ÁLGEBRA LINEAR PARA DATA SCIENCE E MACHINE LEARNING

EXPERT ACADEMY

CONTEÚDO DO CURSO

- Definição – equação linear
- Escalares, vetores, matrizes, tensores
- Sistemas lineares
- Vetores
- Matrizes
- Operações
- Transformações
- Aplicações

NumPy

DEFINIÇÃO

- **Álgebra linear** é uma subárea da **Álgebra**
- A **Álgebra** trata de problemas matemáticos envolvendo **variáveis**

variáveis

constantes

$$\left.\begin{aligned} &x + 2 = 5 \\ &4x - \frac{1}{2}y^2 = 1 \\ &7x + \sqrt{5}y^2 - \frac{1}{3}\log z = 10 \end{aligned}\right\} \textbf{equações}$$

- Objetivo da Álgebra é encontrar os valores das variáveis ➔ **resolver as equações**

DEFINIÇÃO

- **Álgebra linear** trata de problemas matemáticos envolvendo **variáveis lineares**

  - ☑ $x + 2 = 5$
  - ☒ $4x - \frac{1}{2}y^2 = 1$
  - ☒ $7x + \sqrt{5}y^2 - \frac{1}{3}\log z = 10$
  - ☑ $4x - \frac{1}{2}y = 1$
  - ☑ $7x + \sqrt{5}y - \frac{1}{3}z = 10$

- **Equações lineares** podem ser representadas como **retas** (linhas)

EQUAÇÃO LINEAR

$$4x - \frac{1}{2}y = 1$$

$$\downarrow$$

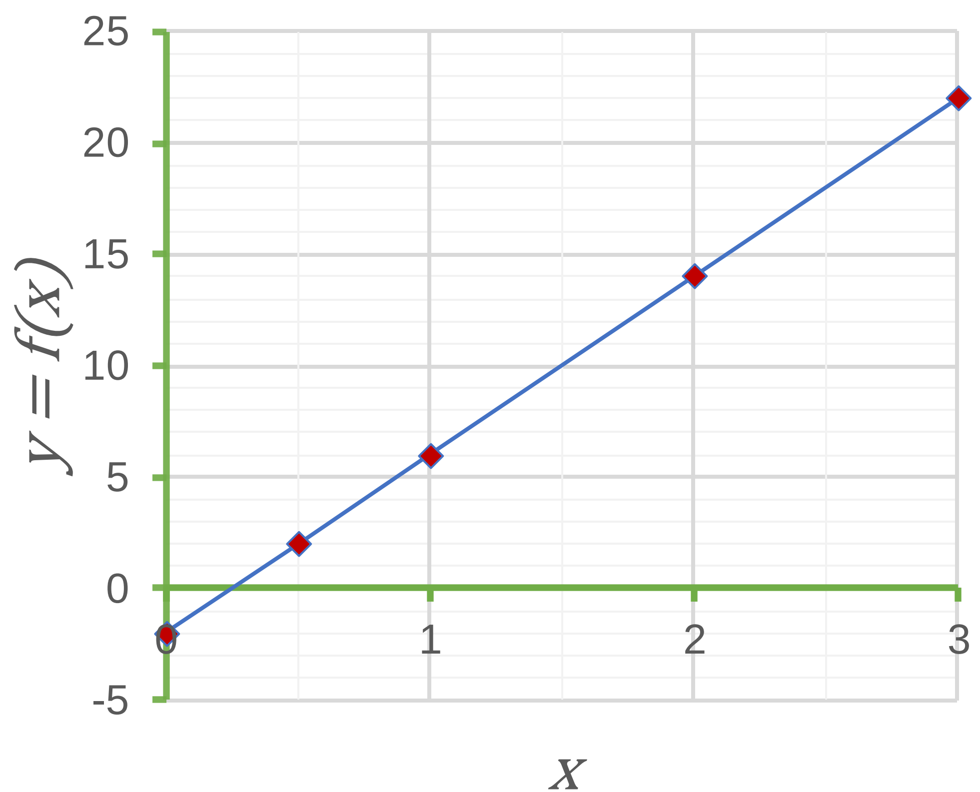
$$-\frac{1}{2}y = -4x + 1$$

$$\downarrow$$

$$y = 8x - 2 = f(x)$$

$$\downarrow$$

equação de uma reta



EQUAÇÃO LINEAR

variável dependente

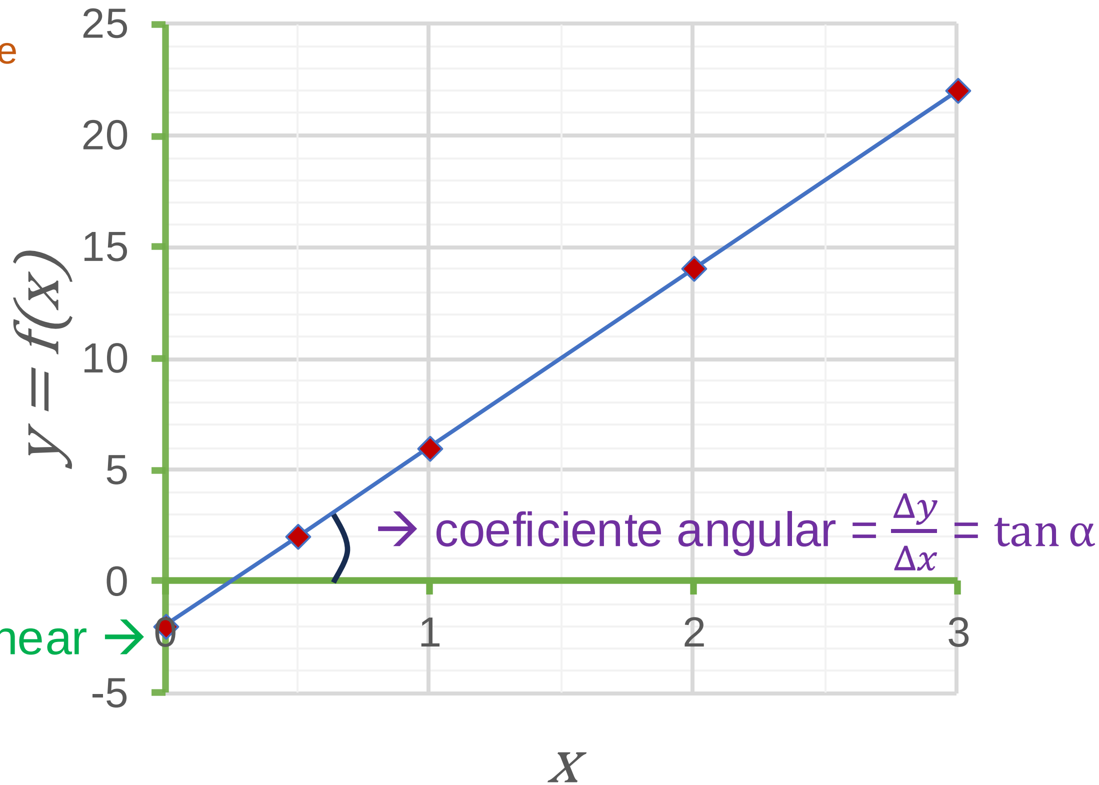
variável independente

$$y = ax + b$$

coeficiente angular

coeficiente linear (intercepto)

$$y = 8x - 2$$



EQUAÇÃO LINEAR

- **Equações lineares com mais de uma variável independente**

$$y = a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n + b$$

$a_1, a_2, a_3, \ldots, a_n$ ➔

coeficientes angulares de cada variável $x_1, x_2, x_3, \ldots, x_n$

b ➔ coeficiente linear da reta

EQUAÇÃO LINEAR

$$y = 7.5x_1 + 8.25x_2 + 35.8x_3 + 15x_4 + 10$$

b = R$ 10 = couvert artístico

a_4 = R$ 15 = preço da sobremesa
x_4 = quantidade de sobremesas

a_3 = R$ 35.80 = preço do prato principal
x_3 = quantidade de pratos principais

a_2 = R$ 8.25 = preço da entrada
x_2 = quantidade de entradas

a_1 = R$ 7.50 = preço da bebida
x_1 = quantidade de bebidas

DESPESA TOTAL

ESCALARES, VETORES, MATRIZES, TENSORES

- **Escalares**

$$y = \boxed{7.5}x_1 + \boxed{8.25}x_2 + \boxed{35.8}x_3 + \boxed{15}x_4 + \boxed{10}$$

float *integer*

$$x + 2 = 5$$

$$4x - \frac{1}{2}y = 1$$

$$7x + \sqrt{5}y - \frac{1}{3}z = 10$$

ESCALARES, VETORES, MATRIZES, TENSORES

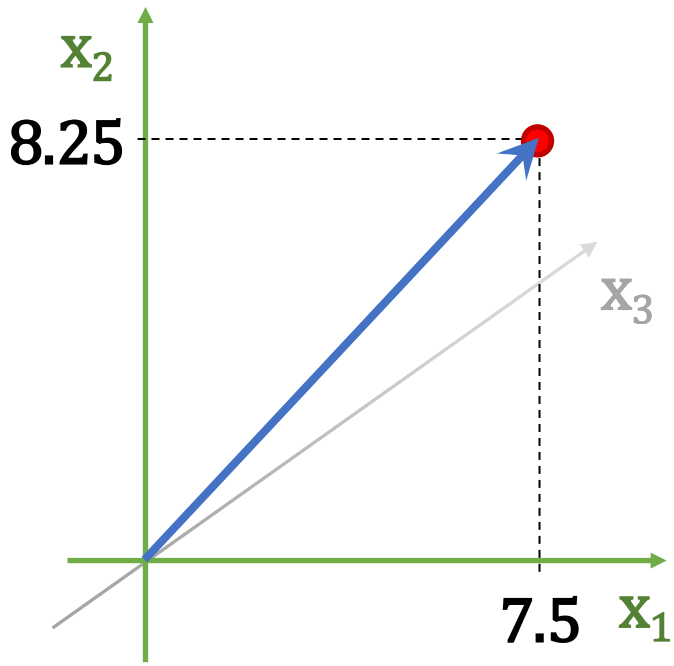
- **Vetores**

Row:
[7.5, 8.25, 35.8, 15, 10] *array*

ndarray 1-D

Column:
[7.5, → x_1
8.25, → x_2
35.8, → x_3
15, → x_4
10] → x_5

Notação:
v, $\vec{v}$



ESCALARES, VETORES, MATRIZES, TENSORES

- **Matrizes**

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

array de array

ndarray 2-D

$[[a_{11},\ a_{12},\ ...,\ a_{1n}],$
$[a_{21},\ a_{22},\ ...,\ a_{2n}],$
$...$
$[a_{m1},\ a_{m2},\ ...,\ a_{mn}]]$

Forma genérica dos elementos de uma matriz: a_{ij}

Notação: A

ESCALARES, VETORES, MATRIZES, TENSORES

- ## Tensores

Objeto matemático	ÁLGEBRA	PYTHON (NumPy) ndarray | np.array	PYTHON (TensorFlow/PyTorch) tensor
Escalar	Tensor de rank = 0	Array de ndim = 0	Tensor de ndim = 0
Vetor	Tensor de rank = 1	Array de ndim = 1	Tensor de ndim = 1
Matriz	Tensor de rank = 2	Array de ndim = 2	Tensor de ndim = 2
“Tensor”	Tensor de rank ≥ 3	Array de ndim ≥ 3	Tensor de ndim ≥ 3

método numpy

TENSORES EM ÁLGEBRA LINEAR
Sistemas lineares

$$y = 7.5x_1 + 8.25x_2 + 35.8x_3 + 15x_4 + 10$$

$$a_1 x_1 \begin{cases} a_1 = preço \\ x_1 = quantidade \end{cases} \updownarrow \begin{cases} a_1 = quantidade \\ x_1 = preço \end{cases}$$

$$\begin{cases} 127.35 = x_1 + x_2 + 2x_3 + 2x_4 + 10 \\ 128.10 = 2x_1 + 2x_2 + 2x_3 + x_4 + 10 \\ 134.85 = 2x_1 + x_2 + 2x_3 + 2x_4 + 10 \\ 119.85 = 2x_1 + x_2 + 2x_3 + x_4 + 10 \end{cases}$$

➔ sistema linear

- uma solução (completo, não-singular)
- nenhuma solução (singular)
- infinitas soluções (singular)
- ~~mais que uma solução?~~

