L04: BLUE under linear restrictions

1. Linear model under linear restrictions

   (1) Linearly restricted model
   $y = X\beta + e$, $e \sim (0, \sigma^2 I_n)$, under $G\beta = 0$ is a linearly restricted model. The restriction

   $$G\beta = 0 \Longleftrightarrow \beta \in \mathcal{N}(G) = \mathcal{N}(G^+G) = \mathcal{R}(I - G^+G)$$

   confines β in a linear space of R^p. Thus the restriction is a linear restriction. We consider point estimation for $A\beta$ in this model.

   (2) One-way ANOVA
   Sample of one-way ANOVA $y_i = \mu_i + \epsilon_i$ can be written as $y = M\mu + e$ where the p columns of $M \in R^{n\times p}$ are values of p indicators for p treatments. So M has full column rank.
   Consequently $A\mu$ is estimable for all A. With minimum norm LSE for β, $\widehat{\beta} = M^+y$, AM^+y is BLUE for $A\beta$.
   Let $\mu_{.} = \frac{\sum_i \mu_i}{p}$ and $\alpha_i = \mu_i - \mu_{.}$. Then $\theta = (\mu_{.}, \alpha_1, .., \alpha_p)' = A\mu$. So BLUE of θ can be obtained. However θ can be modeled in $y = X\theta + e$, $e \sim (0, \sigma^2 I_n)$, under $\alpha_1 + \cdots + \alpha_p = 0$. We need a way to estimator θ in this restricted model.

   (a)](3)] Two-way ANOVA with interactions
   Sample of two-way ANOVA with interactions $y_{ij} = \mu_{ij} + \epsilon_{ij}$ can be written as $y = M\mu + e$ where the ab columns of $M \in R^{n\times ab}$ are values of ab indicators for ab treatments. So M has full column rank.
   Consequently $A\mu$ is estimable for all A. With minimum norm LSE for β, $\widehat{\beta} = M^+y$, AM^+y is BLUE for $A\mu$.
   Let $\mu_{..}$ be the average of μ_{ij}, $\mu_{i.}$ be the average of $\mu_{i1}, .., \mu_{ib}$, $\mu_{.j}$ be the average of $\mu_{1j}, .., \mu_{aj}$, $\alpha_i = \mu_{i.} - \mu_{..}$, $\beta_j = \mu_{.j} - \mu_{..}$ and $(\alpha)\beta_{ij} = \mu_{ij} - \mu_{i.} - \mu_{.j} + \mu_{..}$. Then

   $$\theta = (\mu_{..}, \alpha_1, .., \alpha_a, \beta_1, .., \beta_b, (\alpha\beta)_{11}, .., (\alpha\beta)_{ab})' = A\mu.$$

   So BLUE of θ can be obtained. However θ can be modeled in $y = X\theta + e$, under $\sum_i \alpha_i = \sum_i (\alpha\beta)_{ij} = 0$ and $\sum_j \beta_j = \sum_j (\alpha\beta)_{ij} = 0$. How can we estimate θ in this restricted model?

2. Concepts
   Consider model $y = X\beta + e$, $e \sim (0, \sigma^2 I_n)$, under $G\beta = 0$.

   (1) LSE of β under $G\beta = 0$.

   $$\begin{aligned} & \widehat{\beta} \text{ is a LSE for } \beta \text{ under } G\beta = 0 \\ \overset{def}{\Longleftrightarrow}\ & \widehat{\beta} \in \mathcal{N}(G) \text{ and } \|y - X\widehat{\beta}\|^2 \le \|y - X\beta\|^2 \text{ for all } \beta \in \mathcal{N}(G) \\ \Longleftrightarrow\ & \cdots \Longleftrightarrow \widehat{\beta} \in [X(I - G^+G)]^+ y + \mathcal{N}(X) \cap \mathcal{N}(G). \end{aligned}$$

   Remarks: (i) $A' = A = A^2 \Longrightarrow A^+ = A$ and $AA^+ = A = A^+A$.
   Thus $(I - G^+G)^+(I - G^+G) = I - G^+G$.
   (ii) $(AB)^+ = B^+B(AB)^+$ because with $H = B^+B(AB)^+$, $(AB)H(AB) = AB$; $H(AB)H = H$; $(AB)H$ is symmetric; and $H(AB)$ is symmetric.
   (iii) By (i) and (ii) $[X(I - G^+G)]^+ = (I - G^+G)[X(I - G^+G)]^+$.

