Matrix Approach to Simple Linear Regression Analysis

Recall:

In the simple linear regression model:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \ i = 1,2,\ldots,n$$
$$E(\varepsilon_i) = 0, Var(\varepsilon_i) = \sigma^2 \quad and \quad Cov(\varepsilon_i, \varepsilon_j) = 0 \ for \ all \ i \neq j.$$

Then

$$E(Y_i) = \beta_0 + \beta_1 X_i \qquad and \qquad Var(Y_i) = \sigma^2.$$

The point estimates of β_0, β_1 are

$$\widehat{\beta}_1 = b_1 = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2} = \frac{S_{xy}}{S_{xx}}, \quad \widehat{\beta}_0 = b_0 = \bar{Y} - b_1\bar{X}$$

$$\widehat{\beta}_1 = b_1 = \sum_{i=1}^{n} K_i Y_i, \quad K_i = \frac{\left(X_i - \bar{X}\right)}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}, \ \widehat{\beta}_0 = \sum_{i=1}^{n} L_i Y_i, \quad L_i = \frac{1}{n} - \bar{X} K_i$$

$$Var(\widehat{\beta}_1) = \frac{\sigma^2}{\sum (X_i - \bar{X})^2}, \ Var(\widehat{\beta}_0) = \sigma^2\left[\frac{1}{n} + \frac{\bar{X}^2}{\sum (X_i - \bar{X})^2}\right],$$

The unbiased estimate of σ^2 is

$$s^2 = MSE = \widehat{\sigma^2} = \frac{SSE}{n-2}, \quad SSE = \sum_{i=1}^{n} e_i^2, \quad e_i = Y_i - \widehat{Y}_i, \quad i = 1,2,\ldots$$

Basic Definitions of the matrices and operations

The matrix A of with r rows and c columns will be represented either in full

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1c} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2c} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{ic} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rj} & \cdots & a_{rc} \end{bmatrix}$$

or in abbreviated form:

$$\mathbf{A} = [a_{ij}] \qquad i = 1, \ldots, r; j = 1, \ldots, c$$

or simply by a boldface symbol, such as **A**.

Transpose

The transpose of a matrix **A** is another matrix, denoted by $\mathbf{A}'$, that is obtained by interchanging corresponding columns and rows of the matrix **A**.

For example, if:

$$\underset{3\times 2}{\mathbf{A}} = \begin{bmatrix} 2 & 5 \\ 7 & 10 \\ 3 & 4 \end{bmatrix}$$

then the transpose $\mathbf{A}'$ is:

$$\underset{2\times 3}{\mathbf{A}'} = \begin{bmatrix} 2 & 7 & 3 \\ 5 & 10 & 4 \end{bmatrix}$$

Equality of Matrices

Two matrices **A** and **B** are said to be equal if they have the same dimension and if all corresponding elements are equal. Conversely, if two matrices are equal, their corresponding elements are equal. For example, if:

$$\underset{3\times 1}{\mathbf{A}} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \qquad \underset{3\times 1}{\mathbf{B}} = \begin{bmatrix} 4 \\ 7 \\ 3 \end{bmatrix}$$

then $\mathbf{A} = \mathbf{B}$ implies:

$$a_1 = 4 \qquad a_2 = 7 \qquad a_3 = 3$$

Similarly, if:

$$\underset{3\times 2}{\mathbf{A}} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \qquad \underset{3\times 2}{\mathbf{B}} = \begin{bmatrix} 17 & 2 \\ 14 & 5 \\ 13 & 9 \end{bmatrix}$$

then $\mathbf{A} = \mathbf{B}$ implies:

$$\begin{aligned} a_{11} &= 17 \qquad & a_{12} &= 2 \\ a_{21} &= 14 \qquad & a_{22} &= 5 \\ a_{31} &= 13 \qquad & a_{32} &= 9 \end{aligned}$$

Regression Examples

In regression analysis, one basic matrix is the vector **Y**, consisting of the n observations on the response variable:

$$\underset{n\times 1}{\mathbf{Y}} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$$

Note that the transpose $\mathbf{Y}'$ is the row vector:

$$\underset{1\times n}{\mathbf{Y}'} = [Y_1 \quad Y_2 \quad \cdots \quad Y_n]$$

Another basic matrix in regression analysis is the **X** matrix, which is defined as follows for simple linear regression analysis:

Another basic matrix in regression analysis is the **X** matrix, which is defined as follows for simple linear regression analysis:

$$\underset{n\times 2}{\mathbf{X}} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix}$$

The matrix **X** consists of a column of 1s and a column containing the n observations on the predictor variable X. Note that the transpose of **X** is:

$$\underset{2\times n}{\mathbf{X}'} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_n \end{bmatrix}$$

The **X** matrix is often referred to as the *design matrix*.

