiplied because the number of columns in **A** is 3, which is equal to the number of rows in **B**. The product matrix **C** is a 2×2, that is, the outer dimensions of **A** and **B**. To obtain the element c_{11} (in the first row and first column), we multiply corresponding elements of the first row of **A** by the elements of the first column of **B**. Then, we simply add the sum of these products. To obtain c_{12} we take the sum of products of the corresponding elements of the first row of **A** by the second column of **B**. This procedure is presented next for all four elements of **C**:

Element

$$c_{11} \quad (2,1,3)\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = 2(1) + 1(2) + 3(-1) = 1$$

$$c_{12} \quad (2,1,3)\begin{pmatrix} 0 \\ 4 \\ 5 \end{pmatrix} = 2(0) + 1(4) + 3(5) = 19$$

$$c_{21} \quad (4,5,6)\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = 4(1) + 5(2) + 6(-1) = 8$$

$$c_{22} \quad (4,5,6)\begin{pmatrix} 0 \\ 4 \\ 5 \end{pmatrix} = 4(0) + 5(4) + 6(5) = 50$$

Therefore, the product matrix **C** is:

Example 2.6

$$\overset{\mathbf{A}}{\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}} \overset{\mathbf{B}}{\begin{bmatrix} 3 & 5 \\ 5 & 6 \end{bmatrix}} = \begin{bmatrix} 2\cdot 3+1\cdot 5 & 2\cdot 5+1\cdot 6 \\ 1\cdot 3+4\cdot 5 & 1\cdot 5+4.6 \end{bmatrix} = \overset{\mathbf{AB}}{\begin{bmatrix} 11 & 16 \\ 23 & 29 \end{bmatrix}}$$

$$\overset{\mathbf{B}}{\begin{bmatrix} 3 & 5 \\ 5 & 6 \end{bmatrix}} \overset{\mathbf{A}}{\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}} = \begin{bmatrix} 3\cdot 2+5\cdot 1 & 3\cdot 1+5\cdot 4 \\ 5\cdot 2+6\cdot 1 & 5\cdot 1+6\cdot 4 \end{bmatrix} = \overset{\mathbf{BA}}{\begin{bmatrix} 11 & 23 \\ 16 & 29 \end{bmatrix}}$$

Notice that $\mathbf{AB} \neq \mathbf{BA}$; that is, the *order* in which matrices are multiplied makes a difference. The mathematical statement of this is to say that multiplication of matrices is not commutative. Multiplying matrices in two different orders (assuming they are conformable both ways) in general yields different results.

Example 2.7

$$\underset{(3\times 3)}{\overset{\mathbf{A}}{\begin{bmatrix} 3 & 1 & 2 \\ 1 & 4 & 5 \\ 2 & 5 & 2 \end{bmatrix}}} \underset{(3\times 1)}{\overset{\mathbf{x}}{\begin{bmatrix} 2 \\ 6 \\ 3 \end{bmatrix}}} = \underset{(3\times 1)}{\overset{\mathbf{Ax}}{\begin{bmatrix} 18 \\ 41 \\ 40 \end{bmatrix}}}$$

Notice that multiplying a matrix on the right by a column vector takes the matrix into a column vector.

$$(2,5)\begin{bmatrix} 3 & 1 \\ 1 & 4 \end{bmatrix} = (11,22)$$

Multiplying a matrix on the left by a row vector results in a row vector. If we are multiplying more than two matrices, then we may *group at will*. The mathematical statement of this is that multiplication of matrices is associative. Thus, if we are considering the matrix product **ABC**, we get the same result if we multiply **A** and **B** first (and then the result of that by **C**) as if we multiply **B** and **C** first (and then the result of that by **A**), i.e.,

$$\mathbf{A\,B\,C} = (\mathbf{A\,B})\,\mathbf{C} = \mathbf{A}\,(\mathbf{B\,C})$$

A matrix product that is of particular interest to us in Chapter 4 is of the following form:

$$\underset{1\times p}{\mathbf{x}'} \quad \underset{p\times p}{\mathbf{S}} \quad \underset{p\times 1}{\mathbf{x}}$$

Note that this product yields a number, i.e., the product matrix is 1×1 or a number. The multivariate test statistic for two groups is of this form (except for a scalar constant in front).

