L14: Fundamental equation for BLUE

1. LUE and LUP

   (1) Setting

   For $y \in R^n$ and $y_* \in R^m$, $\begin{pmatrix} y \\ y_* \end{pmatrix} \sim N\left(\begin{pmatrix} X \\ X_* \end{pmatrix}\beta,\, \sigma^2 \begin{pmatrix} \Sigma & C \\ C' & V \end{pmatrix}\right)$
   where $y \sim N(X\beta,\, \sigma^2\Sigma)$ is a linear model with unknown parameters $\beta \in R^p$ and $\sigma^2 > 0$. $y_* \sim N(X_*\beta,\, \sigma^2 V)$ is unobserved random vector. We are interested in estimating $X_*\beta$ and predicting y_*.

   (2) Linear unbiased estimator for $X_*\beta$

   $$Ly \text{ is a linear unbiased estimator (LUE) for } X_*\beta \overset{def}{\iff} E(Ly - X_*\beta) = 0 \ \forall\ \beta$$
   $$\iff LX\beta = X_*\beta \text{ for all } \beta \iff LX = X_*$$
   So the collection of all LUEs for $X_*\beta$ is $\quad \text{LUE}(X_*\beta) = \{Ly : \, LX = X_*\}$.

   (3) Linear unbiased predictor for y_*

   $$Ly \text{ is a linear unbiased predictor (LUP) for } y_* \overset{def}{\iff} E(Ly - y_*) = 0 \ \forall\ \beta$$
   $$\iff LX\beta = X_*\beta \text{ for all } \beta \iff LX = X_*$$
   So the collection of all LUPs for y_* is $\quad \text{LUP}(y_*) = \{Ly : \, LX = X_*\}$.
   Comment: $\text{LUP}(y_*) = \text{LUE}(E(y_*))$, i.e., $\text{LUP}(y_*) = \text{LUE}(X_*\beta)$.

2. BLUE

   (1) Definition

   $$By \text{ is a best linear unbiased estimator (BLUE) for } X_*\beta$$
   $$\overset{def}{\iff} By \in \text{LUE}(X_*\beta) \text{ and } \text{Cov}(By) \le \text{Cov}(Ly) \text{ for all } Ly \in LUE(X_*\beta).$$

   (2) A sufficient and necessary condition

   $$By \text{ is a BLUE for } X_*\beta$$
   $$\iff \begin{cases} BX = X_* \text{ and with } M = B\Sigma(I - XX^+) \\ MH' + HM' + H'(I - XX^+)\Sigma(I - XX^+)H' \ge 0 \text{ for all } H \in R^{m\times n}. \end{cases}$$

   Proof $LX = BX \iff (L - B)X = 0 \overset{*}{\iff} L - B \in \{H(I - XX^+) : \, H \in R^{m\times n}\}$.
   $\overset{*}{\Rightarrow}$: $(L - B)X = 0 \Longrightarrow L - B = (L - B)(I - XX^+) \in \{H(I - XX^+) : \, H \in R^{m\times n}\}$
   $\overset{*}{\Leftarrow}$: $(L - B)X = H(I - XX^+)X = H0 = 0$.

   $$By \text{ is BLUE for } X_*\beta \iff \begin{cases} BX = X_* \\ \text{Cov}(Ly) - \text{Cov}(By) \ge 0 \ \forall\, LX = BX \end{cases}$$
   $$\iff \begin{cases} BX = X_* \\ L\Sigma L' - B\Sigma B' \ge 0 \ \forall\ L \in \{B + H(I - X^+X) : \, H \in R^{m\times n}\} \end{cases}.$$
   $$\iff \begin{cases} BX = X_* \\ MH' + HM' + H'(I - XX^+)\Sigma(I - XX^+)H' \ge 0 \text{ for all } H \in R^{m\times n}. \end{cases}$$

   (3) Fundamental equation for BLUE

   By is a BLUE for $X_*\beta \iff B[X,\, \Sigma(I - XX^+)] = (X_*,\, 0)$

   Proof Let $M = B\Sigma(I - XX^+)$.

   $\Leftarrow$: By the equation, $BX = X_*$ and $M = 0$. So
   $MH' + HM' + H(I - XX^+)\Sigma(I - XX^+)H' = H(I - XX^+)\Sigma(I - XX^+)H' \ge 0$
   for all $H \in R^{m\times n}$. Hence By is a BLUE for $X_*\beta$.

