

11. ÁLGEBRA LINEAR

Álgebra linear é a matemática do século XXI."

– Gilbert Strang

"O advento da álgebra linear foi o desenvolvimento matemático mais importante do século XX."

– Steven Strogatz

INTRODUÇÃO A ÁLGEBRA LINEAR

Os fundamentos da álgebra linear são:

- **Vetores**

Um vetor é uma lista ordenada de números que representa magnitude e direção.

$$v = \begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix}$$

Este é um vetor tridimensional. Ele pode representar um ponto no espaço, uma velocidade, ou até mesmo dados abstratos.

Vetores são os objetos básicos da álgebra linear — eles vivem em espaços vetoriais, que são coleções de vetores que podem ser somados e escalados.

- **Matrizes**

Uma matriz é um arranjo retangular de números dispostos em linhas e colunas.

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

Esta é uma matriz 2×2 .

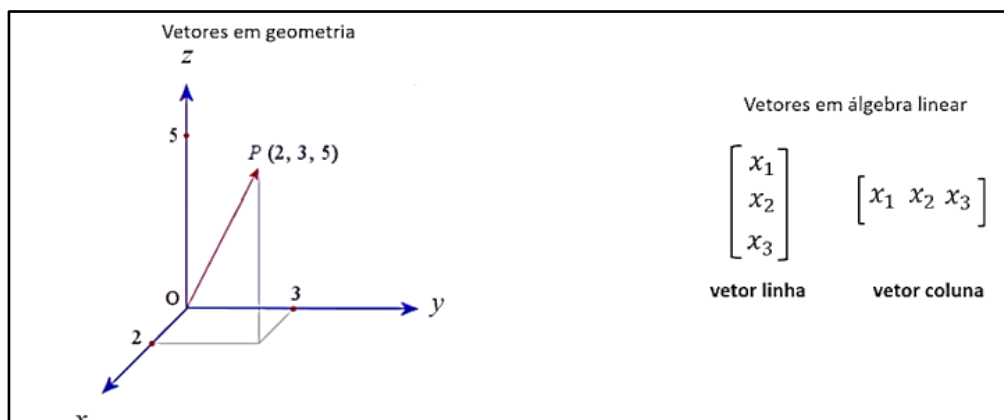
Matrizes são usadas para representar sistemas de equações lineares de forma compacta. Elas codificam transformações lineares (como rotações, reflexões, escalonamentos) e organizam dados de forma estruturada.

A álgebra linear fornece a estrutura para resolver sistemas lineares, analisar espaços vetoriais e entender como as transformações atuam sobre eles. Portanto, o nome reflete sua essência: a álgebra de sistemas e transformações lineares.

De modo geral, os dados em álgebra linear são representados na forma de equações lineares. Estas equações lineares são, por sua vez, representadas na forma de matrizes e vetores.

Como muitos conceitos matemáticos, os vetores podem ser compreendidos e investigados de diferentes maneiras. Existem pelo menos duas formas de olhar para vetores:

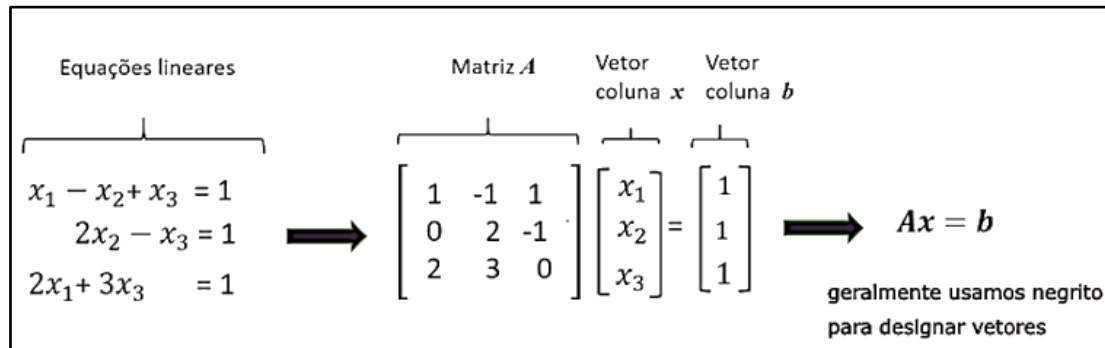
- **Geométrica** — Um vetor representa uma quantidade com magnitude e direção. Descrevemos um vetor como uma linha.



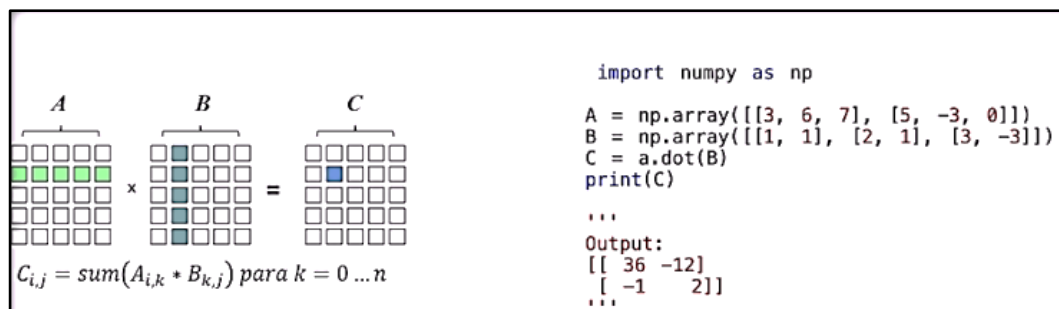
- **Algébrica** — Trata um vetor como um conjunto de valores escalares como uma única entidade, com adição, subtração e multiplicação escalar operando sobre o vetor inteiro.

Um vetor pode ser visto como uma matriz unidimensional.

Por exemplo, sem entrar em detalhes, veja como um conjunto de três equações lineares é transformado em matrizes e vetores e então reduzido a uma equação.



Geralmente, uma função é aplicada a todo um **conjunto de dados ao mesmo tempo**, sem focar em observações individuais. Matrizes nos ajudam a olhar e processar um lote de dados como uma única entidade. Por exemplo, a multiplicação de matrizes é realizada em uma única etapa.



Digamos que temos um conjunto de dados contendo 100.000 amostras e queremos realizar algum tipo de operação em cada um desses pontos. Do ponto de vista computacional, se usarmos um loop regular para isso, as coisas rapidamente se tornarão muito ineficientes. Se pudermos, em vez disso, vetorizar nossa operação e apenas realizar uma (ou algumas) grandes operações de vetor ou matriz usando algo como NumPy, nosso código será executado muito mais rápido.

Linguagens de programação como Python e R oferecem maneiras eficientes de implementar álgebra linear diretamente.

Abaixo, o banco de dados tem 7 milhões de observações e 100 variáveis dependentes, com cada observação representando uma equação linear com 100 parâmetros desconhecidos ($\theta_0, \theta_1, \dots$), que queremos descobrir, por exemplo, através de uma regressão linear.

ID	Target(y)	x_1	x_2	x_3	...	x_{100}	7.000.000 de equações
1	88.016	0.637	8.442	2.116	2.772	8.448	$\theta_1 0,637 + \theta_2 8,442 + \dots + \theta_{100} 8,448 = 88,016$
2	138.755	4.595	4.388	4.926	2.682	8.618	$\theta_1 4,595 + \theta_2 4,388 + \dots + \theta_{100} 8,618 = 138,755$
3	38.284	4.292	0.861	9.158	6.607	4.085	.
4	82.148	8.560	3.856	1.038	1.688	2.893	.
5	33.790	2.272	1.834	6.301	1.938	2.070	.
...	7.389	3.956	0.344	4.337	7.563	8.591	.
7.000.000	191.453	2.220	7.843	6.463	8.732	9.282	$\theta_1 2,220 + \theta_2 7,843 + \dots + \theta_{100} 9,282 = 191,453$
7.000.001	15.349	6.105	1.075	1.163	9.241	7.239	

Veremos como as matrizes são usadas para resolver este tipo de problema em uma única etapa.

TIPOS DE MATRIZES

Compreender a teoria de matrizes é crucial, pois frequentemente encontramos diferentes tipos de matrizes para várias finalidades. Estas incluem matrizes de dados, matrizes de pesos e coeficientes, matrizes de transformação, matrizes de payoff e muitas outras.

Notação

Você sabe que computadores são projetados para processamento binário, ou seja, apenas 0 e 1. Uma imagem é armazenada usando as intensidades de cada pixel em uma estrutura chamada matriz. Esta matriz pode então ser processada para identificar o objeto.

O estilo de notação para matrizes e vetores varia entre livros didáticos, e pode ser confuso no início.

