

ATIVIDADE 1 – Regressão Linear

A primeira atividade do TCC consiste em um exercício de regressão linear.

A)

Seja o conjunto de 5 medições (x_i, y_i) , onde $i = 1, 2, \dots, 5$:

$$\begin{aligned}(x_1, y_1) &= (0, 0.2) \\(x_2, y_2) &= (0.1, 0.3) \\(x_3, y_3) &= (0.2, 0.45) \\(x_4, y_4) &= (0.3, 0.7) \\(x_5, y_5) &= (0.4, 0.8)\end{aligned}$$

Deseja-se obter a melhor reta que relaciona (x_i, y_i) . Essa reta é definida pelo corpo de uma função afim, onde a e b são constantes ainda a serem determinadas. Obtenha os valores ótimos dos coeficientes ajustáveis a e b , usando mínimos quadrados.

$$y_i = ax_i + b$$

B)

Fazendo as contas no papel.

1. Defina, para cada par (x_i, y_i) , um erro como sendo a diferença entre o valor medido (y_i) e o valor fornecido pela equação da reta ($ax_i + b$). Após isso, realizar a soma dos quadrados dos erros. Essa equação, chamada de MSE, dependerá dos coeficientes ajustáveis a e b .
2. O objetivo aqui é escolher coeficientes ajustáveis a e b que minimizem o MSE. Para isso, buscam-se os pontos de mínimos de MSE em relação aos coeficientes ajustáveis a e b :
 - (a) Derivar a equação do MSE em relação ao coeficiente a e, na sequência, igualar o resultado a zero;
 - (b) Derivar a equação do MSE em relação ao coeficiente b e, na sequência, igualar o resultado a zero;
 - (c) Resolver o sistema algébrico linear de 2 equações (resultantes dos itens anteriores) nas 2 incógnitas a e b .
3. Use os coeficientes a e b obtidos e, para cada x_i , obter a estimativa usando a equação da reta. Em um mesmo gráfico, visualizar:
 - (y_i) em função de (x_i) ;
 - $(ax_i + b)$ em função de (x_i) .

C)

Usando mínimos quadrados no MATLAB (ou Python).

1. Usando notação matricial/vetorial, definir:

(a) Vetor de saídas desejadas de dimensão 5×1 :

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix}$$

(b) Vetor de entradas aplicadas de dimensão 5×1 :

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$$

(c) Matriz de regressão de dimensão 5×2 :

$$XX = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ x_4 & 1 \\ x_5 & 1 \end{bmatrix}$$

(d) Vetor de coeficientes ajustáveis de dimensão 2×1 :

$$COEF = \begin{bmatrix} a \\ b \end{bmatrix}$$

2. Admitindo que todos os erros sejam nulos, tem-se o seguinte sistema algébrico linear:

$$Y = XX \cdot COEF$$

Em um sistema algébrico linear onde a quantidade de equações é maior que a quantidade de incógnitas, só há solução se houver equações redundantes. Sendo 5 medições diferentes, espera-se que não haja redundância. Dessa forma, não há solução exata para esse sistema. Contudo, é possível obter a solução que minimize o erro quadrático médio (MSE). No MATLAB, isso é feito através do comando de mínimos quadrados:

$$COEF = XX \setminus Y$$

3. Use o vetor de coeficientes obtido para obter o vetor de saídas estimadas pela reta:

$$\hat{Y} = XX \cdot COEF$$

Em um mesmo gráfico, visualizar:

- Y em função de X ;
- \hat{Y} em função de X .