Análise Estatística de dados

Inteligência Artificial

PECE Programa de Educação Continuada
Escola Politécnica da USP

USP

"Intelligence is not just about pattern recognition and function approximation. It's about modeling the world".
— Josh Tenenbaum, NeurIPS 2021.

AULA 02 – NOÇÕES DE ÁLGEBRA LINEAR E GEOMETRIA ANALÍTICA

Larissa Driemeier

PROGRAMA DO CURSO

Aula	Data	Conteúdo da Aula
01	18/02	Aula Inaugural
02	25/02	Introdução ao Curso. Noções de Álgebra Linear , Geometria Analítica Parte I
03	11/03	Noções de Álgebra Linear , Geometria Analítica Parte II
04	18/03	Decomposição de valor singular (SVD)
05	25/03	Otimização: derivadas, derivadas parciais (operadores gradiente, Jacobiano, Hessiano e Laplaciano), algoritmos de gradiente, noções de algoritmos genéticos
06	01/04	Variáveis independentes e não independentes. Estatística Descritiva e Indutiva. Definições de medidas de dispersão e tendência central.
07	08/04	Probabilidade e Teorema de Bayes
08	15/04	Modelos de probabilidade discretos/contínuos
09	22/04	Intervalo de confiança e teste de hipóteses
10	29/04	Modelo de Markov. Modelo de Markov oculto.

AVALIAÇÃO

Soma das notas dos Trabalhos/Testes

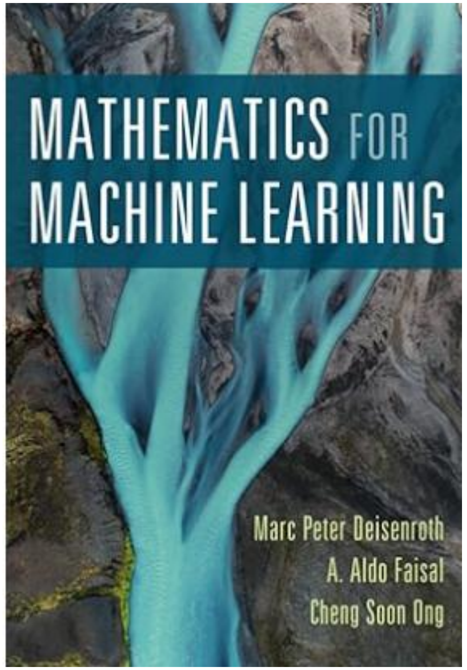
média final

$$M = \sum_i T_i \leq 10$$

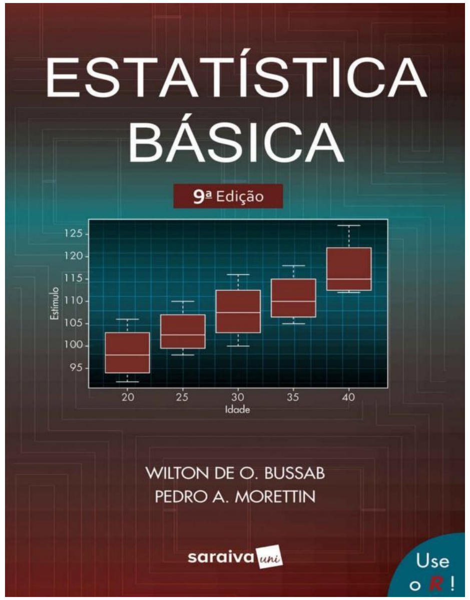
Critério de Aprovação: $M \geq 7{,}0$

Recuperação: Para alunos cuja média final estiver contida no intervalo $4{,}0 \leq M < 7{,}0$ será oferecida uma prova de recuperação, que **substituirá a nota total**, **e vale, no máximo, 7,0.**

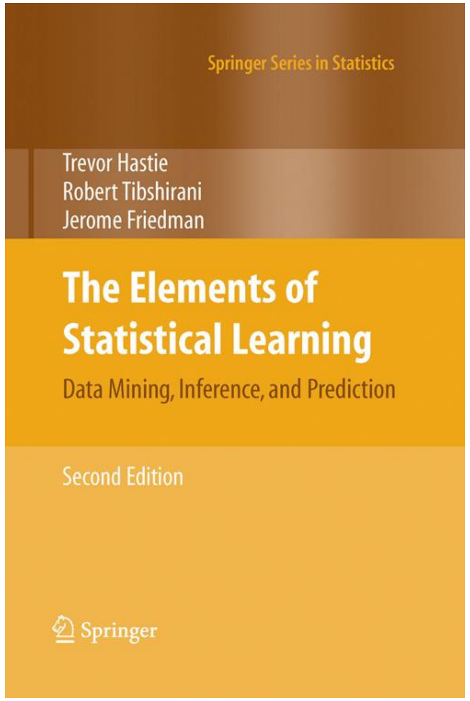
REFERÊNCIAS BIBLIOGRÁFICAS



Deisenroth, M.P.; Faisal, A.A.; Ong, C.S. Mathematics For Machine Learning, Cambridge University Press, 2020.



Hastie, T; Tibshirani, R.; Friedman, J. The Elements of Statistical Learning, 2nd edition, Springer-Verlag, 2017 (download grátis)



Bussab, W.; Morettin, P.L. Estatística Básica, Saraiva, 9ª Edição, 2023

COMUNICADOS IMPORTANTES

- Aluno Web é nosso canal de comunicação. Todo material será depositado ali: notas de aula, listas de exercícios, notas, comunicados, provas, testes ...

APRENDIZADO DE MÁQUINAS

O Aprendizado de Máquinas pode resolver:

- Regressão
- Classificação
- Agrupamento (Clustering)
- Redução de Dimensionalidade
- Estimativa de Densidade
- Previsão de Séries Temporais
- Detecção de Anomalias
- Aprendizado por Reforço
- Modelagem Generativa
- Aprendizado Baseado em Grafos
- Otimização & Tomada de Decisão

O Aprendizado de Máquinas pode resolver:

- ☑ Regressão — O modelo prevê o preço de um imóvel com base em seu tamanho e localização
- ☑ Classificação — O algoritmo identifica se um e-mail é spam ou não com base no seu conteúdo.
- ☑ Agrupamento (Clustering) — Os clientes foram agrupados em diferentes perfis de consumo sem necessidade de rótulos prévios.
- ☑ Redução de Dimensionalidade — O PCA ajudou a visualizar dados de alta dimensão
- ☑ Estimativa de Densidade — A distribuição dos dados foi estimada para entender padrões nos dados e criar modelos generativos.
- ☑ Previsão de Séries Temporais — O modelo prevê a demanda de energia para os próximos dias analisando o histórico de consumo
- ☑ Detecção de Anomalias — O sistema identificou uma transação suspeita que pode indicar fraude no cartão de crédito.
- ☑ Autonomia em robótica — Um agente de IA aprendeu a jogar xadrez sozinho, recebendo recompensas por boas jogadas
- ☑ Modelagem Generativa — Uma rede generativa criou imagens realistas de rostos que não existem na vida real.
- ☑ Aprendizado Baseado em Grafos — O algoritmo analisou conexões em uma rede social para recomendar novas amizades.
- ☑ Otimização & Tomada de Decisão — O modelo encontrou a melhor rota de entrega para minimizar custos e tempo de viagem.

ESTATÍSTICA-PROBABILIDADE-OTIMIZAÇÃO-CÁLCULO-ÁLGEBRA LINEAR

ML E ÁLGEBRA LINEAR

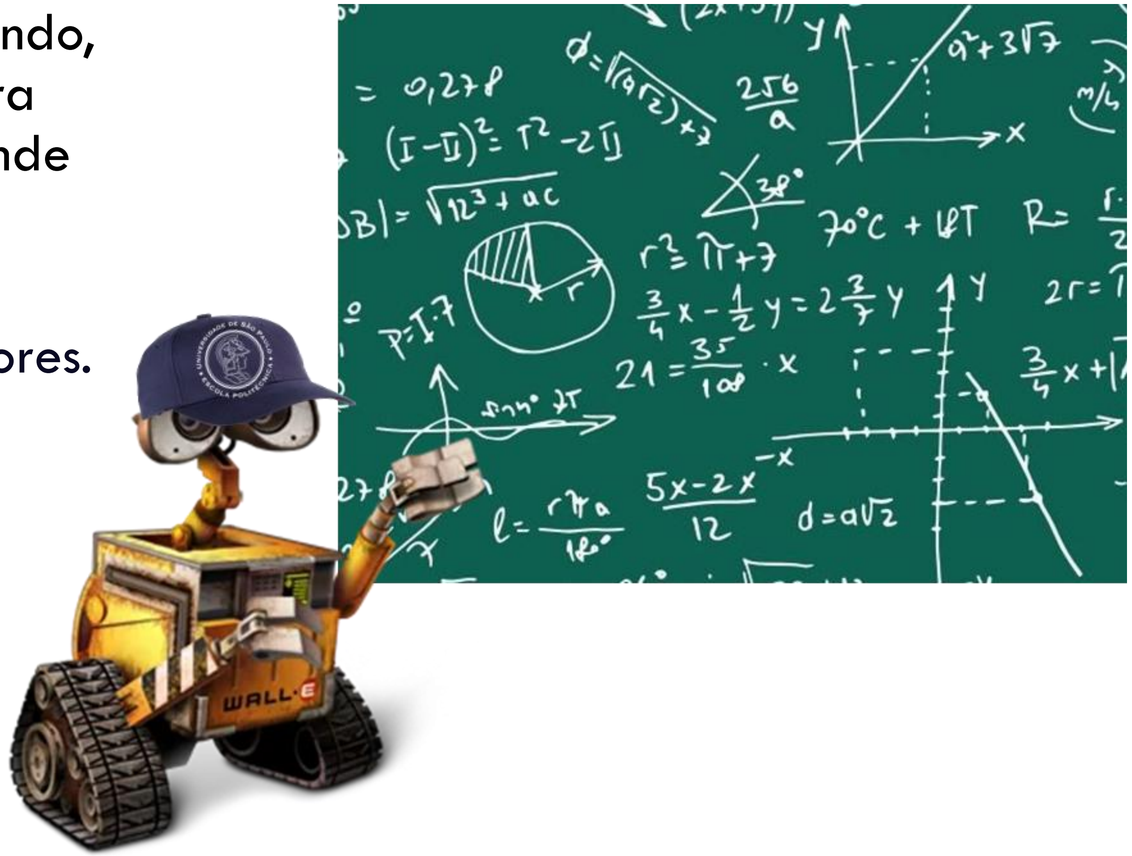
(ML, de *Machine Learning*, em inglês)

Nós temos dados. Nós sabemos que no fundo, no fundo, tudo o que o computador entende são números. Para sermos mais exatos, tudo o que o computador entende são 0s e 1s.

A Álgebra Linear é **a matemática dos arrays** — tecnicamente chamados de vetores, matrizes e tensores.

Ela é essencial em ML e Deep Learning (DL).
É a base matemática que resolve o problema de representação de dados e cálculos em modelos de ML.

Portanto, o primeiro passo para aprender matemática para ML é aprender álgebra linear.



ESTOU PERDIDO!

Felizmente, não precisamos conhecer álgebra linear em toda sua amplitude e profundidade para melhorar nossa compreensão e aplicação do aprendizado de máquina.

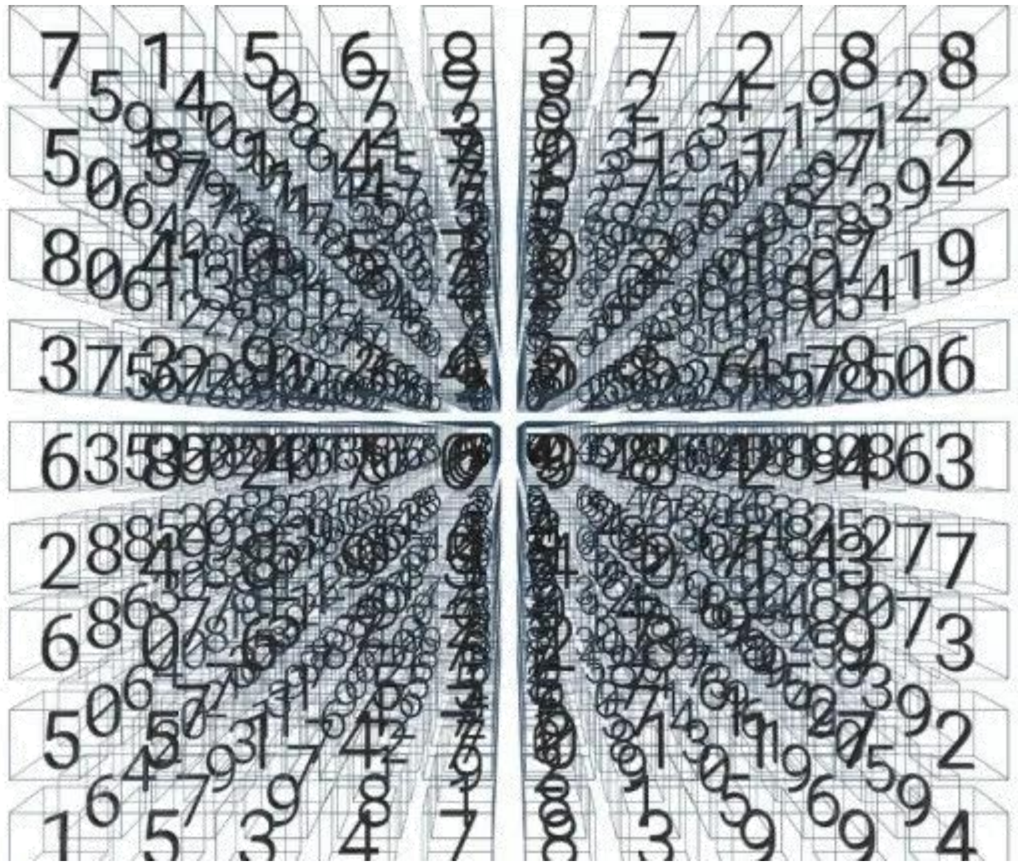
Então não será difícil!!! Você só precisará relembrar alguns conceitos!

Você não tem tempo, e nem deve, mergulhar em livros didáticos de álgebra linear e cursos on-line projetados para estudantes de graduação.

Ela deve ser encarada como mais um conjunto de ferramentas que podemos aproveitar em nossa jornada em direção ao domínio do aprendizado de máquina.

Vamos rever os conceitos básicos rapidamente e cobrir alguns tópicos importantes com maior profundidade.

PECE Programa de Educação Continuada
Escola Politécnica da USP



ESCALARES, VETORES, MATRIZES, TENSORES

Notação básica
Operações e propriedades
Cálculo matricial

$$\mathcal{A} \in \mathbb{R}^{N_1 \times N_2 \times \cdots \times N_d}$$

'A' é um tensor de ordem-d

$d = 0$
Tensor 0D (Escalar)
Shape: ()

$$= 24\ (anos)$$

$$\mathcal{A} \in \mathbb{R}^{N_1 \times N_2 \times \cdots \times N_d}$$

'A' é um tensor de ordem-d

$d = 0$
Tensor 0D (Escalar)
Shape: ()

$d = 1$
Tensor 1D (Vetor)
Shape: (3,)

Lista de atributos de um objeto!

24 anos
feminino
0 filhos
0 casa própria
3 anos com emprego fixo
2200 reais de renda mensal

$$= \begin{bmatrix} 24 \\ 0 \\ 0 \\ 0 \\ 3 \\ 2200 \end{bmatrix}$$

```python
x = np.array(42)
print(x)
print('Um escalar tem rank %d' %(x.ndim))
x.ndim,x.shape
```

O número de dimensões (eixos) do *array* e sua *forma*, respectivamente.

```
42
Um escalar tem rank 0
(0, ())
```

```python
x = np.array([24, 0, 0, 0, 3, 2200])
print(x)
print('Um vetor tem rank {:d}'.format(x.ndim))
x.ndim,x.shape
```

```
[  24    0    0    0    3 2200]
Um vetor tem rank 1
(1, (6,))
```

Retorna uma tupla onde cada índice representa o número de elementos correspondentes àquela dimensão.

```python
x.size
```

O número total de elementos

```
6
```

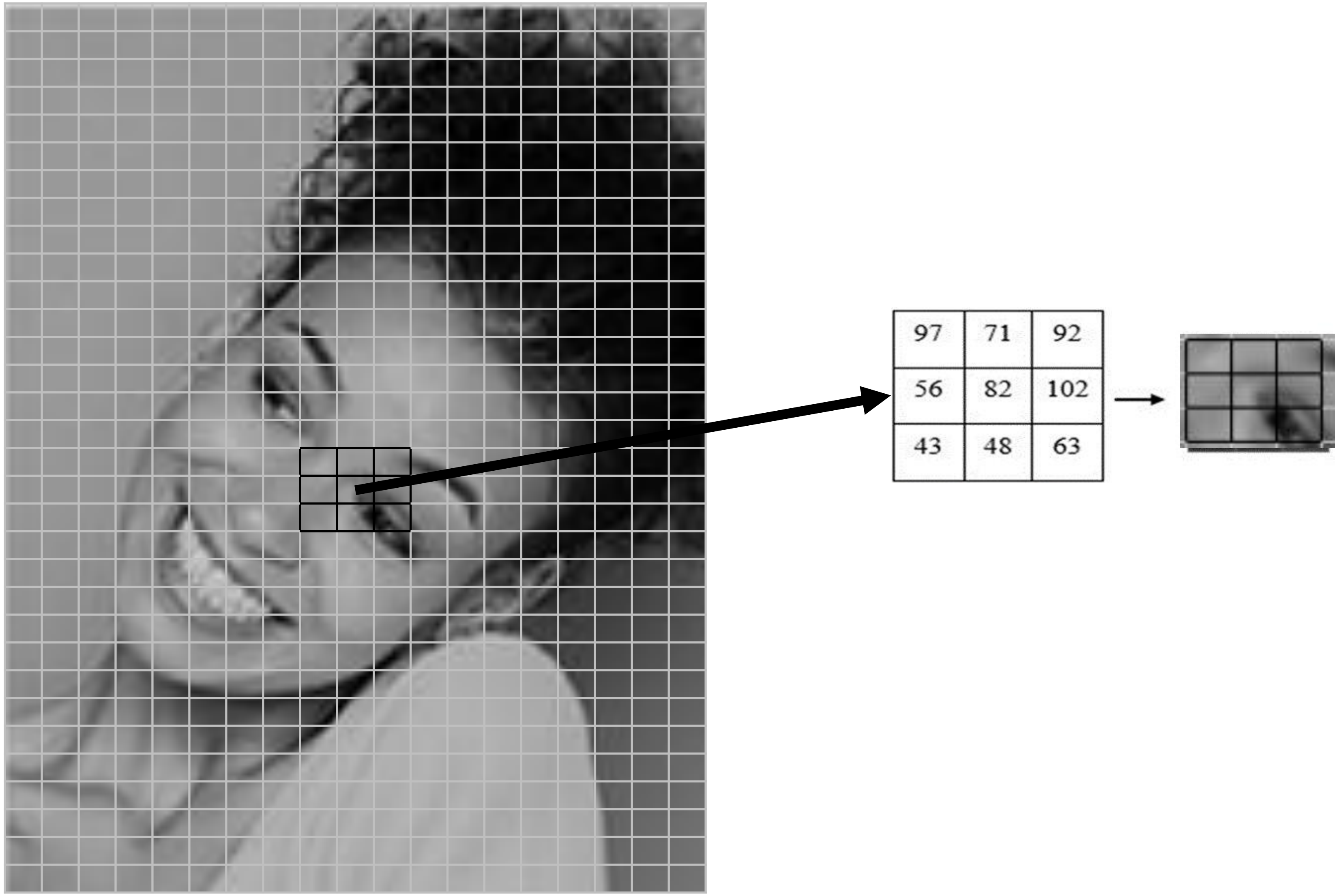
$$\mathcal{A} \in \mathbb{R}^{N_1 \times N_2 \times \cdots \times N_d}$$

'A' é um tensor de ordem-d

$d = 0$
Tensor 0D (Escalar)
Shape: ()

$d = 1$
Tensor 1D (Vetor)
Shape: (3,)

$d = 2$
Tensor 2D (Matriz)
Shape: (3,3)


97 71 92
56 82 102
43 48 63

```
[6]
x = np.array([[1, 4, 7],
              [2, 5, 8],
              [3, 6, 9]])
print(x)
print('A matriz tem rank {:d}'.format(x.ndim))
x.ndim,x.shape
```

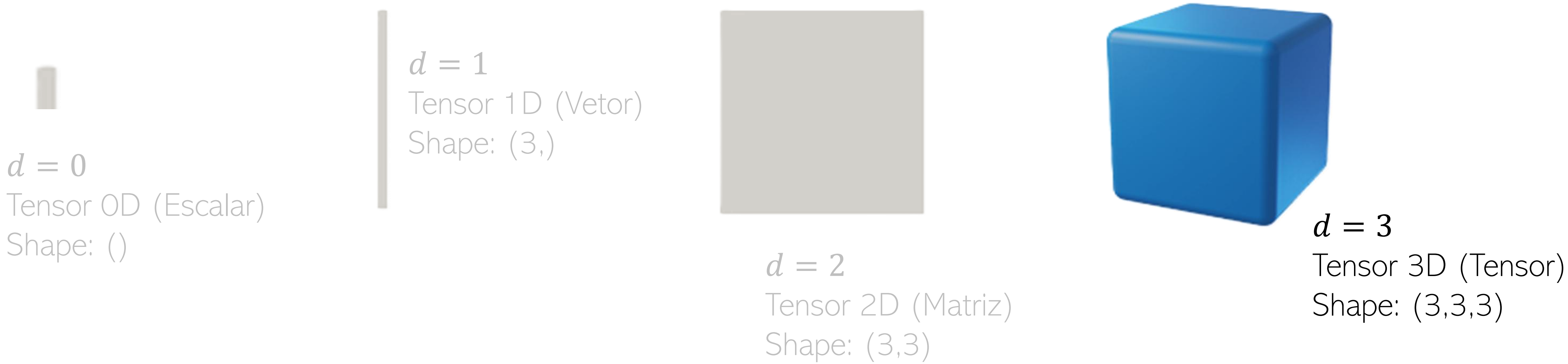
```
[[1 4 7]
 [2 5 8]
 [3 6 9]]
A matriz tem rank 2
(2, (3, 3))
```

```
[7]
x.size
```

