L04 BLUE and BLUP

1. Best linear unbiased estimators (BLUEs)

   (1) Setting

   Model: $Y = X\beta + \epsilon$, $\epsilon \sim (0, \sigma^2\Sigma)$, i.e., $Y \sim \left(X\beta, \sigma^2 I\right)$

   Estimable $H\beta$: $H = LX$ for some $L \Longleftrightarrow \text{LUE}(H\beta) \neq \emptyset$.

   Estimators for β: With $U = \Sigma^{-1}$, $\text{GLSE}_{\Sigma^{-1}}(\beta) = \left(\Sigma^{-1/2}X\right)^{+}\Sigma^{-1/2}Y + \mathcal{N}(X)$

   Minimum norm GLSE: $\widehat{\beta} = \left(\Sigma^{-1/2}X\right)^{+}\Sigma^{-1/2}Y$

   LUE($H\beta$): $\text{LUE}(H\beta) = H\left(\Sigma^{-1/2}X\right)^{+}\Sigma^{-1/2}Y + \mathcal{N}(I_q, X)Y$.

   (2) Task

   We need to select $\widehat{\eta} \in \text{LUE}(H\beta)$ such that $r(\widehat{\eta}, H\beta) \leq r(\widetilde{\eta}, H\beta)$ for all $\widetilde{\eta} \in \text{LUE}(H\beta)$. This $\widehat{\eta}$ is the best linear unbiased estimator for $H\beta$ by the specified criterion and is denoted as the BLUE for $H\beta$. Note that $r(\widetilde{\eta}, H\beta) = \text{Cov}(\widetilde{\eta})$.

   (3) BLUE

   $H\widehat{\beta} = H\left(\Sigma^{-1/2}X\right)^{+}\Sigma^{-1/2}Y$ is the BLUE for $H\beta$

   Proof. $\widehat{\eta} = H\widehat{\beta} = L_0Y \in \text{LUE}(H\beta)$ where $L_0 = H\left(\Sigma^{-1/2}X\right)^{+}\Sigma^{-1/2}$.

   $\widetilde{\eta} = LY \in \text{LUE}(H\beta)$ where $L = L_0 + T$ with $T \in \mathcal{N}(I_q, X)$.

   $$\begin{aligned} r(\widetilde{\eta}, H\beta) &= \text{Cov}(LY) = \text{Cov}((L_0 + T)Y) \\ &= \text{Cov}(L_0Y) + \text{Cov}(TY) + \text{Cov}(L_0Y, TY) + \text{Cov}(TY, L_0Y) \end{aligned}$$

   $$\begin{aligned} \text{But } \text{Cov}(L_0Y, TY) &= \sigma^2 L_0\Sigma T' = \sigma^2 H\left(\Sigma^{-1/2}X\right)^{+}\Sigma^{-1/2}\Sigma T' \\ &= \sigma^2 H\left(\Sigma^{-1/2}X\right)^{+}\left(\Sigma^{-1/2}X\right)\left(\Sigma^{-1/2}X\right)^{+}\Sigma^{1/2}T' \\ &= \sigma^2 H\left(\Sigma^{-1/2}X\right)^{+}\left[\left(\Sigma^{-1/2}X\right)^{+}\right]' X'\Sigma^{-1/2}\Sigma^{1/2}T' \\ &= \sigma^2 H\left(\Sigma^{-1/2}X\right)^{+}\left[\left(\Sigma^{-1/2}X\right)^{+}\right]'(I_qTX)' = 0 \end{aligned}$$

   and $\text{Cov}(TY, L_0Y) = [\text{Cov}(L_0Y, TY)]'$. So

   $r(\widetilde{\eta}, H\beta) = r(\widehat{\eta}, H\beta) + \text{Cov}(TY) + 0 + 0 \geq r(\widehat{\eta}, H\beta)$. Hence $\widehat{\eta} = H\widehat{\beta}$ is a BLUE.

   Ex1: In model $Y \sim \left(X\beta, \sigma^2 I_n\right)$ the estimable $H\beta$ has BLUE $H\widehat{\beta} = HX^{+}Y$. When X has full column rank, $H\widehat{\beta} = H(X'X)^{-1}X'Y$.

   (4) $\widehat{Y}$ and other statistics

   $E(Y) = X\beta$ has BLUE $\widehat{Y} = X\widehat{\beta} = X\left(\Sigma^{-1/2}X\right)\Sigma^{-1/2}Y$.

