Statistical Linear Models

Assignment 4

Hanbin Liu 11912410

1.

We find the least squares estimates of θ and ϕ by minimizing

$$S(\theta,\phi) = \Big[(Y_1-\theta)^2 + (Y_2 - 2\theta+\phi)^2 + (Y_3-\theta-2\phi)^2\Big].$$

Solving

$$\begin{cases} \frac{\partial S(\theta,\phi)}{\partial\theta} = -2(Y_1-\theta) - 4(Y_2-2\theta+\phi) - 2(Y_3-\theta-2\phi) = 0 \\ \frac{\partial S(\theta,\phi)}{\partial\phi} = 2(Y_2-2\theta+\phi) - 4(Y_3-\theta-2\phi) = 0 \end{cases}$$

yields that

$$\hat{\theta} = \frac{Y_1+2Y_2+Y_3}{6}, \quad \hat{\phi} = \frac{2Y_3-Y_2}{5}$$

$\square$

2.

Let $x_1,...,x_m$ denote the m observations of type (a), $y_1,...,y_m$ denote the m observations of type (b) and $z_1,...,z_n$ denote the n observations of type (c). To find the least squares estimates $\hat{\theta}$ and $\hat{\phi}$, we need to minimize

$$\begin{aligned} S(\theta,\phi) &= \sum_{i=1}^m (x_i - E(X_i))^2 + \sum_{i=1}^m (y_i - E(Y_i))^2 + \sum_{i=1}^n (z_i - E(Z_i))^2 \\ &= \sum_{i=1}^m (x_i-\theta)^2 + \sum_{i=1}^m (y_i-\theta-\phi)^2 + \sum_{i=1}^n (z_i-\theta+2\phi)^2, \end{aligned}$$

where X_i, Y_i and Z_i are corresponding random variables. Solving

$$\begin{cases} \frac{\partial S(\theta,\phi)}{\partial\theta} = 0 \\ \frac{\partial S(\theta,\phi)}{\partial\phi} = 0 \end{cases}$$

yields that

$$\hat{\theta} = \frac{(m+4n)\bar{x} + 6n\bar{y} + 3n\bar{z}}{m+13n}$$

$$\hat{\phi} = \frac{(2n-m)\bar{x} + (m+3n)\bar{y} - 5n\bar{z}}{m+13n}$$

where $\bar{x} = \frac{1}{m}\sum_{i=1}^m x_i, \bar{y} = \frac{1}{m}\sum_{i=1}^m y_i$ and $\bar{z} = \frac{1}{n}\sum_{i=i}^n z_i$.
If $m = 2n$, then

$$\hat{\theta} = \frac{2\bar{x}+2\bar{y}+\bar{z}}{5} = \mathbf{A}\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} =: \mathbf{AW}$$

$$\hat{\phi} = \frac{\bar{y}-\bar{z}}{3} = \mathbf{B}\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} =: \mathbf{BW},$$

where

$$\mathbf{A} = \left(\tfrac{1}{5n}\mathbf{1}'_{2n} \quad \tfrac{1}{5n}\mathbf{1}'_{2n} \quad \tfrac{1}{5n}\mathbf{1}'_{n}\right), \quad \mathbf{B} = \left(\mathbf{0}'_{2n} \quad \tfrac{1}{6n}\mathbf{1}'_{2n} \quad -\tfrac{1}{3n}\mathbf{1}'_{n}\right), \quad \mathbf{1}'_n = \begin{pmatrix}1 & 1 & 1 & \dots & 1\end{pmatrix}_{1\times n}$$
$$\mathbf{X} = \begin{pmatrix}X_1 & X_2 & \dots & X_{2n}\end{pmatrix}', \quad \mathbf{Y} = \begin{pmatrix}Y_1 & Y_2 & \dots & Y_{2n}\end{pmatrix}', \quad \mathbf{Z} = \begin{pmatrix}Z_1 & Z_2 & \dots & Z_{n}\end{pmatrix}'$$

Then, the covariance of $\hat{\theta}$ and $\hat{\phi}$ is

$$Cov(\mathbf{AW}, \mathbf{BW}) = \mathbf{A}Cov(\mathbf{W}, \mathbf{W})\mathbf{B}' = A\sigma^2\mathbf{IB}' = \sigma^2\mathbf{IAB}' = \mathbf{0}$$

since $\mathbf{AB}' = 0$. Therefore, $\hat{\theta}$ and $\hat{\phi}$ are uncorrelated if $m = 2n$. □

3.

