Universidade do Estado do Rio de Janeiro

Instituto Politécnico do Rio de Janeiro

**Análise da Resposta de Sensores de Baixo Custo Aplicados na Agricultura**

Uso de tecnologia Python aplicado no cálculo de correlação entre duas variáveis regressoras

Graduação em Engenharia de Computação

Aluno: Kauan Peçanha Lira

Professor: Bernardo Sotto Maior Peralva

Nova Friburgo, Rio de Janeiro

**SUMÁRIO**

[**1 - Introdução 3**](#_eaf65ompog2u)

[**2 - Desenvolvimento 3**](#_o8w0y7uopp7s)

# 1 - Introdução

Ao longo desse período, a turma de 2023.2 da disciplina de Modelos Lineares aprendeu sobre as técnicas mais utilizadas de Análise de Dados da Estatística, sendo elas o Modelo de Regressão Linear Simples, e o Modelo de Regressão Linear Múltiplo.

Vimos que a utilização do Modelo de Regressão Linear Múltiplo permite que dois eventos distintos possam ser correlacionados, de forma que se calcule matematicamente um modelo capaz de predizer, com o máximo de precisão possível, a dispersão dos dados. Dessa forma, pode-se tentar compreender o comportamento deles para que seja possível sua previsão dados alguns cenários que possam vir a acontecer.

Tendo em mente toda a teoria que circunda a importância dessa estatística, para esse trabalho, foi utilizado o Modelo de Regressão Linear Múltiplo com o intuito de correlacionar duas variáveis regressoras com suas dispersões definidas aleatoriamente em torno de uma faixa de valores pré-definida. Através desse uso, espera-se que um modelo matemático possa ser desenvolvido com o intuito de prever, com máxima precisão, o comportamento de ambas as dispersões, com seus dados relacionados entre si.

# 2 - Desenvolvimento

1 - a) Faça o gráfico da dispersão bidimensional (XY) entre a variável y e cada uma das variáveis regressoras X1, … , Xp (p gráficos separados).

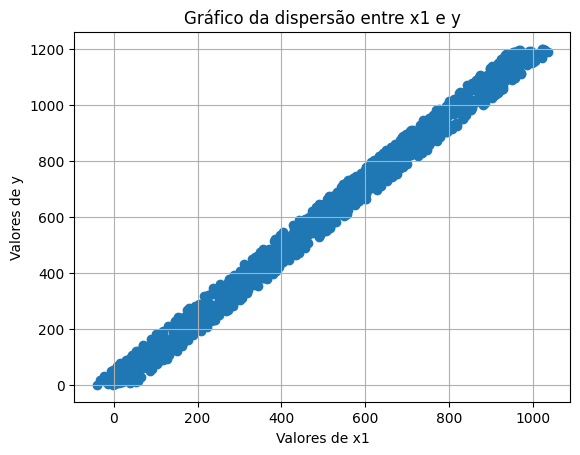


Gráfico 1 - Dispersão dos valores (X1, Y)

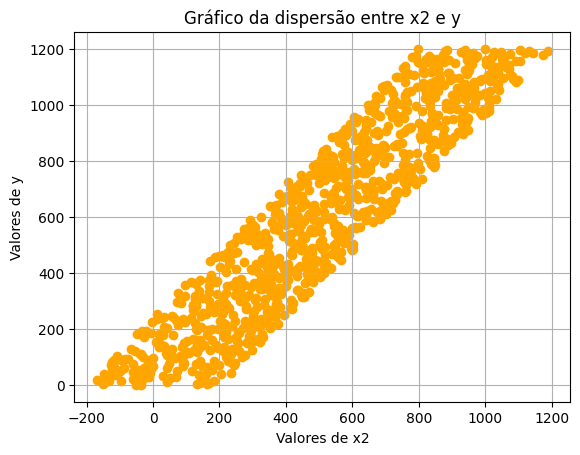


Gráfico 2 - Dispersão dos valores (X2, Y)

1 - b) Estime a correlação entre a variável regressora Y e cada uma das variáveis regressoras (p(Y, Xp)). Comente.

Solução:

O coeficiente de correlação entre x1 e y é 0.9952

O coeficiente de correlação entre x2 e y é 0.9290

Resultado Computacional da Correlação entre Y e as variáveis regressoras X1 e X2, respectivamente.

Comentário: Note que a primeira variável regressora detém um coeficiente de correlação de aproximadamente 0,9952, o que significa que 99% de sua dispersão está relacionada com Y, ou seja, fortemente relacionada. Já a variável regressora X2 detém um coeficiente de correlação de 0,9290, ou seja, uma correlação de aproximadamente 93%.

1 - c) Calcule o coeficiente de determinação R²(Y, Xp) entre a variável Y e cada uma de suas variáveis regressoras.