Example 2.8

$$(4,2)\begin{bmatrix}10 & 3\\ 3 & 4\end{bmatrix}\begin{bmatrix}4\\ 2\end{bmatrix} = (46,20)\begin{bmatrix}4\\ 2\end{bmatrix} = 184 + 40 = 224$$

2.3 Obtaining the Matrix of Variances and Covariances

Now, we show how various matrix operations introduced thus far can be used to obtain a very important quantity in statistical work, i.e., the matrix of variances and covariances for a set of variables. Consider the following set of data

x_1	x_2
1	1
3	4
2	7
$\bar{x}_1 = 2$	$\bar{x}_2 = 4$

First, we form the matrix X_d of deviation scores, that is, how much each score deviates from the mean on that variable:

$$\mathbf{X}_d = \overset{\mathbf{X}}{\begin{bmatrix}1 & 1\\ 3 & 4\\ 2 & 7\end{bmatrix}} - \overset{\bar{\mathbf{X}}}{\begin{bmatrix}2 & 4\\ 2 & 4\\ 2 & 4\end{bmatrix}} = \begin{bmatrix}-1 & -3\\ 1 & 0\\ 0 & 3\end{bmatrix}$$

Next we take the transpose of $\mathbf{X}_d$:

$$\mathbf{X}_d' = \begin{bmatrix}-1 & 1 & 0\\ -3 & 0 & 3\end{bmatrix}$$

Now we can obtain the so-called matrix of sums of squares and cross products **(SSCP)** as the product of $\mathbf{X}_d'$ and $\mathbf{X}_d$:

Deviation scores for x_1

Deviation scores for x_2

$$\mathbf{SSCP} = \overset{\mathbf{X}_d'}{\begin{bmatrix}-1 & 1 & 0\\ -3 & 0 & 3\end{bmatrix}} \overset{\mathbf{X}_d}{\begin{bmatrix}-1 & -3\\ 1 & 0\\ 0 & 3\end{bmatrix}} = \begin{bmatrix}ss_1 & ss_{12}\\ ss_{21} & ss_2\end{bmatrix}$$

Notice that these deviation sums of squares are the numerators of the variances for the variables, because the variance for a variable is

$$s^2 = \sum_i (x_{ii} - \bar{x})^2/(n-1).$$

The sum of deviation cross products (ss_{12}) for the two variables is

$$ss_{12} = ss_{21} = (-1)(-3) + 1(0) + (0)(3) = 3$$

This is just the numerator for the covariance for the two variables, because the definitional formula for covariance is given by:

$$s_{12} = \frac{\sum_{i=1}^{n} (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2)}{n-1},$$

where $(x_{i1} - \bar{x}_1)$ is the deviation score for the ith subject on x_1 and $(x_{i2} - \bar{x}_2)$ is the deviation score for the ith subject on x_2.

Finally, the matrix of variances and covariances **S** is obtained from **SSCP** matrix by multiplying by a constant, namely, $1/(n-1)$:

$$\mathbf{S} = \frac{\mathbf{SSCP}}{n-1}$$

Variance for variable 1

$$\mathbf{S} = \frac{1}{2}\begin{bmatrix} 2 & 3 \\ 3 & 18 \end{bmatrix} = \begin{bmatrix} 1 & 1.5 \\ 1.5 & 9 \end{bmatrix}$$

Variance for variable 2

Covariance

Thus, in obtaining **S** we have:

1. Represented the scores on several variables as a matrix.
2. Illustrated subtraction of matrices—to get $\mathbf{X}_d$.
3. Illustrated the transpose of a matrix—to get $\mathbf{X}'_d$.
4. Illustrated multiplication of matrices, i.e., $\mathbf{X}'_d\,\mathbf{X}_d$, to get **SSCP**.
5. Illustrated multiplication of a matrix by a scalar, i.e., by $1/(n-1)$, to finally obtain **S**.

