
INTRODUÇÃO À MATEMÁTICA E IA

Exercícios: Geometria Analítica

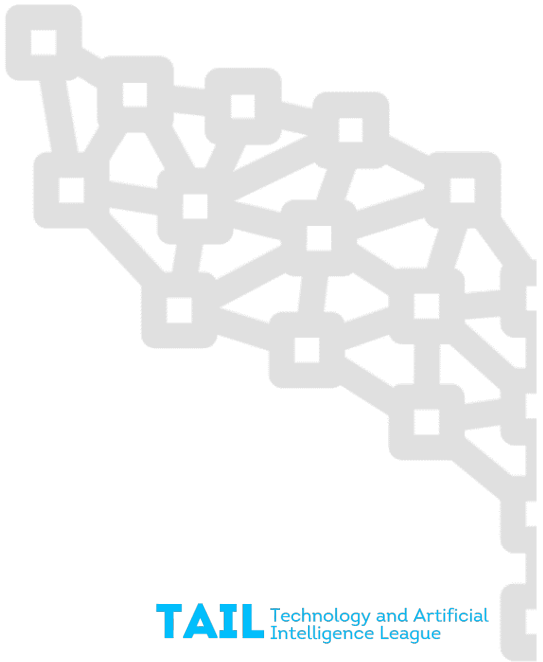
Diretoria de Matemática

Diretor: Manuel F. Junior

Membro: Yann N. G. Nóbrega

`tail.ufpb@gmail.com`

`https://aria.ci.ufpb.br/tail/`



Sumário

Ponto, Reta e Plano

Distância Euclidiana

Distância Manhattan

Exemplos

Métodos de Avaliação de Regressão

R^2

MSE

RMSE

MAE

Aplicação Prática



Ponto, Reta e Plano

- ▶ **Pontos:** Começamos com o objeto mais simples, o ponto, sendo este com função de determinar localização no espaço, porém com dimensão **nula** (0), ou seja, por definição, podemos dizer que um ponto não possui **Volume**, **Área** ou **Comprimento**.



Ponto, Reta e Plano

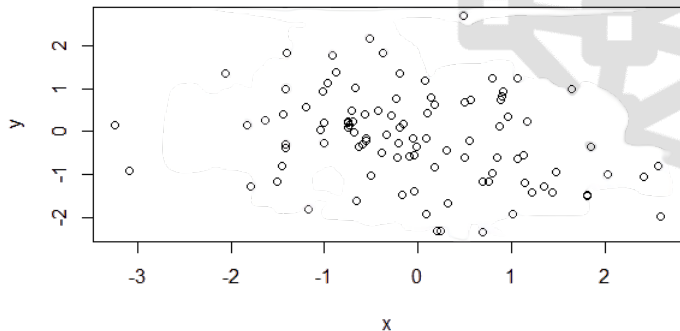


Figura: Gráfico de pontos em um espaço bidimensional



Ponto, Reta e Plano

► Retas:

$$ax + b = y$$

As retas são conjuntos de pontos, de tal forma que não formem uma curva. Dentro dessas retas, podemos calcular a distância entre eles, agora definindo um **segmento de reta** entre eles. Encontramos as retas, muitas vezes, como a equação vista acima, $ax + b = y$, essa forma é como veremos bem comum no cálculo de regressões, um tópico de muita importância.