Matrizes são tipicamente denotadas por letras maiúsculas em negrito, como $\mathbf{A}$ ou $\mathcal{A}$. O negrito distingue matrizes de variáveis escalares, que são frequentemente representadas por letras maiúsculas ou minúsculas regulares.

Similarmente, letras minúsculas em negrito (como $\mathbf{x}$ ou $\mathbf{x}$) são frequentemente usadas para representar vetores.

Uma matriz $\mathcal{A}$ é um conjunto de elementos, organizados em linhas e colunas. Geralmente, as linhas são denotadas por i e as colunas por j (ou m e n).

A forma explícita é:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, a_{ij} \in \mathbb{R}$$

Aqui afirmamos que cada entrada a_{ij} é um número real.

Uma forma abreviada padrão é:

$$\mathbf{A} \in \mathbb{R}^{m \times n}$$

Isto se lê: Matriz $\mathcal{A}$ pertence ao conjunto de todas as matrizes $m \times n$ com entradas reais. Isso deixa explícito que as entradas da matriz são números reais.

Por exemplo:

$$\mathbf{A} \in \mathbb{R}^{3 \times 2} \Rightarrow \mathcal{A} \text{ é uma matriz 3 por 2 com entradas reais.}$$

Notação de Vetor Coluna

$$\mathbf{A} = (\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n), \mathbf{a}_j \in \mathbb{R}^m$$

Cada coluna é um vetor em $\mathbb{R}^m$.

Notação de Vetor Linha

$$\mathbf{A} = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_m \end{pmatrix}, \mathbf{r}_i \in \mathbb{R}^n$$

Cada linha é um vetor em $\mathbb{R}^n$.

Matriz Quadrada

Uma matriz com o mesmo número de linhas e colunas: $m = n$

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}, \text{Tamanho: } 3 \times 3$$

Matriz Retangular

Uma matriz onde o número de linhas e colunas difere: $m \neq n$

$$\begin{pmatrix} a & b \\ d & e \\ g & h \end{pmatrix}, \text{Tamanho: } 3 \times 2$$

Matriz Diagonal

Uma matriz quadrada onde todos os elementos fora da diagonal são zero:

$$\begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix}$$

Matriz Identidade

Uma matriz diagonal especial com uns na diagonal:

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

Matriz Transposta

A transposta de uma matriz $\mathbf{A}$, denotada $\mathbf{A}^T$, é obtida trocando linhas por colunas:
Se

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

então

$$\mathbf{A}^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

Matriz Simétrica

Uma matriz quadrada que é igual à sua transposta: $\mathbf{A} = \mathbf{A}^T$.

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}$$

Aqui, as entradas são espelhadas através da diagonal.

Tabela Resumo

Tipo	Formato	Propriedade Chave
Quadrada	$m = n$	Linhas e colunas iguais
Retangular	$m \neq n$	Linhas e colunas diferentes
Diagonal	Quadrada	Zeros fora da diagonal
Identidade	Quadrada	Diagonal de uns, age como "1"
Transposta	Qualquer	Linhas $\leftrightarrow$ Colunas
Simétrica	Quadrada	Igual à sua transposta ($A = A^T$)

DETERMINANTES

O determinante é uma **função matricial** que associa um escalar a cada matriz quadrada. Ele transforma essa matriz em um número real. O determinante fornece informações valiosas sobre a matriz. Esta função permite saber se a matriz tem inversa ou não, porque aquelas que não têm são precisamente aquelas cujo determinante é igual a 0.

O determinante fornece informações valiosas sobre a matriz, especialmente sobre sua invertibilidade:

- $\det(A) \neq 0 \rightarrow$ A matriz é **invertível** (não-singular): possui inversa única e o sistema $Ax = b$ tem solução única.
- $\det(A) = 0 \rightarrow$ A matriz é **singular**: não possui inversa, e o sistema pode ter infinitas soluções ou nenhuma.

Veremos mais adiante por que isso é importante.

Abaixo mostramos o determinante de uma matriz de segunda ordem, que é a diferença entre o produto dos termos da diagonal principal e o produto dos termos da diagonal secundária.

Dada uma matriz A :

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

O determinante é calculado como:

$$\det(A) = ad - bc$$

Exemplo: Calcule o determinante de:

$$\begin{pmatrix} -5 & -4 \\ -2 & -3 \end{pmatrix}$$

$$\det = (-5)(-3) - (-4)(-2) = 15 - 8 = 7$$

MATRIZ INVERSA

Da álgebra elementar sabemos que se existem dois números cuja multiplicação é 1, ou seja, $x \cdot \frac{1}{x} = 1$, então esses dois números são chamados inversos um do outro. Da mesma forma, em álgebra linear, se duas matrizes multiplicadas juntas produzem uma matriz identidade (I), então as matrizes são chamadas inversas uma da outra.

$$AB = I$$

Uma matriz invertível — não-singular — representa uma transformação que pode ser revertida. Isso significa que cada saída da transformação corresponde a exatamente uma entrada, e podemos recuperar a entrada original a partir da saída. Uma matriz invertível também garante que o sistema de equações tenha uma solução única. Se a matriz não for invertível, isso indica que as equações são dependentes (levando a infinitas soluções) ou inconsistentes (levando a nenhuma solução).

Exemplo 1: Seja a matriz A :

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

Então:

$$\det(\mathbf{A}) = (1)(4) - (2)(3) = 4 - 6 = -2$$

Como o determinante é diferente de zero, $\mathbf{A}$ é invertível, e podemos encontrar sua inversa.

Exemplo 2: Considere a matriz:

$$\mathbf{B} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$$

$$\det(\mathbf{B}) = (1)(4) - (2)(2) = 4 - 4 = 0$$

Em contraste, para esta matriz o determinante é zero, **$\mathbf{B}$ não é invertível**, indicando que o sistema de equações que ela representa não tem uma solução única. **$\mathbf{B}$ não é invertível** — é uma matriz **singular**. Isso significa que suas linhas (ou colunas) são linearmente dependentes.

Por exemplo, existem 10 matrizes singulares usando apenas 0 e 1:

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

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

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

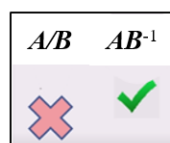
Todas essas matrizes têm determinante igual a 0.

O processo de inversão de uma matriz é mostrado abaixo.

Processo de inversão de uma matriz

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \xrightarrow{\text{inversão de } \mathbf{A}} \mathbf{A}^{-1} = \frac{1}{\text{determinante}} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \quad \mathbf{A}\mathbf{A}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \xrightarrow{\text{matriz identidade}}$$

A inversão de matrizes é usada extensivamente. Isso ocorre porque não podemos dividir matrizes. Não existe conceito de divisão por uma matriz, mas podemos multiplicar uma matriz por uma inversa, o que resulta essencialmente na mesma coisa.



MATRIZES ORTONORMAL E ORTOGONAL

1) Uma **matriz ortonormal** é uma matriz cujas colunas (ou linhas) formam um conjunto ortonormal, ou seja:

- Cada coluna tem **norma 1**.

Se uma coluna v da matriz é um vetor:

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

então sua norma é:

$$\|v\| = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}$$

Para a matriz ser ortonormal, cada coluna deve satisfazer:

$$\|v\| = 1$$

Isso garante que cada vetor tem comprimento unitário.

- **Colunas diferentes são ortogonais entre si.**

Isso implica a propriedade central:

$$Q^T Q = I$$

onde Q é a matriz ortonormal e I é a matriz identidade.

- As colunas de Q são vetores unitários e mutuamente perpendiculares.
- A inversa de Q é simplesmente sua transposta:

$$Q^{-1} = Q^T$$

São muito usadas em transformações rígidas, decomposições (SVD), computação gráfica e em física. Considere:

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

As colunas são ortogonais e têm norma 1. Verificando:

$$Q^T Q = I$$

Portanto, Q é ortonormal.

Geometricamente, vetores ortonormais são perpendiculares entre si, porque “ortonormal” combina duas ideias:

- Orto $\rightarrow$ ortogonais (ângulo de 90°)
- Normal $\rightarrow$ cada vetor tem norma 1

2) Na **matriz ortogonal** as colunas perpendiculares entre si, mas **não necessariamente de norma 1**:

$$Q^T Q = D \text{ (diagonal, não necessariamente } I)$$

Por exemplo:

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

As colunas são $\mathbf{c}_1 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ e $\mathbf{c}_2 = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$

Verificar ortogonalidade (produto escalar = 0):

$$\mathbf{c}_1 \cdot \mathbf{c}_2 = (2)(0) + (0)(3) = 0$$

As colunas são **perpendiculares** — condição de ortogonalidade satisfeita.

