3ª Lista de Exercícios - SME0812 Modelos Lineares

Exercício 1. Considere o modelo $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ com $\boldsymbol{\epsilon} \sim N_n(\mathbf{0}, \sigma^2\mathbf{I}_n)$ e $\mathbf{X}$ de posto completo $k+1$. Se $\widehat{\boldsymbol{\epsilon}} = \mathbf{Y} - \widehat{\mathbf{Y}} = \mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}$, prove que

(i) $\mathrm{Cov}(\widehat{\boldsymbol{\epsilon}}, \widehat{\boldsymbol{\beta}}) = \mathbf{0}$

(ii) $\mathrm{Cov}(\boldsymbol{\epsilon}, \widehat{\boldsymbol{\beta}}) = \mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\sigma^2$

(iii) $\mathrm{Cov}(\boldsymbol{\epsilon}, \mathbf{Y}) = [\mathbf{I} - \mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top]\sigma^2$

(iv) $\mathrm{Cov}(\boldsymbol{\epsilon}, \widehat{\mathbf{Y}}) = \mathbf{0}$.

Com base nos itens (iii) e (iv), discuta por que é preferível a análise gráfica de $\boldsymbol{\epsilon} \times \widehat{\mathbf{Y}}$ em relação à análise gráfica de $\boldsymbol{\epsilon} \times \mathbf{Y}$.

Exercício 2. Seja $\mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}_n)$ com $\mathbf{X}$ de posto completo p. Prove que

$$\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)^\top \frac{\mathbf{X}^\top\mathbf{X}}{\sigma^2} \left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \sim \chi^2_p.$$

Exercício 3. No modelo $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, com $\boldsymbol{\epsilon} \sim N_n(\mathbf{0}, \sigma^2\mathbf{I}_n)$, determine a distribuição do vetor de resíduos $\widehat{\boldsymbol{\epsilon}} = \mathbf{Y} - \widehat{\mathbf{Y}} = \mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}$.

Exercício 4. Para o modelo $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, $i = 1, \ldots, 20$, onde $\sum_{i=1}^{20} y_i^2 = 90$ e $\epsilon_i \sim N(0, \sigma^2)$ independentes, o sistema de equações normais obtido foi

$$\begin{bmatrix} 20 & 10 \\ 10 & 6 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 40 \\ 22 \end{bmatrix}$$

(a) Calcule $\widehat{\boldsymbol{\beta}}^\top \mathbf{X}^\top\mathbf{Y}$ e SQRes=$\mathbf{Y}^\top\mathbf{Y} - \widehat{\boldsymbol{\beta}}^\top\mathbf{X}^\top\mathbf{Y}$.

(b) Construa um intervalo de confiança com coeficiente de confiança 0,95 para $E(Y|x=3)$.

(c) Construa o correspondente intervalo de predição.

Exercício 5. Sejam

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

onde $E(\epsilon_i) = 0$, $i = 1, \ldots, 3$. Determine os estimadores de mínimos quadrados de θ e γ.

Exercício 6. Considere um modelo linear em que

$$\begin{aligned} \mathrm{E}(Y_{1i}) &= \theta,\ i = 1, \ldots, m, \\ \mathrm{E}(Y_{2i}) &= \theta + \gamma,\ i = 1, \ldots, m, \\ \mathrm{E}(Y_{3i}) &= \theta - 2\gamma,\ i = 1, \ldots, n \end{aligned}$$

em que todas as observações estão sujeitas a erros independentes com média 0 e variância σ^2.

(a) Determine os estimadores de mínimos quadrados de θ e γ.

(b) Prove que esses estimadores são não correlacionados se $m = 2n$.

Exercício 7. Mostre que $(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})^\top(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta}) = (\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}})^\top(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}}) + (\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})^\top\mathbf{X}^\top\mathbf{X}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})$ com $\widehat{\boldsymbol{\beta}} = (\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{Y}$ e conclua que $(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})^\top(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})$ é minimizada para $\boldsymbol{\beta} = \widehat{\boldsymbol{\beta}}$.

Exercício 8. Prove que o coeficiente de explicação do modelo, R^2, em um modelo de regressão linear simples é o quadrado do coeficiente de correlação entre $\mathbf{Y}$ e $\widehat{\mathbf{Y}}$.