$\Leftarrow$: By EVD $\Sigma = P\Lambda P$. Let $0 < t < \frac{2}{\max(\lambda_1,..,\lambda_n)}$. Then $t^2\Sigma - 2tI = P\Gamma P'$ where $\Gamma = \text{diag}(t(t\lambda_1 - 2), ..., t(t\lambda_n - 2)) < 0$. Now By is a BLUE for $X_*\beta$. Then $BX = X_*$ and $MH' + HM' + H(I - XX^+)\Sigma(I - XX^+)H' \geq 0$ for all $H \in R^{m\times n}$. Let $H = -tM$. Then $0 \leq M(t^2\Sigma - 2tI)M' = MP\Gamma P'M' \leq 0$ since $\Gamma < 0$. So $MP\Gamma P'M' = 0$. Thus $MP(-\Gamma)P'M' = 0$. Hence $MP(-\Gamma)^{1/2} = 0$. Therefore $M = 0$. Consequently $B[X, \Sigma(I - XX^+)] = (X_*, 0)$.

3. Two examples

(1) In linear model $y = X\beta + e$, $e \sim N(0, \sigma^2 I)$, $X_*\beta$ is estimable.
Show that X_*X^+y is BLUE for $X_*\beta$.

Proof $X_*\beta$ is estimable $\Longleftrightarrow X_* = LX$ for some L.
$X_*X^+y = By$ where $B = X_*X^+$.
With $\Sigma = I$, we show that $B[X, \Sigma(I - XX^+)] = B(X, I - XX^+) = (X_*, 0)$.

$BX = X_*X^+X = LXX^+X = LX = X_*$.
$B(I - XX^+) = X_*X^+(I - XX^+) = X_*(X^+ - X^+) = 0$.

So $B(X, I - XX^+) = (X_*, 0)$. Therefore $By = X_*X^+y$ is BLUE for $X_*\beta$.

(2) In linear model $y = X\beta + e$, $X \in R^{n\times p}$ has full column rank and $e \sim N(0, \sigma^2\Sigma)$.
Show that for all $X_* \in R^{m\times p}$, $X_*(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}y$ is BLUE for $X_*\beta$.

Proof $X_*(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}y = By$ where $B = X_*(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}$.
We show that the fundamental equation for By to be BLUE for $X_*\beta$ holds, i.e., $B[X, \Sigma(I - XX^+)] = (X_*, 0)$.

By direct computation, $BX = [X_*(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}]X = X_*$.

$$\begin{aligned} B\Sigma(I - XX^+) &= [X_*(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}]\Sigma(I - XX^+) \\ &= X_*(X'\Sigma^{-1}X)^{-1}X'(I - XX^+) = X_*(X'\Sigma^{-1}X)^{-1}(X' - X') \\ &= 0. \end{aligned}$$

So $B[X, \Sigma(I - XX^+)] = (X_*, 0)$. Hence By is BLUE for $X_*\beta$.

L15: Fundamental equation for BLUP

1. Predictors

   (1) Predictors
   y_* is an unknown random vector and γ is an unknown parameter vector.
   When γ is estimated by statistic $\widehat{\gamma}$, $\widehat{\gamma}$ is called an estimator for γ.
   When y_* is predicted by statistic $\widehat{y}_*$, $\widehat{y}_*$ is called a predictor for y_*.

   (2) Linear predictors
   $L(y) = \{Ly : L\}$ is the collection of all linear functions of $y \in R^n$, a random sample.
   $\widehat{\gamma}$ is a linear estimator for γ if $\widehat{\gamma}$ is an estimator for γ and $\widehat{\gamma} \in L(y)$.
   $\widehat{y}_*$ is a linear predictor for y_* if $\widehat{y}_*$ is a predictor for y_* and $\widehat{y}_* \in L(y)$.