Verificar normas:

$$\|\mathbf{c}_1\| = \sqrt{2^2 + 0^2} = 2 \neq 1$$

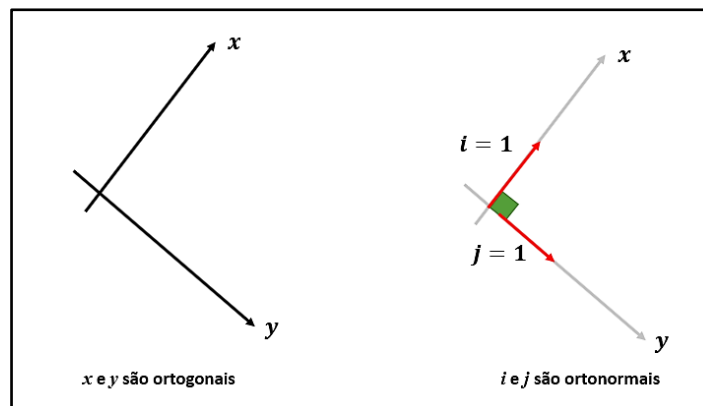
$$\|\mathbf{c}_2\| = \sqrt{0^2 + 3^2} = 3 \neq 1$$

As normas **não são 1** — condição de normalidade **não satisfeita**.

Verificar $A^T A$:

$$A^T A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix} \neq I$$

Propriedade	Ortogonal	Ortonormal
Colunas perpendiculares	sim	sim
Colunas com norma 1	não (qualquer norma)	sim (norma = 1)
$Q^T Q$	Diagonal	Identidade
Inversa = Transposta	não	sim ($Q^{-1} = Q^T$)



Nesta fase, pode não estar claro por que essas operações serão úteis. No entanto, a álgebra linear é a linguagem na qual o **machine learning** é escrita, em que conceitos como ortonormalidade, inversão, *eigenvectors* são essências. Se pudermos entender os fundamentos da linguagem, estaremos em uma posição muito melhor para entender as ideias mais complexas que formam a espinha dorsal dos modelos quando estudarmos machine learning no Volume VI.

OPERAÇÕES MATRICIAIS

Já vimos as três entidades básicas usadas em álgebra linear, a saber: o escalar, o vetor e a matriz. Na verdade, todos são versões específicas de uma entidade mais geral conhecida como tensor. Agora veremos como formar operações entre essas entidades. Essas operações incluem adição e multiplicação.

Adição e Subtração de Matrizes

Para adicionar duas ou mais matrizes, elas devem ter o mesmo número de linhas e colunas. A soma (ou subtração) dessas matrizes resultará em outra matriz que também terá o mesmo número de linhas e colunas.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} 1+3 & 2+1 \\ 3+2 & 4+5 \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 5 & 9 \end{pmatrix}$$

Produto Escalar (Dot Product)

Existem várias maneiras de multiplicar matrizes (produto escalar, produto vetorial, etc.), mas basicamente usamos o produto escalar (*dot product*). O produto escalar e a multiplicação de matrizes são operações relacionadas, mas distintas, na álgebra linear:

Envolve dois vetores do mesmo tamanho. É a soma dos produtos de suas entradas correspondentes, resultando em um valor escalar. O produto escalar é usado principalmente para encontrar ângulos entre vetores ou a projeção de um vetor sobre outro.

$$\mathbf{a} = \begin{pmatrix} 1 \\ 3 \\ -5 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 4 \\ -2 \\ -1 \end{pmatrix}$$

$$\mathbf{a} \cdot \mathbf{b} = (1 \times 4) + (3 \times -2) + (-5 \times -1) = 4 - 6 + 5 = 3$$

O resultado é um escalar: 3.

Multiplicação de Matrizes

Generaliza o produto escalar para matrizes. Ao multiplicar duas matrizes, calcula-se o produto escalar das linhas da primeira matriz com as colunas da segunda matriz. O resultado é uma nova matriz onde o número de linhas vem da primeira matriz e o número de colunas da segunda.

Seja:

$$\mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{B} \in \mathbb{R}^{n \times p}$$

Então o produto $\mathbf{C} = \mathbf{AB}$ é definido e:

$$\mathbf{C} \in \mathbb{R}^{m \times p}$$

A multiplicação de matrizes só é possível quando o número de colunas em $\mathbf{A}$ é igual ao número de linhas em $\mathbf{B}$.

Exemplos

Dimensões incompatíveis:

$$\mathbf{A} \in \mathbb{R}^{2 \times 3}, \mathbf{B} \in \mathbb{R}^{4 \times 3}$$

A multiplicação $\mathbf{AB}$ não é definida porque as dimensões internas não coincidem: $3 \neq 4$

Dimensões compatíveis:

$$\mathbf{A} \in \mathbb{R}^{2 \times 3}, \mathbf{B} \in \mathbb{R}^{3 \times 2}$$

Neste caso

$$\mathbf{C} = \mathbf{AB} \in \mathbb{R}^{2 \times 2}$$

Vejamos um exemplo numérico:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} 10 & 11 \\ 20 & 21 \\ 30 & 31 \end{pmatrix}$$

$$\mathbf{AB} = \begin{pmatrix} 1 \cdot 10 + 2 \cdot 20 + 3 \cdot 30 & 1 \cdot 11 + 2 \cdot 21 + 3 \cdot 31 \\ 4 \cdot 10 + 5 \cdot 20 + 6 \cdot 30 & 4 \cdot 11 + 5 \cdot 21 + 6 \cdot 31 \end{pmatrix} = \begin{pmatrix} 140 & 146 \\ 320 & 335 \end{pmatrix}$$

$$\mathbf{AB} \in \mathbb{R}^{2 \times 2}$$

Produto de Vetores (Outer Product):

Sejam:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \mathbf{y} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

A multiplicação $\mathbf{xy}$ não é válida como escrita, porque ambos são vetores coluna — suas dimensões são incompatíveis para multiplicação matricial padrão.

Para tornar a multiplicação válida, transpomos o vetor $\mathbf{y}$ em um vetor linha:

$$\mathbf{y}^T = (1 \quad 2 \quad 3)$$

Agora o produto se torna:

$$\mathbf{xy}^T = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} (1 \quad 2 \quad 3) = \begin{pmatrix} x_1 \cdot 1 & x_1 \cdot 2 & x_1 \cdot 3 \\ x_2 \cdot 1 & x_2 \cdot 2 & x_2 \cdot 3 \\ x_3 \cdot 1 & x_3 \cdot 2 & x_3 \cdot 3 \end{pmatrix}$$

Isso é conhecido como **produto externo** de dois vetores.

Não Comutatividade

O símbolo para o produto escalar normalmente é: $\mathbf{A} \cdot \mathbf{B}$, $\mathbf{A} \times \mathbf{B}$ ou simplesmente $\mathbf{AB}$. Portanto, o produto escalar no contexto de matrizes geralmente é chamado de multiplicação de matrizes.

Em geral, $\mathbf{AB} \neq \mathbf{BA}$, ou seja, o produto de matrizes **não é comutativo**.

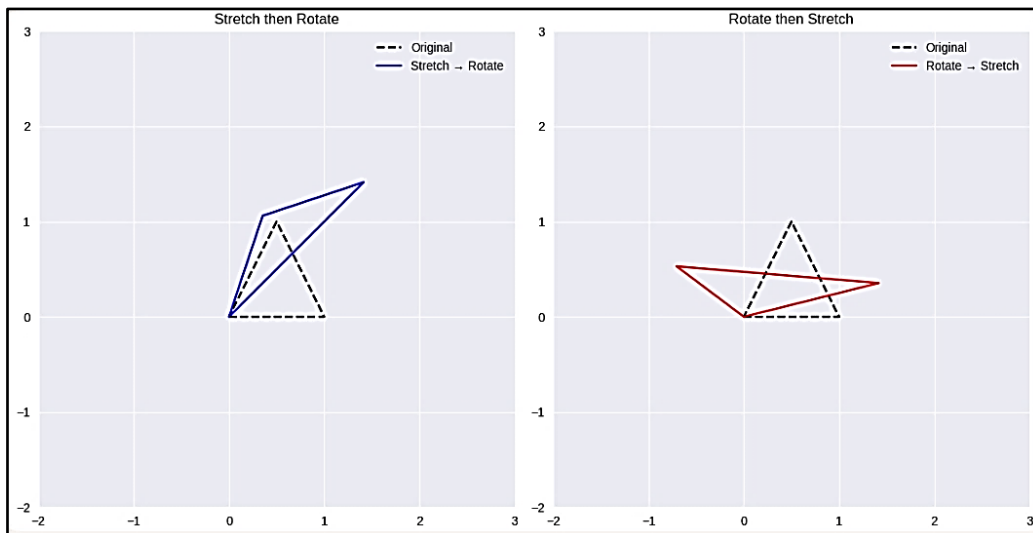
Exemplo:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \times \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}$$

$$\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \times \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}$$

O produto escalar de matrizes (multiplicação de matrizes) não é comutativo porque a operação representa a composição de transformações (que veremos no próximo capítulo) e a ordem em que as transformações são aplicadas importa.