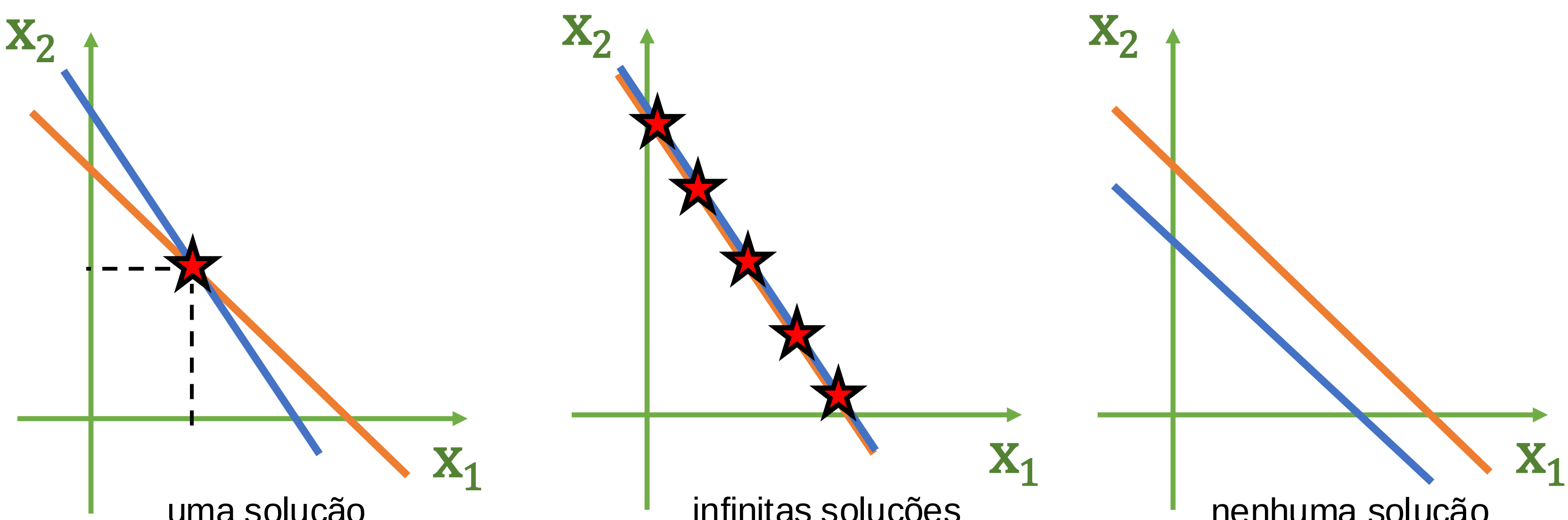
$$\begin{bmatrix} 127.35 \\ 128.10 \\ 134.85 \\ 119.85 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & 2 & 2 & 1 \\ 2 & 1 & 2 & 2 \\ 2 & 1 & 2 & 1 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 10 \\ 10 \\ 10 \\ 10 \end{bmatrix}$$

➔ $\boxed{\boldsymbol{y} = A\boldsymbol{x} + \boldsymbol{b}}$

TENSORES EM ÁLGEBRA LINEAR
Sistemas lineares

$$\begin{cases} 127.35 = x_1 + x_2 + 2x_3 + 2x_4 + 10 \\ 128.10 = 2x_1 + 2x_2 + 2x_3 + x_4 + 10 \\ 134.85 = 2x_1 + x_2 + 2x_3 + 2x_4 + 10 \\ 119.85 = 2x_1 + x_2 + 2x_3 + x_4 + 10 \end{cases}$$

casal_1
casal_2



TENSORES EM MACHINE LEARNING
Representação de dados

Dados tabulares (Pandas → `DataFrame`)

	Número de bebidas	Número de entradas	Número de pratos principais	Número de sobremesas	Valor total de couvert	Valor da conta
casal_1	1	1	2	2	10	127.35
casal_2	2	2	2	1	10	128.10
casal_3	2	1	2	2	10	134.85
casal_4	2	1	2	1	10	119.85

atributo `values`

Arrays (NumPy → `ndarray`)

```
[[1, 1, 2, 2, 10, 127.35],
 [2, 2, 2, 1, 10, 128.10],
 [2, 1, 2, 2, 10, 134.85],
 [2, 1, 2, 1, 10, 119.85]]
```

X y

```
[[1, 1, 2, 2, 10, 127.35]]
```

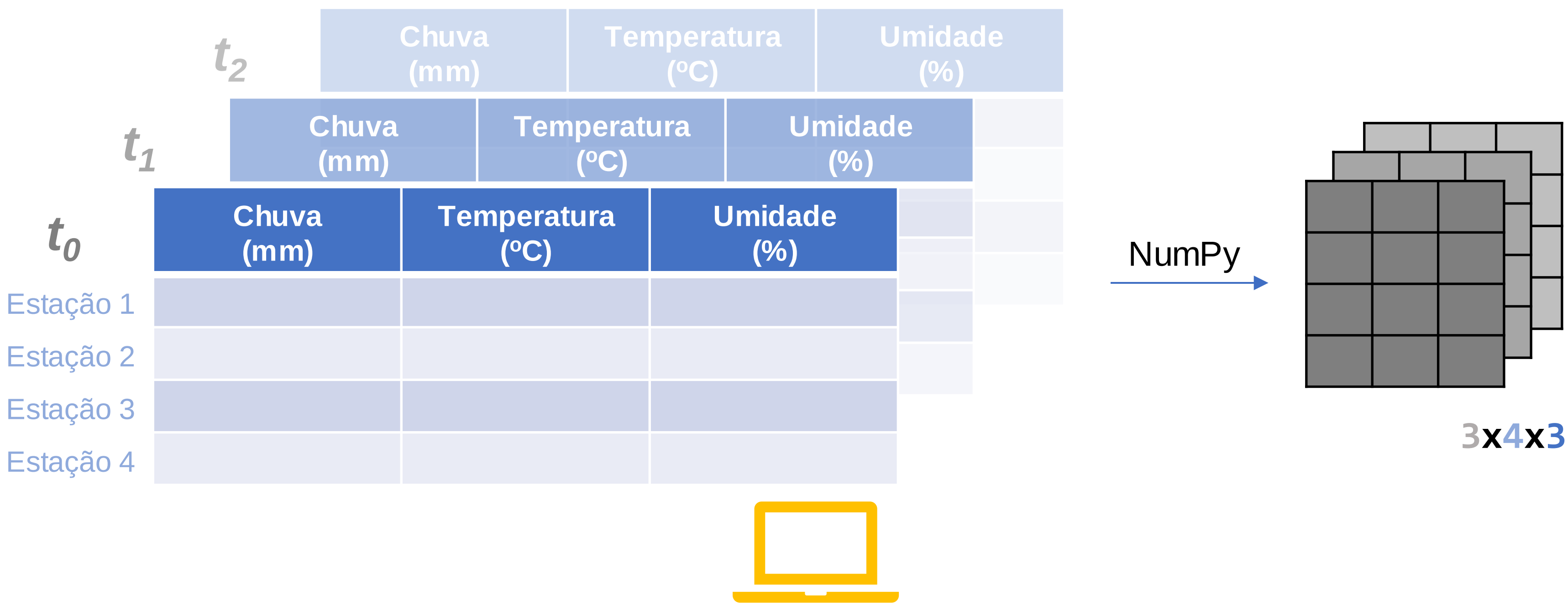
1x6

instância

EXPERT ACADEMY

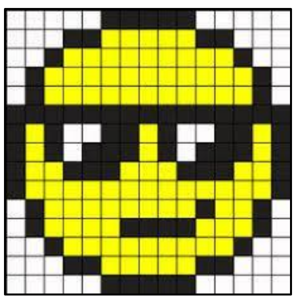
TENSORES EM MACHINE LEARNING
Representação de dados

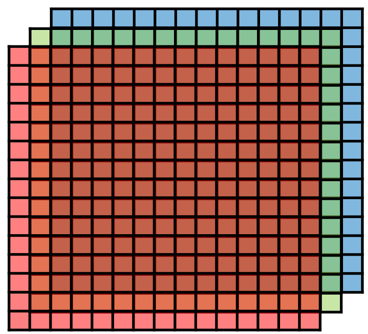
Séries temporais



TENSORES EM MACHINE LEARNING
Representação de dados

Imagens





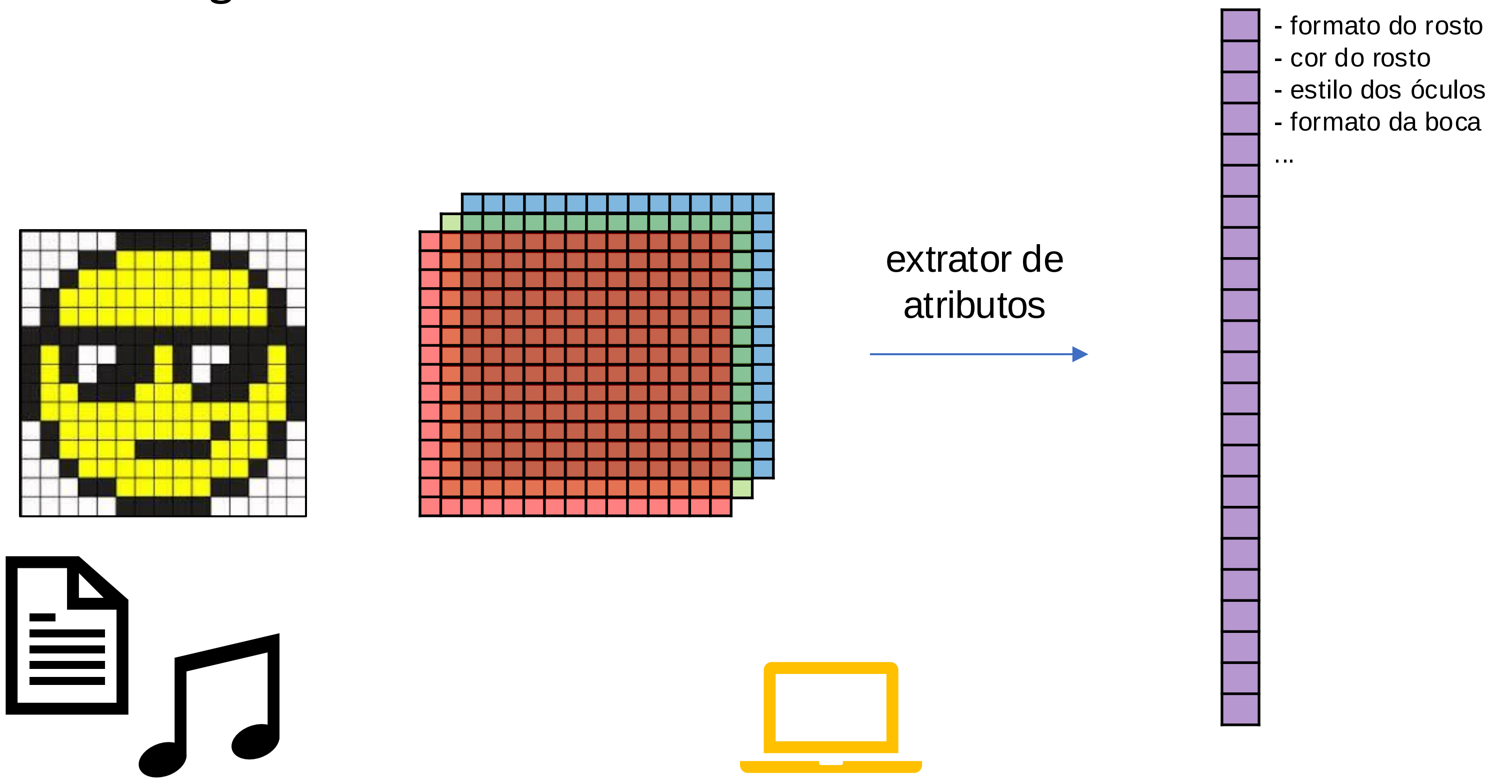
```
[l, h, c]

[b, l, h, c]
```

NumPy →

```
[
 [
  [11r, 12r, 13r, ...],
  [21r, 22r, 23r, ...],
   ...
 ],
 [
  [11g, 12g, 13g, ...],
  [21g, 22g, 23g, ...],
    ...
 ],
 [
  [11b, 12b, 13b, ...],
  [21b, 22b, 23b, ...],
    ...
 ]
]
```

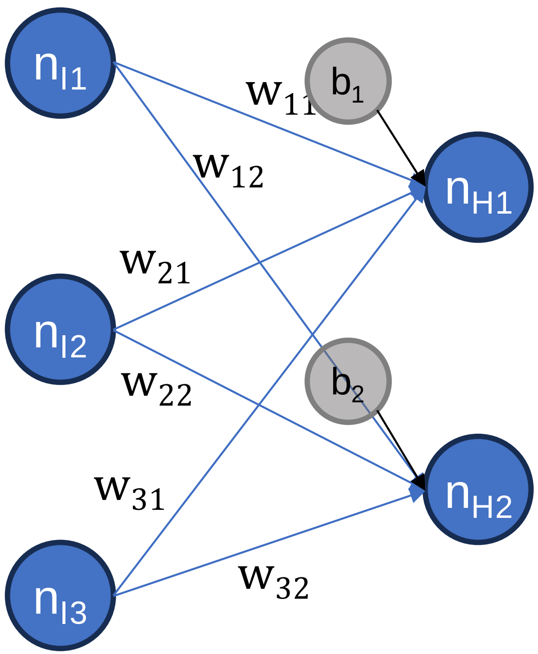
TENSORES EM MACHINE LEARNING
Representação de dados

