Let $\hat{\beta} \;=\; (X^T X)^{-1} X^T y$ *be the LS* $-$ *estimator and* $\hat{\sigma}^2{}_{ad} = \frac{1}{n-p-1}\ \hat{\in}^T \in$ the REML-estimator. Show that the following properties hold:

1. $E[\hat{\beta}] = \beta$
2. $Cov(\hat{\beta}) = \sigma^2 (X^T X)^{-1}$
3. $E[\hat{\sigma}^2{}_{ad}] = \sigma^2$

Solution:

i)

The LS-estimator may be written as follows:

$\hat{\beta} \;=\; (X^T X)^{-1} X^T y$

$= (X^T X)^{-1} X^T y \; * \; (X\,\beta + \in)$ [For linear regression model Yi = $\beta_0 \;+\; \beta_1 X_i \;+ \in_i$, where I = 1, 2, …., n]

$= \; (X^T X)^{-1} (X^T X)\beta \;+\; (X^T X)^{-1} X^T \in$

$= \; \beta \;+\; (X^T X)^{-1} X^T \in$

We can now get the expectation vector and the covariance matrix from the LS-estimator:

$E[\hat{\beta}] \;=\; E[\beta \;+\; (X^T X)^{-1} X^T \in]$

$= \; \beta \;+\; (X^T X)^{-1} X^T E[\in]$

$= \; \beta \; [E[\in] \;=\; 0;\ \textit{the ordinary multiple linear regression model}. (Lecture\_10, Page\ 3)]$

ii)

Cov$(\hat{\beta}] \;=\; E[(\hat{\beta} \;-\; \beta)(\hat{\beta} \;-\; \beta)^T]$ ………………………….. (1)

Now,

$\hat{\beta} \;-\; \beta \;=\; (X^T X)^{-1} X^T y \;-\; \beta$

$= (X^T X)^{-1} X^T (X\,\beta + \in) \;-\; \beta$

$= (X^T X)^{-1} (X^T X)\beta \;+\; (X^T X)^{-1} X^T \in \; -\beta$

$= (X^T X)^{-1} (X^T X)\beta \;+\; (X^T X)^{-1} X^T \in \; -\beta$

$= \beta \;+\; (X^T X)^{-1} X^T \in \; -\beta$

$= (X^T X)^{-1} X^T \in$

Now from equation (1) [2],

$$Cov(\hat{\beta}) = E[((X^TX)^{-1}X^T \in)((X^TX)^{-1}X^T \in)^T]$$

$$= E[(X^TX)^{-1}X^TX \in\in^T (X^TX)^{-1}]$$

$$= (X^TX)^{-1}X^TX(X^TX)^{-1}E[\in\in^T]$$

$$= (X^TX)^{-1}X^TX(X^TX)^{-1}\sigma^2 \; [E[\in\in^T] = \sigma^2 I, \textit{here I is the identity } m * m \textit{ matrix}]$$

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

iii)

we know,

$$\hat{\in} = y - X\hat{\beta}$$

$$= y - X((X^TX)^{-1}X^Ty)$$

$$= \left(I_n - X((X^TX)^{-1}X^T)\right)y$$

$$= \left(I_n - X((X^TX)^{-1}X^T)\right)(X\beta + \in)$$

$$= X\beta - X(X^TX)^{-1}(X^TX)\beta + \left(I_n - X((X^TX)^{-1}X^T)\right) \in$$

$$= 0 + \left(I_n - X((X^TX)^{-1}X^T)\right) \in$$

$$= M \in \left[take\ M = \left(I_n - X((X^TX)^{-1}X^T)\right)\right]$$

M is a (deterministic) symmetric and idempotent matrix; Hence, we can write:

$$\hat{\in}^T\hat{\in} = \in^T \left(I_n - X((X^TX)^{-1}X^T)\right) \in = \in^T M \in$$

Also, obtain a quadratic form in $\in$, with other words a scalar. With the help of the trace operator *tr* we obtain [1],

$$E[\hat{\in}^T\hat{\in}] = E[\in^T M \in]$$

$$= E[tr(\in^T M \in)] \; [\in^T M \in \textit{ is a scalar}]$$

$$= E[tr(M \in^T \in)] \; [use\ tr(XY) = tr(YX)]$$

$$= tr(ME[\in^T \in])$$

$$= tr(M\sigma^2 I_n)$$

$$= \sigma^2 tr(M)$$

$$= \sigma^2 tr\left(I_n - X((X^TX)^{-1}X^T)\right)$$

$$= \sigma^2\left[tr(I_n) - tr\left(X((X^TX)^{-1}X^T)\right)\right] \; [use\ tr(X+Y) = tr(Y) + tr(X)]$$

$= \sigma^2(n - p - 1)$ [The matrix $I_n - H$ is also symmetric and idempotent with $rk(I_n - H) = n - p - 1$]

Hence,

$$E[\sigma^2] = E\left[\frac{\hat{\epsilon}^T \hat{\epsilon}}{n-p-1}\right] = E\left[\frac{\sigma^2(n-p-1)}{n-p-1}\right] = \sigma^2$$

[1]. https://www.fm.mathematik.uni-muenchen.de/teaching/teaching_ss15/lectures/regression/notes.pdf

[2]. Proofs involving ordinary least squares, https://en.wikipedia.org/wiki/Proofs_involving_ordinary_least_squares