   (3) Unbiased predictors
   $\widehat{\gamma}$ is an unbiased estimator (UE) for γ if $E(\widehat{\gamma} - \gamma) \equiv 0 \Longleftrightarrow E(\widehat{\gamma}) = \gamma$.
   $\widehat{y}_*$ is an unbiased predictor (UP) for y_* if $E(\widehat{y}_* - y_*) \equiv 0 \Longleftrightarrow E(\widehat{y}_*) = E(y_*)$.
   Let $\mathrm{UE}(\gamma)$ and $\mathrm{UP}(y_*)$ be the collections of all UEs for γ and all UPs for y_*.
   Then $\mathrm{UP}(y_*) = \mathrm{UE}(E(y_*))$.

   (4) Linear unbiased predictors
   $\widehat{\gamma}$ is a linear UE (LUE) for γ if $\widehat{\gamma} \in L(y) \cap \mathrm{UE}(\gamma) = \{Ly : L[E(y)] = \gamma\}$.
   $\widehat{y}_*$ is a linear UP (LUP) for y_* if $\widehat{y}_* \in L(y) \cap \mathrm{UP}(y_*) = \{Ly : L[E(y)] = E(y_*)\}$.
   Let $\mathrm{LUE}(\gamma)$ and $\mathrm{LUP}(y_*)$ be the collections of all LUEs of γ and all LUPs of $y*$.
   Then $\mathrm{LUP}(y_*) = \mathrm{LUE}(E(y_*))$.

2. Best predictors

   (1) Risks
   The estimator $\widehat{\gamma}$ for γ has matrix-valued risk $E[(\widehat{\gamma} - \gamma)(\widehat{\gamma} - \gamma)']$.
   The predictor $\widehat{y}_*$ for y_* has matrix-valued risk $E[(\widehat{y}_* - y_*)(\widehat{y}_* - y_*)']$.

   (2) Risks for UP
   For $\widehat{\gamma}$, an UE for γ, the matrix-valued risk is $\mathrm{Cov}(\widehat{\gamma})$.
   For $\widehat{y}_*$, an UP for y_*, the matrix-valued risk is $\mathrm{Cov}(\widehat{y}_* - y_*)$.

   (3) Best UP
   $\widehat{\gamma}$ is best UE for γ if $\widehat{\gamma} \in \mathrm{UE}(\gamma)$ and $\mathrm{Cov}(\widehat{\gamma}) \leq \mathrm{Cov}(\widetilde{\gamma})\ \forall \widetilde{\gamma} \in \mathrm{UE}(\gamma)$.
   $\widehat{y}_*$ is best UP for y_* if $\widehat{y}_* \in \mathrm{UP}(y_*)$ and $\mathrm{Cov}(\widehat{y}_* - y_*) \leq \mathrm{Cov}(\widetilde{y}_* - y_*)$, $\forall \widetilde{y}_* \in \mathrm{UP}(y_*)$.

   (4) Best LUP (BLUP)
   By is a BLUE for γ if $By \in \mathrm{UE}(\gamma)$ and $\mathrm{Cov}(By) \leq \mathrm{Cov}(Ly)$, $\forall Ly \in \mathrm{UE}(\gamma)$.
   By is a BLUP for y_* if $By \in \mathrm{UP}(y_*)$ and $\mathrm{Cov}(By - y_*) \leq \mathrm{Cov}(Ly - y_*)$, $\forall Ly \in \mathrm{UP}(y_*)$.

3. Fundamental equation for BLUP

   (1) Setting
   $$\begin{pmatrix} y \\ y_* \end{pmatrix} \sim N\left(\begin{pmatrix} X \\ X_* \end{pmatrix}\beta,\, \sigma^2 \begin{pmatrix} \Sigma & C \\ C' & V \end{pmatrix}\right).$$
   $y \sim N(X\beta, \sigma^2\Sigma)$ is a linear model. $y_* \sim N(X_*\beta, \sigma^2 V)$ is a unobserved random vector.
   For completeness we assume $\mathrm{Cov}(y, y_*) = C$.