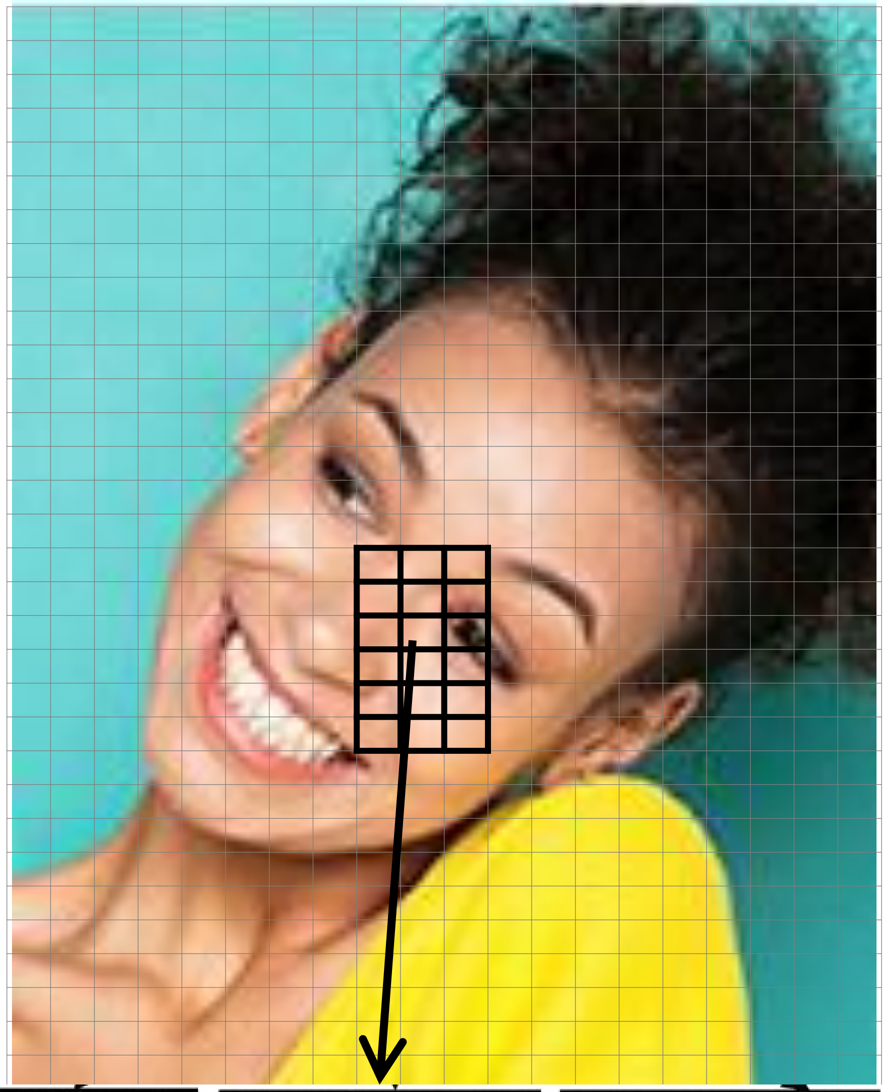
```
9
```

$$\mathcal{A} \in \mathbb{R}^{N_1 \times N_2 \times \cdots \times N_d}$$

'A' é um tensor de ordem-d



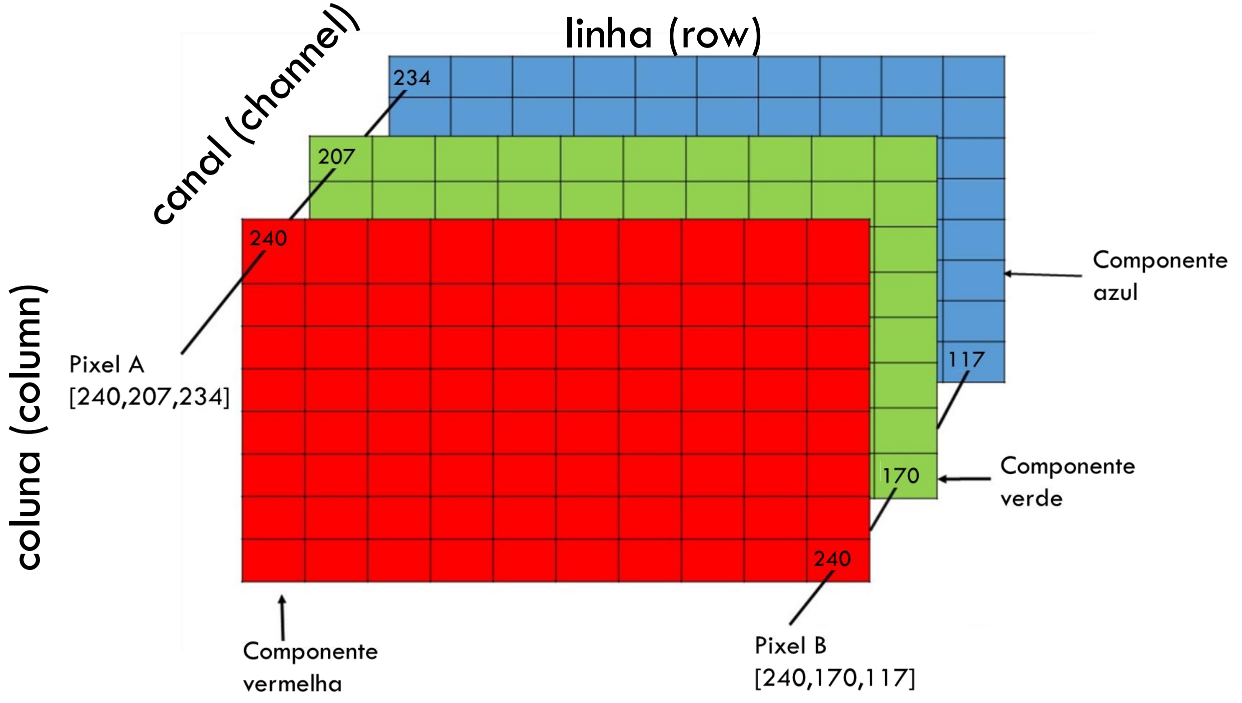
PECE Programa de Educação Continuada
Escola Politécnica da USP



240	241	241
240	237	238
239	240	240
238	237	240
240	240	239
239	240	240

207	199	196
183	163	195
183	166	184
176	172	181
184	167	176
182	180	170

234	231	225
223	213	225
219	211	195
176	205	189
168	141	117
160	142	117



```
[8]
x = np.array([[[1, 4, 7],
               [2, 5, 8],
               [3, 6, 9]],
              [[10, 40, 70],
               [20, 50, 80],
               [30, 60, 90]]])
print(x)
print('O tensor mostrado é de rank {:d}'.format(x.ndim))
x.ndim,x.shape
```

```
[[[ 1  4  7]
  [ 2  5  8]
  [ 3  6  9]]

 [[10 40 70]
  [20 50 80]
  [30 60 90]]]
O tensor mostrado é de rank 3
(3, (2, 3, 3))
```

```
print(x[1,:,:])
```

```
[[10 40 70]
 [20 50 80]
 [30 60 90]]
```

```
[9]
x.size
```

```
18
```

```
[8]
x = np.array([[[1, 4, 7],
               [2, 5, 8],
               [3, 6, 9]],
              [[10, 40, 70],
               [20, 50, 80],
               [30, 60, 90]]])
print(x)
print('O tensor mostrado é de rank {:d}'.format(x.ndim))
x.ndim,x.shape
```

```
[[[ 1  4  7]
  [ 2  5  8]
  [ 3  6  9]]

 [[10 40 70]
  [20 50 80]
  [30 60 90]]]
O tensor mostrado é de rank 3
(3, (2, 3, 3))
```

```
print(x[:,0,:])???
```

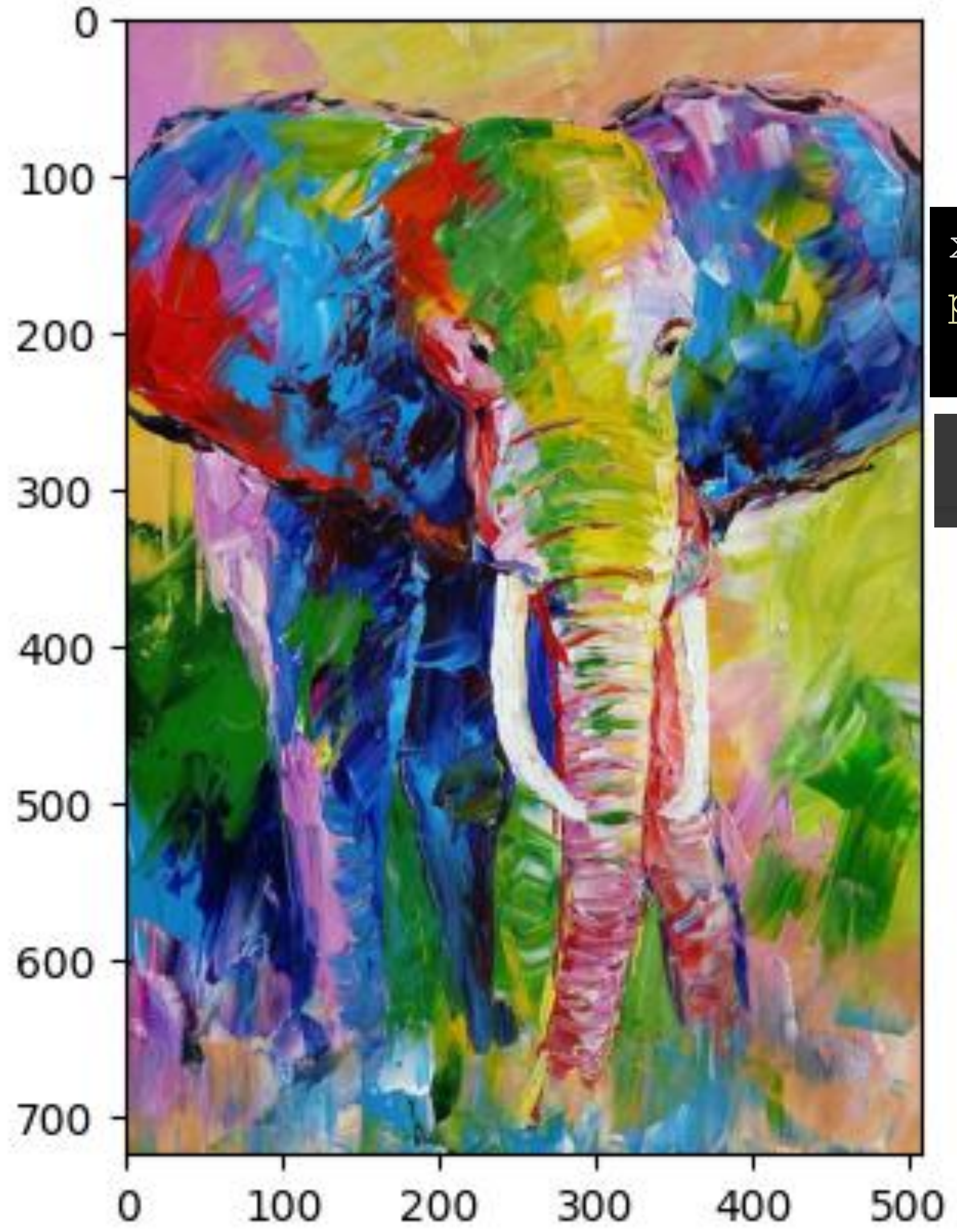
```
[9]
x.size
```

```
18
```

```
pix = img.load()
print(img.size)   # Para sabermos o número de
                    pixels (largura e altura)
data=np.array(img)
print(data.shape)
```

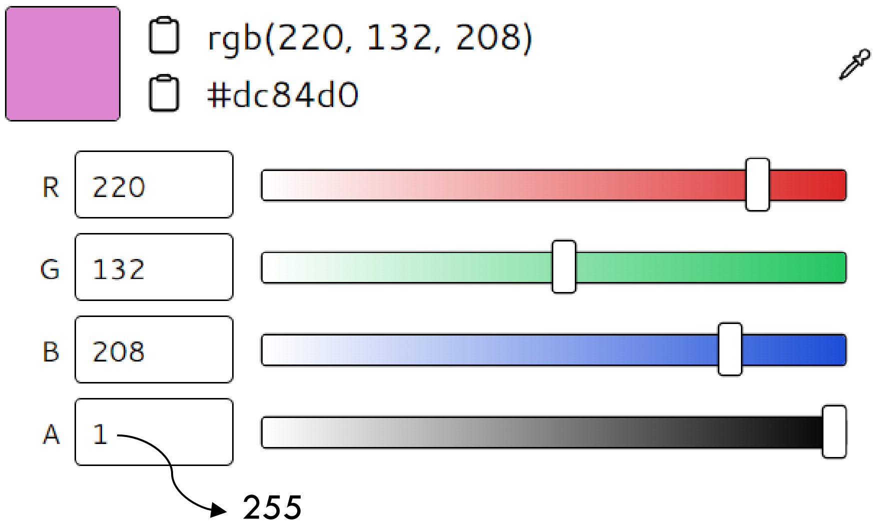
```
(507, 725)
(725, 507, 4)
```

A imagem no formato PNG possui 4 canais, R (red) B (blue), G (green) e A (alpha). O valor alfa indica a transparência ou opacidade daquele pixel específico. Os valores alfa variam de 0 a 255, onde 0 representa uma cor totalmente transparente e 255 representa uma cor totalmente opaca.




```python
x,y=0,0
print(pix[x,y])   # selecionar o valor RGBA de
                   um pixel da imagem
```

```
(220, 132, 208, 255)
```



$$\mathcal{A} \in \mathbb{R}^{N_1 \times N_2 \times \cdots \times N_d}$$

'A' é um tensor de ordem-d

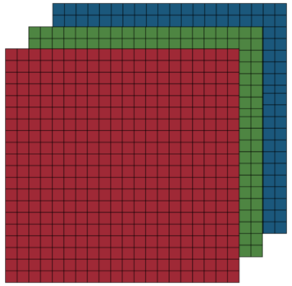
$d = 0$
Tensor 0D (Escalar)
Shape: ()

$d = 1$
Tensor 1D (Vetor)
Shape: (3,)

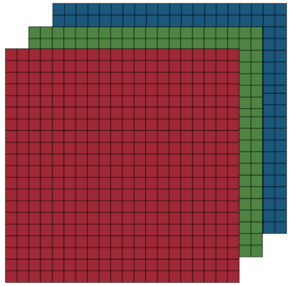
$d = 2$
Tensor 2D (Matriz)
Shape: (3,3)

$d = 3$
Tensor 3D (Tensor)
Shape: (3,3,3)

$d = 4$
Tensor 4D (Tensor)
Shape: (3,3,3,3)



$Feliz\ (0)$



$Triste\ (1)$

Tensores 4-D são muito úteis para armazenar dados para reconhecimento de imagens em deep learning.

$$\mathcal{A} \in \mathbb{R}^{N_1 \times N_2 \times \cdots \times N_d}$$

'A' é um tensor de ordem-d

$d = 0$
Tensor 0D (Escalar)
Shape: ()

$d = 1$
Tensor 1D (Vetor)
Shape: (3,)

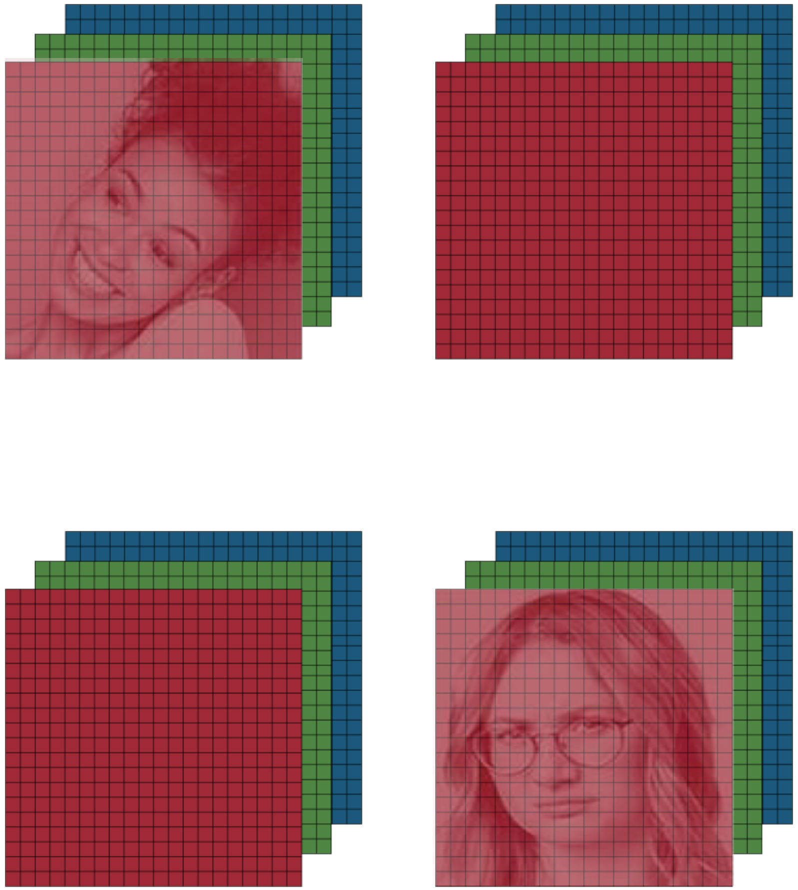
$d = 2$
Tensor 2D (Matriz)
Shape: (3,3)

$d = 3$
Tensor 3D (Tensor)
Shape: (3,3,3)

$d = 4$
Tensor 4D (Tensor)
Shape: (3,3,3,3)

$d = 5$
Tensor 5D (Tensor)
Shape: (3,3,3,3,3)

$pessoa\ 0$ $\quad$ $pessoa\ 1$

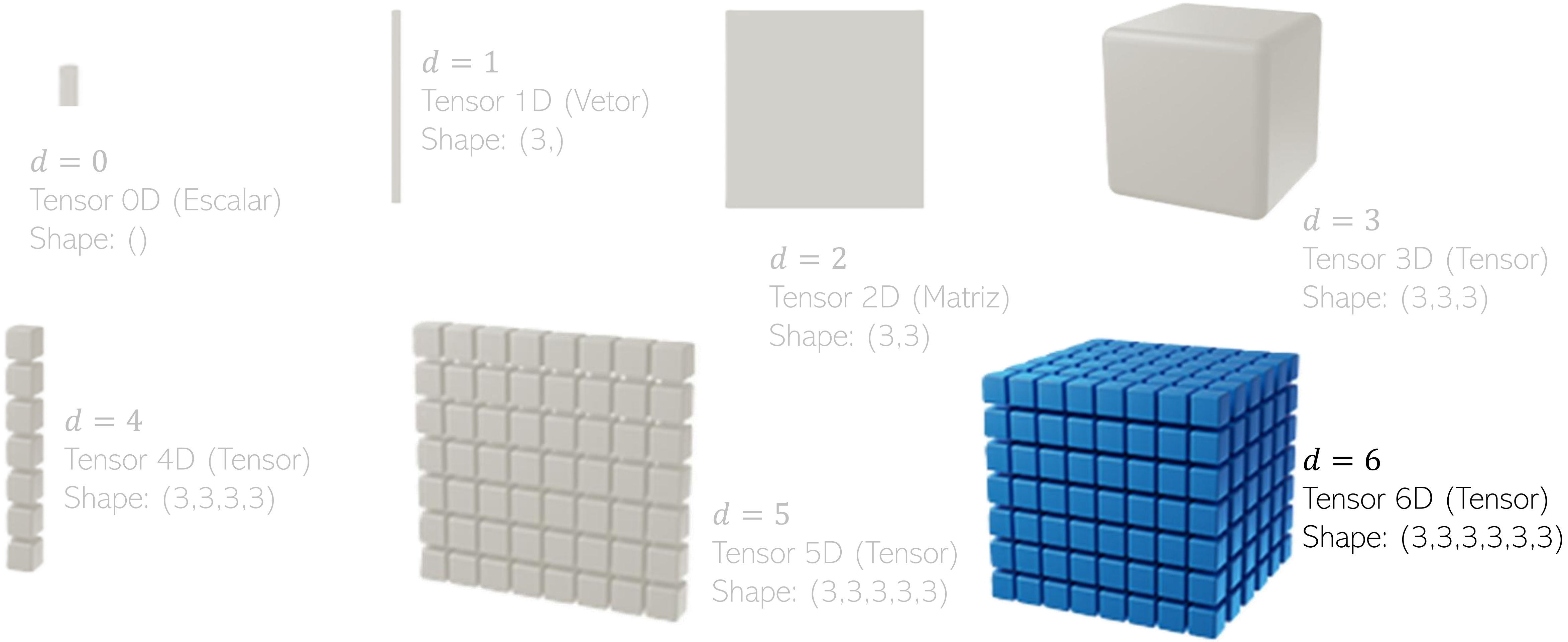


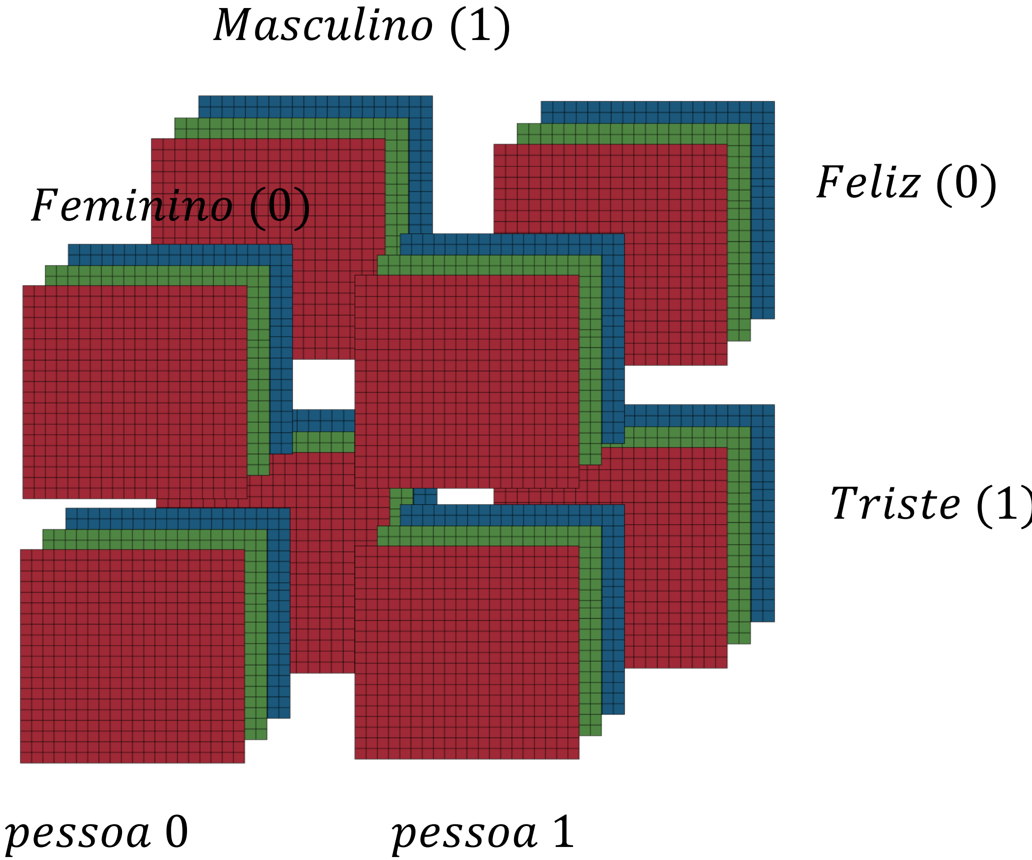
$Feliz\ (0)$

$Triste\ (1)$

$$\mathcal{A} \in \mathbb{R}^{N_1 \times N_2 \times \cdots \times N_d}$$

'A' é um tensor de ordem-d





$Masculino\ (1)$
$Feminino\ (0)$
$Feliz\ (0)$
$Triste\ (1)$
$pessoa\ 0$
$pessoa\ 1$