(a)

$$\begin{aligned}
MSE(\tilde{\beta}) &= E(\tilde{\beta} - \beta)'(\tilde{\beta} - \beta) = E(c\hat{\beta} - \beta)'(c\hat{\beta} - \beta) \\
&= c^2E(\hat{\beta}'\hat{\beta}) - 2cE(\beta'\hat{\beta}) + E(\beta'\beta) \\
&= c^2E(\hat{\beta}'\hat{\beta}) - 2c\beta'\beta + \beta'\beta
\end{aligned}$$

Since $y \sim N(X\beta, \sigma^2 I)$, it then follows that

$$\begin{aligned}
E(\hat{\beta}'\hat{\beta}) &= E(y'X(X'X)^{-1}(X'X)^{-1}X'y) \\
&= tr(X(X'X)^{-1}(X'X)^{-1}X'\sigma^2 I) + \beta'X'X(X'X)^{-1}(X'X)^{-1}X'X\beta \\
&= tr((X'X)^{-1}(X'X)^{-1}X'X\sigma^2 I) + \beta'\beta \\
&= tr((X'X)^{-1}\sigma^2 I) + \beta'\beta \\
&= \sigma^2 tr(X'X)^{-1} + \beta'\beta
\end{aligned}$$

Thus,

$$\begin{aligned}
MSE(\tilde{\beta}) &= c^2(\sigma^2 tr(X'X)^{-1} + \beta'\beta) - 2c\beta'\beta + \beta'\beta \\
&= c^2\sigma^2 tr(X'X)^{-1} + (c-1)^2\beta'\beta
\end{aligned}$$

(b)

Note that $f(c) := MES(\tilde{\beta})$ is a function of c and it is convex. Thus, $f'(c^*) = 0$. That is

$$0 = 2c\sigma^2 tr(X'X)^{-1} + 2(c-1)\beta'\beta,$$

which implies that $c^* = \frac{\beta'\beta}{\sigma^2 tr(X'X)^{-1} + \beta'\beta}$.

(c)

Since the eigenvalues of $X'X$ are $1, 2, 3, 4, 5$, the eigenvalues of $(X'X)^{-1}$ are $1, 1/2, 1/3, 1/4, 1/5$ and thus the trace is $tr(X'X)^{-1} = 1 + 1/2 + 1/3 + 1/4 + 1/5 = 137/60$. $\beta'\beta = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 = 55$. Then

$$c^* = \frac{\beta'\beta}{\sigma^2 tr(X'X)^{-1} + \beta'\beta} = \frac{55}{\frac{137}{60} + 55} = \frac{3300}{3437}$$

□

4.

(a)

Since $\widehat{\beta}^* = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{Y}_1$, we have $(\mathbf{X}_1'\mathbf{X}_1)\widehat{\beta}^* = \mathbf{X}_1'\mathbf{Y}_1$. Since

$$\mathbf{X} = \begin{bmatrix}\mathbf{X}_1\\ \mathbf{X}_2\end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix}\mathbf{Y}_1\\ \mathbf{Y}_2\end{bmatrix},$$

it follows that

$$\begin{aligned}\widehat{\beta} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}\\ &= \left[(\mathbf{X}_1' \quad \mathbf{X}_2')\begin{pmatrix}\mathbf{X}_1\\ \mathbf{X}_2\end{pmatrix}\right]^{-1}[\mathbf{X}_1' \quad \mathbf{X}_2']\begin{bmatrix}\mathbf{Y}_1\\ \mathbf{Y}_2\end{bmatrix}\\ &= (\mathbf{X}_1'\mathbf{X}_1 + \mathbf{X}_2'\mathbf{X}_2)^{-1}(\mathbf{X}_1'\mathbf{Y}_1 + \mathbf{X}_2'\mathbf{Y}_2)\end{aligned}$$

Then,

$$(\mathbf{X}_1'\mathbf{X}_1 + \mathbf{X}_2'\mathbf{X}_2)\widehat{\beta} = (\mathbf{X}_1'\mathbf{Y}_1 + \mathbf{X}_2'\mathbf{Y}_2),$$

which implies that

$$\begin{aligned}\mathbf{X}_1'\mathbf{X}_1\widehat{\beta} &= (\mathbf{X}_1'\mathbf{Y}_1 + \mathbf{X}_2'\mathbf{Y}_2) - \mathbf{X}_2'\mathbf{X}_2\widehat{\beta}\\ &= (\mathbf{X}_1'\mathbf{X}_1)\widehat{\beta}^* + \mathbf{X}_2'(\mathbf{Y}_2 - \mathbf{X}_2\widehat{\beta}).\end{aligned}$$

Thus,

$$\widehat{\beta} - \widehat{\beta}^* = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_2'(\mathbf{Y}_2 - \mathbf{X}_2\widehat{\beta}) = \mathbf{M}_1^{-1}\mathbf{X}_2'(\mathbf{Y}_2 - \mathbf{X}_2\widehat{\beta})$$

It suffices to show that $\mathbf{Y}_2 - \mathbf{X}_2\widehat{\beta} = e_2$. This holds since

$$\hat{\mathbf{Y}} = \mathbf{X}\widehat{\beta} \Longrightarrow \hat{\mathbf{Y}}_2 = \mathbf{X}_2\widehat{\beta}, \quad e_2 = \mathbf{Y}_2 - \hat{\mathbf{Y}}_2 = \mathbf{Y}_2 - \mathbf{X}_2\widehat{\beta}$$

(b)

