

MA329 Statistical linear models

Assignment 4 (Due date: Nov 26, 11pm. For late submission, each day costs 10 percent)

1. (10 marks) Let

$$\begin{aligned}Y_1 &= \theta + \epsilon_1 \\Y_2 &= 2\theta - \phi + \epsilon_2 \\Y_3 &= \theta + 2\phi + \epsilon_3\end{aligned}$$

where $E[\epsilon_i] = 0$ ($i = 1; 2; 3$). Find the least squares estimates of θ and ϕ .

2. (15 marks) In order to estimate two parameters θ and ϕ , it is possible to make observations of three types:
- (a) the first type have expectation θ ,
 - (b) the second type have expectation $\theta + \phi$, and
 - (c) the third type have expectation $\theta - 2\phi$.

All observations are subject to independent normal errors with zero means and common variance σ^2 . If m observations of type (a), m observations of type (b) and n observations of type (c) are made, find the least squares estimates $\hat{\theta}$ and $\hat{\phi}$. Prove that these estimates are uncorrelated if $m = 2n$.

3. (15 marks) Consider the linear regression model

$$y = \mathbf{X}_{n \times p} \boldsymbol{\beta}_{p \times 1} + \boldsymbol{\epsilon}$$

where $\boldsymbol{\epsilon} \sim N(\mathbf{0}; \sigma^2 \mathbf{I})$. Let $\hat{\boldsymbol{\beta}}$ be the least squares estimate of $\boldsymbol{\beta}$. Define $\tilde{\boldsymbol{\beta}} = c\hat{\boldsymbol{\beta}}$ where $c \leq 1$. The mean squared error (MSE) of $\tilde{\boldsymbol{\beta}}$ is

$$MSE(\tilde{\boldsymbol{\beta}}) = E(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})'(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}).$$

- (a) Prove that $MSE(\tilde{\boldsymbol{\beta}}) = c^2 \sigma^2 \text{tr}(\mathbf{X}'\mathbf{X})^{-1} + (c - 1)^2 \boldsymbol{\beta}'\boldsymbol{\beta}$.
 - (b) Let c^* be the value of c such that $MSE(\tilde{\boldsymbol{\beta}})$ is a minimum. Find c^* .
 - (c) Let $p = 5$, $\sigma^2 = 1$, $\boldsymbol{\beta}' = (1; 2; 3; 4; 5)$ and the eigenvalues of $\mathbf{X}'\mathbf{X}$ be 1, 2, 3, 4, 5. Evaluate c^* .
4. (20 marks) Let

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

and assume that \mathbf{X} and \mathbf{X}_1 have full column rank. Consider the linear model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where $\epsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$. Let $\hat{\beta}$ be the least squares estimator of β and $\hat{Y} = \mathbf{X}\hat{\beta} = (\hat{Y}_1, \hat{Y}_2)'$. Further, for the linear model

$$\mathbf{Y}_1 = \mathbf{X}_1 \beta^* + \epsilon^*$$

where $\epsilon^* \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$, the least squares estimator of β^* is $\hat{\beta}^*$. Let

$$\hat{Y}^* = \mathbf{X} \hat{\beta}^* = \begin{bmatrix} \hat{Y}_1^* \\ \hat{Y}_2^* \end{bmatrix}$$

Define

$$\begin{aligned} \mathbf{Y} - \hat{Y} &= \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix} - \begin{bmatrix} \hat{Y}_1 \\ \hat{Y}_2 \end{bmatrix} = \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \\ \mathbf{Y} - \hat{Y}^* &= \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix} - \begin{bmatrix} \hat{Y}_1^* \\ \hat{Y}_2^* \end{bmatrix} = \begin{bmatrix} e_1^* \\ e_2^* \end{bmatrix} \end{aligned}$$

(a) Prove that

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

where $\mathbf{M}_1 = \mathbf{X}_1' \mathbf{X}_1$.

- (b) Express e_2 in terms of e_2^* and rewrite the expression of $\hat{\beta} - \hat{\beta}^*$ in Part (a)
(c) The following is a data set with sample size = 7

x	-3	-2	-1	0	1	2	3
y	14	7	-2

For the above data and with a simple linear regression model, the parameter estimate $\hat{\beta}^* = (6, -2)'$.

Suppose an additional observation $(x, y) = (4, 4)$ is obtained (You now have 8 pairs of (x, y) in your updated dataset), compute the new parameter estimate $\hat{\beta}$. (Hint: use Parts (a) and (b))