Chapter 5 – Matrix Approach to Simple Linear Regression Analysis

Po-Wei Chen

Institute of Statistics, National Chengchi University

May 1, 2023

1 Definition of Matrix in Regression Analysis

In regression analysis, one basic matrix is the vector $\mathbf{Y}$, 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:

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

Another basic matrix in regression analysis is the $\mathbf{X}$ 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 $\mathbf{X}$ consists of a column of 1s and a column containing the n observations on the predictor variable $\mathbf{X}$. Note that the transpose of $\mathbf{X}$ is:

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

2 Regression Examples

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

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

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

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

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

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

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

3 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:

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

In general, for a random vector $\mathbf{Y}$ the expectation is:

$$\underset{n\times 1}{\mathbf{E}\{\mathbf{Y}\}} = [E\{Y_i\}] \quad i = 1, ..., n$$

and for a random matrix $\mathbf{Y}$ with dimension $n \times p$, the expectation is:

$$\underset{n\times p}{\mathbf{E}\{\mathbf{Y}\}} = [E\{Y_{ij}\}] \quad i = 1, ..., n; j = 1, ..., p$$

4 Variance-Covariance Matrix of Random Vector

Consider again the random vector $\mathbf{Y}$ consisting of three observations Y_1, Y_2, Y_3. The variances of the three random variables, $\sigma^2\{Y_i\}$, and the covariances between any two of the random variables, $\sigma\{Y_i, Y_j\}$, are assembled in the $variance - covariance\ matrix$ of $\mathbf{Y}$, denoted by $\boldsymbol{\sigma}^2\{\mathbf{Y}\}$, in the following form:

$$\underset{3\times 3}{\boldsymbol{\sigma}^2\{\mathbf{Y}\}} = \begin{bmatrix} \sigma^2\{Y_1\} & \sigma\{Y_1, Y_2\} & \sigma\{Y_1, Y_3\} \\ \sigma\{Y_2, Y_1\} & \sigma^2\{Y_2\} & \sigma\{Y_2, Y_3\} \\ \sigma\{Y_3, Y_1\} & \sigma\{Y_3, Y_2\} & \sigma^2\{Y_3\} \end{bmatrix}$$

To generalize, the variance-covariance matrix for an $n \times 1$ random vector $\mathbf{Y}$ is:

$$\underset{n\times n}{\boldsymbol{\sigma}^2\{\mathbf{Y}\}} = \begin{bmatrix} \sigma^2\{Y_1\} & \sigma\{Y_1, Y_2\} & \cdots & \sigma\{Y_1, Y_n\} \\ \sigma\{Y_2, Y_1\} & \sigma^2\{Y_2\} & \cdots & \sigma\{Y_2, Y_n\} \\ \vdots & \vdots & & \vdots \\ \sigma\{Y_n, Y_1\} & \sigma\{Y_n, Y_2\} & \cdots & \sigma^2\{Y_n\} \end{bmatrix}$$

5 Some Basic Results

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

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

Some basic results for this case are:

$$\begin{aligned} \mathbf{E}\{\mathbf{A}\} &= \mathbf{A} \\ \mathbf{E}\{\mathbf{W}\} &= \mathbf{E}\{\mathbf{AY}\} = \mathbf{AE}\{\mathbf{Y}\} \\ \boldsymbol{\sigma}^2\{\mathbf{W}\} &= \boldsymbol{\sigma}^2\{\mathbf{AY}\} = \mathbf{A}\boldsymbol{\sigma}^2\{\mathbf{Y}\}\mathbf{A}' \end{aligned}$$

6 Exercises

***5.4. Flavor deterioration.** The results shown below were obtained in a small-scale experiment to study the relation between $^\circ F$ of storage temperature (X) and number of weeks before flavor deterioration of a food product begins to occur (Y).

i	1	2	3	4	5
X_i	8	4	0	-4	-8
Y_i	7.8	9.0	10.2	11.0	11.7

Assume that first-order regression model (2.1) is applicable. Using matrix methods, find (1) $\mathbf{Y'Y}$, (2) $\mathbf{X'X}$, (3) $\mathbf{X'Y}$.

