ON AN EXTENSIVE PROPERTY OF BLUE ESTIMATOR

Hsiang-Chuan Liu

Keyword:

General linear model, BLUE, BLE $(\underset{\sim}{\beta} \mid \underset{\sim}{B})$

Abstract:

We provide an extensive theorem of the best linear unbiased estimator in general linear model

1. Introduction:

Let $Y = \mathbf{X}\underset{\sim}{\beta} + \underset{\sim}{\epsilon}$, $\underset{\sim}{\epsilon} \sim Nn(\underset{\sim}{0}, \Sigma_n)$ be a general Gaussian linear model

where $\underset{\sim}{Y}$ is an $n \times 1$ observable vector of random variables

$\mathbf{X}$ is an $n \times P$ full rank design matrix(n>P)

β is a $p \times 1$ vector of unknown parameters

ϵ is an $n \times 1$ unobservable vector of normal random variable

Σ_n is a known positive definite $n \times n$ matrix

It is well known that $\underset{\sim}{\hat{\beta}} = (\mathbf{X}^T \Sigma_n^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Sigma_n^{-1} \underset{\sim}{Y} \sim N_p(\beta, (\mathbf{X}^T \Sigma_n^{-1} \mathbf{X})^{-1})$ and $\hat{\beta}$ is the best linear unbiased estimator of β i.e $\hat{\beta}$ =BLUE(β), Here we develope the best linear estimator with the given bias vector based on the BLUE estimator

2. Definition:

Let $\hat{\hat{\beta}}$ be a linear estimator of $\underset{\sim}{\beta}$ with given bias $\underset{\sim}{B}$, we define $\hat{\hat{\beta}}$ =BLE($\beta \mid \underset{\sim}{B}$) to be the best linear estimator of $\underset{\sim}{\beta}$ with given Bias $\underset{\sim}{B}$, if for any linear estimator of $\underset{\sim}{\beta}$, $\underset{\sim}{\breve{\beta}}$ with the same bias $\underset{\sim}{B}$ such that the matrix〔MSE($\breve{\beta}$)-MSE($\hat{\hat{\beta}}$)〕 is non-negative definite where MSE($\breve{\beta}$)=E($\breve{\beta} - \underset{\sim}{\beta}$)($\breve{\beta} - \underset{\sim}{\beta}$)^T. Specially we say that $\hat{\hat{\beta}}$ = BLE($\underset{\sim}{\beta}$ $\underset{\sim}{B}$) is the best linear unbiased estimator if $\underset{\sim}{B} = \underset{\sim}{0}$, that is, BLE($\beta \mid \underset{\sim}{0}$) = BLUE(β)

3. Theoremm:

〔Theorem〕 Let $\underset{n\times 1}{Y} = \underset{n\times p}{\mathbf{X}} \underset{p\times 1}{\beta} + \underset{n\times 1}{\epsilon}$, $\epsilon \sim N_n(0, \Sigma_n)$ be general Gaussian linear madel

$\underset{\sim}{B} = \underset{p\times p}{T} \underset{\sim}{\beta}$ ……………………………………………………①

$C = T\, I_p$

$\underset{\sim}{\hat{\beta}} = (\mathbf{X}^T \Sigma_n^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Sigma^{-1} \underset{\sim}{Y} = \text{BLUE}(\underset{\sim}{\beta})$

if and only if $\hat{\underset{\sim}{\beta}}* = C\hat{\underset{\sim}{\beta}}$

then $\hat{\underset{\sim}{\beta}}* = BLE(\underset{\sim}{\beta} \mid \underset{\sim}{B})$

〔proof〕 (→) If $\hat{\underset{\sim}{\beta}}* = C\hat{\underset{\sim}{\beta}}$ now to prove that

$\hat{\underset{\sim}{\beta}}* = BLE(\underset{\sim}{\beta} \mid \underset{\sim}{B})$

Let $\breve{\underset{\sim}{\beta}} = H\underset{\sim}{Y}$ ……………………………………………………………②

where $\begin{cases} \underset{p\times n}{H} = \underset{p\times n}{C}(\mathbf{X}^T\Sigma_n^{-1}\mathbf{X})^{-1}\mathbf{X}^T\Sigma_n^{-1} + \underset{p\times n}{D} \\ D\text{: arbitrary known } p\times n \text{ matrix} \end{cases}$

then $\breve{\underset{\sim}{\beta}}$ is any linear estimator of $\underset{\sim}{\beta}$

Let $E(\breve{\underset{\sim}{\beta}}) = \underset{\sim}{\beta} + \underset{\sim}{B} = \underset{\sim}{\beta} + T\underset{\sim}{\beta} = (T + I_p)\underset{\sim}{\beta}$

$= C\underset{\sim}{\beta}$ ……………………………………………………………③

then

$E(\breve{\underset{\sim}{\beta}}) = E(H\underset{\sim}{Y}) = E[H(\mathbf{X}\underset{\sim}{\beta} + \underset{\sim}{\epsilon})]$

$= H\mathbf{X}\underset{\sim}{\beta} + HE(\underset{\sim}{\epsilon}) = H\mathbf{X}\underset{\sim}{\beta}$ ……………………………………………④