Since $e_2^* = \mathbf{Y}_2 - \hat{\mathbf{Y}}_2^* = \mathbf{Y}_2 - \mathbf{X}_2\widehat{\beta}^*$, we have

$$\begin{aligned}e_2 &= \mathbf{Y}_2 - \hat{\mathbf{Y}}_2 = \mathbf{Y}_2 - \mathbf{X}_2\widehat{\beta}\\ &= \mathbf{Y}_2 - \mathbf{X}_2(\widehat{\beta}^* + \mathbf{M}_1^{-1}\mathbf{X}_2'e_2)\\ &= \mathbf{Y}_2 - \mathbf{X}_2\widehat{\beta}^* - \mathbf{X}_2\mathbf{M}_1^{-1}\mathbf{X}_2'e_2\\ &= e_2^* - \mathbf{X}_2\mathbf{M}_1^{-1}\mathbf{X}_2'e_2\end{aligned}$$

Thus,

$$(\mathbf{I} + \mathbf{X}_2\mathbf{M}_1^{-1}\mathbf{X}_2')e_2 = e_2^*.$$

Note that $\mathbf{X}_2\mathbf{M}_1^{-1}\mathbf{X}_2' = \mathbf{X}_2(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_2'$ is symmetric and semi positive definite, thus $\mathbf{I} + \mathbf{X}_2\mathbf{M}_1^{-1}\mathbf{X}_2'$ is symmetric and positive definite. $\mathbf{I} + \mathbf{X}_2\mathbf{M}_1^{-1}\mathbf{X}_2'$ is invertible. Then,

$$e_2 = (\mathbf{I} + \mathbf{X}_2\mathbf{M}_1^{-1}\mathbf{X}_2')^{-1}e_2^* \tag{1}$$

so that the expression $\widehat{\beta} - \widehat{\beta}^*$ can be rewritten as

$$\widehat{\beta} - \widehat{\beta}^* = \mathbf{M}_1^{-1}\mathbf{X}_2'(\mathbf{I} + \mathbf{X}_2\mathbf{M}_1^{-1}\mathbf{X}_2')^{-1}e_2^* \tag{2}$$

Another expression

$$\widehat{\beta} - \widehat{\beta}^* = \mathbf{M}_1^{-1}\mathbf{X}_2'e_2 = \mathbf{M}_1^{-1}\mathbf{X}_2'(e_2^* - \mathbf{X}_2\mathbf{M}_1^{-1}\mathbf{X}_2'e_2),$$

which implies that

$$(\mathbf{I} + \mathbf{M}_1^{-1}\mathbf{X}_2'\mathbf{X}_2)(\hat{\beta} - \hat{\beta}^*) = \mathbf{M}_1^{-1}\mathbf{X}_2' e_2^*$$

Note that $\mathbf{I} + \mathbf{M}_1^{-1}\mathbf{X}_2'\mathbf{X}_2 = (\mathbf{X}_1'\mathbf{X}_1)^{-1}(\mathbf{X}_1'\mathbf{X}_1 + \mathbf{X}_2'\mathbf{X}_2) = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}'\mathbf{X}$ is invertible. Therefore, the expression $\hat{\beta} - \hat{\beta}^*$ can be rewritten as

$$\begin{aligned} \hat{\beta} - \hat{\beta}^* &= ((\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}'\mathbf{X})^{-1}\mathbf{M}_1^{-1}\mathbf{X}_2' e_2^* \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_2' e_2^* \end{aligned} \tag{3}$$

(c)

Let

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} = \begin{bmatrix} 1 & -3 \\ 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix},$$

with $\mathbf{X}_2 = \begin{bmatrix} 1 & 4 \end{bmatrix}$ and $\mathbf{Y}_2 = 4$. Then, by (a), we have

$$\begin{aligned} \hat{\beta} - \hat{\beta}^* &= (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_2' e_2 \\ &= (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_2'(\mathbf{Y}_2 - \mathbf{X}_2\hat{\beta}) \\ &= \begin{bmatrix} \frac{1}{7} & 0 \\ 0 & \frac{1}{28} \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} \left(4 - \begin{bmatrix} 1 & 4 \end{bmatrix} \hat{\beta}\right) \\ &= \begin{bmatrix} \frac{4}{7} \\ \frac{4}{7} \end{bmatrix} - \begin{bmatrix} \frac{1}{7} & \frac{4}{7} \\ \frac{1}{7} & \frac{4}{7} \end{bmatrix} \hat{\beta} \end{aligned}$$

Thus,

$$\begin{bmatrix} \frac{8}{7} & \frac{4}{7} \\ \frac{1}{7} & \frac{11}{7} \end{bmatrix} \hat{\beta} = \hat{\beta}^* + \begin{bmatrix} \frac{4}{7} \\ \frac{4}{7} \end{bmatrix} = \begin{bmatrix} \frac{46}{7} \\ -\frac{10}{7} \end{bmatrix},$$

which implies that $\hat{\beta} = \begin{pmatrix} 6.5 & -1.5 \end{pmatrix}'$. $\square$