Sol:

$$(1) \because \underset{5\times1}{\mathbf{Y}} = \begin{bmatrix} 7.8 \\ 9.0 \\ 10.2 \\ 11.0 \\ 11.7 \end{bmatrix} \text{ and so } \underset{1\times5}{\mathbf{Y}'} = \begin{bmatrix} 7.8 & 9.0 & 10.2 & 11.0 & 11.7 \end{bmatrix}$$

$$\Rightarrow \underset{1\times5}{\mathbf{Y'Y}} = \begin{bmatrix} 7.8 & 9.0 & 10.2 & 11.0 & 11.7 \end{bmatrix} \begin{bmatrix} 7.8 \\ 9.0 \\ 10.2 \\ 11.0 \\ 11.7 \end{bmatrix}$$

$$= 7.8^2 + 9.0^2 + 10.2^2 + 11.0^2 + 11.7^2 = 503.77$$

(2) $\because \underset{5\times 2}{\mathbf{X}} = \begin{bmatrix} 1 & 8 \\ 1 & 4 \\ 1 & 0 \\ 1 & -4 \\ 1 & -8 \end{bmatrix}$ and so $\underset{2\times 5}{\mathbf{X}'} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 8 & 4 & 0 & -4 & -8 \end{bmatrix}$

$$\Rightarrow \underset{2\times 2}{\mathbf{X}'\mathbf{X}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 8 & 4 & 0 & -4 & -8 \end{bmatrix} \begin{bmatrix} 1 & 8 \\ 1 & 4 \\ 1 & 0 \\ 1 & -4 \\ 1 & -8 \end{bmatrix}$$

$$= \begin{bmatrix} 1+1+1+1+1 & 8+4+0-4-8 \\ 8+4+0-4-8 & 64+16+0+16+64 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 160 \end{bmatrix}$$

(3) $\because \underset{2\times 5}{\mathbf{X}'} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 8 & 4 & 0 & -4 & -8 \end{bmatrix}$ and $\underset{5\times 1}{\mathbf{Y}} = \begin{bmatrix} 7.8 \\ 9.0 \\ 10.2 \\ 11.0 \\ 11.7 \end{bmatrix}$

$$\Rightarrow \underset{2\times 5}{\mathbf{X}'\mathbf{Y}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 8 & 4 & 0 & -4 & -8 \end{bmatrix} \begin{bmatrix} 7.8 \\ 9.0 \\ 10.2 \\ 11.0 \\ 11.7 \end{bmatrix}$$

$$= \begin{bmatrix} 7.8+9.0+10.2+11.0+11.7 \\ 62.4+36+0-44-93.6 \end{bmatrix} = \begin{bmatrix} 49.7 \\ -39.2 \end{bmatrix}$$

5.17. Consider the following functions of the random variables Y_1, Y_2, and Y_3:

$$\begin{aligned} W_1 &= Y_1 + Y_2 + Y_3 \\ W_2 &= Y_1 - Y_2 \\ W_3 &= Y_1 - Y_2 - Y_3 \end{aligned}$$

a. State the above in matrix notation.

b. Find the expectation of the random vector $\mathbf{W}$.

c. Find the variance-covariance matrix of $\mathbf{W}$.

Sol:

a. $$\begin{bmatrix} W_1 \\ W_2 \\ W_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix}$$

b. Let $\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & -1 & -1 \end{bmatrix}$.

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

$$\Rightarrow \mathbf{E}\{\mathbf{W}\} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} E\{Y_1\} \\ E\{Y_2\} \\ E\{Y_3\} \end{bmatrix} = \begin{bmatrix} E\{Y_1\} + E\{Y_2\} + E\{Y_3\} \\ E\{Y_1\} - E\{Y_2\} \\ E\{Y_1\} - E\{Y_2\} - E\{Y_3\} \end{bmatrix}$$

c. $\because \boldsymbol{\sigma^2}\{\mathbf{W}\} = \boldsymbol{\sigma^2}\{\mathbf{AY}\} = \mathbf{A}\boldsymbol{\sigma^2}\{\mathbf{Y}\}\mathbf{A}'$