2.4 Determinant of a Matrix

The determinant of a matrix **A**, denoted by $|\mathbf{A}|$, is a unique number associated with each *square* matrix. There are two interrelated reasons that consideration of determinants is quite important for multivariate statistical analysis. First, the determinant of a covariance matrix represents the *generalized* variance for several variables. That is, it characterizes in a single number how much variability is present on a set of variables. Second, because the determinant represents variance for a set of variables, it is intimately involved in several multivariate test statistics. For example, in Chapter 3 on regression analysis, we use a test statistic called Wilks' Λ that involves a ratio of two determinants. Also, in k group, multivariate analysis of variance the following form of Wilks' Λ ($\Lambda = |\mathbf{W}|/|\mathbf{T}|$) is the most widely used test statistic for determining whether several groups differ on a set of variables. The **W** and **T** matrices are multivariate generalizations of SS_w (sum of squares within) and SS_t (sum of squares total) from univariate ANOVA, and are defined and described in detail in Chapters 4 and 5.

There is a formal definition for finding the determinant of a matrix, but it is complicated and we do not present it. There are other ways of finding the determinant, and a convenient method for smaller matrices (4 × 4 or less) is the method of cofactors. For a 2 × 2 matrix, the determinant could be evaluated by the method of cofactors; however, it is evaluated more quickly as simply the difference in the products of the diagonal elements.

Example 2.9

$$\mathbf{A} = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix} \Rightarrow |\mathbf{A}| = 4 \cdot (2) - 1 \cdot (1) = 7$$

In general, for a 2 × 2 matrix $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then $|\mathbf{A}| = ad - bc$.

To evaluate the determinant of a 3 × 3 matrix we need the method of cofactors and the following definition.

Definition: The *minor* of an element a_{ij} is the determinant of the matrix formed by deleting the ith row and the jth column.

Example 2.10

Consider the following matrix

a_{12} a_{13}

The minor of $a_{12} = 2$ is the determinant of the matrix $\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}$ obtained by deleting the first row and the second column. Therefore, the minor of 2 is $\begin{vmatrix} 2 & 1 \\ 3 & 4 \end{vmatrix} = 8 - 3 = 5$.

The minor of $a_{13} = 3$ is the determinant of the matrix $\begin{bmatrix} 2 & 2 \\ 3 & 1 \end{bmatrix}$ obtained by deleting the first row and the third column. Thus, the minor of 3 is $\begin{vmatrix} 2 & 2 \\ 3 & 1 \end{vmatrix} = 2 - 6 = -4$.

Definition: The cofactor of $a_{ij} = (-1)^{i+j} \times$ minor

Thus, the cofactor of an element will differ at most from its minor by sign. We now evaluate $(-1)^{i+j}$ for the first three elements of the **A** matrix given:

$$a_{11} : (-1)^{1+1} = 1$$

$$a_{12} : (-1)^{1+2} = -1$$

$$a_{13} : (-1)^{1+3} = 1$$

Notice that the signs for the elements in the first row alternate, and this pattern continues for all the elements in a 3×3 matrix. Thus, when evaluating the determinant for a 3×3 matrix it will be convenient to write down the pattern of signs and use it, rather than figuring out what $(-1)^{i+j}$ is for each element. That pattern of signs is:

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$$

We denote the matrix of cofactors **C** as follows:

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}$$

Now, *the determinant is obtained by expanding along any row or column of the matrix of cofactors*. Thus, for example, the determinant of **A** would be given by

$$|\mathbf{A}| = a_{11}c_{11} + a_{12}c_{12} + a_{13}c_{13}$$

(expanding along the first row)

or by

$$|\mathbf{A}| = a_{12}c_{12} + a_{22}c_{22} + a_{32}c_{32}$$

(expanding along the second column)

We now find the determinant of **A** by expanding along the first row:

Element	**Minor**	**Cofactor**	**Element × Cofactor**
$a_{11} = 1$	$\begin{vmatrix} 2 & 1 \\ 1 & 4 \end{vmatrix} = 7$	7	7
$a_{12} = 2$	$\begin{vmatrix} 2 & 1 \\ 3 & 4 \end{vmatrix} = 5$	−5	−10
$a_{13} = 3$	$\begin{vmatrix} 2 & 2 \\ 3 & 1 \end{vmatrix} = -4$	−4	−12

Therefore, $|\mathbf{A}| = 7 + (-10) + (-12) = -15$.

