Proof: LSE

In vector form, we rewrite the multiple regression model as

$$Y = X\beta + e,$$

where $Y = (y_1, \ldots, y_N)'$, $\beta = (\beta_1, \ldots, \beta_K)'$, $e = (e_1, \ldots, e_N)'$, and

$$X = \begin{pmatrix} 1 & x_{1,2} & \cdots & x_{1,K} \\ 1 & x_{2,2} & \cdots & x_{2,K} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N,2} & \cdots & x_{N,K} \end{pmatrix},$$

and $e \sim N_N(0, \sigma^2 I)$, where I is the identity matrix.

Note we have an equal expression.

$$Y \sim N(X\beta, \sigma^2 I)\,.$$

We write the sum of squared errors as

$$SSE(\beta) = (Y - X\beta)'(Y - X\beta)\,.$$

The first-order condition requires that

$$\frac{\partial SSE(\beta)}{\partial \beta} = -2X'(Y - X\beta) = 0.$$

Thus, $-2X'(Y - X\beta) = 0$ implies that $X'Y = X'X\beta$ and $\beta = (X'X)^{-1}(X'Y)$. Thus, the LSE for β is

$$b = (X'X)^{-1}(X'Y)\,.$$

Q1: Let $K=2$, show that (b1, b2) in p. 29 of slides in Ch 5 reduces to the formula of (b1, b2) in (2.7) - (2.8).

模型：$y_i = b_1 + b_2 x_i + e_i$

矩陣形式為：$Y = X\beta + e$，其中 $Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}$，$\beta = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$，$X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$

$$X'X = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} = \begin{bmatrix} N & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} \Sigma y_i \\ \Sigma x_i y_i \end{bmatrix}$$

$$(X'X)^{-1} = \frac{1}{N\Sigma x_i^2 - (\Sigma x_i)^2} \begin{bmatrix} \Sigma x_i^2 & -\Sigma x_i \\ -\Sigma x_i & N \end{bmatrix}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{\underset{\text{行列式}}{ad-bc}} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = (X'X)^{-1} X'Y = \frac{1}{N\Sigma x_i^2 - (\Sigma x_i)^2} \begin{bmatrix} \Sigma x_i^2 & -\Sigma x_i \\ -\Sigma x_i & N \end{bmatrix} \cdot \begin{bmatrix} \Sigma y_i \\ \Sigma x_i y_i \end{bmatrix}$$

$$b_2 = \frac{(-\Sigma x_i)(\Sigma y_i) + N(\Sigma x_i y_i)}{N\Sigma x_i^2 - (\Sigma x_i)^2} = \frac{N\Sigma x_i y_i - \Sigma x_i \Sigma y_i}{N\Sigma x_i^2 - (\Sigma x_i)^2}$$

$\because \bar{x} = \frac{1}{N}\Sigma x_i$，$\bar{y} = \frac{1}{N}\Sigma y_i$　　$\therefore \Sigma x_i = N \cdot \bar{x}$，$\Sigma y_i = N \cdot \bar{y}$

b_2 分子：$N\Sigma x_i y_i - \Sigma x_i \Sigma y_i = N(\Sigma x_i y_i - \bar{x}\Sigma y_i) = N(\Sigma x_i y_i - N\bar{x}\bar{y})$

b_2 分母：$N\Sigma x_i^2 - (\Sigma x_i)^2 = N\Sigma x_i^2 - (N\bar{x})^2 = N(\Sigma x_i^2 - N\bar{x}^2)$

$$\therefore b_2 = \frac{N\Sigma x_i y_i - \Sigma x_i \Sigma y_i}{N\Sigma x_i^2 - (\Sigma x_i)^2} = \frac{N(\Sigma x_i y_i - N\bar{x}\bar{y})}{N(\Sigma x_i^2 - N\bar{x}^2)} = \frac{\Sigma x_i y_i - N\bar{x}\bar{y}}{\Sigma x_i^2 - N\bar{x}^2}$$

Equation (2.7)：$b_2 = \dfrac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2}$

Equation (2.7) 分子：$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = \Sigma x_i y_i - \bar{x}\Sigma y_i - \bar{y}\Sigma x_i + N\bar{x}\bar{y}$
$= \Sigma x_i y_i - \bar{x} \cdot N\bar{y} - \bar{y} \cdot N\bar{x} + N\bar{x}\bar{y} = \Sigma x_i y_i - N\bar{x}\bar{y}$

Equation (2.7) 分母：$\Sigma(x_i - \bar{x})^2 = \Sigma(x_i^2 - 2x_i\bar{x} + \bar{x}^2) = \Sigma x_i^2 - 2\bar{x}\Sigma x_i + N\bar{x}^2$
$= \Sigma x_i^2 - 2\bar{x} \cdot N\bar{x} + N\bar{x}^2 = \Sigma x_i^2 - N\bar{x}^2$

$\therefore$ Equation (2.7) 可改寫成：

$$b_2 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{\Sigma x_i y_i - N\bar{x}\bar{y}}{\Sigma x_i^2 - N\bar{x}^2}$$ 故得證

$$b_1 = \frac{(\Sigma x_i^2)(\Sigma y_i) - (\Sigma x_i)(\Sigma x_i y_i)}{N\Sigma x_i^2 - (\Sigma x_i)^2} = \frac{(\Sigma x_i^2)(N\bar{y}) - (N\bar{x})(\Sigma x_i y_i)}{N\Sigma x_i^2 - (N\bar{x})^2} = \frac{\bar{y}\Sigma x_i^2 - \bar{x}\Sigma x_i y_i}{\Sigma x_i^2 - N\bar{x}^2}$$

Equation (2.8)：$b_1 = \bar{y} - b_2\bar{x}$　　代入 $b_2 = \dfrac{\Sigma x_i y_i - N\bar{x}\bar{y}}{\Sigma x_i^2 - N\bar{x}^2}$

$$= \bar{y} - \frac{\Sigma x_i y_i - N\bar{x}\bar{y}}{\Sigma x_i^2 - N\bar{x}^2}\bar{x} = \frac{1}{\Sigma x_i^2 - N\bar{x}^2}\left[\bar{y}\Sigma x_i^2 - \cancel{\bar{y}N\bar{x}^2} - (\bar{x}\Sigma x_i y_i - \cancel{N\bar{x}^2\bar{y}})\right]$$

$\therefore$ Equation (2.8) 可改寫成：

$$b_1 = \bar{y} - b_2\bar{x} = \frac{\bar{y}\Sigma x_i^2 - \bar{x}\Sigma x_i y_i}{\Sigma x_i^2 - N\bar{x}^2}$$ 故得證

The Ordinary Least Squares (OLS) Estimators

$$b_2 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} \tag{2.7}$$

$$b_1 = \bar{y} - b_2\bar{x} \tag{2.8}$$

where $\bar{y} = \sum y_i/N$ and $\bar{x} = \sum x_i/N$ are the sample means of the observations on y and x.