TAIL — INTRODUÇÃO À MATEMÁTICA PARA INTELIGÊNCIA ARTIFICIAL

Diretoria de Matématica - Tail

11/09/2020

1 Introdução a matrizes

Uma matriz $A_{m \times n}$ é um arranjo retangular com m linhas e n colunas da forma

$$A_{m\times n} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix} \tag{1}$$

onde $a_{ij} \in \mathbb{R}$, $i = 1, 2, ..., m$ e $j = 1, 2, ..., n$.

A i-ésima linha da matriz A é matriz 1 × n, sendo considerado como vetor

$$Li = \begin{vmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{vmatrix} \tag{2}$$

e a j-ésima coluna da matriz A é matriz m × 1, também considerado como um vetor

$$Cj = \begin{vmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{vmatrix} \tag{3}$$

Podemos também dizer que uma matriz é formada por m vetores horizontais e n vetores verticais, normalmente, vemos a primeira forma, m vetores com n elementos cada.

Definição 1.1 (diagonal principal)**.** A *diagonal principal* de uma matriz quadrada A é formada pelos componentes a_{ij} onde $i = j$.

Exemplo 1.1.1. :

$$\begin{vmatrix} \mathbf{a_{11}} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & \mathbf{a_{22}} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & \mathbf{a_{33}} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & \mathbf{a_{mn}} \end{vmatrix} \tag{4}$$

Exemplo 1.1.2. :

$$\begin{vmatrix} \mathbf{1} & -45 & 10 \\ 4 & \mathbf{2} & 3 \\ 7 & 25 & \mathbf{3} \end{vmatrix} \tag{5}$$

Exemplo 1.1.3. :

$$\begin{vmatrix} \mathbf{2} & -45 & 10 & 3 \\ 4 & \mathbf{3} & 3 & 7 \\ 7 & 25 & \mathbf{8} & 0 \\ 1 & 2 & 5 & \mathbf{1} \end{vmatrix} \tag{6}$$

Definição 1.2 (matriz diagonal)**.** É uma matriz quadrada formada por componentes $a_{ij} = 0$ onde $i \neq j$.

Exemplo 1.2.1. :

$$\begin{vmatrix} \mathbf{a_{11}} & 0 & 0 & \cdots & 0 \\ 0 & \mathbf{a_{22}} & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{a_{33}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \mathbf{a_{mn}} \end{vmatrix} \tag{7}$$

Exemplo 1.2.2. :

$$\begin{vmatrix} \mathbf{3} & 0 & 0 \\ 0 & \mathbf{4} & 0 \\ 0 & 0 & \mathbf{2} \end{vmatrix} \tag{8}$$

Exemplo 1.2.3. :

$$\begin{vmatrix} \mathbf{3} & 0 & 0 & 0 \\ 0 & \mathbf{4} & 0 & 0 \\ 0 & 0 & \mathbf{0} & 0 \\ 0 & 0 & 0 & \mathbf{7} \end{vmatrix} \tag{9}$$

Definição 1.3 (Matriz Identidade)**.** A matriz identidade é uma matriz diagonal especial, formada por apenas 1s e 0s. Desse modo: $\forall\ a_{ij} \in \mathbb{R}$ se $i = j,\ a_{ij} = 1$ e se $i \neq j,\ a_{ij} = 0$.

Exemplo 1.3.1.

$$I_M = \begin{vmatrix} \mathbf{1} & 0 & 0 & \cdots & 0 \\ 0 & \mathbf{1} & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \mathbf{1} \end{vmatrix} \tag{10}$$

Exemplo 1.3.2.

$$I_2 = \begin{vmatrix} \mathbf{1} & 0 \\ 0 & \mathbf{1} \end{vmatrix} \tag{11}$$

Exemplo 1.3.3.

$$I_3 = \begin{vmatrix} \mathbf{1} & 0 & 0 \\ 0 & \mathbf{1} & 0 \\ 0 & 0 & \mathbf{1} \end{vmatrix} \tag{12}$$

Além das definições, podemos realizar operações básicas com matrizes como **Multiplicação por escalar** e **Soma de matrizes**. Vamos definir $\mathbf{A}_{mxn}$, com m linhas e n colunas e $\mathbf{B}_{mxn}$, com as mesmas dimensões de $\mathbf{A}$, $\forall \alpha \in \mathbb{R}$, ou seja, alpha pertencente ao conjunto dos números *Reais*, tal que:

$$\mathbf{A} = \begin{vmatrix} a_{11} & a_{21} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m1} & \cdots & a_{mn} \end{vmatrix} \tag{13}$$

$$\mathbf{B} = \begin{vmatrix} b_{11} & b_{21} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m1} & \cdots & b_{mn} \end{vmatrix} \tag{14}$$

Podemos definir $\alpha \cdot \mathbf{A}$ da seguinte forma:

$$\alpha \cdot \mathbf{A} = \alpha \cdot \begin{vmatrix} a_{11} & a_{21} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m1} & \cdots & a_{mn} \end{vmatrix} = \begin{vmatrix} \alpha \cdot a_{11} & \alpha \cdot a_{21} & \cdots & \alpha \cdot a_{1n} \\ \alpha \cdot a_{21} & \alpha \cdot a_{22} & \cdots & \alpha \cdot a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha \cdot a_{m1} & \alpha \cdot a_{m1} & \cdots & \alpha \cdot a_{mn} \end{vmatrix} \tag{15}$$

Como pode ver, a multiplicação de um escalar pela matriz dar-se termo a termo, multiplicado pelo escalar. Abaixo definimos a soma das matrizes **A** e **B** também termo a termo, da seguinte forma:

$$\mathbf{A}+\mathbf{B}=\begin{vmatrix} a_{11} & a_{21} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m1} & \cdots & a_{mn} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{21} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m1} & \cdots & b_{mn} \end{vmatrix} = \tag{16}$$

$$= \begin{vmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{vmatrix} \tag{17}$$

Abaixo seguem as propriedades para operações com matrizes e valores escalares.