$= [C(\mathbf{X}^T\Sigma_n^{-1}\mathbf{X})^{-1}\mathbf{X}^T\Sigma_n^{-1} + D]\mathbf{X}\beta$

$= C\underset{\sim}{\beta} + D\mathbf{X}\underset{\sim}{\beta}$ ……………………………………………………⑤

compare ⑤ to ③ and ④ we get

$\begin{cases} D\mathbf{X} = 0 \\ H\mathbf{X} = C \end{cases}$ ……………………………………………………⑥

i.e

$\breve{\underset{\sim}{\beta}}$ is a linear estimator of $\underset{\sim}{\beta}$ with bias

$\underset{\sim}{B} = T\underset{\sim}{\beta}$ if $D\mathbf{X} = \underset{\sim}{0}$ and $H\mathbf{X} = C$

then

$Var(\breve{\underset{\sim}{\beta}})$

$= E(\breve{\underset{\sim}{\beta}} - C\underset{\sim}{\beta})(\breve{\underset{\sim}{\beta}} - C\underset{\sim}{\beta})^T$

$= E(\breve{\underset{\sim}{\beta}}\breve{\underset{\sim}{\beta}}^T) - C\underset{\sim}{\beta}(C\underset{\sim}{\beta})^T$

$= E\{[H(\mathbf{X}\underset{\sim}{\beta} + \underset{\sim}{\epsilon})][\underset{\sim}{\epsilon}^T + \underset{\sim}{\beta}^T\mathbf{X}^T)H^T]\} - C\underset{\sim}{\beta}\underset{\sim}{\beta}^TC^T$

$= H\mathbf{X}\underset{\sim}{\beta}\underset{\sim}{\beta}^T\mathbf{X}^TH^T + E[H\underset{\sim}{\epsilon}\underset{\sim}{\epsilon}^TH^T] - C\underset{\sim}{\beta}\underset{\sim}{\beta}^TC^T$ ……………………………⑦

Using $H\mathbf{X} = C$ and $E[\underset{\sim}{\epsilon}\underset{\sim}{\epsilon}^T] = \Sigma_n$

formula ⑦ may be rewritten as:

$Var(\breve{\underset{\sim}{\beta}}) = H\Sigma_nH^T$

$= [C(\mathbf{X}^T\Sigma_n^{-1}\mathbf{X})^{-1}\mathbf{X}^T\Sigma_n^{-1} + D]\Sigma_n[C(\mathbf{X}^T\Sigma_n^{-1}\mathbf{X})^{-1}\mathbf{X}^T\Sigma_n^{-1} + D]^T$

$= [C(\mathbf{X}^T\Sigma_n^{-1}\mathbf{X})^{-1}\mathbf{X}^T\Sigma_n^{-1} + D]\Sigma_n[D^T + \Sigma_n^{-1}\mathbf{X}(\mathbf{X}^T\Sigma_n^{-1}\mathbf{X})^{-1}C^T]$

$= [C(\mathbf{X}^T\Sigma_n^{-1}\mathbf{X})^{-1}\mathbf{X}^TD^T + C(\mathbf{X}^T\Sigma_n^{-1}\mathbf{X})^{-1}C^T + D\Sigma_nD^T$

$+ D\mathbf{X}(\mathbf{X}^T\Sigma_n^{-1}\mathbf{X})^{-1}C^T$ ……………………………………⑧

Since $\begin{cases} D\mathbf{X} = 0 \\ \mathbf{X}^TD^T = (D\mathbf{X})^T = \underset{\sim}{0} \end{cases}$

formula ⑧ may be reduced to

$Var(\breve{\underset{\sim}{\beta}}) = C(\mathbf{X}^T\Sigma_n^{-1}\mathbf{X})^{-1}C^T + D\Sigma_nD^T$ ……………………………………⑨

Hence

$MSE(\breve{\beta}) = Var(\breve{\beta}) + \underset{\sim}{B}\underset{\sim}{B}^T$

$= C(\mathbf{X}^T \Sigma_n^{-1}\mathbf{X})^{-1}C^T + D\Sigma_n D + \underset{\sim}{B}\underset{\sim}{B}^T$ ………………………………………⑩

now consider $MSE(\hat{\beta}*)$ Since $\hat{\beta}* = C\hat{\beta}$ We get

$MES(\hat{\beta}*) = MSE(C\hat{\beta})$

$= Var(C\hat{\beta}) + \underset{\sim}{B}\underset{\sim}{B}^T$

$= C\ Var(\hat{\beta})C^T + \underset{\sim}{B}\underset{\sim}{B}^T$

$= C(\mathbf{X}^T \Sigma_n^{-1}\mathbf{X})^{-1}C^T + \underset{\sim}{B}\underset{\sim}{B}^T$ ………………………………………⑪

Since $D\Sigma_n D^T$ is non-negative difinite

compare ⑪ to ⑩ we get

$MSE(\hat{\beta}*) \leq MSE(\breve{\beta})$ i.e$[MSE(\breve{\beta}) - MSE(\hat{\beta}*)]$ is non-negative definite

where $\breve{\beta}$ is any linear estimator of β with bias $\underset{\sim}{B} = T\beta$

