L02: Estimable parameter functions

1. Estimable parameter functions

   (1) Estimable parameter functions
   Recall: In linear model $y = X\beta + e$, $e \sim (0, \sigma^2 I_n)$

$$\begin{aligned} A\beta \text{ is estimable } &\overset{def}{\Longleftrightarrow} A\beta \text{ has a linear unbiased estimator} \Longleftrightarrow A = LX \text{ for some } L \\ &\Longleftrightarrow A \cdot \text{LSE}(\beta) \text{ contains a unique estimator} \\ &\Longrightarrow \text{The unique } A \cdot \text{LSE}(\beta) = AX^+y \text{ is the BLUE for } A\beta. \end{aligned}$$

   (2) $X\beta$
   $E(y) = X\beta$ is always estimable since $X = I_n X$. Thus $X \cdot \text{LSE}(\beta) = \{XX^+y\}$ and XX^+y is the BLUE for $X\beta$.
   $X\beta$ is the essential estimable parameter function since

$$A\beta \text{ is estimable} \Longleftrightarrow A\beta \text{ is a linear function of } X\beta$$

   $\Rightarrow$: If $A\beta$ is estimable, then $A = LX$ for some L. So $A\beta = LX\beta$ is a linear function of $X\beta$.

   $\Leftarrow$: If $A\beta$ is a linear function of $X = beta$, then $A\beta = LX\beta$ for some L and all β. So $A = LX$ for some L. Hence $A\beta$ is estimable.

   (3) β

$$\begin{aligned} \beta \text{ is estimable } &\Longleftrightarrow I_p = LX \text{ for some L} \Longleftrightarrow X \text{ has a left-inverse} \\ &\Longleftrightarrow X \text{ has full column rank} \Longleftrightarrow \mathbf{N}(X) = \{0\} \\ &\Longleftrightarrow \text{LSE}(\beta) = \{X^+y\} = \{(X'X)^{-1}X'y\}. \end{aligned}$$

   The above condition is the most restricted one. Under that condition $A\beta$ is estimable for all A since $A \cdot \text{LSE}(\beta)$ contains a unique estimator. This unique vector is the BLUE for $A\beta$.

   Ex1: For regression $y = X\beta + e$, $e \sim (0, \sigma^2 I_n)$ where X has full column rank. There is one and only one LSE for β, $\widehat{\beta} = X^+y = (X'X)^{-1}X'y$. All $A\beta$ are estimable with BLUE $AX^+y = A(X'X)^{-1}X'y$.

2. Estimable functions in One-way ANOVA

   (1) One-way ANOVA

   $y = M\mu + e$, $e \sim (0, \sigma^2 I_n)$ is one-way ANOVA model where $\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix}$ contains the mean response to p treatments due to p levels of a factor. $M = (m_{ij})_{n\times p}$ where

$$m_{ij} = \begin{cases} 1, & y_i \text{ is the response to the } j\text{th treatment with mean } \mu_j \\ 0, & \text{otherwise} \end{cases}$$

   $M = \begin{pmatrix} 1_{n_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1_{n_p} \end{pmatrix} \in R^{n\times p}$ for example.

(2) BLUE of μ
With response to all levels, M has full column rank and hence $\widehat{\mu} = M^+ y = (M'M)^{-1}M'y$ is BLUE for μ.

Note that $M'M = \text{diag}(n_1, .., n_p)$, $(M'M)^{-1} = \text{diag}(1/n_1, .., 1/n_p)$, $M'y = \begin{pmatrix} y_{1.} \\ \vdots \\ y_{p.} \end{pmatrix}$. Here $y_{i.}$ is the summation of all responses to the treatment i. Hence $\widehat{y} = \begin{pmatrix} \overline{y}_1 \\ \vdots \\ \overline{y}_p \end{pmatrix}$.

(3) BLUE of θ

Let $\mu_i = \mu_. + \alpha_i$ with $\alpha_1 + \cdots + \alpha_p = 0$. Then $\alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_p \end{pmatrix}$ is called the factor effects.