For a 4×4 matrix the pattern of signs is given by:

$$\begin{matrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{matrix}$$

and the determinant is again evaluated by expanding along any row or column. However, in this case the minors are determinants of 3×3 matrices, and the procedure becomes quite tedious. Thus, we do not pursue it any further here.

In the example in 2.3 we obtained the following covariance matrix:

$$\mathbf{S} = \begin{bmatrix} 1.0 & 1.5 \\ 1.5 & 9.0 \end{bmatrix}$$

We also indicated at the beginning of this section that the determinant of **S** can be interpreted as the generalized variance for a set of variables.

Now, the generalized variance for the above two variable example is just $|\mathbf{S}| = 1 \times (9) - (1.5 \times 1.5) = 6.75$. Because for this example there is a covariance, the generalized variance is reduced by this. That is, some of the variance in variable 2 is accounted for by variance in variable 1. On the other hand, if the variables were uncorrelated (covariance = 0), then we would expect the generalized variance to be larger (because none of the variance in variable 2 can be accounted for by variance in variable 1), and this is indeed the case:

$$|\mathbf{S}| = \begin{vmatrix} 1 & 0 \\ 0 & 9 \end{vmatrix} = 9$$

variables, each of which has a variance. In addition, each pair of variables has a covariance. Thus, to represent variance in the multivariate case, we must take into account all the variances and covariances. This gives rise to a matrix of these quantities. Consider the simplest case of two dependent variables. The population covariance matrix Σ looks like this:

2.5 Inverse of a Matrix

The inverse of a square matrix **A** is a matrix $\mathbf{A}^{-1}$ that satisfies the following equation:

$$\mathbf{A}\,\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n$$

where $\mathbf{I}_n$ is the identity matrix of order n. The identity matrix is simply a matrix with 1's on the main diagonal and 0's elsewhere.

$$\mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Why is finding inverses important in statistical work? Because we do not literally have division with matrices, *inversion for matrices is the analogue of division for numbers.* This is why finding inverses is so important. An analogy with univariate ANOVA may be helpful here. In univariate ANOVA, recall that the test statistic $F = MS_b/MS_w = MS_b\,(MS_w)^{-1}$, that is, a ratio of between to within variability. The analogue of this test statistic in multivariate analysis of variance is $\mathbf{BW}^{-1}$, where **B** is a matrix that is the multivariate generalization of SS_b (sum of squares between); that is, it is a measure of how differential the effects of treatments have been on the set of dependent variables. In the multivariate case, we also want to "divide" the between-variability by the within-variability, but we don't have division per se. However, multiplying the **B** matrix by $\mathbf{W}^{-1}$ accomplishes this for us, because inversion is the analogue of division. Also, as shown in the next chapter, to obtain the regression coefficients for a multiple regression analysis, it is necessary to find the inverse of a matrix product involving the predictors.

2.5.1 Procedure for Finding the Inverse of a Matrix

1. Replace each element of the matrix **A** by its minor.
2. Form the matrix of cofactors, attaching the appropriate signs from the pattern of signs.
3. Take the transpose of the matrix of cofactors, forming what is called the adjoint.
4. Divide each element of the adjoint by the determinant of **A**.

For symmetric matrices (with which this text deals almost exclusively), taking the transpose is *not* necessary, and hence, when finding the inverse of a symmetric matrix, Step 3 is omitted.

We apply this procedure first to the simplest case, finding the inverse of a 2 × 2 matrix.