(2) LUE under $G\beta = 0$

$$Ly \text{ is LUE for } A\beta \text{ under } G\beta = 0 \overset{def}{\iff} E(Ly) = A\beta \text{ for all } \beta \in \mathcal{N}(G)$$
$$\iff \cdots \iff A(I - G^+G) = LX(I - G^+G)$$

(3) Estimable parameter under $G\beta = 0$

$$A\beta \text{ is estimable under } G\beta = 0 \overset{def}{\iff} \exists \text{ a LUE } Ly \text{ for } A\beta \text{ under } G\beta = 0$$
$$\iff \exists L \text{ such that } A(I - G^+G) = LX(I - G^+G)$$

(4) BLUE under $G\beta = 0$

By is BLUE for $A\beta$ under $G\beta = 0$

$\overset{def}{\iff}$ By is a LUE for $A\beta$ and
Cov$(By) \leq$ Cov(Ly) for all LUE Ly for $A\beta$ under $G\beta = 0$

$\iff$ $A(I - G^+G) = BX(I - G^+G)$ and Cov$(By) \leq$ Cov(Ly) for all L such that $A(I - G^+G) = LX(I - G^+G)$

3. BLUE under $G\beta = 0$

(1) A Theorem
Suppose $A\beta$ is estimable under $G\beta = 0$.
With LSE $\widehat{\beta} = [X(I - G^+G)]^+y$ for β under $G\beta = 0$,
$A\widehat{\beta}$ is BLUE for $A\beta$ under $G\beta = 0$.

(2) Proof
First, $A\beta$ is estimable under $G\beta = 0$. So there is a LUE Ly for $A\beta$ under $G\beta = 0$.
By (2) of 2,

$$A(I - G^+G) = LX(I - G^+G).$$

This Ly will be a representative of all LUEs for $A\beta$ under $G\beta = 0$.
Secondly we show $A\widehat{\beta} = A[X(I - G^+G)]^+y = By$ is also a LUE for $A\beta$ under $G\beta = 0$.
By (2) of 2 we show $A(I - G^+G) = BX(I - G^+G)$.
Note that by the definition of $A\widehat{\beta}$, $B = A[X(I - G^+G)]^+$.
By Remark (iii) in (2) of 1, $B = A(I - G^+G)[X(I - G^+G)]^+$.
By $A(I - G^+G) = LX(I - G^+G)$, $B = LHH^+$ where $H = X(I - G^+G)$.
By $A(I - G^+G) = LX(I - G^+G)$ again,

$$BX(I - G^+G) = BH = LHH^+H = LH = LX(I - G^+G) = A(I - G^+G).$$

Finally, we show Cov$(Ly) -$ Cov$(By) \geq 0$.

$$\begin{aligned}
\text{Cov}(Ly) - \text{Cov}(By) &= \sigma^2 LL' - \sigma^2 BB' \\
&= \sigma^2 LL' - \sigma^2 LHH^+L' = \sigma^2 L(I - HH^+)L' \\
&= [\sigma L(I - HH^+)][\sigma L(I - HH^+)]' \geq 0
\end{aligned}$$

L05: Generalized least square estimators

1. Generalized least square estimators

Consider linear model $y = X\beta + e$ with $E(e) = 0$.