   (2) Definition
   $$By \text{ is BLUP for } y_* \overset{def}{\Longleftrightarrow} \begin{cases} By \in \mathrm{LUP}(y_*) \\ \mathrm{Cov}(By - y_*) \leq \mathrm{Cov}(Ly - y_*)\ \forall\ Ly \in \mathrm{LUP}(y_*). \end{cases}$$

(3) A sufficient and necessary condition
Let $M = B\Sigma(I - XX^+) - C'(I - XX^+)$. Then

$$By \text{ is BLUP for } y_* \Longleftrightarrow \begin{cases} BX = X_* \\ HM' + MH' + H(I - XX^+)\Sigma(I - XX^+)H' \geq 0 \ \forall \ H \end{cases}$$

Proof Note that in the definition in (2),

$$By \in \text{LUP}(y_*) \Longleftrightarrow E(By - y_*) \equiv 0 \Longleftrightarrow BX - X_* = 0 \Longleftrightarrow BX = X_*.$$

Under the above condition,

$$Ly \in \text{LUP}(y_*) \Longleftrightarrow (L - B)X = 0 \Longleftrightarrow L - B \in \{H(I - XX^+) : H\}.$$

Thus for all $Ly \in \text{LUP}(y_*) \Longleftrightarrow L = B + H(I - XX^+)$ for all H.
Also

$$\begin{aligned}
& \text{Cov}(Ly - y_*) - \text{Cov}(By - y_*) \\
= \ & [\text{Cov}(Ly) - \text{Cov}(Ly, y_*) - \text{Cov}(y_*, Ly) + \text{Cov}(y_*)] \\
& \quad - [\text{Cov}(By) - \text{Cov}(By, y_*) - \text{Cov}(y_*, By) + \text{Cov}(y_*)] \\
= \ & [\text{Cov}(Ly) - \text{Cov}(By)] - \text{Cov}((L - B)y, y_*) - \text{Cov}(y_*, (L - B)y) \\
= \ & [B\Sigma(I - XX^+)H' + H(I - XX^+)\Sigma B' + H(I - XX^+)\Sigma(I - XX^+)H'] \\
& \quad - H(I - XX^+)C - C'(I - XX^+)H' \\
= \ & MH' + HM' + H(I - XX^+)\Sigma(I - XX^+)H'.
\end{aligned}$$

Thus

$$\begin{cases} By \in \text{LUP}(y_*) \\ \text{Cov}(By - y_*) \leq \text{Cov}(Ly - y_*) \, \forall \, Ly \in \text{LUP}(y_*) \end{cases}.$$

$$\Longleftrightarrow \begin{cases} BX = X_* \\ HM' + MH' + H(I - XX^+)\Sigma(I - XX^+)H' \geq 0 \, \forall \, H \end{cases}$$

(4) Fundamental equation for BLUP
By is BLUP for $y_* \Longleftrightarrow B[X, \Sigma(I - XX^+)] = [X_*, C'(I - XX^+)]$

Proof $\Leftarrow$**:** $B[X, \Sigma(I - XX^+)] = [X_*, C'(I - XX^+)] = 0$ implies that $BX = X_*$ and $M = B\Sigma(I - XX^+) - C'(I - XX^+) = 0$. So for all $H \in R^{m \times n}$,
$HM' + MH' + H(I - XX^+)\Sigma(I - XX^+)H' = H(I - XX^+)\Sigma(I - XX^+)H' \geq 0$.
Thus By the sufficient and necessary condition, By is BLUP for y_*.

$\Rightarrow$**:** When By is BLUP for y_*, by the sufficient and necessary conditions, $BX = X_*$ and $HM' + MH' + H(I - XX^+)\Sigma(I - XX^+)H' \geq 0$ for all H.
Let $H = -dM$ and by the EVD $\Sigma = P\Lambda P'$,

$$\begin{aligned}
& HM' + MH' + H(I - XX^+)\Sigma(I - XX^+)H' \\
& \qquad = M(d^2\Sigma - 2dI)M' = MP(\Gamma)P'M' \geq 0 \, \forall \, d
\end{aligned}$$

where $\Gamma = \text{diag}(d^2\lambda_1 - 2d, ..., d^2\lambda_n - 2d)$. But with $0 < d < \frac{2}{\max(\lambda_1, .., \lambda_n)}$,

$$\begin{aligned}
\Gamma < 0 \ & \Longrightarrow \ MP\Gamma P'M' \leq 0 \Longrightarrow MP\Gamma P'M' = 0 \\
& \Longrightarrow \ MP(-\Gamma)P'M' = 0 \Longrightarrow MP(-\Gamma)^{1/2} = 0 \Longrightarrow M = 0.
\end{aligned}$$

Thus $B[X, \Sigma(I - XX^+)] = [X_*, C'(I - XX^+)]$.