Assignment 3 - Problem 56 Dec-2018 Paper

Consider a linear model $Y_l = \theta_1 + \theta_2 + \epsilon_l$ for $l = 1, 2$ and $Y_l = \theta_1 - \theta_3 + \epsilon$ for l = 3,4, where ϵ_l are independent with $R(\epsilon_i) = 0$, $Var(\epsilon_i) = \sigma^2 > 0$ for i = 1, ... 4, and $\theta_1, ...\theta_3 \in R$. Which of the following parametric function is estimable?
(A) $\theta_1 + \theta_3$
(B) $\theta_2 - \theta_3$
(C) $\theta_2 + \theta_3$
(D) $\theta_1 + \theta_2 + \theta_3$

Answer: C, $\theta_2 + \theta_3$

1 Solution:

Given,

$$Y_i = \theta_1 + \theta_2 + \epsilon_l \quad \textit{for l=1,2} \tag{1}$$

$$Y_i = \theta_1 - \theta_3 + \epsilon_l \quad \textit{for l=3,4} \tag{2}$$

Which means,

$$\begin{aligned} Y_1 &= \theta_1 + \theta_2 + \epsilon_1 \\ Y_2 &= \theta_1 + \theta_2 + \epsilon_2 \\ Y_3 &= \theta_1 - \theta_3 + \epsilon_3 \\ Y_4 &= \theta_1 - \theta_3 + \epsilon_4 \end{aligned}$$

Matrix Notation

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix}$$

Which can be represented as

$$Y = X\beta + \epsilon_i$$

Given R(ϵ_i) = 0, ϵ_i is In dependant, Var(ϵ_i) = $\sigma^2 > 0$ for all $i = 1, 2, 3, 4$

In dependence implies that all of them un-correlated and given $\text{Var}(\epsilon_i) = \sigma^2$. So co-variance matrix is of form σ^2 I. Which means the co-variance is spherical.

1.1 Estimable parametric function

A parametric functional $\varphi(\beta)$ is estimable if it is uniquely determined by $X\beta$ in the sense that $\varphi(\beta_1) = \varphi(\beta_2)$ whenever $\beta_1, \beta_2 \in R^k$ satisfy $X\beta_1 = X\beta_2$. This is known as Gauss-Markov property of estimable parameters

1.2 Evaluating Options

For $\beta_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ and $\beta_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

$X\beta_1 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$

$X\beta_2 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$

1. **Option A** $\theta_1 + \theta_3$

$$\varphi(\beta_1) = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = 0$$

$$\varphi(\beta_2) = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 2$$

$$\varphi(\beta_1) \neq \varphi(\beta_2)$$

2. **Option B** $\theta_2 - \theta_3$

$$\varphi(\beta_1) = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = 1$$

$$\varphi(\beta_2) = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = -1$$

$$\varphi(\beta_1) \neq \varphi(\beta_2)$$

3. **Option D** $\theta_1 + \theta_2 + \theta_3$

$$\varphi(\beta_1) = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = 1$$

$$\varphi(\beta_2) = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 2$$

$$\varphi(\beta_1) \neq \varphi(\beta_2)$$

4. **Option C** $\theta_2 + \theta_3$

   Let $\beta_1 = \begin{bmatrix} i_1 \\ j_1 \\ k_1 \end{bmatrix} \beta_2 = \begin{bmatrix} i_2 \\ j_2 \\ k_2 \end{bmatrix}$

$$X\beta_1 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} i_1 \\ j_1 \\ k_1 \end{bmatrix} = \begin{bmatrix} i_1 + j_1 \\ i_1 + j_1 \\ i_1 - k_1 \\ i_1 - k_1 \end{bmatrix}$$

$$X\beta_2 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} i_2 \\ j_2 \\ k_2 \end{bmatrix} = \begin{bmatrix} i_2 + j_2 \\ i_2 + j_2 \\ i_2 - k_2 \\ i_2 - k_2 \end{bmatrix}$$

   If β_1, β_2 satisfy $X\beta_1 = X\beta_2$

then, $i_1 + j_1 = i_2 + j_2, \quad i_1 - k_1 = i_2 - k_2 \implies j_1 + k_1 = j_2 + k_2$

$$\varphi(\beta_1) = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} i_1 \\ j_1 \\ k_1 \end{bmatrix} = j_1 + k_1$$

$$\varphi(\beta_1) = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} i_2 \\ j_2 \\ k_3 \end{bmatrix} = j_2 + k_2$$

$\varphi(\beta_1) = \varphi(\beta_2)$ for any β_1, β_2 so **Option C** is correct answer

1.3 Generalised equation for least squares

Let X be input matrix, $\bar{\theta}$ be the parameter vector to be estimated. Then

$$Y = X\bar{\theta} \tag{3}$$

Let $\bar{t}$ be estimated targets.
Then error

$$\bar{e} = [\bar{t} - Y]$$

$$\bar{e}^2 = \frac{1}{2} Tr[[\bar{t} - Y][\bar{t} - Y]^T]$$

Taking derivative w.r.t $\bar{\theta}$ and equating it to 0 to minimise the error and find best value of $\bar{\theta}$. Let the objective function be represented as $J(\theta)$
By chain rule

$$\begin{aligned} \nabla J(\bar{\theta}) &= \frac{\partial J(\bar{\theta})}{\partial \bar{\theta}} \\ &= \Big[\frac{\partial Y)}{\partial \bar{\theta}}\Big]\Big[\frac{\partial \bar{e}}{\partial Y}\Big]\Big[\frac{\partial J(\bar{\theta})}{\partial \bar{e}}\Big] \end{aligned}$$

$$\frac{\partial Y}{\partial \bar{\theta}} = \frac{\partial X\bar{\theta}}{\partial \bar{\theta}} = X^T$$

$$\frac{\partial \bar{e}}{\partial Y} = \frac{\partial}{\partial Y}[\bar{t} - Y] = -I$$

$$\frac{\partial J(\bar{\theta})}{\partial \bar{e}} = \frac{\partial}{\partial \bar{e}}[\bar{e}\bar{e}^T] = 2\bar{e}$$

$$\nabla J(\bar{\theta}) = X^T \bar{e} = 0$$
$$\implies X^T[\bar{t} - Y] = 0$$
$$\implies X^T[\bar{t}] = X^T Y$$
$$\implies \bar{\theta} = [X^T X]^{-1}[X^T \bar{t}]$$

1.4 Evaluating options

Subtracting equation (2) from equation (1), we get parametric function

$$\theta_2 + \theta_3$$

Whose Parameters can be estimated by finding least squares estimation of (1) - (2). But for all the other options parametric estimation can not be done given (1) and (2), as we can not write them as linear combinations of (1) and (2).
Hence the answer is C