(1) A general metric system

With positive definite $D \in R^{n\times n}$ for $x,\, y \in R^n$ $\langle x,\, y\rangle_D = y'Dx$ is an inner product with norm $\|x\|_D = \sqrt{\langle x,\, x\rangle_D} = \sqrt{x'Dx}$.
When $D = I$, $\langle x,\, y\rangle = y'x$ is popular Frobenius inner product with induced Frobenius norm $\|x\| = \sqrt{\langle x,\, x\rangle} = \sqrt{x^x}$. Thus $\langle \cdot, \cdot\angle_D$ is a general inner product.

(2) Relations

$\langle x,\, y\rangle_D$ and $\|x\|_D^2$ can be expressed via Frobenius inner product and Frobenius norm.
$\langle x,\, y\rangle_D = y'Dx = (D^{1/2}y)'(D^{1/2}x) = \langle D^{1/2}x,\, D^{1/2}y\rangle$ and $\|x\|_D^2 = \|D^{1/2}x\|^2$.

(3) Generalized least square estimators (GLSE)

$$\begin{aligned} & \widehat{\beta} \text{ is a generalized least square estimaotor (GLSE) with respect to } \langle\cdot,\cdot\rangle_D \\ \overset{def}{\Longleftrightarrow}\quad & \|y - X\widehat{\beta}\|_D^2 \le \|y - X\beta\|_D^2 \text{ for all } \beta. \end{aligned}$$

(4) The collection of all GLSEs

Let $\text{GLSE}_D(\beta)$ be the collection of all GLSEs for β with respect to $\langle\cdot,\cdot\rangle_D$.
Then $\text{GLSE}_D(\beta) = (D^{1/2}X)^+(D^{1/2}y) + \mathcal{N}(X)$.

Proof
$$\begin{aligned} \widehat{\beta} \in \text{GLSE}_D(\beta) \quad \overset{def}{\Longleftrightarrow}\quad & \|y - X\widehat{\beta}\|_D^2 \le \|y - X\beta\|_D^2 \text{ for all } \beta \\ \Longleftrightarrow\quad & \|D^{1/2}y - D^{1/2}X\widehat{\beta}\|^2 \le \|D^{1/2}y - D^{1/2}X\beta\|^2 \text{ for all } \beta \\ \Longleftrightarrow\quad & D^{1/2}X\widehat{\beta} = \pi(D^{1/2}y \mid \mathcal{R}(D^{1/2}X)) \\ \Longleftrightarrow\quad & D^{1/2}X\widehat{\beta} = (D^{1/2}X)(D^{1/2}X)^+(D^{1/2}y) \\ \Longleftrightarrow\quad & D^{1/2}X\left[\widehat{\beta} - (D^{1/2}X)^+(D^{1/2}y)\right] = 0 \\ \Longleftrightarrow\quad & \widehat{\beta} - (D^{1/2}X)^+(D^{1/2}y) \in \mathcal{N}(D^{-1/2}X) \\ \Longleftrightarrow\quad & \widehat{\beta} \in (D^{1/2}X)^+(D^{1/2}y) + \mathcal{N}(D^{1/2}X) \\ \Longleftrightarrow\quad & \widehat{\beta} \in (D^{1/2}X)^+(D^{1/2}y) + \mathcal{N}(X). \end{aligned}$$

Comment: $\text{GLSE}_I(\beta) = X^+y + \mathcal{N}(X) = \text{LSE}(\beta)$ also denoted as $\text{OLSE}(\beta)$ for ordinary LSE.

2. A subset of LUE($A\beta$).

Consider $y = X\beta + e$ with $E(e) = 0$.

(1) Recall: $A\beta$ is estimable $\iff A = LX$ for some L.
Ly is a LUE for $A\beta \iff A = LX$.

(2) A Lemma
If $A\beta$ is estimable, then (i) $A \cdot \mathrm{GLSE}_D(\beta) = \left\{A(D^{1/2}X)^+(D^{1/2}y)\right\}$.
(ii) $A(D^{1/2}X)^+(D^{1/2}y) \in \mathrm{LUE}(A\beta)$.

