

Álgebra Linear

Aula 01: Introdução

Kelvin Rafael Duarte Machado

setembro de 2025

O que é Álgebra Linear?

É o ramo da matemática cujo objeto de estudo são os vetores, espaços vetoriais, transformações lineares e sistemas de equações lineares.

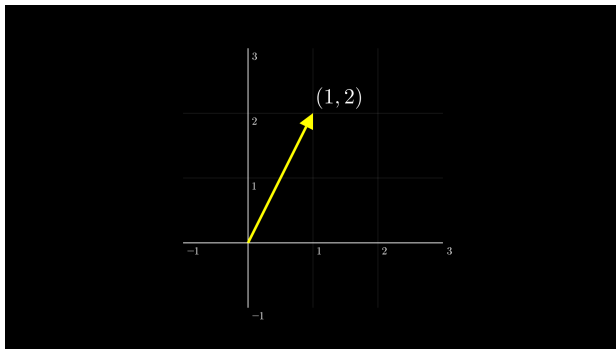
Onde a Álgebra Linear é importante?

- Ciências da computação: essencial na inteligência artificial, machine learning, processamento de imagens e computação gráfica;
- Engenharia: Usada na análise estrutural, processamento de sinais e robótica;
- Física: Fundamental para compreender fenômenos em diversas áreas desde a mecânica até a teoria de campos;
- Finanças: Aplicada para modelagem de sistemas e análise de dados complexos.

Vetores

Vetores são segmentos de reta orientados, usados na matemática e na física para representar grandezas que possuam módulo (intensidade), direção e sentido.

Exemplo de um vetor em $\mathbb{R}^2$



Sistema de equações lineares

É um conjunto finito de equações lineares aplicadas a um mesmo conjunto, igualmente finito de variáveis, onde a solução para o sistema é uma atribuição de números às incógnitas que satisfazem simultaneamente todas as equações do sistema.

Exemplo básico:

$$\begin{cases} x + 3y = 1 \\ 2x - y = -2 \end{cases}$$

Para resolver o sistema, resolvemos a equação superior para x em termos de y :

$$x = 1 - 3y$$

Agora, substituindo essa expressão para x na equação inferior:

$$2(1 - 3y) - y = -2$$

Isto resulta em uma equação envolvendo apenas a variável y . Resolvendo, obtemos $y = 4/7$ e ao substituir este valor de y , encontramos a variável $x = -5/7$.

Vetores como combinações lineares

A relação entre vetores e sistemas de equações lineares é intrínseca, já que um sistema de equações lineares pode ser expresso como a combinação linear de vetores, onde os vetores e os coeficientes se relacionam diretamente com as incógnitas do sistema.

Encontrar a solução de um sistema de equações lineares equivale a determinar os escalares que formam uma combinação linear específica de vetores.

Uma combinação linear de vetores é uma soma de múltiplos desses vetores, onde os múltiplos são escalares (números).

$$v = a_1 \vec{v}_1 + a_2 \vec{v}_2 + \cdots + a_n \vec{v}_n$$

Forma matricial

Uma forma extremamente útil de representar um sistema de equações lineares é através da forma matricial, em que a equação descrita anteriormente pode ser representada da seguinte forma:

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

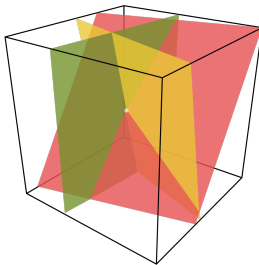
Onde A é uma matriz $m \times n$, $\mathbf{x}$ é um *vetor coluna* com n elementos e $\mathbf{b}$ é um *vetor coluna* com m elementos.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Sistema de equações lineares: interpretação geométrica

Um sistema de equações lineares pode ser interpretado como objetos geométricos (retas em 2D ou planos em 3D) em um gráfico, e a solução do sistema correspondente aos pontos de intersecção desses objetos.

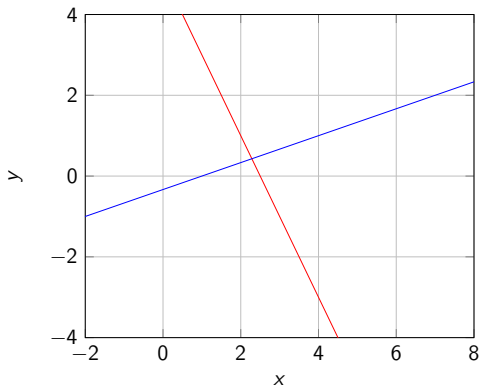
Representação da intersecção de 3 planos (solução de um sistema de 3 incógnitas):



Sistema de equações lineares possível e determinado

Exemplo:

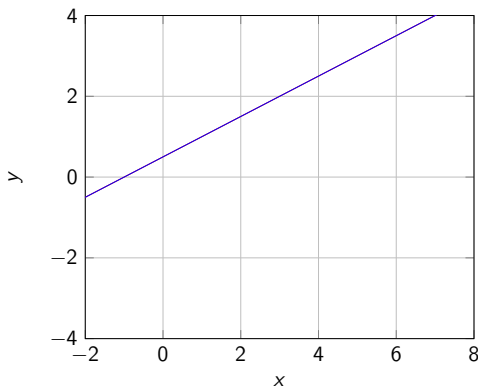
$$\begin{cases} 2x + y = 5 \\ x - 3y = 1 \end{cases}$$



Sistema de equações lineares possível e indeterminado

Exemplo:

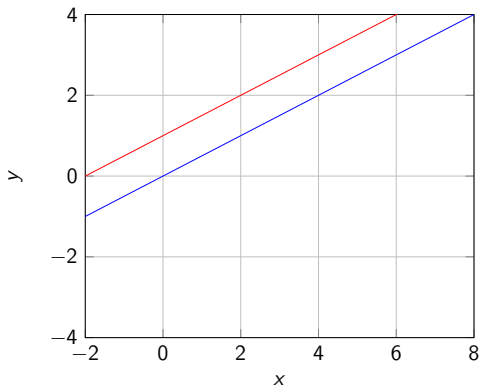
$$\begin{cases} -x + 2y = 1 \\ 3x - 6y = -3 \end{cases}$$



Sistema de equações lineares impossível

Exemplo:

$$\begin{cases} x - 2y = 2 \\ 2x - 4y = 0 \end{cases}$$



Solução computacional de um sistema de equações lineares

Resolução do sistema de equações lineares:

$$\begin{cases} x - 3y + z = 15 \\ x + y + 3z = 27 \\ 2x + 3y - 4z = 2 \end{cases}$$



```
[1]: import numpy as np
[2]: A = np.array([[1, -3, 1],
                  [1, 1, 3],
                  [2, 3, -4]])
[3]: b = np.array([[15], [27], [2]])
[4]: x = np.linalg.solve(A,b)
[5]: print(x)

[[10.9047619 ]
 [ 0.38095238]
 [ 5.23809524]]
```



Command Window

```
>> A = [1, -3, 1; 1, 1, 3; 2, 3, -4];
>> b = [15; 27; 2];
>> x = linsolve(A, b);
>> disp(x)
    10.9048
     0.3810
     5.2381
>> |
```

Solução computacional de um sistema de equações lineares

- Eliminação de Gauss
- Eliminação de Gauss-Jordan
- Decomposição LU
- Decomposição QR
- Decomposição de Cholesky

Álgebra Linear

Aula 02: Espaços vetoriais, Subespaço vetorial, Base e Dimensão

Kelvin Rafael Duarte Machado

setembro de 2025

Espaços vetoriais são conjuntos cujos elementos são chamados vetores, no qual são definidas duas operações: a *adição*, na qual para cada par de vetores u e v faz corresponder um terceiro vetor $u + v \in E$ chamado soma de u e v e a *multiplicação* por um número real, $\alpha \in \mathbb{R}$ por $v \in E$, que faz corresponder o vetor $\alpha.v = \alpha v$. Essas operações devem satisfazer para quaisquer $\alpha, \beta \in \mathbb{R}$ e $u, v, w \in E$, os axiomas apresentados a seguir:

Axiomas do espaço vetorial

- **comutatividade:** $u + v = v + u$;
- **associatividade:** $(u + v) + w = u + (v + w)$;
- **vetor nulo:** existe um vetor $0 \in E$ chamado vetor nulo, ou vetor zero, tal que $v + 0 = 0 + v = v$ para quaisquer $v \in E$;
- **inverso aditivo:** dado um vetor $v \in E$, existe um vetor $-v \in E$, tal que $-v + v = v + (-v) = 0$;
- **distributividade:** $(\alpha + \beta)v = \alpha v + \beta v$ e $\alpha(u + v) = \alpha u + \alpha v$;
- **multiplicação por 1:** $1.v = v$.

Exemplos de espaços vetoriais

- O conjunto dos números reais $\mathbb{R}$
- O conjunto $\mathbb{R}^n$
- O conjunto das matrizes reais $m \times n$ denotado $M_{m \times n}(\mathbb{R})$
- O conjunto das funções $(F) = f : \mathbb{R} \rightarrow \mathbb{R}$

Subespaço vetorial

Um subespaço de E é um subconjunto $F \subset E$ com as seguintes propriedades:

- $0 \in F$;
- Se $u, v \in F$ então $u + v \in F$;
- Se $v \in F$, então para todo $\alpha \in \mathbb{R}$, $\alpha v \in F$.

Uma combinação linear de um conjunto de vetores $v_1, \dots, v_n \in V$ é um vetor na forma:

$$v = a_1 v_1 + \dots + a_n v_n$$

Onde: $a_1, \dots, a_n \in \mathbb{R}$

Subespaço gerador (span)

O conjunto de todas as combinações lineares dos vetores $v_1, \dots, v_n$ em V é chamado de subespaço gerador de V ou, span de V , denotado por $S(v_1, \dots, v_m)$. Em outras palavras:

$$S(v_1, \dots, v_m) = \{a_1 v_1 + \dots + a_m v_m : a_1, \dots, a_m \in \mathbb{R}\}$$

Vetores linearmente independentes

Seja E um espaço vetorial. Diz-se que um conjunto $X \subset E$ é linearmente independente (L.I.), quando nenhum vetor $v \in X$ é combinação linear de outros elementos de X . Se um conjunto X é linearmente independente no espaço E e $\alpha_1 v_1 + \cdots + \alpha_m v_m = 0$ com $v_1, \dots, v_m \in X$ então $\alpha_1, \dots, \alpha_m = 0$. Sejam $v_1, \dots, v_m /$ vetores não nulos do espaço vetorial E . Se nenhum deles é combinação linear dos anteriores ou posteriores, então o conjunto $X = \{v_1, \dots, v_m\}$ é L.I..

Uma base de um espaço vetorial E é um conjunto $B \subset E$ linearmente independente que gera E .

Assim, todo vetor $v \in E$ se exprime, de modo único, como combinação linear $v = \alpha_1 v_1 + \cdots + \alpha_m v_m$ de elementos $v_1, \dots, v_m$ da base B . Se $B = \{v_1, \dots, v_m\}$ é uma base de E e $v = \alpha_1 v_1 + \cdots + \alpha_m v_m$, então os números $\alpha_1, \dots, \alpha_m$ são chamados *coordenadas do vetor v na base B* .

Base canônica

Os vetores:

$$e_1 = (1, 0, 0, \dots, 0)$$

$$e_2 = (0, 1, 0, \dots, 0)$$

$$\vdots$$

$$e_n = (0, 0, 0, \dots, 1)$$

Constituem um conjunto de geradores do espaço $\mathbb{R}^n$, linearmente independentes, sendo, portanto, uma base deste espaço.

Diz-se que o espaço vetorial E tem dimensão finita quando admite uma base $B = \{v_1, \dots, v_m\}$ com um número finito n de elementos. Este número, que é o mesmo para todas as bases de E , chama-se a *dimensão* do espaço vetorial E : $n = \dim E$.