Para ter uma intuição da não comutatividade, pense em uma matriz não como uma grade de números, mas como uma regra que transforma o espaço — alongando, rotacionando ou projetando vetores. Na figura abaixo, um triângulo é transformado usando duas operações: alongamento e rotação.



O gráfico à esquerda mostra o resultado de alongar primeiro e depois rotacionar. O gráfico à direita mostra o resultado de rotacionar primeiro e depois alongar. Esses dois resultados são claramente diferentes, confirmando visualmente que a multiplicação de matrizes não é comutativa.

É por isso que, em geral, a multiplicação de matrizes não é comutativa — $\mathbf{AB} \neq \mathbf{BA}$.

Transposta de um Produto

Para quaisquer duas matrizes compatíveis $\mathbf{A}$ e $\mathbf{B}$:

$$(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$$

Isso funciona porque a operação de transposição inverte a ordem da multiplicação.

Exemplo:

Sejam

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & 5 \\ 6 & 7 \end{pmatrix}$$

1. Calcule $\mathbf{AB}$:

$$\mathbf{AB} = \begin{pmatrix} 1 \cdot 0 + 2 \cdot 6 & 1 \cdot 5 + 2 \cdot 7 \\ 3 \cdot 0 + 4 \cdot 6 & 3 \cdot 5 + 4 \cdot 7 \end{pmatrix} = \begin{pmatrix} 12 & 19 \\ 24 & 43 \end{pmatrix}$$

2. Transposta de $\mathbf{AB}$:

$$\mathbf{A}^T \mathbf{B}^T = \begin{pmatrix} 12 & 24 \\ 19 & 43 \end{pmatrix}$$

3. Calcule $\mathbf{B}^T \mathbf{A}^T$:

$$\mathbf{B}^T = \begin{pmatrix} 0 & 6 \\ 5 & 7 \end{pmatrix}, \quad \mathbf{A}^T = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$$

$$\mathbf{B}^T \mathbf{A}^T = \begin{pmatrix} 0 \cdot 1 + 6 \cdot 2 & 0 \cdot 3 + 6 \cdot 4 \\ 5 \cdot 1 + 7 \cdot 2 & 5 \cdot 3 + 7 \cdot 4 \end{pmatrix} = \begin{pmatrix} 12 & 24 \\ 19 & 43 \end{pmatrix}$$

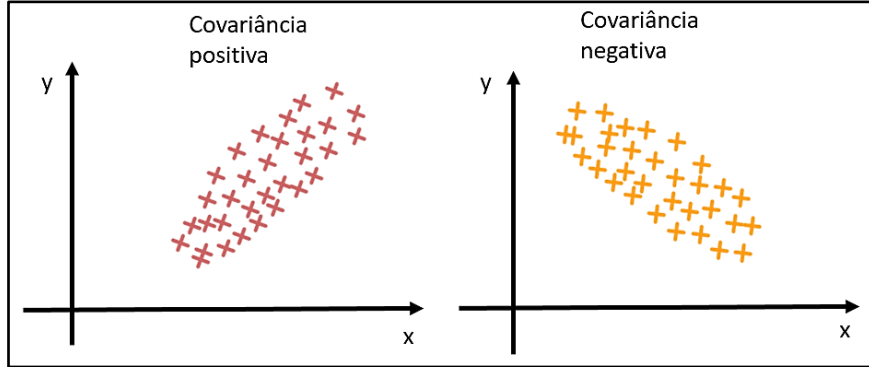
Corresponde exatamente a:

$$(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$$

A MATRIZ DE COVARIÂNCIA

Já vimos que a variância de uma variável descreve como os valores estão distribuídos e que a covariância é uma medida que nos informa sobre a dependência entre duas variáveis.

A matriz de covariância é um conceito fundamental em estatística, álgebra linear e machine learning. Ela captura como múltiplas variáveis aleatórias variam em conjunto — em outras palavras, como mudanças em uma variável estão relacionadas a mudanças nas outras.



Suponha que temos um conjunto de dados com n observações e p variáveis:

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix}$$

Seja $\bar{x}_j$ a média da j -ésima variável (coluna).

A matriz de covariância Σ é definida como:

$$\Sigma = \frac{1}{n-1} (\mathbf{X} - \bar{\mathbf{X}})^T (\mathbf{X} - \bar{\mathbf{X}}) \quad (1)$$

onde:

- $\mathbf{X} - \bar{\mathbf{X}}$ é a matriz dos dados centrados na média,
- cada entrada Σ_{ij} representa a covariância entre as variáveis i e j .
- Elementos diagonais Σ_{ii} : variância da variável i (quanto ela flutua).
- Elementos fora da diagonal Σ_{ij} : covariância entre as variáveis i e j :

Se Σ_{ij} :

- **positivo** → as variáveis aumentam juntas.
- **negativo** → uma aumenta quando a outra diminui.
- **zero** → as variáveis não são correlacionadas.

Por exemplo, se você tem duas variáveis: $\mathbf{X}$ = altura e $\mathbf{Y}$ = peso, medidas entre indivíduos, a matriz de covariância é:

$$\Sigma = \begin{pmatrix} \text{Var}(\mathbf{X}) & \text{Cov}(\mathbf{X}, \mathbf{Y}) \\ \text{Cov}(\mathbf{Y}, \mathbf{X}) & \text{Var}(\mathbf{Y}) \end{pmatrix}$$

Uma covariância positiva grande entre X e Y indica que indivíduos mais altos tendem a pesar mais.

Exemplo:

Vamos pegar um banco de dados com 3 linhas e 3 colunas correspondentes à matriz quadrada A e calcular a matriz de covariância.

ID	x	y	z
1	1	3	5
2	5	4	1
3	3	8	6

Extraindo apenas os dados numéricos das colunas $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{z}$, formamos a matriz $A \in \mathbb{R}^{3 \times 3}$:

$$A = \begin{pmatrix} 1 & 3 & 5 \\ 5 & 4 & 1 \\ 3 & 8 & 6 \end{pmatrix}$$

Usamos a expressão (1) acima.

1. Calcular a média de cada coluna:

$$\bar{x} = \frac{1+5+3}{3} = 3, \bar{y} = \frac{3+4+8}{3} = 5, \bar{z} = \frac{5+1+6}{3} = 4$$

2. Subtrair a média (centralizar os dados)

$$A_c = \begin{pmatrix} 1-3 & 3-5 & 5-4 \\ 5-3 & 4-5 & 1-4 \\ 3-3 & 8-5 & 6-4 \end{pmatrix} = \begin{pmatrix} -2 & -2 & 1 \\ 2 & -1 & -3 \\ 0 & 3 & 2 \end{pmatrix}$$

3. Calcular a matriz de covariância:

$$\Sigma = \frac{1}{n-1} A_c^T A_c$$

$$A_c^T A_c = \begin{pmatrix} 8 & 2 & -8 \\ 2 & 14 & 8 \\ -8 & 8 & 14 \end{pmatrix}$$

Então dividimos por $n-1=2$:

$$\Sigma = \frac{1}{2} \begin{pmatrix} 8 & 2 & -8 \\ 2 & 14 & 8 \\ -8 & 8 & 14 \end{pmatrix} = \begin{pmatrix} 4 & 1 & -4 \\ 1 & 7 & 4 \\ -4 & 4 & 7 \end{pmatrix}$$

Onde:

- $\Sigma_{xx} = 4$ (variância de x)
- $\Sigma_{yy} = 7$ (variância de y)
- $\Sigma_{zz} = 7$ (variância de z)
- $\Sigma_{xy} = 1$ (covariância entre x e y)
- $\Sigma_{xz} = -4$ (covariância entre x e z)
- $\Sigma_{yz} = 4$ (covariância entre y e z)

Descrição das relações		
Par	Covariância	Interpretação
x, y	+1	Pequena relação positiva – x e y aumentam levemente juntos.
x, z	-4	Forte relação negativa – x e z tendem a se mover em direções opostas.
y, z	+4	Relação positiva moderada – y e z aumentam juntos.

Matriz de Covariância vs. Matriz de Correlação (r)

Como sabemos, ambas capturam relações entre variáveis, mas em escalas diferentes.

$$r = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

A matriz de correlação é simplesmente a covariância **normalizada** pelos desvios padrão.