Matrix Addition and Subtraction

Adding or subtracting two matrices requires that they have the same dimension. The sum, or difference, of two matrices is another matrix whose elements each consist of the sum, or difference, of the corresponding elements of the two matrices. Suppose:

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

then:

$$\underset{3\times 2}{\mathbf{A}+\mathbf{B}} = \begin{bmatrix} 1+1 & 4+2 \\ 2+2 & 5+3 \\ 3+3 & 6+4 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ 4 & 8 \\ 6 & 10 \end{bmatrix}$$

Similarly:

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

In general, if:

$$\underset{r\times c}{\mathbf{A}} = [a_{ij}] \qquad \underset{r\times c}{\mathbf{B}} = [b_{ij}] \qquad i = 1, \ldots, r; j = 1, \ldots, c$$

then:

$$\underset{r\times c}{\mathbf{A}+\mathbf{B}} = [a_{ij} + b_{ij}] \qquad \text{and} \qquad \underset{r\times c}{\mathbf{A}-\mathbf{B}} = [a_{ij} - b_{ij}]$$

Formula generalizes in an obvious way to addition and subtraction of more than two matrices. Note also that $\mathbf{A}+\mathbf{B} = \mathbf{B}+\mathbf{A}$, as in ordinary algebra.

Regression Examples

The regression model:

$$Y_i = E\{Y_i\} + \varepsilon_i \qquad i = 1, \ldots, n$$

can be written compactly in matrix notation. First, let us define the vector of the mean responses:

$$\underset{n\times 1}{\mathbf{E}\{\mathbf{Y}\}} = \begin{bmatrix} E\{Y_1\} \\ E\{Y_2\} \\ \vdots \\ E\{Y_n\} \end{bmatrix}$$

and the vector of the error terms:

$$\underset{n\times 1}{\boldsymbol{\varepsilon}} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Recalling the definition of the observations vector **Y** , we can write the regression model as follows:

$$\underset{n\times 1}{\mathbf{Y}} = \underset{n\times 1}{\mathbf{E}\{\mathbf{Y}\}} + \underset{n\times 1}{\boldsymbol{\varepsilon}}$$

because:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} E\{Y_1\} \\ E\{Y_2\} \\ \vdots \\ E\{Y_n\} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} = \begin{bmatrix} E\{Y_1\} + \varepsilon_1 \\ E\{Y_2\} + \varepsilon_2 \\ \vdots \\ E\{Y_n\} + \varepsilon_n \end{bmatrix}$$

Thus, the observations vector **Y** equals the sum of two vectors, a vector containing the expected values and another containing the error terms.

Multiplication of a Matrix by a Scalar

A *scalar* is an ordinary number or a symbol representing a number. In multiplication of a matrix by a scalar, every element of the matrix is multiplied by the scalar. For example, suppose the matrix **A** is given by:

$$\mathbf{A} = \begin{bmatrix} 2 & 7 \\ 9 & 3 \end{bmatrix}$$

Then 4**A**, where 4 is the scalar, equals:

$$4\mathbf{A} = 4\begin{bmatrix} 2 & 7 \\ 9 & 3 \end{bmatrix} = \begin{bmatrix} 8 & 28 \\ 36 & 12 \end{bmatrix}$$

Similarly, $k\mathbf{A}$ equals:

$$k\mathbf{A} = k\begin{bmatrix} 2 & 7 \\ 9 & 3 \end{bmatrix} = \begin{bmatrix} 2k & 7k \\ 9k & 3k \end{bmatrix}$$

where k denotes a scalar.

Matrix multiplications

Here is example of matrix multiplication:

$$\mathbf{AB} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{bmatrix}$$

In general, if **A** has dimension $r \times c$ and **B** has dimension $c \times s$, the product **AB** is a matrix of dimension $r \times s$ whose element in the ith row and jth column is:

$$\sum_{k=1}^{c} a_{ik}b_{kj}$$

so that:

$$\underset{r\times s}{\mathbf{AB}} = \left[\sum_{k=1}^{c} a_{ik}b_{kj}\right] \qquad i = 1, \ldots, r; j = 1, \ldots, s$$

Thus, in the foregoing example, the element in the first row and second column of the product **AB** is:

$$\sum_{k=1}^{3} a_{1k}b_{k2} = a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32}$$

as indeed we found by taking the cross products of the elements in the first row of **A** and second column of **B** and summing.