1D

2D

3D

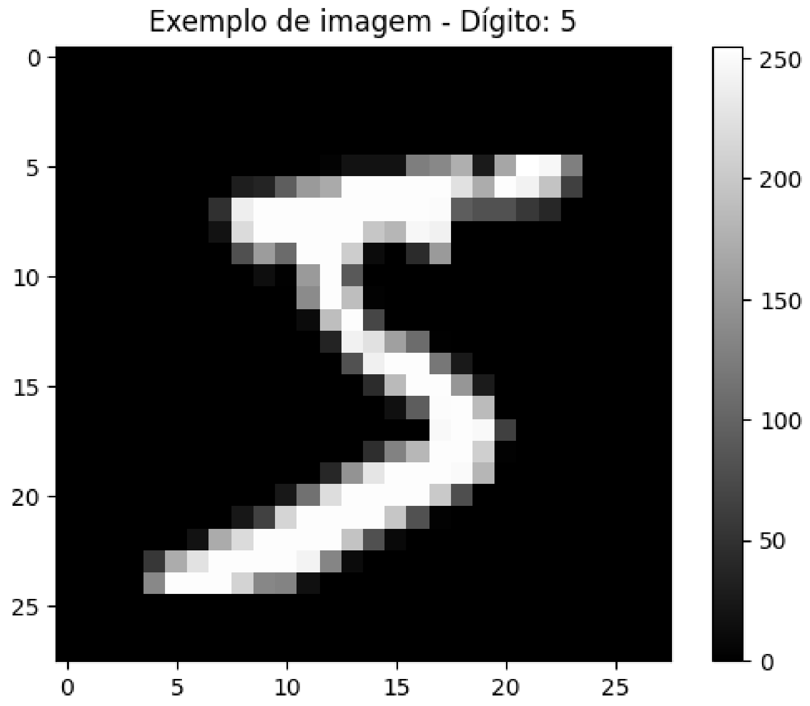
4D

5D

6D

MNIST é um banco de dados de 60.000 imagens de dígitos manuscritos, em escala de cinza (28x28 pixels). Ele é armazenado em um tensor 3-D,

$$60000 \times 28 \times 28$$



CIFAR-10 é um banco de dados de imagens coloridas de objetos. Contém 60.000 imagens coloridas (32x32 pixels) divididas em 10 categorias: avião, automóvel, pássaro, gato, cervo, cachorro, sapo, cavalo, navio, caminhão. Ele é armazenado em um tensor 3-D,

$$60000 \times 32 \times 32 \times 3$$

Vamos considerar um exemplo de um vídeo de 5 minutos com resolução HD 1080. A dimensão da estrutura de dados pode ser calculada da seguinte forma:

- O tamanho do quadro (frame) é de **1080 x 1920 pixels.**
- A duração do vídeo é de **5 minutos**, ou seja, **5 × 60 = 300 segundos**.
- Se o vídeo for amostrado a **10 quadros por segundo**, o número total de quadros será **300 × 10 = 3000 quadros.**
- Suponha que o vídeo tenha uma **profundidade de cor de 3 canais (RGB)**

...

Isso significa que temos um total de **3000 quadros**, cada um com uma matriz de **1080 x 1920 pixels**, e cada pixel é representado por **3 valores** (um para cada canal de cor: vermelho, verde e azul). **Portanto, um único vídeo pode ser representado como um tensor 4D**

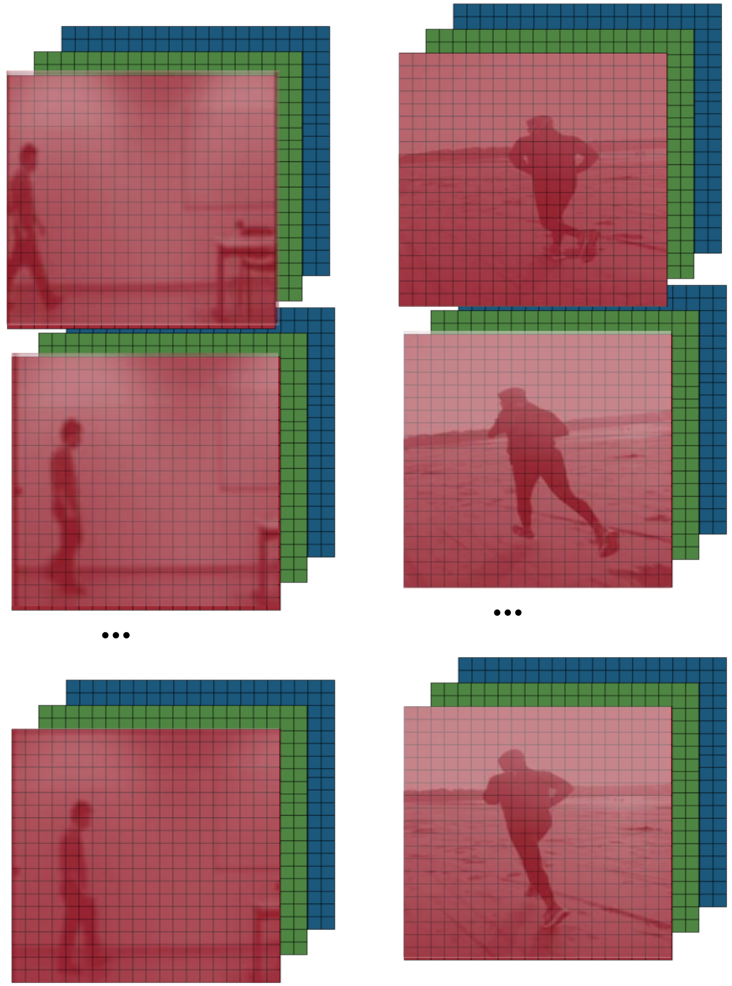
(3000, 1080, 1920, 3)

Se quisermos armazenar múltiplos vídeos, como **10 clipes de vídeo** com resolução **HD 1080**, precisaríamos de um **tensor 5D:**

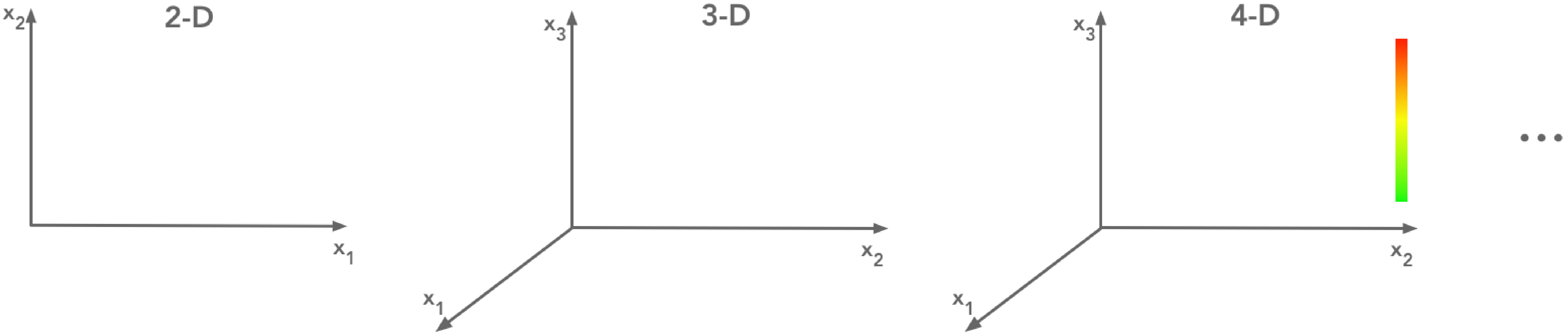
(10, 3000, 1080, 1920, 3)

Isso significa que o tensor contém **10 vídeos**, cada um sendo um tensor **4D (3000, 1080, 1920, 3).**

- **10** representa o número de vídeos,
- **3000** é o número de quadros por vídeo,
- **1080** e **1920** são a altura e largura de cada quadro,
- **3** representa os canais de cor (RGB).



REPRESENTAÇÃO ESPACIAL



DESAFIO

1. Qual o nível de cor verde do pixel (350,250) do elefante?

2. Como tirar a componente de cor vermelha de todos os pixels do elefante?

3. Como deixar o elefante transparente?

4. Rode o seguinte código e descreva o que a função `Image.thumbnail` fez com a imagem do elefante:

```
img.thumbnail((64, 64), Image.LANCZOS)
# resizes image in-place
imgplot = plt.imshow(img);
data=np.array(img)
data.shape
```

VETORES

VETORES

O vetor $\boldsymbol{x} \in \mathbb{R}^n$ é dito vetor coluna de dimensão n se possuir n linhas e 1 coluna

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$$

x_i é o i-ésimo elemento do vetor x

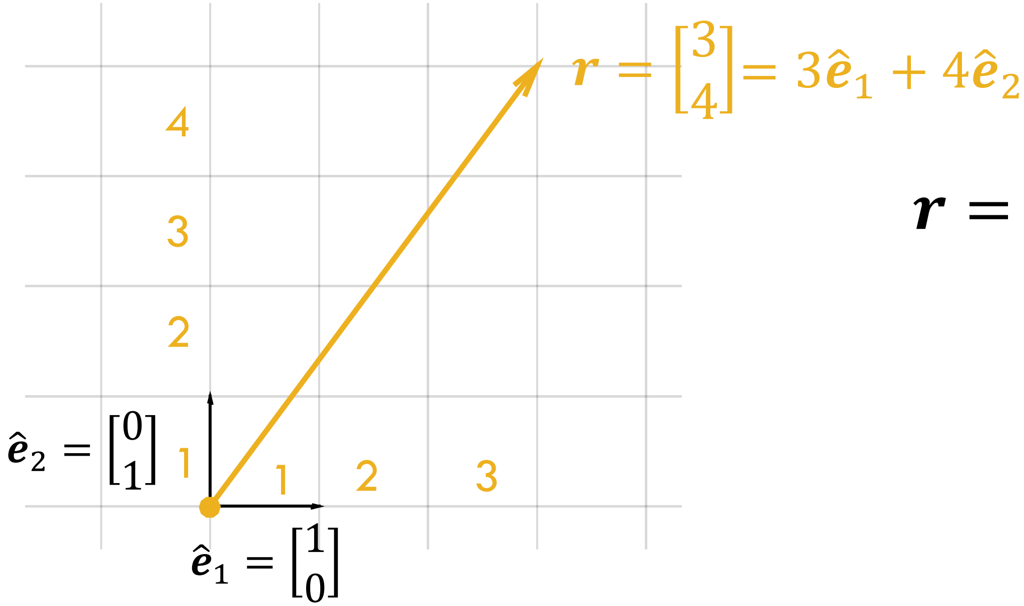
$$\begin{bmatrix} 1 \\ 2 \end{bmatrix}^T = \begin{bmatrix} 1 & 2 \end{bmatrix}$$

Para expressar o vetor linha,

$$\boldsymbol{x}^T = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}$$

calcula-se o transposto de $\boldsymbol{x}$, isto é, um vetor de 1 linha e n colunas.

BASE DE UM CONJUNTO DE VETORES



$$\boldsymbol{r} = r_1\hat{\boldsymbol{e}}_{\boldsymbol{1}} + r_2\hat{\boldsymbol{e}}_{\boldsymbol{2}}$$

Base de vetores (bidimensional) de nosso sistema de coordenadas.

"VETOR DE CARACTERÍSTICAS"

Para nosso universo de IA, vetor é uma lista de atributos de um objeto!

450 m^2
5 quartos
6 banheiro
850000 reais

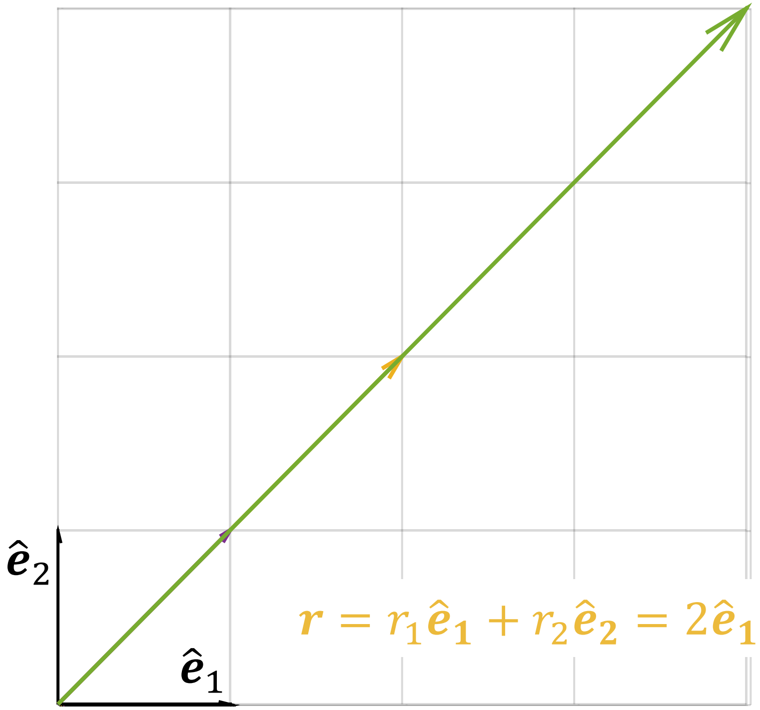
$$\begin{bmatrix} 450 \\ 5 \\ 6 \\ 850 \end{bmatrix}$$

OPERAÇÕES BÁSICAS COM VETORES

$\frac{\boldsymbol{r}}{2} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\boldsymbol{r} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ #quartos #banheiros

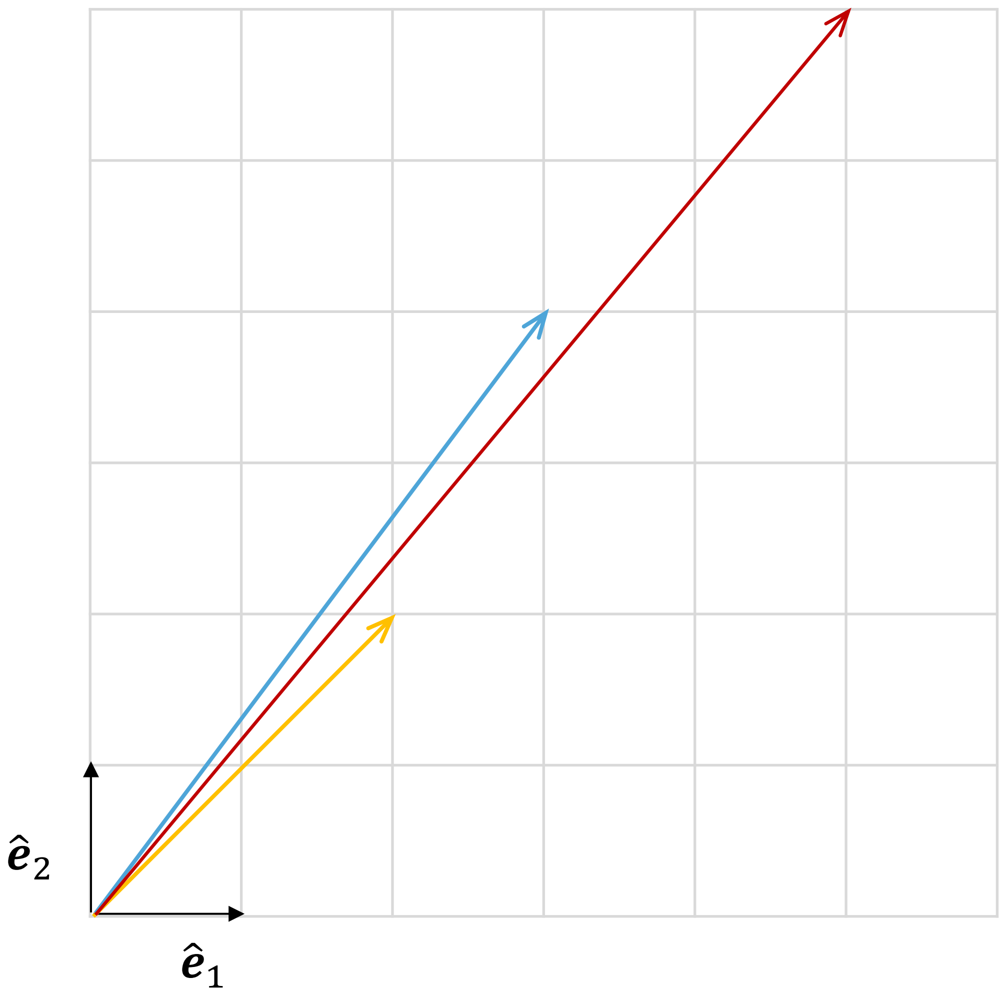
$2\boldsymbol{r} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$

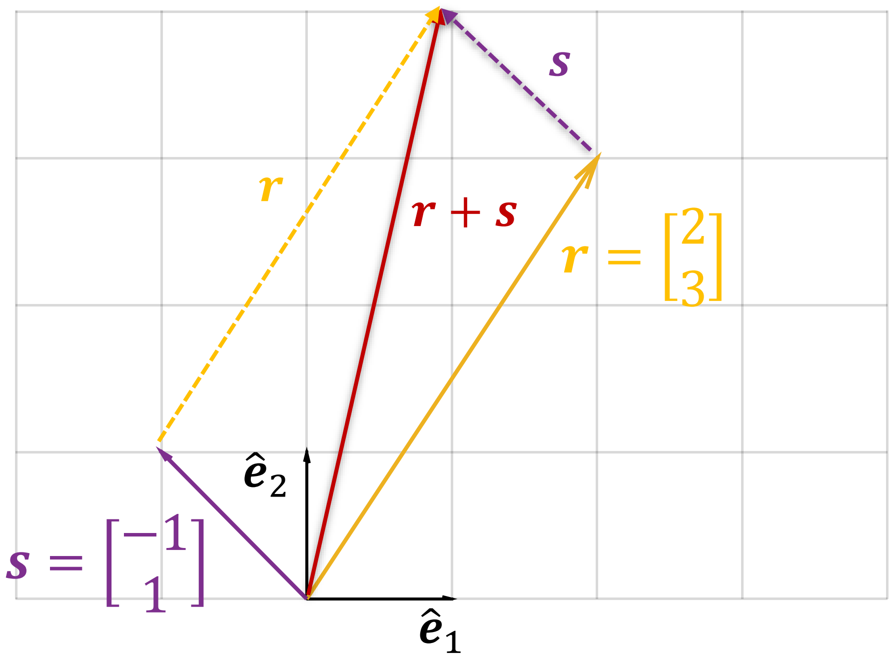


$\boldsymbol{r} = r_1\hat{\boldsymbol{e}}_1 + r_2\hat{\boldsymbol{e}}_2 = 2\hat{\boldsymbol{e}}_1 + 2\hat{\boldsymbol{e}}_2$

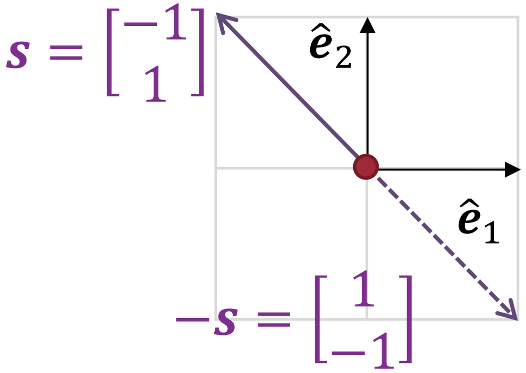
$$4 \times \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \times 4 \\ 2 \times 4 \\ 3 \times 4 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \\ 12 \end{bmatrix}$$

$$\boldsymbol{r}' = \begin{bmatrix}2\\2\end{bmatrix} + \begin{bmatrix}3\\4\end{bmatrix} = \begin{bmatrix}5\\6\end{bmatrix}$$