Embeddings



TENSORES EM MACHINE LEARNING
Estrutura de redes neurais

Pesos e bias de redes neurais



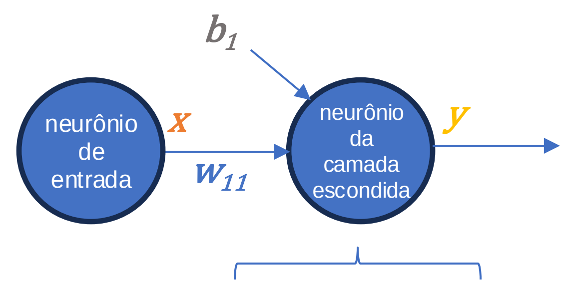

```
W =  [
       [w11, w12],
       [w21, w22],
       [w31, w32]
     ]
```

```
b = [b11, b12]
```

APLICAÇÕES EM MACHINE LEARNING

- **Treinamento de algoritmos**
- Redução de dimensionalidade
- Processamento de imagens
- Ranqueamento de resultados
- Sistemas de recomendação
- Processamento de linguagem natural

$$\begin{cases} 127.35 = x_1 + x_2 + 2x_3 + 2x_4 + 10 \\ 128.10 = 2x_1 + 2x_2 + 2x_3 + x_4 + 10 \\ 134.85 = 2x_1 + x_2 + 2x_3 + 2x_4 + 10 \\ 119.85 = 2x_1 + x_2 + 2x_3 + x_4 + 10 \end{cases}$$



$$y = Wx + b$$

ÁLGEBRA LINEAR
Vetores

DEFINIÇÃO (revisão)

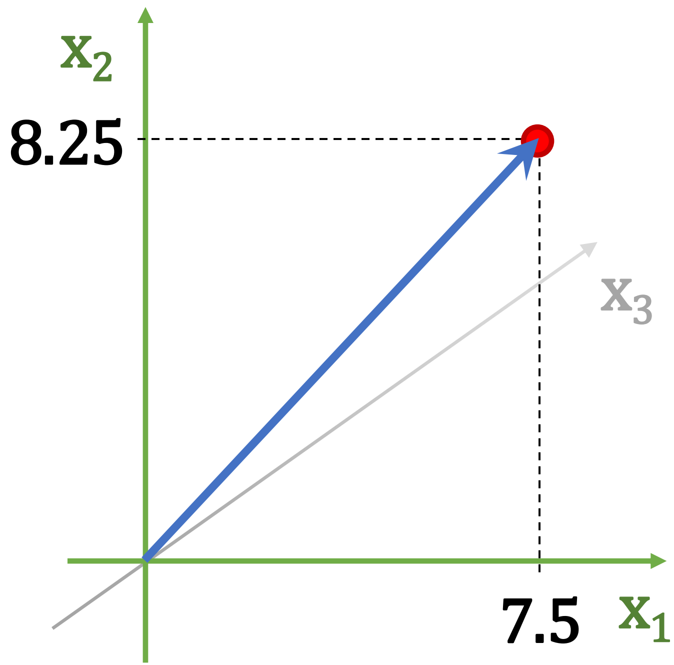
Row:

[7.5, 8.25, 35.8, 15, 10]

Column:

[7.5, → x_1
8.25, → x_2
35.8, → x_3
15, → x_4
10] → x_5

Notação:
v, $\vec{v}$



Álgebra: tensor de rank 1
Python:
- NumPy: ndarray (1-D)
- TF/PyTorch: tensor (1-D)

NORMAS

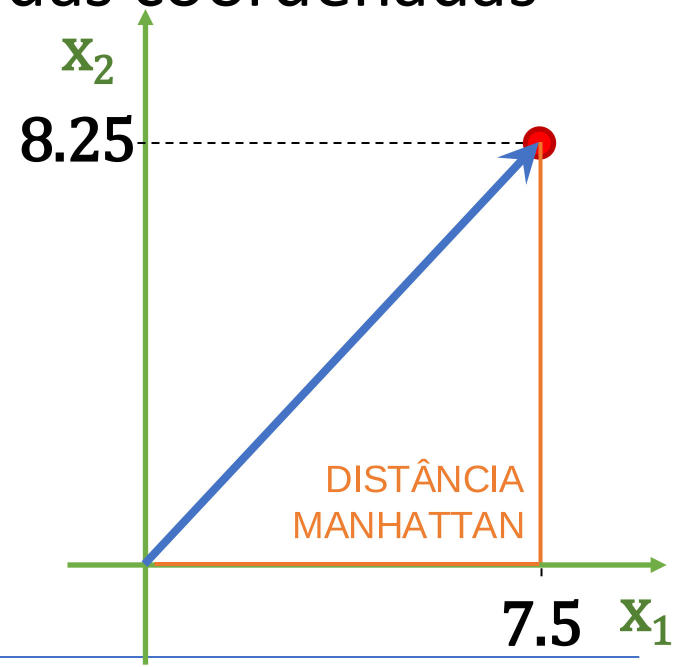
- Normas: funções que caracterizam o vetor
  - *Norma L1* (*L1-norm*): soma das coordenadas absolutas

$$\|\boldsymbol{x}\|_1 = \sum_i |x_i|$$

$$\|\boldsymbol{x}\|_1 = |x_1| + |x_2|$$

$$\|\boldsymbol{x}\|_1 = 7.5 + 8.25$$

$$\|\boldsymbol{x}\|_1 = \mathbf{15.75}$$



NORMAS

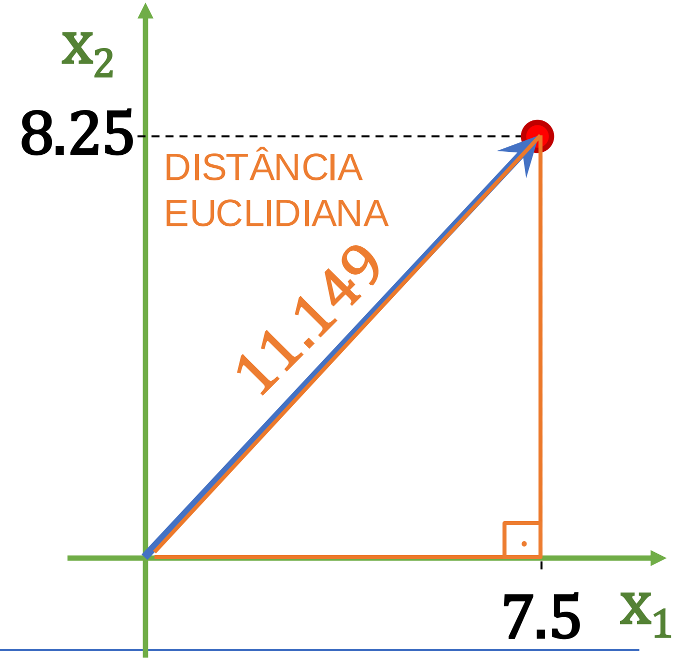
- Normas: funções que caracterizam o vetor
  - *Norma L2* (*L2-norm*): comprimento do vetor

$$\|\boldsymbol{x}\|_2 = \sqrt{\sum_i {x_i}^2}$$

$$\|\boldsymbol{x}\| = \sqrt{{x_1}^2 + {x_2}^2}$$

$$\|\boldsymbol{x}\| = \sqrt{7.5^2 + 8.25^2}$$

$$\|\boldsymbol{x}\| = \mathbf{11.149}$$



NORMAS

- Normas L1 e L2 são usadas para **regularizar** funções de custo.
- **Regressão linear:**

$$y = Ax + b \rightarrow RSS = \sum_{i=1}^{n} (y_i - \widehat{y_i})^2$$

- **Regressão lasso** (L1)

$$RSS = \sum_{i=1}^{n} (y_i - \widehat{y_i})^2 + \alpha||A||_1$$

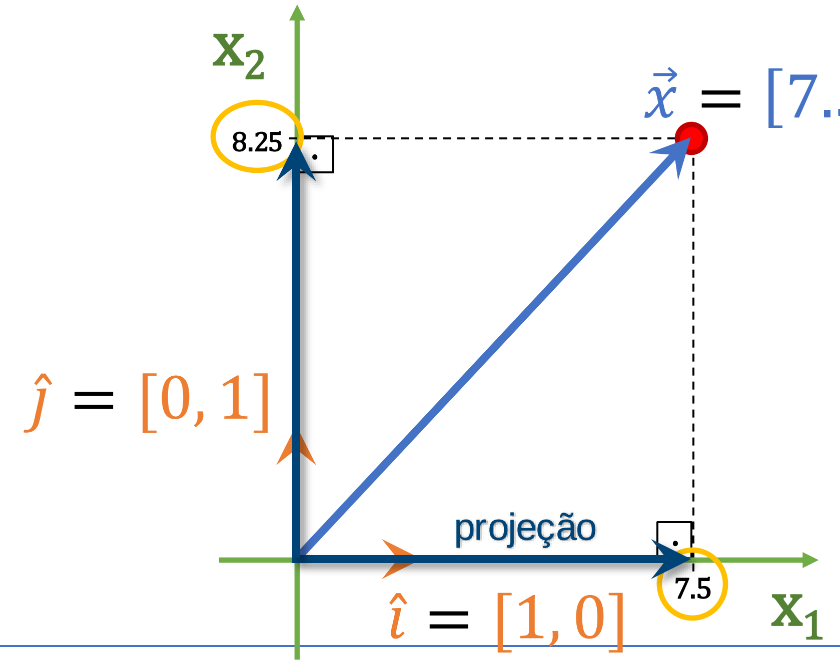
  - Remove correlações (zera alguns coeficientes) → **seleção de atributos**
- **Regressão ridge** (L2)

$$RSS = \sum_{i=1}^{n} (y_i - \widehat{y_i})^2 + \alpha||A||_2$$

  - Coeficientes menores (próximos de zero)

VETOR UNITÁRIO E BASE

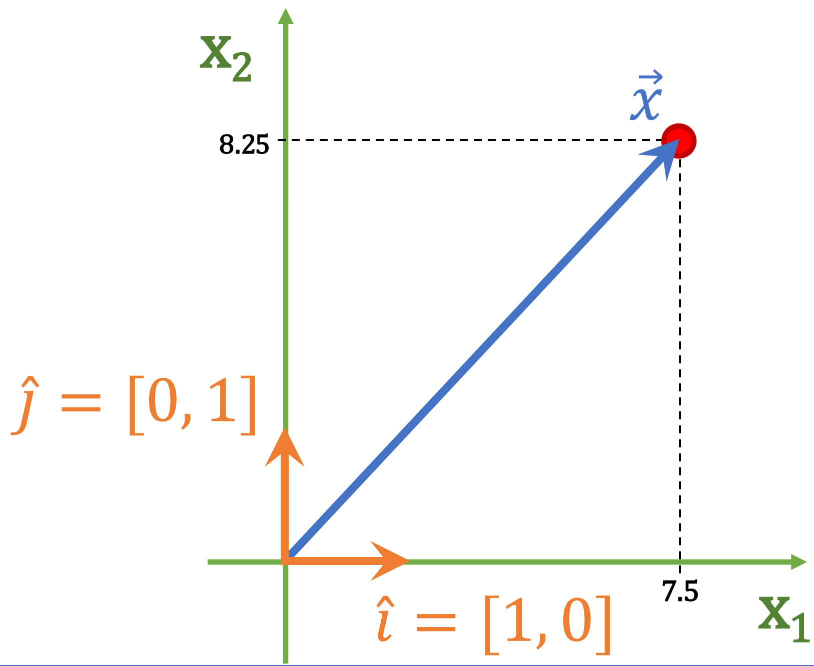
- **Vetor unitário**: vetor cujo comprimento (L2-norm) é igual a 1.
- **Vetor base**: vetor unitário que reside nos eixos do sistema cartesiano.



$$\vec{x} = 7.5\hat{\imath} + 8.25\hat{j}$$

VETOR TRANSPOSTO

- **Vetor transposto**: vetor linha transformado em vetor coluna, ou vice-versa.



$$\vec{x} = [7.5 \quad 8.25] \quad \Rightarrow \quad \vec{x}^T = \begin{bmatrix} 7.5 \\ 8.25 \end{bmatrix}$$