Quando usar cada uma:

Matriz de Covariância (Σ):

- Variáveis já estão na mesma escala (ex: pixels de imagem, componentes PCA)
- Queremos preservar a magnitude das variâncias
- Base para *eigenvectors* em PCA

Matriz de Correlação (r):

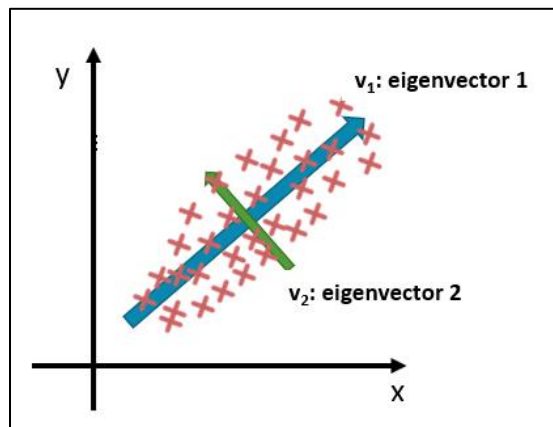
- Variáveis em escalas muito diferentes (ex: renda em R\$ vs. anos de estudo)
- Queremos comparar **força** das relações, não magnitude
- Análise exploratória e heatmaps de correlação

Na prática, se suas variáveis têm escalas muito diferentes, use correlação. Se já estão padronizadas (*z-score*), covariância e correlação coincidem.

AUTOVETORES E AUTOVALORES (*EIGENVECTORS E EIGENVALUES*)

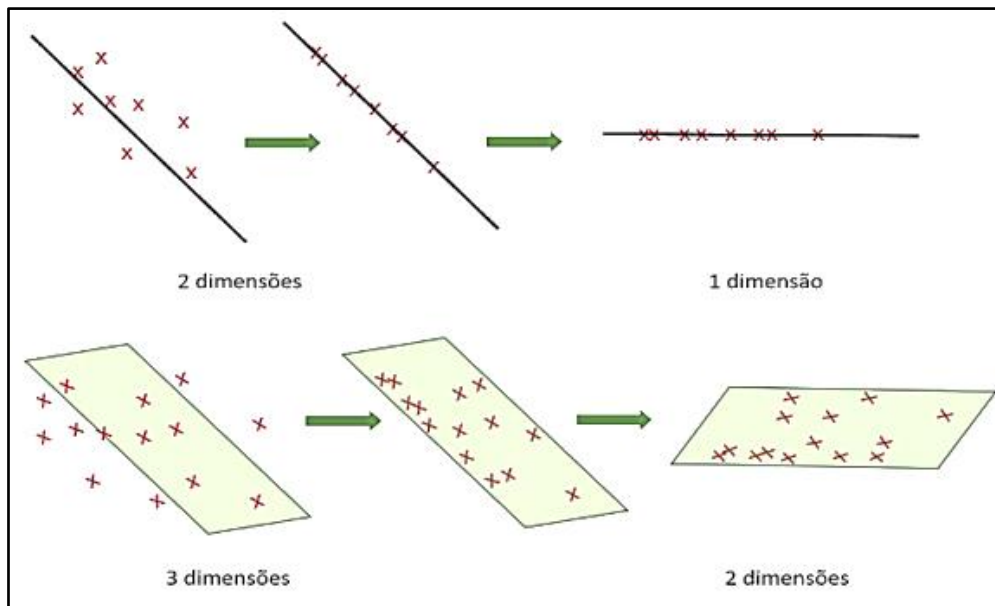
Para descobrir as direções principais de variação — isto é, as direções em que os dados mudam mais — calculamos seus *eigenvectors* e *eigenvalues*.

A ideia principal de *eigenvectors* e *eigenvalues* é transformar um conjunto de dados em um novo sistema de coordenadas onde as direções de máxima variância são representadas. Essas direções são dadas pelos *eigenvectors* da matriz de covariância dos dados.



Informalmente, podemos imaginar os dois vetores abaixo como dois *eigenvectors*. Vemos que a direção do *eigenvector* mais importante está associada à máxima variância dos dados, e o segundo *eigenvector* mais importante está associado à segunda maior variância dos dados. Essas direções onde há mais variância também são onde há mais informação.

A dimensionalidade é reduzida selecionando os *eigenvectors* principais, reduzindo-a, por exemplo, de 2 dimensões para 1 (uma linha) ou 3 dimensões para 2 (um plano) ou, no caso de mais dimensões, para hiperplanos com uma dimensão a menos (fig. abaixo).



Os *eigenvectors* produzem componentes principais de forma que a covariância entre eles seja zero. No entanto, observe que quando usamos os componentes principais, não sabemos exatamente o que eles significam.

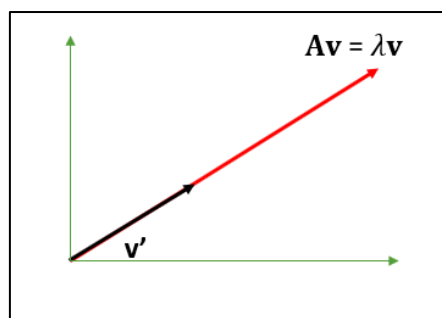
Interpretação Matemática

Considere as duas multiplicações entre uma matriz e um vetor abaixo. Temos a matriz A multiplicando os vetores $\mathbf{v}$ e $\mathbf{v}'$, o que resulta nos vetores $\mathbf{w}$ e $\mathbf{w}'$. Graficamente podemos representar vetores no plano coordenado como mostrado.

$$\begin{aligned}
 A\mathbf{v} &= \mathbf{w} \quad \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1.2 & + & 1.1 \\ 2.1 & + & 1.1 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \end{pmatrix} \\
 A\mathbf{v}' &= \mathbf{w}' \quad \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1.2 & + & 1.3 \\ 2.1 & + & 1.3 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}
 \end{aligned}$$

Em ambos os casos, o vetor resultante $\mathbf{w}$ não é colinear com o vetor original $\mathbf{v}$, ou seja, $\mathbf{w}$ não é um múltiplo de $\mathbf{v}$.

Agora considere a multiplicação na figura abaixo. Neste caso, a matriz A atua esticando o vetor $\mathbf{v}$, sem mudar sua direção. Então dizemos que $\mathbf{v}$ é um *eigenvector* de A e $\lambda\mathbf{v}$, um vetor com a mesma direção de $\mathbf{v}$ e com módulo igual a um escalar λ .



Mais formalmente, um *eigenvector* ou (autovetor ou vetor característico) de uma matriz A é apenas o vetor v tal que:

$$Av = \lambda v$$

para alguma constante λ que chamamos de *eigenvalue*.

Em um conjunto de observações, o *eigenvector* com o maior *eigenvalue* é a direção ao longo da qual o conjunto tem a máxima variância.

Considere a equação matricial acima. Nesta equação, A é uma matriz (n, n) , v é um vetor não nulo $(n, 1)$ e λ é um escalar. Qualquer valor para o qual esta equação tem solução é um *eigenvalue* da matriz A .

Podemos reescrever esta equação da seguinte forma:

$$\begin{aligned} A \cdot v - \lambda \cdot v &= 0 \\ A \cdot v - \lambda \cdot I \cdot v &= 0 \\ (A - \lambda \cdot I) \cdot v &= 0 \end{aligned}$$

Se v não é zero, esta equação só terá solução se:

$$\det(A - \lambda \cdot I) = 0$$

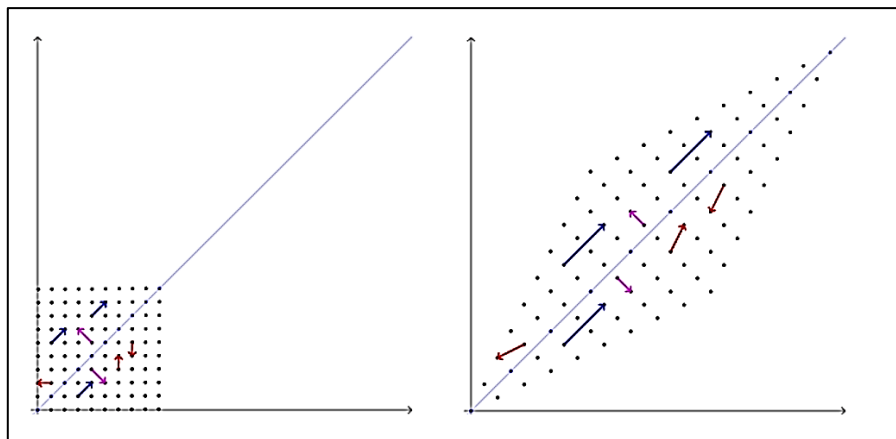
Esta equação é chamada de **equação característica** de A e é um polinômio em λ com n raízes. Essas raízes são os *eigenvalues* de A . Para cada *eigenvalue*, haverá um *eigenvector* para o qual a equação acima é verdadeira.