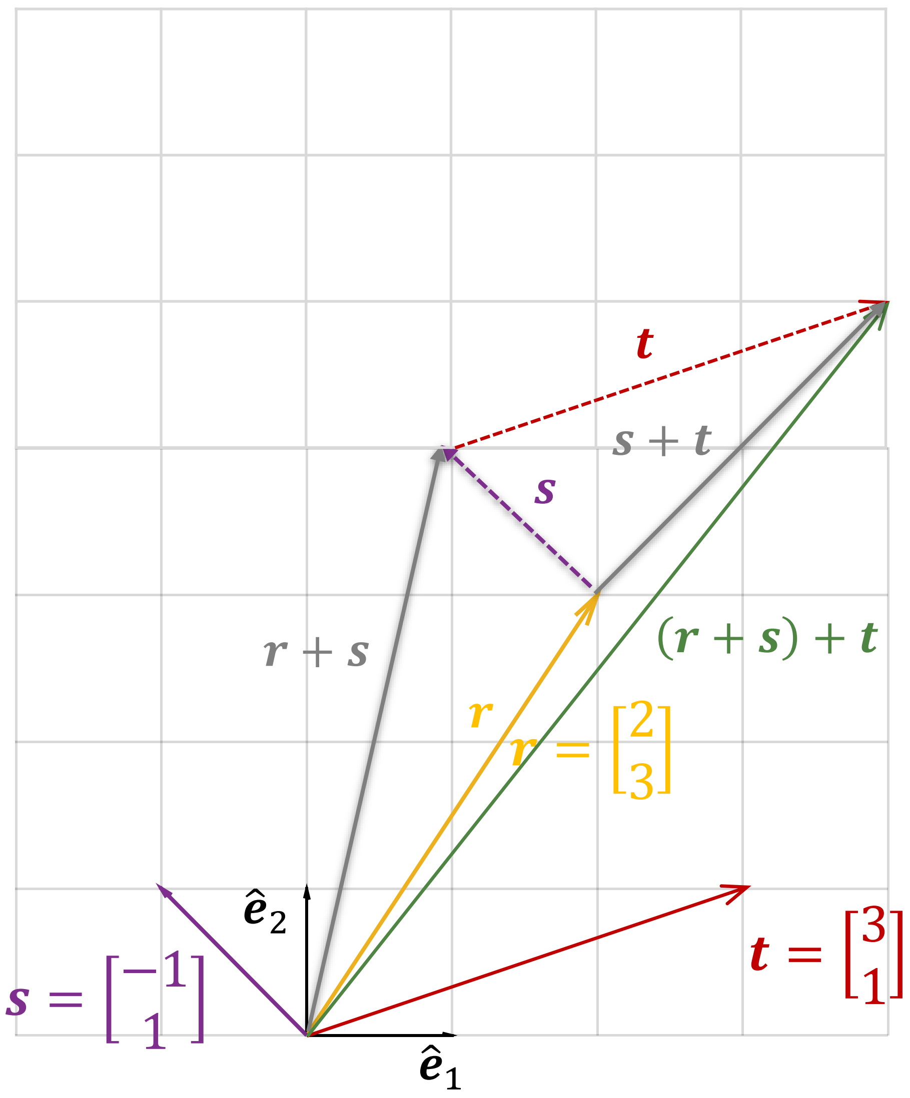




$$s + r = r + s = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$



$$s + (-s) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$



$$\boldsymbol{r}+\boldsymbol{s}+\boldsymbol{t}=\begin{bmatrix}-1\\1\end{bmatrix}+\begin{bmatrix}2\\3\end{bmatrix}+\begin{bmatrix}3\\1\end{bmatrix}=\begin{bmatrix}4\\5\end{bmatrix}$$

$$(\boldsymbol{r}+\boldsymbol{s})+\boldsymbol{t}\ =\boldsymbol{r}+(\boldsymbol{s}+\boldsymbol{t})$$

$$\left(\begin{bmatrix}2\\3\end{bmatrix}+\begin{bmatrix}-1\\1\end{bmatrix}\right)+\begin{bmatrix}3\\1\end{bmatrix}=\begin{bmatrix}2\\3\end{bmatrix}+\left(\begin{bmatrix}-1\\1\end{bmatrix}+\begin{bmatrix}3\\1\end{bmatrix}\right)$$

PRODUTO INTERNO ENTRE VETORES

O produto escalar é a principal ferramenta para calcular projeções de vetores, decomposições de vetores e determinar ortogonalidade.

$\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ calcula-se $\boldsymbol{x}^T \boldsymbol{y} \in \mathbb{R}$

$$\boldsymbol{v} = \boldsymbol{x}^T \boldsymbol{y} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \sum_{i=1}^{n} x_i y_i$$

$$\boldsymbol{x}^T \boldsymbol{y} = \boldsymbol{y}^T \boldsymbol{x}$$

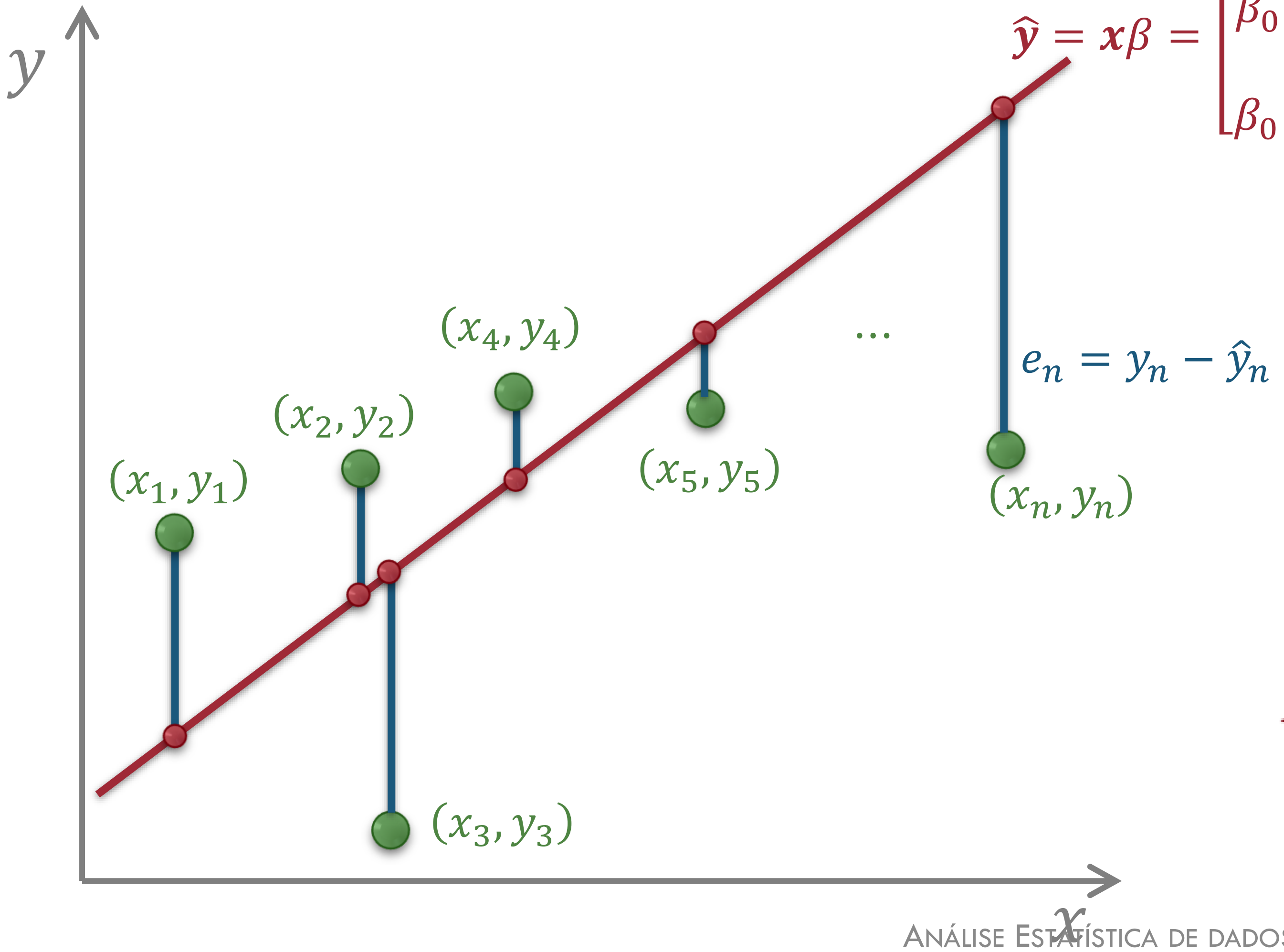
POR EXEMPLO...

Ache o produto interno entre os seguintes vetores,

$$\boldsymbol{r} = \begin{bmatrix} -5 \\ 3 \\ 2 \\ 8 \end{bmatrix} \quad \boldsymbol{s} = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix}$$

ERRO MÉDIO QUADRÁTICO E O PRODUTO INTERNO

PECE Programa de Educação Continuada
Escola Politécnica da USP



$$\hat{\boldsymbol{y}} = \boldsymbol{x}\beta = \begin{bmatrix} \beta_0 + \beta_1 x_1 \\ \beta_0 + \beta_1 x_2 \\ \vdots \\ \beta_0 + \beta_1 x_n \end{bmatrix}$$

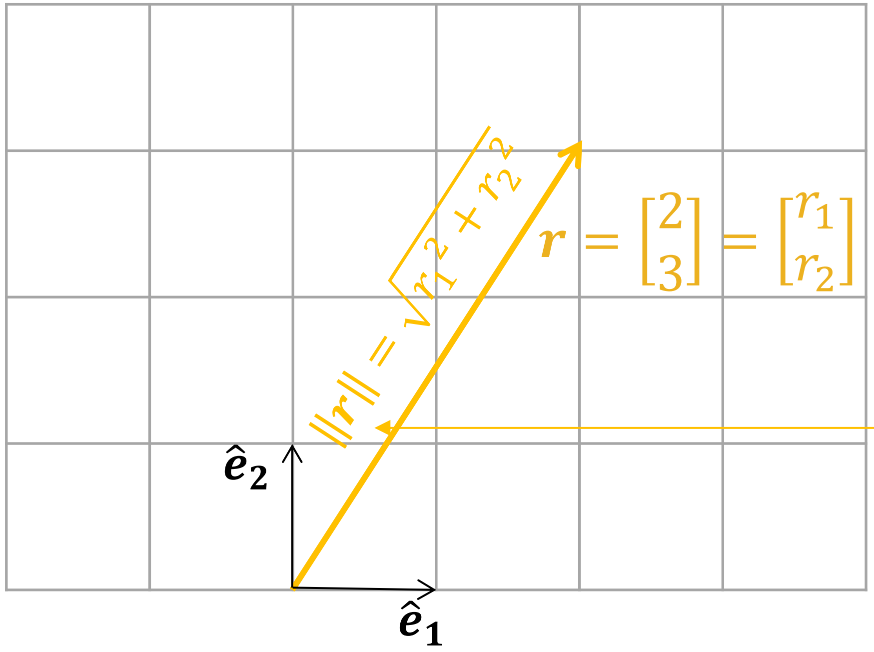
Vetor $n \times 1$ que contém as previsões pontuais

$$\boldsymbol{e}(\beta_0, \beta_1) = \boldsymbol{y} - \boldsymbol{x}\beta = \boldsymbol{y} - \hat{\boldsymbol{y}}$$

$$MSE(\beta) = \frac{1}{n}\sum_{i=1}^{n} e_i^2(\beta)$$

$$MSE(\beta) = \frac{1}{n}[e_1 e_2 \ldots e_n]\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

CARACTERÍSTICAS DO PRODUTO INTERNO



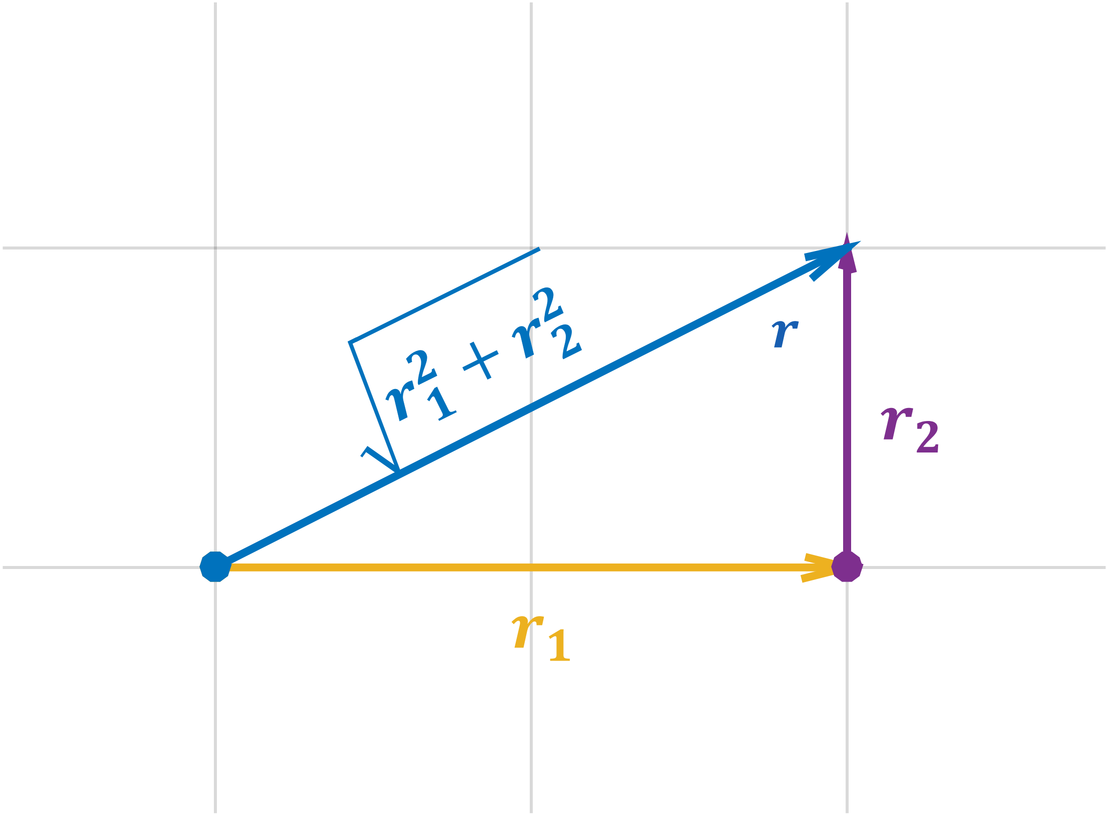
$$\boldsymbol{r} = r_1 \hat{\boldsymbol{e}}_1 + r_2 \hat{\boldsymbol{e}}_2 = 2\hat{\boldsymbol{e}}_1 + 3\hat{\boldsymbol{e}}_2$$

$$\boldsymbol{r} \cdot \boldsymbol{r} = \begin{bmatrix} r_1 & r_2 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}$$

$$\boldsymbol{r} \cdot \boldsymbol{r} = r_1^2 + r_2^2 = \left(\boxed{\sqrt{r_1^2 + r_2^2}} \right)^2 = \|\boldsymbol{r}\|^2$$

Se queremos saber a dimensão de um vetor, basta tirar a raiz do seu produto interno.

EXTRAPOLANDO...

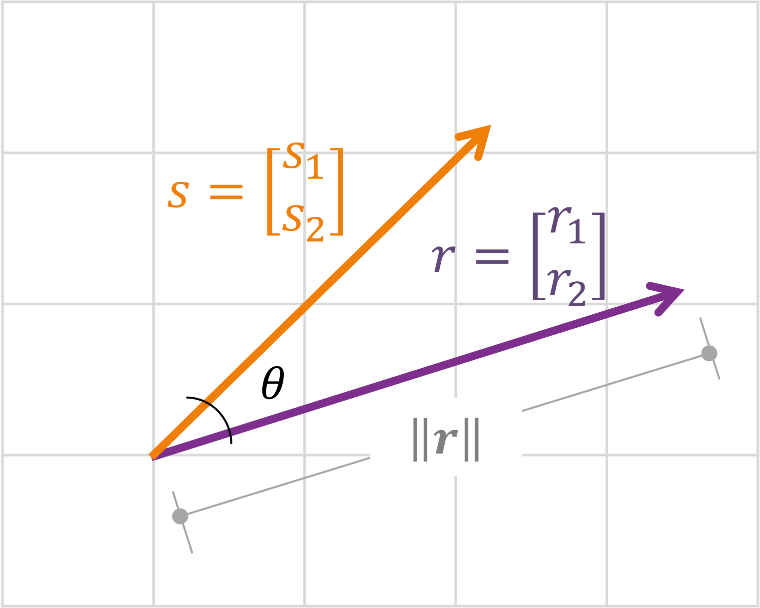


Podemos extender para qualquer dimensão: o tamanho de um vetor é a raíz quadrada da soma dos quadrados de seus componentes. Por exemplo,

$$\boldsymbol{s} = \begin{bmatrix} 1 \\ 3 \\ 5 \\ 1 \end{bmatrix}$$

$$\|\boldsymbol{s}\| = \sqrt{36} = 6$$

PRODUTO INTERNO E O ÂNGULO ENTRE VETORES



$$\boldsymbol{r} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}, \boldsymbol{s} = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

$$\boldsymbol{r}^T \boldsymbol{s} = \begin{bmatrix} r_1 & r_2 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = r_1\, s_1 + r_2\, s_2$$

$$\boldsymbol{r} \cdot \boldsymbol{r} = \boldsymbol{r}^T \boldsymbol{r} = \begin{bmatrix} r_1 & r_2 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = r_1 r_1 + r_2 r_2 = r_1^2 + r_2^2 = \|\boldsymbol{r}\|^2$$

Relação entre produto interno e ângulo,

$$\boldsymbol{r} \cdot \boldsymbol{s} = \boldsymbol{r}^T \boldsymbol{s} = \|\boldsymbol{r}\| \|\boldsymbol{s}\| \cos\theta$$

$$(\boldsymbol{r} \cdot \boldsymbol{s} = \boldsymbol{r}^T \boldsymbol{s} = \boldsymbol{s}^T \boldsymbol{r})$$

VETORES ORTOGONAIS E ORTONORMAIS

Dois vetores são ortogonais, $\cos\theta = 90^o$, se seu produto interno é nulo,

$$\frac{\boldsymbol{r}\cdot\boldsymbol{s}}{\|\boldsymbol{r}\|\|\boldsymbol{s}\|} = \cos\theta \rightarrow \boldsymbol{r}\cdot\boldsymbol{s} = 0$$

Um vetor é normalizado se sua dimensão é unitária, ié,

$$\|\boldsymbol{r}\| = 1$$

Podemos normalizar um vetor...

$$\boldsymbol{r}' = \frac{\boldsymbol{r}}{\|\boldsymbol{r}\|}, \boldsymbol{s}' = \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|}, \boldsymbol{r}'\cdot\boldsymbol{s}' = 0$$

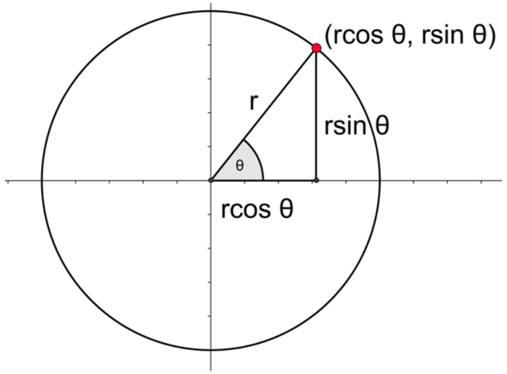
Dois vetores $\boldsymbol{r}, \boldsymbol{s}$ são ditos ortogonais quando separados por 90^o.

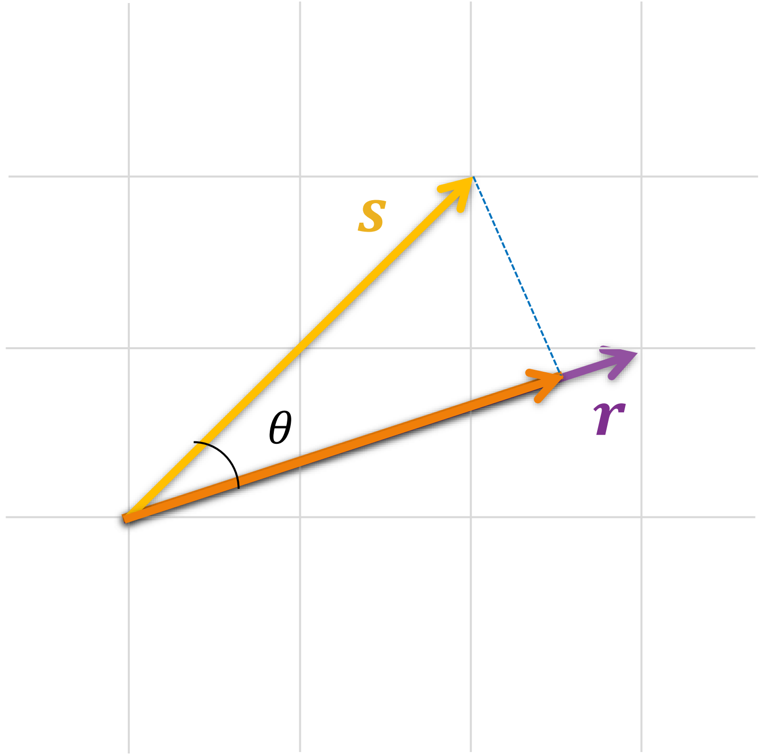
Vetor unitário ou versor é aquele cuja norma é 1 .