Example 2.11

$$\mathbf{D} = \begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix}$$

The minor of 4 is the determinant of the matrix obtained by deleting the first row and the first column. What is left is simply the number 6, and the determinant of a number is that number. Thus we obtain the following matrix of minors:

$$\begin{bmatrix} 6 & 2 \\ 2 & 4 \end{bmatrix}$$

Now the pattern of signs for any 2×2 matrix is

$$\begin{bmatrix} + & - \\ - & + \end{bmatrix}$$

Therefore, the matrix of cofactors is

$$\begin{bmatrix} 6 & -2 \\ -2 & 4 \end{bmatrix}$$

The determinant of $\mathbf{D} = 6(4) - 2(2) = 20$.

Finally then, the inverse of $\mathbf{D}$ is obtained by dividing the matrix of cofactors by the determinant, obtaining

$$\mathbf{D}^{-1} = \begin{bmatrix} \frac{6}{20} & \frac{-2}{20} \\ \frac{-2}{20} & \frac{4}{20} \end{bmatrix}$$

To check that $\mathbf{D}^{-1}$ is indeed the inverse of $\mathbf{D}$, note that

$$\mathbf{D} \quad \mathbf{D}^{-1} \quad \mathbf{D}^{-1} \quad \mathbf{D} \quad \mathbf{I}_2$$

Example 2.12

Let us find the inverse for the 3 × 3 **A** matrix that we found the determinant for in the previous section. Because **A** is a symmetric matrix, it is not necessary to find nine minors, but only six, since the inverse of a symmetric matrix is symmetric. Thus we just find the minors for the elements on and above the main diagonal.

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 2 & 1 \\ 3 & 1 & 4 \end{bmatrix}$$ Recall again that the minor of an element is the determinant of the matrix obtained by deleting the row and column that the element is in.

Element	Matrix	Minor
$a_{11} = 1$	$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$	$2 \times 4 - 1 \times 1 = 7$
$a_{12} = 2$	$\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}$	$2 \times 4 - 1 \times 3 = 5$
$a_{13} = 3$	$\begin{bmatrix} 2 & 2 \\ 3 & 1 \end{bmatrix}$	$2 \times 1 - 2 \times 3 = -4$
$a_{22} = 2$	$\begin{bmatrix} 1 & 3 \\ 3 & 4 \end{bmatrix}$	$1 \times 4 - 3 \times 3 = -5$
$a_{23} = 1$	$\begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}$	$1 \times 1 - 2 \times 3 = -5$
$a_{33} = 4$	$\begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix}$	$1 \times 2 - 2 \times 2 = -2$

Therefore, the matrix of minors for **A** is

$$\begin{bmatrix} 7 & 5 & -4 \\ 5 & -5 & -5 \\ -4 & -5 & -2 \end{bmatrix}$$

Recall that the pattern of signs is

$$\begin{matrix} + & - & + \\ - & + & - \\ + & - & + \end{matrix}$$

Thus, attaching the appropriate sign to each element in the matrix of minors and completing Step 2 of finding the inverse we obtain:

$$\begin{bmatrix} 7 & -5 & -4 \\ -5 & -5 & 5 \\ -4 & 5 & -2 \end{bmatrix}$$

Now the determinant of **A** was found to be –15. Therefore, to complete the final step in finding the inverse we simply divide the preceding matrix by –15, and the inverse of **A** is

$$\mathbf{A}^{-1} = \begin{bmatrix} \frac{-7}{15} & \frac{1}{3} & \frac{4}{15} \\ \frac{1}{3} & \frac{1}{3} & \frac{-1}{3} \\ \frac{4}{15} & \frac{-1}{3} & \frac{2}{15} \end{bmatrix}$$

Again, we can check that this is indeed the inverse by multiplying it by **A** to see if the result is the identity matrix.

Note that for the inverse of a matrix to exist the determinant of the matrix must *not* be equal to 0. This is because in obtaining the inverse each element is divided by the determinant, and division by 0 is not defined. If the determinant of a matrix $\mathbf{B} = 0$, we say **B** is *singular*. If $|\mathbf{B}| \neq 0$, we say **B** is nonsingular, and its inverse does exist.