It can be shown that $\mu_. = \frac{\mu_1 + \cdots + \mu_p}{p}$ and $\alpha_i = \mu_i - \mu_.$. This process can be written as $\theta = \begin{pmatrix} \mu_. \\ \alpha \end{pmatrix} = A\mu$.

Then $\theta = A\mu$ is estimable with BLUE $\widehat{\theta} = A\widehat{\mu}$. Specifically $\widehat{\mu}_. = \frac{\overline{y}_1 + \cdots + \overline{y}_p}{p}$ and $\widehat{\alpha}_i = \overline{y}_i - \widehat{\mu}_.$.

3. Estimable functions in Two-way ANOVA with interactions

(1) Two-way ANOVA
With factor A of a levels and factor B of b levels, $y = M\mu + e$, $e \sim (0, \sigma^2 I_n)$ is two-way ANOVA model where $\mu \in R^{ab}$ with components μ_{ij}, the mean response to the treatment formed by the combination of ith level of A and jth level of B, $i = 1, , , .a$; $j = 1, ..., b$. Model design matrix $M = (m_{st})_{n \times ab}$ such that

$$m_{st} = \begin{cases} 1, & y_s \text{ is the response to the treatment with mean, the } t\text{the component is } \mu \\ 0, & \text{otherwise} \end{cases}$$

(2) BLUE of μ
With response to all treatments, M has full column rank. So $\widehat{\mu} = M^+ y = (M'M)^{-1}M'y$ is BLUE for μ. $\widehat{\mu}$ can be obtained by replacing μ_{ij} in μ by $\overline{y}_{ij}$.

(3) BLUE of θ
Let $\mu_{ij} = \mu_{..} + \alpha_i + \beta_j + (\alpha\beta)_{ij}$ with $\sum_i \alpha_i = \sum_i (\alpha\beta)_{ij} = 0$ and $\sum_j \beta_j = \sum_j (\alpha\beta)_{ij} = 0$. Then $\mu_{..} = \frac{\sum_i \sum_j \mu_{ij}}{ab}$. With $\mu_{i.} = \frac{\sum_j \mu_{ij}}{b}$ and $\mu_{.j} = \frac{\sum_i \mu_{ij}}{a}$, $\alpha_i = \mu_{i.} - \mu_{..}$, $\beta_j = \mu_{.j} - \mu_{..}$ and $(\alpha\beta)_{ij} = \mu_{ij} - \mu_{i.} - \mu_{.j} + \mu_{..}$. So with

$$\theta = (\mu_{..}, \alpha_1, .., \alpha_a, \beta_1, .., \beta_b, (\alpha\beta)_{11}, .., (\alpha\beta)_{ab})' \in R^{1+a+b+ab}$$

there exists A such that $\theta = A\mu$. This θ is estimable with BLUE $A\widehat{\mu}$.
Here $\widehat{\mu}_{..} = \frac{\sum_i \sum_j \overline{y}_{ij}}{ab}$. With $\widehat{\mu}_{i.} = \frac{\sum_j \overline{y}_{ij}}{b}$ and $\widehat{\mu}_{.j} = \frac{\sum_i \overline{y}_{ij}}{a}$, $\widehat{\alpha}_i = \widehat{\mu}_{i.} - \widehat{\mu}_{..}$, $\widehat{\beta}_j = \widehat{\mu}_{.j} - \widehat{\mu}_{..}$ and $\widehat{(\alpha\beta)}_{ij} = \widehat{\mu}_{ij} - \widehat{\mu}_{i.} - \widehat{\mu}_{.j} + \widehat{\mu}_{..}$.

L03: Conditional LSE

1. Sufficient and necessary conditions for estimability

   (1) Sufficient and necessary conditions for estimability
   In $y = X\beta + e$, $e \sim (0, \sigma^2 I_n)$ for the estimability of $A\beta$, besides $A = LX$ for some L, there are other sufficient and necessary conditions. Here we claim that
   the followings are equivalent.