Quando a norma dos vetores ortogonais é unitária, eles são chamados ortonormais $(\boldsymbol{r}', \boldsymbol{s}')$.

É impossível ter mais de n vetores independentes mutuamente ortogonais em $\mathbb{R}^n$

PROJEÇÃO ESCALAR E VETORIAL





Projeção escalar é o *tamanho* do vetor laranja.

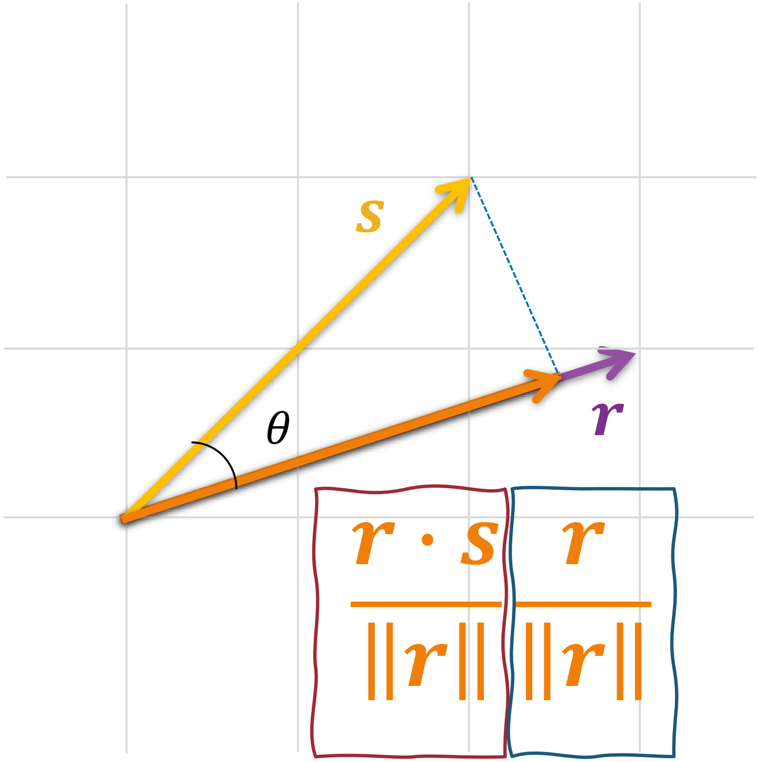
$$\frac{\boldsymbol{r}\cdot\boldsymbol{s}}{\|\boldsymbol{r}\|} = \|\boldsymbol{s}\|\cos\theta$$

Projeção vetorial é um escalonamento do vetor $\boldsymbol{r}$

$$\boxed{\frac{\boldsymbol{r}\cdot\boldsymbol{s}}{\|\boldsymbol{r}\|}\frac{\boldsymbol{r}}{\|\boldsymbol{r}\|} = \frac{\boldsymbol{r}\cdot\boldsymbol{s}}{\boldsymbol{r}\cdot\boldsymbol{r}}\boldsymbol{r}}$$

A definição de projeção escalar é o comprimento da projeção vetorial.

RESUMINDO...



$$\frac{\mathbf{r}\cdot\mathbf{s}}{\|\mathbf{r}\|}\quad\frac{\mathbf{r}}{\|\mathbf{r}\|}$$

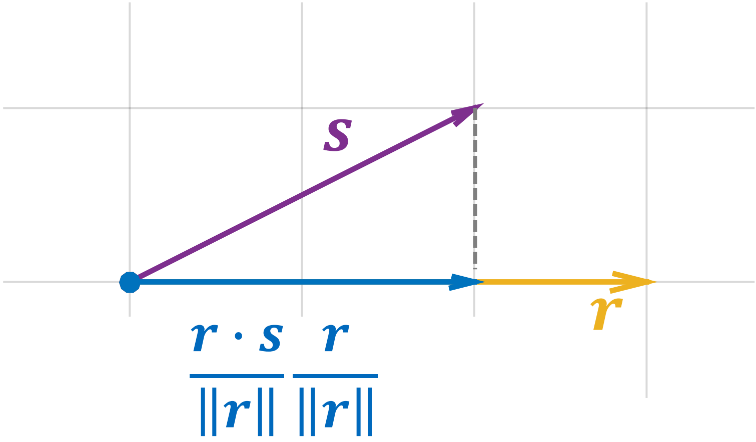
Dimensão do vetor

Direção do vetor

EXEMPLO

A projeção pode ser feita em qualquer dimensão. Por exemplo,

$$\boldsymbol{r} = \begin{bmatrix} 3 \\ -4 \\ 0 \end{bmatrix} \quad \boldsymbol{s} = \begin{bmatrix} 10 \\ 5 \\ -6 \end{bmatrix}$$



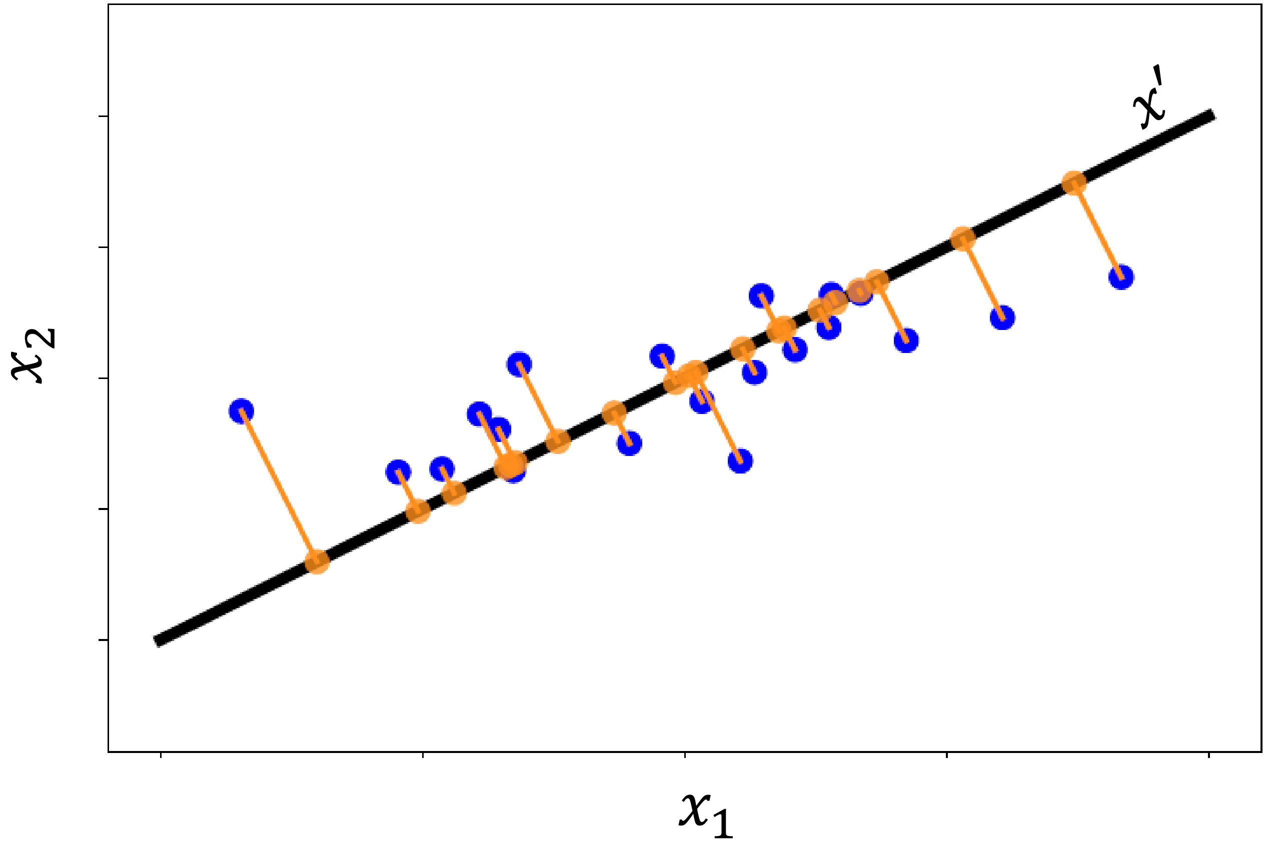
Qual a projeção escalar e vetorial de s em r?

REDUÇÃO DE DIMENSIONALIDADE

As projeções são uma classe importante de transformações lineares.

No aprendizado de máquina, geralmente lidamos com dados de alta dimensão, que são geralmente são difíceis de analisar ou visualizar.

No entanto, dados de alta dimensão muitas vezes possuem a propriedade de que apenas algumas dimensões contêm mais informações, e a maioria das outras dimensões não são essenciais para descrever as principais propriedades dos dados.

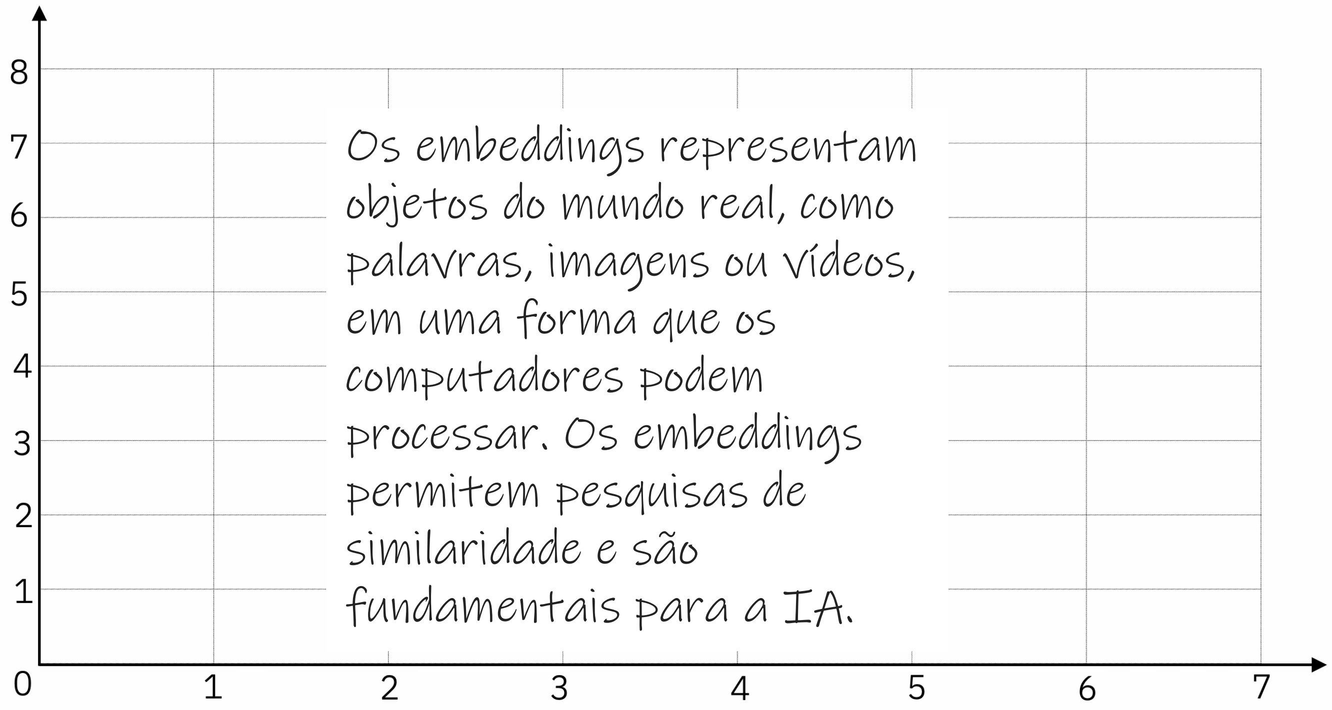


EMBEDDING

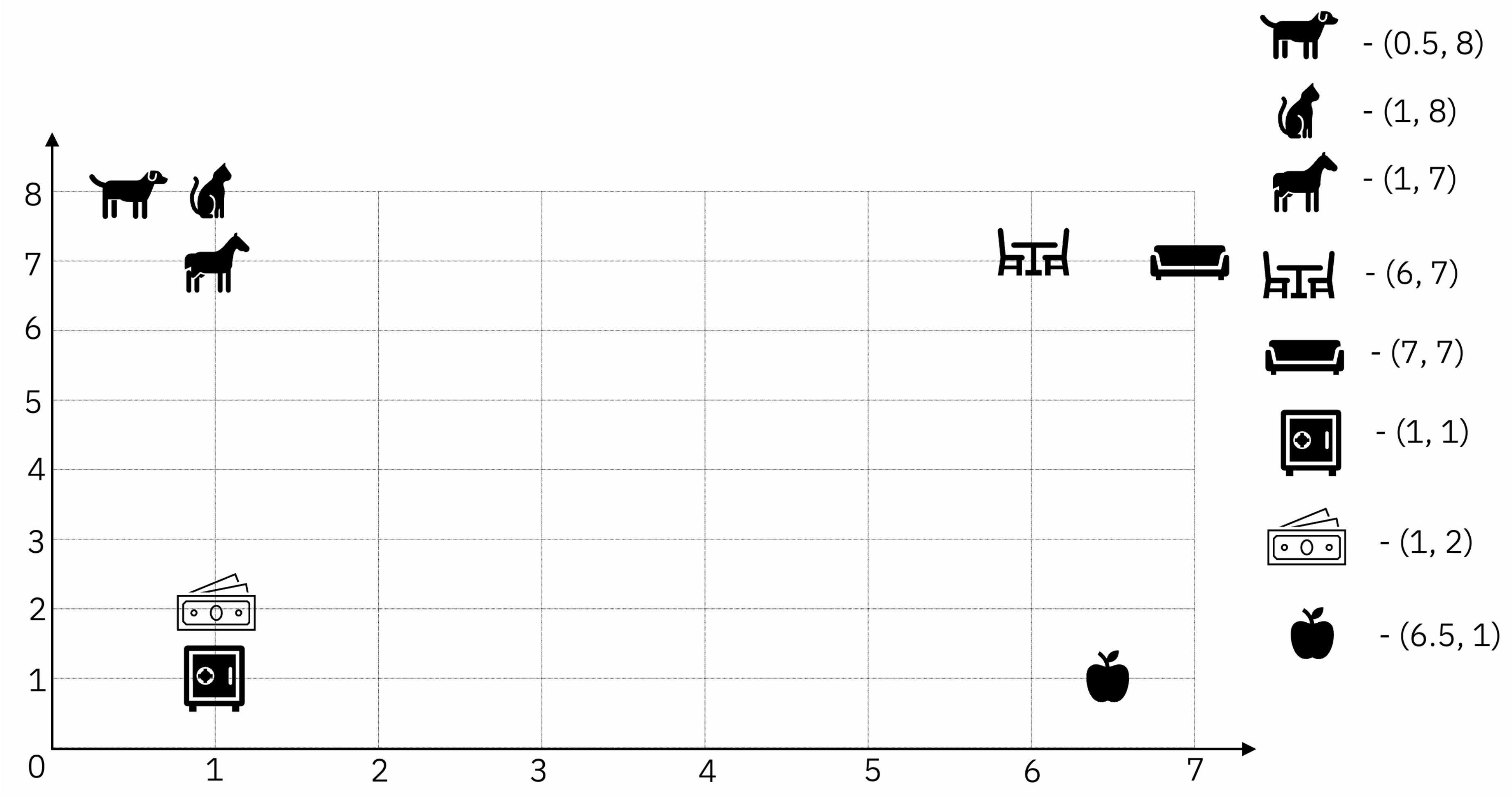
Todo modelo é, por baixo dos panos, um modelo matemático. E nós só fazemos matemática com números. Isso se aplica aos modelos estatísticos clássicos.

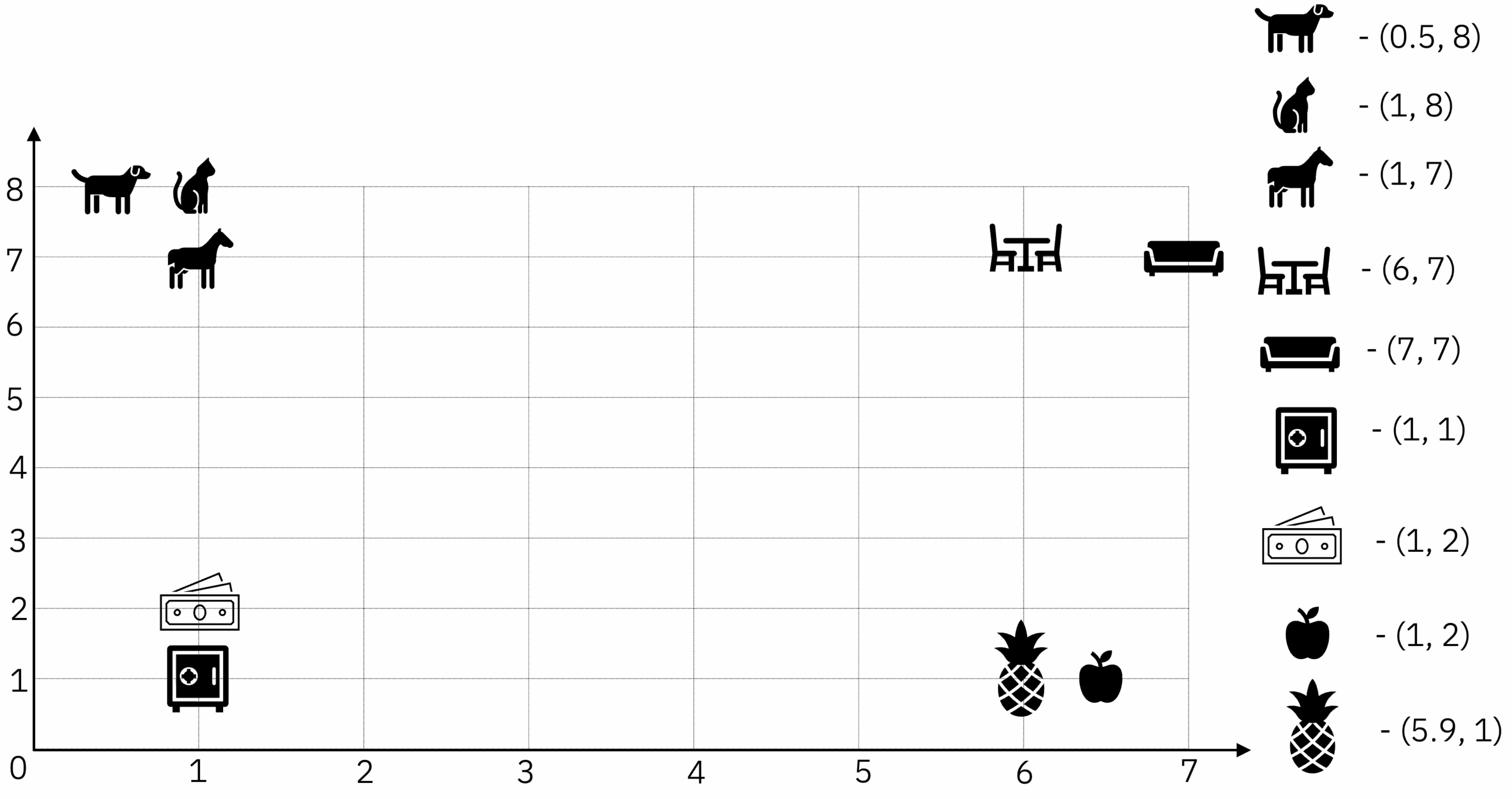
Se aplica aos modelos de Visão Computacional, que transformam pixels em números.

Se aplica também a modelos de linguagem transformando texto em números com Token e Embedding.



https://brains.dev/2024/token-e-embedding-conceitos-da-ia-e-llms/


- (0.5, 8)
- (1, 8)
- (1, 7)
- (6, 7)
- (7, 7)
- (1, 1)
- (1, 2)
- (6.5, 1)
0
1
2
3
4
5
6
7
8



Se a sua resposta foi "**depende**", parabéns! Você acertou. E vai depender justamente do contexto. E para entender todo o contexto, é onde o Mecanismo de Atenção faz toda a diferença.

Mas, isso é assunto para nosso curso...

PRODUTO EXTERNO ENTRE VETORES

$\boldsymbol{x} \in \mathbb{R}^m, \boldsymbol{y} \in \mathbb{R}^n$ calcula-se $\boldsymbol{x}\boldsymbol{y}^T \in \mathbb{R}^{m \times n}$

$$\boldsymbol{A} = \boldsymbol{x}\boldsymbol{y}^T = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_m \end{bmatrix} \begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix} = \boxed{\begin{bmatrix} x_1 y_1 & x_1 y_2 & \dots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & \dots & x_2 y_n \\ \dots & \dots & \dots & \dots \\ x_m y_1 & x_m y_2 & \dots & x_m y_n \end{bmatrix}}$$

Matriz

MATRIZES

MATRIZ DE CARACTERÍSTICAS

Para nosso universo de IA, matriz é constituída de várias listas de atributos de diferentes objetos!

450 m^2
5 quartos
6 banheiro
850000 reais

$$\begin{bmatrix} 450 & 5 & 6 & 850 \end{bmatrix}$$

A linha contém valores para todas as variáveis para aquela observação específica, representando uma única observação ou ponto de dados.

MATRIZ DE CARACTERÍSTICAS

Para nosso universo de IA, matriz é constituída de várias listas de atributos de diferentes objetos!

200 m^2
3 quartos
4 banheiro
450000 reais

$$\begin{bmatrix} 450 & 5 & 6 & 850 \\ 200 & 3 & 4 & 450 \end{bmatrix}$$

Cada coluna contém valores de dados para uma variável ou característica. Pense nas colunas como representando os atributos ou características do conjunto de dados.

MATRIZES

Seja $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ uma matriz com $\boldsymbol{m}$ **linhas e** $\boldsymbol{n}$ **colunas** ($m \times n$ **é a dimensão da matriz)**

Cada linha é uma observação/dado/amostra

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

Cada coluna é uma característica/variável

a_{ij} é o elemento da *i-ésima* linha da *j-ésima* coluna de $\boldsymbol{A}$

a_{ij} é *j-ésima* característica da *i-ésima* observação

Neste exemplo cada uma das m observações tem n características.