Reta

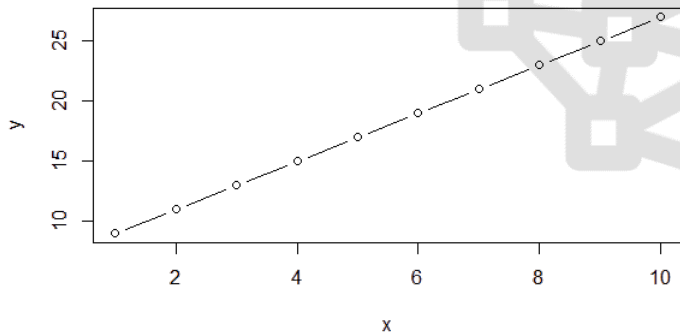


Figura: Gráfico da Reta e sua composição de pontos



Distância Euclidiana

- ▶ A distância Euclidiana será a mais utilizada!
- ▶ Muitos a conhecem, no plano, como Teorema de Pitágoras.
- ▶ Na 3° ou maiores dimensões utilizamos o Teorema várias vezes.
- ▶ Vamos mostrar a generalização para n dimensões



Distância Euclidiana

Essa é a forma mais usual de se calcular a distância entre dois pontos, normalmente calculamos (muitas vezes sem nem se dar conta) com a decomposição do vetor posição em planos e com o uso sequencial do Teorema de Pitágoras. O cálculo da distância euclidiana é dado a partir da generalização do teorema para n dimensões, mostrado na seguinte expressão:

$$d(P, Q) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Sendo, $P = (x_1, x_2, x_3, \dots, x_n)$ e $Q = (y_1, y_2, y_3, \dots, y_n)$. Onde n é o número de dimensões.



Distância Euclidiana

- ▶ **Unidimensional:** para o caso unidimensional ($n = 1$), temos o mais trivial, onde calculamos a distância entre dois pontos. Pela fórmula, temos o seguinte:

$$d(P, Q) = \sqrt{(x_1 - y_1)^2} = |x_1 - y_1|$$



Distância Euclidiana

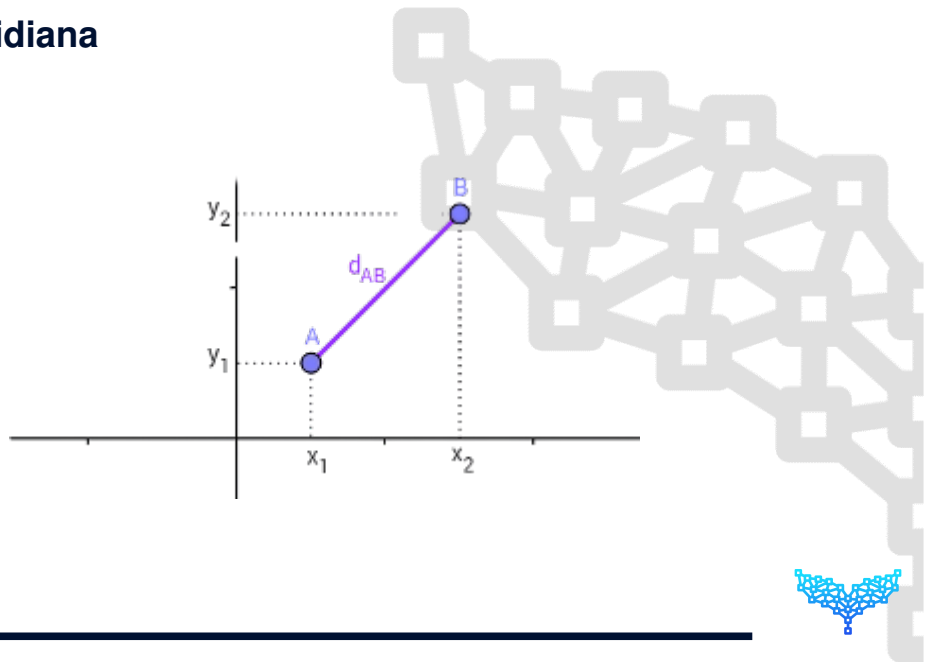
- **Bidimensional:** Para o caso bidimensional ($n = 2$), a fórmula torna-se uma junção aplicação mais ampla do teorema de Pitágoras, tal que:

$$d(A, B) = d_{AB} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

Sendo $A = (x_1, x_2)$ e $B = (y_1, y_2)$.