2.6 SPSS Matrix Procedure

The SPSS matrix procedure was developed at the University of Wisconsin at Madison. It is described in some detail in *SPSS Advanced Statistics 7.5* (1997, pp. 469–512). Various matrix operations can be performed using the procedure, including multiplying matrices, finding the determinant of a matrix, finding the inverse of a matrix, etc. To indicate a matrix you must: (a) enclose the matrix in braces, (b) separate the elements of each row by commas, and (c) separate the rows by semicolons.

The matrix procedure *must* be run from the syntax window. To get to the syntax window, recall that you first click on FILE, then click on NEW, and finally click on SYNTAX. Every matrix program must begin with MATRIX. and end with END MATRIX. The periods are crucial, as each command *must* end with a period. To create a matrix A, use the following

```
MATRIX.
COMPUTE A= { 2, 4, 1; 3,-2, 5} .
COMPUTE B= { 1, 2; 2, 1; 3, 4} .
COMPUTE C= A*B.
COMPUTE E= { 1,-1, 2;-1, 3, 1;2, 1, 10} .
COMPUTE DETE= DET(E).
COMPUTE EINV= INV(E).
PRINT A.
PRINT B.
PRINT C.
PRINT E.
PRINT DETE.
PRINT EINV.
END MATRIX.
```

The A, B, and E matrices are taken from the exercises. Notice in the preceding program that we have all commands, and in SPSS for Windows each command must end with a period. Also, note that each matrix is enclosed in braces, and rows are separated by semicolons. Finally, a separate PRINT command is required to print out each matrix.

To run (or EXECUTE) the above program, click on RUN and then click on ALL from the dropdown menu. When you do, the following output will appear:

Matrix

Run Matrix procedure:

```
A
     2    4    1
     3   -2    5

B
     1    2
     2    1
     3    4

C
    13   12
    14   24

E
     1   -1    2
    -1    3    1
     2    1   10

DETE
     3

EINV
    9.666666667    4.000000000   -2.333333333
    4.000000000    2.000000000   -1.000000000
   -2.333333333   -1.000000000     .666666667

----End Matrix----
```

2.7 SAS IML Procedure

The SAS IML procedure replaced the older PROC MATRIX procedure that was used in version 5 of SAS. SAS IML is documented thoroughly in *SAS/IML: Usage and Reference, Version 6* (1990). There are several features that are very nice about SAS IML, and these are spelled out on pages 2 and 3 of the manual. We mention just three features:

1. SAS/IML is a programming language.
2. SAS/IML software uses operators that apply to entire matrices.
3. SAS/IML software is interactive.

IML is an acronym for Interactive Matrix Language. You can execute a command as soon as you enter it. We do not illustrate this feature, as we wish to compare it with the SPSS Matrix procedure. So we collect the SAS IML commands in a file (or module as they call it) and run it that way.

To indicate a matrix, you (a) enclose the matrix in braces, (b) separate the elements of each row by a blank(s), and (c) separate the columns by commas.

To illustrate use of the SAS IML procedure, we create the same matrices as we did with the SPSS matrix procedure and do the same operations and print out everything. Here is the file and the printout:

```
proc iml;
a= { 2 4 1, 3 -2 5} ;
b= { 1 2, 2 1, 3 4} ;
c= a*b;
e= { 1 -1 2, -1 3 1, 2 1 10} ;
dete= det(e);
einv= inv(e);
print a b c e dete einv;
```

```
 A                   B                      C
 2     4     1       1            2        13           12
 3    -2     5       2            1        14           24
                     3            4
 E                 DETE         EINV
 1    -1     2       3    9.6666667         4    -2.333333
-1     3     1                    4         2           -1
 2     1    10           -2.333333        -1    0.6666667
```

tedious. Finding the determinant is important because the determinant of a covariance matrix represents the generalized variance for a set of variables. Finding the inverse of a matrix is important since inversion for matrices is the analogue of division for numbers. Fortunately, SPSS MATRIX and SAS IML will do various matrix operations, including finding the determinant and inverse.

