Recall:

Definition: Suppose $(Y, X\beta, \sigma^2 I)$ be a linear model and $p \in \mathbb{R}^m$.

The linear parametric function $p^T\beta$ is said to be estimable if

$\exists\ \ell \in \mathbb{R}^n$ s.t.

$$E[\ell^T Y] = p^T\beta.$$

(i.e. in other words $\exists\ \ell \in \mathbb{R}^n$ s.t. $\ell^T Y$ is unbiased estimate of $p^T\beta$)

Theorem: let $(Y, X\beta, \sigma^2 I_n)$ be a linear model. & $p \in \mathbb{R}^m$.

Then $p^T\beta$ is estimable $\Leftrightarrow$ $p \in \mathcal{C}(X^T)$

(where $\mathcal{C}(X^T) = \{ z : \exists y \quad X^T y = z \}$)

Independent Parametric linear functions

- From the Theorem (above : Feb 9^{th} 2023 week) it is clear that the number of linearly independent estimable linear parametric functions

$$= \text{dimension of } \mathcal{C}(X^T)$$

$$= \text{rank}(X)$$

$\Rightarrow$ $\underline{\beta} \in \mathbb{R}^m$ (Ex.), all components of $\underline{\beta}$ are estimable $\Leftrightarrow$ rank $(X) = m$.

- X has full column rank $\Leftrightarrow$

  X^TX is non-singular

  In this case, Normal equations have a unique solution, ie.

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

(unique least square estimate of $\underline{\beta}$)

Definition: (BLUE ; Best linear Unbiased Estimate)

A linear estimate $\underline{a}^T \underline{y}$ is called BLUE for estimating an estimable linear parametric function $\underline{p}^T \underline{\beta}$ if

(i) $E[\underline{a}^T \underline{y}] = \underline{p}^T \underline{\beta} \quad \forall\ \underline{\beta} \in \mathbb{R}^n$

(ii) $Var(\underline{a}^T \underline{y}) \leq var(\underline{l}^T \underline{y}) \quad \forall\ l \in \mathbb{R}^n$

s.t $\underline{l}^T \underline{y}$ is unbiased estimator of $\underline{p}^T \underline{\beta}$.

Theorem (Gauss-Markov Theorem): Suppose $(\underline{y}, X\underline{\beta}, \sigma^2 I_{n\times n})$ be a linear model and $\underline{p}^T \underline{\beta}$ be an estimable linear parametric function. Then the least square estimate of $\underline{p}^T \underline{\beta}$, written as $\underline{p}^T \hat{\underline{\beta}}$ for some solution of the normal equations $(X^T X)\hat{\underline{\beta}} = X^T \underline{y}$ is the BLUE of $\underline{p}^T \underline{\beta}$.

Proof :-

Since $\underline{p}^T \underline{\beta}$ is estimable then exists

$\underline{b} \in \mathbb{R}^m$ $\quad (X^T X)\underline{b} = \underline{p}$ $\quad (\because \underline{p} \in \mathcal{C}(X^T X))$ ⊛

$\therefore$ the least square estimate for $\underline{p}^T \underline{\beta}$

is $\underline{p}^T \underline{\hat{\beta}} = \underline{b}^T (X^T X) \underline{\hat{\beta}}$

$= \underline{b}^T X^T \underline{y}$

⊛

(l.s.e.)

$$\Rightarrow \quad E(\underline{p}^T \underline{\hat{\beta}}) = E(\underline{b}^T X^T \underline{y})$$

$$= E((X b)^T \underline{y})$$

Expectation is linear | linear model

$$\overset{⊛}{=} (X b)^T X \underline{\beta}$$

$\boxed{\text{unbiased}}$

$$= b^T X^T X \underline{\beta}$$

$$⊛= b^T \underline{\beta}$$

Thus $\underline{p}^T \underline{\hat{\beta}}$ is unbiased.