   (i) $A\beta$ is estimable

   (ii) $\mathcal{R}(A') \subset \mathcal{R}(X')$

   (iii) $\mathcal{R}[(A',\, X')] = \mathcal{R}(X')$

   (iv) $\text{rank}\left[\begin{pmatrix} A \\ X \end{pmatrix}\right] = \text{rank}(X)$

   (2) **Proofs**

   (i)⇒(ii): $A\beta$ is estimable$\Longrightarrow A = LX$ for some $L \Longrightarrow A' = X'L'$ for some L.
   So $\mathcal{R}(A') \subset \mathcal{R}(X')$.

   (ii)⇒(iii): First, $\mathbf{R}(X') \subset \mathcal{R}[(A',\, X')]$.
   Under (ii) $\mathcal{R}(A') \subset \mathcal{R}(X')$, $\mathbf{R}[(A',\, X')] = \mathcal{R}(A') + \mathcal{R}(X') \subset \mathcal{R}(X')$.
   So $\mathcal{R}[(A',\, X')] = \mathcal{R}(X')$.

   (iii)⇒(iv): If $\mathcal{R}[(A',\, X')] = \mathcal{R}(X')$, then $\dim[\mathcal{R}[(A',\, X')]] = \dim[\mathcal{R}(X')]$.
   So $\text{rank}[(A',\, X')] = \text{rank}(X')$, i.e., $\text{rank}\left[\begin{pmatrix} A \\ X \end{pmatrix}\right] = \text{rank}(X)$.

   (iv)⇒(i): Suppose $\text{rank}\left[\begin{pmatrix} A \\ X \end{pmatrix}\right] = \text{rank}(X) = r$. Then there exists a sub-matrix of X that contains r linearly independent rows of X such that other rows of $\begin{pmatrix} A \\ X \end{pmatrix}$ are linear combinations of these r rows. Let PX be this sub-matrix. All rows of A are linear combinations of the r rows of this sum-matrix. So $A = QPX$, i.e.,$A = LX$ for some L. Hence $A\beta$ is estimable.

2. Linear model under a linear restriction

   (1) Linear model under linear restriction
   Consider linear model $y = X\beta + e$, $e \sim (0, \sigma^2 I_n)$ under the restriction $G\beta = 0$.
   Note that

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

   So β is confined in a linear space of R^p. Hence we call it a linear restriction.

   (2) Restricted linear unbiased estimator
   Ly is a linear unbiased estimator for $A\beta$ under the restriction $G\beta$ if $E(Ly) = A\beta$ for all β satisfying $G\beta = 0$. So
   Ly is a restricted linear unbiased estimator if and only if $(L - AX)(I - G^+G) = 0$.

   Proof $\quad$ Ly is a restricted linear unbiased estimator under $G\beta = 0$

   $\Longleftrightarrow$ $E(Ly) = A\beta$ for all β satisfying $G\beta = 0$

   $\Longleftrightarrow$ $LX\beta = A\beta$ for all $\beta \in \mathcal{R}(I - G^+G)$

   $\Longleftrightarrow$ $(LX - A)\beta = 0$ for $\beta = (I - G^+G)\gamma$ for all γ

   $\Longleftrightarrow$ $(A - LX)(I - G^+G)\gamma = 0$ for all $\gamma \Longleftrightarrow (A - LX)(I - G^+G) = 0$

(3) Restricted estimable parameter functions
$A\beta$ is estimable under the restriction $G\beta = 0$ if $A\beta$ has one linear unbiased estimator under the restriction $G\beta = 0$.
So $A\beta$ is estimable under the restriction $G\beta = 0$ if and only if $(A - LX)(I - G^+G) = 0$ for some L.

3. Restricted least square estimators

(1) Definition

$$\begin{aligned} & \widehat{\beta} \text{ is a least square estimator (LSE) for } \beta \text{ under } G\beta = 0 \\ \overset{def}{\Longleftrightarrow}\ & G\widehat{\beta} = 0 \text{ and } \|y - X\widehat{\beta}\|^2 \leq \|y - X\beta\|^2 \text{ for all } \beta \text{ satisfying } G\beta = 0. \end{aligned}$$

(2) The collection of all LSE under $G\beta = 0$.
The collection of all LSE for β under $G\beta = 0$ is

$$[X(I - G^+G)]^+y + \mathcal{N}(X) \cap \mathcal{N}(G).$$

Proof

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