

9ª Lista de exercícios – Teoria dos Modelos Lineares

Na lista 8, foi utilizada a base de dados "boston_corrected.txt", na qual estão disponíveis 13 potenciais variáveis preditoras. Nesta lista (9), vamos utilizar todas as potenciais variáveis preditoras para o ajuste da variável resposta, ou seja, do preço dos imóveis. O foco do ajuste não será na relevância estatística, mas na capacidade preditiva do modelo. Nesse sentido, o objetivo do exercício é encontrar o modelo que apresente o maior valor do coeficiente de determinação preditivo (R^2_{pred}). Uma opção é realizar o ajuste utilizando o estimador do tipo *ridge regression*, ou inserindo uma penalização do tipo l_2 . O estimador *ridge regression* pode ser escrito na forma:

$$\tilde{\beta} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^T \mathbf{Y}$$

Faça o ajuste do estimador *ridge regression* de forma a maximizar o coeficiente de determinação preditivo. Faça uma comparação entre os coeficientes estimados com o método de mínimos quadrados (opção padrão) e faça uma comparação do coeficiente de determinação preditivo utilizando as 13 variáveis preditoras sem a penalização do tipo l_2 .

Na sua opinião, quais são as vantagens e desvantagens de utilizar o estimador *ridge regression*?

Elabore o seu relatório utilizando o **Rmarkdown** e envie o documento em formato **PDF** para avaliação pelo sistema minha.ufmg. O documento deve conter, no máximo, 6 páginas.