Exemplos:

- $\mathbb{R}^2$ é um espaço vetorial com $\dim = 2$.
- Os monômios $1, x, \dots, x^n$ constituem uma base do espaço vetorial P_n dos polinômios de grau $\leq n$, logo P_n tem dimensão finita e $\dim P_n = n + 1$

Manipulação computacional de arrays

A implementação computacional da Álgebra Linear é beneficiada pela forma matricial, onde cada elemento é guardado na memória em um endereço específico, de acordo com o tipo de dado utilizado.

Vamos verificar através do Numpy e do SymPy como operar com matrizes.



Álgebra Linear

Aula 03: Transformações lineares

Kelvin Rafael Duarte Machado

setembro de 2025

Transformação linear

Sejam E, F espaços vetoriais. Uma transformação linear $A : E \rightarrow F$ é uma correspondência que associa a cada vetor $v \in E$ um vetor $A(v) = A \cdot v = Av \in F$ de modo que valham para quaisquer $u, v \in E$ e $\alpha \in \mathbb{R}$, as relações:

$$A(u + v) = Au + Av,$$

$$A(\alpha \cdot v) = \alpha \cdot Av$$

O vetor Av chama-se a imagem de v pela transformação A .

As transformações lineares $A : E \rightarrow E$ do espaço vetorial E em si mesmo são chamadas *operadores lineares* em E .

Por sua vez, as transformações lineares $\varphi : E \rightarrow \mathbb{R}$, com valores numéricos, são chamados *funcionais lineares*.

A imagem de A é o subconjunto $Im(A) \subset F$, formado por todos os vetores $w = Av \in F$ que são imagens de elementos de E pela transformação A .
O núcleo da transformação $A : E \rightarrow F$ é o subconjunto $N(A)$ dos vetores $v \in E$ tais que $Av = 0$.

Forma matricial

$$Av_j = a_{1j}w_1 + a_{2j}w_2 + \cdots + a_{mj}w_m = \sum_{n=1}^m a_{nj}w_n$$

Assim, a transformação linear $A : E \rightarrow F$ juntamente com as bases $V \subset E$ e $W \subset F$ determinam uma matriz $\mathbf{a} = [a_{ij}] \in M(m \times n)$ chamada a matriz de A relativamente a essas bases (ou nas bases V, W)

$$Av_j = \sum_{i=1}^n a_{ij}v_i, \quad (j = 1, \dots, n).$$

Transformações matriciais

Um sistema de equações lineares pode ser escrito na forma de matricial:

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

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Transformações matriciais

Exemplos:

- Transformação do $\mathbb{R}^4$ para o $\mathbb{R}^3$:

$$\begin{bmatrix} 2 & -3 & 1 & -5 \\ 4 & 1 & -2 & 1 \\ 5 & -1 & 4 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

- Transformação nula

$$A_0(\mathbf{x}) = 0\mathbf{x} = \mathbf{0}$$

- Operador identidade

$$A_1(\mathbf{x}) = I\mathbf{x} = \mathbf{x}$$

Operadores de rotação, reflexão e projeção

- Rotação pelo ângulo θ

$$w_1 = x \cos \theta - y \sin \theta$$

$$w_2 = x \sin \theta + y \cos \theta$$

$$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

- Reflexão no eixo y

$$A(x, y) = (-x, y)$$

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

- Reflexão no eixo x

$$A(x, y) = (x, -y)$$

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

- Projeção ortogonal sobre o eixo x

$$A(x, y) = (x, 0)$$

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

Operadores de dilatação, contração e cisalhamento

- Contração de fator k em $\mathbb{R}^2$

$$(0 \leq k \leq 1)$$

- Dilatação de fator k em $\mathbb{R}^2$

$$k > 1$$

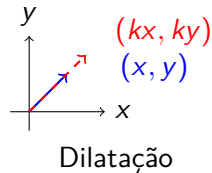
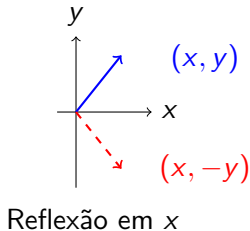
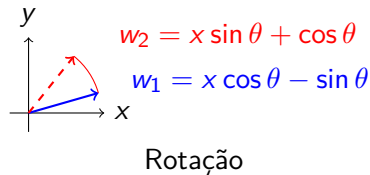
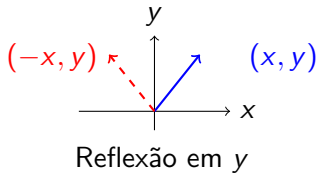
$$\begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$$

- Cisalhamento de $\mathbb{R}^2$ de fator k na direção x

$$A(x, y) = (x + ky, y)$$

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

Transformações em $\mathbb{R}^2$



Escreve-se E^* em vez de $\mathcal{L}(E, \mathbb{R})$ para representar o conjunto das transformações lineares do espaço vetorial E em $\mathbb{R}$ e o conjunto E^* representa os funcionais lineares $\varphi : E \rightarrow \mathbb{R}$, chamando-se o espaço vetorial *dual* de E .

Vetor É um elemento do espaço vetorial E . Pode ser interpretado como uma grandeza geométrica ou física que possui direção e magnitude.

Covetor É um elemento do espaço dual E^* . Ele atua como uma aplicação linear que associa a cada vetor um número real.

Usualmente, vetores são representados na forma matricial em colunas, enquanto os covetores aparecem na forma matricial em linhas (ou seja, na forma transposta).

Álgebra Linear

Aula 04: Matrizes

Kelvin Rafael Duarte Machado

setembro de 2025

Matriz relativa a bases

Já vimos que uma transformação linear $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ fica inteiramente determinada pela matriz $\mathbf{a} = [a_{ij}] \in M(m \times n)$, cujo ij -ésimo termo a_{ij} é a i -ésima coordenada do vetor $A \cdot e_j \in \mathbb{R}^m$.

Conhecendo essa matriz tem-se, para cada $v = (x_1, \dots, x_n) \in \mathbb{R}^n$, o valor $A \cdot v = y_1, \dots, y_m$ dado por:

$$y_i = a_{i1}x_1 + \dots + a_{in}x_n \quad (i = 1, \dots, m).$$

Sejam E, F espaços vetoriais de dimensão finita e $A : E \rightarrow F$ uma transformação linear. Fixadas bases $V = v_1, \dots, v_n \subset E$ e $W = w_1, \dots, w_m \subset F$, para cada $j = 1, \dots, n$ o vetor Av_j se exprime como combinação linear dos vetores da base W .

$$Av_j = a_{1j}w_1 + \dots + a_{mj}w_m = \sum_{i=1}^m a_{ij}w_i$$

Matriz de um operador linear

No caso em que $A : E \rightarrow E$ é um operador linear considera-se apenas uma base $V = v_1, \dots, v_n \subset E$ e a matriz $\mathbf{a} = [a_{ij}]$ do operador A relativamente à base V (ou na base V) é definida pelas n igualdades:

$$Av_j = \sum_{i=1}^n a_{ij} v_i \quad (j = 1, \dots, n).$$

Neste caso, $\mathbf{a} \in M(n \times n)$ é a matriz quadrada $n \times n$ cuja j -ésima coluna é formada pelas coordenadas do vetor

$$Av_j = a_{1j}v_1 + \dots + a_{nj}v_n$$

na base V .

Operações matriciais

Entre as transformações lineares, além das operações de soma $A + B$ e de multiplicação por escalar αA , existe também a multiplicação BA .

Mas antes, precisamos definir o *produto interno*.

Sejam $u = (\alpha_1, \dots, \alpha_n)$ e $v = (\beta_1, \dots, \beta_n)$ vetores em $\mathbb{R}^n$. O *produto interno* de u por v é definido como o número:

$$\langle u, v \rangle = \alpha_1 \beta_1 + \dots + \alpha_n \beta_n.$$

Sejam $\mathbf{b} = [b_{ij}] \in M(m \times n)$ e $\mathbf{a} = [a_{ij}] \in M(n \times p)$ matrizes tais que o número de colunas de $\mathbf{b}$ é igual ao número de linhas de $\mathbf{a}$. O produto da matriz $\mathbf{b}$ pela matriz $\mathbf{a}$ é a matriz $\mathbf{ba} = \mathbf{c} = [c_{ij}] \in M(m \times p)$, cujo ij -ésimo elemento

$$c_{ij} = b_{i1}a_{1j} + b_{i2}a_{2j} + \dots + b_{in}a_{nj} = \sum_{k=1}^n b_{ik}a_{kj}$$

é o produto interno do i -ésimo vetor-linha de $\mathbf{b}$ pelo j -ésimo vetor-coluna de $\mathbf{a}$

Matriz identidade

Representamos a matriz identidade com o símbolo $\mathbf{I}$.

Tem-se $\mathbf{I} = [\delta_{ij}]$ onde δ é o símbolo de Kronecker, onde são válidas as seguintes características:

$$\delta_{ij} = \begin{cases} 1 & \text{se } i = j, \\ 0 & \text{se } i \neq j. \end{cases}$$

As propriedades abaixo são decorrentes de outras já apresentadas:

- ① $(\mathbf{cb})\mathbf{a} = \mathbf{c}(\mathbf{ba})$
- ② $\mathbf{c}(\mathbf{a} + \mathbf{b}) = \mathbf{ca} + \mathbf{cb}$
- ③ $\mathbf{a} \cdot \mathbf{I} = \mathbf{a}$
- ④ $\mathbf{b}(\alpha\mathbf{a}) = \alpha(\mathbf{ba})$
- ⑤ Uma matriz $\mathbf{a}$ chama-se invertível quando é quadrada e existe uma matriz $\mathbf{a}^{-1}$, chamada inversa de $\mathbf{a}$, tal que $\mathbf{a}^{-1}\mathbf{a} = \mathbf{aa}^{-1} = \mathbf{I}$

Exemplos de operações

- Soma

$$\begin{bmatrix} 1 & -1 & \pi & 0 \\ 11 & 11 & -2 & 0 \\ 9 & 0 & 4 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 4 & 1 & 9 \\ -5 & 10 & 2 & 1 \\ -3 & 8 & 0 & -6 \end{bmatrix} = \begin{bmatrix} 1+1 & -1+4 & \pi+1 & 0+9 \\ 11+(-5) & 11+10 & -2+2 & 0+1 \\ 9+(-3) & 0+8 & 4+0 & 4+(-6) \end{bmatrix}$$
$$= \begin{bmatrix} 2 & 3 & \pi+1 & 9 \\ 6 & 21 & 0 & 11 \\ 6 & 8 & 4 & -2 \end{bmatrix}$$

- Multiplicação por escalar

$$5 \cdot \begin{bmatrix} 1 & 4 & 1 & 9 \\ -5 & 10 & 2 & 11 \\ -3 & 8 & 0 & -6 \end{bmatrix} = \begin{bmatrix} 5 \cdot 1 & 5 \cdot 4 & 5 \cdot 1 & 5 \cdot 9 \\ 5 \cdot -5 & 5 \cdot 10 & 5 \cdot 2 & 5 \cdot 11 \\ 5 \cdot -3 & 5 \cdot 8 & 5 \cdot 0 & 5 \cdot -6 \end{bmatrix} = \begin{bmatrix} 5 & 20 & 5 & 45 \\ -25 & 50 & 10 & 55 \\ -15 & 40 & 0 & -30 \end{bmatrix}$$

