

6/11/24

- $l^T \beta$ is estimable, then $l^T G y$ where G is a LS g -inverse of X is the BLUE of $l^T \beta$.

$$\begin{aligned} \text{Var}(l^T G y) &= \sigma^2 l^T G G^T l \\ &= \sigma^2 l^T (X^T X)^- l \quad \text{for all } g\text{-inverse of } (X^T X) \end{aligned}$$

$$X^T X G G^T X^T X = \underbrace{X^T G^T X^T}_{X^T} \underbrace{X G X}_X = X^T X = X^T X (X^T X)^- X^T X$$

for all g -inverse of $X^T X$

$$\begin{aligned} X^T X G G^T X^T X &= X^T X (X^T X)^- X^T X \\ \Rightarrow X G G^T X^T X &= X (X^T X)^- X^T X \\ \Rightarrow \underline{X G G^T X^T} &= X (X^T X)^- X^T \end{aligned} \quad \left\{ \begin{array}{l} \text{w } X^T X = X \\ \text{for some } W \\ \\ X^T X Y = X^T \\ \text{for some } Y \end{array} \right.$$

$$\begin{aligned} l^T G G^T l &= l^T X \underline{G G^T X^T} u \\ &= l^T X (X^T X)^- X^T u \\ &= l^T (X^T X)^- l \end{aligned} \quad \left\{ \begin{array}{l} \text{as } l^T \in R(X) \\ \Rightarrow l^T = u^T X \end{array} \right.$$

- $l^T G y$ is the BLUE of $l^T \beta$ given $l^T \beta$ is estimable.

Suppose $c^T y$ and $d^T y$ are two BLUEs of $l^T \beta$

$\alpha c^T y + (1-\alpha) d^T y$ is an unbiased estimator

$$\begin{aligned}
 & \text{of } l^T \beta \\
 & \text{Var} \left((dC^T + (1-\alpha)d^T) y \right) \\
 & V(a^T y) = \sigma^2 a^T a \\
 & \boxed{E(a^T y) = l^T \beta}
 \end{aligned}$$

is a quadratic in α and is minimum when $\alpha=0$ and $\alpha=1$. This means in the expression, the coefficient of α^2 is 0. Conclude that $C=d$.

$$\begin{aligned}
 V(C^T y) &= \sigma^2 \sum C_i^2 \\
 E(C^T y) &= l^T \beta \Rightarrow C^T X \beta = l^T \beta \\
 &\Rightarrow \boxed{C^T X = l^T}
 \end{aligned}$$

• If $R(x) = p$ i.e., $R(x) = \mathbb{R}^p$ then all linear functions are estimable.

Theorem Let $R(x) = p$ and let $\hat{\beta}$ be the

BLUE of the column vectors β , $\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}$.

Then $\hat{\beta} = (X^T X)^{-1} X^T y$ and $\text{Var}(\hat{\beta}) = \sigma^2 (X^T X)^{-1}$.

Further, BLUE of $l^T \beta$ is $l^T \hat{\beta}$ and $\text{Var}(l^T \hat{\beta})$ is $\sigma^2 l^T (X^T X)^{-1} l$.

Gauss-Markov Theorem

Proof.

$$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix} = \begin{pmatrix} e_1^T \beta \\ e_2^T \beta \\ \vdots \\ e_p^T \beta \end{pmatrix}$$

$$\text{BLUE of } \beta \text{ is } \begin{pmatrix} e_1^T G y \\ e_2^T G y \\ \vdots \\ e_p^T G y \end{pmatrix} = I_p G y = \underline{G y}$$

$$XGX = X \Rightarrow X^T \underline{G^T X^T} = X^T$$

$$\Rightarrow X^T X G = X^T$$

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

$$\text{BLUE of } \beta \text{ is } \hat{\beta} = G y = (X^T X)^{-1} X^T y$$

$$\text{Var}((X^T X)^{-1} X^T y) = \sigma^2 (X^T X)^{-1} X^T X (X^T X)^{-1} \\ = \sigma^2 (X^T X)^{-1}$$

$$\text{For } l^T \beta, \text{ the BLUE is } l^T G y = l^T \underbrace{(X^T X)^{-1} X^T y}_{= l^T \hat{\beta}}$$

$$\text{Var}(l^T \hat{\beta}) = \text{Var}(l^T (X^T X)^{-1} X^T y) \\ = \sigma^2 l^T (X^T X)^{-1} X^T X (X^T X)^{-1} l \quad \left| \begin{array}{l} l^T G y \end{array} \right. \\ = \sigma^2 l^T (X^T X)^{-1} l$$

This theorem is known as Gauss-Markov theorem

Method of least square

$$y = X\beta + \epsilon$$

$$\epsilon^T \epsilon = (y - X\beta)^T (y - X\beta) = \|y - X\beta\|^2$$

$$\|y - P_X y\|^2 \leq \|y - X\beta\|^2 \quad \forall \beta \in \mathbb{R}^p$$

$\|y - X\beta\|^2$ is minimized when $X\beta$ is the orthogonal projection of y into $C(X)$, i.e. $X\beta = P_X y$

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

$$\Rightarrow \Delta F \approx X(X^T X)^{-1} X^T y$$

we take $(X^T X)^{-1} X^T y$ as the least square estimator of β , denoted by β_{LS}

$$\beta_{LS} = (X^T X)^{-1} X^T y$$

Alternative way

$$\begin{aligned} E^T E &= (y - X\beta)^T (y - X\beta) \\ &\hookrightarrow y^T y - \underbrace{y^T X\beta} - \underbrace{\beta^T X^T y} + \beta^T X^T X \beta \end{aligned}$$

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

$$\frac{\partial E^T E}{\partial \beta} = 0 - 2(y^T X)^T + 2(X^T X)\beta \quad \left| \begin{array}{l} \frac{\partial}{\partial y} a^T y \\ = a \end{array} \right.$$

if we equate it to 0.

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

$X^T y \in \mathcal{C}(X^T) = \mathcal{C}(X^T X)$, so $(X^T X)\beta = X^T y$ has a solution,

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