Regression Examples

A product frequently needed is $\mathbf{Y'Y}$, where $\mathbf{Y}$ is the vector of observations on the response variable as defined in

$$\underset{1\times 1}{\mathbf{Y'Y}} = [Y_1 \quad Y_2 \quad \cdots \quad Y_n]\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \left[Y_1^2 + Y_2^2 + \cdots + Y_n^2\right] = \left[\sum Y_i^2\right]$$

Note that $\mathbf{Y'Y}$ is a 1×1 matrix, or a scalar. We thus have a compact way of writing a sum of squared terms: $\mathbf{Y'Y} = \sum Y_i^2$.

We also will need $\mathbf{X'X}$, which is a 2×2 matrix, \

$$\underset{2\times 2}{\mathbf{X'X}} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_n \end{bmatrix}\begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} = \begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix}$$

and $\mathbf{X'Y}$, which is a 2×1 matrix:

$$\underset{2\times 1}{\mathbf{X'Y}} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_n \end{bmatrix}\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix}$$

Identity Matrix. The identity matrix or unit matrix is denoted by $\mathbf{I}$. It is a diagonal matrix whose elements on the main diagonal are all 1s. Premultiplying or postmultiplying any $r \times r$ matrix $\mathbf{A}$ by the $r \times r$ identity matrix $\mathbf{I}$ leaves $\mathbf{A}$ unchanged. For example:

$$\mathbf{IA} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Similarly, we have:

$$\mathbf{AI} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Note that the identity matrix $\mathbf{I}$ therefore corresponds to the number 1 in ordinary algebra, since we have there that $1 \cdot x = x \cdot 1 = x$.

In general, we have for any $r \times r$ matrix $\mathbf{A}$:

$$\mathbf{AI} = \mathbf{IA} = \mathbf{A}$$

Vector and Matrix with All Elements Unity

A column vector with all elements 1 will be denoted by $\mathbf{1}$:

$$\underset{r\times 1}{\mathbf{1}} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

and a square matrix with all elements 1 will be denoted by $\mathbf{J}$:

$$\underset{r\times r}{\mathbf{J}} = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{bmatrix}$$

For instance, we have:

$$\underset{3\times 1}{\mathbf{1}} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \underset{3\times 3}{\mathbf{J}} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Note that for an $n \times 1$ vector $\mathbf{1}$ we obtain:

$$\underset{1\times 1}{\mathbf{1}'\mathbf{1}} = [1 \ \cdots \ 1]\begin{bmatrix}1\\ \vdots \\ 1\end{bmatrix} = [n] = n$$

and:

$$\underset{n\times n}{\mathbf{1}\mathbf{1}'} = \begin{bmatrix}1\\ \vdots \\ 1\end{bmatrix}[1 \ \cdots \ 1] = \begin{bmatrix}1 & \cdots & 1\\ \vdots & & \vdots \\ 1 & \cdots & 1\end{bmatrix} = \underset{n\times n}{\mathbf{J}}$$

Zero Vector

A zero vector is a vector containing only zeros. The zero column vector will be denoted by $\mathbf{0}$:

$$\underset{r\times 1}{\mathbf{0}} = \begin{bmatrix}0\\ 0\\ \vdots \\ 0\end{bmatrix}$$

Finding the Inverse

Up to this point, the inverse of a matrix $\mathbf{A}$ has been given, and we have only checked to make sure it is the inverse by seeing whether or not $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. But how does one find the inverse, and when does it exist?

An inverse of a square $r \times r$ matrix exists if the rank of the matrix is r. Such a matrix is said to be *nonsingular* or of *full rank*. An $r \times r$ matrix with rank less than r is said to be *singular* or *not of full rank*, and does not have an inverse. The inverse of an $r \times r$ matrix of full rank also has rank r.