Let $\underline{l} \in \mathbb{R}^n$ be s.t. $\underline{l}^T \underline{y}$ is also an unbiased estimate of $\underline{p}^T \underline{\beta}$;

i.e $\quad \underline{p} = X^T \underline{l}$

$$\mathrm{Cov}(\ell^T \underline{y} - \underline{p}^T \underline{\hat{\beta}},\ \underline{p}^T \underline{\hat{\beta}})$$

$$= \mathrm{Cov}(\ell^T \underline{y} - b^T X^T \underline{y},\ b^T X^T \underline{y})$$

$$= \mathrm{Cov}(\ell^T \underline{y} - (Xb)^T \underline{y},\ (Xb)^T \underline{y})$$

$$= \mathrm{Cov}((\ell^T - (Xb)^T)\underline{y},\ (Xb)^T \underline{y})$$

$$\overset{(Ex)}{=} \sigma^2 (\ell - Xb)^T Xb$$

$$= \sigma^2 (\ell^T X b - b^T X^T X b)$$

$$= \sigma^2 (p^T b - p^T b)$$

$$= 0$$

$$\mathrm{Var}(\ell^T \underline{y}) = \mathrm{Var}(\underline{p}^T \underline{\hat{\beta}}) + \mathrm{Var}(\ell^T y - p^T \hat{\beta}^T)$$

$$(\because\ \ell^T \underline{y} = \ell^T \underline{y} - \underline{p}^T \hat{\beta}^T + p^T \hat{\beta}^T)$$

$$\Rightarrow \quad \mathrm{Var}(\ell^T \underline{y}) \geq \mathrm{Var}(\underline{p}^T \underline{\hat{\beta}}) \quad (\forall \ell \in \mathbb{R}^n \text{ s.t. } \underline{p} = X^T \ell)$$

$\therefore$ we have show $p^T \hat{\beta}$ is BLUE. $\square$

Variance of BLUE :

let $\underline{p}^T \underline{\beta}$ be estimable i.e. $\underline{p} \in \mathcal{C}(X^T X)$

$\Rightarrow \exists \ \underline{b} \in \mathbb{R}^m$ s.t. $(X^T X)\underline{b} = \underline{p}$

$\hat{\theta} = \underline{p}^T \underline{\hat{\beta}}$ – be the BLUE of $\underline{p}^T \underline{\beta}$

$$\begin{aligned}
\mathrm{Var}(\hat{\theta}) &= \mathrm{var}(\underline{p}^T \underline{\hat{\beta}}) \\
&= \mathrm{var}(\underline{b}^T (X^T X) \underline{\hat{\beta}}) \\
&= \mathrm{var}(b^T X^T \underline{y}) \\
&\overset{\text{Ex}}{=} \sigma^2 b^T (X^T X) \underline{b} = \sigma^2 p^T b
\end{aligned}$$

Recall from earlier week:

Proposition : $X \underline{\beta}$ is estimable and its least square estimate is given by $P_X \underline{Y}$ where

P_X is the orthogonal projection onto $\mathcal{C}(X)$

$\underline{p} \in \mathcal{C}(X^T) \Rightarrow \exists\, \underline{c} \in \mathbb{R}^n$ s.t. $\underline{p} = X^T \underline{c}$

$$\therefore \hat{\theta} = p^T \hat{\beta} = \underline{c}^T X \hat{\beta}$$

$$\therefore \operatorname{var}(\hat{\theta}) = \operatorname{var}(\underline{c}^T X \hat{\beta})$$

$$= c^T \operatorname{var}(X\hat{\beta})\, \underline{c}$$

$$= \underline{c}^T \operatorname{var}(P_X \underline{Y})\, \underline{c}$$

$$\overset{(Ex)}{=} \sigma^2\, \underline{c}^T P_X \underline{c}$$

$\boxed{\begin{array}{l}\text{As } X\hat{\beta} = P_X \underline{Y} \\ \operatorname{var}(X\hat{\beta}) = \sigma^2 P_X\end{array}}$

linear zero Estimators :

A linear function of the data vector $\underline{y}$ says

$l^T \underline{y}$ where $l \in \mathbb{R}^n$ is called a

linear zero estimator if $E(l^T \underline{y}) = 0 \;\; \forall\, \beta \in \mathbb{R}^n$