2.9 Exercises

1. Given:

$$\mathbf{A} = \begin{bmatrix} 2 & 4 & 1 \\ 3 & -2 & 5 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 4 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 1 & 3 & 5 \\ 6 & 2 & 1 \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix} \quad \mathbf{E} = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 3 & 1 \\ 2 & 1 & 10 \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 4 & 6 \\ 5 & 7 \end{bmatrix}$$

$$\mathbf{u}' = (1,3), \mathbf{v} = \begin{bmatrix} 2 \\ 7 \end{bmatrix}$$

Find, where meaningful, each of the following:

(a) $\mathbf{A} + \mathbf{C}$
(b) $\mathbf{A} + \mathbf{B}$
(c) $\mathbf{AB}$
(d) $\mathbf{AC}$
(e) $\mathbf{u}'\mathbf{D}\,\mathbf{u}$
(f) $\mathbf{u}'\mathbf{v}$
(g) $(\mathbf{A} + \mathbf{C})'$
(h) $3\,\mathbf{C}$
(i) $|\mathbf{D}|$
(j) $\mathbf{D}^{-1}$
(k) $|\mathbf{E}|$
(l) $\mathbf{E}^{-1}$
(m) $\mathbf{u}'\mathbf{D}^{-1}\mathbf{u}$
(n) $\mathbf{BA}$ (compare this result with [c])
(o) $\mathbf{X}'\mathbf{X}$

2. In Chapter 3, we are interested in predicting each person's score on a dependent variable y from a linear combination of their scores on several predictors (x_i's). If there were three predictors, then the prediction equations for N subjects would look like this:

$$
\begin{aligned}
y_1 &= e_1 + b_0 + b_1 x_{11} + b_2 x_{12} + b_3 x_{13} \\
y_2 &= e_2 + b_0 + b_1 x_{21} + b_2 x_{22} + b_3 x_{23} \\
y_3 &= e_3 + b_0 + b_1 x_{31} + b_2 x_{32} + b_3 x_{33} \\
&\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
y_N &= e_N + b_0 + b_1 x_{N1} + b_2 x_{N2} + b_3 x_{N3}
\end{aligned}
$$

Note: The e_i's are the portion of y not predicted by the x's, and the b's are the regression coefficient. Express this set of prediction equations as a single matrix equation. Hint: The right hand portion of the equation will be of the form:

vector + matrix times vector

3. Using the approach detailed in section 2.3, find the matrix of variances and covariances for the following data:

x_1	x_2	x_3
4	3	10
5	2	11
8	6	15
9	6	9
10	8	5

4. Consider the following two situations:

(a) $s_1 = 10,\ s_2 = 7,\ r_{12} = .80$

(b) $s_1 = 9,\ s_2 = 6,\ r_{12} = .20$

For which situation is the generalized variance larger? Does this surprise you?

5. Calculate the determinant for

$$\begin{bmatrix} 9 & 2 & 1 \end{bmatrix}$$

6. Using SPSS MATRIX or SAS IML, find the inverse for the following 4 x 4 symmetric matrix:

$$\begin{matrix} 6 & 8 & 7 & 6 \\ 8 & 9 & 2 & 3 \\ 7 & 2 & 5 & 2 \\ 6 & 3 & 2 & 1 \end{matrix}$$

7. Run the following SPSS MATRIX program and show that the output yields the matrix, determinant and inverse.

```
MATRIX.
COMPUTE A={6,2,4;2,3,1;4,1,5}.
COMPUTE DETA=DET(A).
COMPUTE AINV=INV(A).
PRINT A.
PRINT DETA.
PRINT AINV.
END MATRIX.
```

8. Consider the following two matrices:

$$\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 3 & 6 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Calculate the following products: **AB** and **BA**

What do you get in each case? Do you see now why B is called the identity matrix?