A. MATRIX STRUCTURE AND NOTATION

1) A matrix is a rectangular arrangement of numbers. The matrix is usually denoted by a capital letter.

$$A = \begin{bmatrix} 1 & 3 \\ 7 & 9 \end{bmatrix} \qquad\qquad D = \begin{bmatrix} 4 & 2 & 4 \\ 1 & 6 & 0 \\ 3 & 0 & 5 \\ 2 & 3 & 0 \end{bmatrix}$$

2) The dimensions of a matrix are given by the number of rows and columns in the matrix (i.e. the dimensions are r by c). For the matrices above,

$$A \text{ is } 2 \text{ by } 2$$

$$D \text{ is } 4 \text{ by } 3$$

3) The individual elements of a matrix can be referred to by specifying the row and column in which it occurs. Lower case numbers are used to represent individual elements, and should match the upper case letter used to denote matrix. For example, individual elements from matrices A and D above can be referred to as,

$$a_{11} = 1$$

$$a_{21} = 7$$

$$d_{22} = 6$$

$$d_{12} = 2$$

B. TYPES OF MATRICES

1) Square matrix - the number of rows and columns are equal. Matrix A above is a square matrix (2 by 2), matrix D is not (4 by 3). A symmetric matrix is an important variation of the square matrix. In a symmetric matrix, the value in position "ij" equals the value in position "ji" (where $i \neq j$). For example, if $c_{31} = 5$ then c_{13} is also 5.

2) Scalar - a single number can be thought of as a 1 by 1 matrix and is called a scalar.

3) Vector - a single column or single row of numbers is called a vector. The dimensions of a row vector are (1 by c), where "c" is the number of columns, and the dimensions of a column vector (r by 1), where "r" is the number of rows.

4) Identity matrix - this special square matrix consists of all ones on the main diagonal, or principal diagonal, and zeros in all the off diagonal positions. The following are examples of identity matrices,

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad\qquad F = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The diagonal matrix is a generalization of the identity matrix. A diagonal matrix can have any value on the main diagonal, but also has zeros in the off diagonal positions.

C. MATRIX TRANSPOSE

The transpose of a matrix consists of a new matrix such that the rows of the original matrix become the columns of the transpose matrix. The transpose matrix is denoted with the same letter as the original matrix followed by a prime (e.g. the transpose of X is X).

$$D = \begin{bmatrix} 4 & 2 & 4 \\ 1 & 6 & 0 \\ 3 & 0 & 5 \\ 2 & 3 & 0 \end{bmatrix} \qquad\qquad D' = \begin{bmatrix} 4 & 1 & 3 & 2 \\ 2 & 6 & 0 & 3 \\ 4 & 0 & 5 & 0 \end{bmatrix}$$

D. MATRIX ADDITION AND SUBTRACTION

Matrices to be added or subtracted must be of the same dimensions. Each element of the first matrix, (a) is added (or subtracted) from the corresponding element of the second matrix, (b).

$$A = \begin{bmatrix} 1 & -2 \\ 3 & 4 \\ 9 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 4 \\ 1 & 4 \\ -4 & 4 \end{bmatrix} \quad A+B = \begin{bmatrix} 1+1 & -2+4 \\ 3+1 & 4+4 \\ 9-4 & 0+4 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 4 & 8 \\ 5 & 4 \end{bmatrix}$$

E. MATRIX MULTIPLICATION

Multiplication by a scalar - in this type of multiplication each element of the matrix is simply multiplied, element by element, by the scalar value.

$$A = \begin{bmatrix} 1 & -2 \\ 3 & 4 \\ 9 & 0 \end{bmatrix} \quad B = [7] \qquad A * B = 7 * \begin{bmatrix} 1 & -2 \\ 3 & 4 \\ 9 & 0 \end{bmatrix} = \begin{bmatrix} 7 & -14 \\ 21 & 28 \\ 63 & 0 \end{bmatrix}$$

Element by element multiplication - matrix multiplication is not usually done by matching each i,j^{th} element of one matrix with the corresponding ij^{th} element of the second matrix. This is called elementwise multiplication and it is not the normal mode of matrix multiplication and should not be used unless specifically requested.