Exemplos de operações

- Multiplicação de matrizes

$$\mathbf{a} = \begin{bmatrix} 3 & 2 \\ 5 & -5 \end{bmatrix}$$

$$\mathbf{b} = \begin{bmatrix} 6 & 4 & -2 \\ 0 & 7 & -2 \end{bmatrix}$$

$$\mathbf{c} = \mathbf{a} \cdot \mathbf{b}$$

$$\mathbf{c} = \begin{bmatrix} (3 \cdot 6 + 2 \cdot 0) & (3 \cdot 4 + 2 \cdot 7) & (3 \cdot (-2) + 2 \cdot (-2)) \\ (5 \cdot 6 + (-5) \cdot 0) & (5 \cdot 4 + (-5) \cdot 7) & (5 \cdot (-2) + (-5) \cdot (-2)) \end{bmatrix}$$

$$\mathbf{c} = \begin{bmatrix} 18 & 26 & -10 \\ 30 & -15 & 0 \end{bmatrix}$$

Exemplos de operações

- Matriz inversa

Pela propriedade $\mathbf{a}\mathbf{a}^{-1} = \mathbf{I}$

$$\mathbf{a} = \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix}$$

$$\mathbf{a}^{-1} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$\begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix} \cdot \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 3a_{11} + 1a_{21} & 3a_{12} + 1a_{22} \\ 5a_{11} + 2a_{21} & 5a_{12} + 2a_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Resolvendo o sistema linear:

$$\mathbf{a}^{-1} = \begin{bmatrix} 2 & -1 \\ -5 & 3 \end{bmatrix}$$

Álgebra Linear

Aula 05: Produto interno

Kelvin Rafael Duarte Machado

setembro de 2025

Produto interno

Um *produto interno* em um espaço vetorial E é um funcional bilinear simétrico e positivo em E . Mais precisamente, é uma função $E \times E \rightarrow \mathbb{R}$, que associa a cada par de vetores $u, v \in E$ um número real $\langle u, v \rangle$, chamado o produto interno de u por v , de modo que sejam válidas as seguintes propriedades, para quaisquer $u, u', v, v' \in E$ e $\alpha \in \mathbb{R}$:

Bilinearidade: $\langle u + u', v \rangle = \langle u, v \rangle + \langle u', v \rangle$, $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$,
 $\langle u, v + v' \rangle = \langle u, v \rangle + \langle u, v' \rangle$, $\langle u, \alpha v \rangle = \alpha \langle u, v \rangle$

Comutatividade (simetria): $\langle u, v \rangle = \langle v, u \rangle$

Positividade: $\langle u, u \rangle > 0$ se $u \neq 0$

Lembrando que o produto interno de u por v é definido por:

$$\langle u, v \rangle = \alpha_1 \beta_1 + \cdots + \alpha_n \beta_n.$$

Norma

O número não negativo $|u| = \sqrt{\langle u, u \rangle}$ chama-se a *norma* ou o *comprimento* do vetor u .

Quando $|u| = 1$ diz que o vetor $u \in E$ é um vetor unitário.

Seja $E = C^0([a, b])$ o espaço vetorial cujos elementos são as funções contínuas $g, f : [a, b] \rightarrow \mathbb{R}$. Um produto interno em E pode ser definido pondo:

$$\langle f, g \rangle = \int_a^b f(x)g(x) dx$$

Neste caso, a norma da função é:

$$|f| = \sqrt{\int_a^b f(x)^2 dx}$$

Dados $u = (\alpha_1, \alpha_2)$ e $v = (\beta_1, \beta_2)$, os números:

$$|u| = \sqrt{\alpha_1^2 + \alpha_2^2}$$

e

$$|v| = \sqrt{\beta_1^2 + \beta_2^2}$$

medem os comprimentos das flechas que representam esses vetores. Supondo que $u \neq 0$, $v \neq 0$ e chamando de θ o ângulo formado por essas flechas, afirmamos que o produto interno $\langle u, v \rangle = \alpha_1\beta_1 + \alpha_2\beta_2$ é igual a:

$$|u||v| \cos \theta$$

Vetores ortogonais

Seja E um espaço vetorial com produto interno. Dois vetores $u, v \in E$ chamam-se ortogonais (ou perpendiculares) quando $\langle u, v \rangle = 0$. Escreve-se, então, $u \perp v$. Em particular, 0 é ortogonal a qualquer vetor de E . Um conjunto $X \subset E$ diz-se ortogonal quando dois vetores distintos quaisquer em X são ortogonais.

Se além disso, todos os vetores de X são unitários então X chama-se um conjunto ortonormal.

Portanto, o conjunto $X \subset E$ é ortonormal se, e somente se, dados $u, v \in X$ tem-se $\langle u, v \rangle = 0$ e $\langle u, v \rangle = 1$ se $v = u$. Uma base ortonormal é uma base de E que é um conjunto ortonormal.

Projeção ortogonal

Num espaço vetorial E com produto interno, seja u um vetor unitário. Dado qualquer $v \in E$, o vetor $\langle u, v \rangle \cdot u$ chama-se a projeção de ortogonal de v sobre o eixo que contém u .

$$pr_u(v) = \frac{\langle u, v \rangle}{\langle u, u \rangle} \cdot u$$

Desigualdade de Cauchy-Schwarz

Se $z = pr_u(v)$, tem-se $v = z + w$, com $w \perp z$. Pelo Teorema de Pitágora, $|v|^2 = |z|^2 + |w|^2$. Em particular vemos que $|z| \leq |v|$, isto é, o comprimento da projeção $pr_u(v)$ é menor do que ou igual ao comprimento de v .

Sendo a norma do vetor $pr_u(v)$ igual a $|\langle u, v \rangle|/|u|$. Segue-se então que, para quaisquer $u, v \in E$, tem-se $|\langle u, v \rangle|/|u| \leq |v|$, ou seja

$$|\langle u, v \rangle| \leq |u| \cdot |v|$$

que é chamada de desigualdade de Cauchy-Schwarz.

Processo de ortonormalização de Gram-Schmidt

Existem bases ortonormais em todo espaço vetorial de dimensão finita provido de um produto interno.

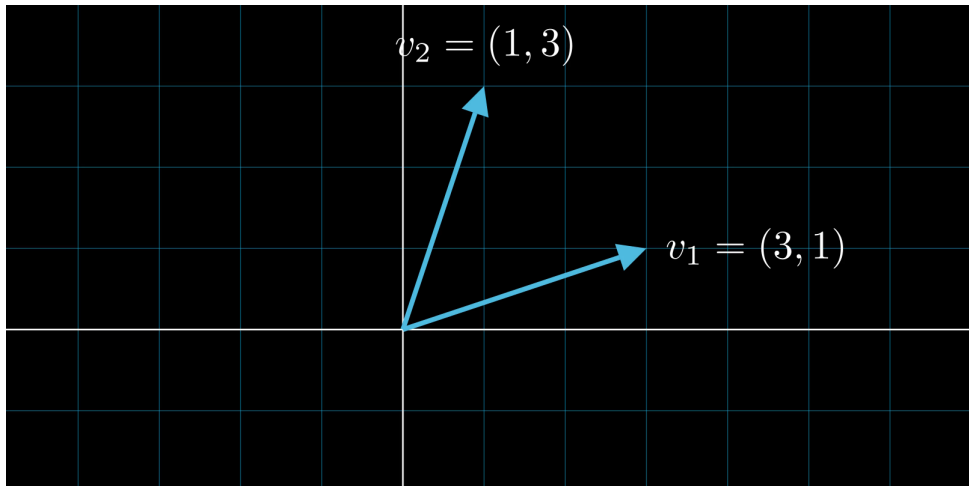
O processo de ortonormalização de Gram-Schmidt é um algoritmo que ensina a passar de uma base qualquer $v_1, \dots, v_n \subset E$ para uma base ortonormal $u_1, \dots, u_n \subset E$, com a importante propriedade de que, para $m = 1, \dots, n$, os vetores $u_1, \dots, u_m$ pertencem ao subespaço F_m , gerado por $v_1, \dots, v_m$.

Dada uma base $v_1, \dots, v_n \subset E$, obteremos primeiro uma base ortogonal $w_1, \dots, w_n \subset E$ e depois poremos $u_1 = w_1/|w_1|, \dots, u_n = w_n/|w_n|$ para chegar à base ortonormalizada $u_1, \dots, u_n \subset E$.

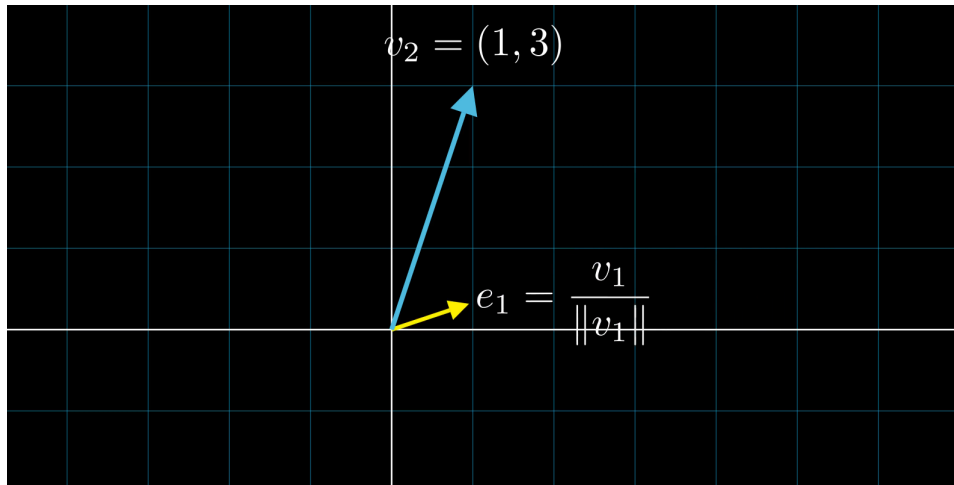
Começamos o processo tomando $w_1 = v_1$ e prosseguimos por indução.

$$w_{m+1} = v_{m+1} - \sum_{i=1}^m \frac{\langle w_i, v_{m+1} \rangle}{\langle w_i, w_i \rangle} w_i$$
$$u_n = \frac{w_n}{|w_n|}$$

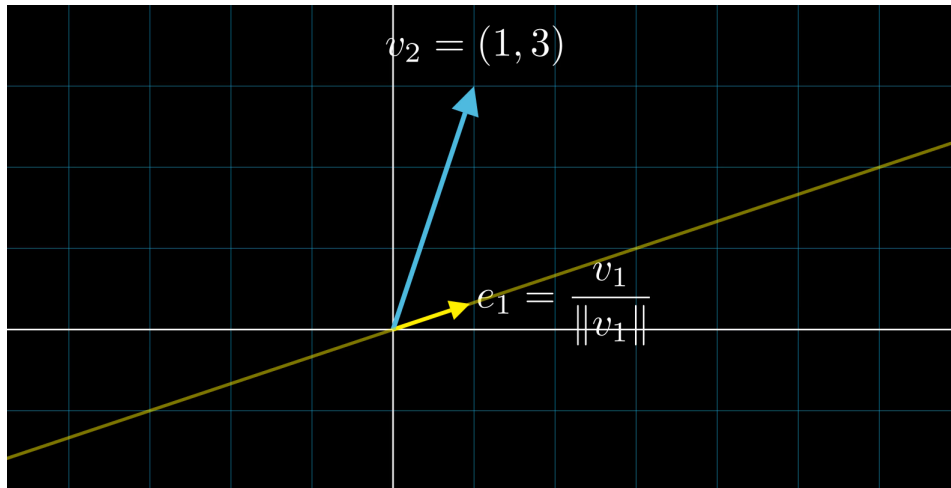
Processo de ortonormalização de Gram-Schmidt