A *i-ésima* coluna, $\boldsymbol{a}_i = a_{:,i}$

$$\boldsymbol{A} = \begin{bmatrix} | & | & & | \\ \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \\ | & | & & | \end{bmatrix}$$

A *j-ésima* linha, $\boldsymbol{a}'_j = a_{j,:}$

$$\boldsymbol{A} = \begin{bmatrix} - & \boldsymbol{a}'_1 & - \\ - & \boldsymbol{a}'_2 & - \\ - & \vdots & - \\ - & \boldsymbol{a}'_m & - \end{bmatrix}$$

PECE Programa de Educação Continuada
Escola Politécnica da USP

MATRIZES ESPECIAIS

Simétricas, ortogonais, identidade...

MATRIZ IDENTIDADE, $\boldsymbol{I} \in \mathbb{R}^{n \times n}$

É uma matriz quadrada de dimensões $n \times n$, definida como valor 1 nos termos da diagonal principal e 0 nos demais,

$$I_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Para $\boldsymbol{A} \in \mathbb{R}^{m \times n}$,

$$\boldsymbol{AI} = \boldsymbol{A} = \boldsymbol{IA}$$

MATRIZ DIAGONAL

Uma matriz diagonal é uma matriz quadrada em que todos os termos fora da diagonal são 0. Nos demais,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

$$I_{ij} = \begin{cases} d_i & i = j \\ 0 & i \neq j \end{cases}$$

Denomina-se, normalmente, $\boldsymbol{D} = \mathrm{diag}\,(d_1, d_2, \dots, d_n)$. A matriz identidade é um caso particular da matriz diagonal $\boldsymbol{I} = \mathrm{diag}\,(1,1, \dots, 1)$

MATRIZES ORTOGONAIS

Uma matriz quadrada $\boldsymbol{U} \in \mathbb{R}^{n \times n}$ é **ortogonal** se todas as suas colunas forem ortogonais entre si e normalizadas (as colunas são chamadas de ortonormais). Segue-se imediatamente da definição de ortogonalidade e normalidade que,

$$\boldsymbol{U}^T \boldsymbol{U} = \boldsymbol{I} = \boldsymbol{U}\boldsymbol{U}^T$$

Exemplo clássico é a matriz de rotação... $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

OPA, E O QUE É VETOR ORTONORMAL???

$$\boldsymbol{U} = \begin{pmatrix} \boxed{\begin{matrix}\cos\theta \\ \sin\theta\end{matrix}} & \boxed{\begin{matrix}-\sin\theta \\ \cos\theta\end{matrix}} \end{pmatrix} = \begin{pmatrix} | & | \\ \boldsymbol{u}_1 & \boldsymbol{u}_2 \\ | & | \end{pmatrix}$$

São normalizados?

$$\boldsymbol{u}_1 = \begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix} \quad \|\boldsymbol{u}_1\|^2 = \boldsymbol{u}_1 \cdot \boldsymbol{u}_1 = {u_1}_1^2 + {u_1}_2^2 = \cos^2\theta + \sin^2\theta = 1$$

$$\boldsymbol{u}_2 = \begin{pmatrix}-\sin\theta \\ \cos\theta\end{pmatrix} \quad \|\boldsymbol{u}_2\|^2 = \boldsymbol{u}_2 \cdot \boldsymbol{u}_2 = {u_2}_1^2 + {u_2}_2^2 = (-\sin\theta)^2 + \cos^2\theta = 1$$

OPA, E O QUE É VETOR ORTONORMAL???

$$\boldsymbol{U} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} | & | \\ \boldsymbol{u}_1 & \boldsymbol{u}_2 \\ | & | \end{pmatrix}$$

São ortogonais?

$$\boldsymbol{u}_1 = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$$

$$\boldsymbol{u}_1 \cdot \boldsymbol{u}_2 = u_{1_{\mathbf{1}}} u_{2_{\mathbf{1}}} + u_{1_{\mathbf{2}}} u_{2_{\mathbf{2}}}$$

$$\boldsymbol{u}_1 \cdot \boldsymbol{u}_2 = \cos\theta\,(-\sin\theta) + \sin\theta\cos\theta = 0$$

$$\boldsymbol{u}_2 = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$$

OPERAÇÕES

Adição de matrizes
Multiplicação por um escalar
Multiplicação entre:
- matriz e vetor
- entre matrizes

PECE Programa de Educação Continuada
Escola Politécnica da USP

ADIÇÃO

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

Adição elemento a elemento, somente matrizes com as mesmas dimensões podem ser adicionadas

MULTIPLICAÇÃO POR ESCALAR

$$\begin{bmatrix} \boxed{2} & 3 & 4 \\ 5 & 2 & 0 \end{bmatrix} \times \boxed{2} = \begin{bmatrix} \boxed{4} & 6 & 8 \\ 10 & 4 & 0 \end{bmatrix}$$

Multiplicação de cada elemento por um escalar

PRODUTO ENTRE MATRIZ E VETOR

$\boldsymbol{A} \in \mathbb{R}^{m \times n}$ e $\boldsymbol{v}, \boldsymbol{x} \in \mathbb{R}^{n \times 1}$

$$\boldsymbol{Ax} = \boldsymbol{v}$$

$$\underset{2 \times 2}{\begin{bmatrix} \boxed{1 \quad 2} \\ \boxed{3 \quad 4} \end{bmatrix}} \times \underset{2 \times 1}{\begin{bmatrix} \boxed{\begin{matrix} x_1 \\ x_2 \end{matrix}} \end{bmatrix}} = \underset{2 \times 1}{\begin{bmatrix} \boxed{1x_1 + 2x_2} \\ \boxed{3x_1 + 4x_2} \end{bmatrix}}$$

$$\boldsymbol{v} = \boldsymbol{Ax} = \begin{bmatrix} - & \boldsymbol{a}_1' & - \\ - & \boldsymbol{a}_2' & - \\ - & \vdots & - \\ - & \boldsymbol{a}_m' & - \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} \boldsymbol{a}_1'\boldsymbol{x} \\ \boldsymbol{a}_2'\boldsymbol{x} \\ \vdots \\ \boldsymbol{a}_m'\boldsymbol{x} \end{bmatrix}$$

a *i-ésima* entrada de $\boldsymbol{v}$ é igual ao produto da *i-ésima* linha de $\boldsymbol{A}$ por $\boldsymbol{x}$, $v_i = \boldsymbol{a}_i'\boldsymbol{x}$

MULTIPLICAÇÃO ENTRE MATRIZ E VETOR

$$\begin{bmatrix} \boxed{\begin{matrix}1\\3\end{matrix}} & \boxed{\begin{matrix}2\\4\end{matrix}} \end{bmatrix} \times \begin{bmatrix} \boxed{x_1} \\ \boxed{x_2} \end{bmatrix} = \begin{bmatrix} \boxed{\begin{matrix}1x_1\\3x_1\end{matrix}} & \begin{matrix}+\\+\end{matrix} & \boxed{\begin{matrix}2x_2\\4x_2\end{matrix}} \end{bmatrix}$$

$$\boldsymbol{v} = \boldsymbol{A}\boldsymbol{x} = \begin{bmatrix} \boxed{\begin{matrix}|\\\boldsymbol{a}_1\\|\end{matrix}} & \boxed{\begin{matrix}|\\\boldsymbol{a}_2\\|\end{matrix}} & \cdots & \boxed{\begin{matrix}|\\\boldsymbol{a}_n\\|\end{matrix}} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \boxed{\begin{bmatrix} \boldsymbol{a}_1 \end{bmatrix} x_1} + \boxed{\begin{bmatrix} \boldsymbol{a}_2 \end{bmatrix} x_2} + \cdots + \boxed{\begin{bmatrix} \boldsymbol{a}_n \end{bmatrix} x_n}$$

Características/variáveis

$\boldsymbol{v}$ é uma combinação linear das características, onde os coeficientes da combinação linear são dados pelas entradas de $\boldsymbol{x}$.

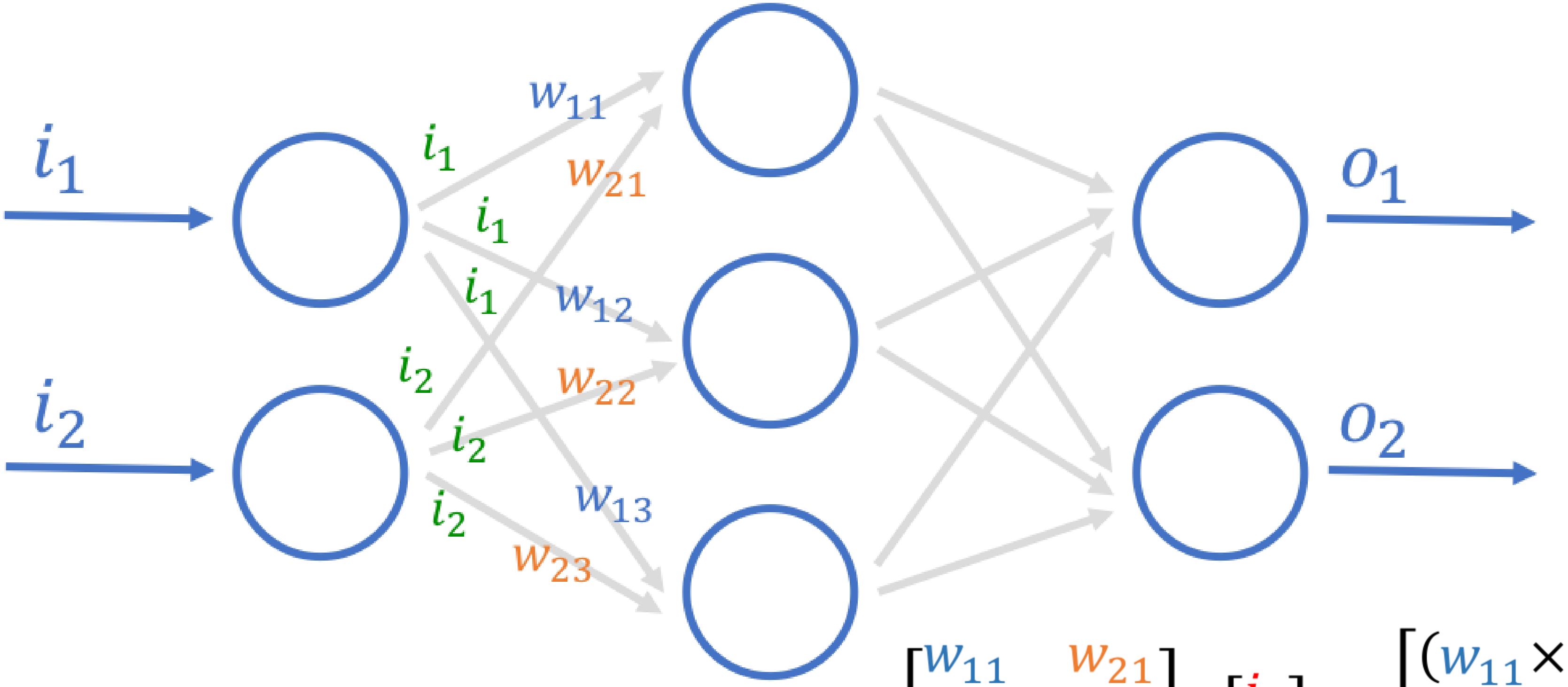
PECE Programa de Educação Continuada
Escola Politécnica da USP

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

$$= \begin{bmatrix} 23 \\ 12 \end{bmatrix}$$

PECE Programa de Educação Continuada
Escola Politécnica da USP

PORQUE ESSE PRODUTO É IMPORTANTE...



$$\begin{bmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \\ w_{13} & w_{23} \end{bmatrix} \cdot \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} (w_{11} \times i_1) + (w_{21} \times i_2) \\ (w_{12} \times i_1) + (w_{22} \times i_2) \\ (w_{13} \times i_1) + (w_{23} \times i_2) \end{bmatrix}$$

MULTIPLICAÇÃO ENTRE MATRIZES

$\boldsymbol{A} \in \mathbb{R}^{m \times n}, \boldsymbol{B} \in \mathbb{R}^{n \times p} \Longrightarrow \boldsymbol{C} \in \mathbb{R}^{m \times p}$

$$\boldsymbol{C} = \boldsymbol{AB} = \underset{m \times n}{\begin{bmatrix} - & \boldsymbol{a}_1' & - \\ - & \boldsymbol{a}_2' & - \\ - & \vdots & - \\ - & \boldsymbol{a}_m' & - \end{bmatrix}} \underset{n \times p}{\begin{bmatrix} | & | & & | \\ \boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_p \\ | & | & & | \end{bmatrix}} = \underset{m \times p}{\begin{bmatrix} \boldsymbol{a}_1'\boldsymbol{b}_1 & \boldsymbol{a}_1'\boldsymbol{b}_2 & \cdots & \boldsymbol{a}_1'\boldsymbol{b}_p \\ \boldsymbol{a}_2'\boldsymbol{b}_1 & \boldsymbol{a}_2'\boldsymbol{b}_2 & \cdots & \boldsymbol{a}_2'\boldsymbol{b}_p \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{a}_m'\boldsymbol{b}_1 & \boldsymbol{a}_m'\boldsymbol{b}_2 & \cdots & \boldsymbol{a}_m'\boldsymbol{b}_p \end{bmatrix}}$$

$$\begin{bmatrix} 2 & 3 & 4 \\ 5 & 2 & 0 \end{bmatrix}_{2x3} \times \begin{bmatrix} 2 & 1 \\ 2 & 3 \\ 0 & 1 \end{bmatrix}_{3x2} = \begin{bmatrix} \boxed{10} & 15 \\ 14 & 11 \end{bmatrix}_{2x2}$$

Resultando em uma matriz com dimensões iguais ao número de linhas da primeira matriz e ao número de colunas da segunda matriz

EXEMPLO

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{bmatrix} \quad \boldsymbol{B} = \begin{bmatrix} 2 & 7 \\ 1 & 2 \\ 3 & 6 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{bmatrix} \begin{bmatrix} 2 & 7 \\ 1 & 2 \\ 3 & 6 \end{bmatrix} = \quad ?$$

VÁRIAS VISÕES DE $\boldsymbol{C} = \boldsymbol{AB}$

$$\boldsymbol{C} = \boldsymbol{AB} = \boldsymbol{A}\begin{bmatrix} | & | & & | \\ \boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_p \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \boldsymbol{Ab}_1 & \boxed{\boldsymbol{Ab}_2} & \cdots & \boldsymbol{Ab}_p \\ | & | & & | \end{bmatrix}$$

$$\boldsymbol{c}_i = \boldsymbol{Ab}_i$$

Colunas de $\boldsymbol{C}$ vistas como produtos entre a matriz $\boldsymbol{A}$ e os vetores formados pelas colunas de $\boldsymbol{B}$.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \times \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1 \times 5 + 2 \times 7 & 1 \times 6 + 2 \times 8 \\ 3 \times 5 + 4 \times 7 & 3 \times 6 + 4 \times 8 \end{bmatrix} = \begin{bmatrix} 5 + 14 & 6 + 16 \\ 15 + 28 & 18 + 32 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$$

Cada coluna de $\boldsymbol{C}$ é uma combinação linear diferente das características, onde os coeficientes de cada combinação linear são dados pelas colunas de $\boldsymbol{B}$.

TREINE O RACIOCÍNIO...

Suponha que você tenha um conjunto de dados com as notas de 4 alunos em 3 áreas do conhecimento (Matemática, Ciências e Humanidades). Cada universidade tem uma matriz de pesos diferente para calcular a nota final de ingresso. A escola deseja saber em quais universidades seus alunos têm mais chances de sucesso.

a) Determine em qual universidade cada aluno tem a maior nota final.
b) Interprete os resultados para identificar onde os alunos têm mais chances de sucesso.

Aluno	Matemática	Ciências	Humanidades
01	8,0	7,5	9,0
02	6,0	8,0	7,0
03	9,0	9,5	8,5
04	7,5	6,5	8,0

Universidade	Peso Matemática	Peso Ciências	Peso Humanidades
A	0,5	0,3	0,2
B	0,3	0,4	0,3
C	0,2	0,5	0,3

VÁRIAS VISÕES DE $\boldsymbol{C} = \boldsymbol{AB}$

PECE Programa de Educação Continuada
Escola Politécnica da USP

$$\boldsymbol{C} = \boldsymbol{AB} = \begin{bmatrix} \boxed{\begin{matrix} | \\ \boldsymbol{a}_1 \\ | \end{matrix}} & \begin{matrix} | \\ \boldsymbol{a}_2 \\ | \end{matrix} & \cdots & \begin{matrix} | \\ \boldsymbol{a}_n \\ | \end{matrix} \end{bmatrix} \begin{bmatrix} \boxed{- \quad \boldsymbol{b}_1' \quad -} \\ - \quad \boldsymbol{b}_2' \quad - \\ \vdots \\ - \quad \boldsymbol{b}_n' \quad - \end{bmatrix} = \sum_{i=1}^{n} \boldsymbol{a}_i \boldsymbol{b}_i'$$

Multiplicação de matrizes vista como **produto externo**

$$\begin{bmatrix} \boxed{\begin{matrix} 1 \\ 3 \end{matrix}} & \begin{matrix} 2 \\ 4 \end{matrix} \end{bmatrix} \times \begin{bmatrix} \boxed{5 \quad 6} \\ 7 \quad 8 \end{bmatrix} = \begin{bmatrix} \boxed{1 \times 5} & \boxed{1 \times 6} \\ \boxed{3 \times 5} & \boxed{3 \times 6} \end{bmatrix} = \begin{bmatrix} \boxed{5} & \boxed{6} \\ \boxed{15} & \boxed{18} \end{bmatrix}$$

VÁRIAS VISÕES DE $\boldsymbol{C = AB}$

$$\boldsymbol{C} = \boldsymbol{AB} = \begin{bmatrix} | & | & & | \\ \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} - & \boldsymbol{b}_1' & - \\ - & \boldsymbol{b}_2' & - \\ - & \vdots & - \\ - & \boldsymbol{b}_n' & - \end{bmatrix} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{a}_i \boldsymbol{b}_i'$$

Multiplicação de matrizes vista como **produto externo**

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \times \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1 \times 5 + \boxed{2 \times 7} & 1 \times 6 + \boxed{2 \times 8} \\ 3 \times 5 + \boxed{4 \times 7} & 3 \times 6 + \boxed{4 \times 8} \end{bmatrix} = \begin{bmatrix} 5 + \boxed{14} & 6 + \boxed{16} \\ 15 + \boxed{28} & 18 + \boxed{32} \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$$

PECE Programa de Educação Continuada
Escola Politécnica da USP

$$\boldsymbol{C} = \boldsymbol{AB} = \begin{bmatrix} - & \boldsymbol{a}_1' & - \\ - & \boldsymbol{a}_2' & - \\ - & \vdots & - \\ - & \boldsymbol{a}_m' & - \end{bmatrix} \boldsymbol{B} = \begin{bmatrix} - & \boldsymbol{a}_1'\boldsymbol{B} & - \\ - & \boldsymbol{a}_2'\boldsymbol{B} & - \\ - & \vdots & - \\ - & \boldsymbol{a}_m'\boldsymbol{B} & - \end{bmatrix}, \qquad \boldsymbol{c}_i' = \boldsymbol{a}_i'\boldsymbol{B}$$

Linhas de $\boldsymbol{C}$ vistas como produtos entre os vetores formados pelas linhas de $\boldsymbol{A}$ e a matriz $\boldsymbol{B}$.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \times \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1 \times 5 + 2 \times 7 & 1 \times 6 + 2 \times 8 \\ 3 \times 5 + 4 \times 7 & 3 \times 6 + 4 \times 8 \end{bmatrix} = \begin{bmatrix} 5 + 14 & 6 + 16 \\ 15 + 28 & 18 + 32 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$$

MATRIZ TRANSPOSTA

A transposição de uma matriz resulta na troca das linhas pelas colunas. Dada uma matriz $\boldsymbol{A} \in \mathbb{R}^{mxn}$, sua transposição, escrita $\boldsymbol{A}^T \in \mathbb{R}^{nxm}$, é a matriz $n \times m$ cujas entradas são dadas por

$$(\boldsymbol{A}^T)_{ij} = a_{ji}$$

$$\begin{array}{cc} & n \\ m & \begin{bmatrix}1 & 2\\ 3 & 4\\ 5 & 6\end{bmatrix}^T \end{array} = \begin{array}{cc} & m \\ n & \begin{bmatrix}1 & 3 & 5\\ 2 & 4 & 6\end{bmatrix} \end{array}$$

Quando $\boldsymbol{A}^T = \boldsymbol{A}$ a matriz é dita **simétrica** (somente matrizes quadradas, $m = n$, podem ser simétricas).

