Homework 7 - Answers

- From your textbook:

1. From the class notes, $\hat{\beta} = (X^T X)^{-1} X^T Y$. But

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{where} \quad x_i = (1 \ \ x_{t1} \cdots x_{tK}) \ \text{a} \ 1 \times (K+1) \ \text{vector.}$$

Hence, $$X^T X = [x_1^T \cdots x_n^T] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{t=1}^{n} x_t^T x_t \quad \text{and}$$

$$X^T Y = [x_1^T \cdots x_n^T] \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} = \sum_{t=1}^{n} x_t^T y_t. \quad \text{Consequently,}$$

$$\hat{\beta} = \left(\Sigma_{t=1}^{n} x_t^T x_t \right)^{-1} \left(\Sigma_{t=1}^{n} x_t^T y_t \right)$$

$$= \left(\frac{1}{n} \Sigma_{t=1}^{n} x_t^T x_t \right)^{-1} \left(\frac{1}{n} \Sigma_{t=1}^{n} x_t^T y_t \right)$$

2. (i) $y - Xb = y - Xb + X\hat{\beta} - X\hat{\beta}$

$$= y + X(\hat{\beta} - b) - X\hat{\beta}$$

$$= \hat{u} + X(\hat{\beta} - b)$$

Then, $(y - Xb)^T(y - Xb) = [\hat{u}^T + (\hat{\beta} - b)^T X^T][\hat{u} + X(\hat{\beta} - b)]$

$$= \hat{u}^T\hat{u} + \hat{u}^T X(\hat{\beta} - b) + (\hat{\beta} - b)X^T\hat{u} + (\hat{\beta} - b)^T X^T X(\hat{\beta} - b)$$

But $X^T\hat{u} = X^T(y - X\hat{\beta}) = X^T y - X^T X(X^T X)^{-1}X^T y = 0.$

$$(y - Xb)^T(y - Xb) = \hat{u}^T\hat{u} + (\hat{\beta} - b)^T X^T X(\hat{\beta} - b). \qquad (1)$$

(ii) Recall that $\hat{\beta}$ minimizers $S_n(\beta) = (y - X\beta)^T(y - X\beta)$.

Equation (1) shows that for any vector b,

$$S_n(b) = S_n(\hat{\beta}) + (\hat{\beta} - b)^T X^T X(\hat{\beta} - b).$$

$$= S_n(\hat{\beta}) + [X(\hat{\beta} - b)]^T X(\hat{\beta} - b)$$

Letting $v = X(\hat{\beta} - b)$, we have that

$$S_n(b) = S_n(\hat{\beta}) + \sum_{i=1}^{n} v_i^2$$

Since, $\sum_{i=1}^{n} v_i^2 \geq 0$, $S_n(b) \geq S_n(\hat{\beta})$. Since X is of full column rank, equality happens only when $b = \hat{\beta}$. Otherwise, i.e.,

$b \neq \hat{\beta} \Rightarrow S_n(b) > S_n(\hat{\beta})$.

5. (i) $E(\tilde{\beta}|X) = (Z^TX)^{-1}Z^T E(Y|X)$, since Z is a function of X.

$= (Z^TX)^{-1}Z^TX\beta = \beta$, since $E(Y|X) = X\beta$.

(ii) $V(\tilde{\beta}|X) = E\{(\tilde{\beta}-\beta)(\tilde{\beta}-\beta)^T|X\}$.

But $\tilde{\beta} = (Z^TX)^{-1}Z^T(X\beta + u) = \beta + (Z^TX)^{-1}Z^Tu$. This implies

$\tilde{\beta}-\beta = (Z^TX)^{-1}Z^Tu$. Then,

$$V(\tilde{\beta}|X) = E\{(Z^TX)^{-1}Z^Tuu^TZ(X^TZ)^{-1}|X\}$$

$$= \sigma^2 (Z^TX)^{-1}Z^TZ(X^TZ)^{-1}$$

We would prefer the estimator with smaller variance, given that both are unbiased. Consider,

$$V(\tilde{\beta}|X) - V(\hat{\beta}|X) = \sigma^2(Z^TX)^{-1}Z^TZ(X^TZ)^{-1} - \sigma^2(X^TX)^{-1}$$

$$= \sigma^2\{(Z^TX)^{-1}Z^TZ(X^TZ)^{-1} - (X^TX)^{-1}\}$$

$$= \sigma^2\{(Z^TX)^{-1}Z^TZ(X^TZ)^{-1} - \underbrace{(Z^TX)^{-1}Z^TX}_{I}(X^TX)^{-1}\underbrace{X^TZ(X^TZ)^{-1}}_{I}\}$$

$$= \sigma^2\{(Z^TX)^{-1}Z^T[I - X(X^TX)^{-1}X^T]Z(X^TZ)^{-1}\}$$

$$= \sigma^2\{A^TM_XA\}, \text{ where } A = Z(X^TZ)^{-1}$$

Because M_X is symmetric and idempotent, $V(\tilde{\beta}|X) - V(\hat{\beta}|X) \geq 0$. The inequality is in the sense that $\sigma^2\{A^TM_XA\}$ is positive semi-definite.

8.

(i) $M^T = \left(I - X(X^TX)^{-1}X^T\right)^T = I - X(X^TX)^{-1}X^T = M.$

$$MM = (I - X(X^TX)^{-1}X^T)(I - X(X^TX)^{-1}X^T)$$
$$= I - X(X^TX)^{-1}X^T - X(X^TX)^{-1}X^T + X(X^TX)^{-1}X^TX(X^TX)^{-1}X^T$$
$$= I - X(X^TX)^{-1}X^T = M$$

(ii) Let $X_{i\cdot}$ be the i^{th} rows of X. Recall that

$\hat{u} = y - X\hat{\beta} = y - X(X^TX)^{-1}X^Ty = My = M(X\beta + u) = Mu$. Then,

$$V(\hat{u}|X) = E\{(\hat{u} - E(\hat{u}|X))(\hat{u} - E(\hat{u}|X))^T | X\}$$

$= E\{Mu(Mu)^T|X\}$, since $E(\hat{u}|X) = E(Mu|X) = 0$

$= M\sigma^2 I M = \sigma^2 M.$ (2)

Since $\sigma^2 > 0$ and M is a variance matrix, all elements of the diagonal of M, i.e., M_{ii} for all $i = 1, 2, \dots, n$ are such that $M_{ii} \geq 0$.

Also, the i^{th} diagonal element of $X(X^TX)^{-1}X^T$ is given by $\underbrace{X_{i\cdot}}_{1\times(K+1)}(X^TX)^{-1}X_{i\cdot}^T$. But since $(X^TX)^{-1}$ is also a variance matrix (Recall $V(\hat{\beta}|X) = \sigma^2(X^TX)^{-1}$), $X_{i\cdot}(X^TX)^{-1}X_{i\cdot}^T \geq 0$ for all i.

Thus, $-X_{i\cdot}(X^TX)^{-1}X_{i\cdot}^T \leq 0$ and $1 - X_{i\cdot}(X^TX)^{-1}X_{i\cdot}^T \leq 1$. Since $M_{ii} = 1 - X_{i\cdot}(X^TX)^{-1}X_{i\cdot}^T$, $M_{ii} \leq 1$.

(iii) Proved in part (ii). See (2).

(iv) Follows from the fact that $M_{ii} \neq M_{jj}$ and $M_{ij} \neq 0$.

2. MATLAB question.

See code HOLSWage2.m.