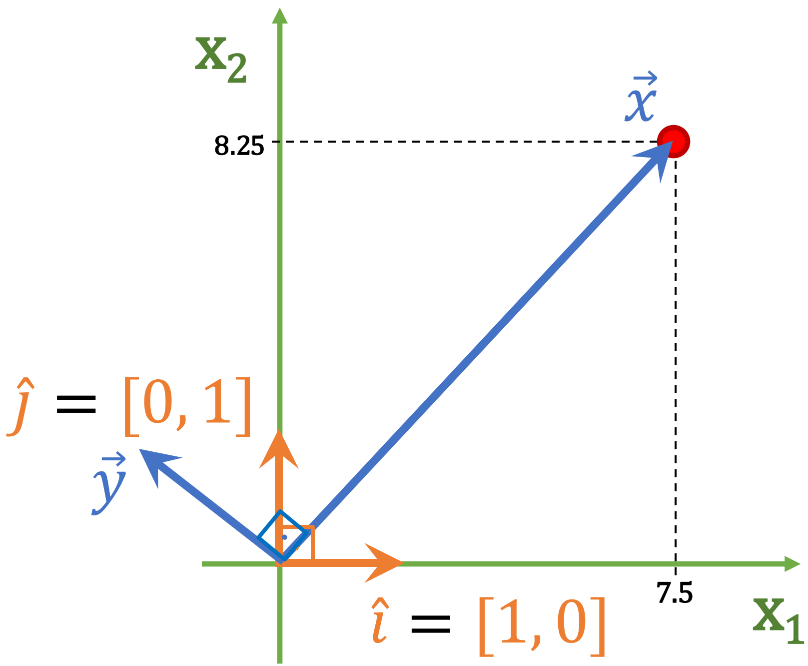
$$\vec{x} = [[7.5 \quad 8.25]]_{1\times2} \quad \Rightarrow \quad \vec{x}^T = \begin{matrix} [[7.5\ ] \\ [8.25]]_{2\times1} \end{matrix}$$

$$\hat{\imath} = [1 \quad 0] \quad \Rightarrow \quad \hat{\imath}^T = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\hat{\jmath} = [0 \quad 1] \quad \Rightarrow \quad \hat{\jmath}^T = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

VETOR ORTOGONAL E ORTONORMAL

- **Vetores ortogonais**: vetores que mantêm 90 graus entre si.
- **Vetores ortonormais:** vetores ortogonais e também unitários.



$$\vec{y} \neq f(\vec{x})$$
$$\vec{x} \neq f(\vec{y})$$
$$\hat{j} \neq f(\hat{\imath})$$
$$\hat{\imath} \neq f(\hat{j})$$

$$\vec{x} \cdot \vec{y} = 0$$

	$\vec{b}$	$\vec{e}$	$\vec{p}$	$\vec{s}$		
	Número de bebidas	Número de entradas	Número de pratos principais	Número de sobremesas	Valor total de couvert	Valor da conta
casal_1	1	1	2	2	10	127.35
casal_2	2	2	2	1	10	128.10
casal_3	2	1	2	2	10	134.85
casal_4	2	1	2	1	10	119.85

ÁLGEBRA LINEAR
Matrizes

DEFINIÇÃO (revisão)

- **Matrizes**

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Notação:
A

Forma genérica dos elementos de uma matriz: a_{ij}

Álgebra: tensor de rank 2
Computação: array de array
Python (NumPy): ndarray (2-D)

NORMAS

Norma Frobenius:

$$\|\boldsymbol{X}\|_F = \sqrt{\sum_{i,j} {x_{i,j}}^2}$$

Norma L2 do vetor

$$\|\boldsymbol{x}\|_2 = \sqrt{\sum_{i} {x_i}^2}$$

- análoga à norma L2 do vetor
- medida de “tamanho” (distância euclidiana)

Item do cardápio

```
[[1, 1, 2, 2, 10, 127.35],
 [2, 2, 2, 1, 10, 128.10],
 [2, 1, 2, 2, 10, 134.85],
 [2, 1, 2, 1, 10, 119.85]]
```

casal

MATRIZ TRANSPOSTA

- **Matrizes transpostas**: análogas aos vetores transpostos; no caso de matrizes, podemos pensar em uma **rotação** da matriz original.

$$\vec{x} = [7.5 \quad 8.25] \Rightarrow \vec{x}^T = \begin{bmatrix} 7.5 \\ 8.25 \end{bmatrix}$$

$$\vec{x} = [[7.5 \quad 8.25]]_{1\times2} \Rightarrow \vec{x}^T = \begin{matrix} [[7.5\ \ ] \\ [8.25]] \end{matrix}_{2\times1}$$

n_item_cardapio_1

casal_1

```
R  =  [[1, 1, 2, 2, 10, 127.35],
        [2, 2, 2, 1, 10, 128.10],
        [2, 1, 2, 2, 10, 134.85],
        [2, 1, 2, 1, 10, 119.85]]
                                  4x6
```

$\Rightarrow$

```
R^T  =  [[  1   ,   2   ,   2   ,   2   ],
         [  1   ,   2   ,   1   ,   1   ],
         [  2   ,   2   ,   2   ,   2   ],
         [  2   ,   1   ,   2   ,   1   ],
         [ 10   ,  10   ,  10   ,  10   ],
         [127.35, 128.10, 134.85, 119.85]]
                                          6x4
```

MATRIZ SIMÉTRICA, DIAGONAL E IDENTIDADE

- **Matriz simétrica**: somente matrizes quadradas, os elementos são "espelhados" em torno da diagonal principal.

```
[[1, 3, 2, 5],
 [3, 2, 8, 1],
 [2, 8, 6, 7],
 [5, 1, 7, 4]]
```

$$A = A^T$$

- **Matriz diagonal**: caso especial da matriz simétrica, todos os valores fora da diagonal principal são 0.

$\vec{a}$

```
[[1, 0, 0, 0],
 [0, 2, 0, 0],
 [0, 0, 6, 0],
 [0, 0, 0, 4]]
```

$$diag(\vec{a})$$

$$diag(\vec{a}) \times B = diag(\vec{a}) \odot B$$

$$diag(\vec{a})^{-1} = diag([a_1{}^{-1}, \dots, a_n{}^{-1}])$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{6} & 0 \\ 0 & 0 & 0 & \frac{1}{4} \end{bmatrix}$$

- **Matriz identidade**: caso especial da matriz diagonal, a diagonal principal é igual a 1.

```
[[1, 0, 0, 0],
 [0, 1, 0, 0],
 [0, 0, 1, 0],
 [0, 0, 0, 1]]
```

$$I_n \; (I_4)$$

$$I \times A = A$$

$$(1 \times n = n)$$

MATRIZ INVERSA

- **Matriz inversa**: matriz que, se multiplicada pela matriz original (que deve ser <u>quadrada</u>), resulta na matriz identidade.
- Matrizes *singulares* não têm inversa.
- **Matriz pseudoinversa de Moore-Penrose**: permite calcular a matriz "pseudo" inversa de matrizes retangulares ou singulares.

$$A^{-1}A = I$$

$$(n^{-1}n = 1)$$

```
[[1, 3, 2,  5],
 [3, 2, 8,  1],
 [2, 6, 4, 10],
 [5, 1, 7,  4]]
```

$\times 2$ **dependente**

DETERMINANTES

- **Determinante**: representação escalar da matriz (quadrada)

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \rightarrow \det(A) = |A| = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$$

$$A = \begin{bmatrix} \boldsymbol{a_{11}} & \boldsymbol{a_{12}} & \boldsymbol{a_{13}} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \rightarrow |A| = \boldsymbol{a_{11}} \cdot C_{11} + \boldsymbol{a_{12}} \cdot C_{12} + \boldsymbol{a_{13}} \cdot C_{13}$$

pivô

$$C_{ij} = (-1)^{i+j} \times M_{ij}$$

$$C_{11} = (-1)^{1+1} \times M_{11}$$
$$C_{11} = (-1)^2 \times \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$$
$$C_{11} = + \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$$

$$C_{12} = (-1)^{1+2} \times M_{12}$$
$$C_{12} = (-1)^3 \times \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$$
$$C_{11} = - \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$$

$$C_{13} = (-1)^{1+3} \times M_{13}$$
$$C_{13} = (-1)^4 \times \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$
$$C_{13} = + \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

EXPERT ACADEMY

DETERMINANTES

$$C = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} \rightarrow \quad C^T = \mathrm{adj}(A)$$

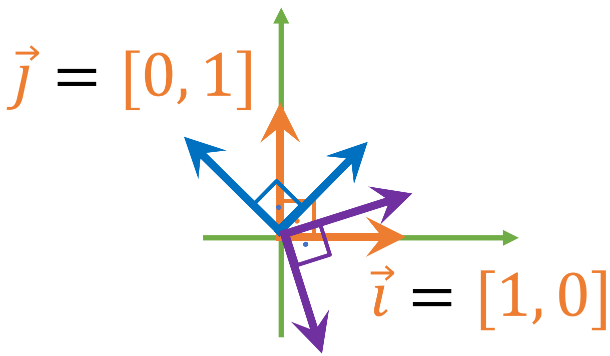
$$\boldsymbol{A^{-1} = \frac{1}{\det(A)}\mathrm{adj}(A)}$$

O determinante de matrizes singulares é zero

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

MATRIZ ORTOGONAL

- **Matrizes ortogonais**: matrizes onde os vetores que representam as linhas, ou os vetores que representam as colunas, são todos ortonormais (ou seja, ortogonais e unitários).



$$A^T A = A A^T = I$$
$$A^T = A^{-1} I = A^{-1}$$

```
[[1, 0],    [[0, 1],
 [0, 1]]     [1, 0]]
```

ÁLGEBRA LINEAR
Operações

OPERAÇÕES COM ESCALARES

- A operação é "transmitida" (*broacasted*) para todos os elementos

$$\begin{bmatrix}1\\2\\3\end{bmatrix} + 4 = \begin{bmatrix}5\\6\\7\end{bmatrix}$$

$$\begin{bmatrix}9 & 8 & 7\\6 & 5 & 4\\3 & 2 & 1\end{bmatrix} - 5 = \begin{bmatrix}4 & 3 & 2\\1 & 0 & -1\\-2 & -3 & -4\end{bmatrix}$$

$$\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\\7 & 8 & 9\end{bmatrix} \times 2 = \begin{bmatrix}2 & 4 & 6\\8 & 10 & 12\\14 & 16 & 18\end{bmatrix}$$

$$\begin{bmatrix}8\\10\\12\end{bmatrix} \div 2 = \begin{bmatrix}4\\5\\6\end{bmatrix}$$

OPERAÇÕES DENTRO DE TENSORES
Redução

- Representação "reduzida" do tensor original.
- A operação pode ser realizada ao longo de uma dimensão (`axis`) ou do tensor inteiro.

$$\text{sum}\left(\begin{bmatrix}1\\2\\3\end{bmatrix}\right) = 6$$

$$sum\left(\begin{bmatrix}9 & 8 & 7\\6 & 5 & 4\\3 & 2 & 1\end{bmatrix}\right) = 45$$

$$sum\left(\begin{bmatrix}9 & 8 & 7\\6 & 5 & 4\\3 & 2 & 1\end{bmatrix}, axis = 0\right) = \begin{bmatrix}18 & 15 & 12\end{bmatrix}$$

Algumas funções comuns:

$$sum, mean, prod, min, max$$

$$sum\left(\begin{bmatrix}9 & 8 & 7\\6 & 5 & 4\\3 & 2 & 1\end{bmatrix}, axis = 1\right) = \begin{bmatrix}24\\15\\6\end{bmatrix}$$

OPERAÇÕES ENTRE TENSORES (elemento a elemento)

- Os tensores devem ter o mesmo número de dimensões (shape).
- Resultado mantém o mesmo formato.

$$\begin{bmatrix}2\\1\\0\end{bmatrix}+\begin{bmatrix}3\\4\\5\end{bmatrix}=\begin{bmatrix}5\\5\\5\end{bmatrix}$$

$$\begin{bmatrix}9&8\\7&6\\5&4\end{bmatrix}-\begin{bmatrix}1&2\\3&4\\5&6\end{bmatrix}=\begin{bmatrix}8&6\\4&2\\0&-2\end{bmatrix}$$

$$\begin{bmatrix}1&3\\4&6\\7&9\end{bmatrix}\odot\begin{bmatrix}2&2\\3&3\\4&4\end{bmatrix}=\begin{bmatrix}2&6\\12&18\\28&36\end{bmatrix}$$

produto Hadamard

$$\begin{bmatrix}10\\9\\8\end{bmatrix}\div\begin{bmatrix}2\\3\\4\end{bmatrix}=\begin{bmatrix}5\\3\\2\end{bmatrix}$$