The standard method of matrix multiplication requires that the number of columns in the first matrix equal the number of rows in the second matrix. If the first matrix is (r by c) and the second is (r by c), in order to multiply the matrices, c must equal r. The resulting matrix will have the dimensions (r by c).
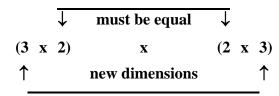
Multiplication is accomplished by summing the cross products of each row of the first matrix and each column of the second matrix.

$$A = \begin{bmatrix} 1 & -2 \\ 3 & 4 \\ 9 & 0 \end{bmatrix} \quad X = \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix}$$

Since A is 3 rows by 2 columns, and X is 2 by 2, then the columns of the first matrix equals the rows of the second matrix, and the matrices may be multiplied.

$$A*X = \begin{bmatrix} 1 & -2 \\ 3 & 4 \\ 9 & 0 \end{bmatrix} * \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} (1*1)+(-2*3) & (1*-2)+(-2*4) \\ (3*1)+(4*3) & (3*-2)+(4*4) \\ (9*1)+(0*3) & (9*-2)+(0*4) \end{bmatrix} = \begin{bmatrix} -5 & -10 \\ 15 & 10 \\ 9 & -18 \end{bmatrix}$$

the new dimensions for the product of A * X are,

$$\downarrow \quad \textbf{must be equal} \quad \downarrow$$

$$\textbf{(3 x 2)} \qquad \textbf{x} \qquad \textbf{(2 x 3)}$$

$$\uparrow \qquad \textbf{new dimensions} \qquad \uparrow$$

Note that though we can multiply A * X, we could not have done the multiplication the other way (i.e. X * A), since the dimensions would not have matched. That is, we could pre-multiply by A, but could not pre-multiply by X.

F. SIMPLE MATRIX INVERSION (2 by 2 matrix only)

Matrices are not "divided", but may be inverted. Instead of "dividing" A by B, one would multiply A by the inverse of B. The inverse of a (2 by 2) matrix is given by,

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad\qquad A^{-1} = \frac{1}{(a \times b) - (b \times c)} \begin{bmatrix} d & \text{-}b \\ \text{-}c & a \end{bmatrix}$$

The scalar value resulting from the calculation "$(a \times d) - (b \times c)$" is called the determinant. The matrix cannot be inverted unless the inverse of the determinant exists (is defined). It will not exist in a case such as the one below since $(1 \div 0)$ is not defined.

$$A = \begin{bmatrix} 1 & 4 \\ 2 & 8 \end{bmatrix} \qquad A^{-1} = \frac{1}{\text{Determinant of A}} = \frac{1}{(1 \times 8) - (2 \times 4)} = \frac{1}{0}$$

This occurs in regression when two variables are linearly related.

An example of the inversion of a 2 * 2 matrix is given below.

$$B = \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix} \quad B^{-1} = \frac{1}{(2 \times 4) - (1 \times 3)} \begin{bmatrix} 4 & -3 \\ -1 & 2 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 4 & -3 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 0.8 & -0.6 \\ -0.2 & 0.4 \end{bmatrix}$$

Note that a matrix times its inverse (i.e. $B \times B^{-1}$) results in an identity matrix. By definition, the inverse of a matrix G is a matrix which when multiplied by G produces an identity matrix, or $G \times G^{-1} = I$.

G. SIMPLE LINEAR REGRESSION

Solving a simple linear regression with matrices requires the same values used for an algebraic solution from summation notation formulas. These are;

$$n, \qquad \sum_{i=1}^{n} X_i, \qquad \sum_{i=1}^{n} Y_i, \qquad \sum_{i=1}^{n} X_i^2, \qquad \sum_{i=1}^{n} Y_i^2, \qquad \sum_{i=1}^{n} X_i Y_i$$

where n is the size of the sample of data. To obtain these values in the matrix form we start with the matrix equivalent of the individual values of X and Y, the raw data matrices.

$$X = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ 1 & X_3 \\ 1 & X_4 \\ 1 & X_5 \\ 1 & X_6 \\ 1 & X_7 \end{bmatrix} \qquad Y = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \\ Y_6 \\ Y_7 \end{bmatrix}$$