Finding the inverse of a matrix can often require a large amount of computing. We shall take the approach in this book that the inverse of a 2×2 matrix and a 3×3 matrix can be calculated by hand. For any larger matrix, one ordinarily uses a computer to find the inverse, unless the matrix is of a special form such as a diagonal matrix. It can be shown that the inverses for 2×2 and 3×3 matrices are as follows:

1. If:

$$\underset{2\times 2}{\mathbf{A}} = \begin{bmatrix}a & b\\ c & d\end{bmatrix}$$

then:

$$\underset{2\times2}{\mathbf{A}^{-1}} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \begin{bmatrix} \dfrac{d}{D} & \dfrac{-b}{D} \\ \dfrac{-c}{D} & \dfrac{a}{D} \end{bmatrix}$$

where:

$$D = ad - bc$$

D is called the *determinant* of the matrix **A**. If **A** were singular, its determinant would equal zero and no inverse of **A** would exist.

2. If:

$$\underset{3\times3}{\mathbf{B}} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}$$

then:

$$\underset{3\times3}{\mathbf{B}^{-1}} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}^{-1} = \begin{bmatrix} A & B & C \\ D & E & F \\ G & H & K \end{bmatrix}$$

Example

$$\mathbf{A} = \begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix}$$

We have:

$$a = 2 \quad b = 4$$
$$c = 3 \quad d = 1$$

$$D = ad - bc = 2(1) - 4(3) = -10$$

Hence:

$$\mathbf{A}^{-1} = \begin{bmatrix} \dfrac{1}{-10} & \dfrac{-4}{-10} \\ \dfrac{-3}{-10} & \dfrac{2}{-10} \end{bmatrix} = \begin{bmatrix} -.1 & .4 \\ .3 & -.2 \end{bmatrix}$$

Regression Example

The principal inverse matrix encountered in regression analysis is the inverse of the matrix **X′X**

$$\underset{2\times 2}{\mathbf{X'X}} = \begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix}$$

Using rule , we have:

$$a = n \qquad b = \sum X_i$$
$$c = \sum X_i \qquad d = \sum X_i^2$$

so that:

$$D = n\sum X_i^2 - \left(\sum X_i\right)\left(\sum X_i\right) = n\left[\sum X_i^2 - \frac{(\sum X_i)^2}{n}\right] = n\sum(X_i - \bar{X})^2$$

Hence:

$$\underset{2\times 2}{(\mathbf{X'X})^{-1}} = \begin{bmatrix} \dfrac{\sum X_i^2}{n\sum(X_i - \bar{X})^2} & \dfrac{-\sum X_i}{n\sum(X_i - \bar{X})^2} \\ \dfrac{-\sum X_i}{n\sum(X_i - \bar{X})^2} & \dfrac{n}{n\sum(X_i - \bar{X})^2} \end{bmatrix}$$

Since $\sum X_i = n\bar{X}$ and $\sum(X_i - \bar{X})^2 = \sum X_i^2 - n\bar{X}^2$, we can simplify

$$\underset{2\times 2}{(\mathbf{X'X})^{-1}} = \begin{bmatrix} \dfrac{1}{n} + \dfrac{\bar{X}^2}{\sum(X_i - \bar{X})^2} & \dfrac{-\bar{X}}{\sum(X_i - \bar{X})^2} \\ \dfrac{-\bar{X}}{\sum(X_i - \bar{X})^2} & \dfrac{1}{\sum(X_i - \bar{X})^2} \end{bmatrix}$$

Some Basic Results for Matrices

We list here, without proof, some basic results for matrices which we will utilize in later work.

$$\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$$
$$(\mathbf{A}+\mathbf{B})+\mathbf{C}=\mathbf{A}+(\mathbf{B}+\mathbf{C})$$
$$(\mathbf{AB})\mathbf{C}=\mathbf{A}(\mathbf{BC})$$
$$\mathbf{C}(\mathbf{A}+\mathbf{B})=\mathbf{CA}+\mathbf{CB}$$
$$k(\mathbf{A}+\mathbf{B})=k\mathbf{A}+k\mathbf{B}$$
$$(\mathbf{A}')'=\mathbf{A}$$
$$(\mathbf{A}+\mathbf{B})'=\mathbf{A}'+\mathbf{B}'$$
$$(\mathbf{AB})'=\mathbf{B}'\mathbf{A}'$$

$$(\mathbf{ABC})'=\mathbf{C}'\mathbf{B}'\mathbf{A}'$$
$$(\mathbf{AB})^{-1}=\mathbf{B}^{-1}\mathbf{A}^{-1}$$
$$(\mathbf{ABC})^{-1}=\mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}$$
$$(\mathbf{A}^{-1})^{-1}=\mathbf{A}$$
$$(\mathbf{A}')^{-1}=(\mathbf{A}^{-1})'$$

