Section 5.9. Simple Linear Regression Model in Matrix Terms

Note. In this section we express the simple linear regression model based on n trials from Chapters 1 and 2 in terms of matrices.

Note. Recall from Section 1.8. Normal Error Regression Model, the normal error regression model based on n trials of observations Y_i based on the values X_i of the predictor variable, is

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

where:

Y_i is the observed response in the ith trial

X_i is a known constant, the value of the predictor variable in the ith trial

β_0 and β_1 are parameters

ε_i are independent random error terms with distribution $N(0, \sigma^2)$,

for $i = 1, 2, \ldots, n$. This is easily expressed as a linear equation in terms of vectors an matrices as follows.

Definition. The *matrix form of the normal error regression model* based on n trials is $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ where

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

Note. Since

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

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

Since $E\{Y_i\} = \beta_0 + \beta_1 X_i$ by Note 1.3.A, then $\mathbf{E}\{\mathbf{Y}\} = \mathbf{XB}$. We can use the alternative regression model (see Note 1.3.C) where we take $Y_i = \beta_0 X_0 + \beta_1 X_i + \varepsilon_i$ with $X_0 = 1$ to see the role played by matrix $\mathbf{X}$ (in particular, the role played by the column of 1's). Since each ε_i has distribution $N(0, \sigma^2)$ then $\{\varepsilon_i\} = 0$, $\sigma^2\{\varepsilon_i\} = \sigma^2$, and $\mathbf{E}\{\boldsymbol{\varepsilon}\} = [E\{\varepsilon_i\}] = \mathbf{0}$. Since the ε_i are assumed to be independent, then $\sigma\{\varepsilon_i, \varepsilon_j\} = 0$ for $i \neq j$. Therefore the variance-covariance matrix of the error terms is

$$\boldsymbol{\sigma}^2\{\boldsymbol{\varepsilon}\} = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix} = \sigma^2\mathbf{I}.$$

Revised: 9/17/2022