The column of ones is necessary, and represents the intercept. Omitting this column would force the regression through the origin. The next step in the calculations is to obtain the X′X, X′Y and Y′Y matrices. These calculations provide the sums of squares and cross products.

$$X'X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ X_1 & X_2 & X_3 & X_4 & X_5 & X_6 & X_7 \end{bmatrix} \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ 1 & X_3 \\ 1 & X_4 \\ 1 & X_5 \\ 1 & X_6 \\ 1 & X_7 \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} X_i & \sum_{i=1}^{n} X_i^2 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ X_1 & X_2 & X_3 & X_4 & X_5 & X_6 & X_7 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \\ Y_6 \\ Y_7 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} Y_i \\ \sum_{i=1}^{n} X_i Y_i \end{bmatrix}$$

$$Y'Y = \begin{bmatrix} Y_1 & Y_2 & Y_3 & Y_4 & Y_5 & Y_6 & Y_7 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \\ Y_6 \\ Y_7 \end{bmatrix}$$

The regression coefficients, b_0 and b_1, are then given by, $B = (X'X)^{-1} X'Y$, where

$$X'X = \frac{1}{\left(n\sum X_i^2\right) - \left(n\sum X_i\right)^2} \begin{bmatrix} \sum X_i^2 & -\sum X_i \\ -\sum X_i & n \end{bmatrix}$$

and since

$$\frac{1}{\text{Determinant } (X'X)^{-1}} = \frac{1}{\left(n\sum X_i^2\right) - \left(n\sum X_i\right)^2} = \frac{1}{nS_{XX}}$$

where S_{XX} is the corrected sum of squares of X.

Then
$$X'X = \begin{bmatrix} \dfrac{\sum X_i^2}{nS_{XX}} & \dfrac{-\sum X_i}{nS_{XX}} \\ \dfrac{-\sum X_i}{nS_{XX}} & \dfrac{n}{nS_{XX}} \end{bmatrix}$$

and the regression coefficients can be obtained by,

$$(X'X)^{-1}X'Y = \begin{bmatrix} \dfrac{\sum X_i^2}{nS_{XX}} & \dfrac{-\sum X_i}{nS_{XX}} \\ \dfrac{-\sum X_i}{nS_{XX}} & \dfrac{n}{nS_{XX}} \end{bmatrix} \times \begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix} = \begin{bmatrix} \dfrac{\sum X_i^2}{nS_{XX}}\sum Y_i & \dfrac{-\sum X_i}{nS_{XX}}\sum X_i Y_i \\ \dfrac{-\sum X_i}{nS_{XX}}\sum Y_i & \dfrac{n}{nS_{XX}}\sum X_i Y_i \end{bmatrix}$$

$$= \begin{bmatrix} \overline{Y} - b_1\overline{X} \\ \dfrac{\sum X_i Y_i - \dfrac{\sum X_i \sum Y_i}{n}}{\sum X_i^2 - \dfrac{\left(\sum X_i\right)^2}{n}} \end{bmatrix} = \begin{bmatrix} \overline{Y} - b_1\overline{X} \\ \dfrac{\sum (X_i - \overline{X})(Y_i - \overline{Y})}{\sum (X_i - \overline{X})^2} \end{bmatrix} = \begin{bmatrix} \overline{Y} - b_1\overline{X} \\ \dfrac{S_{XY}}{S_{XX}} \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

The remaining calculations usually needed to complete the compliment of calculations for the simple linear regression is the sum of squared deviations or error term. The matrix formula is

$$SSE = Y'Y - B'X'Y = \Sigma Y^2 - \begin{bmatrix} b_0 & b_1 \end{bmatrix}\begin{bmatrix} \Sigma Y \\ \Sigma XY \end{bmatrix}$$

$$= \Sigma Y^2 - (b_0 * \Sigma Y + b_1 * \Sigma XY) = \text{UCSSTotal} - \text{UCSSReg}$$

These calculations produce the same algebraic equations for b_0, b_1, and SSE that are given in most statistics texts. The advantage of using the matrix version of the formulas is that the matrix equations given above will work equally well for multiple regression with two or more independent variables.