Random Vectors and Matrices

Expectation of Random Vector or Matrix

Suppose we have $n=3$ observations in the observations vector $\mathbf{Y}$:

$$\underset{3\times 1}{\mathbf{Y}}=\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix}$$

The expected value of $\mathbf{Y}$ is a vector, denoted by $\mathbf{E}\{\mathbf{Y}\}$, that is defined as follows:

$$\underset{3\times 1}{\mathbf{E}\{\mathbf{Y}\}}=\begin{bmatrix} E\{Y_1\} \\ E\{Y_2\} \\ E\{Y_3\} \end{bmatrix}$$

Regression Example

Suppose the number of cases in a regression application is $n = 3$. The three error terms ε_1, ε_2, ε_3 each have expectation zero. For the error terms vector:

$$\underset{3\times 1}{\boldsymbol{\varepsilon}} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}$$

we have:

$$\underset{3\times 1}{\mathbf{E}\{\boldsymbol{\varepsilon}\}} = \underset{3\times 1}{\mathbf{0}}$$

since:

$$\begin{bmatrix} E\{\varepsilon_1\} \\ E\{\varepsilon_2\} \\ E\{\varepsilon_3\} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Variance-Covariance Matrix of Random Vector

If we have p random variables we can put them into a random vector as

$Y = [Y_1, \ldots, Y_n]'$, *then*

$$Var(Y) = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{p1} & \sigma_{p} & \cdots & \sigma_{pp} \end{bmatrix}, \ where$$

$$\sigma_{ij} = \sigma_{ji} = \operatorname{cov}(Y_i, Y_j) \quad and \quad \sigma_{ii} = \operatorname{var}(Y_i).$$

Regression Example

$$Var(\varepsilon) = \begin{bmatrix} \sigma_{11} & 0 & \dots & 0 \\ 0 & \sigma_{22} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_{pp} \end{bmatrix} = \sigma^2 I .$$

Some Basic Results

Frequently, we shall encounter a random vector **W** that is obtained by premultiplying the random vector **Y** by a constant matrix **A** (a matrix whose elements are fixed):

$$\mathbf{W} = \mathbf{AY}$$

Some basic results for this case are:

$$\mathbf{E\{A\} = A}$$

$$\mathbf{E\{W\} = E\{AY\} = AE\{Y\}}$$

$$Var(AW) = A \operatorname{var}(W) A'.$$

Simple Linear Regression Model in Matrix Terms

We are now ready to develop simple linear regression in matrix terms. Remember again that we will not present any new results, but shall only state in matrix terms the results obtained earlier. We begin with the normal error regression model

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \qquad i = 1, \ldots, n$$

This implies:

$$\begin{aligned} Y_1 &= \beta_0 + \beta_1 X_1 + \varepsilon_1 \\ Y_2 &= \beta_0 + \beta_1 X_2 + \varepsilon_2 \\ &\vdots \\ Y_n &= \beta_0 + \beta_1 X_n + \varepsilon_n \end{aligned}$$

Let us repeat these definitions and also define the β vector of the regression

$$\underset{n\times 1}{\mathbf{Y}} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \quad \underset{n\times 2}{\mathbf{X}} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} \quad \underset{2\times 1}{\boldsymbol{\beta}} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \quad \underset{n\times 1}{\boldsymbol{\varepsilon}} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Then

$$\underset{n\times 1}{\mathbf{Y}} = \underset{n\times 2}{\mathbf{X}}\ \underset{2\times 1}{\boldsymbol{\beta}} + \underset{n\times 1}{\boldsymbol{\varepsilon}}$$

since:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$= \begin{bmatrix} \beta_0 + \beta_1 X_1 \\ \beta_0 + \beta_1 X_2 \\ \vdots \\ \beta_0 + \beta_1 X_n \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 X_1 + \varepsilon_1 \\ \beta_0 + \beta_1 X_2 + \varepsilon_2 \\ \vdots \\ \beta_0 + \beta_1 X_n + \varepsilon_n \end{bmatrix}$$

and the conditions are

$$E(\varepsilon) = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = 0 \quad \textit{and} \quad \mathrm{var}(\varepsilon) = \sigma^2 I$$

Then

$$E(Y) = X\beta \quad \textit{and} \quad Var(Y) = \sigma^2 I$$