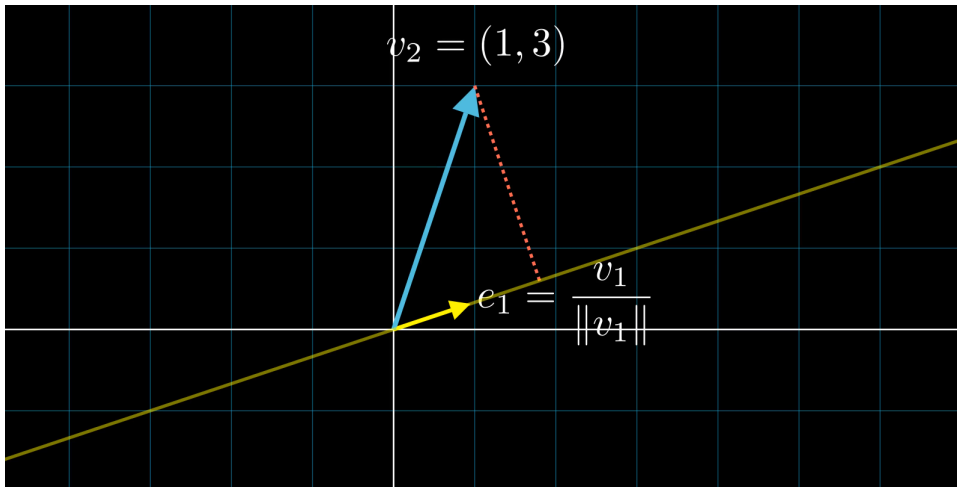
Processo de ortonormalização de Gram-Schmidt



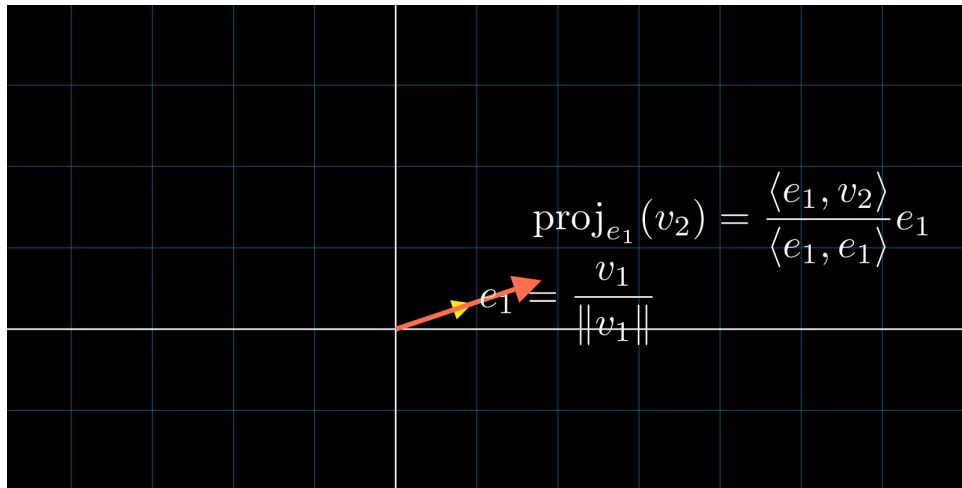
Processo de ortonormalização de Gram-Schmidt



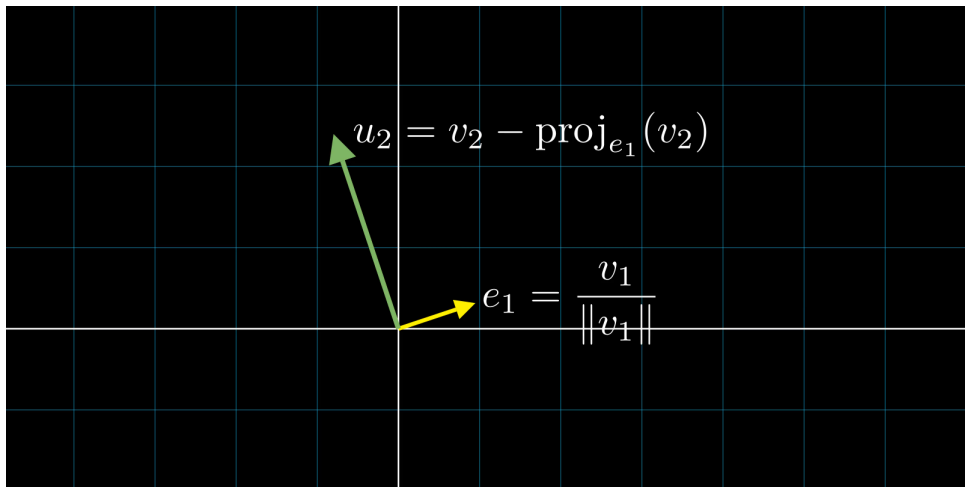
Processo de ortonormalização de Gram-Schmidt



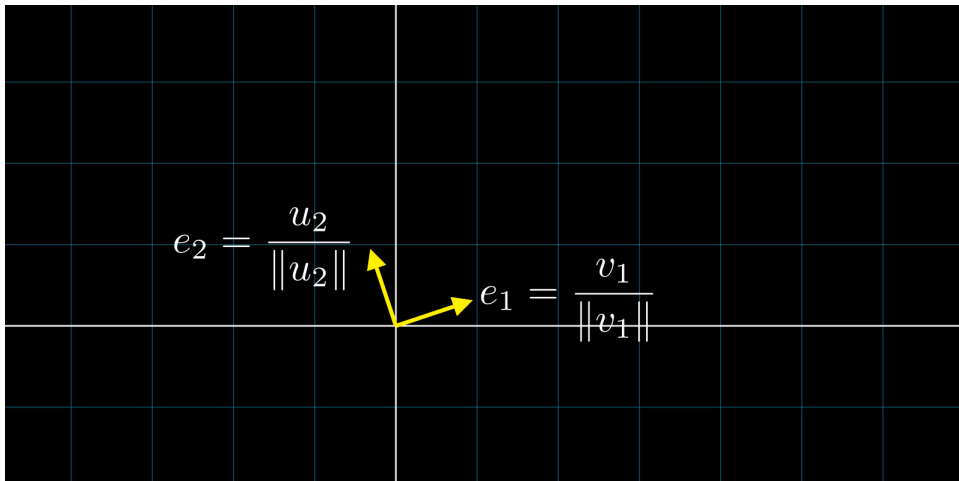
Processo de ortonormalização de Gram-Schmidt



Processo de ortonormalização de Gram-Schmidt



Processo de ortonormalização de Gram-Schmidt



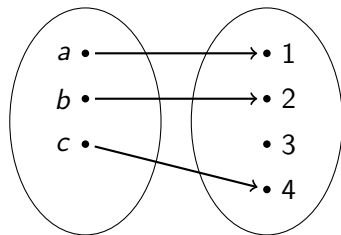
Álgebra Linear

Aula 06: A adjunta

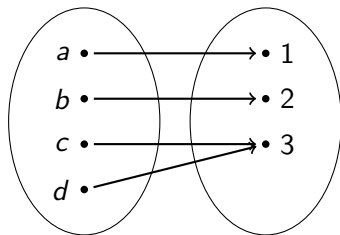
Kelvin Rafael Duarte Machado

setembro de 2025

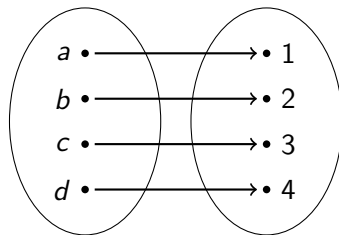
Injetividade, Sobrejetividade e Bijetividade



Função injetiva



Função sobrejetiva



Função bijetiva

A adjunta de A deve ser uma transformação linear $A^* : F \rightarrow E$ tal que, para $v \in E$ e $w \in F$ quaisquer se tenha:

$$\langle Av, w \rangle = \langle v, Aw \rangle$$

Matriz da adjunta

Theorem

Sejam $\mathcal{U} = u_1, \dots, u_n \subset E$ e $\mathcal{V} = v_1, \dots, v_m \subset F$ bases ortonormais. Se $\mathbf{a} = [a_{ij}]$ é a matriz de transformação linear $A: E \rightarrow F$ nas bases $\mathcal{U}, \mathcal{V}$, então a matriz da adjunta $A^: F \rightarrow E$ nas bases $\mathcal{U}, \mathcal{V}$ é a transposta $\mathbf{a}^T = [a_{ji}] \in \mathcal{M}(n \times m)$ de $\mathbf{a}$*

Corollary

Uma transformação linear A e sua adjunta A^ têm o mesmo posto.*

$$I^* = I$$

$$(A + B)^* = A^* + B^*$$

$$(\alpha A)^* = \alpha A^*$$

$$(BA)^* = A^* B^*$$

$$A^{**} = A$$

$$\mathbf{I}_n^T = \mathbf{I}_n$$

$$(\mathbf{a} + \mathbf{b})^T = \mathbf{a}^T + \mathbf{b}^T$$

$$(\alpha \mathbf{a})^T = \alpha \mathbf{a}^T$$

$$(\mathbf{ba})^T = \mathbf{a}^T \mathbf{b}^T$$

$$(\mathbf{a}^T)^T = \mathbf{a}$$

Complemento ortogonal

Seja E um espaço vetorial com produto interno. O complemento ortogonal de um conjunto não-vazio $X \subset E$ é o conjunto $X^\perp$ formado pelos vetores $v \in E$ que são ortogonais a todos os vetores $x \in X$. Portanto

$$v \in X^\perp \Leftrightarrow \langle v, x \rangle = 0$$

para todo $x \in X$

Seja E um espaço vetorial de dimensão finita, munido de produto interno. Para todo subespaço vetorial $F \subset E$ tem-se a decomposição em soma direta $E = F \oplus F^\perp$. Logo, $\dim F + \dim F^\perp = \dim E$. Para todo subespaço vetorial $F \subset E$, tem-se $(F^\perp)^\perp = F$.

Teorema fundamental

Theorem

Dada a transformação linear $A : E \rightarrow F$, entre espaços vetoriais de dimensão finita munidos de produto interno, tem-se

$$\mathcal{N}(A^*) = \mathcal{I}m(A)^\perp$$

$$\mathcal{I}m(A^*) = \mathcal{N}(A)^\perp$$

$$\mathcal{N}(A) = \mathcal{I}m(A^*)^\perp$$

$$\mathcal{I}m(A) = \mathcal{N}(A^*)^\perp$$

Corollary

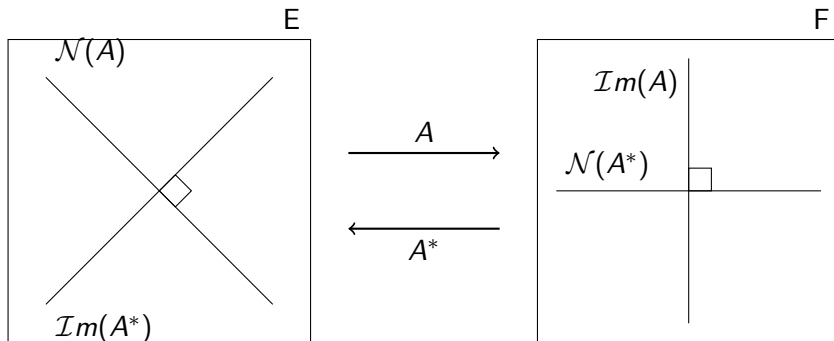
A fim de que o sistema de m equações lineares com n incógnitas

$$\sum_{j=1}^n a_{ij}x_j = b_i, \quad (i = 1, \dots, m)$$

possua solução é necessário e suficiente que o vetor $b = (b_1, \dots, b_m) \in \mathbb{R}^m$ seja perpendicular a toda solução $y = (y_1, \dots, y_m)$ do sistema homogêneo transposto

$$\sum_{j=1}^m a_{ji}y_j = 0, \quad (i = 1, \dots, n)$$

Interpretação geométrica



Álgebra Linear

Aula 07: Subespaços invariantes

Kelvin Rafael Duarte Machado

novembro de 2025

Subespaço invariante

Diz-se que um subespaço vetorial $F \subset E$ é invariante pelo operador linear $A : E \rightarrow E$ quando $A(F) \subset F$, isto é, quando a imagem Av de qualquer vetor $v \in F$ é ainda um vetor em F .