The ANOVA table calculated with matrix formulas is

	Uncorrected		Corrected	
Source	d.f.	Sum of Squares	d.f.	Sum of Squares
Regression	2	B'X'Y	1	B'X'Y – CF
Error	n–2	Y'Y–B'X'Y	n–2	Y'Y – B'X'Y
Total	n	Y'Y	n–1	Y'Y – CF

where the correction factor is calculated as usual, $CF = \dfrac{(\Sigma Y)^2}{n} = n\overline{Y}^2$.

The value for R^2 is calculated as $\dfrac{\text{SSRegression}}{\text{SSTotal}} = \dfrac{B'X'Y - CF}{Y'Y - CF}$, and is often expressed as a percent. Note that this calculation employs corrected sums of squares for both SSRegression and SSTotal.

The Mean Squares (MS) for the SSRegression and SSError are calculated by dividing the SS (corrected sum of squares) by their d.f. (degrees of freedom). The test of hypotheses for [H_0:= β_1] is then calculated as;

$$F = \frac{\text{MSRegression}}{\text{MSTotal}} = \frac{\left(B'X'Y - CF\right)\Big/ \text{dfReg}}{\left(Y'Y - CF\right)\Big/ \text{dfError}}$$

or

$$t = \frac{\left(b_1 - 0\right)}{S_{b_1}} = \sqrt{F \text{ value}}$$

where S_{b_1} is obtained from the VARIANCE COVARIANCE matrix.

The VARIANCE COVARIANCE matrix is calculated as from the $(X'X)^{-1}$ matrix.

$$\left(X'X\right)^{-1} = \begin{bmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{bmatrix}$$

where the c_{ij} values are called Gaussian multipliers. The VARIANCE-COVARIANCE matrix is then calculated from this matrix by multiplying by the MSError.

$$\text{MSE}\left(X'X\right)^{-1} = \begin{bmatrix} \text{MSE}c_{00} & \text{MSE}c_{01} \\ \text{MSE}c_{10} & \text{MSE}c_{11} \end{bmatrix}$$

The individual values then provide the variances and covariances such that

$\text{MSE}*c_{00} = $ Variance of $b_0 = $ VAR(b_0)

$\text{MSE}*c_{11} = $ Variance of $b_1 = $ VAR(b_1), so $S_{b_1} = \sqrt{\text{MSE}*c_{11}}$

$\text{MSE}*c_{01} = \text{MSE}*c_{10} = $ Covariance of b_0 and $b_{0\backslash 1} = $ COV(b_0,b_1)

It is important to note that the variances and covariances calculated from the $(X'X)^{-1}$ are for the b_i (β_i estimates), not for the X_i values. Also, COV(b_0,b_1) $\neq$ COV(X_0,X_1).

Application of matrix procedures to multiple regression first requires calculation of the X′X, X′Y and Y′Y matrices, where for dependent variable Y and independent variables X_1 and X_2. For a 2 factor multiple regression, these matrices are;

$$X'X = \begin{pmatrix} n & \sum_{i=1}^{n}X_{1i} & \sum_{i=1}^{n}X_{2i} & \sum_{i=1}^{n}X_{3i} \\ \sum_{i=1}^{n}X_{1i} & \sum_{i=1}^{n}X_{1i}^{2} & \sum_{i=1}^{n}X_{1i}X_{2i} & \sum_{i=1}^{n}X_{1i}X_{3i} \\ \sum_{i=1}^{n}X_{2i} & \sum_{i=1}^{n}X_{1i}X_{2i} & \sum_{i=1}^{n}X_{2i}^{2} & \sum_{i=1}^{n}X_{2i}X_{3i} \\ \sum_{i=1}^{n}X_{3i} & \sum_{i=1}^{n}X_{1i}X_{3i} & \sum_{i=1}^{n}X_{2i}X_{3i} & \sum_{i=1}^{n}X_{3i}^{2} \end{pmatrix} \quad X'Y = \begin{pmatrix} \sum_{i=1}^{n}Y_{i} \\ \sum_{i=1}^{n}X_{1i}Y_{i} \\ \sum_{i=1}^{n}X_{2i}Y_{i} \\ \sum_{i=1}^{n}X_{3i}Y_{i} \end{pmatrix} \quad Y'Y = \left(\sum_{i=1}^{n}Y_{i}^{2} \right)$$

As with the simple linear regression, these sums, sums of squares and cross products are required by any method of fitting multiple regression. Once these values are obtained, application of formulas for an algebraic solution is relatively easy for a two-factor model. However, matrix procedures are more easily expanded to more than two independent variables than are summation notation formulas.