Solução:

O valor de R² para a primeira variável regressora é 0.99

O valor de R² para a segunda variável regressora é 0.86

Valores de R² para X1 e X2, respectivamente.

Comentário: Note que R², em ambos os casos, está muito perto de 1,00, o que significa que boa parte da variação da variável dependente é explicado pela variável independente.

1 - d) Encontre o hiperplano de quadrados mínimos (estime b0, b1 e b2)

Solução:

Equação do MRLM: 5.2461 + 1.1099 \* x1 + 0.0776 \* x2

Equação 1 - Hiperplano composto pelos coeficientes b0, b1 e b2, sendo os dois últimos acompanhando x1 e x2.

1 - e) Calcule os valores estimados de y(ou seja, y^) através do hiperplano estimado no item anterior.

Solução:

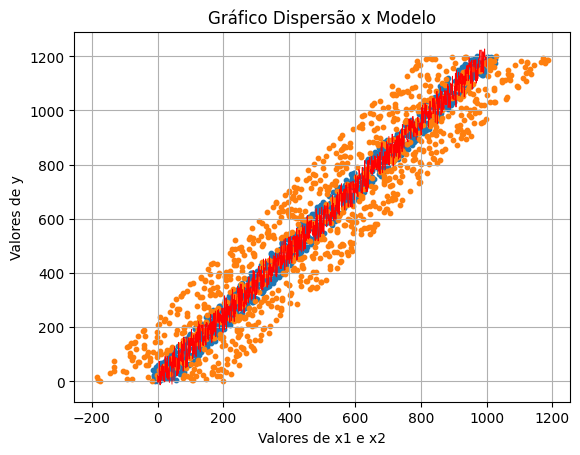


Gráfico 3 - Valores calculados através do Modelo de Regressão Linear Múltiplo

1 - f) Calcule os resíduos, ou seja, o erro de estimação y - y^

Solução:

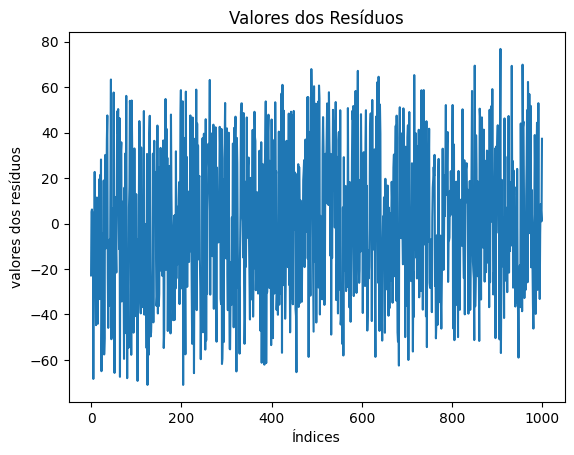


Gráfico 4 - Gráfico dos valores dos resíduos

1 - g) Faça o histograma dos erros de estimação encontrados no item anterior. Comente sobre sua estatística.

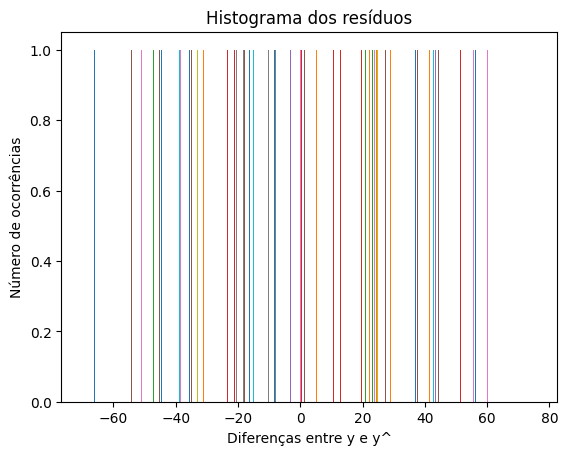


Gráfico 5 - Gráfico do histograma dos resíduos

Comentário: Note que o histograma delata que nenhum valor advindo da diferença entre y e y^ ocorre mais do que uma vez. Além disso, é possível perceber através de uma análise rápida e sem viés matemático que eles estão concentrados no intervalo [-60, 60].

# 3 - Software

O software desenvolvido para a solução do trabalho proposto pode ser acessado através do link de acesso abaixo, no qual pode-se acessar o arquivo jupyter notebook que compõem todos os passos necessários para a solução de cada um de seus respectivos itens, conforme pode ser observado em seus comentários, denotados pelo começo “#” .

<https://drive.google.com/drive/folders/1tp-yRwi3ur450fDFCdKZIEFl495KQiKb?usp=sharing>

# 4 - Agradecimento

Meu agradecimento ao professor Bernardo por ter conduzido a disciplina de Modelos Lineares de forma excelente. O conhecimento adquirido aqui será parte de um ótimo uso na engenharia, futuramente. Obrigado.