Therefore $\hat{\beta}* = BLE(\beta \mid \underset{\sim}{B})$

$(\leftarrow)$If $\hat{\beta}* = BLE(\beta \mid \underset{\sim}{B})$

where $\underset{\sim}{B} = T\beta$, $E(\hat{\beta}*) = C\beta$ ………………………………………⑫

then we get $MSE(\hat{\beta}*) \leq MSE(\breve{\beta}*)$

where $\breve{\beta}*$ is any linear estimator of β with bias $\underset{\sim}{B} = T\beta$

Hence $MSE(\hat{\beta}*) \leq MSE(C\hat{\beta})$ ………………………………………⑬

where $\hat{\beta} = (\mathbf{X}^T\Sigma_n^{-1}\mathbf{X})^{-1}\mathbf{X}^T\Sigma_n^{-1}\underset{\sim}{Y}$

now to prove that $\hat{\beta}* = C\hat{\beta}$

Let $\hat{\beta}* = \underset{p\times n}{H*}\underset{\sim}{Y}$ ………………………………………⑭

and

$\hat{\beta}* - C\hat{\beta} = H*\underset{\sim}{Y} - C(\mathbf{X}^T\Sigma_n^{-1}\mathbf{X})^{-1}\mathbf{X}^T\Sigma_n^{-1}\underset{\sim}{Y}$

$= [H* - C(\mathbf{X}^T\Sigma_n^{-1}\mathbf{X})^{-1}\mathbf{X}^T\Sigma_n^{-1}]\underset{\sim}{Y} \triangleq D*\underset{\sim}{Y}$

i.e $H* - C(\mathbf{X}^T\Sigma_n^{-1}\mathbf{X})^{-1}\mathbf{X}^{-1}\Sigma_n^{-1} = D*$

or $H* = C(\mathbf{X}^T\Sigma_n^{-1}\mathbf{X})^{-1}\mathbf{X}^T\Sigma_n^{-1} + D*$ ………………………………………⑮

we have:

$E(\hat{\beta}*) = E(H*Y) = E[H*(\mathbf{X}\beta + \underset{\sim}{\epsilon})]$

$= H*\mathbf{X}\beta + H*E(\underset{\sim}{\epsilon}) = H*\mathbf{X}\beta$ ………………………………………⑯

$= [C(\mathbf{X}^T\Sigma_n^{-1}\mathbf{X})^{-1}\mathbf{X}^T\Sigma_n^{-1}D*]\mathbf{X}\beta$

$= C\beta + D*\mathbf{X}\beta$ ………………………………………⑰

compare ⑰ to ⑫ and ⑯

we must have

$\begin{cases} D*\mathbf{X} = \underset{\sim}{0} \\ H*\mathbf{X} = \underset{\sim}{C} \end{cases}$ ………………………………………⑱

By the similarly derivation, replace D.H. by D*, H* respectively,we may obtain

$MSE(\hat{\beta}*) = C(\mathbf{X}^T\Sigma_n^{-1}\mathbf{X})^{-1}C^T + D*\Sigma_n D*^T + \underset{\sim}{B}\underset{\sim}{B}^T$

$= [C(\mathbf{X}^T\Sigma_n^{-1}\mathbf{X})^{-1}C^T + \underset{\sim}{B}\underset{\sim}{B}^T] + D*\Sigma_n D*^T$ ………………………………………⑲

Since $D*\Sigma_n D*^T$ is non-negative definite

$$\text{MSE}(\hat{\beta}^*) \geq \text{MSE}(C\hat{\beta}) \cdots\cdots ⑳$$

from ⑬ and ⑳ we have

$$\text{MSE}(\hat{\beta}^*) = \text{MSE}(C\hat{\beta})$$

i.e $\text{MSE}(\hat{\beta}^*) = C(\mathbf{X}^{-1}\Sigma_n^{-1}\mathbf{X})^{-1}C + \underset{\sim}{B}\underset{\sim}{B}^T$ ……… ㉑

compare ㉑ to ⑲ we get

$$D^*\Sigma_n D^{*T} = 0 \ \text{ but } \Sigma_n \neq 0$$

$$\text{i.e } D^* = 0$$

Hence $\hat{\beta}^* - C\hat{\beta} = D^*\underset{\sim}{Y} = \underset{\sim}{0}$

Therefore $\hat{\beta}^* = C\hat{\beta}$

It completes the proof Q.E.D.

4. Conclusion:

In general linear model, always we have a hope of finding an estimator at least as good as BLUE in some specified criterion, such as MSE risk funtion matrix, we may consider whether there exists any improved biased estimator. The theorem mentioned above may be helpful for performing this work.

References:

1. Anderson, T.W.(1984) An inrtoduction to multivariate statistical analysis 2nd ed, P.86-96.
2. Graybill, F.A(1976). Theory and application of the linear model P.171-225.

中文摘要

本文提出了線性模式BLUE估計式之一擴張定理，將偏誤爲0（亦即不偏）之情況推廣至偏誤爲任意值之一般情況。