Joseph High
Hopkins ID: 9E1FDC

Homework # 3 EN.550.413

Problem 1: Let $y = X\beta + \varepsilon$ be a linear model where X is of size $n \times p$ and the error terms ε are independent, normally dist. with mean 0 and variance σ^2. Suppose furthermore that the columns of X can be partitioned as $X = [W \ Z]$ where W is of size $n \times q$ and is of full-column rank and Z is of size $n \times (p-q)$ and is of full-column rank, for some q satisfying $1 \le q \le p$, and $W^T Z = 0$. We now partition β as $\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$ where β_1 is of size $q \times 1$ and β_2 is of size $(p-q) \times 1$. Let $\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix}$ be the least square estimate of β.

(a) Show that $\hat{\beta}_1 = (W^T W)^{-1} W^T y$ and $\hat{\beta}_2 = (Z^T Z)^{-1} Z^T y$.

Pf: Using the methods of least squares to determine the least square estimates, with the model $y = X\beta + \varepsilon$, we have

$$Q = (y - X\beta)^T (y - X\beta) = y^T y - y^T X\beta - \beta^T X^T y + \beta^T X^T X \beta$$

(and $y^T X\beta = (y^T X\beta)^T = \beta^T X^T y$ since it is a 1×1 matrix.)

$$= y^T y - 2\beta^T X^T y + \beta^T X^T X \beta$$

Then, $$\frac{\partial Q}{\partial \beta} = -2X^T y + 2X^T X\beta = 0 \Rightarrow X^T X \beta = X^T y$$

Now, $X^T X$ is not necessarily invertible, so we consider the partitioned/block matrix of $X = [W \ Z]$ and $\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$

We have,

$$X^T X \beta = \begin{bmatrix} W^T \\ Z^T \end{bmatrix} [W \ Z] \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} W^T \\ Z^T \end{bmatrix} [y] = X^T y$$

$$\Rightarrow \begin{bmatrix} W^T W & W^T Z \\ Z^T W & Z^T Z \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} W^T y \\ Z^T y \end{bmatrix}$$

Now, since $W^T Z = 0_{p \times (p-q)} \Rightarrow (W^T Z)^T = Z^T W = 0_{(p-q) \times p}$

$\Longrightarrow$