Um vetor $v \neq 0$ em E chama-se um *autovetor* do operador $A : E \rightarrow E$ quando existe $\lambda \in \mathbb{R}$ tal que

$$Av = \lambda v$$

O número $\lambda \in \mathbb{R}$, por sua vez, chama-se um *autovalor* do operador A quando existe um vetor não-nulo $v \in E$ tal que $Av = \lambda v$. Diz-se então que o autovalor λ corresponde, ou pertence, ao autovetor v e, vice-versa, que o autovetor v também corresponde, ou pertence, ao autovalor λ .

Autovetores e autovalores

Encontrar um autovetor ou autovalor do operador A é, portanto, o mesmo que achar um subespaço de dimensão 1 invariante por A .

Analogamente, diz que o número real λ é um autovalor da matriz $\mathbf{a} \in \mathcal{M}(n \times n)$ quando λ é um autovalor do operador $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$, cuja matriz na base canônica é $\mathbf{a}$. Isto significa que existe um vetor $v \neq 0$ em $\mathbb{R}^n$ tal que $Av = \lambda v$ ou, o que é o mesmo, uma matriz não-nula $\mathbf{x} \in \mathcal{M}(n \times 1)$ tal que $\mathbf{a}\mathbf{x} = \lambda\mathbf{x}$

Um operador linear num espaço vetorial de dimensão n admite no máximo n autovalores distintos.

Theorem

A autovalores diferentes do mesmo operador correspondem autovetores linearmente independentes.

Corollary

Seja $\dim E = n$. Se um operador linear $A : E \rightarrow E$ possui n autovalores diferentes então existe uma base $\{v_1, \dots, v_n\} \subset E$ em relação à qual a matriz de A é diagonal (isto é, tem a forma a_{ij} com $a_{ij} = 0$ se $i \neq j$.)

Logo, se $Av_1 = \lambda_1 v_1, \dots, Av_n = \lambda_n v_n$, a matriz de A nesta base é:

$$\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

na qual os termos que não aparecem são iguais a zero.

Polinômio característico

A igualdade $Av = \lambda v$ equivale a $(A - \lambda I)v = 0$, logo v é um autovetor do operador $A : E \rightarrow E$ se, e somente se, é um elemento não-nulo do núcleo ($\mathcal{N}(A - \lambda I)$). A fim de que λ seja um autovalor de A é necessário e suficiente que o operador $A - \lambda I : E \rightarrow E$ não possua inverso.

Polinômio característico

Dado o caso particular que ocorre quando $\dim E = 2$, se $u, v \subset E$ é uma base, então os vetores $\alpha u + \beta v$ e $\gamma u + \delta v$ são linearmente dependentes se, e somente se, $\alpha\delta - \beta\gamma = 0$. Dados o operador $A : E \rightarrow E$ e a base $u, v \subset E$, sejam $Au = au + cv$ e $Av = bu + dv$. A matriz do operador A na base u, v é

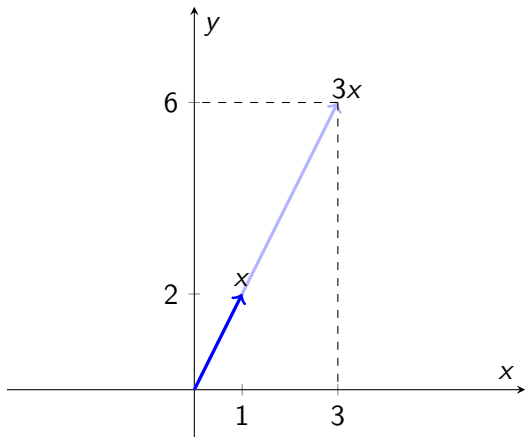
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Então $(A - \lambda I)u = (a - \lambda)u + cv$ e $(A - \lambda I)v = bu + (d - \lambda)v$. A fim de que $A - \lambda I$ não seja invertível é necessário e suficiente que os vetores $(A - \lambda)u$ e $(A - \lambda)v$ sejam linearmente dependentes, ou seja $(a - \lambda)(d - \lambda) - bc = 0$ ou ainda, que λ seja a raiz do polinômio

$$p(\lambda) = \lambda^2 - (a + d)\lambda + ad - bc$$

chamado de *polinômio característico*.

Exemplo



O vetor $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ é um autovetor de

$$A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$$

Associado ao autovalor $\lambda = 3$, pois

$$A \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix} = 3\mathbf{x}$$

Álgebra Linear

Aula 08: Operadores auto-adjuntos, ortogonais e normais

Kelvin Rafael Duarte Machado

novembro de 2025

Operador auto-adjunto

Um operador linear $A : E \rightarrow E$, num espaço vetorial munido de produto interno, chama-se *auto-adjunto* quando $A = A^*$, ou seja, quando $\langle Au, v \rangle = \langle u, Av \rangle$ para quaisquer $u, v \in E$.

Se $A, B : E \rightarrow E$ são operadores auto-adjuntos e $\alpha \in \mathbb{R}$ então $(A + B)^* = A^* + B^* = A + B$ e $(\alpha A)^* = \alpha A^* = \alpha A$, logo $A + B$ e αA são auto-adjuntos.

O produto AB dos operadores auto-adjuntos A, B é auto-adjunto se, e somente se, A e B comutam, isto é, $AB = BA$. Com efeito, sendo A e B auto-adjuntos, temos

$$(AB)^* = B^* A^* = BA$$

Matrizes simétricas

Uma matriz quadrada $\mathbf{a} = [a_{ij}]$ se diz simétrica quando é igual a sua transposta $\mathbf{a}^T$, isto é, quando $a_{ij} = a_{ji}$ para todo i e todo j .

Theorem

$A : E \rightarrow E$ é auto-adjunto se, e somente se, sua matriz $\mathbf{a} = [a_{ij}]$ relativamente a uma (e portanto a qualquer) base ortonormal $\mathcal{U} = \{u_1, \dots, u_n\} \subset E$ é uma matriz simétrica.

Operadores auto-adjuntos

Theorem

Seja $A : E \rightarrow E$ um operador auto-adjunto. Se o subespaço $F \subset E$ é invariante por A , seu complemento ortogonal $F^\perp$ também é.

Theorem

Se o subespaço $F \subset E$ é invariante pelo operador linear $A : E \rightarrow E$ então seu complemento ortogonal $F^\perp$ é invariante pelo operador adjunto $A^ : E \rightarrow E$.*

Theorem

Se $\lambda_1, \dots, \lambda_m$ são autovalores dois a dois diferentes do operador auto-adjunto $A : E \rightarrow E$, os autovetores correspondentes $v_1, \dots, v_m$ são dois a dois ortogonais.

Teorema espectral

Um problema importante sobre operadores num espaço vetorial de dimensão finita é o de encontrar uma base em relação à qual a matriz desse operador seja a mais simples possível. Se $A : E \rightarrow E$ é um operador auto-adjunto num espaço vetorial finito com produto interno, existe uma base ortonormal em E , relativamente à qual a matriz de A é uma matriz diagonal $\mathbf{a} = [a_{ij}]$, isto é, $a_{ij} = 0$ se $i \neq j$.

Theorem

Para todo operador auto-adjunto $A : E \rightarrow E$, num espaço vetorial de dimensão finita munido de produto interno, existe uma base ortonormal $u_1, \dots, u_n \subset E$ formado por autovetores de A .

Teorema espectral

Theorem

Um operador auto-adjunto $A : E \rightarrow E$ é não negativo se, e somente se, seus autovalores são todos ≥ 0 . A é positivo se, e somente se, todos os seus autovalores são números positivos.

Corollary

Seja $A \geq 0$. Se, para um certo $v \in E$, vale $\langle Av, v \rangle = 0$ então $Av = 0$.

Corollary

Um operador é positivo se, e somente se, é não negativo e invertível.

Teorema dos valores singulares

Um operador $X : E \rightarrow E$ chama-se *raiz quadrada* do operador $A : E \rightarrow E$ quando $X^2 = A$.

Se λ é um autovalor do operador $A : E \rightarrow E$ o conjunto

$$E_\lambda = \{v \in E; Av = \lambda v\}$$

é um subespaço vetorial de E , invariante por A , chamado auto-subespaço. Restrito a E_λ , o operador A é simplesmente a multiplicação por λ .

Quando A é auto-adjunto e $\lambda_1, \dots, \lambda_r$ são seus autovalores distintos então, para $i \neq j$, todo vetor em E_{λ_i} é ortogonal a qualquer vetor em E_{λ_j} .

Theorem

Todo operador não-negativo $A : E \rightarrow E$, num espaço vetorial de dimensão finita munido de produto interno, possui uma única raiz quadrada não-negativa, a qual é positiva se, e somente se, A é positivo.

Teorema dos valores singulares

Sejam $\lambda_1 \geq 0, \dots, \lambda_r \geq 0$ os autovalores distintos de A . Todo vetor $v \in E$ se escreve, de modo único, como $v = v_1 + \dots + v_r$, onde $v_1 \in E_{\lambda_1}, \dots, v_r \in E_{\lambda_r}$, logo $Av = \lambda_1 v_1, \dots, \lambda_r v_r$. Definimos o operador linear $B : E \rightarrow E$ pondo

$$Bv = \sqrt{\lambda_1} v_1 + \dots + \sqrt{\lambda_r} v_r$$

Se $w = w_1 + \dots + w_r$ com $w_1 \in E_{\lambda_1}, \dots, w_r \in E_{\lambda_r}$ então $\langle v_i, w_j \rangle = 0$ para $i \neq j$, portanto

$$\langle Bv, w \rangle = \sum_{i=1}^r \sqrt{\lambda_i} \langle v_i, w_i \rangle = \langle v, Bw \rangle \quad e \quad \langle Bv, v \rangle = \sum_{i=1}^r \sqrt{\lambda_i} |v_i|^2.$$

Logo, B é um operador não-negativo. Além disso, $B^2 v = Av$ para todo $v \in E$, logo, B é uma raiz quadrada de A .

Teorema dos valores singulares

Theorem

Seja $A : E \rightarrow F$ uma transformação linear de posto r entre espaços de dimensão finita com produto interno. Existem bases ortonormais $\{u_1, \dots, u_n\} \subset E$ e $\{v_1, \dots, v_m\} \subset F$ tais que $Au_i = \sigma_i v_i$, $A^ v_i = \sigma_i u_i$ com $\sigma_i > 0$ para $i = 1, \dots, r$ e $Au_i = 0$, $A^* v_i = 0$ se $i > r$.*

$$A^* v_i = \frac{1}{\sigma_i} A^* A u_i = \frac{1}{\sigma_i} \sigma_i^2 u_i = \sigma_i u_i$$

Os números positivos $\sigma_1, \dots, \sigma_r$ chamam-se os *valores singulares* da transformação linear $A : E \rightarrow F$, de posto r .

Teorema dos Valores Singulares

Podemos observar que $v_1, \dots, v_m \subset F$ é uma base ortonormal de autovetores para o operador $AA^* : F \rightarrow F$, pois

$$(AA^*)v_i = A(\sigma_i u_i) = \sigma_i A u_i = \sigma_i^2 v_i$$

Analogamente, $\{u_1, \dots, u_r\}$ é uma base ortonormal de autovetores do operador $A^*AE \rightarrow E$.