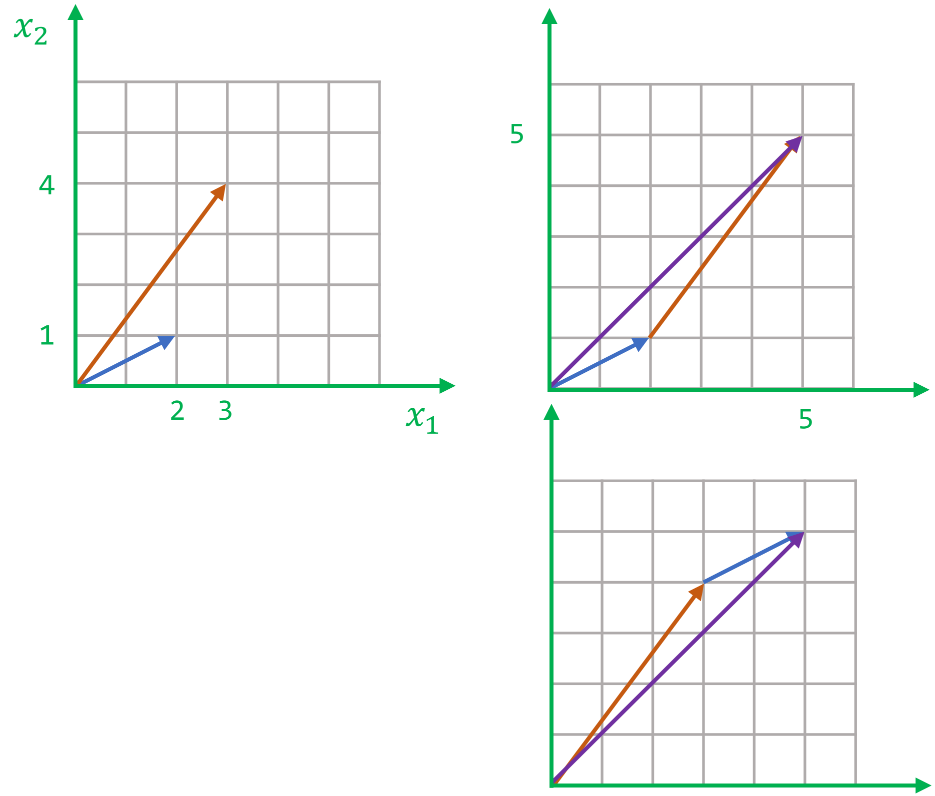
$$\begin{bmatrix}10\\9\\8\end{bmatrix}\odot\begin{bmatrix}1/2\\1/3\\1/4\end{bmatrix}=\begin{bmatrix}5\\3\\2\end{bmatrix}$$

OPERAÇÕES ENTRE TENSORES (elemento a elemento)

- Soma/subtração entre vetores tem um significado espacial.

$$\begin{bmatrix}2\\1\\0\end{bmatrix} + \begin{bmatrix}3\\4\\5\end{bmatrix} = \begin{bmatrix}5\\5\\5\end{bmatrix}$$

$$\begin{bmatrix}3\\4\\5\end{bmatrix} + \begin{bmatrix}2\\1\\0\end{bmatrix} = \begin{bmatrix}5\\5\\5\end{bmatrix}$$

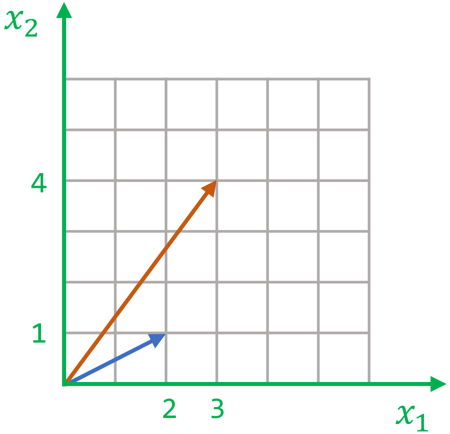


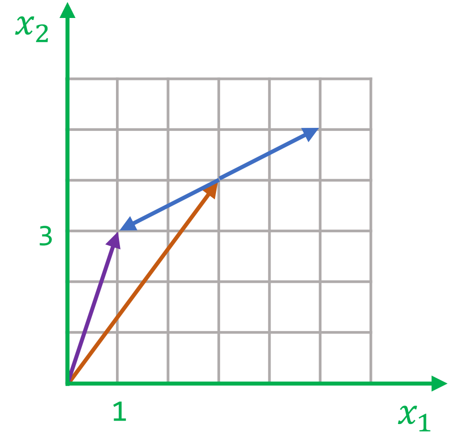
OPERAÇÕES ENTRE TENSORES (elemento a elemento)

- Soma/subtração entre vetores tem um significado espacial.

$$\boxed{\begin{bmatrix}3\\4\\5\end{bmatrix} - \begin{bmatrix}2\\1\\0\end{bmatrix} = \begin{bmatrix}1\\3\\5\end{bmatrix}}$$

$$\begin{bmatrix}3\\4\\5\end{bmatrix} + \begin{bmatrix}-2\\-1\\0\end{bmatrix} = \begin{bmatrix}-2\\-1\\0\end{bmatrix} + \begin{bmatrix}3\\4\\5\end{bmatrix} = \begin{bmatrix}1\\3\\5\end{bmatrix}$$





OPERAÇÕES ENTRE TENSORES
Multiplicação matricial

- Diferente da multiplicação elemento-a-elemento (produto Hadamard).
- Deve haver compatibilidade no número de dimensões ($\mathtt{shape}$).
- Não é comutativa: $A \times B \neq B \times A$

$$\begin{bmatrix}1 & 3\\4 & 6\\7 & 9\end{bmatrix} \odot \begin{bmatrix}2 & 2\\3 & 3\\4 & 4\end{bmatrix} = \begin{bmatrix}2 & 6\\12 & 18\\28 & 36\end{bmatrix}$$

$$\begin{bmatrix}1 & 3\\4 & 6\\7 & 9\end{bmatrix}_{3\times2} \times \begin{bmatrix}2 & 2\\3 & 3\\4 & 4\end{bmatrix}_{3\times2} =$$

$$\begin{bmatrix}1 & 3\\4 & 6\\7 & 9\end{bmatrix}_{3\times2} \times \begin{bmatrix}2 & 3 & 4\\2 & 3 & 4\end{bmatrix}_{2\times3} =$$

OPERAÇÕES ENTRE TENSORES
Multiplicação matricial

$$\begin{bmatrix} 1 & 3 \\ 4 & 6 \\ 7 & 9 \end{bmatrix}_{3\times 2} \times \begin{bmatrix} 2 & 3 & 4 \\ 2 & 3 & 4 \end{bmatrix}_{2\times 3} =$$

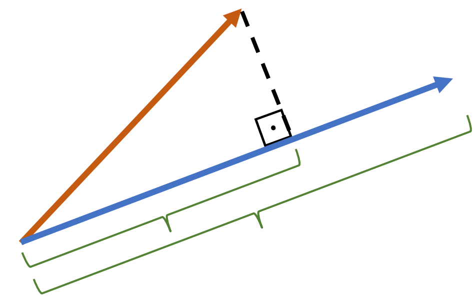
$$\begin{bmatrix} 1*2+3*2 & 1*3+3*3 & 1*4+3*4 \\ 4*2+6*2 & 4*3+6*3 & 4*4+6*4 \\ 7*2+9*2 & 7*3+9*3 & 7*4+9*4 \end{bmatrix}_{3\times 3} =$$

$$\begin{bmatrix} 8_{11} & 12_{12} & 16_{13} \\ 20_{21} & 30_{22} & 40_{23} \\ 32_{31} & 48_{32} & 64_{33} \end{bmatrix}_{3\times 3}$$

OPERAÇÕES ENTRE TENSORES
Produto vetorial/escalar/interno – *dot product*

- Multiplicação elemento-a-elemento seguida de redução por soma
- Resultado é um escalar

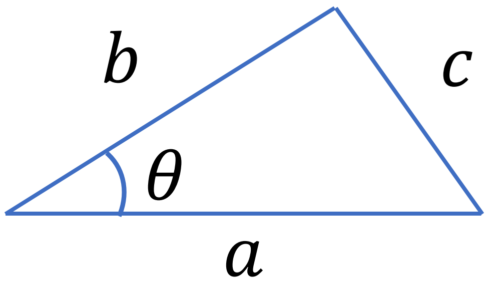
$$\begin{bmatrix}2\\1\\0\end{bmatrix} \cdot \begin{bmatrix}3\\4\\5\end{bmatrix} = 2 * 3 + 1 * 4 + 0 * 5 = 10$$



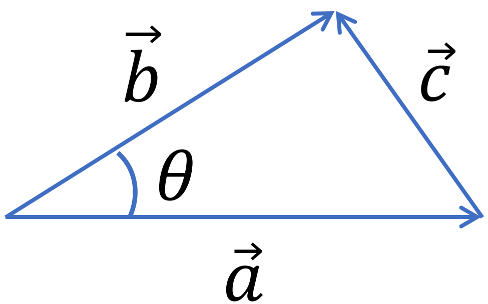
$$\begin{bmatrix}[2 & 1 & 0]\end{bmatrix}_{1\times 3} \times \begin{bmatrix}[3]\\ [4]\\ [5]\end{bmatrix}_{3\times 1} = [2 * 3 + 1 * 4 + 0 * 5]_{1\times 1} = 10$$

$$\boxed{\vec{x} \cdot \vec{x} = \|\vec{x}\|^2}$$

OPERAÇÕES ENTRE TENSORES
Regra de cosseno



$$c^2 = a^2 + b^2 - 2ab\cos\theta$$



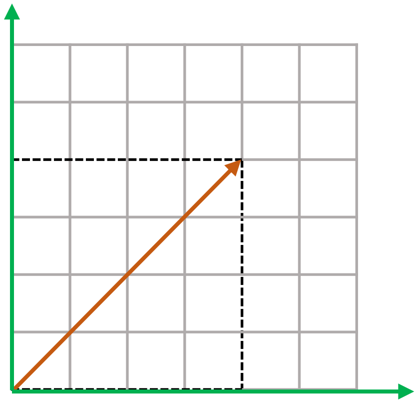
$$\vec{b} = \vec{a} + \vec{c} \quad \rightarrow \quad \vec{c} = \vec{b} - \vec{a}$$

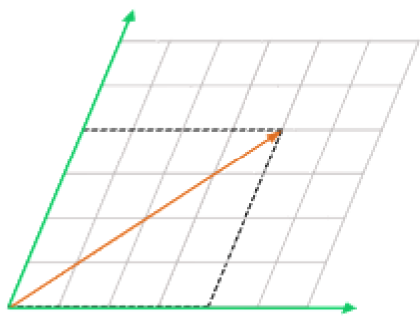
$$\boxed{\cos\theta = \frac{\vec{a}\cdot\vec{b}}{ab}}$$

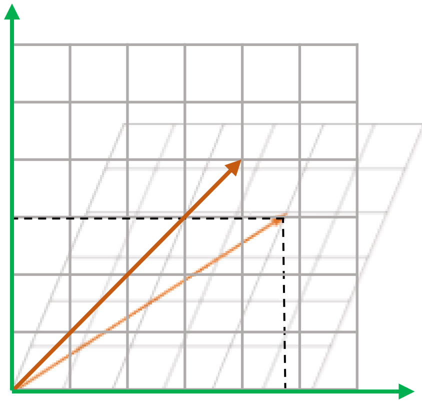
ÁLGEBRA LINEAR
Transformações

TRANSFORMAÇÕES AFINS

- É possível “transformar” um vetor através da manipulação do espaço cartesiano.



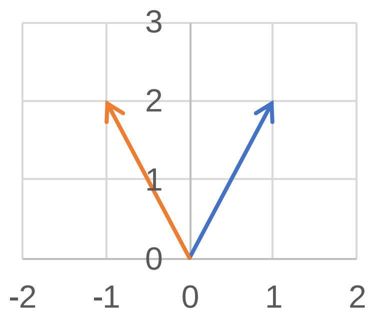


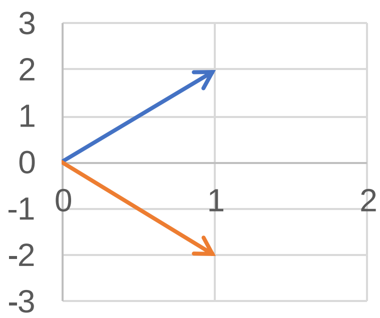


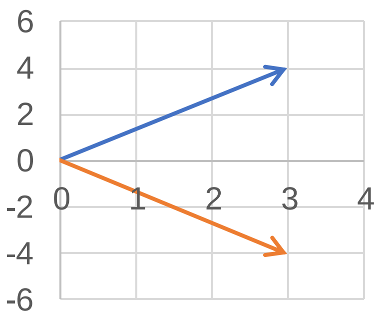
TRANSFORMAÇÕES AFINS

- Um vetor pode ser **transformado** em outro através da "aplicação" de uma matriz.
- A e B são chamadas matrizes de *espelhamento*.

$$A \times \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \qquad B \times \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \qquad B \times \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$$



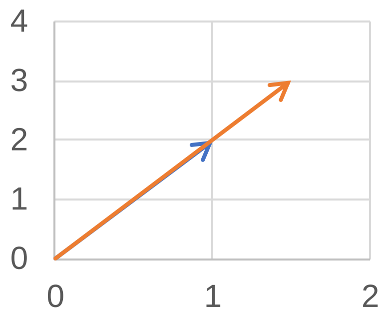




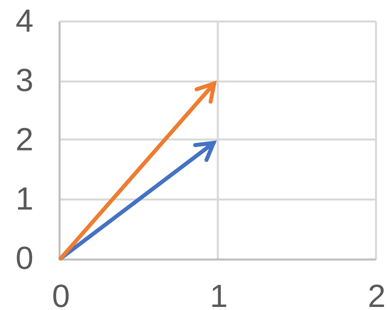
TRANSFORMAÇÕES AFINS

- Outras transformações afins:

