L25 Estimable parameters and BLUEs in ANCOVA

1. Recall

   For ANCOVA model $y = (X,\, Z)\begin{pmatrix}\delta\\ \gamma\end{pmatrix} + \epsilon = (X,\, Z)\beta + \epsilon$, $\epsilon \sim N(0,\, \sigma^2 I_n)$

   (1) $E(y) \in S = \mathcal{R}[(X,\, Z)] = \mathcal{R}(X) + \mathcal{R}(Z)$
   Here S can be written as orthogonal sums.

   $$\begin{aligned} S &= \mathcal{R}[(X,\, (I - XX^+)Z)] &&= \mathcal{R}(X)\dot{\oplus}\mathcal{R}[(I - XX^+)Z] \\ &= \mathcal{R}[((I - ZZ^+)X,\, Z)] &&= \mathcal{R}[(I - ZZ^+)X]\dot{\oplus}\mathcal{R}(Z) \end{aligned}$$

   (2) $E(y) = (X,\, Z)\beta$ is estimable with BLUE $(X,\, Z)\widehat{\beta} = (X,\, Z)(X,\, Z)^+ y$
   Here the projection matrix $(X,\, Z)(X,\, Z)^+$ can be decomposed as the sums of two projection matrices onto two perpendicular spaces.

   $$\begin{aligned} (X,\, Z)(X,\, Z)^+ &= XX^+ + [(I - XX^+)Z][(I - XX^+)Z]^+ \\ &= [(I - ZZ^+)Z][(I - ZZ^+)Z]^+ + ZZ^+. \end{aligned}$$

   (3) Assumption of $\mathcal{R}(X) \cap \mathcal{R}(Z) = \{0\}$
   Note that
   $mathcal R(X) \oplus \mathcal{R}(Z)$ is a direct sum $\iff$ $\text{rank}[(X,\, Z)] = \text{rank}(X) + \text{rank}(Z)$
   $\iff$ $\mathcal{R}(X) \cap \mathcal{R}(Z) = \{0\}$.
   Under this assumption, $\mathcal{R}(X') = \mathcal{R}[X'(I - ZZ^+)]$ and $\mathcal{R}(Z') = \mathcal{R}[Z'(I - XX^+)]$.
   Thus there exist T_1 and T_2 such that $X' = X'(I - ZZ^+)T_1'$ and $Z' = Z'(I - XX^+)T_2'$.
   Therefore,

   $$X = T_1(I - ZZ^+)X \text{ and } Z = (I - XX^+)ZT_2.$$

2. Estimable parameters and BLUEs

   (1) Essential estimable parameter and its BLUE
   $E(y) = (X,\, Z)\beta$ in $y = (X,\, Z)\beta + \epsilon \sim N((X,\, Z)\beta,\, \sigma^2 I_n)$ is estimable with BLUE

   $$\widehat{y} = (X,\, Z)\widehat{\beta} = (X,\, Z)(Z,\, Z)^+ y \sim N((X,\, Z)\beta,\, \sigma^2 (X,\, Z)(X,\, Z)^+).$$

   $H\beta$ is estimable $\iff H\beta = L(X,\, Z)\beta$ for some $L \iff H\beta = L[E(y)]$ for some L.
   Estimable $H\beta$ has BLUE

   $$\begin{aligned} H\widehat{\beta} = L(X,\, Z)\widehat{\beta} &= L(X,\, Z)(X,\, Z)^+ y = H(X,\, Z)^+ y \\ &\sim L\, N((X,\, Z)\beta,\, \sigma^2 (X,\, Z)(X,\, Z)^2) \\ &= N(H\beta,\, \sigma^2 L(X,\, Z)(X,\, Z)^2 L'). \end{aligned}$$

   (2) $X\delta$ is estimable with BLUE $(X,\, 0)\widehat{\beta} \sim N\left(X\delta,\, \sigma^2 (X,\, 0)\begin{pmatrix} X'X & X'Z \\ Z'Z & Z'Z \end{pmatrix}^+ \begin{pmatrix} X' \\ 0 \end{pmatrix}\right)$

   Proof. $X\delta = (X,\, 0)\beta = [T_1(I - ZZ^+)](X,\, Z)\beta$. Thus $X\delta$ is estimable with BLUE

   $$\begin{aligned} (X,\, 0)\widehat{\beta} &= [T_1(I - ZZ^+)](X,\, Z)\widehat{\beta} = [T_1(I - ZZ^+)](X,\, Z)(X,\, Z)^+ y \\ &\sim [T_1(I - ZZ^+)]\, N\left((X,\, Z)\beta,\, \sigma^2 (X,\, Z)(X,\, Z)^+\right) \\ &= N(X\delta,\, \sigma^2 [T_1(I - ZZ^+)](X,\, Z)(X,\, Z)^+ [T_1(I - ZZ^+)]') \\ &= N\left(X\delta,\, (X,\, 0)(X,\, Z)^+ [(X,\, Z)']^+ \begin{pmatrix} X' \\ 0 \end{pmatrix}\right) \\ &= N\left(X\delta,\, \sigma^2 (X,\, 0)\begin{pmatrix} X'X & X'Z \\ Z'X & Z'Z \end{pmatrix}^+ \begin{pmatrix} X' \\ 0 \end{pmatrix}\right). \end{aligned}$$