Operador ortogonal

Uma matriz $\mathbf{u} \in \mathcal{M}(m \times n)$ cujas n colunas forma um conjunto ortonormal em $\mathbb{R}^m$ chama-se uma matriz ortogonal.

Se $u_1, \dots, u_n \in \mathbb{R}^m$ são as colunas da matriz $\mathbf{u} = [a_{ij}] \in \mathcal{M}(m \times n)$, a condição para que $\mathbf{u}$ seja ortogonal é que $\langle u_i, u_j \rangle = 0$ se $i \neq j$ e $\langle u_i, u_i \rangle = 1$, onde $i, j = 1, \dots, n$.

Noutras palavras, deve-se ter

$$\sum_{k=1}^m a_{ki} a_{kj} = \delta_{ij}$$

Onde δ_{ij} é o conhecido como "delta de Kronecker", que vale 0 se $i \neq j$ e 1 se $i = j$.
A igualdade acima significa que o produto $\mathbf{u}^T \mathbf{u} = \mathbf{I}_n$.

Operador ortogonal

Se $\mathbf{u} \in \mathbb{R}^{m \times n}$ é ortogonal, então seu posto é n (logo $m \geq n$), pois suas n colunas são vetores linearmente independentes no espaço $\mathbb{R}^m$.

Quando $m = n$ e $\mathbf{u} \in \mathbb{R}^{n \times n}$ é uma matriz quadrada ortogonal, a igualdade

$$\mathbf{u}^T \mathbf{u} = \mathbf{I}_n$$

implica

$$\mathbf{u} \mathbf{u}^T = \mathbf{I}_n \quad \Rightarrow \quad \mathbf{u}^T = \mathbf{u}^{-1}.$$

Assim, matrizes quadradas ortogonais são aquelas cuja transposta é igual à inversa.

Operador ortogonal

Exemplo (rotação no plano):

$$\mathbf{u} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

Então:

$$\mathbf{u}^T = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$$

e, multiplicando:

$$\mathbf{u}^T \mathbf{u} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}$$

Matriz de passagem ortogonal

Sejam $\mathcal{U} = \{u_1, \dots, u_n\} \in E$ e $\mathcal{U}' = \{u'_1, \dots, u'_n\} \in E$ bases ortonormais. A matriz de passagem $\mathbf{p} = [p_{ij}]$ de $\mathcal{U}$ para $\mathcal{U}'$ é uma matriz ortogonal. Com efeito, para quaisquer $i, j = 1, \dots, n$, temos

$$u'_i = \sum_{k=1}^n p_{ki} u_k \quad e \quad u'_j = \sum_{k=1}^n p_{kj} u_k$$

logo

$$\delta_{ij} = \langle u'_i, u'_j \rangle = \sum_{k=1}^n p_{ki} p_{kj}$$

portanto $\mathbf{p}^T \cdot \mathbf{p} = \mathbf{I}_n$.

Operador ortogonal

As seguintes afirmações a respeito de uma transformação linear $A : E \rightarrow F$, entre espaços vetoriais de dimensão finita providos de produto interno, são equivalentes

- A preserva norma: $|Av| = |v|$ para todos $v \in E$;
- A preserva distância: $|Au - Av| = |u - v|$ para quaisquer $u, v \in E$;
- A preserva produto interno: $\langle Au, Av \rangle = \langle u, v \rangle$ para quaisquer $u, v \in E$;
- A matriz A relativa a qualquer par de bases ortonormais $\mathcal{U} \subset E$, $\mathcal{V} \subset F$ é uma matriz ortogonal;
- A matriz A relativa a um certo par de bases ortonormais $\mathcal{U} \subset E$, $\mathcal{V} \subset F$ é uma matriz ortogonal;
- A transforma uma certa base ortonormal $\mathcal{U} \subset E$ num conjunto ortonormal $\mathcal{X} \subset F$;
- A transforma toda base ortonormal $\mathcal{W} \subset E$ num conjunto ortonormal $\mathcal{Z} \subset F$.

Operador ortogonal

Seja $A : E \rightarrow E$ um operador ortogonal em um espaço vetorial de dimensão finita com produto interno.

Pela definição, A preserva o produto interno:

$$\langle Au, Av \rangle = \langle u, v \rangle, \quad \forall u, v \in E.$$

Isto implica que:

- Todos os autovalores de A têm módulo 1;
- Autovetores associados a autovalores distintos são ortogonais;
- Autovalores reais podem ser $+1$ ou -1 , enquanto os complexos aparecem em pares conjugados.

Consequentemente, existe uma base ortonormal de E que diagonaliza os autovalores reais e transforma os pares complexos em blocos de rotação 2×2 , levando à forma canônica demonstrada a seguir.

Operador ortogonal

Theorem

Seja $A : E \rightarrow E$ um operador ortogonal num espaço vetorial de dimensão finita munido de produto interno. Existe uma base ortonormal de E relativamente à qual a matriz de A tem a forma

$$\begin{bmatrix} I_p & & & & \\ & \ddots & & & \\ & & -I_q & & \\ & & & R(\alpha_1) & \\ & & & & \ddots \\ & & & & & R(\alpha_k) \end{bmatrix}$$

- I_p : autovalores $+1$ (direções fixas) e $-I_q$: autovalores -1 (reflexões);
- $R(\alpha_i)$: rotação de ângulo α_i em subespaço 2D $R(\alpha_i) = \begin{bmatrix} \cos \alpha_i & -\sin \alpha_i \\ \sin \alpha_i & \cos \alpha_i \end{bmatrix}$;

Theorem

Seja E um espaço vetorial de dimensão finita munido de produto interno. Todo operador linear $A : E \rightarrow E$ admite uma decomposição da forma $A = PU$, onde $U : E \rightarrow E$ é ortogonal e $P : E \rightarrow E$ é não negativo.

De acordo com o teorema, existem bases ortonormais $\{u_1, \dots, u_n\} \subset E$, $\{v_1, \dots, v_n\} \subset E$ e números $\lambda_i \geq 0$ tais que $Au_i = \lambda_i v_i$ para $i = 1, \dots, n$. Definimos os operadores $P, U : E \rightarrow E$ impondo que $Pv_i = \lambda_i v_i$ e $Uu_i = v_i$. Evidentemente, P é auto-adjunto, ≥ 0 , U é ortogonal e $PU = A$.

Decomposição polar

A expressão $A = PU$ chama-se uma decomposição polar do operador A , por analogia a forma polar $z = re^{i\theta}$ de um número complexo.

De $A = PU$ tiramos $A^* = U^*P$ e, multiplicando membro a membro estas igualdades, vem $AA^* = PUU^*P$, portanto P é a raiz quadrada positiva do operador positivo AA^* . Multiplicando a igualdade $A = PU$ por P^{-1} à esquerda, vem $U = P^{-1}A$. Isto mostra que no caso em que A é invertível, $P = \sqrt{AA^*}$ e $U = P^{-1}A$ são univocamente determinados a partir da igualdade $A = PU$.

Operador normal

Um operador linear $A : E \rightarrow E$ chama-se normal quando comuta com seu adjunto, isto é, quando $AA^* = A^*A$.

Uma matriz quadrada $\mathbf{a}$ diz-se normal quando comuta com sua transposta, isto é, quando $\mathbf{a}\mathbf{a}^T = \mathbf{a}^T\mathbf{a}$. Portanto um operador é normal se, e somente se, sua matriz relativamente a uma base ortonormal é uma matriz normal.

Assim, os operadores auto-adjuntos e os ortogonais são normais. Bem como as matrizes simétricas e as ortogonais quadradas são normais.

Operador normal

Um operador linear $A : E \rightarrow E$ chama-se anti-simétrico quando $A^* = -A$, ou seja, $\langle Au, v \rangle = -\langle u, Av \rangle$ para $u, v \in E$ quaisquer. Para que A seja anti-simétrico é necessário e suficiente que sua matriz $[a_{ij}]$ em relação a uma base ortonormal de E seja anti-simétrica, isto é, $a_{ij} = -a_{ji}$ para quaisquer $i, j = 1, 2, \dots, n$ em particular, $a_{ii} = 0$. Num espaço de dimensão 1, todo operador anti-simétrico é igual a zero; logo o único autovalor possível de um operador anti-simétrico é 0.

Operador normal

Theorem

Seja $A : E \rightarrow E$ um operador normal. Então existe uma base ortonormal de E na qual a matriz de A tem a forma

$$\begin{bmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_r & & \\ & & & R_1 & \\ & & & & \ddots \\ & & & & & R_k \end{bmatrix}$$

,
onde

- $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ são autovalores reais;
- cada bloco $R_i = \begin{bmatrix} \alpha_i & -\beta_i \\ \beta_i & \alpha_i \end{bmatrix}$ representa um par de autovalores complexos $\alpha_i \pm i\beta_i$, $\beta_i \neq 0$.

Álgebra Linear

Aula 09: Formas bilineares e quadráticas

Kelvin Rafael Duarte Machado

novembro de 2025

Sejam E, F espaços vetoriais. Uma *forma bilinear* $b : E \times F \rightarrow \mathbb{R}$ é uma função $b(u, v)$, linear em cada uma das duas variáveis $u \in E, v \in F$. Mais precisamente, para quaisquer $u, u' \in E, v, v' \in F$ e $\alpha \in \mathbb{R}$ devem valer:

$$\begin{aligned} b(u + u', v) &= b(u, v) + b(u', v), & b(\alpha u, v) &= \alpha b(u, v), \\ b(u, v + v') &= b(u, v) + b(u, v'), & b(u, \alpha v) &= \alpha b(u, v) \end{aligned}$$

Representação matricial de formas bilineares

Tomando bases $\mathcal{U} = \{u_1, \dots, u_m\} \subset E$ e $\mathcal{V} = \{v_1, \dots, v_n\} \subset F$, os números $b_{ij} = b(u_i, v_j)$ definem uma matriz $\mathbf{b} = [b_{ij}] \in \mathcal{M}(m \times n)$, chamada *matriz da forma bilinear b relativamente às bases $\mathcal{U}$ e $\mathcal{V}$* .

Conhecidos os valores $b_{ij} = b(u_i, v_j)$, os quais podem ser tomados arbitrariamente, a forma bilinear $b : E \times F \rightarrow \mathbb{R}$ fica inteiramente determinada pois, para $u = \sum x_i u_i \in E$ e $v = \sum y_j v_j \in F$ quaisquer, tem-se

$$b(u, v) = \sum_{i,j} x_i y_j b(u_i, v_j) = \sum_{i,j} b_{ij} x_i y_j \quad 1 \leq i \leq m, 1 \leq j \leq n$$

Theorem

As matrizes $\mathbf{b}$ e $\mathbf{b}'$, da forma bilinear b nas bases $\mathcal{U}$, $\mathcal{V}$ e $\mathcal{U}'$, $\mathcal{V}'$ respectivamente, se relacionam pela igualdade $\mathbf{b}' = \mathbf{p}^T \mathbf{b} \mathbf{q}$, onde $\mathbf{p}$ é a matriz de passagem de $\mathcal{U}$ para $\mathcal{U}'$ e $\mathbf{q}$ é a matriz de passagem de $\mathcal{V}$ para $\mathcal{V}'$.

Formas simétricas e antissimétricas

Uma forma bilinear $b : E \times E \rightarrow \mathbb{R}$ chama-se simétrica quando $b(u, v) = b(v, u)$ para quaisquer $u, v \in E$. Para que b seja simétrica é suficiente que sua matriz em relação a uma base $\mathcal{U} \subset E$ seja simétrica e é necessário que sua matriz em relação a qualquer base de E seja simétrica.