Proof $A\beta$ is estimable $\iff$ there esists L such that $A = LX$.

(i) By $A = LX$,
$$\begin{aligned} & A \cdot \mathrm{GLSE}_D(\beta) = A\left\{(D^{1/2}X)^+(D^{1/2}y) + \mathcal{N}(X)\right\} \\ = \; & A(D^{1/2}X)^+(D^{1/2}y) + A\mathcal{N}(X) = A(D^{1/2}X)^+(D^{1/2}y) + LX\mathcal{N}(X) \\ = \; & A(D^{1/2}X)^+(D^{1/2}y) + \{0\} = \left\{A(D^{1/2}X)^+(D^{1/2}y)\right\}. \end{aligned}$$

(ii) By $A = LX$, $A(D^{1/2}X)^+(D^{1/2}y) = By$ with
$$B = A(D^{1/2}X)^+D^{1/2} = LX(D^{1/2}X)^+D^{1/2} = LD^{-1/2}(D^{1/2}X)(D^{1/2}X)^+D^{1/2}.$$
For $By \in \mathrm{LUE}(A\beta)$ we show $A = BX$.
$BX = LD^{-1/2}(D^{1/2}X)(D^{1/2}X)^+(D^{1/2}X) = LD^{-1/2}(D^{1/2}X) = LX = A$.

(3) A subset of LUE($A\beta$)
By (2) for estimable $A\beta$,
$$\{A(D^{1/2}X)^+(D^{1/2}y : \; D \in R^{n\times n} \text{ is positive definite}\} \subset \mathrm{LUE}(A\beta).$$

3. BLUE
Consider $y = X\beta + e$, $e \sim (0, \, \sigma^2 V)$.

(1) Definition for BLUE
For estimable $A\beta$, By is BLUE if $By \in \mathrm{LUE}(A\beta)$ and $\mathrm{Cov}(By) \le \mathrm{Cov}(Ly)$ for all $Ly \in \mathrm{LUE}(A\beta)$.

(2) Theorem
For estimable $A\beta$ in model $y = X\beta + e$, $e \sim (0, \, \sigma^2 V)$, $A(V^{-1/2}X)^+(V^{-1/2}y)$ is BLUE.

Proof By (3) of 2, with $D = V^{-1}$, $A(V^{-1/2}X)^+(V^{-1/2}y) \in \mathrm{LUE}(A\beta)$.
Suppose Ly is also a LUE for $A\beta$. Then $A = LX$. We need to show
$$\mathrm{Cov}(A(V^{-1/2}X)^+(V^{-1/2}y)) \le \mathrm{Cov}(Ly).$$
$\mathrm{Cov}(Ly) = \sigma^2 LVL' = \sigma^2(LV^{1/2})(LV^{1/2})'$.
Write $A(V^{-1/2}X)^+(V^{-1/2}y) = By$. By $A = LX$,
$$\begin{aligned} B &= A(V^{-1/2}X)^+V^{-1/2} = LX(V^{-1/2}X)^+V^{-1/2} \\ &= LV^{1/2}(V^{-1/2}X)(V^{-1/2}X)^+V^{-1/2} = (LV^{1/2})HH^+ \end{aligned}$$
where $H = V^{-1/2}X$. So $\mathrm{Cov}(By) = \sigma^2 BB' = \sigma^2(LV^{1/2})HH^+(LV^{1/2})'$. Thus
$$\begin{aligned} \mathrm{Cov}(Ly) = \mathrm{Cov}(By) &= \sigma^2(LV^{1/2}(I - HH^+)(LV^{1/2})' \\ &= [\sigma(LV^{1/2})(I - HH^+)][\sigma(LV^{1/2})(I - HH^+)]' \ge 0. \end{aligned}$$