Propriedade 1.1 (Propriedade Associativa)**.** $(A+B)+C = A+(B+C),\ \ \forall\ A,\ B,\ C \in \mathbb{R}^{m\times n}$ (para todos).

Propriedade 1.2 (Matriz Nula)**.** Existe $O \in \mathbb{R}^{m\times n}$ tal que $A+O=A,\ \forall\ A \in \mathbb{R}^{m\times n}$.

Propriedade 1.3 (Matriz Oposta)**.** $\forall A \in \mathbb{R}^{m\times n}$, existe $-A \in \mathbb{R}^{m\times n}$ tal que $A+(-A)=O$, onde $-A=[-a_{ij}]$.

Propriedade 1.4 (Propriedade da Soma)**.** $A+B=B+A$, $\forall\ A,\ B \in \mathbb{R}^{m\times n}$.

Propriedade 1.5 (Propriedade Distributiva da Multiplicação)**.** $(a+b)A = aA+bA$, $\forall\ a,b \in \mathbb{R}$ e $A \in \mathbb{R}^{m\times n}$.

Propriedade 1.6 (Propriedade Distributiva da Multiplicação 2)**.** $a(A+B)=aA+aB$, $\forall\ a \in R$ e $A,\ B \in \mathbb{R}^{m\times n}$.

Propriedade 1.7 (Propriedade da igualdade)**.** $a(A+B)=aA+aB$, $\forall\ a \in R$ e $A,\ B \in \mathbb{R}^{m\times n}$.

Propriedade 1.8 (Propriedade da igualdade)**.** $1 \cdot A = A$, $\forall\ A \in \mathbb{R}^{m\times n}$.

1.1 Multiplicação de matrizes

Sejam $A=[a_{ij}] \in \mathbb{R}^{m\times n}$ e $B=[b_{ij}] \in \mathbb{R}^{n\times p}$. O produto de A por B, em símbolos, AB, é definido como:

$$A_{\ 3\times 2} \times B_{\ 2\times 1} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \times \begin{vmatrix} b_{11} \\ b_{21} \end{vmatrix} = \begin{vmatrix} (a_{11}\cdot b_{11})+(a_{12}\cdot b_{21}) \\ (a_{21}\cdot b_{11})+(a_{22}\cdot b_{21}) \\ (a_{31}\cdot b_{11})+(a_{32}\cdot b_{21}) \end{vmatrix} \tag{18}$$

A operação pode parecer desafiadora a primeira vista, porém o passo a passo a ser seguido é simples

O mais importante é ter atenção nas dimensões das matrizes, dim(A) é 3x2, e dim(B) é 2x1, para podermos multiplicá-las o número de colunas da primeira deve ser igual ao número de linhas da segunda, em termos matemáticos é comum ver: $A_{m\times n}\cdot B_{n\times p}$, o que resulta numa matriz Cm×p, ou seja, nesse exemplo a matriz resultante tem dimensões dim(C) = 3x1.

O passo a passo a ser seguido é:

1. Identificar a primeira linha da primeira matriz e a primeira coluna da segunda matriz
2. Multiplicar o primeiro elemento da linha com o primeiro elemento da coluna
3. Somar as multiplicações até o n-ésimo termo das matrizes
4. Repetir com o restante da matriz

Exemplo 1.3.4.

$$A_{\ 3\times 2} \times B_{\ 2\times 1} = \begin{vmatrix} 3 & -1 \\ 0 & 2 \\ 1 & 1 \end{vmatrix} \times \begin{vmatrix} 1 \\ 2 \end{vmatrix} = \begin{vmatrix} (3 \cdot 1) + (-1 \cdot 2) \\ \\ \\ \end{vmatrix} \tag{19}$$

$$A_{\ 3\times 2} \times B_{\ 2\times 1} = \begin{vmatrix} 3 & -1 \\ 0 & 2 \\ 1 & 1 \end{vmatrix} \times \begin{vmatrix} 1 \\ 2 \end{vmatrix} = \begin{vmatrix} 1 \\ (0 \cdot 1) + (2 \cdot 2) \\ \\ \end{vmatrix} \tag{20}$$

$$A_{\ 3\times 2} \times B_{\ 2\times 1} = \begin{vmatrix} 3 & -1 \\ 0 & 2 \\ 1 & 1 \end{vmatrix} \times \begin{vmatrix} 1 \\ 2 \end{vmatrix} = \begin{vmatrix} 1 \\ 4 \\ (1 \cdot 1) + (1 \cdot 2) \end{vmatrix} \tag{21}$$

$$A_{\ 3\times 2} \times B_{\ 2\times 1} = \begin{vmatrix} 3 & -1 \\ 0 & 2 \\ 1 & 1 \end{vmatrix} \times \begin{vmatrix} 1 \\ 2 \end{vmatrix} = \begin{vmatrix} 1 \\ 4 \\ 3 \end{vmatrix} \tag{22}$$