MATRIZ SIMÉTRICA E ANTI-SIMÉTRICA, $A \in \mathbb{R}^{m\times m}$

Simétrica

$$\boldsymbol{A} = \boldsymbol{A}^T$$

Anti-simétrica

$$\boldsymbol{A} = -\boldsymbol{A}^T$$

Toda matriz pode ser escrita como uma soma de uma parcela simétrica e outra anti-simétrica,

$$\boldsymbol{A} = \frac{1}{2}(\boldsymbol{A} + \boldsymbol{A}^T) + \frac{1}{2}(\boldsymbol{A} - \boldsymbol{A}^T)$$

SÓ PARA CONSTAR...

$$\boldsymbol{A} = \begin{bmatrix} 3 & 3 & -1 \\ -2 & -2 & 1 \\ -4 & -5 & 2 \end{bmatrix} \quad \boldsymbol{A}^T = \begin{bmatrix} 3 & -2 & -4 \\ 3 & -2 & -5 \\ -1 & 1 & 2 \end{bmatrix}$$

$$\boldsymbol{D} = \boldsymbol{A} + \boldsymbol{A}^T = \begin{bmatrix} 3+3 & 3-2 & -1-4 \\ -2+3 & -2-2 & 1-5 \\ -4-1 & -5+1 & 2+2 \end{bmatrix} = \begin{bmatrix} 6 & 1 & -5 \\ 1 & -4 & -4 \\ -5 & -4 & 4 \end{bmatrix}$$ Simétrica... $(\boldsymbol{D} = \boldsymbol{D}^T)$

$$\boldsymbol{S} = \boldsymbol{A} - \boldsymbol{A}^T = \begin{bmatrix} 3-3 & 3+2 & -1+4 \\ -2-3 & -2+2 & 1+5 \\ -4+1 & -5-1 & 2-2 \end{bmatrix} = \begin{bmatrix} 0 & 5 & 3 \\ -5 & 0 & 6 \\ -3 & -6 & 0 \end{bmatrix}$$ Anti-Simétrica... $(\boldsymbol{S} = -\boldsymbol{S}^T)$

$$\frac{1}{2}(\boldsymbol{A}+\boldsymbol{A}^T) = \frac{1}{2}\begin{bmatrix} 6 & 1 & -5 \\ 1 & -4 & -4 \\ -5 & -4 & 4 \end{bmatrix}$$

$$\frac{1}{2}(\boldsymbol{A}-\boldsymbol{A}^T) = \frac{1}{2}\begin{bmatrix} 0 & 5 & 3 \\ -5 & 0 & 6 \\ -3 & -6 & 0 \end{bmatrix}$$

$$\frac{1}{2}\begin{bmatrix} 6 & 1 & -5 \\ 1 & -4 & -4 \\ -5 & -4 & 4 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 0 & 5 & 3 \\ -5 & 0 & 6 \\ -3 & -6 & 0 \end{bmatrix} = \begin{bmatrix} \frac{6}{2}+0 & \frac{1}{2}+\frac{5}{2} & \frac{-5}{2}+\frac{3}{2} \\ \frac{1}{2}+\frac{-5}{2} & \frac{-4}{2} & \frac{-4}{2}+\frac{6}{2} \\ \frac{-5}{2}+\frac{-3}{2} & \frac{-4}{2}+\frac{-6}{2} & \frac{4}{2} \end{bmatrix} = \underbrace{\boxed{\begin{bmatrix} 3 & 3 & -1 \\ -2 & -2 & 1 \\ -4 & -5 & 2 \end{bmatrix}}}_{\boldsymbol{A}}$$

PECE Programa de Educação Continuada
Escola Politécnica da USP

TRAÇO DE UMA MATRIZ

O traço de uma matriz é a soma de sua diagonal principal,

$$\operatorname{tr} \boldsymbol{A} = \sum_{i=1}^{n} a_{ii}$$

$\boldsymbol{A} \in \mathbb{R}^{\boldsymbol{n}\times\boldsymbol{n}}$, $\quad \operatorname{tr} \boldsymbol{A} = \operatorname{tr} \boldsymbol{A}^T$

$\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{\boldsymbol{n}\times\boldsymbol{n}}$, $\quad \operatorname{tr}(\boldsymbol{A} + \boldsymbol{B}) = \operatorname{tr} \boldsymbol{A} + \operatorname{tr} \boldsymbol{B}$

$\boldsymbol{A} \in \mathbb{R}^{\boldsymbol{n}\times\boldsymbol{n}}, k \in \mathbb{R} \quad \operatorname{tr}(k\boldsymbol{A}) = k \operatorname{tr} \boldsymbol{A}$

$\boldsymbol{A}, \boldsymbol{B}$ tal que $\boldsymbol{AB}$ é quadrada $\operatorname{tr}(\boldsymbol{AB}) = \operatorname{tr}(\boldsymbol{BA})$

$\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ tal que $\boldsymbol{ABC}$ é quadrada $\operatorname{tr}(\boldsymbol{ABC}) = \operatorname{tr}(\boldsymbol{BCA}) = \operatorname{tr}(\boldsymbol{CAB})$

$$\begin{bmatrix} 0 & 2 & 3 \\ 1 & 6 & 4 \\ 2 & 6 & 1 \end{bmatrix}$$

Traço

$$0 + 6 + 1 = 7$$

PROPRIEDADES DA MULTIPLICAÇÃO DE MATRIZES

Multiplicação de matrizes é distributiva e associativa,

$$(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$$
$$\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{AC}$$

mas somente em casos especiais é comutativa. Em geral,

$$\boldsymbol{AB} \neq \boldsymbol{BA}$$

IMPORTANTE... $ABC \neq CAB \neq BCA$

$$\boldsymbol{A} = \begin{bmatrix} 4 & 12 \\ 7 & 6 \end{bmatrix}$$

$$\boldsymbol{B} = \begin{bmatrix} 1 & -3 \\ 4 & 3 \end{bmatrix}$$

$$\boldsymbol{C} = \begin{bmatrix} 6 & 6 \\ 2 & 5 \end{bmatrix}$$

$$\boldsymbol{ABC} = \begin{bmatrix} 360 & 432 \\ 180 & 171 \end{bmatrix}$$

$$\boldsymbol{CAB} = \begin{bmatrix} 498 & 126 \\ 259 & 33 \end{bmatrix}$$

$$\boldsymbol{BCA} = \begin{bmatrix} -63 & -54 \\ 393 & 594 \end{bmatrix}$$

Lembram-se da propriedade do traço?

$$\mathrm{Tr}(\boldsymbol{ABC}) = \mathrm{Tr}(\boldsymbol{CAB}) = \mathrm{Tr}(\boldsymbol{BCA}) = 531$$

MATRIZ IDENTIDADE E MATRIZ INVERSA

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

$\boldsymbol{A}^{-1}$ é a matriz inversa de $\boldsymbol{A}$

$$\begin{bmatrix} -3 & 1 \\ 5 & 0 \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{5} \\ 1 & \frac{3}{5} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

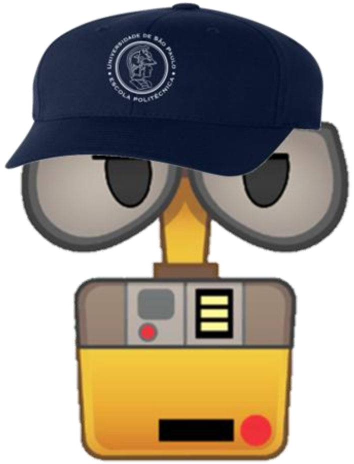
$\boldsymbol{A}$

$\boldsymbol{A}^{-1}$

OPERAÇÕES BÁSICAS

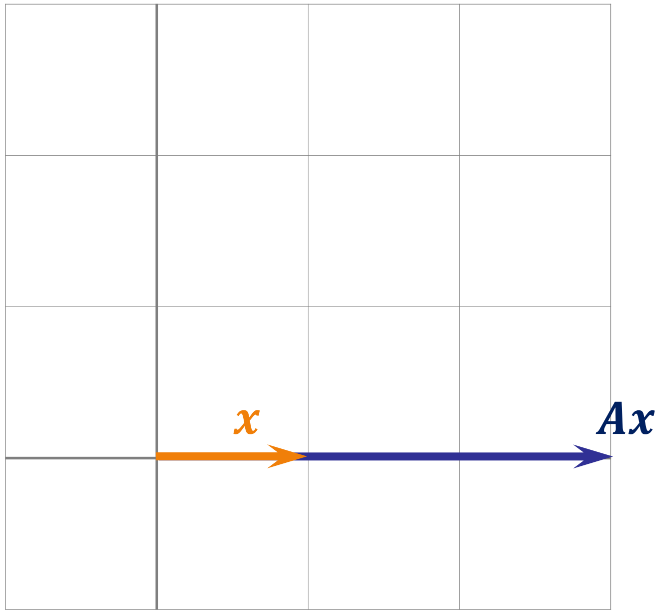
LEMBRAM-SE DA MATRIZ $\boldsymbol{I}$?

Primeiro, vamos pensar em uma matriz que não muda nada.

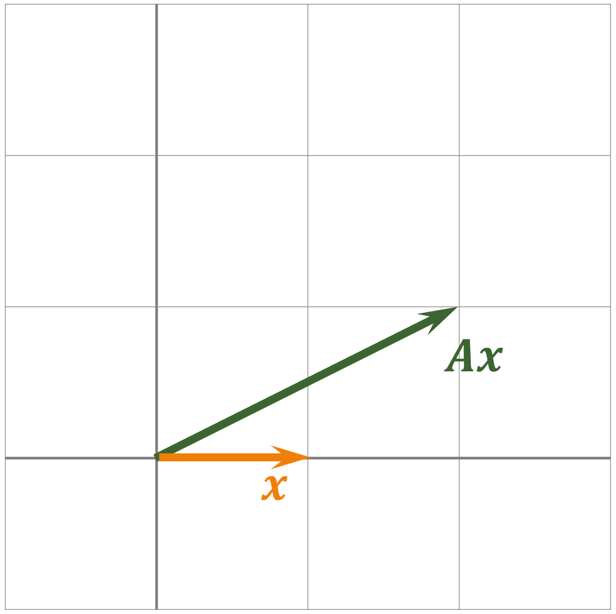


$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \times 1 \\ 1 \times 2 \\ 1 \times 3 \\ 1 \times 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

SE $\boldsymbol{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, QUAL RESULTADO $\boldsymbol{Ax}$?

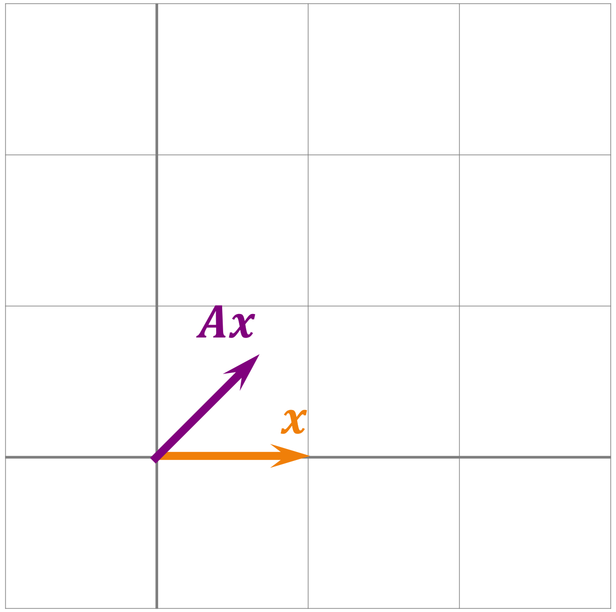


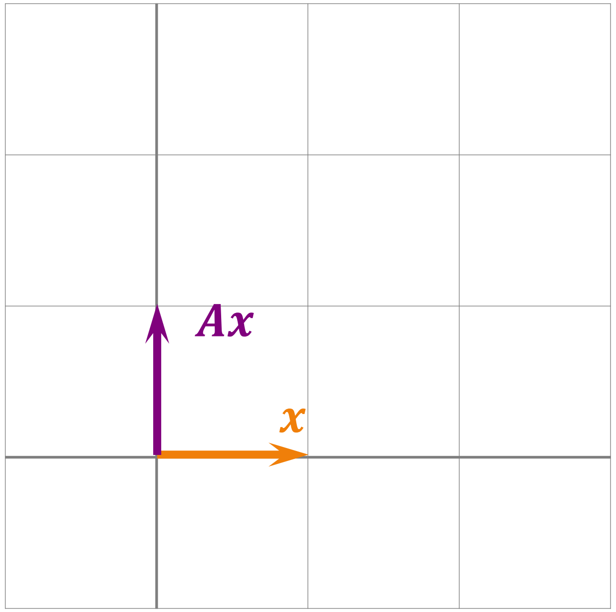
$$A = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$$



$$A = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}$$

SE $\boldsymbol{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, QUAL RESULTADO $\boldsymbol{Ax}$?





$$A = \begin{pmatrix} \cos 45° & -\sin 45° \\ \sin 45° & \cos 45° \end{pmatrix} = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}$$

$$A = \begin{pmatrix} \cos 90° & -\sin 90° \\ \sin 90° & \cos 90° \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

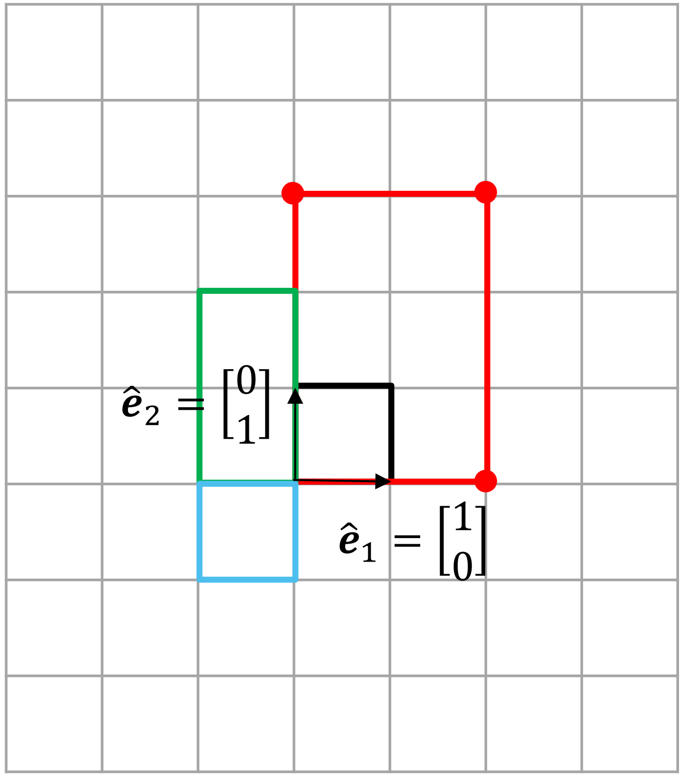
MODIFICANDO A IDENTIDADE...

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ Escalona em torno do ponto $(0,0)$

$\begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}$ Escalona em torno do ponto $(0,0)$ na direção y e gira em torno de y

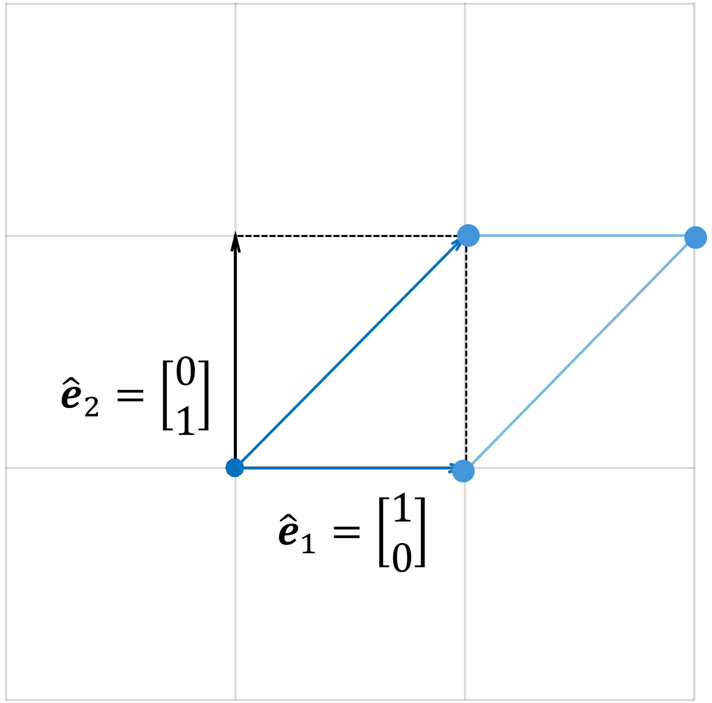
$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ Gira em torno do ponto $(0,0)$, ie, em torno de x e y.



$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$



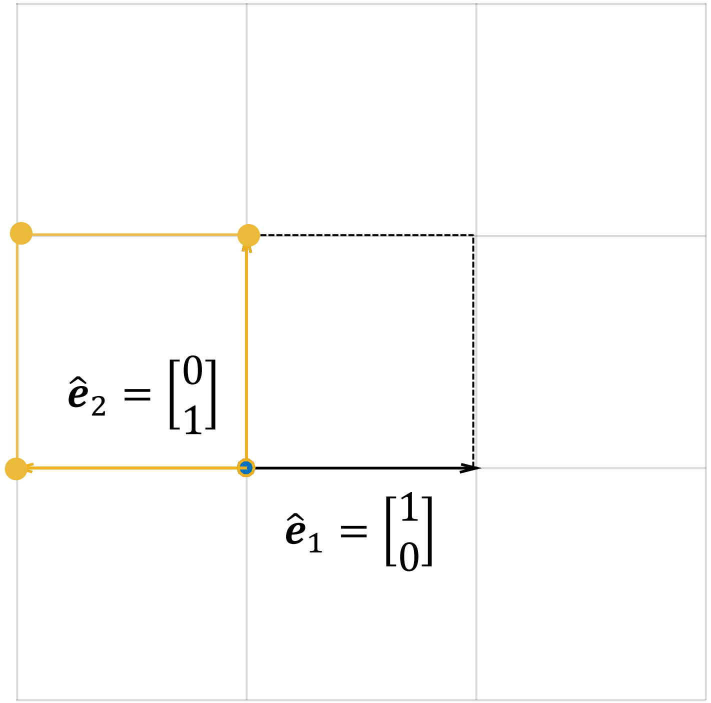
Distorção

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

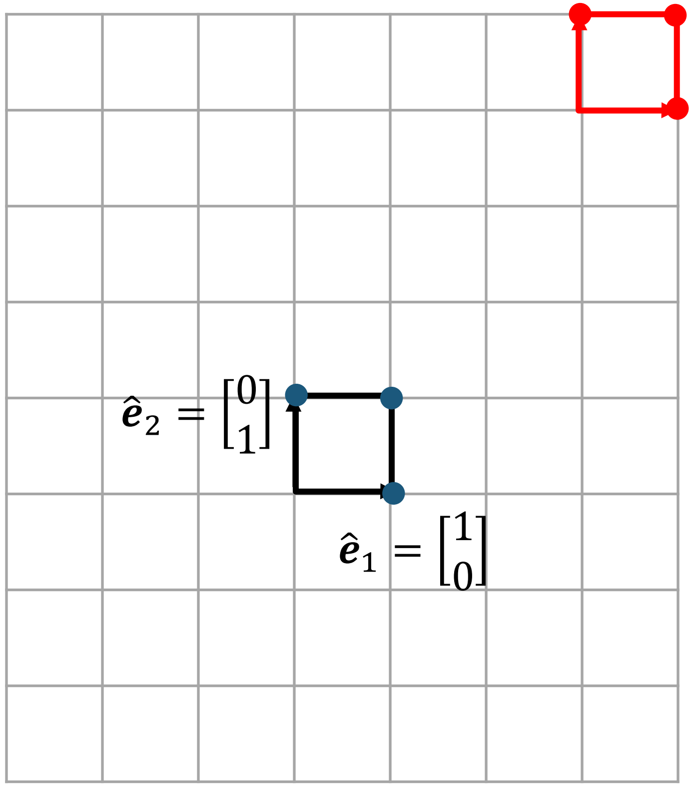


Rotação

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Rotação: $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

E SE EU QUISER **TRANSLADAR** ???



Ao multiplicar um vetor 2D por uma matriz 2x2 qualquer, você não consegue transladá-lo.

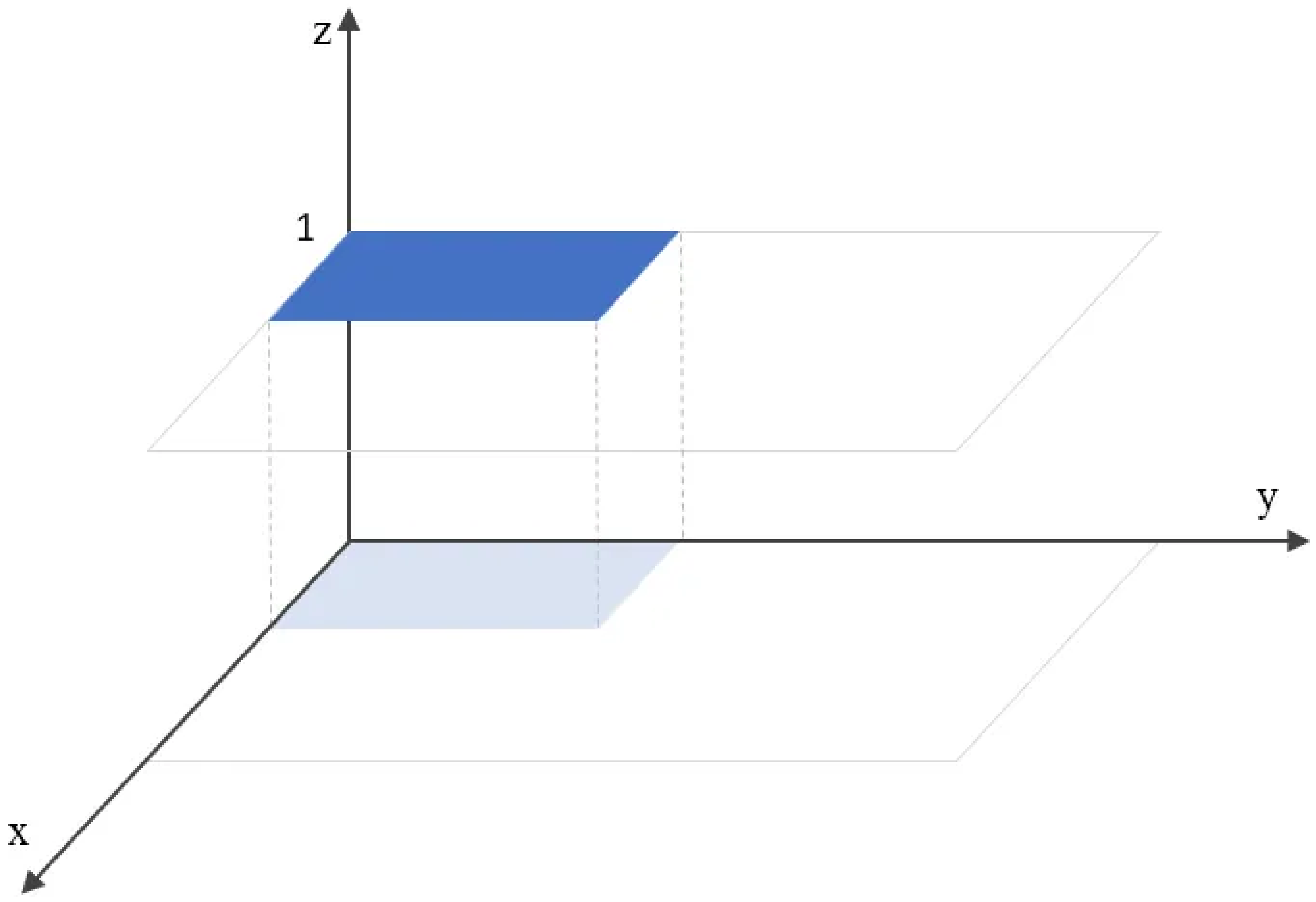
Na transformação linear: escalonamento, distorção e rotação, os vetores de base permanecem todos na mesma origem (0,0) antes e depois da transformação.