The inversion technique we will use is called the sweepout technique, and it requires the application of "row operations". Row operations consist of **(1) multiplying any row by a scalar value, and (2) adding or subtracting any row from any other row**. These are the only operations required to complete the sweepout technique after the matrices have been obtained and augmented.

Obtaining a maximum of information from the technique requires reducing the X′X matrix one column at a time to an identity matrix. However, values of the regression coefficients, error sum of squares and inverse matrix will be correct even of the row operations are not applied in a column by column reduction.

By "sweeping" out each column of the X′X matrix one by one to obtain an identity matrix, the sequentially adjusted sums of squares error can also be obtained. This requires augmenting the X′X matrix with the X′Y matrix and an identity matrix prior to applying the row operations. The complete augmented matrix is given below. The matrix has separate sections that are recognizable as matrices seen earlier. This type of sectioned matrix is called a partitioned matrix.

$$\begin{bmatrix} X'X & X'Y & I \\ X'Y & Y'Y & 0 \end{bmatrix} \xrightarrow{\text{row operations}} \begin{bmatrix} I & B & (X'X)^{-1} \\ 0 & SSE & -B' \end{bmatrix}$$

Sections of the matrix may be left off if less information is required. For example, if only the regression coefficients are needed, then the sweepout technique need be applied only to the matrix, $\begin{bmatrix} X'X & X'Y \end{bmatrix} \xrightarrow{\text{row operations}} \begin{bmatrix} I & B \end{bmatrix}$,

and if only the inverse is required, the only matrix needed is

$$\begin{bmatrix} X'X & I \end{bmatrix} \xrightarrow{\text{row operations}} \begin{bmatrix} I & (X'X)^{-1} \end{bmatrix} .$$

The regression coefficients and sum of squares error can be obtained by sweeping out the matrix,

$$\begin{bmatrix} X'X & X'Y \\ X'Y & Y'Y \end{bmatrix} \xrightarrow{\text{row operations}} \begin{bmatrix} I & B \\ 0 & SSE \end{bmatrix}.$$

If the above matrix is swept out column by column, then it will also provide the sequentially adjusted sums of squares. Only the use of the complete augmented matrix provides the inverted X′X matrix necessary to obtain the variance - covariance matrix, confidence limits and other types of sums of squares.

The technique will be illustrated with an example using data from Snedecor and Cochran (1981; ex. 17.2.1). The example will employ the complete augmented matrix. The original data matrices are;

$$X'X = \begin{bmatrix} 17 & 188.2 & 700 \\ 188.2 & 3602.78 & 8585.1 \\ 700 & 8585.1 & 31712 \end{bmatrix} \quad X'Y = \begin{bmatrix} 1295 \\ 16203.8 \\ 54081 \end{bmatrix} \quad Y'Y = [103075]$$

The augmented matrix to be swept is then,

$$\begin{pmatrix} 17 & 188.2 & 700 & 1295 & 1 & 0 & 0 \\ 188.2 & 3602.78 & 8585.1 & 16203.8 & 0 & 1 & 0 \\ 700 & 8585.1 & 31712 & 54081 & 0 & 0 & 1 \\ 1295 & 16203.8 & 54081 & 103075 & 0 & 0 & 0 \end{pmatrix}$$

The first step in the sweepout technique is to multiply through the first row by the inverse of 17. This will result in a value of 1 in the first row - first column. A multiple of this new first row is then subtracted from each of the other rows (2, 3 and 4). The multiplier should be such that value(i,1)–[value(1,1)*multiplier] = 0 for i ≠ 1.