Desse modo, se AxB for uma multiplicação válida não implica que BxA também a seja, apenas se A e B forem matrizes quadradas

1.2 Determinante de uma matriz

Definição 1.4 (Determinante)**.** É um número associado à matriz quadrada, que envolve operações incluindo todos os elementos desta

Exemplo 1.4.1 (Determinante de matriz 2x2)**.** *Simplesmente multiplicamos os valores da diagonal principal e subtraimos o valor da multiplicação dos elementos restantes*

$$det[A] = det \begin{vmatrix} 1 & 2 \\ 2 & 0 \end{vmatrix} = (1 \cdot 0) - (2 \cdot 2) = -4 \tag{23}$$

Exemplo 1.4.2 (Determinante de matriz 3x3)**.** Método de Sarrus

$$det[A] = det \begin{vmatrix} 1 & 2 & 3 \\ 2 & 0 & 1 \\ 0 & 1 & 1 \end{vmatrix} \tag{24}$$

"Duplica-se" as duas primeiras colunas a direita da matriz para facilitar o agrupamento dos elementos

$$\left|\begin{array}{ccc|cc} 1 & 2 & 3 & 1 & 2 \\ 2 & 0 & 1 & 2 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{array}\right| \tag{25}$$

Multiplica-se as três diagonais da esquerda para a direita e as três da direita para a esquerda. Por fim, soma-se os três primeiros resultados e subtrai-se a soma dos outros três.

$$\Big[(1 \cdot 0 \cdot 1) + (2 \cdot 1 \cdot 0) + (3 \cdot 2 \cdot 1)\Big] - \Big[(2 \cdot 2 \cdot 1) + (1 \cdot 1 \cdot 1) + (3 \cdot 0 \cdot 0)\Big] \tag{26}$$

$$\Big(0 + 0 + 6\Big) - \Big(4 + 1 + 0\Big) = 6 - 5 = 1 \tag{27}$$

1.3 Outras operações de matrizes

Definição 1.5 (Matriz transposta)**.** A matriz transposta de uma matriz é obtida escrevendo as linhas da matriz como colunas, ou simplesmente,

Exemplo 1.5.1.

$$A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \quad A^T = \begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \end{vmatrix} \tag{28}$$

Definição 1.6 (Matriz Adjunta)**.** A matriz adjunta de uma matriz A - adj(A) é a transposta da matriz dos cofatores A_{ij} de A.

Definição 1.7 (Matriz Inversa)**.** A matriz inversa tem a propriedade de quando multiplicada a matriz base resultar na matriz identidade.

$$A^{-1} \times A = I \tag{29}$$

Ela é definida como:

$$A^{-1} = \frac{1}{det(A)} \cdot adj(A) \tag{30}$$

Há outra forma de adquirir a matriz inversa, de maneira bem mais simples e intuitiva:

Exemplo 1.7.1. Começamos com o seguinte sistema linear: (Sendo A uma matriz quadrada e I a matriz identidade de mesma dimensão que A)

$$A \times A^{-1} = I \tag{31}$$

Sabemos os valores de A, sabemos também os valores de I, então usaremos as regras simples da multiplação de matriz para achar A^{-1}

$$A \times A^{-1} = \begin{vmatrix} 3 & -1 & 0 \\ 0 & 2 & 1 \\ 1 & 1 & 0 \end{vmatrix} \times \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \tag{32}$$

O resultado da multiplicação é esse:

$$\begin{vmatrix} (3 \cdot a_{11} + (-1) \cdot a_{21} + 0 \cdot a_{31}) & (3 \cdot a_{12} + (-1) \cdot a_{22} + 0 \cdot a_{32}) & (3 \cdot a_{13} + (-1) \cdot a_{23} + 0 \cdot a_{33}) \\ (0 \cdot a_{11} + 2 \cdot a_{21} + 1 \cdot a_{31}) & (0 \cdot a_{12} + 2 \cdot a_{22} + 1 \cdot a_{32}) & (0 \cdot a_{13} + 2 \cdot a_{23} + 1 \cdot a_{33}) \\ (1 \cdot a_{11} + 1 \cdot a_{21} + 0 \cdot a_{31}) & (1 \cdot a_{12} + 1 \cdot a_{22} + 0 \cdot a_{32}) & (1 \cdot a_{13} + 1 \cdot a_{23} + 0 \cdot a_{33}) \end{vmatrix} \tag{33}$$

Agora igualamos cada coluna da matriz multiplicação a cada coluna da matriz identidade, formando os três sistemas lineares.
Primeira coluna:

$$\begin{vmatrix} (3 \cdot a_{11} + (-1) \cdot a_{21} + 0 \cdot a_{31}) = 1 \\ (0 \cdot a_{11} + 2 \cdot a_{21} + 1 \cdot a_{31}) = 0 \\ (1 \cdot a_{11} + 1 \cdot a_{21} + 0 \cdot a_{31}) = 0 \end{vmatrix} \tag{34}$$