$l^T \underline{y}$ is a linear zero estimator

$$\Leftrightarrow \quad l^T X \underline{\beta} = 0 \quad \forall\, \beta \in \mathbb{R}^m$$

$$\Leftrightarrow \quad X^T l = 0 \quad \Leftrightarrow \quad l \in \text{Null}(X^T) \quad \text{—} \; + $$

$$\overset{(Ex.)}{\text{||}} \quad \mathcal{C}(X)^{\perp}$$

$\therefore$ $\underline{p}^T \hat{\underline{\beta}}$ is BLUE of estimable linear parametric function $\underline{p}^T \underline{\beta}$

when $\underline{p} \in \mathcal{C}(X^T) = \mathcal{C}(X^T X)$

$\exists\, \underline{b} \in \mathbb{R}^m \quad (X^T X)\underline{b} = \underline{p}$

$$\text{Cov}(\underline{p}^T \hat{\beta}, l^T \underline{y}) = \text{Cov}(b^T X^T X \hat{\beta}, l^T y)$$

$$= \text{Cov}(b^T X^T \underline{y}, l^T \underline{y})$$

$$= \sigma^2 b^T X^T l^T \quad - (*)$$

If $l^T \underline{y}$ is a linear zero estimator then $(*)$

$$\Rightarrow \text{Cov}(\underline{p}^T \hat{\beta}, l^T \underline{y}) \overset{(+)}{=} 0$$

$\Rightarrow$ Every BLUE is uncorrelated with every linear zero estimator.

Theorem :- Suppose $\underline{p}^T \underline{\beta}$ is estimable. A linear unbiased estimate $\eta^T \underline{y}$ is the BLUE $\iff$ $\text{Cov}(\eta^T \underline{y}, l^T \underline{y}) = 0$

$\forall\, l \in \mathbb{R}^n$ s.t. $l^T \underline{y}$ is the linear zero estimator.

Proof:- Above we showed the "only if".

$\boxed{\Leftarrow}$

let $\lambda^T \underline{y}$ be another unbiased estimate of $p^T \underline{\beta}$

$\Rightarrow$ (Ex) $(\eta - \lambda)^T \underline{y}$ is a linear zero estimate

$\Rightarrow$ (Assumption) $Cov(\eta^T \underline{y}, (\eta - \lambda)^T \underline{y}) = 0$ $\perp$ ⋆

$\Rightarrow$ $var(\eta^T \underline{y}) = var(\eta^T y - \lambda^T y + \lambda^T y)$

$\underset{⋆}{=} var((\eta - \lambda)^T y) + var(\lambda^T y)$

$\Rightarrow$ $var(\eta^T \underline{y}) \leq var(\lambda^T \underline{y})$

$\therefore$ $\eta^T \underline{y}$ is BLUE.

$\square$

Corollary:- For an estimable linear parametric function BLUE is unique & is given by least square estimate.

Assumption of Normality :

Suppose we assume

$$\underline{y} \sim \text{Normal}(X\underline{\beta}, \sigma^2 I_{n\times n})$$

(i.e. $(\underline{y}, X\underline{\beta}, \sigma^2 I_{n\times n})$ linear model has an extra assumption that errors are
$$\underline{\varepsilon} \sim \text{Normal}(\underline{0}, \sigma^2 I_{n\times n})$$

Recall : The likehood function for β, σ^2 is given

$$L(\underline{\beta}, \sigma^2 \mid \underline{y}) = \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left(-\frac{1}{2\sigma^2}(\underline{Y}-X\underline{\beta})^T(\underline{Y}-X\underline{\beta})\right)$$

$$= \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left(-\frac{1}{2\sigma^2}\|\underline{Y}-X\underline{\beta}\|^2\right)$$

Then if $\sigma^2 > 0$ is given

$$\max_{\beta \in \mathbb{R}^m} L(\underline{\beta}, \sigma^2 \mid \underline{y}) = \min_{\beta \in \mathbb{R}^m} \|\underline{Y}-X\underline{\beta}\|^2$$

M.L.E. of $\underline{\beta}$

l.s.e of $\underline{\beta}$

Remarks:

① M.L.E. of $\underline{\beta}$ exists but may not be unique.

② Unique iff $\text{rank}(X) = m$.

③ If $\underline{b}^T \underline{\beta}$ is estimable the M.L.E. of $\underline{b}^T \underline{\beta}$ is unique, given by its L.S.E.

(i.e. $\underline{b}^T \hat{\underline{\beta}}$ for any solution of normal equation $(X^T X)\hat{\beta} = X^T \underline{y}$)

④ $\underline{\eta} = X\underline{\beta}$ has a unique M.L.E.

$$\hat{\eta} = X\hat{\beta} = P_x \underline{Y}.$$

⑤ $\hat{\eta} \sim \text{Normal}(X\underline{\beta}, \sigma^2 P_x)$