The multiplier which accomplishes this is simply the value(i,1) since the new value(1,1) is unity (1). Therefore, every value(i,j) will be processed in the same way. The calculations would be,

for row 2: value(2,j) – (value(1,j) * 118.2)

for row 3: value(3,j) – (value(1,j) * 700)

for row 4: value(4,j) – (value(1,j) * 1295)

After applying these transformations we obtain the following matrix,

COLUMN 1 SWEEP

$$\begin{pmatrix} 1 & 11.0706 & 41.1765 & 76.1765 & 0.05882 & 0 & 0 \\ 0 & 1519.30 & 835.688 & 1867.39 & -11.0706 & 1 & 0 \\ 0 & 835.688 & 2888.47 & 757.471 & -41.1765 & 0 & 1 \\ 0 & 1867.39 & 757.471 & 4426.47 & -76.1765 & 0 & 0 \end{pmatrix}$$

At this point the effect of X (the intercept) has been removed from the model. The value replacing Y′Y is 4426.471. This is the corrected sum of squares of Y (i.e. Y was 103075, and has now been corrected for the mean, yielding 4426.47).

The sweepout now proceeds to the second column. A value of 1 is needed in the second column - second row to proceed with the development of the identity matrix. This is obtained by multiplying through the second row by the inverse of the value presently in that position (i.e. 1519.30). Then, the appropriate multiple of the new row 2 is subtracted from each of the other rows. Note that the first column remains unchanged since the value subtracted is always a multiple of zero.

COLUMN 2 SWEEP

$$\begin{pmatrix} 1 & 0 & 35.08709 & 62.5694 & 0.13949 & -0.00729 & 0 \\ 0 & 1 & 0.550050 & 1.22911 & -0.00729 & 0.00066 & 0 \\ 0 & 0 & 2428.800 & -269.686 & -35.0871 & -0.55005 & 1 \\ 0 & 0 & -269.686 & 2131.236 & -62.5694 & 1.22911 & 0 \end{pmatrix}$$

The sweep then proceeds with the third column. Once again a value of 1 is required in row 3, column 3, and all rows other than row 3 will have a multiple of row 3 subtracted from them.

COLUMN 3 SWEEP

$$\begin{pmatrix} 1 & 0 & 0 & 66.4654 & 0.646369 & 0.000660 & -0.014446 \\ 0 & 1 & 0 & 1.29019 & 0.000660 & 0.000783 & 0.000226 \\ 0 & 0 & 1 & -0.11104 & -0.014446 & 0.000226 & 0.000412 \\ 0 & 0 & 0 & 2101.291 & -66.46541 & -1.290191 & -0.11104 \end{pmatrix}$$

Once this swept out matrix has been obtained, most commonly desired calculations follow easily. Some of these results are discussed below.

There are also several checks which can be done on the calculations. As the matrix is swept out, the null matrix (matrix of zeroes in the original augmented matrix) is replaced by the negative values of the regression coefficients if the calculations have been done correctly. As a second check, the product of the original $X'X$ matrix and its inverse should produce an identity matrix. (i.e. $X'X * (X'X)^{-1} = I$)

REGRESSION COEFFICIENTS

The regression coefficients are produced during the sweepout, replacing the $X'Y$ matrix. The model for the analysis above is,

$$\hat{Y} = b_0 + b_1 X_{1i} + b_2 X_{2i}$$

$$\hat{Y} = 66.4654 + 1.2902 X_{1i} + 0.1110 X_{2i}$$

SEQUENTIALLY ADJUSTED SUMS OF SQUARES

As each column is swept out, the sums of squares are "adjusted" for the factor removed. The first sweep adjusts for the intercept (i.e. $1 = n$) on the diagonal of $X'X$, so the reduction in the Y is the correction factor or the adjustment for the mean.

e.g. C.F. = 103075 – 4426.470 = 98648.530

The second sweep adjusts for the second term in the X matrix, usually X, and the reduction in the error term is that sum of squares attributable to X (given that X is already in the model).

e.g. SS(X|X) = 4426.470 – 2131.236 = 2295.234

The third sweep adjusts for X and the reduction in the sum of squares is attributable to X (given that X and X are already in the model).

e.g. SS(X|X X) = 2131.236 – 2101.291 = 29.945

Finally, the remaining sum of squares is the error sum of squares

SSE = 2101.291

Note that since the variables are adjusted sequentially, the sums of squares obtained are dependent on the order in which the variables are entered. That is, if we had entered X first and X second, the sums of squares attributable to these two variables would not be the same as the results obtained above. Only the correction factor would be the same (since it would have been entered first in both models).