Segunda coluna:

$$\begin{vmatrix} (3 \cdot a_{12} + (-1) \cdot a_{22} + 0 \cdot a_{32}) = 0 \\ (0 \cdot a_{12} + 2 \cdot a_{22} + 1 \cdot a_{32}) = 1 \\ (1 \cdot a_{12} + 1 \cdot a_{22} + 0 \cdot a_{32}) = 0 \end{vmatrix} \tag{35}$$

Terceira coluna:

$$\begin{vmatrix} (3 \cdot a_{13} + (-1) \cdot a_{23} + 0 \cdot a_{33}) = 0 \\ (0 \cdot a_{13} + 2 \cdot a_{23} + 1 \cdot a_{33}) = 0 \\ (1 \cdot a_{13} + 1 \cdot a_{23} + 0 \cdot a_{33}) = 1 \end{vmatrix} \tag{36}$$

1.4 Exemplos

Com tantos conceitos expostos, tantas definições, fica a pergunta: Onde eu usarei isso? Aqui está um exemplo claro do uso da matemática, de modo bem simples, de um método estatístico muito frequente para resolução de problemas lineares: A Regressão Linear.

1.4.1 Regressão Linear

Na estatística, a regressão é uma equação utilizada para estimar valores para uma variável Y (Variável dependente ou explicada), dado um valor para X (Variável independente ou explicativa). A regressão linear tem esse nome "linear", pelo fato de querer assumir uma relação Linear entre as variáveis, ou seja, Y é uma função linear, tendo alguns valores como parâmetros. A equação abaixo define a expressão para regressão linear múltipla, escrita em sua forma matricial:

$$Y = X \cdot \hat{\beta}$$

Onde podemos definir Y como uma matriz n **x** 1, ou também, como simplesmente um vetor com n entradas; X, uma matriz de dimensões n **x** $p+1$ (sendo a primeira coluna preenchida sempre com 1s); β, uma matriz de dimensões $p+1$ x 1, ou um vetor com $p+1$ entradas.

$$Y = \begin{vmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{vmatrix}, \; X = \begin{vmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1p} \\ 1 & X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{np} \end{vmatrix}, \; \vec{\beta} = \begin{vmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{vmatrix} \tag{37}$$

Dado que conhecemos os valores de X e queremos estimar Y, apenas nos resta estimar os β, então utilizamos a seguinte expressão para o caso da regressão múltipla:

$$\hat{\beta} = (X^T \cdot X)^{-1} \cdot X^T \cdot Y \tag{38}$$

Exemplo 1.7.2. Uma equipe de segurança de redes, decidiu desenvolver várias formas de conter ataques aos servidores e rede, e então decidiram encontrar uma forma de avaliar os mecanismos e estimar um valor para o índice de sucesso dos meios desenvolvidos. Para isso foi desenvolvido um **Índice de sucesso** baseado em dois fatores:

- Tempo de experimento (duração);
- Número de ataques no período.

Logo em seguida, definindo como seria feito o experimento, observaram os seguintes dados amostrais:

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

Logo, dado os dados amostrais, podemos definir Y e X da seguinte forma matricial:

$$X = \begin{vmatrix} 1 & 5 & 118 \\ 1 & 13 & 132 \\ 1 & 20 & 119 \\ 1 & 28 & 153 \\ 1 & 41 & 91 \\ 1 & 49 & 118 \\ 1 & 61 & 132 \\ 1 & 62 & 105 \end{vmatrix}, Y = \begin{vmatrix} 8.1 \\ 6.8 \\ 7.0 \\ 7.4 \\ 7.7 \\ 7.5 \\ 7.6 \\ 8.0 \end{vmatrix}$$

Note agora que além desses valores, precisamos realizar operações para a estimativa de β, que são evidenciadas pela equação (38). Vamos por partes, primeiro vamos encontrar a transposta de X:

$$X = \begin{vmatrix} 1 & 5 & 118 \\ 1 & 13 & 132 \\ 1 & 20 & 119 \\ 1 & 28 & 153 \\ 1 & 41 & 91 \\ 1 & 49 & 118 \\ 1 & 61 & 132 \\ 1 & 62 & 105 \end{vmatrix} \Longrightarrow X^T = \begin{vmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 5 & 13 & 20 & 28 & 41 & 49 & 61 & 62 \\ 118 & 132 & 119 & 153 & 91 & 118 & 132 & 105 \end{vmatrix}$$

Então, vamos calcular primeiro $(X^T \cdot X)^{-1}$, uma multiplicação matricial como na seção 1.1, logo em seguida, basta calcularmos a inversa do resultado, como na seção 1.3.

$$X^T \cdot X = \begin{vmatrix} 8 & 279 & 968 \\ 279 & 13025 & 33045 \\ 968 & 33045 & 119572 \end{vmatrix} \Longrightarrow (X^T \cdot X)^{-1} = \begin{vmatrix} 7.7134 & -0.0227 & -0.0562 \\ -0.0227 & 0.0003 & 0.0001 \\ -0.0562 & 0.0001 & 0.0004 \end{vmatrix}$$

E então, calculamos agora $X^T \cdot Y$.

$$X^T \cdot Y = \begin{vmatrix} 60.1 \\ 2118.9 \\ 7247.5 \end{vmatrix}$$

Por fim, temos que $\hat{\beta}$ é estimado da seguinte forma:

$$\hat{\beta} = (X^T \cdot X)^{-1} \cdot X^T \cdot Y = \begin{vmatrix} 8.37 \\ 0.005 \\ -0.009 \end{vmatrix}$$