Exemplo:

Considere a matriz:

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

A figura abaixo mostra o efeito desta transformação nos pontos no plano. Os *eigenvectors* v desta transformação satisfazem $Av = \lambda v$, e os valores de λ para os quais o determinante da matriz é igual a zero são os *eigenvalues*.



Note na figura que na transformação alguns pontos se moveram ao longo da diagonal (setas azuis), outros perpendicularmente à diagonal (setas vermelhas) e outros inclinados em relação à diagonal (setas roxas).

Calculamos o determinante para encontrar a equação característica de A :

$$\det(A - \lambda I) = \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} - \lambda \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = \begin{vmatrix} 2 - \lambda & 1 \\ 1 & 2 - \lambda \end{vmatrix} = \lambda^2 - 4\lambda + 3$$

O polinômio característico é:

$$\lambda^2 - 4\lambda + 3 = 0$$

Igualando a 0, obtemos as raízes $\lambda = 1$ e $\lambda = 3$, que são os dois *eigenvalues* de A .

Para $\lambda = 1$:

$$(A - 1 \cdot I) \cdot \mathbf{v} = 0$$

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

Seja $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$. Então:

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_1 + v_2 \\ v_1 + v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$v_1 + v_2 = 0$$

Esta equação aparece duas vezes (redundantemente), então a solução é:

$$v_2 = -v_1$$

Os *eigenvectors* correspondentes a $\lambda = 1$ são todos os múltiplos escalares de:

$$\mathbf{v} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Para $\lambda = 3$:

$$(A - 3 \cdot I) \cdot \mathbf{v} = 0$$

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

Seja $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$. Então:

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} -v_1 + v_2 \\ v_1 - v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$-v_1 + v_2 = 0 \Rightarrow v_2 = v_1$$

Os *eigenvectors* correspondentes a $\lambda = 3$ são todos os múltiplos escalares de:

$$\mathbf{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Agora volte à figura acima e veja que a matriz de transformação A preserva a direção dos vetores roxos (paralelos a $\mathbf{v}_{\lambda=1}$) e azuis (paralelos a $\mathbf{v}_{\lambda=3}$). Os vetores vermelhos não são paralelos a nenhum dos *eigenvectors*, então suas direções são alteradas pela transformação. Os comprimentos dos vetores roxos permanecem inalterados após a transformação (devido ao seu *eigenvalue* de 1), enquanto os vetores azuis têm três vezes o comprimento do original (devido ao seu *eigenvalue* de 3).

Programas como Python, MATLAB ou R podem produzir *eigenvectors* que parecem diferentes daqueles que calculamos manualmente. No entanto, isso não é um erro — reflete uma verdade fundamental sobre *eigenvectors*: Eles não são únicos em magnitude, apenas em direção.

Se $\mathbf{v}$ é um *eigenvector* de uma matriz A , então qualquer múltiplo escalar $k\mathbf{v}$ (onde $k \neq 0$) também é um *eigenvector* associado ao mesmo *eigenvalue*. Assim, embora os valores numéricos dos componentes possam variar, a razão entre seus elementos — e, portanto, sua direção — permanece a mesma.

Softwares modernos como NumPy ou Scikit-learn geralmente normalizam os *eigenvectors* de modo que seu comprimento (norma) seja igual a 1:

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2} = 1$$

Os *eigenvectors* e *eigenvalues* de uma matriz de covariância estão no coração da **Análise de Componentes Principais (PCA)** — uma pedra angular da ciência de dados e machine learning que exploraremos no Volume VI. Eles também desempenham um papel central na teoria de redes, onde definem a centralidade por *Eigenvector*, uma medida de influência que sustenta o algoritmo *PageRank* do Google e outros modelos de propagação.

Em suma, os *eigenvalues* quantificam quanto de variação ou influência existe em uma determinada direção, e os *eigenvectors* mostram onde essa variação ou influência ocorre. Juntos, eles formam uma das pontes mais elegantes entre álgebra linear, geometria e análise de dados — um verdadeiro exemplo de como a estrutura matemática revela a ordem oculta de sistemas complexos.

SISTEMAS DE EQUAÇÕES LINEARES

Considere o sistema linear com 3 equações e 3 incógnitas:

$$\begin{aligned} 1x - 1y + 1z &= 1 \\ 0x + 2y - 1z &= 1 \\ 2x + 3y + 0z &= 1 \end{aligned}$$

Denotando as matrizes por A , X e B , respectivamente, o sistema original é:

$$A \cdot X = B$$

$$A = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 2 & -1 \\ 2 & 3 & 0 \end{pmatrix}, X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, B = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Onde a matriz de coeficientes é A , a matriz coluna das incógnitas é X e a matriz coluna das constantes é B . Para encontrar X , usamos a inversa da matriz A :

1. Comece com a equação:

$$A \cdot X = B$$

2. Multiplique ambos os lados à esquerda por A^{-1} :

$$A^{-1} \cdot A \cdot X = A^{-1} \cdot B$$

3. Use a identidade $A^{-1} \cdot A = I$:

$$I \cdot X = A^{-1} \cdot B$$

4. Simplifique:

$$X = A^{-1} \cdot B$$

No entanto, encontrar manualmente a inversa de uma matriz é trabalhoso e, na verdade, não faz sentido atualmente. Vetores e matrizes são os elementos básicos das linguagens de programação científica, portanto

não há razão para usar operações de álgebra matricial manualmente. Em outras palavras, você precisa entender o que essas operações fazem, mas não precisa realizá-las ou programá-las manualmente.

O sistema acima é resolvido em Python usando o pacote "symbols".

```
from sympy import symbols, Eq, solve

# Step 1: Define the variables
x1, x2, x3 = symbols('x1 x2 x3')

# Step 2: Define the equations
eq1 = Eq(x1 - x2 + x3, 1)
eq2 = Eq(2*x2 - x3, 1)
eq3 = Eq(2*x1 + 3*x3, 1)

# Step 3: Solve the system
solution = solve((eq1, eq2, eq3), (x1, x2, x3))

# Step 4: Print the solution
print("✅ Solution to the system of equations:")
for var, val in solution.items():
    print(f"{var} = {val}")
```

✅ Solution to the system of equations:
x1 = 2
x2 = 0
x3 = -1

TENSORES

Em Deep Learning, é comum ver tensores usados como a estrutura básica de dados. Tensores até aparecem no nome da principal biblioteca de ML do Google: "TensorFlow".

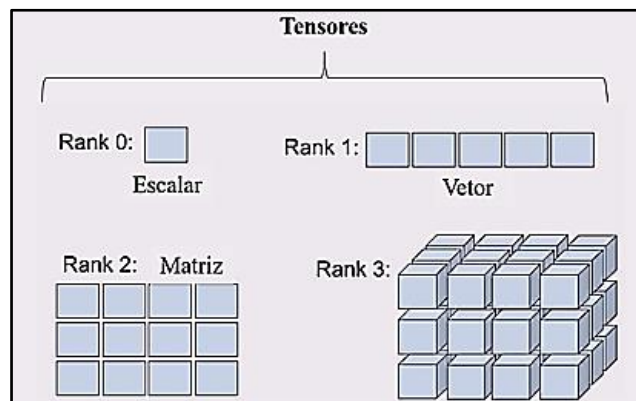


Tensores são um tipo de estrutura de dados usada em álgebra linear e, como vetores e matrizes, você pode calcular operações aritméticas com tensores. Mas o que é um tensor?

Embora tensores tenham sido inventados por físicos para poder descrever interações, no campo da Inteligência Artificial eles podem ser entendidos simplesmente como **contêineres de números**, assim como vetores ou matrizes.

Um tensor é uma **generalização de vetores e matrizes** e pode ser facilmente entendido como uma **matriz multidimensional**:

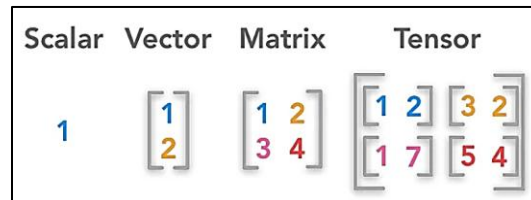
- Um tensor unidimensional pode ser representado como um **vetor**
- Um tensor bidimensional, como você pode adivinhar, pode ser representado como uma **matriz**
- Para ordens (*ranks*) superiores, simplesmente designamos o tensor como de ordem N



•

Tensores geralmente oferecem representações mais naturais de dados. Por exemplo:

- Uma **imagem** é geralmente representada por três dimensões: largura, altura e profundidade. Faz todo sentido codificá-la como um tensor de rank **3**
- Um **vídeo**, que consiste em imagens obviamente correlacionadas ao longo do tempo. Neste caso, precisamos de um tensor de rank **4**



Tipos de Tensores

- Tensores 0D / Escalares

Um tensor com zero dimensões é conhecido como tensor 0D ou escalar. Por exemplo, $[1]$ ou $[5,020]$ são tensores 0D.