   The error vector $\epsilon = Y - X\beta$ is predicted by the residual vector

   $$\widehat{e} = Y - X\widehat{\beta} = Y - \widehat{Y} = \Sigma^{1/2}\left[I - \left(\Sigma^{-1/2}X\right)\left(\Sigma^{-1/2}X\right)^{+}\right]\left(\Sigma^{-1/2}Y\right)$$

   The minimized $\|Y - X\beta\|^2_{V^{-1}}$ is

   $$\text{SSE} = \|Y - X\widehat{\beta}\|^2_{\Sigma^{-1}} = \left(\Sigma^{-1/2}Y\right)'\left[I - \left(\Sigma^{-1/2}X\right)\left(\Sigma^{-1/2}X\right)^{+}\right]\left(\Sigma^{-1/2}Y\right).$$

   (4) An UE for σ^2

   $Y \sim (X\beta, \sigma^2\Sigma) \Longrightarrow \Sigma^{-1/2}Y \sim \left(\Sigma^{-1/2}X\beta, \sigma^2 I_n\right)$,

   $$\begin{aligned} E(\text{SSE}) &= \left(\Sigma^{-1/2}X\beta\right)\left[I - \left(\Sigma^{-1/2}X\right)\left(\Sigma^{-1/2}X\right)^{+}\right]\left(\Sigma^{-1/2}X\beta\right) \\ &\quad + \text{tr}\left\{\left[I - \left(\Sigma^{-1/2}X\right)\left(\Sigma^{-1/2}X\right)^{+}\right]\sigma^2 I_n\right\} = 0 + \sigma^2(n-r) = \sigma^2(n-r) \end{aligned}$$

   where $r = \text{rank}(\Sigma^{-1/2}X) = \text{rank}(X)$.

   Let $\text{MSE} = \frac{\text{SSE}}{n-r}$. Then MSE is an UE for σ^2.

2. BLUP for Y_f

(1) A special case

If Y and Y_f are independent, and $E(Y_f) = H\beta$, then

$$L_0Y \text{ is a BLUP for } Y_f \iff L_0Y \text{ is a BLUE for } H\beta$$

Proof. $$L_0Y \text{ is a BLUP for } Y_f \iff \begin{cases} L_0Y \in \text{LUP}(Y_f). \forall\, LY \in \text{LUP}(Y_f) \\ r(L_0Y,\, Y_f) \le r(LY,\, Y_f) \end{cases}$$

$$\iff \begin{cases} L_0Y \in \text{LUE}(H\beta). \forall\, LY \in \text{LUE}(H\beta) \\ \text{Cov}(L_0Y - Y_f) \le \text{Cov}(LY - Y_f) \end{cases}$$

$$\iff \begin{cases} L_0Y \in \text{LUE}(H\beta). \forall\, LY \in \text{LUE}(H\beta) \\ \text{Cov}(L_0Y) \le \text{Cov}(LY) \end{cases}$$

$$\iff L_0Y \text{ is a BLUE for } H\beta.$$

(2) A closed form

Suppose Y and Y_f are independent, Y_f is predictable and $E(Y_f) = H\beta$. Then

$$H\left(\Sigma^{-1/2}X\right)^{+}\Sigma^{-1/2}Y \text{ is a BLUP for } Y_f.$$

3. A more general setting

(1) Model

Iff-conditions for BLUE and BLUP can be derived in more genera setting

$$\begin{pmatrix} Y \\ Y_f \end{pmatrix} \sim \left(\begin{pmatrix} X \\ H \end{pmatrix}\beta,\, \sigma^2\begin{pmatrix} \Sigma & C \\ C' & V \end{pmatrix}\right) \text{ where } \Sigma \ge 0$$

(2) Preparation I

Suppose λ_i, $i = 1, .., n$ are the eigenvalues of $\Sigma \ge 0$, and $\lambda_{max} = \max(\lambda_1, ..., \lambda_n) > 0$. Let $0 < c < \frac{2}{\lambda_{max}}$. Then $c^2\Sigma - 2cI_n < 0$.

Proof. By EVD $\Sigma = P\Lambda P'$, $c^2\Sigma - 2cI = P(c^2\Lambda - 2cI)P'$ is the EVD for $c^2\Sigma - 2cI$ with eigenvalues $c^2\lambda_i - 2c$, $i = 1, ..., n$. We show that $c^2\lambda_i - 2c < 0$ for all i.

If $\lambda_i = 0$, then $c^2\lambda_i - 2c = -2c < 0$.

If $\lambda_i > 0$, then $c^2\lambda_i - 2c \le c^2\lambda_{max} - 2c = c(c\lambda_{max} - 2) < c\left(\frac{2}{\lambda_{max}}\lambda_{max} - 2\right) = 0$.