(3) $Y\gamma$ is estimable with BLUE $(0,\, Z)\widehat{\beta} \sim N\left(Z\gamma,\, \sigma^2(0,\, Z)\begin{pmatrix} X'X & X'Z \\ Z'Z & Z'Z \end{pmatrix}^+ \begin{pmatrix} 0 \\ Z' \end{pmatrix}\right)$

Proof. $Z\gamma = (0,\, Z)\beta = [T_2(I - XX^+)](X,\, Z)\beta$. Thus $Z\gamma$ is estimable with BLUE

$$\begin{aligned}
(0,\, Z)\widehat{\beta} &= [T_2(I - XX^+)](X,\, Z)\widehat{\beta} = [T_2(I - XX^+)](X,\, Z)(X,\, Z)^+ y \\
&\sim [T_2(I - XX^+)]\, N\left((X,\, Z)\beta,\, \sigma^2(X,\, Z)(X,\, Z)^+\right) \\
&= N(Z\gamma,\, \sigma^2[T_2(I - XX^+)](X,\, Z)(X,\, Z)^+[T_2(I - XX^+)]') \\
&= N\left(Z\gamma,\, (0,\, Z)(X,\, Z)^+[(X,\, Z)']^+ \begin{pmatrix} 0 \\ Z' \end{pmatrix}\right) \\
&= N\left(Z\gamma,\, \sigma^2(0,\, Z)\begin{pmatrix} X'X & X'Z \\ Z'X & Z'Z \end{pmatrix}^+ \begin{pmatrix} 0 \\ Z' \end{pmatrix}\right).
\end{aligned}$$

3. Variance component model

(1) A model for random vector $y \in R^n$
Consider model $y = X\beta + \sum_{i=1}^k U_i\xi_i$ where $X \in R^{n\times p}$ is a known non-random matrix, $\beta \in R^p$ is an unknown parameter vector, $U_i \in R^{n\times r_i}$, $i = 1, ..., k$, are known non-random matrices, and $\xi_i \sim N(0,\, \sigma_i^2 I_{r_i})$, $i = 1, ..., k$, are independent random error vectors.

(2) Variance component model
For y in (1)

$$y \sim N(X\beta,\, \sigma_1^2 U_1U_1' + \cdots + \sigma_k^2 U_kU_k')$$

is a linear model since $E(y) = X\beta$ is a linear function of β and hence $E(y) \in \mathcal{R}(X)$. But the $\mathrm{Cov}(y) = \sigma_1^2 U_1U_1' + \cdots + \sigma_k^2 U_kU_k'$ has components $\sigma_i^2 U_iU_i'$, $i = 1, ..., k$. Hence the model is called a variance component model.

(3) The problems considered
For the variance component model we need to estimate parameters $\sigma_1^2, ..., \sigma_k^2$ and parameter vector β.

L26: Point estimators in variance component model

1. Parameters in variance component model

   (1) Parameters in variance component model
   For variance component model $y = X\beta + \sum_{i=1}^{k} U_i\xi_i$ where random errors $\xi_i \sim N(0, \sigma_i^2 I_{r_i})$, $i = 1, ..., k$, are independent,

   $$y \sim N(X\beta, \Sigma) \text{ with } \Sigma = \sum_{i=1}^{k} \sigma_i^2 U_i U_i'.$$

   Thus there are parameters $\sigma_1^2, ..., \sigma_k^2$ and β that must be estimated.

   (2) Plan for the estimation order
   We will first estimate $\sigma_1^2, ..., \sigma_k^2$ by $\widehat{\sigma}_1^2, ..., \widehat{\sigma}_k^2$ such that

   $$\widehat{\Sigma} = \sum_{i=1}^{k} \widehat{\sigma}_i^2 U_i U_i'.$$

   Note that for $y \sim N(X\beta, \Sigma)$, β is estimated by $\widehat{\beta} = \text{GLSE}_{\Sigma^{-1}}(\beta) = \text{MLE}(\beta)$ determined by the projection equation $X\widehat{\beta} = X(\Sigma^{-1/2}X)^{+}\Sigma^{-1/2}y$. With Σ estimated by $\widehat{\Sigma}$, β is estimated by $\widehat{\beta}$ determined by

   $$X\widehat{\beta} = X(\widehat{\Sigma}^{-1/2}X)^{+}\widehat{\Sigma}^{-1/2}y$$

   Thus, the key is to estimate $\sigma_1^2, ..., \sigma_k^2$.