- Tensores 1D / Vetores

Um tensor com uma dimensão é um tensor 1D ou vetor. Por exemplo, $[1, 2, 3, 4]$ pode ser considerado um tensor 1D.

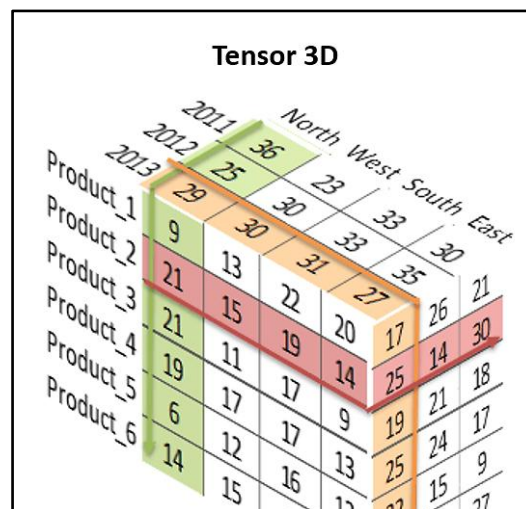
- Tensores 2D / Matrizes

Uma coleção de múltiplos vetores forma um tensor 2D, uma matriz. Por exemplo, se temos $[1, 2, 3]$, $[4, 5, 6]$ e $[7, 8, 9]$, combinar esses vetores cria um tensor 2D. É representado como:

$$M = [[1, 2, 3], [4, 5, 6], [7, 8, 9]]$$

- Tensores 3D

Um array tridimensional é um tensor 3D. Um tensor 3D também é conhecido como cubo.



- Tensores 4D

Um vetor de tensores 3D é um tensor 4D. Tensores 4D são frequentemente usados em análise de imagem. Neste caso, os quatro eixos denotam:

Por exemplo, $(3, 640, 640, 2)$ significa que o array contém 2 amostras de imagens de tamanho 640×640 pixels com 3 canais de cor.



Vemos que a **forma de um tensor** está relacionada às suas dimensões. É um array de números que representa o comprimento de cada dimensão. Esta é a característica mais importante e interessante de um tensor.

A maneira de identificar a forma de um tensor é descobrir o número de elementos em cada dimensão.



No exemplo acima, o tensor tem 3 dimensões e em cada dimensão há dois elementos, por isso a forma do tensor é $[2, 2, 2]$. Não importa se os elementos são arrays ou valores — ambos são contados como elementos ao determinar a forma do tensor.

Tipo	Dimensões	Exemplo em Python	Nome comum
0D	0	5 ou [5]	Escalar
1D	1	[1, 2, 3, 4]	Vetor
2D	2	[[1,2], [3,4]]	Matriz
3D	3	[[[1,2], [3,4]], [[5,6], [7,8]]]	Cubo/Array 3D
ND	N	Arrays aninhados	Tensor de ordem N

RESOLVENDO À MÃO!

Durante a Segunda Guerra Mundial, a resolução de sistemas lineares era uma tarefa crucial para cálculos balísticos e criptografia. Equipes de “computadores humanos” — muitas vezes formadas por mulheres — trabalhavam horas com papel, lápis e tabelas para determinar trajetórias de projéteis ou decifrar mensagens codificadas. Um dos métodos empregados era a regra de Cramer, que utiliza determinantes, e a matriz de cofatores para obter a inversa de uma matriz. A seguir, apresentamos um exemplo numérico passo a passo, ilustrando como esses cálculos eram realizados manualmente.

Problema: Cálculo de Trajetória Balística

Suponha que um engenheiro precise determinar três parâmetros x_1 , x_2 e x_3 (como ângulo de disparo, velocidade inicial e tempo de voo) a partir das seguintes equações lineares:

$$\begin{cases} 2x_1 + 3x_2 - x_3 = 5 \\ 4x_1 - x_2 + 2x_3 = 6 \\ -3x_1 + 2x_2 + 3x_3 = 4 \end{cases}$$

Na forma matricial, temos $\mathbf{Ax} = \mathbf{b}$, com:

$$\mathbf{A} = \begin{bmatrix} 2 & 3 & -1 \\ 4 & -1 & 2 \\ -3 & 2 & 3 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 5 \\ 6 \\ 4 \end{bmatrix}.$$

Passo 1: Calcular o determinante de $\mathbf{A}$

O determinante é necessário para verificar se a matriz é invertível e para o cálculo da inversa.

$$\begin{aligned} \det(\mathbf{A}) &= 2 \cdot \det \begin{bmatrix} -1 & 2 \\ 2 & 3 \end{bmatrix} - 3 \cdot \det \begin{bmatrix} 4 & 2 \\ -3 & 3 \end{bmatrix} + (-1) \cdot \det \begin{bmatrix} 4 & -1 \\ -3 & 2 \end{bmatrix} \\ &= 2((-1) \cdot 3 - 2 \cdot 2) - 3(4 \cdot 3 - 2 \cdot (-3)) - 1(4 \cdot 2 - (-1) \cdot (-3)) = -73 \end{aligned}$$

Como $\det(\mathbf{A}) \neq 0$, a matriz é invertível.

Passo 2: Construir a matriz dos cofatores

Para cada elemento a_{ij} , calculamos o menor $\mathbf{M}_{ij}$ (determinante da submatriz obtida eliminando a linha i e a coluna j) e aplicamos o sinal $(-1)^{i+j}$.

- **Cofator C_{11} :** elimina linha 1, coluna 1

$$\mathbf{M}_{11} = \det \begin{bmatrix} -1 & 2 \\ 2 & 3 \end{bmatrix} = (-1) \cdot 3 - 2 \cdot 2 = -3 - 4 = -7$$

Sinal: $+$ $\rightarrow C_{11} = -7$

- **Cofator C_{12} :** elimina linha 1, coluna 2

$$\mathbf{M}_{12} = \det \begin{bmatrix} 4 & 2 \\ -3 & 3 \end{bmatrix} = 4 \cdot 3 - 2 \cdot (-3) = 12 + 6 = 18$$

Sinal: $-$ $\rightarrow C_{12} = -18$

- **Cofator C_{13} :** elimina linha 1, coluna 3

$$\mathbf{M}_{13} = \det \begin{bmatrix} 4 & -1 \\ -3 & 2 \end{bmatrix} = 4 \cdot 2 - (-1) \cdot (-3) = 8 - 3 = 5$$

Sinal: $+$ $\rightarrow C_{13} = 5$

- **Cofator C_{21} :** elimina linha 2, coluna 1

$$\mathbf{M}_{21} = \det \begin{bmatrix} 3 & -1 \\ 2 & 3 \end{bmatrix} = 3 \cdot 3 - (-1) \cdot 2 = 9 + 2 = 11$$

Sinal: $-$ $\rightarrow C_{21} = -11$

- **Cofator C_{22} :** elimina linha 2, coluna 2

$$\mathbf{M}_{22} = \det \begin{bmatrix} 2 & -1 \\ -3 & 3 \end{bmatrix} = 2 \cdot 3 - (-1) \cdot (-3) = 6 - 3 = 3$$

Sinal: $+$ $\rightarrow C_{22} = 3$

- **Cofator C_{23} :** elimina linha 2, coluna 3

$$\mathbf{M}_{23} = \det \begin{bmatrix} 2 & 3 \\ -3 & 2 \end{bmatrix} = 2 \cdot 2 - 3 \cdot (-3) = 4 + 9 = 13$$

Sinal: $-$ $\rightarrow C_{23} = -13$

- **Cofator C_{31} :** elimina linha 3, coluna 1

$$\mathbf{M}_{31} = \det \begin{bmatrix} 3 & -1 \\ 2 & 2 \end{bmatrix} = 3 \cdot 2 - (-1) \cdot (-1) = 6 - 1 = 5$$

Sinal: $+$ $\rightarrow C_{31} = 5$

- **Cofator C_{32} :** elimina linha 3, coluna 2

$$M_{32} = \det \begin{bmatrix} 2 & -1 \\ 4 & 2 \end{bmatrix} = 2 \cdot 2 - (-1) \cdot 4 = 4 + 4 = 8$$

Sinal: $- \rightarrow C_{32} = -8$

- **Cofator C_{33} :** elimina linha 3, coluna 3

$$M_{33} = \det \begin{bmatrix} 2 & 3 \\ 4 & -1 \end{bmatrix} = 2 \cdot (-1) - 3 \cdot 4 = -2 - 12 = -14$$