Each adjustment of the sum of squares takes one degree of freedom. The residual sum of squares has (nk) degrees of freedom, where n is the number of observations, and k is the number of sweeps, or the number of columns in the X′X matrix. The mean square error is then,

$$MSE = \frac{SSE}{(n\text{-}k)} = \frac{2101.291}{(17\text{ - }3)} = 150.092$$

PARTIAL SUMS OF SQUARES

Since the sequentially adjusted sums of squares are dependent on the order in which the variables are entered, another value of interest is the partial sum of squares or the uniquely attributable sum of squares. This is simply the sum of squares that would be accounted for by each variable if it had been entered into the model in last place. This value could be obtained by reversing the sweep operation, and observing the change in the sum of squares as each variable was swept back into the model.

The only change in sum of squares when a variable is swept back into the model is, bc,

So this calculation will give the partial SS due to variable X without actually doing all the calculations necessary to reverse the sweepout technique. The elements (c) are obtained from the $(X′X)^{-1}$ matrix and are called Gaussian multipliers.

The partial SS due to X above does not change since it was the variable in the last position. The partial SS due to X would be calculated as,

$$SS(X_1|X_0\ X_2)\ =\ \frac{(1.29019)^2}{(0.000783)} = 2125.913$$

VARIANCE COVARIANCE MATRIX

Another major result of the sweepout technique is the inverse of the X′X matrix. Multiplying this matrix by the mean square error (MSE) gives the variance - covariance matrix of the regression coefficients.

e.g. $VarCov\ =\ MSE * (X′X)^{-1} =$

$$= 150.092 \begin{pmatrix} 0.64637 & 0.00066 & -0.01445 \\ 0.00066 & 0.00078 & -0.00023 \\ -0.01445 & -0.00023 & 0.00041 \end{pmatrix} = \begin{pmatrix} 97.0149 & 0.0990 & -2.1683 \\ 0.0990 & 0.1175 & -0.0340 \\ -2.1683 & -0.0340 & 0.0618 \end{pmatrix}$$

so, Var(b_0)=97.0149, Var(b_1)=0.1175, Var(b_2)=0.0618, Var(b_{12})=0.0340, etc.

The variance - covariance matrix can also be used to obtain confidence intervals about estimates of $\hat{Y}$ for particular values of X and X. The most versatile approach is to use matrix algebra in these calculations. The equation is

$$S^2_{\hat{Y}} = \text{MSE} (L'(X'X)^{-1}L)$$

where L is a vector of values for X corresponding to $\hat{Y}$. It may also be a vector of hypothesized X values for which a variance is needed.

For example, if we wish to predict the response ($\hat{Y}$) and its variance when X = 4 and X = 24, first we would calculate the response,

$$\hat{Y}=66.4654+1.2902X_{1i}+0.1110X_{2i} = 66.4654+1.2902(4)+0.1110(24) = 68.9622$$

Using L = [1 4 24], (note that a 1 is included for the intercept) the variance of the estimate is then,

$$S^2_{\hat{Y}} =150.092[1 \ 4 \ 24] \begin{pmatrix} 0.64637 & 0.00066 & -0.01445 \\ 0.00066 & 0.00078 & -0.00023 \\ -0.01445 & -0.00023 & 0.00041 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \\ 24 \end{pmatrix} = 24.6782$$

and the standard error is $\sqrt{24.6782}$ = 4.9677.

The sweepout technique is not the only method of matrix inversion. However, its application to the augmented matrix described above is a relatively simple and versatile method of obtaining most of the results commonly desired from a multiple regression analysis.

REFERENCES

Goodnight, J. H. 1978. The Sweep Operator: Its importance in statistical computing. in Proc. Eleventh Annual Symposium on the INTERFACE. Gallant A. R. and T. M. Gerig (ed), Inst. Statistics, N. C. State University, Raleigh, N. C.