Deep Learning Book

Capítulo 2 - Álgebra Linear

Davi Duarte de Paula
davi_duarte@outlook.com

UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"

INSTITUTO DE GEOCIÊNCIAS E CIÊNCIAS EXATAS
CÂMPUS DE RIO CLARO

Escalares, Vetores, Matrizes e Tensores

Escalar: Um único número. Representado aqui em itálico e em caixa baixa. Por exemplo: seja $s \in \mathbb{R}$ e $n \in \mathbb{N}$.

Escalares, Vetores, Matrizes e Tensores

Vetores: Uma sequência de números em ordem e indexáveis. Representado por uma letra em caixa baixa e negrito. Os valores que pertencem ao vetor precisam ser definidos, por exemplo: $\mathbf{x} \in \mathbb{R}^n$, significando que o vetor **x** possui n elementos pertencentes ao conjunto $\mathbb{R}$.

Escalares, Vetores, Matrizes e Tensores

Vetores:

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

Escalares, Vetores, Matrizes e Tensores

Vetores: Os elementos individuais internos são representados com letras em caixa baixa, sem negrito, indexados com um número subscrito.

Um vetor pode ser representado como um ponto no espaço, sendo que cada elemento representa o valor em uma determinada dimensão.

Escalares, Vetores, Matrizes e Tensores

Vetores: É possível também acessar um subconjunto do vetor, definindo um conjunto específico. Por exemplo, para acessar os elementos x_1,x_3 e x_6, podemos definir:

$$S = \{1,2,3\}, \text{ e acessar o vetor como } \mathbf{x}_S$$

Para acessar todos os elementos do vetor, exceto os presente em S, denota-se como $\mathbf{x}_{-s}$.

Escalares, Vetores, Matrizes e Tensores

Matrizes: É uma sequência de valores 2D. Portanto, cada elemento é identificado com 2 índices. As matrizes aqui serão representadas com letras capitais em negrito, como $\boldsymbol{A}$**.**

Considerando que a matriz $\boldsymbol{A}$ seja composta por números reais e tenha *m* linhas e *n* colunas, ela é denotada como: $\boldsymbol{A} \in \mathbb{R}^{m \times n}$.

Os elementos de uma matriz aqui serão denotados com letras capitais em itálico (sem negrito). O elemento mais ao topo e a esquerda é o $A_{1,1}$, enquanto que o elemento mais à direita e em baixo é o $A_{m,n}$.

Escalares, Vetores, Matrizes e Tensores

Matrizes: Uma única linha ou coluna de uma matriz é denotada como $A_{i,:}$ ou $A_{:,i}$, respectivamente, sendo que *i* é a posição da linha que se deseja obter.

A matriz $\boldsymbol{A}$ pode ser denotada explicitamente da seguinte maneira,

$$\boldsymbol{A} = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix}.$$

Escalares, Vetores, Matrizes e Tensores

Tensores: É uma sequência de valores em 3 ou mais dimensões. Supondo que um tensor **A** (repare na mudança da fonte) possui 3 dimensões, um elemento na coordenada (*i*, *j*, *k*) é definido como $A_{i,j,k}$.

Escalares, Vetores, Matrizes e Tensores

Transposição e matrizes:

$$\boldsymbol{A} = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \\ A_{3,1} & A_{3,2} \end{bmatrix} \Rightarrow \boldsymbol{A}^\top = \begin{bmatrix} A_{1,1} & A_{2,1} & A_{3,1} \\ A_{1,2} & A_{2,2} & A_{3,2} \end{bmatrix}$$

Escalares, Vetores, Matrizes e Tensores

Transposição e matrizes:

Um vetor pode ser pensado com uma matriz com apenas 1 coluna. Portanto

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [x_1, x_2, ..., x_n]^\top.$$

Escalares, Vetores, Matrizes e Tensores

Transposição e matrizes:

Um escalar pode ser considerado como uma matriz com somente 1 elemento, portanto

$$a = a^\top$$

Escalares, Vetores, Matrizes e Tensores

Adição e multiplicação de matrizes e vetores:

