Nguyen Xuan Binh 887799 Theory Exercise Week 1

Consider the linear model $y = X\beta + \varepsilon$, where $y, \varepsilon \in R^n$, $X \in R^{n\times(k+1)}$ and $\beta \in R^{k+1}$.
Let the standard assumptions (i) - (v), given in the lecture slides, hold. Let $M = I - X(X^TX)^{-1}X^T$ and recall that $\text{rank}(M) = n - (k+1)$

a) Let e be the estimated residual vector, that is, $e = y - \hat{y}$. Show that $Cov(e) = \sigma^2 M$

We have: $y = X\beta + \epsilon$ and $\hat{y} = Xb$

$\Rightarrow$ Estimated residual vector: $e = y - \hat{y} = X\beta + \epsilon - Xb = y - Xb$

Previously we know: $b = (X^TX)^{-1}X^Ty$

$\Rightarrow e = (I - X(X^TX)^{-1}X^T)y = My$ We need to find $cov(e)$

$cov(e) = cov(My) = cov(y)M^2$ (M is non-random)

$M^2 = (I - X(X^TX)^{-1}X^T)(I - X(X^TX)^{-1}X^T)$

$= I - 2X(X^TX)^{-1}X^T + X(X^TX)^{-1}X^TX(X^TX)^{-1}X^T$

$= I - (2X + X)(X^TX)^{-1}X^T = I - X(X^TX)^{-1}X^T = M$

$\Rightarrow cov(e) = cov(y)M^2 = cov(y)M = cov(X\beta + \epsilon)M$

$= E[(X\beta + \epsilon - E[X\beta + \epsilon])(X\beta + \epsilon - E[X\beta + \epsilon])^T]M$

We have: $E[X\beta + \epsilon] = E[X\beta] + E[\epsilon] = X\beta + 0 = X\beta$ (according to assumptions)

$\Rightarrow cov(e) = E(\epsilon\epsilon^T]M = E[(\epsilon - E[\epsilon])(\epsilon - E[\epsilon])^T]M = cov(\varepsilon)M = \sigma^2 IM$

($cov(\epsilon) = \sigma^2 I$ according to assumptions)

$\Rightarrow cov(e) = \sigma^2 IM = \sigma^2 M$ (proven)

b) Use previous exercises and part (a) and show that $s^2 = \frac{1}{n-k-1}\sum_{i=1}^{n} e_i^2$ is an unbiased estimator for $Var[\varepsilon_i] = \sigma^2$, that is, show that $E[s^2] = \sigma^2$

We have : $E[s^2] = E\left[\frac{1}{n-k-1}\sum_{i=1}^{n} e_i^2\right] = \frac{1}{n-k-1} E\left[\sum_{i=1}^{n} e_i^2\right] \Rightarrow$ Analyze $\sum_{i=1}^{n} e_i^2$

$\sum_{i=1}^{n} e_i^2 = e_1^2 + e_2^2 + \dots + e_n^2$ and $e = [e_1 \ e_2 \ \cdots \ e_n]^T$

$$e^Te = \begin{bmatrix} e_1^2 & \cdots & \cdots \\ \vdots & e_2^2 & \vdots \\ \vdots & \ddots & \vdots \\ \cdots & \cdots & e_n^2 \end{bmatrix} \Rightarrow \text{trace}(e^Te) = \sum_{i=1}^{n} e_i^2 \Rightarrow E[s^2] = \frac{1}{n-k-1} E[\text{trace}(e^Te)]$$

$\Rightarrow$ We need to find $E[\text{trace}(e^Te)]$

Transform e^Te: $e = My$ from (a) $\Rightarrow e^Te = y^TM^TMy = y^TMy$ (M is idempotent)

$\Rightarrow e^Te = (X^T\beta^T + \epsilon^T)M(X\beta + \epsilon)$

$= (X^T\beta^T + \epsilon^T)(MX\beta + M\epsilon)$. We need to find $MX\beta$

$MX\beta = (I - X(X^TX)^{-1}X^T)X\beta = X\beta - X(X^TX)^{-1}X^TX\beta = X\beta - X\beta = 0$

$\Rightarrow e^Te = (X^T\beta^T + \epsilon^T)M\epsilon = X^T\beta^TM\epsilon + \epsilon^TM\epsilon = X^T\beta^TM^T\epsilon + \epsilon^TM\epsilon$ ($M = M^T$)

$= (M\beta X)^T\epsilon + \epsilon^TM\epsilon$. $M\beta X = 0$ by expanding like above

$= \epsilon^TM\epsilon$

$\Rightarrow E[\text{trace}(e^Te)] = E[\text{trace}(\epsilon^TM\epsilon)] = E[\text{trace}(M\epsilon\epsilon^T)]$ (trace is cyclic)

$= \text{trace}(M \cdot E[\epsilon\epsilon^T]) = \text{trace}(M \cdot \sigma^2) = \sigma^2 \,\text{trace}(M)$

Idempotent matrix has eigenvalues of either 0 or 1 $\Rightarrow$ its trace is the number of non-zero eigenvalues $\Rightarrow$ Trace of idempotent matrix is equal to its rank

We know from the original knowledge that M's rank is $n-k-1$ and M is idempotent

$\Rightarrow E[\text{trace}(e^Te)] = \sigma^2 \text{trace}(M) = \sigma^2(n-k-1)$

$\Rightarrow E[s^2] = \frac{1}{n-k-1} E[\text{trace}(e^Te)] = \frac{1}{n-k-1}\sigma^2(n-k-1) = \sigma^2$

$\Rightarrow E[s^2] = \sigma^2$ (proven) $\therefore$ s^2 is unbiased estimator for $Var[\epsilon_i] = \sigma^2$