4ª Lista de Exercícios – Teoria dos Modelos Lineares

Procure desenvolver os exercícios a seguir definindo adequadamente as variáveis aleatórias de interesse, as hipóteses a serem testadas e o desenvolvimento da solução final.

Exercício 1) Obtenha a expressão para o estimador de mínimos quadrados **ponderados** para o modelo linear na forma matricial:

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

onde $\boldsymbol{\epsilon}$ é um vetor de variáveis aleatórias independentes e identicamente distribuídas, $\epsilon_i \sim Normal(0, \sigma^2)$, $E(\boldsymbol{\epsilon}) = \mathbf{0}$ e $Cov(\boldsymbol{\epsilon}) = \sigma^2\mathbf{I}$. Para o estimador de mínimos quadrados ponderados, a equação de otimização é descrita na forma:

$$\widehat{\boldsymbol{\beta}} = \arg min_{\boldsymbol{\beta}} \, (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T \mathbf{W} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$$

onde $\mathbf{W}$ é uma matriz diagonal cujos elementos, $[\mathbf{W}]_{ii}$ são definidos previamente pelo usuário.

(a) Para o modelo acima encontre as expressões para $E[\widehat{\boldsymbol{\beta}}]$, $Cov[\widehat{\boldsymbol{\beta}}]$ e para a matriz de projeção $\mathbf{H}$.

Exercício 2) Para a base de dados "aneel_2014-2016.csv", utilizada na lista 4, ajuste um modelo de regressão linear múltipla no R utilizando as seguinte variáveis preditoras: rsub, log(rdist_a), log(ralta), log(mponderado) e log(cons). Ajuste dois modelos: o primeiro modelo utilizando o log(PMSOaj) como variável preditora e o segundo modelo utilizando a transformação de Boxcox. Faça uma comparação sucinta dos resultados obtidos. Na sua opinião, qual o melhor modelo. Justifique.