Sinal: $+ \rightarrow C_{33} = -14$

Assim, a matriz dos cofatores é:

$$\text{Cof}(\mathbf{A}) = \begin{bmatrix} -7 & -18 & 5 \\ -11 & 3 & -13 \\ 5 & -8 & -14 \end{bmatrix}.$$

Passo 3: Obter a matriz adjunta

A adjunta é a transposta da matriz dos cofatores:

$$\text{adj}(\mathbf{A}) = (\text{Cof}(\mathbf{A}))^T = \begin{bmatrix} -7 & -11 & 5 \\ -18 & 3 & -8 \\ 5 & -13 & -14 \end{bmatrix}.$$

Passo 4: Calcular a inversa de $\mathbf{A}$

A inversa é dada por

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \text{adj}(\mathbf{A}):$$

$$\mathbf{A}^{-1} = \frac{1}{-73} \begin{bmatrix} -7 & -11 & 5 \\ -18 & 3 & -8 \\ 5 & -13 & -14 \end{bmatrix} = \begin{bmatrix} \frac{7}{73} & \frac{11}{73} & -\frac{5}{73} \\ \frac{18}{73} & -\frac{3}{73} & \frac{8}{73} \\ -\frac{5}{73} & \frac{13}{73} & \frac{14}{73} \end{bmatrix}.$$

Passo 5: Resolver o sistema $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$

Multiplicamos a inversa pelo vetor $\mathbf{b}$:

$$\begin{aligned} x_1 &= \frac{7}{73} \cdot 5 + \frac{11}{73} \cdot 6 + \left(-\frac{5}{73}\right) \cdot 4 = \frac{35 + 66 - 20}{73} = \frac{81}{73}, \\ x_2 &= \frac{18}{73} \cdot 5 + \left(-\frac{3}{73}\right) \cdot 6 + \frac{8}{73} \cdot 4 = \frac{90 - 18 + 32}{73} = \frac{104}{73}, \\ x_3 &= \left(-\frac{5}{73}\right) \cdot 5 + \frac{13}{73} \cdot 6 + \frac{14}{73} \cdot 4 = \frac{-25 + 78 + 56}{73} = \frac{109}{73}. \end{aligned}$$

Portanto, a solução do sistema é:

$$x_1 = \frac{81}{73}, \quad x_2 = \frac{104}{73}, \quad x_3 = \frac{109}{73}.$$

Esses valores poderiam ser usados para ajustar os parâmetros de um disparo ou decodificar uma mensagem. Note que todos os cálculos envolvem apenas operações aritméticas básicas, que eram feitas à mão na época, com o auxílio de tabelas de logaritmos e régua de cálculo. A matriz de cofatores e a adjunta são ferramentas poderosas que permitem obter a inversa de uma matriz de forma sistemática, mesmo quando os números são fracionários.

EXERCÍCIOS RESOLVIDOS DE ÁLGEBRA LINEAR

1) Seja A uma matriz quadrada inversível. Mostre que:

$$(A^T)^{-1} = (A^{-1})^T$$

Solução

Comece com a identidade:

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

Tome a transposta de ambos os lados:

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

Aplice a regra da transposta:

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

Isso mostra que $(A^{-1})^T$ é a inversa de A^T , logo:

$$(A^T)^{-1} = (A^{-1})^T$$

2) Dadas duas matrizes conformáveis A e B , isto é, cujas dimensões permitem uma operação matricial específica como soma ou multiplicação, prove que:

$$(A + B)(A - B) = A^2 - B^2$$

Solução

Expanda o lado esquerdo:

$$\begin{aligned} (A + B)(A - B) &= A(A - B) + B(A - B) \\ &= A^2 - AB + BA - B^2 \end{aligned}$$

Isso se simplifica para:

$$A^2 - B^2 + (BA - AB)$$

Logo, a identidade vale se e somente se A e B comutam (são conformáveis), isto é,

$$AB = BA.$$

3) Sejam A e B matrizes inversíveis de mesma ordem. Mostre que AB é inversível e que:

$$(AB)^{-1} = B^{-1}A^{-1}$$

Solução

Considere:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

De modo análogo:

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$$

Como $B^{-1}A^{-1}$ é inversa à esquerda e à direita de AB , concluímos:

$$(AB)^{-1} = B^{-1}A^{-1}.$$

4) Se A é simétrica ($A^T = A$) e B também é simétrica, mostre que AB é simétrica se, e somente se, $AB = BA$.

Solução

Suponha que AB seja simétrica:

$$(AB)^T = AB$$

Mas:

$$(AB)^T = B^T A^T = BA \text{ (pois ambas são simétricas)}$$

Logo:

$$BA = AB$$

Reciprocamente, se $AB = BA$, então:

$$(AB)^T = BA = AB$$

e AB é simétrica.

5) Considere a matriz simétrica

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

Encontre os autovalores de A .

Encontre uma base de autovetores para cada autovalor.

Verifique com Python (NumPy).

Solução à mão

1) Autovalores

Calcule:

$$\det(A - \lambda I) = 0$$

$$\det \begin{bmatrix} 2-\lambda & 0 & 0 \\ 0 & 3-\lambda & 1 \\ 0 & 1 & 3-\lambda \end{bmatrix} = (2-\lambda) \det \begin{bmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{bmatrix}$$

$$= (2-\lambda)((3-\lambda)^2 - 1) = (2-\lambda)(\lambda-2)(\lambda-4)$$

Logo:

$$\lambda_1 = 4, \lambda_2 = 2 \text{ (duplo)}$$

2) Autovetores

• Para $\lambda = 4$:

$$A - 4I = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

Da solução do sistema:

$$\mathbf{v} = (0, a, a)$$

Escolhendo $a = 1$:

$$\mathbf{v} = (0, 1, 1)^T$$

• Para $\lambda = 2$:

$$A - 2I = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

Obtemos:

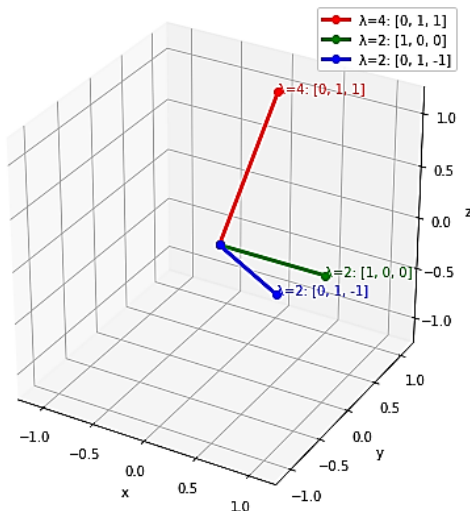
$$\mathbf{v}_1 = (1, 0, 0)^T, \quad \mathbf{v}_2 = (0, 1, -1)^T$$

Verificação no Python

```
import numpy as np
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D

# Define matrix
A = np.array([[2, 0, 0],
              [0, 3, 1],
              [0, 1, 3]], dtype=float)

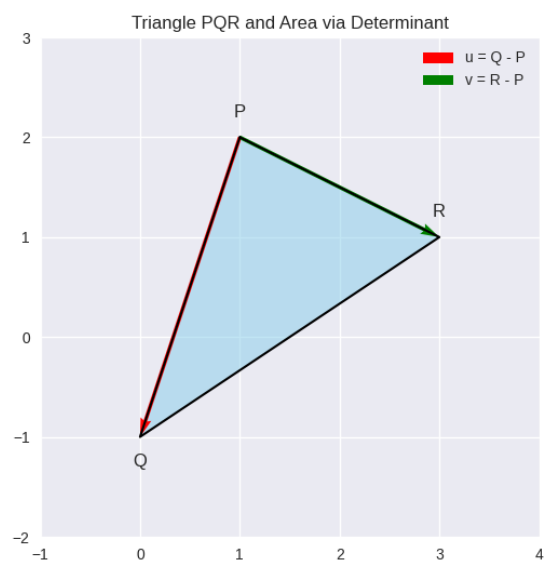
# Compute eigenvalues and eigenvectors
eigvals, eigvecs = np.linalg.eigh(A)
```



6) Mostre que a fórmula

$$\text{Área}(\triangle PQR) = \frac{1}{2} |\det [\mathbf{Q} - \mathbf{P} \quad \mathbf{R} - \mathbf{P}]|$$

fornece a área de um triângulo.



Solução

Defina os vetores:

$$\mathbf{u} = \mathbf{Q} - \mathbf{P}, \quad \mathbf{v} = \mathbf{R} - \mathbf{P}$$

O determinante da matriz formada por $\mathbf{u}$ e $\mathbf{v}$ dá a área do paralelogramo.

A área do triângulo é metade disso:

$$\text{Área} = \frac{1}{2} | \det [\mathbf{u}, \mathbf{v}] |$$