Duas matrizes com as mesmas dimensões podem ser somadas efetuando a soma de seus respectivos elementos ponto a ponto.

$$\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}\text{, onde } C_{i,j} = A_{i,j} + B_{i,j}.$$

Escalares, Vetores, Matrizes e Tensores

Adição e multiplicação de matrizes e vetores:

Um escalar e uma matriz podem ser somados e/ou multiplicados, efetuando a operação com o escalar e cada um dos elementos da matriz.

$$\boldsymbol{D} = a \cdot \boldsymbol{B} + c, \text{ onde } D_{i,j} = a \cdot B_{i,j} + c,$$

sendo a e c escalares e $\boldsymbol{B}$ uma matriz qualquer.

Escalares, Vetores, Matrizes e Tensores

Broadcasting

Em Deep Learning, o conceito de Broadcasting permite somar um vetor por uma matriz, copiando implicitamente o vetor nas outras colunas.

$$\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{b}, \quad C_{i,j} = A_{i,j} + b_j \quad ,$$

sendo b um escalar e $\boldsymbol{A}$ uma matriz qualquer.

Multiplicação de matrizes e vetores

A multiplicação envolvendo duas matrizes resulta em uma terceira matriz. Essa multiplicação é representada como:

$$\boldsymbol{C} = \boldsymbol{A}\boldsymbol{B}$$

Multiplicação de matrizes e vetores

$$\begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix} \times \begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \overbrace{1 \times 2 + 3 \times 0}^{c_{11}} \\ \end{bmatrix}$$

Multiplicação de matrizes e vetores

$$\begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}_{2,2} \times \begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix}_{2,2} = \begin{bmatrix} \overset{c_{11}}{1 \times 2 + 3 \times 0} & \\ & \end{bmatrix}_{2,2}$$

Multiplicação de matrizes e vetores

$$\begin{bmatrix}1 & 3\\ 2 & 5\end{bmatrix}_{2,2} \times \begin{bmatrix}2 & 2\\ 0 & 1\end{bmatrix}_{2,2} = \begin{bmatrix}\overset{c_{11}}{1\times 2+3\times 0} & \\ & \end{bmatrix}_{2,2}$$

Multiplicação de matrizes e vetores

$$\begin{bmatrix}1 & 3\\2 & 5\end{bmatrix}_{2,2} \times \begin{bmatrix}2 & 2\\0 & 1\end{bmatrix}_{2,2} = \begin{bmatrix}\overset{c_{11}}{1\times 2+3\times 0} & \\ & \end{bmatrix}_{2,2}$$

Multiplicação de matrizes e vetores

$$\begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}_{2,2} \times \begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix}_{2,2} = \begin{bmatrix} \overset{c_{11}}{1 \times 2 + 3 \times 0} & \\ & \end{bmatrix}_{2,2}$$

Multiplicação de matrizes e vetores

$$\begin{bmatrix}1 & 3\\ 2 & 5\end{bmatrix}_{2,2} \times \begin{bmatrix}2 & 2\\ 0 & 1\end{bmatrix}_{2,2} = \begin{array}{c} c_{11} \\ \begin{bmatrix}1 \times 2 + 3 \times 0 & \\ & \end{bmatrix}_{2,2}\end{array}$$

Multiplicação de matrizes e vetores

$$\begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}_{2,2} \times \begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix}_{2,2} = \begin{bmatrix} \overset{c_{11}}{\boxed{1 \times 2 + 3 \times 0}} & \\ & \end{bmatrix}_{2,2}$$

Multiplicação de matrizes e vetores

$$\begin{bmatrix}1 & 3\\ 2 & 5\end{bmatrix}_{2,2} \times \begin{bmatrix}2 & 2\\ 0 & 1\end{bmatrix}_{2,2} = \begin{bmatrix}\overset{c_{11}}{\boxed{1 \times 2 + 3 \times 0}} \\ \\ \end{bmatrix}_{2,2}$$

Multiplicação de matrizes e vetores

Hadamard product: É a multiplicação ponto a ponto dos elementos da matriz, e é representado por $C = A \odot B$, sendo A, B e C matrizes de mesma dimensão.