Distância Euclidiana



Distância Euclidiana

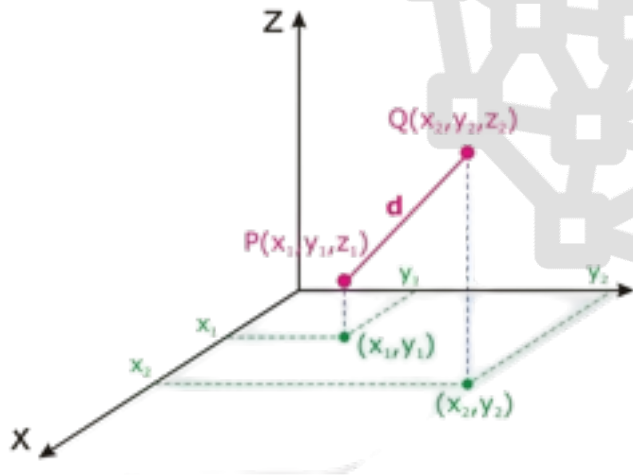
- ▶ **Tridimensional:** Para o caso tridimensional ($n=3$), devemos utilizar uma forma ainda mais geral para o teorema de Pitágoras, dado por:

$$d(P, Q) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}$$

Onde $P = (x_1, x_2, x_3)$ e $Q = (y_1, y_2, y_3)$



Distância Euclidiana



Distância Manhattan

- ▶ Também conhecida como **geometria do táxi**.

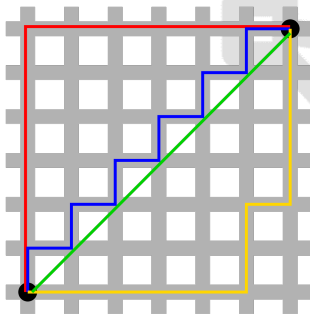


Figura: Distância Manhattan e Euclidiana



Distância Manhattan

- ▶ A distância prática entre 2 pontos não é a reta que os liga
- ▶ Quantas unidades no eixo x mais a quantidade no eixo y.

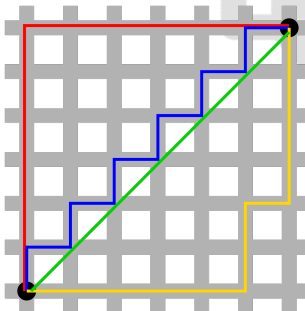


Figura: Distância Manhattan e Euclidiana



Distância Manhattan

$$d(P_0, P_1) = |P_1 - P_0|$$

$$P_0 = (x_0, y_0, z_0, \dots, w_0)$$

$$P_1 = (x_1, y_1, z_1, \dots, w_1)$$

$$d(P_0, P_1) = |x_1 - x_0| + |y_1 - y_0| + |z_1 - z_0| + \dots + |w_1 - w_0|$$



Distância Manhattan

Qual a distância Manhattan entre os pontos:

$$P_1 = (0, 3, 7) \text{ e}$$

$$P_2 = (4, 0, -5)$$



Exemplos

- ▶ Utilizamos a definição da distância:

$$d(P_0, P_1) = |0 - 4| + |3 - 0| + |7 - (-5)|$$

$$d(P_0, P_1) = |4 + 3 + 12|$$

- ▶ Ou, combinando as duas contas antes de operar:

$$d(P_0, P_1) = 19$$



A Regressão

Esquema	Nº Ataques	Duração	Índice
A	5	118	8.1
B	13	132	6.8
C	20	119	7.0
D	28	153	7.4
E	41	91	7.7
F	49	118	7.5
G	61	132	7.6
H	62	105	8.0



Métodos de Avaliação de Regressão

$$Y = \begin{bmatrix} 8.1 \\ 6.8 \\ 7.0 \\ 7.4 \\ 7.7 \\ 7.5 \\ 7.6 \\ 8.0 \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} 7.4 \\ 7.3 \\ 7.4 \\ 7.2 \\ 7.8 \\ 7.6 \\ 7.5 \\ 7.8 \end{bmatrix}$$