Agora que está estimado o $\hat{\beta}$, podemos definir o $\hat{Y}$ estimado pela regressão da seguinte forma:

$$\hat{Y} = X \cdot \hat{\beta} = \begin{vmatrix} 1 & 5 & 118 \\ 1 & 13 & 132 \\ 1 & 20 & 119 \\ 1 & 28 & 153 \\ 1 & 41 & 91 \\ 1 & 49 & 118 \\ 1 & 61 & 132 \\ 1 & 62 & 105 \end{vmatrix} \cdot \begin{vmatrix} 8.37 \\ 0.005 \\ -0.009 \end{vmatrix} = \begin{vmatrix} 7.4 \\ 7.3 \\ 7.4 \\ 7.2 \\ 7.8 \\ 7.6 \\ 7.5 \\ 7.8 \end{vmatrix}$$

1.5 Exercicios

Exemplo 1.7.3 (Soma, Subtração e Multiplicação por escalar)**.** Dadas as matrizes A, B e C

$$A = \begin{vmatrix} 1 & 2 & 3 \\ -4 & 1 & 2 \\ 1 & 1 & 2 \end{vmatrix} \quad B = \begin{vmatrix} 2 & 0 & 3 \\ 5 & 4 & 6 \\ 5 & 1 & 7 \end{vmatrix} \quad C = \begin{vmatrix} 1 & 3 & 4 \\ 1 & 0 & 2 \\ 1 & 2 & 4 \end{vmatrix} \tag{39}$$

Determine:

1. A + B
2. B + C
3. A + B - C
4. 2A - B + 3C

Exemplo 1.7.4 (Multiplicação de matrizes)**.** Dadas as matrizes A, B e C

$$A = \begin{vmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 1 \end{vmatrix} \quad B = \begin{vmatrix} 2 & 1 & 3 \\ 1 & 0 & 2 \end{vmatrix} \quad C = \begin{vmatrix} 2 & 3 & 1 \\ 2 & 0 & 2 \\ 1 & 2 & 4 \end{vmatrix} \tag{40}$$

Determine:

1. $A \times B$
2. $B \times A$
3. $C \times A$
4. $A \times B \times C$

Exemplo 1.7.5 (Determinante)**.** Dadas as matrizes A, B e C

$$A = \begin{vmatrix} 1 & 2 & 3 \\ -4 & 1 & 2 \\ 1 & 1 & 2 \end{vmatrix} \quad B = \begin{vmatrix} 2 & 0 & 3 \\ 5 & 4 & 6 \\ 5 & 1 & 7 \end{vmatrix} \quad C = \begin{vmatrix} 1 & 3 & 4 \\ 1 & 0 & 2 \\ 1 & 2 & 4 \end{vmatrix} \tag{41}$$

Determine:

1. $det(A)$
2. $det(B)$
3. $det(C)$
4. $det(A \times B)$

Exemplo 1.7.6 (Inversa)**.** Dadas as matrizes A, B e C

$$A = \begin{vmatrix} 1 & 2 & 3 \\ -4 & 1 & 2 \\ 1 & 1 & 2 \end{vmatrix} \quad B = \begin{vmatrix} 2 & 0 & 3 \\ 5 & 4 & 6 \\ 5 & 1 & 7 \end{vmatrix} \quad C = \begin{vmatrix} 1 & 3 & 4 \\ 1 & 0 & 2 \\ 1 & 2 & 4 \end{vmatrix} \tag{42}$$

Determine:

1. A^{-1}
2. B^{-1}
3. C^{-1}
4. $(A \times B)^{-1}$

2 Introdução à Geometria Analítica

2.1 Vetores Unitários Ortogonais

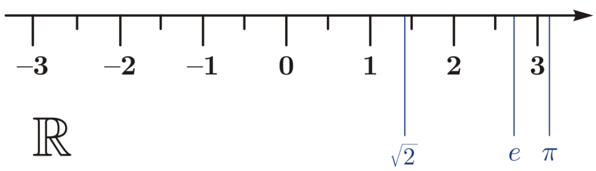
Vamos definir $\hat{\imath}$, $\hat{\jmath}$ e $\hat{k}$ como os vetores unitários ortogonais, ou seja com comprimento de uma unidade de medida (1 u.a) e ortogonais entre eles. Vamos definir a ortogonalidade como vetores que formam um angulo de 90 entre eles, ou seja, $\hat{\imath} \perp \hat{\jmath}$, $\hat{\jmath} \perp \hat{k}$ e $\hat{\imath} \perp \hat{k}$, onde o simbolo $\perp$ significa que um vetor é perpedincular ou ortogonal a outro.

2.2 Sistemas de Coordenadas Cartesianas

O espaço cartesiano é construído por vetores unitários ortogonais entre si. Normalmente estes são chamados $\hat{\imath}$, $\hat{\jmath}$ e $\hat{k}$. Sendo relacionados aos eixos x, y e z, respectivamente. De forma mais simples, podemos dizer que o sistema cartesiano é uma ferramenta que nos auxilia a encontrar um objeto ou grupo de informações em um espaço de até n-dimensões. O Sistema de Coordenadas Cartesianas é largamente usado por diversas áreas como matemática, física e astronomia, além de outras, como a química.