$$\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \odot \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} A_1 \cdot B_1 \\ A_2 \cdot B_2 \end{bmatrix}$$

Multiplicação de matrizes e vetores

Produto interno: Seja 2 vetores de mesma dimensão, **x** e **y**, o produto interno entre eles pode visto como:

$$x^\top y = [x_1, x_2, ..., x_n] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Multiplicação de matrizes e vetores

Algumas propriedades:

- **Distributiva:**

$$A(B + C) = AB + AC.$$

- **Associativa**

$$A(BC) = (AB)C$$

- **Transposta**

$$(AB)^\top = B^\top A^\top$$

Multiplicação de matrizes e vetores

Algumas propriedades:

- **Comutativa (APENAS NO PRODUTO INTERNO!)**

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

Multiplicação de matrizes e vetores

Sistemas Lineares

Um sistema linear pode ser representado na forma

$$\boldsymbol{A}\mathbf{x} = \mathbf{b},$$

Sendo que os elementos de cada linha de $\boldsymbol{A}$ são os coeficientes de uma equação, cada elemento de **x** é uma variável e cada elemento de **b** é um resultado

Multiplicação de matrizes e vetores

$$\boldsymbol{A}_{1,1}x_1 + \boldsymbol{A}_{1,2}x_2 + \cdots + \boldsymbol{A}_{1,n}x_n = b_1$$

$$\boldsymbol{A}_{2,1}x_1 + \boldsymbol{A}_{2,2}x_2 + \cdots + \boldsymbol{A}_{2,n}x_n = b_2$$

$$\cdots$$

$$\boldsymbol{A}_{m,1}x_1 + \boldsymbol{A}_{m,2}x_2 + \cdots + \boldsymbol{A}_{m,n}x_n = b_m.$$

Matriz Inversa e Matriz Identidade

Matriz Identidade

Qualquer matriz multiplicada pela identidade, é a própria matriz

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

Matriz Inversa e Matriz Identidade

Matriz Inversa

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

Matriz Inversa e Matriz Identidade

Matriz Inversa

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

Matriz Inversa e Matriz Identidade

Matriz Inversa

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

$$\boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$$

Matriz Inversa e Matriz Identidade

Matriz Inversa

$$\boldsymbol{Ax} = \boldsymbol{b}$$

$$\boldsymbol{A}^{-1}\boldsymbol{Ax} = \boldsymbol{A}^{-1}\boldsymbol{b}$$

$$\boldsymbol{I}_n\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$$

Matriz Inversa e Matriz Identidade

Matriz Inversa

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

$$\boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$$

$$\boldsymbol{I}_n\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$$

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$$

Matriz Inversa e Matriz Identidade

Identificando soluções

- SPD – Sistema Possível e Determinado – possui uma única solução.
- SPI – Sistema Possível e Indeterminado – possui infinitas soluções.
- SI – Sistema Impossível – não possui solução.

Matriz Inversa e Matriz Identidade

Identificando soluções

- SPD:

$$\left[\begin{array}{cc|c} 1 & 3 & 1 \\ 0 & -7 & -4 \end{array}\right], \left[\begin{array}{ccc|c} 1 & 2 & 1 & 12 \\ 0 & -5 & 4 & -11 \\ 0 & 0 & -3 & -3 \end{array}\right], \left[\begin{array}{cccc|c} 2 & 0 & 0 & 0 & 10 \\ 0 & 1 & 0 & 0 & -12 \\ 0 & 0 & 2 & 0 & 6 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

Matriz Inversa e Matriz Identidade

Identificando soluções

- SPD:

$$\left[\begin{array}{cc|c} 1 & 0 & -5/7 \\ 0 & 1 & 4/7 \end{array}\right], \left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{array}\right], \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 5 \\ 0 & 1 & 0 & 0 & -12 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right].$$

