HW2.

↳ Hardik Prajapati (2678294168)

1. Gauss-Markov Theorem

Let $\tilde{\beta} = Cy$ be another linear Estimator.

where $C = (X^TX)^{-1}X^T + D$, D is an a non-zero matrix

$y = X\beta + \varepsilon$

$$E[\tilde{\beta}] = E[Cy]$$
$$= E\left[\left(\left(X^TX\right)^{-1}X^T + D\right)\left(X\beta + \varepsilon\right)\right]$$
$$= \left(\left(X^TX\right)^{-1}X^T + D\right)X\beta + \left(\left(X^TX\right)^{-1}X^T + D\right)E[\varepsilon]$$
$$= \left(\left(X^TX\right)^{-1}X^T + D\right)X\beta \qquad \rightarrow \left(E[\varepsilon] = 0\right)$$
$$= \underbrace{\left(X^TX\right)^{-1}X^TX}_{I}\beta + DX\beta$$
$$= (I + DX)\beta.$$

Since, $\tilde{\beta}$ is unbiased $\Rightarrow DX = 0$

$$\Rightarrow Var(\tilde{\beta}) = Var(Cy)$$
$$= C\,Var(y)\,C^T$$
$$= \sigma^2 CC^T$$
$$= \sigma^2\left(\left(X^TX\right)^{-1}X^T + D\right)\left(\left(X^TX\right)^{-1}X^T + D\right)^T$$
$$= \sigma^2\left(\left(X^TX\right)^{-1}X^T + D\right)\left(X\left(X^TX\right)^{-1} + D^T\right)$$
$$= \sigma^2\left(\left(X^TX\right)^{-1}X^TX\left(X^TX\right)^{-1} + \left(X^TX\right)^{-1}X^TD^T + DX\left(X^TX\right)^{-1} + DD^T\right)$$
$$= \sigma^2\left(X^TX\right)^{-1} + \sigma^2\left(X^TX\right)^{-1}\cancelto{0}{(DX)^T} + \sigma^2\cancelto{0}{DX}\left(X^TX\right)^{-1} + \sigma^2 DD^T$$

$$\Rightarrow Var(\tilde{\beta}) = \sigma^2(X^TX)^{-1} + \sigma^2 DD^T \qquad \rightarrow \{DX = 0\}$$

$$\Rightarrow Var(\tilde{\beta}) = Var(\hat{\beta}) + \underbrace{\sigma^2 DD^T}_{\geq 0}$$

$$\Rightarrow Var(\tilde{\beta}) \geq Var(\hat{\beta})$$

$$\Rightarrow Var(c^Ty) \geq Var(a^T\hat{\beta})$$

Hence proved

2. Given :- $X = [x_0 \; x_1 \; \cdots \; x_p]$ is an Orthogonal Matrix.

$$\Rightarrow X^{-1} = X^T$$

Now,
$$\begin{aligned}
\hat{\beta} &= (X^TX)^{-1}X^Ty \\
&= (X)^{-1}(X^T)^{-1}X^Ty \\
&= (X^T)\underbrace{(X^T)^{-1}X^T}_{I}y \\
&= X^Ty \\
&= \begin{bmatrix} x_0^T \\ \vdots \\ x_p^T \end{bmatrix} y \\
&= \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{12} & \cdots & x_{1n} \\ \vdots & & & \\ x_{p1} & x_{p2} & \cdots & x_{pn} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}
\end{aligned}$$

$$\Rightarrow \hat{\beta}_j = \hat{\sum_{i=1}} x_{ji} y_i$$

$$\Rightarrow \hat{\beta}_j = x_j^T y$$

3. a) $X = U \Sigma V^T$

$$X^T X = (U \Sigma V^T)^T (U \Sigma V^T)$$

$$= V \Sigma^T \underbrace{U^T U}_{I} \Sigma V^T$$

$$= V \Sigma^T \Sigma V^T$$

$$\Rightarrow (X^T X)^+ = V (\Sigma^T \Sigma)^+ V^T$$

$$= V \begin{bmatrix} 1/\sigma_1^2 & & \\ & \ddots & \\ & & 1/\sigma_r^2 \end{bmatrix} V^T \quad \rightarrow \text{as } \Sigma^T \Sigma \in \mathbb{R}^{(r \times r)}$$

$$\Rightarrow \beta = (X^T X)^+ X^T y$$

$$= V (\Sigma^T \Sigma)^{-1} V^T (U \Sigma V^T)^T y$$

$$= V (\Sigma^T \Sigma)^{-1} V^T V \Sigma^T U^T y$$

$$= V \Sigma^{-1} (\Sigma^T)^{-1} \Sigma^T U^T y$$

$$= V \Sigma^{-1} U^T y$$

$\Rightarrow \beta_{mns} = V \Sigma^{-1} U^T y$ is a solution to normal Equations

Q3. b) Let $\beta = \left[(X^T X)^+ + D\right] X^T y$, $D \in \mathbb{R}^{(p+1)\times(p+1)}$ (non-Zero Matrix

$\Rightarrow \beta = (X^T X)^+ X^T y + D X^T y$

$= V \Sigma^{-1} U^T y + D X^T y$ $\rightarrow$ (from 3(a))

$= \beta_{mns} + b$, $b \in \mathbb{R}^{(p+1)\times 1}$ (non-Zero vector)

$\Rightarrow \|\beta\| = \|\beta_{mns} + b\|$

$\Rightarrow \|\beta_{mns}\| \leq \|\beta\|$

Hence proved.

Q3. c) Penrose Properties

i) $X X^+ X = X$?

$(U \Sigma V^T)(V \Sigma^{-1} U^T)(U \Sigma V^T)$

$= U \Sigma (V^T V) \Sigma^{-1} (U^T U) \Sigma V^T$

$= U \Sigma \Sigma^+ \Sigma V^T$

$= U \Sigma V^T$

$= X$

Yes

ii) $X^+ X X^+ = X^+$?

$$(V\Sigma^{-1}U^T)(U\Sigma V^T)(V\Sigma^{-1}U^T)$$

$$= V\Sigma^{-1}(U^TU)\Sigma(V^TV)\Sigma^{-1}U^T$$

$$= V\Sigma^{-1}\Sigma\Sigma^{-1}U^T$$

$$= V\Sigma^{-1}U^T$$

$$= X^+$$

Yes.

iii) $(XX^+)^* = XX^+$?

$$((U\Sigma V^T)(V\Sigma^{-1}U^T))^*$$

$$= (U\Sigma\Sigma^{-1}U^T)^*$$

$$= (UU^T)^*$$

$$= (\text{conj}(UU^T))^T$$

$$= \boxed{UU^T}^T \qquad \rightarrow U \in \mathbb{R}^{n \times r}$$

$$= UU^T$$

Now, $XX^+ = (U\Sigma V^T)(V\Sigma^{-1}U^T)$

$$= U\Sigma\Sigma^{-1}U^T$$

$$= UU^T$$

Yes

iv) $(X^+X)^* = X^+X$?

$(X^+X)^*$

$= [(V\Sigma^{-1}U^T)(U\Sigma V^T)]^*$

$= [V\Sigma^{-1}\Sigma V^T]^*$

$= [VV^T]^*$

$= [VV^T]^T$

$= VV^T$

Now, $X^+X = (V\Sigma^{-1}U^T)(U\Sigma V^T)$

$= V\Sigma^{-1}\Sigma V^T$

$= VV^T$

Yes.

$\Rightarrow$ All the 4 properties were satisfied.

$\Rightarrow$ Yes, $V\Sigma^{-1}U^T$ is pseudo-inverse of X.