Vamos encontrar os seguintes valores:

- ▶ R^2
- ▶ MSE
- ▶ RMSE
- ▶ MAE



Coefficiente de Determinação (R^2)

- ▶ Avalia a variabilidade da variável dependente (Y) que é explicada pelo nosso modelo, baseado na variável independente (X).
- ▶ Definimos \bar{y} , y_i e \hat{y}_i

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

- ▶ O R^2 é então:

$$R^2 = 1 - \frac{\sum_{i=1}^n |y_i - \hat{y}_i|^2}{\sum_{i=1}^n |y_i - \bar{y}|^2}$$



Coefficiente de Determinação (R^2)

$$\bar{y} = \frac{(8.1 + 6.8 + 7.0 + 7.4 + 7.7 + 7.5 + 7.6 + 8.0)}{8} = 7.5125$$

$$num = \sum_{i=1}^n |y_i - \hat{y}_i|^2 = 1.01$$

$$den = \sum_{i=1}^n |y_i - \bar{y}|^2 = 1.408755$$

$$R^2 = 1 - \frac{\sum_{i=1}^n |y_i - \hat{y}_i|^2}{\sum_{i=1}^n |y_i - \bar{y}|^2} = 0.2830524$$



Erro Quadrático Médio (MSE):

- ▶ O MSE é a distância da estimativa do nosso modelo para os valores verdadeiros dos dados, da seguinte forma:

$$MSE = \frac{1}{n} \cdot d(Y, \hat{Y})^2$$



Erro Quadrático Médio (MSE):

$$\begin{aligned}MSE &= (8.1 - 7.4)^2 + (6.8 - 7.3)^2 + (7.0 - 7.4)^2 + (7.4 - 7.2)^2 + \\&+ (7.7 - 7.8)^2 + (7.5 - 7.6)^2 + (7.6 - 7.5)^2 + (8.0 - 7.8)^2 = 1.01 \\MSE &= 0.12625\end{aligned}$$



Raiz do Erro Quadrático Médio (RMSE):

- ▶ A Raiz do MSE é a raiz da distância da estimativa do nosso modelo para os valores verdadeiros dos dados, da seguinte forma:

$$RMSE = \sqrt{\frac{1}{n} \cdot d(Y, \hat{Y})^2}$$



Raiz do Erro Quadrático Médio (RMSE):

$$RMSE = \sqrt{MSE} = \sqrt{0.12625}$$

$$RMSE = 0.3553$$



Erro Absoluto Médio (MAE):

- ▶ É a média das distâncias entre os valores preditos (\hat{Y}) e reais (Y):

$$MAE = \frac{1}{n} \cdot \sum_{i=1}^n |y_i - \hat{y}_i| = \frac{1}{n} \cdot \sum_{i=1}^n d(y_i, \hat{y}_i)$$



Erro Absoluto Médio (MAE):

$$\begin{aligned} MAE &= |8.1 - 7.4| + |6.8 - 7.3| + |7.0 - 7.4| + |7.4 - 7.2| + \\ &+ |7.7 - 7.8| + |7.5 - 7.6| + |7.6 - 7.5| + |8.0 - 7.8| = 2.3/8 \\ MAE &= 0.2875 \end{aligned}$$



No Python:

- ▶ `from sklearn.metrics import r2_score, mean_squared_error, mean_absolute_error`
- ▶ `import numpy as np`
- ▶ `r2 = r2_score(Y_true, Y_pred)`
- ▶ `mse = mean_squared_error(Y_true, Y_pred)`
- ▶ `rmse = mean_squared_error(Y_true, Y_pred, squared = False)`
- ▶ `mae = mean_absolute_error(Y_true, Y_pred)`

