Tutorial 4

Question 1

Consider a MLR model with two predictors as follows,

Y	X_1	X_2
y_1	x_{11}	x_{12}
y_2	x_{21}	x_{22}
$\vdots$	$\vdots$	$\vdots$
y_n	x_{n1}	x_{n2}

Please derive the detailed form for $\mathbf{X}'\mathbf{X}$, $\mathbf{X}'\mathbf{Y}$ and $\mathbf{Y}'\mathbf{Y}$.

Solution Note that

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix},$$

and

$$\mathbf{Y} = (y_1, y_2, \cdots, y_n)'.$$

In view of the above we obtain

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} n & \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i2} \\ \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i1}^2 & \sum_{i=1}^{n} x_{i1}x_{i2} \\ \sum_{i=1}^{n} x_{i2} & \sum_{i=1}^{n} x_{i1}x_{i2} & \sum_{i=1}^{n} x_{i2}^2 \end{bmatrix}, \qquad \mathbf{X}'\mathbf{Y} = (\sum_{i=1}^{n} y_i, \sum_{i=1}^{n} x_{i1}y_i, \sum_{i=1}^{n} x_{i2}y_i)'.$$

and

$$\mathbf{Y}'\mathbf{Y} = \sum_{i=1}^{n} y_i^2.$$

Question 2

Consider a multiple linear regression (MLR) model with one response (Y) and two predictors (X_1 and X_2), $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$, where $\epsilon \sim N(0, \sigma^2)$. Given that

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} 25 & 1315 & 506 \\ 1315 & 76323.42 & 26353.3 \\ 506 & 26353.3 & 10460 \end{pmatrix} \quad \mathbf{X}'\mathbf{Y} = \begin{pmatrix} 235.6 \\ 11821.432 \\ 4831.86 \end{pmatrix}$$

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} 2.779 & -0.0112 & -0.106 \\ -0.0112 & 0.146 \times 10^{-3} & 0.175 \times 10^{-3} \\ -0.106 & 0.175 \times 10^{-3} & 0.479 \times 10^{-2} \end{pmatrix}$$

$$\mathbf{Y}'\mathbf{Y} = 2784.11$$

Calculate the values of S_{yy}, SSE, $\widehat{\beta}$ and s^2. Why?

Solution It follows from the general expression of $\mathbf{X}'\mathbf{X}$ given in Q1 and the values of $\mathbf{X}'\mathbf{X}$ given in Q2 that $n = 25$. Moreover, by Q1 the first component of $\mathbf{X}'\mathbf{Y}$ is $\sum_{i=1}^{n} y_i$ and therefore $\bar{Y} = 235.6/25 = 9.424$. It follows that

$$S_{yy} = \sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - n(\bar{y})^2 = \mathbf{Y}'\mathbf{Y} - n(\bar{y})^2 = 2284.11 - 259.4242 = 63.8156.$$

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = (10.16, -0.07, 0.24)',$$

and

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \mathbf{Y}'\mathbf{Y} - \mathbf{Y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = 2784.11 - 2756.47 = 27.64.$$

Hence we have

$$s^2 = \frac{SSE}{n - p - 1} = \frac{27.64}{22} = 1.256.$$

Question 3

Write down the $\mathbf{X}$ matrix and β vector for each of the following (transformed) regression model.

- $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i1} x_{i2} + \epsilon_i, \quad i = 1, \cdots, 5;$
- $\sqrt{y}_i = \beta_0 + \beta_1 x_{i1} + \beta_2 \log x_{i2} + \epsilon_i, \quad i = 1, \cdots, 5.$

Solution

- The β vector and the design matrix are given as follows

$$\beta = (\beta_0, \beta_1, \beta_2).$$

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{11}x_{12} \\ 1 & x_{21} & x_{21}x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{51} & x_{51}x_{52} \end{bmatrix},$$

- The β vector and the design matrix are given as follows

$$\beta = (\beta_0, \beta_1, \beta_2).$$

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & \log x_{12} \\ 1 & x_{21} & \log x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{51} & \log x_{52} \end{bmatrix},$$