2.2.1 Reta dos reais ($\mathbb{R}$)

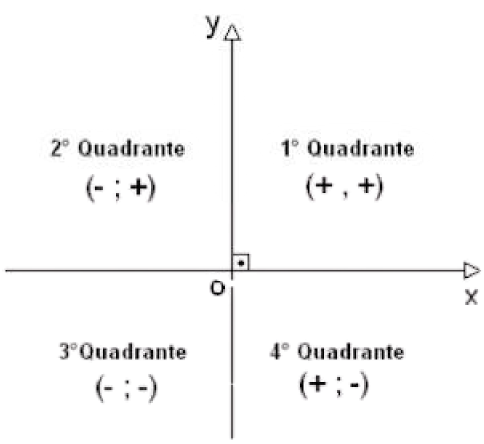
A reta do conjunto dos Reais ($\mathbb{R}$) é o espaço que compreende a primeira dimensão. Nele temos somente as operações simples já conhecidas na álgebra. A distância entre pontos, por exemplo, sempre é trivial, como veremos abaixo. Podemos representar um vetor na reta dos reais como um **segmento orientado da reta**.



Note que, o valor 0 divide a reta dos reais, negativos para a esquerda ($\longleftarrow$) e para a direita, os positivos ($\longrightarrow$).

2.2.2 Espaço Bidimensional ($\mathbb{R}^2$)

No espaço bidimensional, vemos o básico da geometria, a distância entre dois pontos, ou operações de vetores não são mais tão simples, mas na maioria das vezes envolvem somente o Teorema de Pitágoras ou outras operações também bastante conhecidas.

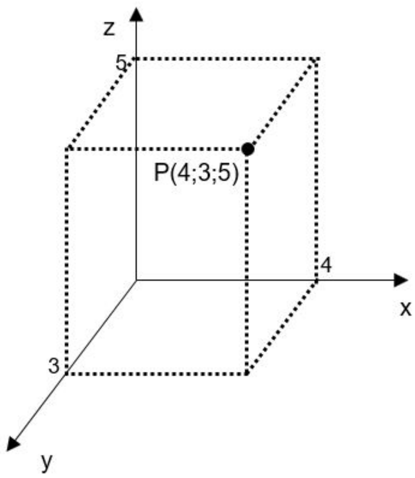


Perceba também que o espaço bidimensional é composto por duas Retas Reais, que possuem uma interseção, determinada como **origem**, suas coordenadas são definidas como $x = 0$ e $y = 0$, ou simplesmente o par de coordenadas (0,0). Note também que esse espaço bidimensional pode ser tratado também como um **Plano infinito**. Além disso, em um espaço bidimensional podemos definir 4 quadrantes fundamentais:

- **Quadrante x+, y+**
- **Quadrante x+, y-**
- **Quadrante x-, y+**
- **Quadrante x-, y-**

2.2.3 Espaço Tridimensional ($\mathbb{R}^3$)

No espaço tridimensional avançamos mais um passo de complexidade, porém ainda temos um objeto de estudo facilmente compreensível espacialmente para o cérebro humano. Podemos resolver muitas situações ainda utilizando o Teorema de Pitágoras.



Ao avançarmos a complexidade do espaço, note que podemos interpretar um ponto em seu espaço como (x,y,z). Note que um espaço tridimensional é composto por retas, geradas a partir da interseção de 3 planos, uma ideia que será explicada mais à frente, que são:

- **(0, y, z)** ;
- **(x, 0, z)** ;
- **(x, y, 0)** .

Espaços *n-dimensionais:* Falando em n-dimensões temos situações que fogem da compreensão humana (tente imaginar um hipercubo em 4 dimensões, por exemplo), tratamos esses exemplos, na maioria das vezes, de forma puramente matemática, e é então onde a álgebra linear é mais aplicada, já que a maioria dos problemas reais que encontraremos possuem inúmeras variáveis independentes de entrada com uma ou mais saídas.

2.3 Sistema de Coordenadas Esféricas e Cilíndricas

Não precisamos necessariamente tratar uma posição no espaço tridimensional (ou em quaisquer n-dimensões) de forma cartesiana (x, y, z). Podemos, utilizar de ângulos para adquirir o mesmo resultado, como por exemplo no sistema de coordenadas esféricas (r, α, θ), que se referem ao raio, ou distância da origem ao ponto, ângulo num primeiro plano e ângulo no segundo plano.

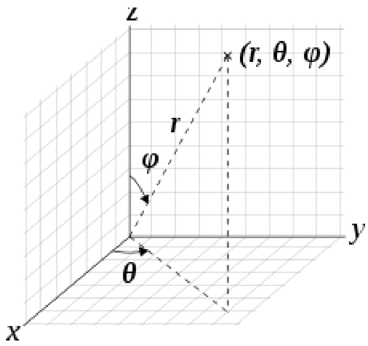


Figura 1: Coordenadas Esféricas

Já no sistema cilíndrico temos (r, Z, θ), onde r é a distância do ponto a origem no primeiro plano, Z a “altura” ou a distância ortogonal do ponto ao segundo plano e θ o ângulo no primeiro plano.