Matriz Inversa e Matriz Identidade

Identificando soluções

- SI:

$$\begin{cases} x + y = 1 \\ x + y = 2 \end{cases}$$

$$\left[\begin{array}{ccccc|c} 0 & 0 & 0 & 0 & 0 & a \end{array}\right]$$

Matriz Inversa e Matriz Identidade

Identificando soluções

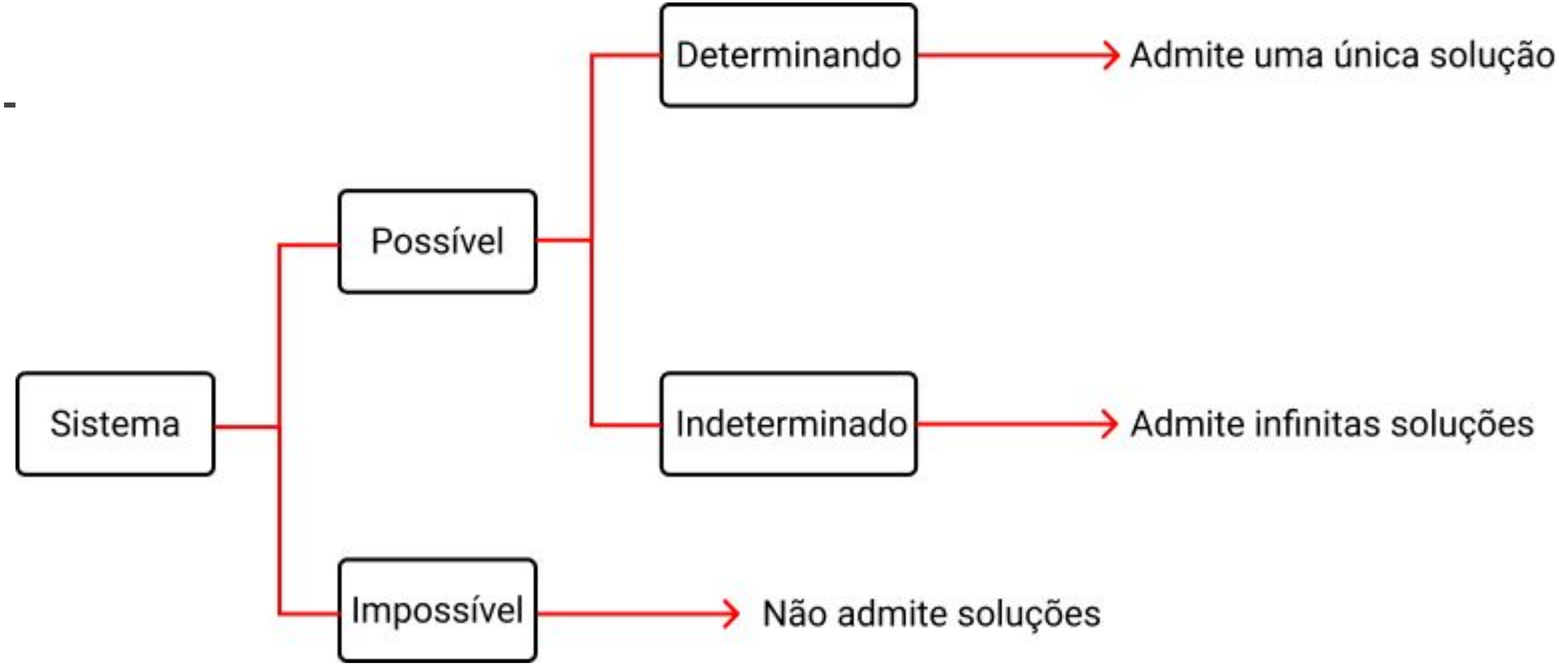
- SPI:

$$\left[\begin{array}{ccccc|c} 1 & 3 & 2 & -6 & 2 & 6 \\ 0 & -8 & 4 & 11 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & -9. \end{array}\right]$$