$$\Rightarrow \boldsymbol{\sigma^2}\{\mathbf{W}\} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} \sigma^2\{Y_1\} & \sigma\{Y_1, Y_2\} & \sigma\{Y_1, Y_3\} \\ \sigma\{Y_2, Y_1\} & \sigma^2\{Y_2\} & \sigma\{Y_2, Y_3\} \\ \sigma\{Y_3, Y_1\} & \sigma\{Y_3, Y_2\} & \sigma^2\{Y_3\} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1^2 + \sigma_{21} + \sigma_{31} & \sigma_{12} + \sigma_2^2 + \sigma_{32} & \sigma_{13} + \sigma_{23} + \sigma_3^2 \\ \sigma_1^2 - \sigma_{21} & \sigma_{21} - \sigma_2^2 & \sigma_{13} - \sigma_{23} \\ \sigma_1^2 - \sigma_{21} - \sigma_{31} & \sigma_{12} - \sigma_2^2 - \sigma_{32} & \sigma_{13} - \sigma_{23} - \sigma_3^2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1^2 + \sigma_2^2 + \sigma_3^2 + 2\sigma_{12} + 2\sigma_{13} + 2\sigma_{23} & \sigma_1^2 - \sigma_2^2 + \sigma_{13} - \sigma_{23} & \sigma_1^2 - \sigma_2^2 - \sigma_3^2 - 2\sigma_{23} \\ \sigma_1^2 - \sigma_2^2 + \sigma_{13} - \sigma_{23} & \sigma_1^2 + \sigma_2^2 - 2\sigma_{12} & \sigma_1^2 + \sigma_2^2 - 2\sigma_{12} - \sigma_{13} + \sigma_{23} \\ \sigma_1^2 - \sigma_2^2 - \sigma_3^2 - 2\sigma_{23} & \sigma_1^2 + \sigma_2^2 - 2\sigma_{12} - \sigma_{13} + \sigma_{23} & \sigma_1^2 + \sigma_2^2 + \sigma_3^2 - 2\sigma_{12} - 2\sigma_{13} + 2\sigma_{23} \end{bmatrix}$$

Method: Least Squares Estimation (LSE)

$$\because Q = \sum_{i=1}^{n}(Y_i - E(Y_i))^2$$

$$= \sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i)^2$$

$$= (\underset{\sim}{Y} - \underset{\approx}{X}\underset{\sim}{\beta})'(\underset{\sim}{Y} - \underset{\approx}{X}\underset{\sim}{\beta})$$

$$= (\underset{\sim}{Y}' - \underset{\sim}{\beta}'\underset{\approx}{X}')(\underset{\sim}{Y} - \underset{\approx}{X}\underset{\sim}{\beta})$$

$$= \underset{\sim}{Y}'\underset{\sim}{Y} - \underset{\sim}{Y}'\underset{\approx}{X}\underset{\sim}{\beta} - \underset{\sim}{\beta}'\underset{\approx}{X}'\underset{\sim}{Y} + \underset{\sim}{\beta}'\underset{\approx}{X}'\underset{\approx}{X}\underset{\sim}{\beta}$$

$$= \underset{\sim}{Y}'\underset{\sim}{Y} - 2\underset{\sim}{\beta}'\underset{\approx}{X}'\underset{\sim}{Y} + \underset{\sim}{\beta}'\underset{\approx}{X}'\underset{\approx}{X}\underset{\sim}{\beta}$$

$$\Rightarrow \frac{\partial Q}{\partial \underset{\sim}{\beta}} = -2\underset{\approx}{X}'\underset{\sim}{Y} + (\underset{\approx}{X}'\underset{\approx}{X})\underset{\sim}{\beta} + (\underset{\approx}{X}'\underset{\approx}{X})'\underset{\sim}{\beta} = -2\underset{\approx}{X}'\underset{\sim}{Y} + 2(\underset{\approx}{X}'\underset{\approx}{X})\underset{\sim}{\beta}$$

Let $\frac{\partial Q}{\partial \underset{\sim}{\beta}} = \underset{\sim}{0}$ and substituting $\underset{\sim}{\mathrm{b}}$ to $\underset{\sim}{\beta}$.

$$\Rightarrow -2\underset{\approx}{X}'\underset{\sim}{Y} + 2(\underset{\approx}{X}'\underset{\approx}{X})\ \underset{\sim}{\mathrm{b}} = \underset{\sim}{0}.$$

$$\Rightarrow (\underset{\approx}{X}'\underset{\approx}{X})\ \underset{\sim}{\mathrm{b}} = \underset{\approx}{X}'\underset{\sim}{Y}$$

Properties of the Matrix Derivatives

(1) $\frac{\partial Ax}{\partial x} = A'$

(2) $\frac{\partial x'A}{\partial x} = A$

(3) $\frac{\partial x'x}{\partial x} = 2x$

(4) $\frac{\partial x'Ax}{\partial x} = Ax + A'x$