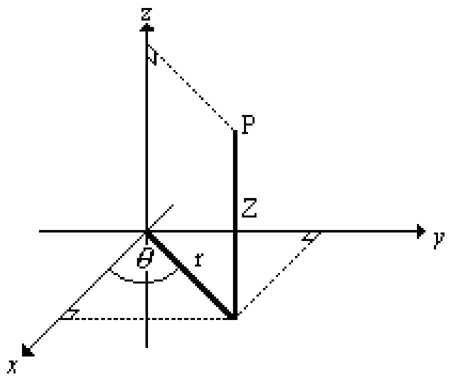


Figura 2: Coordenadas Cilíndricas

2.4 Pontos, Retas e Planos

2.4.1 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**.

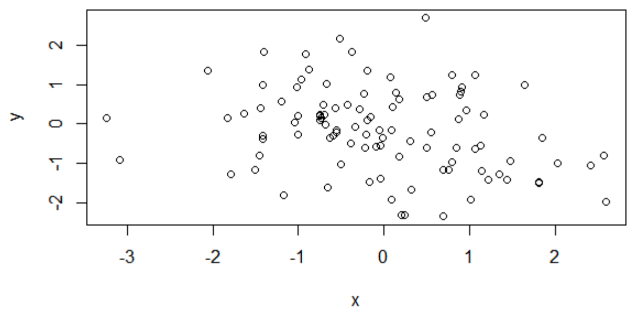


Figura 3: Gráfico de Pontos

2.4.2 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.

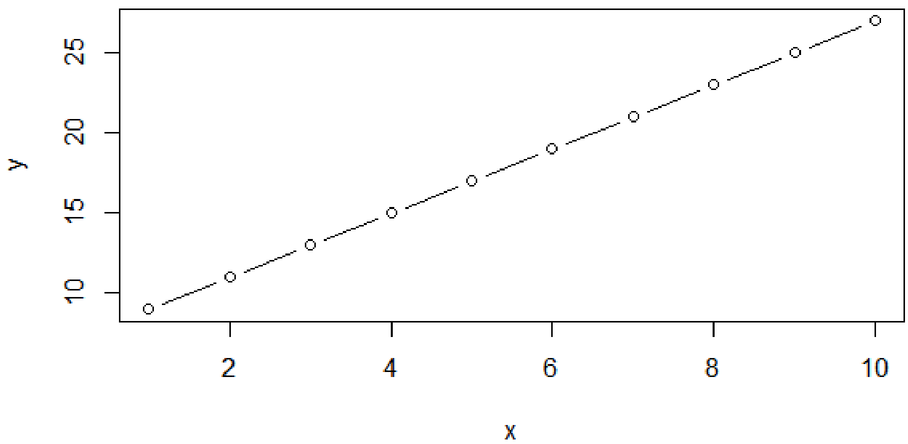


Figura 4: Gráfico da Reta e sua composição de pontos

2.4.3 Planos

O plano é um conjunto de retas e pontos, de modo que apenas um único plano passa por duas retas paralelas não coincidentes.

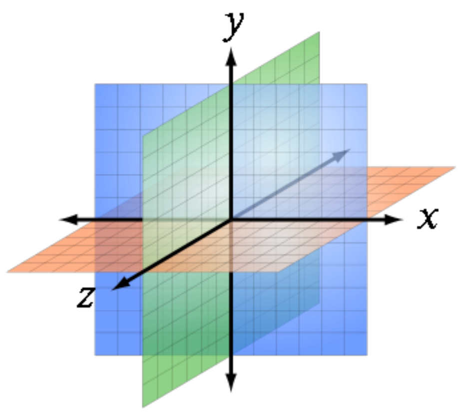


Figura 5: Planos principais

Outra forma de definir um plano é a partir de uma reta e um ponto não coincidentes (i.e. o ponto não pertence à reta). Podemos ter num plano infinitas retas e pontos. De modo contrário aos últimos dois itens, a interseção entre dois planos não paralelos gera uma única reta. Entre três planos define-se um único ponto ou três retas, dependendo da forma que se encontrarem.

2.5 Distância

2.5.1 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,

$$P = (x_1, x_2, x_3, \ldots, x_n)$$

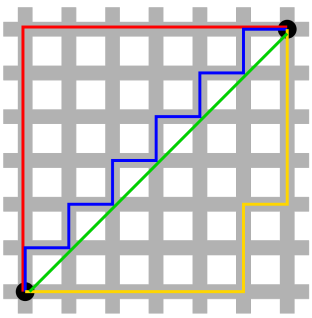
$$Q = (y_1, y_2, y_3, \ldots, x_n)$$

Onde n é o número de dimensões. Então, podemos definir os seguintes casos mais utilizados:

- **Unidimensional:**
$$d(P,Q) = \sqrt{(x_1 - y_1)^2} = |x_1 - y_1|$$
- **Bidimensional:**
$$d(P,Q) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$
- **Tridimensional:**
$$d(P,Q) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}$$
$$\vdots$$
- **n-dimensional:**
$$d(P,Q) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + \cdots + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

2.5.2 Distância Manhattan

Também é conhecida como geometria do táxi.



Imagine um táxi se movendo por uma cidade, a distância prática entre 2 pontos não é a reta que os liga, mas sim quantos quarteirões no eixo x mais a quantidade de quarteirões no eixo y o táxi percorreu.

Definimos P_0 como o vetor posição inicial e P_1 o vetor posição final.