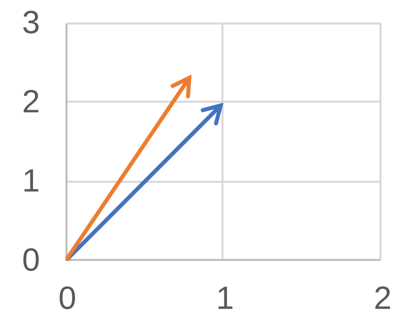
escalonamento



cisalhamento



rotação

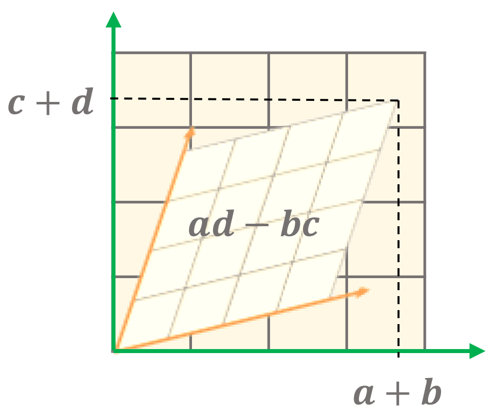


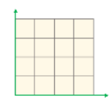
$$\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \qquad \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \qquad \begin{bmatrix} -2 & 0 \\ 0 & -5 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -5 \end{bmatrix}$$

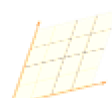
$$\begin{bmatrix} 3 & 5 \\ 1 & -2 \end{bmatrix} \times \underset{\hat{\imath}}{\begin{bmatrix} 1 \\ 0 \end{bmatrix}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \qquad \begin{bmatrix} 3 & 5 \\ 1 & -2 \end{bmatrix} \times \underset{\hat{\jmath}}{\begin{bmatrix} 0 \\ 1 \end{bmatrix}} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$

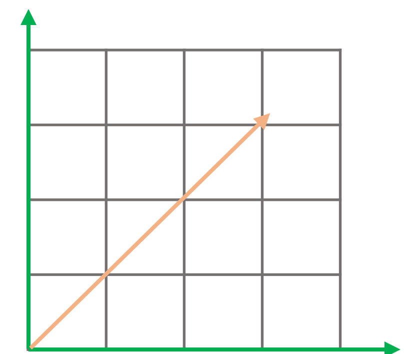
TRANSFORMAÇÕES E DETERMINANTES

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a} & \boldsymbol{b} \\ \boldsymbol{c} & \boldsymbol{d} \end{bmatrix} \rightarrow \det(A) = |\boldsymbol{A}| = \boldsymbol{ad} - \boldsymbol{bc}$$



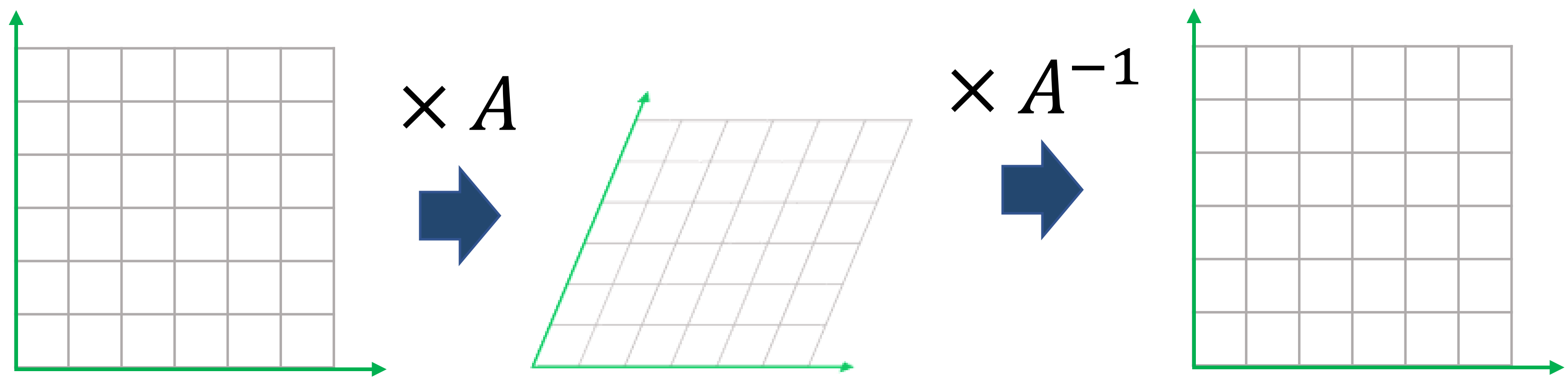

Área


Área $\times$ $\boldsymbol{det}(\boldsymbol{A})$



Matrizes singulares
$\det(A) = 0$

TRANSFORMAÇÕES E MATRIZES INVERSAS



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

TRANSFORMAÇÕES E SISTEMAS LINEARES

$$\begin{cases} 127.35 = x_1 + x_2 + 2x_3 + 2x_4 + 10 \\ 128.10 = 2x_1 + 2x_2 + 2x_3 + x_4 + 10 \\ 134.85 = 2x_1 + x_2 + 2x_3 + 2x_4 + 10 \\ 119.85 = 2x_1 + x_2 + 2x_3 + x_4 + 10 \end{cases} \rightarrow \begin{bmatrix} 127.35 \\ 128.10 \\ 134.85 \\ 119.85 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & 2 & 2 & 1 \\ 2 & 1 & 2 & 2 \\ 2 & 1 & 2 & 1 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 10 \\ 10 \\ 10 \\ 10 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & 2 & 2 & 1 \\ 2 & 1 & 2 & 2 \\ 2 & 1 & 2 & 1 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 127.35 \\ 128.10 \\ 134.85 \\ 119.85 \end{bmatrix} - \begin{bmatrix} 10 \\ 10 \\ 10 \\ 10 \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{2} & \mathbf{2} \\ \mathbf{2} & \mathbf{2} & \mathbf{2} & \mathbf{1} \\ \mathbf{2} & \mathbf{1} & \mathbf{2} & \mathbf{2} \\ \mathbf{2} & \mathbf{1} & \mathbf{2} & \mathbf{1} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{x_1} \\ \boldsymbol{x_2} \\ \boldsymbol{x_3} \\ \boldsymbol{x_4} \end{bmatrix} = \begin{bmatrix} \mathbf{117.35} \\ \mathbf{118.10} \\ \mathbf{124.85} \\ \mathbf{109.85} \end{bmatrix}$$

$$\boldsymbol{A} \times \vec{\boldsymbol{x}} = \vec{\boldsymbol{y}}$$

EIGENVECTORS E EIGENVALUES

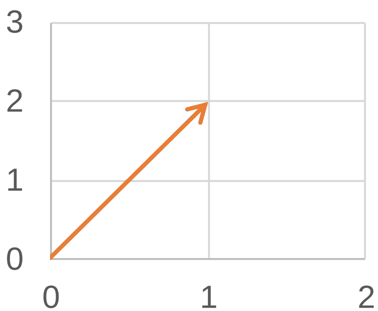
- **Eigen** (alemão): típico, característico, "auto"
- **Eigenvectors**: vetores que, após uma transforr
mudam de direção
- **Eigenvalues (λ)**: valores escalares aplicados a

Matrizes singulares

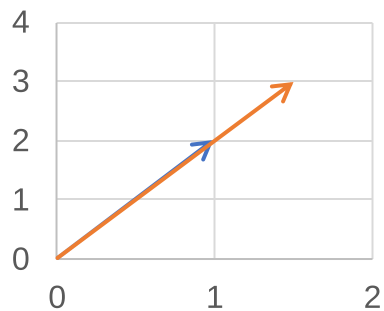
```
[[1, 3, 2,  5],
 [3, 2, 8,  1],
 [2, 6, 4, 10],
 [5, 1, 7,  4]]
```

× 2

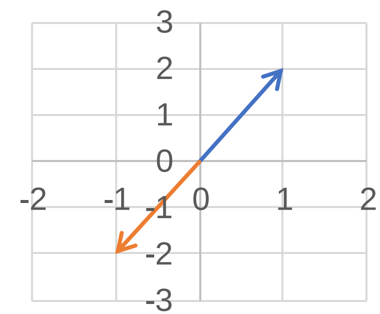
$$A\vec{v} = \lambda\vec{v}$$



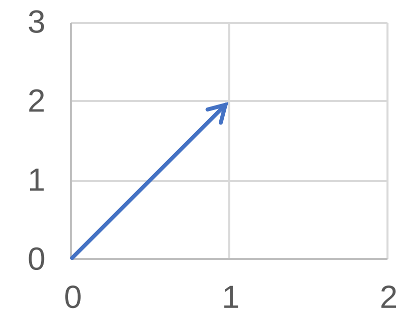
eigenvalue = 1



eigenvalue = 1,5



eigenvalue = -1



eigenvalue = **0**

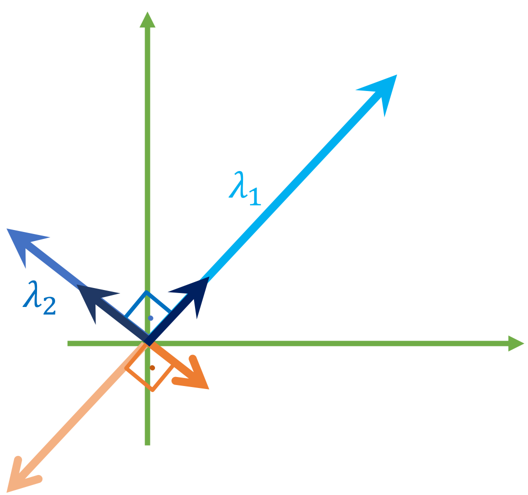
EIGENVECTORS E EIGENVALUES

- Matriz de transformação A_n pode ter vários eigenvectors.
- Mas no máximo n eigenvectors (se A for simétrica) serão linearmente independentes (ortogonais).
- Esses eigenvectors são expressos com tamanho 1 (ortonormais), junto com seus respectivos eigenvalues λ.

Matriz simétrica: somente matrizes quadradas, os elementos são “espelhados” em torno da diagonal principal.

```
[[1, 3, 2, 5],
 [3, 2, 8, 1],
 [2, 8, 6, 7],
 [5, 1, 7, 4]]
```

$$A = A^T$$



$$A\vec{v} = \lambda\vec{v}$$

$$\det(A_n) = \prod_{i=1}^{n} \lambda_i$$

AUTODECOMPOSIÇÃO (*EIGENDECOMPOSITION)*

$$A = Q\Lambda Q^{-1}$$

$$A_n$$

$$Q = \begin{array}{c} \begin{array}{cccc} \overrightarrow{v_1} & \overrightarrow{v_2} & & \overrightarrow{v_n} \end{array} \\ \begin{bmatrix} \boxed{\begin{matrix} q_{11} \\ q_{21} \\ \vdots \\ q_{n1} \end{matrix}} & \boxed{\begin{matrix} q_{12} \\ q_{22} \\ \cdots \\ q_{n2} \end{matrix}} & \begin{matrix} \cdots \\ \cdots \\ \ddots \\ \cdots \end{matrix} & \boxed{\begin{matrix} q_{1n} \\ q_{2n} \\ \vdots \\ q_{nn} \end{matrix}} \end{bmatrix} \end{array}$$

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$$

Se A é uma matriz simétrica, então: $\boldsymbol{A = Q\Lambda Q^T}$

DECOMPOSIÇÃO DE VALOR SINGULAR (SINGULAR VALUE DECOMPOSITION - SVD)

$$A = UDV^T$$

$$A_{m\times n}$$

$$U = \begin{array}{cccc} \overrightarrow{u_1} & \overrightarrow{u_2} & & \overrightarrow{u_m} \\ \end{array}$$

$$U = \begin{bmatrix} \boxed{\begin{matrix} u_{11} \\ u_{21} \\ \vdots \\ u_{m1} \end{matrix}} & \boxed{\begin{matrix} u_{12} \\ u_{22} \\ \cdots \\ u_{m2} \end{matrix}} & \begin{matrix} \cdots \\ \cdots \\ \ddots \\ \cdots \end{matrix} & \boxed{\begin{matrix} u_{1m} \\ u_{2m} \\ \vdots \\ u_{mm} \end{matrix}} \end{bmatrix}_{m\times m}$$

eigenvectors de AA^T

$$\sqrt{\lambda_{AA^T}} = \sqrt{\lambda_{A^TA}}$$

$$D = \begin{bmatrix} s_1 & 0 & \cdots & 0 \\ 0 & s_2 & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_m \end{bmatrix}_{m\times n}$$

$$s_1 > s_2 > \cdots > s_m$$

$$V = \begin{array}{cccc} \overrightarrow{v_1} & \overrightarrow{v_2} & & \overrightarrow{v_n} \\ \end{array}$$

$$V = \begin{bmatrix} \boxed{\begin{matrix} v_{11} \\ v_{21} \\ \vdots \\ v_{n1} \end{matrix}} & \boxed{\begin{matrix} v_{12} \\ v_{22} \\ \cdots \\ v_{n2} \end{matrix}} & \begin{matrix} \cdots \\ \cdots \\ \ddots \\ \cdots \end{matrix} & \boxed{\begin{matrix} v_{1n} \\ v_{2n} \\ \vdots \\ v_{nn} \end{matrix}} \end{bmatrix}_{n\times n}$$