Matriz Inversa e Matriz Identidade

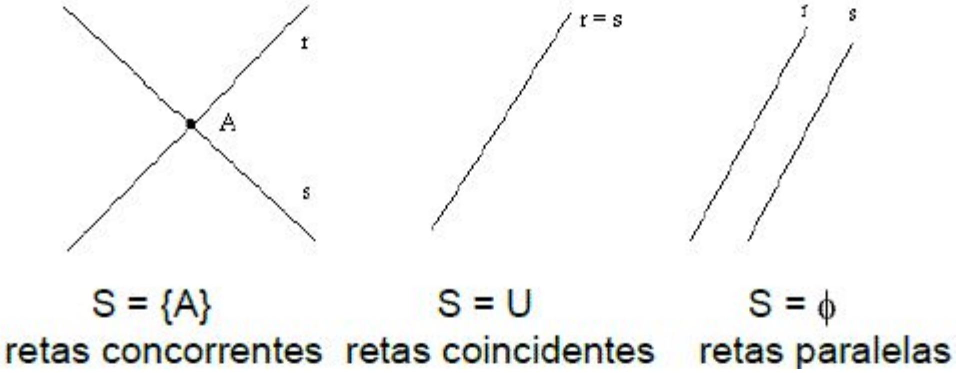
Identificando soluções

-



Matriz Inversa e Matriz Identidade

Identificando soluções



Matriz Inversa e Matriz Identidade

Combinação linear

É uma expressão construída a partir de um conjunto de termos, multiplicando cada termo por uma constante.

Matriz Inversa e Matriz Identidade

Combinação linear

Formalmente, uma combinação linear é a multiplicação de um conjunto de vetores $\mathbf{v}^{(i)}$ pelo correspondente coeficiente escalar, somando os resultados no final.

$$\sum_i c_i \boldsymbol{v}^{(i)}.$$

Matriz Inversa e Matriz Identidade

Dependência Linear

Útil para o estudo do paralelismo entre retas, planos e vetores

Matriz Inversa e Matriz Identidade

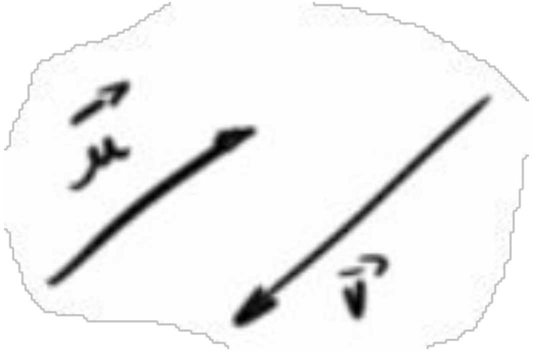
Sendo o conjunto de vetores {**x**}

- LD se x = 0
- LI se x != 0

Matriz Inversa e Matriz Identidade

Sendo o conjunto de vetores {**x**, **y**}

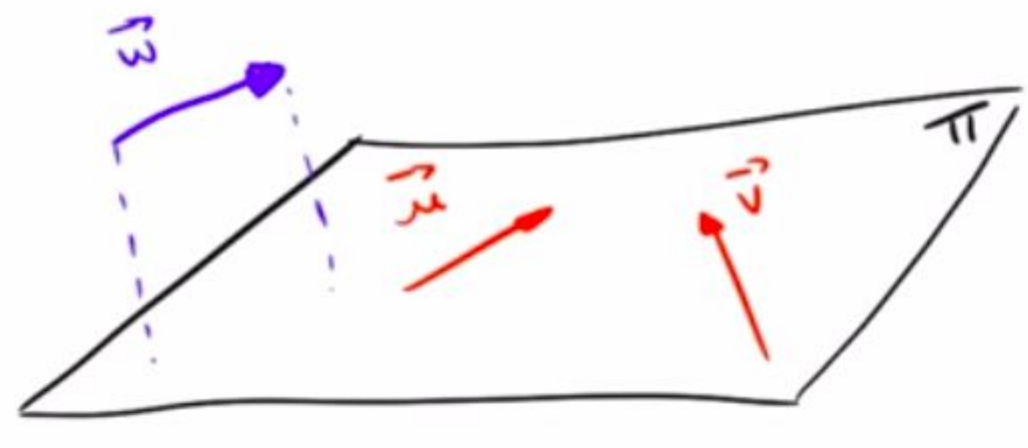
- LD se x e y são paralelos
- LI se x e y não são paralelos



Matriz Inversa e Matriz Identidade

Sendo o conjunto de vetores {**x, y, z**}

- LD se x, y e z são coplanares
- LI se x, y e z não são coplanares



Normas

A norma de um vetor pode ser entendido com a distância entre a origem e o respectivo vetor.

$$||\boldsymbol{x}||_p = \left(\sum_i |x_i|^p \right)^{\frac{1}{p}}$$

Norma L_2 muito usada, sendo que ela é a Distância Euclidiana.