2. A general approach to estimating $\sigma_1^2, ..., \sigma_k^2$

   (1) Moment method
   Suppose a population y has parameters $\theta_1, ..., \theta_k$ and $y_1, ..., y_n$ is a random sample from the population. Then $\frac{1}{n}\sum_{i=1}^{n} y_i^t$ is the tth moment from the sample, and $E(y^t)$ is the tth moment from the population. Thus

   $$\begin{cases} E(y) &= \frac{1}{n}\sum_{i=1}^{n} y_i \\ & \vdots \\ E(y^k) &= \frac{1}{n}\sum_{i=1}^{n} y_i^k \end{cases}$$

   is an equation system of k equations for $\theta_1, ..., \theta_k$ since $E(y^t)$ is a function of $\theta_1, ..., \theta_k$. The solutions $\widehat{\theta}_1, ..., \widehat{\theta}_k$ to the equation system can be used as the estimators for $\theta_1, .., \theta_k$. This method is called moment method.

   (2) A proposed method
   Selecting k statistics $T_1, ..., T_k$ and set up an equation system of k equations for $\theta_1, .., \theta_k$

   $$\begin{cases} E(T_1) &= T_1 \\ & \vdots \\ E(T_k) &= T_k \end{cases}$$

   We propose to use the solutions $\widehat{\theta}_1, ...\widehat{\theta}_k$ to the equation system as the estimators for $\theta_1, ..., \theta_k$.

3. Point estimators for $\sigma_1^2, ..., \sigma_k^2$

(1) $B_1, ..., B_k$ and their properties
Let $A_0 = X,\ A_1 = (X,\ U_1)$,...,$A_k = (X,\ U_1, ..., U_k)$ and

$$B_1 = A_1 A_1^+ - A_0 A_0^+,\ B_2 = A_2 A_2^+ - A_1 A_1^+, \cdots, B_k = A_k A_k^+ - A_{k-1} A_{k-1}^+.$$

Then (i) $B_i X = 0$ for all i (ii) $B_i U_j = 0$ for all $j < i$

Proof. (i) For all $j = 0, ..., k$,

$$A_j A_j^+ X = (X, ..., U_j)(X, ..., U_j)^+ (X, ..., U_j) \begin{pmatrix} I \\ \vdots \\ 0 \end{pmatrix} = (X, ..., U_j) \begin{pmatrix} I \\ \vdots \\ 0 \end{pmatrix} = X.$$

So $B_i X = (A_i A_i^+ - A_{i-1} A_{i-1}^+) X = X - X = 0$ for all $i = 0, ..., k$.

(ii) For all $j \le i$,

$$A_i A_i^+ U_j = (X, .., U_j, .., U_i)(X, .., U_j, .., U_i)^+ (X, .., U_j, ., U_i) \begin{pmatrix} 0 \\ \vdots \\ I \\ \vdots \\ 0 \end{pmatrix} = (X, .., U_j, ., U_i) \begin{pmatrix} 0 \\ \vdots \\ I \\ \vdots \\ 0 \end{pmatrix} = U_j.$$

So for $j < i$, $B_i U_j = (A_i A_i^+ - A_{i-1} A_{i-1}^+) U_j = U_j - U_j = 0$.

(2) Point estimators for $\sigma_1^2, ..., \sigma_k^2$
The solutions $\widehat{\sigma}_1^2, ..., \widehat{\sigma}_k^2$ to the equation system below are proposed as the estimators for $\sigma_1^2, ..., \sigma_k^2$.

$$\begin{cases} E(y' B_1 y) &= y' B_1 y \\ & \vdots \\ E(y' B_k y) &= y' B_k y \end{cases}$$

(3) The equation system in (2) is a systgem for $\sigma_1^2, ..., \sigma_k^2$
Note that

$$\begin{aligned} E(y' B_i y) &= [E(y)]' B_i [E(y)] + \mathrm{tr}[B_i \mathrm{Cov}(y)] = (X\beta)' B_i (X\beta) + \mathrm{tr}[B_i \textstyle\sum_{j=1}^k \sigma_j^2 U_j U_j'] \\ &= 0 + \mathrm{tr}[\textstyle\sum_{j=i}^k \sigma_j^2 B_i U_j U_j'] = \sum_{j=i}^k \mathrm{tr}(U_j' B_i U_j)\, \sigma_j^2. \end{aligned}$$

Therefore the equation system in (2) becomes

$$\begin{pmatrix} \mathrm{tr}(U_1' B_1 U_1) & \mathrm{tr}(U_2' B_1 U_2) & \cdots & \mathrm{tr}(U_k' B_1 U_k) \\ 0 & \mathrm{tr}(U_2' B_2 U_2) & \cdots & \mathrm{tr}(U_k' B_2 U_k) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathrm{tr}(U_k' B_k U_k) \end{pmatrix} \begin{pmatrix} \sigma_1^2 \\ \sigma_2^2 \\ \vdots \\ \sigma_k^2 \end{pmatrix} = \begin{pmatrix} y' B_1 y \\ y' B_2 y \\ \vdots \\ y' B_k y. \end{pmatrix}$$

This is indeed a system for $\sigma_1^2, ..., \sigma_k^2$, and it is easy to solve.