Thus $c^2\Sigma - 2c < 0$

(3) Preparation II

(i) $A \le 0$ and $A \ge 0 \iff A = 0$ (ii) $AA' = 0 \iff A = 0$

(iii) $A < 0$ and $TAT' \ge 0 \Longrightarrow T = 0$

Proof. (i) Only show $\Rightarrow$: First, $A \ge 0 \Longrightarrow A^{1/2}$ exists.

$A \le 0$ and $A \ge 0 \Longrightarrow x'Ax \le 0$ and $x'Ax \ge 0$ for all $x \Longrightarrow x'Ax = 0$ for all x.

So $0 = x'Ax = \left(A^{1/2}x\right)' A^{1/2}x = \|A^{1/2}x\|^2 \Longrightarrow A^{1/2}x = 0$ for all x.

Hence $A^{1/2} = 0 \Longrightarrow A = 0$.

(ii) Only show $\Rightarrow$: $\langle A',\, A'\rangle = \text{tr}(AA') = 0 \Longrightarrow A' = 0 \Longrightarrow A = 0$.

(iii) Only show $\Rightarrow$: $A < 0 \Longrightarrow TAT' \le 0$. But $TAT' \ge 0$. So $TAT' = 0$.

Hence $T(-A)T' = 0$. But $A < 0 \Longrightarrow -A > 0 \Longrightarrow (-A)^{\pm 1/2}$ exist.

$0 = T(-A)T' = [T(-A)^{1/2}][T(-A)^{1/2}]' \Longrightarrow T(-A)^{1/2} = 0 \Longrightarrow T = 0$.

L05 Fundamental theorems for BLUE and BLUP

1. Model

    (1) Model
    $\begin{pmatrix} Y \\ Y_f \end{pmatrix} \sim \left(\begin{pmatrix} X \\ H \end{pmatrix} \beta,\, \sigma^2 \begin{pmatrix} \Sigma & C \\ C' & V \end{pmatrix} \right).$ Here $\Sigma \geq 0$. So Σ^{-1} may not be available.

    (2) Two results
    (i) For $0 \neq \Sigma \geq 0$, there exists $c > 0$ such that $c^2\Sigma - 2cI < 0$.
    (ii) $A < 0$ and $TAT' \geq 0 \Longrightarrow T = 0$.

    (3) $\mathrm{LUE}(H\beta) = \mathrm{LUP}(Y_f)$
    If $BY \in \mathrm{LUE}(H\beta)$, then

$$\mathrm{LUP}(Y_f) = \mathrm{LUE}(H\beta) = \{[B + Z(I - XX^+)]Y,\, Z \in R^{q\times n}\}.$$

Proof. If $BY \in \mathrm{LUE}(H\beta)$, then $\quad \mathrm{LUE}(H\beta) = BY + \mathcal{N}(I_q,\, X)Y$.
But $\mathcal{N}(I_q,\, X) = \mathcal{R}(I_q,\, I_n - XX^+) = \{Z(I - XX^+) : Z \in R^{q\times n}\}$.
So $\quad BY + \mathcal{N}(I_q,\, X)Y = \{[B + Z(I - XX^+)]Y,\, Z \in R^{q\times n}\}$.

2. BLUE

    (1) Theorem
    BY is a BLUE for $H\beta \Longleftrightarrow B\,(X,\, \Sigma(I - XX^+)) = (H,\, 0)$.

Proof. Let $T = B\Sigma(I - XX^+)$ such that

$$B\left(X,\, \Sigma(I - XX^+)\right) = (H,\, 0) \Longleftrightarrow BX = H \text{ and } T = 0.$$

$$BY \text{ is a BLUE for } H\beta \Longleftrightarrow \begin{cases} BY \in \mathrm{LUE}(H\beta) \\ r(BY,\, H\beta) \leq r(LY,\, H\beta) \text{ for all } LY \in \mathrm{LUE}(H\beta) \end{cases}$$

$$\Longleftrightarrow \begin{cases} BX = H \\ \mathrm{Cov}(BY) \leq \mathrm{Cov}(BY + Z(I - XX^+)Y) \text{ for all } Z \in R^{q\times n} \end{cases}$$

$$\Longleftrightarrow \begin{cases} \text{(i) } BX = H \\ \text{(ii) } TZ' + ZT' + Z(I - XX^+)\Sigma(I - XX^+)Z' \geq 0 \text{ for all } Z \end{cases}$$

We need to show that (i) and (ii) $\Longleftrightarrow BX = H$ and $T = 0$. But $BX = H$ is on the both sides. We show

$$\text{(ii)} \Longleftrightarrow T = 0.$$

$\Longleftarrow$: Under $T = 0$, the inequality in (ii) becomes

$$Z(I - XX^+)\Sigma(I - XX^+)Z' \geq 0 \text{ for all } Z$$

which is clearly true since $\Sigma \geq 0$.