- **Unidimensional:** Defina $P_0 = (x_0)$ e $P_1 = (x_1)$
$$d(P_0, P_1) = |x_1 - x_0|$$

- **Bidimensional:** Defina $P_0 = (x_0, y_0)$ e $P_1 = (x_1, y_1)$, então

$$d(P_0, P_1) = |x_1 - x_0| + |y_1 - y_0|$$

- **Tridimensional:** Defina $P_0 = (x_0, y_0, z_0)$ e $P_1 = (x_1, y_1, z_1)$

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

$$\vdots$$

- **n-dimensional:** Defina $P_0 = (x_0, y_0, \dots, w_0)$, $P_1 = (x_1, y_1, \dots, w_1)$

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

3 Métricas de avaliação de modelos

3.1 Métricas de avaliação de modelos de Regressão

Vimos mais definições e mais fórmulas, e agora vamos a questão de onde iremos utilizar. Abaixo veremos alguns meios de avaliação do modelo de regressão visto na última seção. Para as métricas abaixo, vamos utilizar a **distância euclidiana**.

3.1.1 Coeficiente de Determinação (R^2):

O R-quadrado avalia a variabilidade da variável dependente (Y) que é explicada pelo nosso modelo, baseado na variável independente (X). Nele temos $\bar{Y}$, que é a média dos valores dependentes da nossa base:

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

y_i, cada valor real da variável dependente e $\hat{y}_i$ cada valor estimado para y.

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

3.1.2 Erro Quadrático Médio (MSE):

O MSE é a métrica muito utilizada, baseado na distância da estimativa do nosso modelo baseado nos verdadeiros valores dos dados, dado da seguinte forma:

$$MSE = \frac{1}{n} \cdot \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \frac{1}{n} \cdot d(Y, \hat{Y})^2$$

Onde Y é o vetor formado pelos valores da variável dependente e $\hat{Y}$ os valores estimados por nosso modelo. Por fim n é a dimensão, ou a quantidade dos valores de Y, ou seja, $Y = (y_1, y_2, \dots, y_n)$ e $\hat{Y} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n)$.

3.1.3 Raiz do Erro Quadrático Médio (RMSE):

Como forma de correção de unidade de medida, a RMSE surge, corrigindo este fator da unidade de medida. Assim como o anterior, ambos penalizam estimativas fora da realidades. Um exemplo desta correção é dado da seguinte forma: Digamos que temos um modelo que quer estimar o gasto em reais de uma familia, dado o número de membros, supondo uma linearidade nessa dependencia, digamos que o nosso MSE é igual $MSE = \frac{1}{n} \cdot d(Y, \hat{Y})^2 = (143,00 \quad reais)^2$. Perceba que tanto o valor e a unidade de medida dele está elevado ao quadrado, para isso temos o $RMSE$, retirando a raiz desse valor, e deixando o erro como mesma unidade de medida que Y.

$$RMSE = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} = \sqrt{\frac{1}{n} \cdot d(Y, \hat{Y})^2} = \frac{1}{\sqrt{n}} \cdot |d(Y, \hat{Y})|$$

3.1.4 Erro Absoluto Médio (MAE):

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

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

3.2 Exercícios

3.2.1 Distância Euclidiana

Determine as distâncias entre os pontos pelo método Euclidiano

- P(1, 2) e Q(3, 3)

  Resolução:

  Um modo simples é calcular as distâncias de cada eixo (assim como em Manhattan) |(1, 2) - (3, 3)| = (2, 1). agora utilizamos Pitágoras: $d = \sqrt{2^2 + 1^2}$, ou seja, d = 5.

  Ou simplesmente $d = \sqrt{(P_x - Q_x)^2 + (P_y - Q_y)^2}$, então $d = \sqrt{(1-3)^2 + (2-3)^2}$
- P(3, 0, 5) e Q(1, 0, 2)
- P(1, 0, 1) e Q(-2, 0, -1)
- P(1, 2, 3, 4) e Q(3, 3, 0, 3)
- P(3, -1, 0, 2) e Q(1, 1, 4, 1)

3.2.2 Distância Manhattan

Determine agora as distâncias entre os mesmos pontos pelo método Manhattan

- P(1, 2) e Q(3, 3)

  Resolução:

  Como foi visto |(1, 2) - (3, 3)| = (2, 1). Agora somamos os 2 valores, $d_M = 2 + 1 = 3$

  Ou simplesmente $d_M = |P_x - Q_x| + |P_y - Q_y|$, então $d_M = |1 - 3| + |2 - 3|$
- P(3, 0, 5) e Q(1, 0, 2)
- P(1, 0, 1) e Q(-2, 0, -1)
- P(1, 2, 3, 4) e Q(3, 3, 0, 3)
- P(3, -1, 0, 2) e Q(1, 1, 4, 1)

3.2.3 Modelo para avaliação

Utilizando os valores da regressão da seção anterior:

$$Y = \begin{vmatrix} 8.1 \\ 6.8 \\ 7.0 \\ 7.4 \\ 7.7 \\ 7.5 \\ 7.6 \\ 8.0 \end{vmatrix} \qquad \hat{Y} = \begin{vmatrix} 7.4 \\ 7.3 \\ 7.4 \\ 7.2 \\ 7.8 \\ 7.6 \\ 7.5 \\ 7.8 \end{vmatrix}$$

Encontre os valores:

- R^2
- MSE
- RMSE
- MAE

4 Introdução a Funções