Analogamente, uma forma bilinear $b : E \times E \rightarrow \mathbb{R}$ chama-se antissimétrica quando $b(u, v) = -b(v, u)$ para quaisquer $u, v \in E$. Para que b seja antissimétrica é suficiente que se tenha $b(u_i, u_j) = -b(u_j, u_i)$ ou seja, $b_{ij} = -b_{ji}$ numa base $\{u_1, \dots, u_m\} \subset E$. Uma forma bilinear b , ao mesmo tempo simétrica e antissimétrica, deve ser nula. Este fato, juntamente com a igualdade

$$b(u, v) = \frac{1}{2}[b(u, v) + b(v, u)] + \frac{1}{2}[b(u, v) - b(v, u)]$$

mostra que o espaço $\mathcal{B}(E \times E)$ das formas bilineares $b : E \times E \rightarrow \mathbb{R}$ é a soma direta de dois subespaços, um deles formado pelas formas simétricas e outro pelas formas antissimétricas.

Produtos tensorial, simétrico e antissimétrico

Dados os funcionais lineares $f : E \rightarrow \mathbb{R}$, $g : F \rightarrow \mathbb{R}$, a função $b : E \times F \rightarrow \mathbb{R}$, definida por $b(u, v) = f(u)g(v)$, é uma forma bilinear, chamada o *produto tensorial* de f e g . Se E e F são dotados de produto interno, fixados $u_0 \in E$ e $v_0 \in F$, a função $b : E \times F \rightarrow \mathbb{R}$, onde $b(u, v) = \langle u, u_0 \rangle \cdot \langle v, v_0 \rangle$, é uma forma bilinear. No caso particular em que $E = F$, dados $f, g : E \rightarrow \mathbb{R}$, funcionais lineares, então as igualdades

$$(f \cdot g)(u, v) = f(u)g(v) + f(v)g(u),$$

$$(f \wedge g)(u, v) = f(u)g(v) - f(v)g(u).$$

definem formas bilineares $(f \cdot g), (f \wedge g) : E \times E \rightarrow \mathbb{R}$, a primeira simétrica e a segunda antissimétrica.

Theorem

Seja E um espaço vetorial de dimensão finita provido de produto interno. Para cada forma bilinear $b : E \times E \rightarrow \mathbb{R}$ existe um único operador linear $B : E \rightarrow E$ tal que

$$\langle u, Bv \rangle = b(u, v) \quad \text{para } u, v \in E \text{ quaisquer.}$$

A correspondência $b \mapsto B$, assim definida, é um isomorfismo entre os espaços vetoriais $\mathcal{B}(E \times E)$ e $\mathcal{L}(E)$. A forma bilinear b é simétrica (respect. antissimétrica) se, e somente se, o operador B é auto-adjunto (respect. antissimétrico).

Forma quadrática

Uma função $\varphi : E \rightarrow \mathbb{R}$ chama-se uma *forma quadrática* quando existe uma forma bilinear $b : E \times E \rightarrow \mathbb{R}$ tal que $\varphi(v) = b(v, v)$ para todo $v \in E$.

Se, em vez da forma bilinear b tomarmos a forma bilinear simétrica

$$b'(u, v) = \frac{1}{2}[b(u, v) + b(v, u)]$$

teremos ainda

$$\varphi(v) = b(v, v) = \frac{1}{2}[b(v, v) + b(v, v)] = b'(v, v)$$

Diagonalização de formas quadráticas

Theorem

Seja E um espaço vetorial de dimensão finita munido de produto interno. Dada uma forma bilinear simétrica $b : E \times E \rightarrow \mathbb{R}$, existe uma base ortonormal $\mathcal{U} = \{u_1, \dots, u_m\} \subset E$ tal que $b(u_i, u_j) = 0$ se $i \neq j$.

Este teorema é a versão para formas bilineares do Teorema Espectral. A versão matricial correspondente é a seguinte: para toda matriz simétrica $\mathbf{b} = [b_{ij}] \in \mathcal{M}(m \times m)$, existe uma matriz ortogonal $\mathbf{p} \in \mathcal{M}(m \times m)$ tal que $\mathbf{p}^T \mathbf{b} \mathbf{p} = \mathbf{p}^{-1} \mathbf{b} \mathbf{p}$ é uma matriz diagonal. A diagonal de $\mathbf{d} = \mathbf{p}^T \mathbf{b} \mathbf{p}$ é formada pelos autovalores de $\mathbf{b}$, os quais se dizem também autovalores da forma bilinear b e da forma quadrática $\varphi(v) = b(v, v)$.

Coordenadas das formas quadráticas

Em termos de coordenadas dos vetores $u = \sum x_i u_i$ e $v = \sum y_j u_j$ relativamente à base $\mathcal{U}$ do teorema acima a forma bilinear b se exprime como

$$b(u, v) = \sum \lambda_i x_i y_i$$

Em particular, a forma quadrática $\varphi : E \rightarrow \mathbb{R}$, $\varphi(u) = b(u, u)$, para $v = \sum y_i u_i$ (expressão relativa à base $\mathcal{U}$) é usada por uma combinação linear de quadrados:

$$\varphi(v) = \sum \lambda_i y_i^2 = \lambda_1 y_1^2 + \cdots + \lambda_m y_m^2$$

Quádricas

Um subconjunto $\Sigma \subset \mathbb{R}^n$ chama-se uma *quádrica central* quando existe uma forma quadrática $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ tal que Σ é definido pela equação $\varphi(v) = 1$.

Se

$$\varphi(v) = \sum_{i,j} a_{ij}x_i x_j$$

para $v = (x_1, \dots, x_n)$, isto significa que Σ é o conjunto dos pontos $(x_1, \dots, x_n) \in \mathbb{R}^n$ tais que

$$\sum_{i,j=1}^n a_{ij}x_i x_j = 1$$

Segue-se imediatamente que, dada uma quádrica central $\Sigma \subset \mathbb{R}^n$, existe uma base ortonormal $\mathcal{U} = \{u_1, \dots, u_n\}$ em $\mathbb{R}^n$, relativamente à qual a equação de Σ é

$$\lambda_1 y_1^2 + \dots + \lambda_n y_n^2 = 1$$

onde $\lambda_1, \dots, \lambda_n$ são os autovalores da matriz simétrica $[a_{ij}]$.

Eixos principais de uma quádrlica

Noutras palavras, quando se exprime o vetor $v = y_1 u_1 + \cdots + y_n u_n$ como combinação linear dos elementos da base $\mathcal{U}$, tem-se $v \in \Sigma$ se, e somente se, suas coordenadas $y_1, \cdots, y_n$ satisfazem a equação

$$\Sigma \lambda_i y_i^2 = 1$$

As retas que contêm os vetores u_i da base $\mathcal{U}$ chamam-se *os eixos principais* da quádrlica central Σ .

Quando $n = 2$, Σ chama-se uma *cônica*. Portanto, a cônica definida em $\mathbb{R}^2$ pela equação

$$ax^2 + 2bxy + cy^2 = 1$$

pode, numa nova base ortonormal de $\mathbb{R}^2$, ser representada pela equação

$$\lambda s^2 + \mu t^2 = 1$$

onde λ, μ são os autovalores da matriz $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$.

Classificação das cônicas

Segundo os sinais desses autovalores, as seguintes possibilidades podem ocorrer:

- Se $\lambda\mu < 0$ (isto é, $ac < b^2$) tem-se uma *hipérbole*.
- Se $\lambda > 0$ e $\mu > 0$ (isto é, $a > 0$ e $ac > b^2$) tem-se uma *elipse*.
- Se $\lambda < 0$ e $\mu < 0$ (isto é, $a < 0$ e $ac > b^2$) tem-se o conjunto vazio.
- Se $\lambda\mu = 0$ (isto é, $ac = b^2$) tem-se um par de retas paralelas, ou o conjunto vazio, conforme seja $a + c > 0$ ou $a + c \leq 0$.

Quando $\lambda = \mu > 0$ tem-se uma circunferência.

Classificação das superfícies quádricas em 3D

Em dimensão 3, as superfícies quádricas centrais $\Sigma \subset \mathbb{R}^3$ têm, num sistema de coordenadas ortogonais convenientes, uma equação do tipo

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2 = 1$$

As possibilidades são as seguintes:

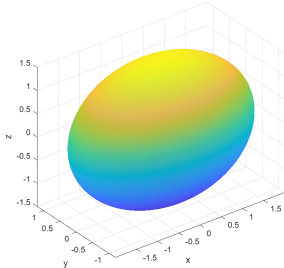
- Se $\lambda_1 > 0, \lambda_2 > 0, \lambda_3 > 0$, Σ é um *elipsóide*.
- Se $\lambda_1 > 0, \lambda_2 > 0, \lambda_3 < 0$, Σ é um *hiperbolóide de revolução*.
- Se $\lambda_1 > 0, \lambda_2 < 0, \lambda_3 < 0$, Σ é um *hiperbolóide de duas folhas*.
- Se $\lambda_1, \lambda_2, \lambda_3$ são ≤ 0 , Σ é o conjunto vazio.
- Se $\lambda_3 = 0$, $\Sigma = C \times \mathbb{R} = \{(v, t); v \in C, t \in \mathbb{R}^2\}$ é definido pela equação

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 = 1$$

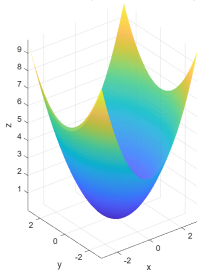
Neste caso, Σ é um cilindro de base C .

Exemplos

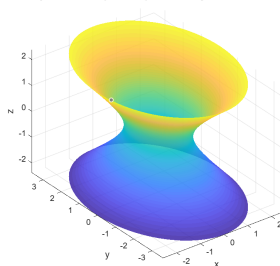
Elipsóide: $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$



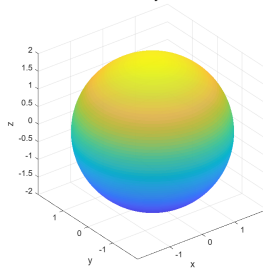
Paraboloide elíptico: $z = x^2/a^2 + y^2/b^2$



Hiperboloide (1 folha): $x^2/a^2 + y^2/b^2 - z^2/c^2 = 1$



Esfera: $x^2 + y^2 + z^2 = R^2$



Álgebra Linear

Aula 10: Determinante

Kelvin Rafael Duarte Machado

dezembro de 2025

Forma r-linear

Uma forma r-linear no espaço vetorial E é uma função

$$f : E \times \cdots \times E \rightarrow \mathbb{R}$$

definida no produto cartesiano $E \times \cdots \times E = E^r$, que é linear em cada uma das suas variáveis, isto é:

$$f(v_1, \cdots, v_i + v'_i, \cdots, v_r) = f(v_1, \cdots, v_i, \cdots, v_r) + f(v_1, \cdots, v'_i, \cdots, v_r),$$

$$f(v_1, \cdots, \alpha v_i, \cdots, v_r) = \alpha f(v_1, \cdots, v_i, \cdots, v_r)$$

Se f é r-linear e um dos vetores $v_1, \cdots, v_r$ é igual a zero então $f(v_1, \cdots, v_r) = 0$, pois

$$f(\cdots, 0, \cdots) = f(\cdots, 0 \cdot 0, \cdots) = 0 \cdot f(\cdots, 0, \cdots) = 0$$

Theorem

Uma forma r -linear $f : E \times \cdots \times E \rightarrow \mathbb{R}$ é alternada se, e somente se, é anti-simétrica, isto é,

$$f(v_1, \cdots, v_i, \cdots, v_j, \cdots, v_r) = -f(v_1, \cdots, v_j, \cdots, v_i, \cdots, v_r)$$

para quaisquer $v_1, \cdots, v_r \in E$.