Normas

- Em muitas aplicações de Machine Learning, a norma L_2 pode ser indesejada pois ela apresenta um crescimento pequeno em valores próximos à 0.

- A norma L_1 é uma solução, pois ela tem um crescimento uniforme nos valores próximos a 0 e mantém a simplicidade.

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

Normas

- A norma L1 é utilizada quando a diferença entre elementos nulos e não-nulos é muito importante.
- Incremento linear nos valores próximos à 0.

Normas

A **norma máxima** representa um determinado vetor pelo seu elemento máximo.

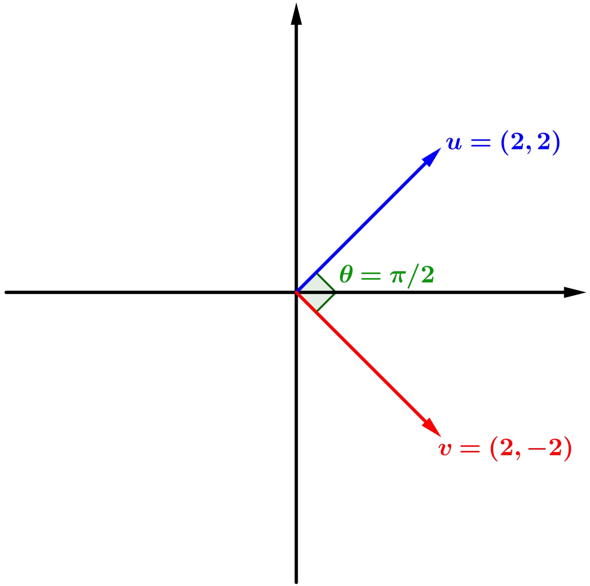
$$||\boldsymbol{x}||_\infty = \max_i |x_i|$$

Normas

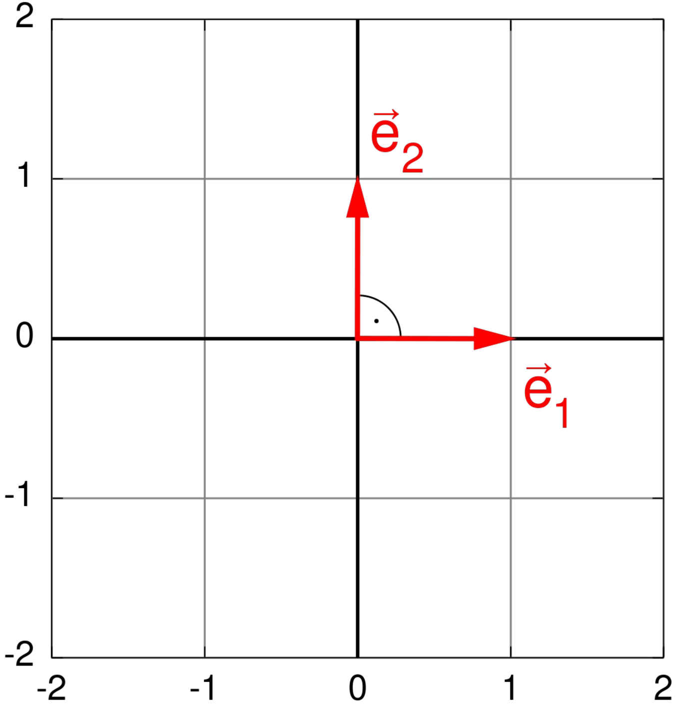
Quando se deseja medir o tamanho de uma matriz, a **norma de Frobenius** fornece uma alternativa.

$$||A||_F = \sqrt{\sum_{i,j} A_{i,j}^2}$$

Vetores Ortogonais



Vetores Ortonormais



Obrigado!