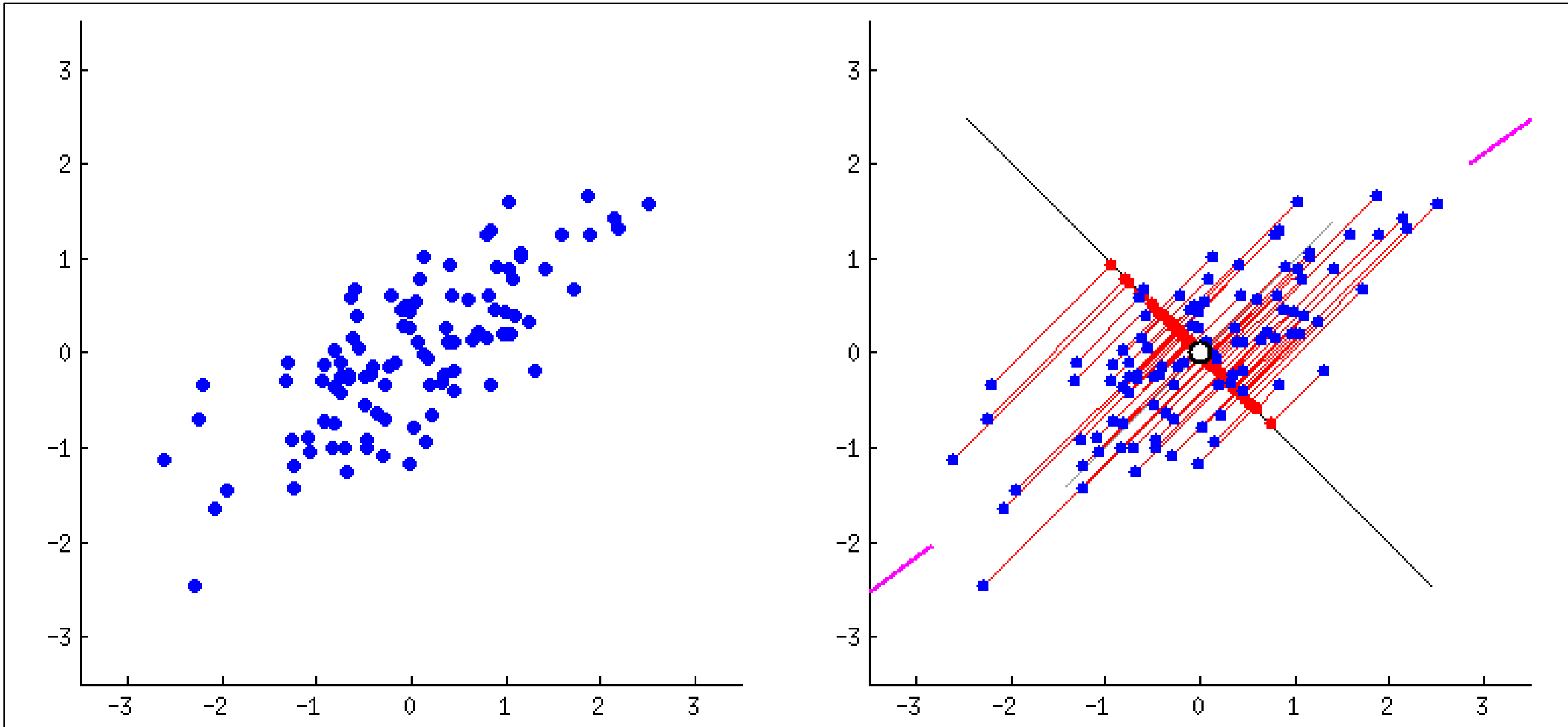
eigenvectors de A^TA

MATRIZ PSEUDOINVERSA DE MOORE-PENROSE

$$A^{+} = UD^{+}V^{T}$$

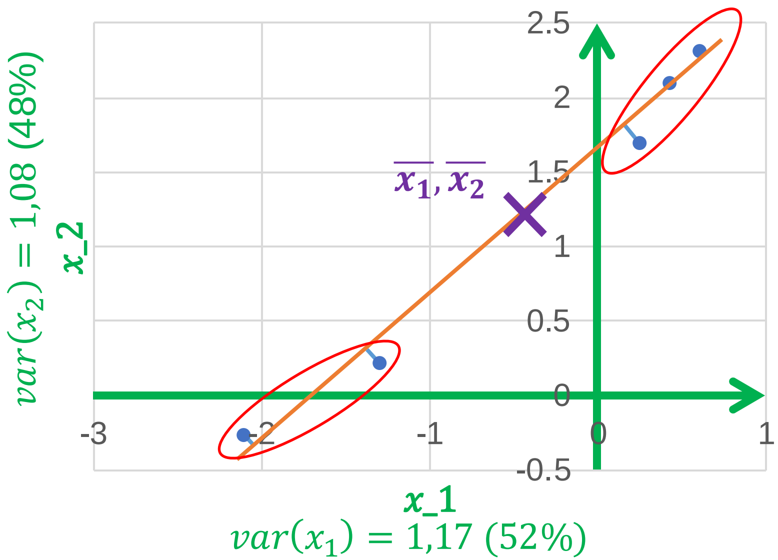
$$D^{+} = \left(\frac{1}{D^{*}}\right)^{T} = D^{-1}$$

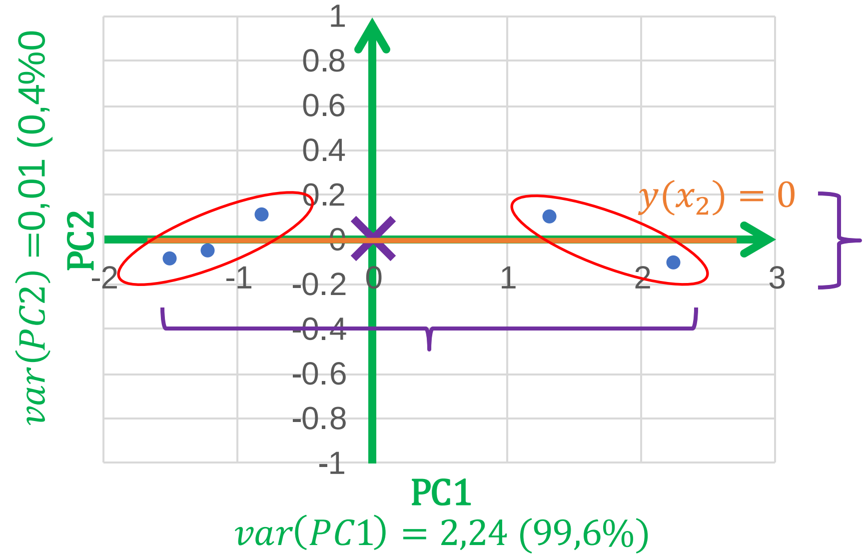
ANÁLISE DO COMPONENTE PRINCIPAL (PRINCIPAL COMPONENT ANALYSIS – PCA)



ANÁLISE DO COMPONENTE PRINCIPAL (*PRINCIPAL COMPONENT ANALYSIS – PCA)*

Transformação que provoca rotação e translação dos eixos, de forma a remover as correlações, equilibrar os valores das variáveis em volta da origem do sistema, e acumular a variância nos primeiros eixos.





ÁLGEBRA LINEAR
Aplicações

Sistemas lineares

- Podem ser representados como operações matriciais:

$$\begin{cases} 127.35 = x_1 + x_2 + 2x_3 + 2x_4 + 10 \\ 128.10 = 2x_1 + 2x_2 + 2x_3 + x_4 + 10 \\ 134.85 = 2x_1 + x_2 + 2x_3 + 2x_4 + 10 \\ 119.85 = 2x_1 + x_2 + 2x_3 + x_4 + 10 \end{cases} \rightarrow \text{sistema linear}$$

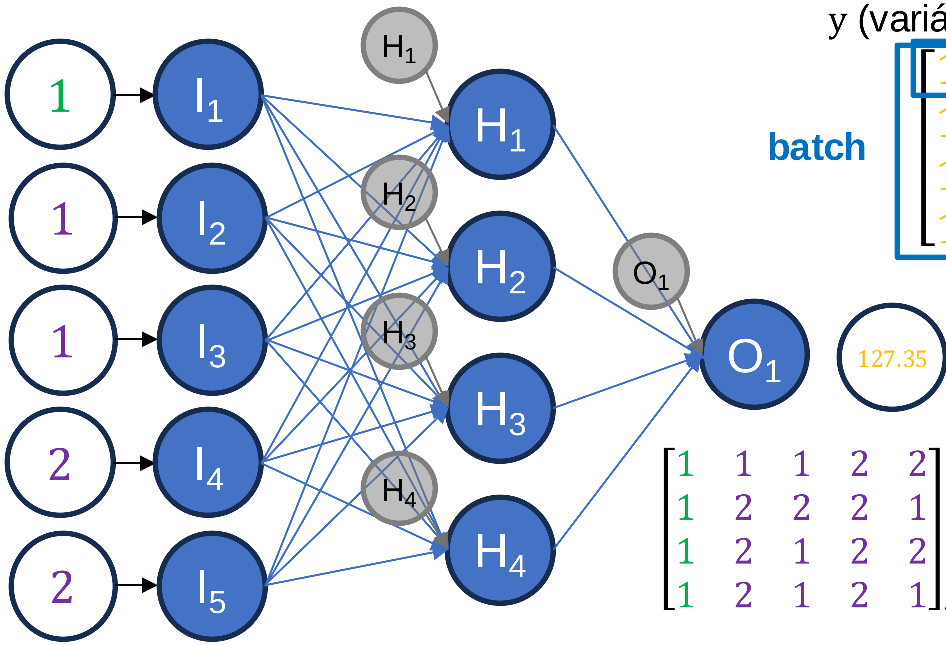
$$\begin{bmatrix} 127.35 \\ 128.10 \\ 134.85 \\ 119.85 \end{bmatrix}_{4\times 1} = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & 2 & 2 & 1 \\ 2 & 1 & 2 & 2 \\ 2 & 1 & 2 & 1 \end{bmatrix}_{4\times 4} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}_{4\times 1} + \begin{bmatrix} 10 \\ 10 \\ 10 \\ 10 \end{bmatrix}_{4\times 1} \rightarrow \boxed{\boldsymbol{y} = A\boldsymbol{x} + \boldsymbol{b}}$$

$m \geq n$

$$\begin{bmatrix} 127.35 \\ 128.10 \\ 134.85 \\ 119.85 \end{bmatrix}_{4\times 1} = \begin{bmatrix} 1 & 1 & 1 & 2 & 2 \\ 1 & 2 & 2 & 2 & 1 \\ 1 & 2 & 1 & 2 & 2 \\ 1 & 2 & 1 & 2 & 1 \end{bmatrix}_{4\times 5} \times \begin{array}{l} \begin{bmatrix} 10 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}_{5\times 1} \end{array} \begin{array}{l} (x_0) \\ \\ \\ \\ \\ \end{array} \rightarrow \boxed{\boldsymbol{y} = A\boldsymbol{x}}$$

$$y = Ax$$
$$A^{-1}y = A^{-1}Ax$$
$$A^{-1}y = Ix$$
$$\boldsymbol{x} = \boxed{\boldsymbol{A^{-1}}}\boldsymbol{y} \rightarrow \boldsymbol{A^{+}}$$

Redes neurais



y (variável alvo, $\mathbf{y}_{\text{true}}$)

batch

$$\begin{bmatrix} 127.35 \\ 128.10 \\ 134.85 \\ 119.85 \end{bmatrix}_{4\times1} = \begin{bmatrix} 1 & 1 & 1 & 2 & 2 \\ 1 & 2 & 2 & 2 & 1 \\ 1 & 2 & 1 & 2 & 2 \\ 1 & 2 & 1 & 2 & 1 \end{bmatrix}_{4\times5} \times \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}_{5\times1}$$

X (variáveis preditoras)