$\Longrightarrow$: Select c as in 1 (2) (i) so that $\quad c^2\Sigma - 2cI < 0$.
The inequality in (ii) is true for all Z. Select $Z = -cT$. We have

$$-cTT' - cTT' + (-cT)\Sigma(-cT)' \geq 0,$$

i.e., $T(c^2\Sigma - 2cI)T' \geq 0$. It follows from 1 (2) (ii) that $T = 0$.

(2) An example

When $\Sigma > 0$, for estimable $H\beta$ such that $H = LX$, let $BY = H\left(\Sigma^{-1/2}X\right)^+\Sigma^{-1/2}Y$.

$$\begin{aligned} \text{Then } BX &= H\left(\Sigma^{-1/2}X\right)^+\Sigma^{-1/2}X = L\Sigma^{1/2}\left(\Sigma^{-1/2}X\right)\left(\Sigma^{-1/2}X\right)^+\left(\Sigma^{-1/2}X\right) \\ &= L\Sigma^{1/2}\Sigma^{-1/2}X = LX = H, \end{aligned}$$

$$\begin{aligned} B\Sigma(I - XX^+) &= H\left(\Sigma^{-1/2}X\right)^+\Sigma^{-1/2}\Sigma(I - XX^+) \\ &= H\left(\Sigma^{-1/2}X\right)^+\left[\left(\Sigma^{-1/2}X\right)^+\right]'\left(\Sigma^{-1/2}X\right)'\Sigma^{-1/2}\Sigma(I - XX^+) \\ &= H\left(\Sigma^{-1/2}X\right)^+\left[\left(\Sigma^{-1/2}X\right)^+\right]'X'(I - XX^+) = 0. \end{aligned}$$

So $B(X,\, \Sigma(I - XX^+)) = (H,\, 0)$ is true. Hence BY is a BLUE for $H\beta$.

3. BLUP

(1) Theorem

BY is a BLUP for $Y_f \Longleftrightarrow B(X,\, \Sigma(I - XX^+)) = (H,\, C'(I - XX^+))$.

Proof. Let $T = (B\Sigma - C')(I - XX^+)$ such that

$$B(X,\, \Sigma(I - XX^+)) = (H,\, C'(I - XX^+)) \Longleftrightarrow BX = H \text{ and } T = 0.$$

$$BY \text{ is a BLUP for } Y_f \Longleftrightarrow \begin{cases} BY \in \text{LUP}(Y_f) \\ r(BY,\, Y_f) \le r(LY,\, Y_f) \text{ for all } LY \in \text{LUP}(Y_f) \end{cases}$$

$$\Longleftrightarrow \begin{cases} BX = H \\ \text{Cov}(BY - Y_f) \le \text{Cov}(BY + Z(I - XX^+)Y - Y_f) \text{ for all } Z \in R^{q\times n} \end{cases}$$

$$\Longleftrightarrow \begin{cases} \text{(i) } BX = H \\ \text{(ii) } TZ' + ZT' + Z(I - XX^+)\Sigma(I - XX^+)Z' \ge 0 \text{ for all } Z \end{cases}$$

We need to show that (i) and (ii) $\Longleftrightarrow BX = H$ and $T = 0$. But $BX = H$ is on the both sides. We show that

$$\text{(ii)} \Longleftrightarrow T = 0.$$

$\Longleftarrow$**:** Under $T = 0$, the inequality in (ii) becomes

$$Z(I - XX^+)\Sigma(I - XX^+)Z' \ge 0 \text{ for all } Z$$

which is clearly true since $\Sigma \ge 0$.

$\Longrightarrow$**:** Select c as in 1 (2) (i) so that $\quad c^2\Sigma - 2cI < 0.$

In (ii) select $Z = -cT$ so that $\quad T(c^2\Sigma - 2cI)T' \ge 0.$

By 1 (2) (ii), $T = 0$.

(2) A comment

$B(X,\, \Sigma(I - XX^+)) = (H,\, C'(I - XX^+))$ and $B(X,\, \Sigma(I - XX^+)) = (H,\, 0)$ are two different equations. So for Y_f with $E(Y_f) = X\beta$ estimable, BLUP for Y_f and BLUE for $H\beta$ may or may not be equal.

But if Y and Y_f are independent, then $C = 0$. Thus the two equations become identical. In that case the BLUP for Y_f and the BLUE for $H\beta$ are equal.