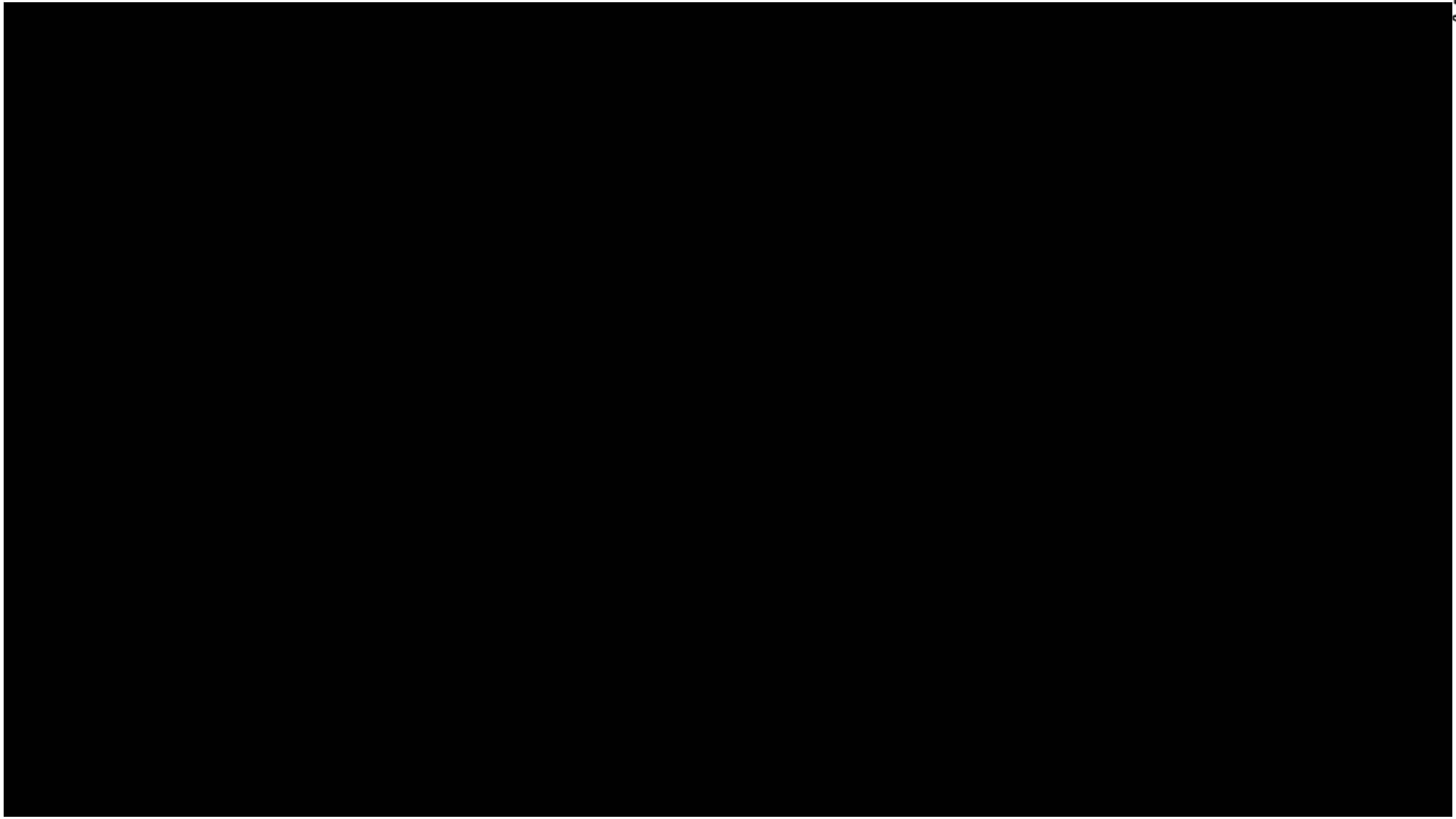
Isso significa que o ponto (0,0) nunca muda de localização. Para transladar, você precisa adicionar um vetor após a transformação da matriz...

$$\boxed{\begin{bmatrix} q_1 \\ q_2 \end{bmatrix}} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \boxed{\begin{bmatrix} p_1 \\ p_2 \end{bmatrix}} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

nova localização após a transformação

localização original


z
1
y
x



TRANSFORMAÇÃO GEOMÉTRICA DA IMAGEM

Envie cada pixel p da posição original (x, y) para a correspondente posição $(x', y') = T(x, y)$ na imagem transformada

Escalonamento, distorção e rotação

Translação

$$\begin{bmatrix} q_1 \\ q_2 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ 1 \end{bmatrix}$$

Fixando a imagem no plano $z = 1$

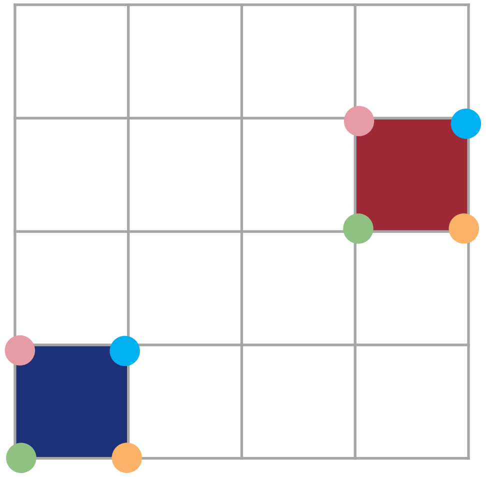
PORTANTO...

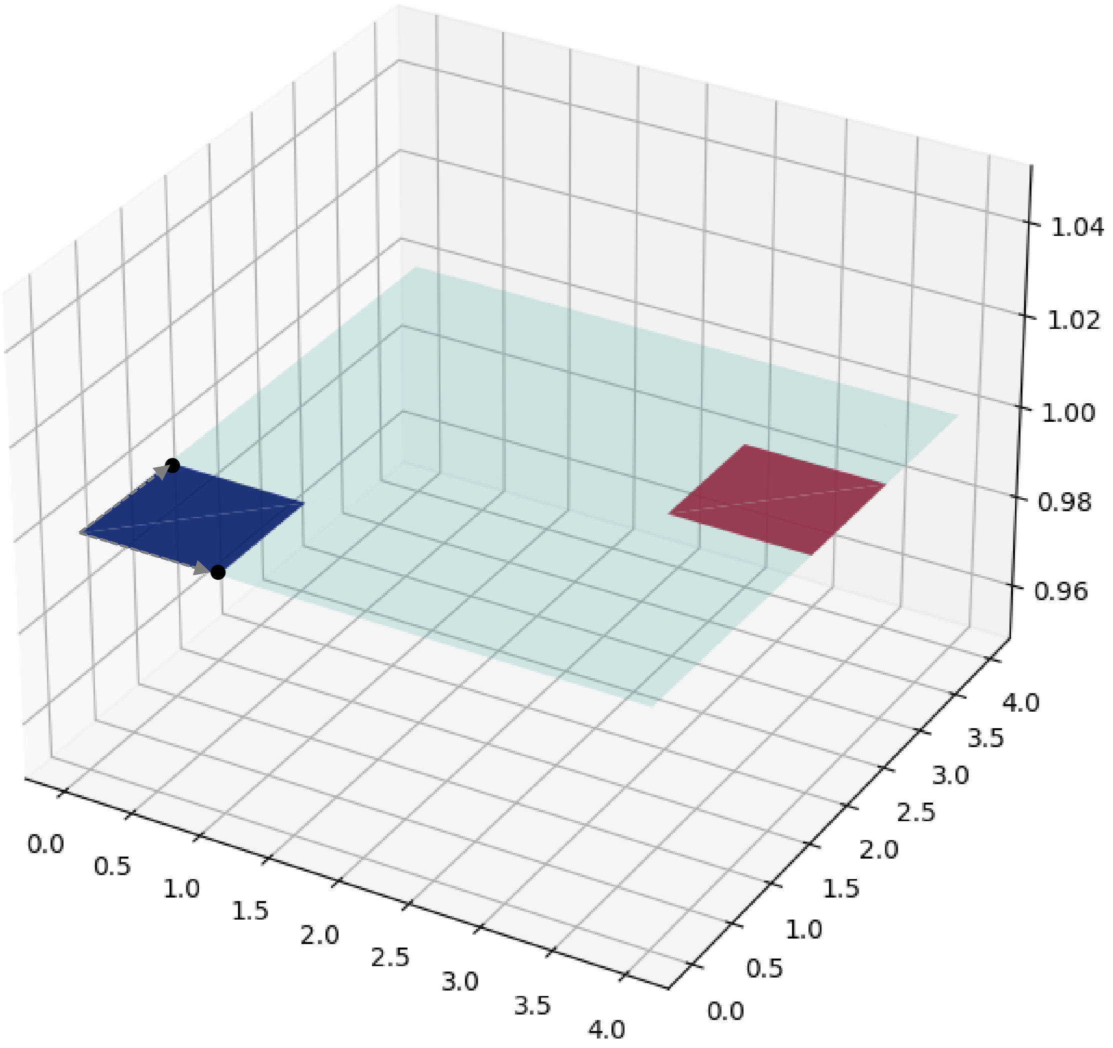
$$\begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix}$$

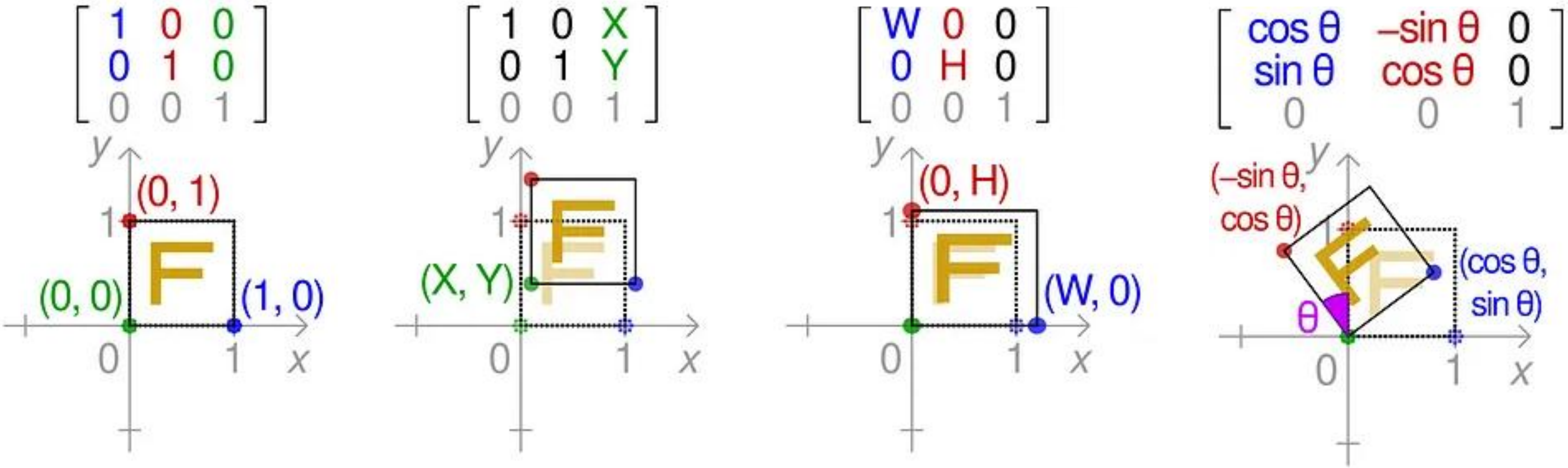
$$\begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 1 \end{pmatrix}$$





TRANSLAÇÃO, ESCALONAMENTO E ROTAÇÃO

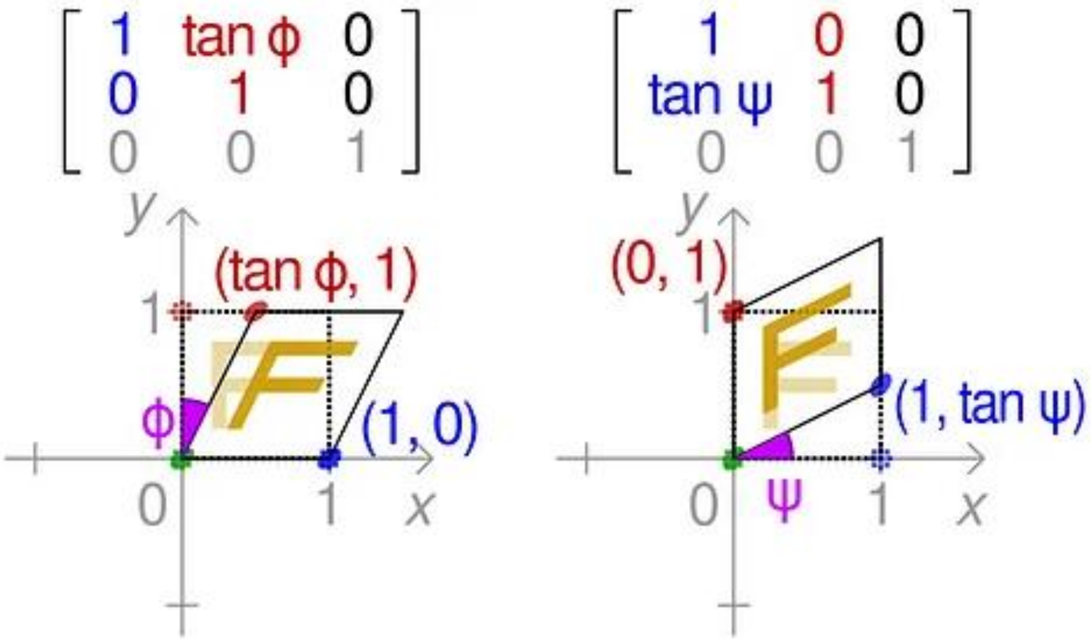
ORIGINAL | TRANSLAÇÃO | ESCALONAMENTO | ROTAÇÃO EM TORNO DA ORIGEM



DISTORÇÃO

DISTORÇÃO EM TORNO EIXO X

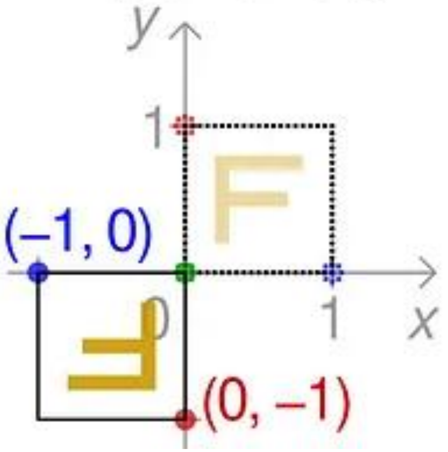
DISTORÇÃO EM TORNO EIXO Y



REFLEXÃO

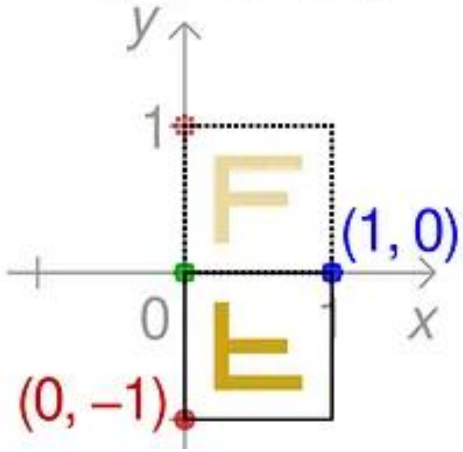
REFLEXÃO EM TORNO DA ORIGEM

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



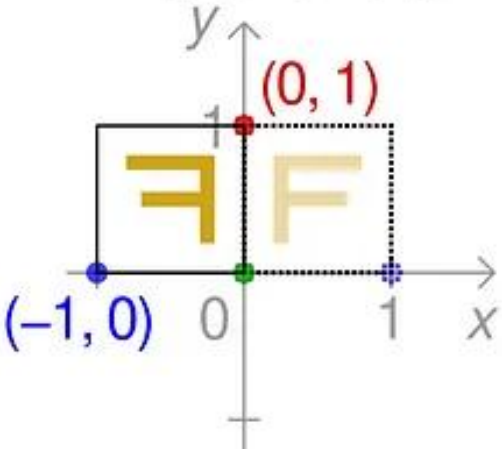
EM TORNO EIXO X

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



EM TORNO EIXO Y

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



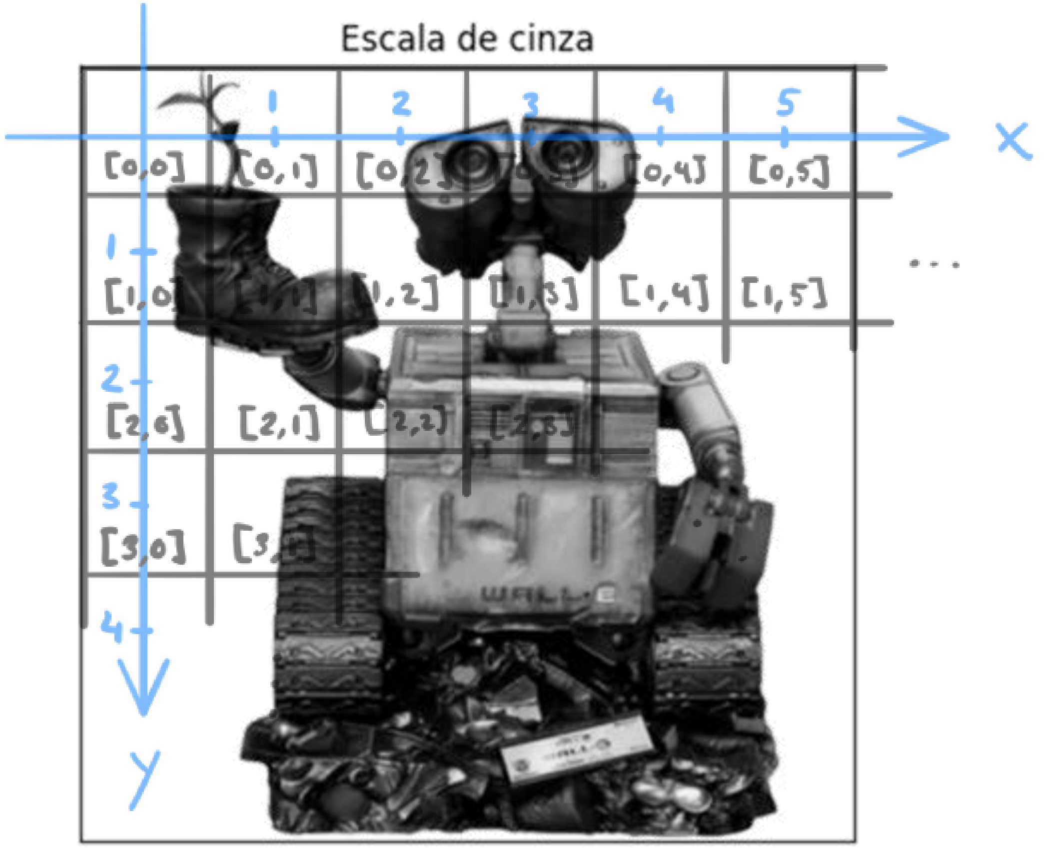
POR EXEMPLO...

Sabendo que a figura do Wall-E ao lado sofreu seguinte transformação,

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \boldsymbol{T} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$ onde $\boldsymbol{T}$ é composta das matrizes de escalonamento e rotação, respectivamente,

$$\boldsymbol{S} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \boldsymbol{R} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Defina a matriz $\boldsymbol{T}$ para as seguintes transformações,

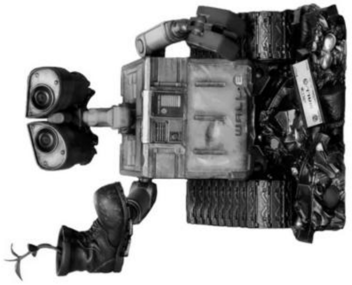


POR EXEMPLO...

$$\boldsymbol{S} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \boldsymbol{R} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Defina a matriz $\boldsymbol{T}$ para as seguintes transformações,

(✓) $\boldsymbol{T} = \boldsymbol{S}$
( ) $\boldsymbol{T} = \boldsymbol{R}$



( ) $\boldsymbol{T} = \boldsymbol{S}$
(✓) $\boldsymbol{T} = \boldsymbol{R}$

(✓) $\boldsymbol{T} = \boldsymbol{SR}$
( ) $\boldsymbol{T} = \boldsymbol{RS}$

( ) $\boldsymbol{T} = \boldsymbol{RS}$
(✓) $\boldsymbol{T} = \boldsymbol{RS}$

Stop waiting to *feel* ready.

Ready is not a *feeling*.

It's a ***decision***.

ACABOU... REVEJAM O MATERIAL DISPONIBILIZADO EM AULA. REFAÇAM EXERCÍCIOS

Bom carnaval!