REDE NEURAL

$$\begin{bmatrix} 1 & 1 & 1 & 2 & 2 \\ 1 & 2 & 2 & 2 & 1 \\ 1 & 2 & 1 & 2 & 2 \\ 1 & 2 & 1 & 2 & 1 \end{bmatrix}_{4\times5} \times \begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \\ w_{41} & w_{42} & w_{43} & w_{44} \\ w_{51} & w_{52} & w_{53} & w_{54} \end{bmatrix}_{5\times4} + \begin{bmatrix} b_1 & b_2 & b_3 & b_4 \end{bmatrix}_{1\times4}$$

broadcasting

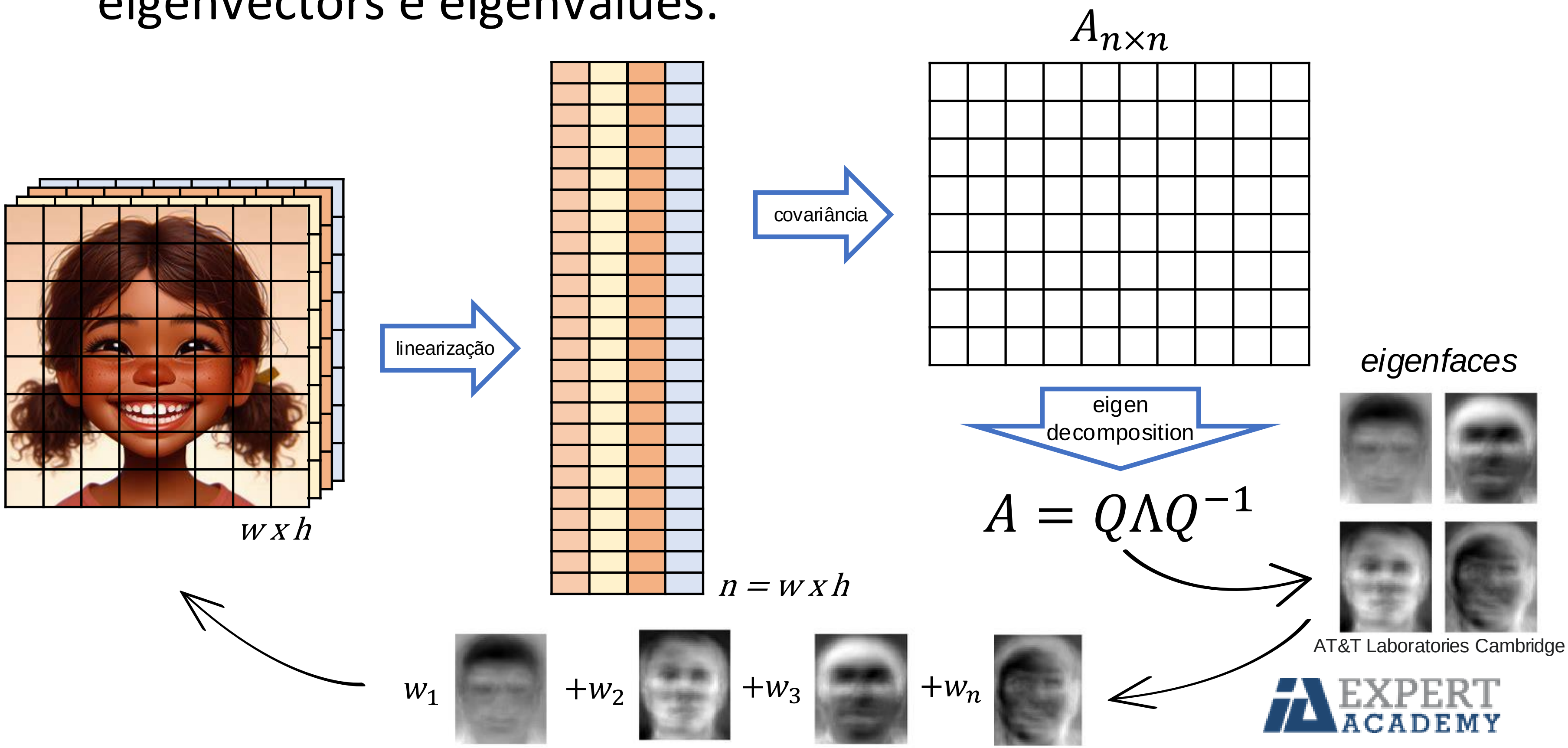
$$\begin{bmatrix} b_1 & b_2 & b_3 & b_4 \\ b_1 & b_2 & b_3 & b_4 \\ b_1 & b_2 & b_3 & b_4 \\ b_1 & b_2 & b_3 & b_4 \end{bmatrix}_{4\times4}$$

$$\boldsymbol{R}_{4\times4} \times \boldsymbol{W}_{4\times1} + \boldsymbol{b}_{1(4)\times1} = \boldsymbol{O}_{4\times1}$$

$\mathbf{y}_{\text{pred}}$

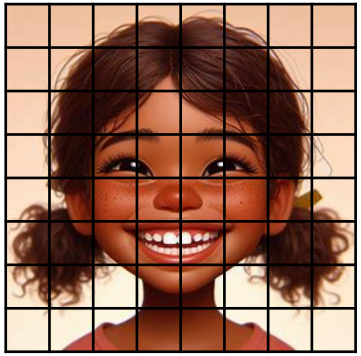
Eigendecomposition

- Um dataset qualquer pode ser decomposto em seus eigenvectors e eigenvalues.



Singular value decomposition

- Comprimir dados



$w\,x\,h$

quantidade de informação necessária para armazenar a matriz

decomposição

$$A_{w\times h} = U_{w\times w} D_{w\times h} {V_{h\times h}}^T$$

reconstrução

$$s_1 > s_2 > \cdots > s_m$$

$w = h = 200$
$40000 \Rightarrow 3208$
(-92%)

$${U_{w\times 1} D_{1\times 1} V_{h\times 1}}^T$$

$${U_{w\times 3} D_{3\times 3} V_{h\times 3}}^T$$

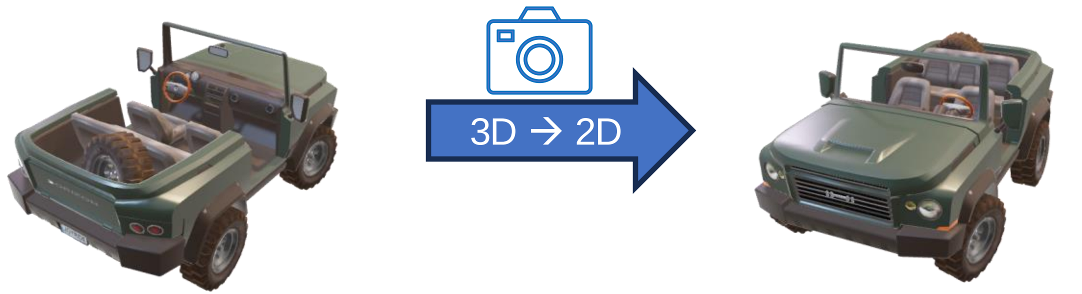
$${U_{w\times 8} D_{8\times 8} V_{h\times 8}}^T$$

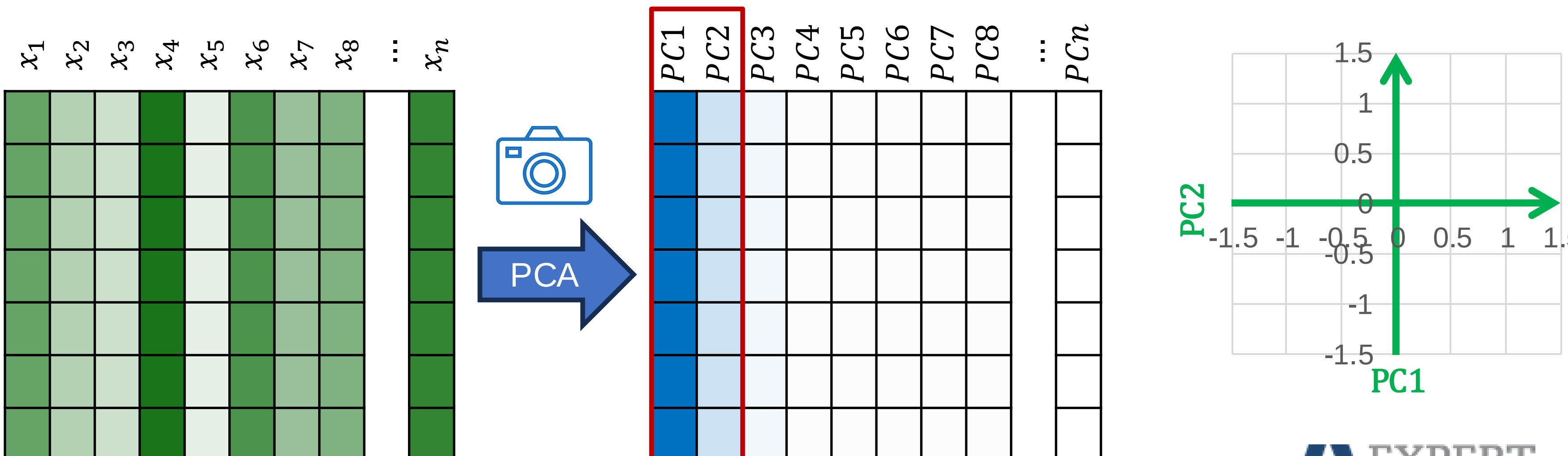
$w \times 8$
$+$
8
$+$
$h \times 8$

EXPERT ACADEMY

PCA

- Redução de dimensionalidade e visualização.

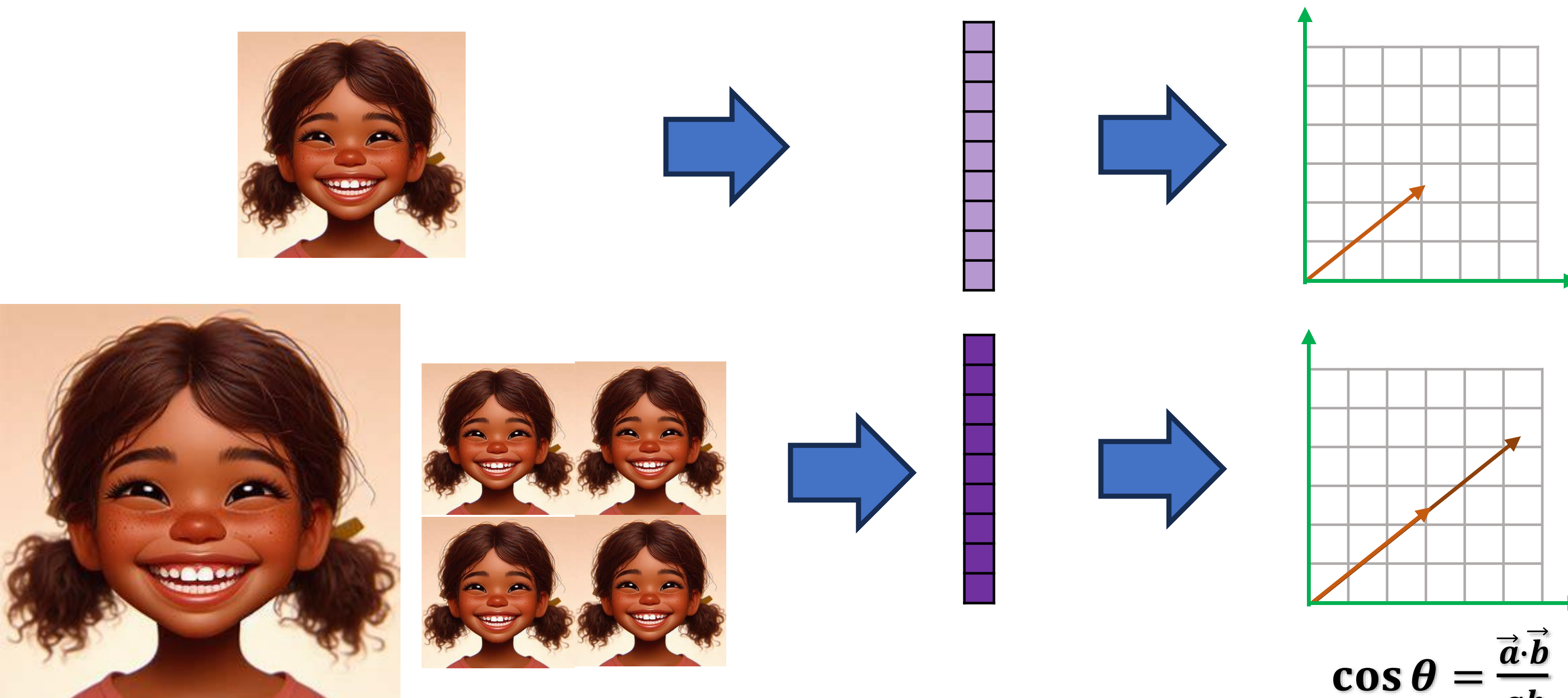




EXPERT ACADEMY

Semelhança entre dados não-estruturados

- Similaridade de cosseno entre embeddings



$$\theta = 0$$
$$\cos\theta = 1$$

$$\cos\theta = \frac{\vec{a}\cdot\vec{b}}{ab}$$