Assignment 3 - Problem 56 Dec-2018 Paper

Consider a linear model $Y_1 = \theta_1 + \theta_2 + \epsilon_i$ for $l = 1, 2$ and $Y_1 = \theta_1 - \theta_3 + \epsilon$ for l = 3,4, where ϵ_i 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) $\theta_2 + \theta_2 + \theta_3$

Answer: C, $\theta_2 + \theta_3$

1 Solution:

1.1 Parametric estimation using Maximum Likelihood

Given Mean of the errors in the both parametric functions is 0 and variance is greater than zero. Assuming the error is Gaussian distributed, the deterministic function can be written as follows

$$t = y(X, \Theta) + \epsilon$$

Which means imposing a error distribution on the deterministic function, then we can write

$$P(t/y(\mathbf{X}, \mathbf{\Theta}), \sigma^2) = N(t/y(\mathbf{X}, \mathbf{\Theta}), \sigma^2)$$

Where $\Theta = \theta_1, \theta_2...\theta_n$ are the parameters to estimated
$X = x_1, x_2..., x_n$ are input data sets
t = $t_1, t_2...t_n$ is the output corresponding to $X = x_1, x_2..., x_n$
This equation can be represented and solved by taking Log on both sides

$$P(t/y(\mathbf{X}, \mathbf{\Theta}), \sigma^2) = \Pi_{n=1}^{N} N(t_n/\Theta^T X_i, \sigma^2)$$
$$\ln P(t/y(\mathbf{X}, \mathbf{\Theta}), \sigma^2) = \Sigma_{n=1}^{N} \ln N(t_n/\Theta^T X_i, \sigma^2)$$

Solving equation on RHS without summation. It is a univariate gaussian

$$\Sigma_{n=1}^{N} N(t_n/\Theta^T X_i, \sigma^2) = \Sigma_{n=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{\frac{-1}{2\sigma^2}(t_n - \Theta^T X_n)^2\}$$

$$\ln N(t_n/\Theta^T X_n, \sigma^2) = -\frac{1}{2}\ln\frac{1}{\sigma^2} - \frac{1}{2}\ln 2\pi - \frac{1}{2\sigma^2}\Sigma_{n=1}^{N}(t_n - \Theta^T X_n)^2$$

To find the minima of the error, we take gradient on both sides, minimising error is equivalent to maximising the Θ
Setting Gradient of $\nabla \ln P(\mathbf{t}/\mathbf{y}(\mathbf{X}, \mathbf{\Theta}), \sigma^2) = 0$ we get

$$0 = \frac{\partial}{\partial\Theta}\frac{1}{2}\Sigma_{n=1}^{N}(t_n - \Theta^T X_n)^2$$

$$= \Sigma_{n=1}^{N}(t_n - \Theta^T X_n)\frac{\partial}{\partial\Theta}\Theta^T X_n$$

$$= \Sigma_{n=1}^{N}(t_n - \Theta^T X_n)X_n{}^T$$

We know that Θ is a vector.

$$\frac{\partial}{\partial a}a^T X = X^T \text{ where a is a vector}$$

Now,

$$0 = \Sigma_{n=1}^{N}(t_n - \Theta^T X_n)X_n{}^T$$

$$= \Sigma_{n=1}^{N}(t_n)(X_n)^T - \Theta^T\Sigma_{n=1}^{N}(X_n)X_n^T$$

$$= (t_n)(X_n)^T - \Theta^T(X_n)X_n^T$$

Taking whole transpose and solving it for Θ we get

$$\Theta = [X^T X]^{-1}[X^T \bar{t}] \tag{1}$$

1.2   Evaluating options for ML

In all the given options ML can be estimated when error associated with
the parametric equations can be minimised by applying derivative on the

error and equating it to zero. Given

$$Y_1 = \theta_1 + \theta_2 + \epsilon_i \ \ l = 1, 2 \tag{2}$$
$$Y_1 = \theta_1 - \theta_3 + \epsilon_j \ \ l = 3, 4 \tag{3}$$

(3) subtracted from (2) gives us

$$\theta_2 + \theta_3 + \epsilon_i - \epsilon_j = 0$$
$$\theta_2 + \theta_3 = \epsilon_j - \epsilon_i$$

Parametric function $\theta_2 + \theta_3$ can be estimated by finding ML for (2) and ML for (3) and then subtracting them both. ML can be found from equation (1)
For all other Parametric equations error is not known to take a derivative and equate it to zero.
Hence the correct answer is C