Forma alternada - permutações

Corollary

Seja $f : E \times \cdots \times E \rightarrow \mathbb{R}$ r -linear alternada. Para toda permutação σ dos inteiros $1, 2, \dots, r$, e toda lista de vetores $v_1, \dots, v_r \in E$ tem-se

$$f(v_{\sigma(1)}, \dots, v_{\sigma(r)}) = \epsilon_{\sigma} f(v_1, \dots, v_r)$$

isto é,

$$f(v_{\sigma(1)}, \dots, v_{\sigma(r)}) = \pm f(v_1, \dots, v_r)$$

onde o sinal é $+$ se σ é uma permutação par e $-$ se σ é uma permutação ímpar.

A notação $A_r(E)$ indica o espaço vetorial das formas r -lineares alternadas $f : E \times \cdots \times E \rightarrow \mathbb{R}$.

Forma n -linear alternada

Theorem

Seja $\mathcal{U} = \{u_1, \dots, u_n\}$ uma base de E . Dado arbitrariamente um número real a , existe uma única forma n -linear alternada $f : E \times \dots \times E \rightarrow \mathbb{R}$ tal que $f(u_1, \dots, u_n) = a$

Corollary

Se $\dim E = n$ então $\dim A_n(E) = 1$

Corollary

Se $\{u_1, \dots, u_n\} \subset E$ é uma base e $0 \neq f \in A_n(E)$ então $f(u_1, \dots, u_n) \neq 0$

Determinante

Toda transformação linear $A : E \rightarrow F$ induz uma transformação $A^\# : A_n(F) \rightarrow A_n(E)$, a qual faz corresponder a cada forma n -linear alternada $f : F \times \cdots \times F \rightarrow \mathbb{R}$ a nova forma $A^\# f : E \times \cdots \times E \rightarrow \mathbb{R}$, definida por

$$(A^\# f)(v_1, \dots, v_n) = f(Av_1, \dots, Av_n)$$

onde $v_1, \dots, v_n \in E$.

Seja agora $A : E \rightarrow E$ um operador linear.

Como $\dim A_n(E) = 1$, o operador linear $A^\# : A_n(E) \rightarrow A_n(E)$ consiste na multiplicação por um número real, que chamaremos de *determinante* do operador $A : E \rightarrow E$ e indicaremos com a notação $\det A$. Assim, por definição, $A^\# f = \det A \cdot f$, isto é

$$f(Av_1, \dots, Av_n) = \det A \cdot f(v_1, \dots, v_n)$$

para toda $f : E \times \cdots \times E \rightarrow \mathbb{R}$ n -linear alternada e quaisquer $v_1, \dots, v_n \in E$.

Theorem

Se $A, B : E \rightarrow E$ são operadores lineares então $\det(BA) = \det B \cdot \det A$.

Theorem

O operador linear $A : E \rightarrow E$ é invertível se, e somente se, $\det A \neq 0$. No caso afirmativo, tem-se $\det(A^{-1}) = (\det A)^{-1}$.

Theorem

O sistema de equações lineares $Ax = b$, com $A \in \mathcal{L}(\mathbb{R}^n)$, possui uma única solução se, e somente se, $\det A \neq 0$.

Regra de Cramer

Seja $\mathbf{a} \in \mathcal{M}(m \times n)$ uma matriz invertível. Dado $\mathbf{b} \in \mathbb{R}^n$, indiquemos com o símbolo $\mathbf{a}[i; \mathbf{b}]$ a matriz obtida de $\mathbf{a}$ quando se substitui sua i -ésima coluna por $\mathbf{b}$. A solução do sistema linear $\mathbf{a}\mathbf{x} = \mathbf{b}$, de n equações a n incógnitas é o vetor $\mathbf{x} = (x_1, \dots, x_n)$ cujas coordenadas são

$$x_i = \frac{\det \mathbf{a}[i; \mathbf{b}]}{\det \mathbf{a}}, \quad i = 1, \dots, n.$$

Determinante da forma matricial

Tradicionalmente, o determinante de uma matriz $\mathbf{a} = [a_{ij}] \in \mathcal{M}(m \times n)$ é definido como a soma de $n!$ parcelas do tipo

$$\pm a_{1j_1}, a_{2j_2}, \dots, a_{mj_n}$$

Essas parcelas são produtos de n fatores que pertencem a linhas e colunas diferentes de $\mathbf{a}$. A ordem em que os fatores são dispostos é a ordem crescente dos índices $1, 2, \dots, n$ das linhas. Os segundos índices (das colunas) exibem uma permutação $\sigma = j_1, \dots, j_n$ dos inteiros $1, 2, \dots, n$. O sinal que precede cada parcela é $+$ ou $-$, conforme a permutação σ seja par ou ímpar respectivamente. Noutras palavras, a definição clássica do determinante é

$$\det \mathbf{a} = \sum_{\sigma} \epsilon_{\sigma} \cdot a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \dots \cdot a_{n\sigma(n)}$$

Determinante: exemplo

Seja a matriz **a**:

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

o determinante de **a** é dado por

$$\det \mathbf{a} = \begin{vmatrix} 2 & 1 & 0 \\ 1 & 3 & 4 \\ 0 & 2 & 1 \end{vmatrix} = -11$$

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

$(-)$ $(-)$ $(-)$ $(+)$ $(+)$ $(+)$

$$-(0) - (16) - (1) + 6 + 0 + 0 = -11$$

Determinante - exemplo

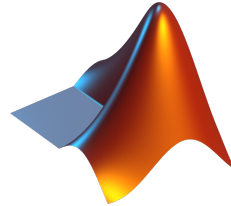


```
[1]: import numpy as np

a = np.array([[2, 1, 0],
              [1, 3, 4],
              [0, 2, 1]])

det = np.linalg.det(a)
print(det)

-11.000000000000002
```



Command Window

```
>> a = [2, 1, 0; 1, 3, 4; 0, 2, 1];
>> det(a)
```

ans =

-11

Determinante da forma matricial

Se $A : E \rightarrow E$ é um operador em um espaço vetorial munido de produto interno e $\mathbf{a}$ é sua matriz numa base ortonormal então

$$\det A = \det \mathbf{a} = \det \mathbf{a}^T = \det A^*$$

Resulta da igualdade $\det \mathbf{a} = \det \mathbf{a}^T$ que o determinante é uma função n-linear alternada das *linhas* da matriz $\mathbf{a} \in \mathcal{M}(m \times n)$. (A única tal que $\det \mathbf{I}_n = 1$.) Como no caso de colunas, uma qualquer função n-linear alternada f das linhas da matriz $\mathbf{a}$ é a forma $f(\mathbf{a}) = c \cdot \det \mathbf{a}$, onde $c = f(\mathbf{I}^n) = f(e_1, \dots, e_n)$.

Sendo o determinante de uma matriz uma função multilinear alternada das linhas ou das colunas dessa matriz, segue-se o teorema a seguir:

Theorem

Se $\mathbf{b} \in \mathcal{M}(m \times r)$, $\mathbf{c} \in \mathcal{M}(r \times (n - r))$, $\mathbf{0} \in \mathcal{M}((n - r) \times r)$ e $\mathbf{d} \in \mathcal{M}((n - r) \times (n - r))$ então o determinante da matriz

$$\mathbf{a} = \begin{bmatrix} \mathbf{b} & \mathbf{c} \\ \mathbf{0} & \mathbf{d} \end{bmatrix} \in \mathcal{M}(n \times n)$$

é igual a $\det \mathbf{b} \cdot \det \mathbf{d}$

O teorema anterior implica em sua versão mais geral, onde se tem uma matriz triangular por blocos, por exemplo, do tipo

$$\mathbf{a} = \begin{bmatrix} \mathbf{b} & \mathbf{c} & \mathbf{d} \\ & \mathbf{e} & \mathbf{f} \\ & & \mathbf{g} \end{bmatrix}$$

Com $\mathbf{b}$, $\mathbf{e}$, $\mathbf{g}$ matrizes quadradas e $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$ tendo o mesmo número de linhas, assim como $\mathbf{e}$, $\mathbf{f}$. Além disso, $\mathbf{d}$, $\mathbf{f}$, $\mathbf{g}$ têm o mesmo número de colunas, assim como $\mathbf{c}$, $\mathbf{e}$. Os lugares em branco são ocupados por zeros.

Nestas condições, $\det \mathbf{a} = \det \mathbf{b} \cdot \det \mathbf{e} \cdot \det \mathbf{g}$.

Menor relativo

Dada $\mathbf{a} = [a_{ij}] \in \mathcal{M}(m \times n)$, indicaremos com a notação $\mathbf{M}_{ij}$ a matriz de ordem $(n-1) \times (n-1)$ obtida mediante a omissão da i -ésima linha da j -ésima coluna da matriz $\mathbf{a}$. O ij -ésimo menor de $\mathbf{a}$ é, por definição, o determinante $\det \mathbf{M}_{ij}$ o qual se chama também *menor relativo ao elemento a_{ij}* .

Daí, tem-se

$$\det \mathbf{a} = \sum_{i=1}^n (-1)^{(i+j)} a_{ij} \det \mathbf{M}_{ij}$$

Daí, nasce a *adjunta clássica*

$$\text{adj } \mathbf{a}_{ij} = (-1)^{(i+j)} \det \mathbf{M}_{ji}$$

significando que

$$\mathbf{a} \cdot \text{adj } \mathbf{a} = (\det \mathbf{a}) \cdot \mathbf{I}_n$$

Segue-se daí que se a matriz $\mathbf{a}$ é invertível ($\det \mathbf{a} \neq 0$) então sua inversa se exprime como

$$\mathbf{a}^{-1} = \frac{1}{\det \mathbf{a}} \cdot \text{adj } \mathbf{a}$$

Polinômio característico

Seja $A : E \rightarrow E$ um operador linear num espaço vetorial E , de dimensão finita. A fim de que um número real λ seja autovalor de A , é necessário e suficiente que exista $v \neq 0$ em E tal que $(A - \lambda I)v = 0$, ou seja, que o operador $A - \lambda I : E \rightarrow E$ tenha um núcleo não-trivial e portanto não seja invertível. Isto acontece se, e somente se, $\det(A - \lambda I) = 0$. Conforme resulta da definição clássica de determinante $\det(A - \lambda I)$ é um polinômio de grau n em λ . Ele é chamado de *polinômio característico* do operador A e é representado por $p_A(\lambda)$

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

As raízes (reais ou complexas) da equação algébrica $p_A(\lambda) = 0$ são chamadas as raízes características do operador A . Segue-se que os autovalores do operador linear A são suas raízes características reais.

Polinômio característico

Um operador $A : E \rightarrow E$ diz-se triangularizável quando existe uma base $\mathcal{U}$ de E não qual a matriz de A é triangular. Se a matriz de A na base $\mathcal{U} = \{u_1, \dots, u_n\}$ é triangular inferior então, na base $\mathcal{U}' = \{u_n, \dots, u_1\}$, a matriz de A é triangular superior. Isto significa que existem subespaços $F_0 \subset F_1 \subset \dots \subset F_n = E$, invariantes por A , tais que $\dim F_i = i$. Se A é triangularizável e, na base $\mathcal{U}$, sua matriz $\mathbf{a} = [a_{ij}]$ é triangular superior então o polinômio característico de A é

$$p_A(\lambda) = \prod_{i=1}^n (a_{ii} - \lambda)$$

Polinômio característico

É uma tarefa muito difícil a determinação dos coeficientes de p_A quando o seu grau é elevado (e mais difícil ainda a determinação de suas raízes). Um desses coeficientes é, entretanto, fácil de calcular: o termo independente de λ é igual a $p_A(0)$, logo é igual a $\det A$.

Por outro lado, se as raízes de p_A são $\lambda_1, \dots, \lambda_n$, tem-se

$$p_A(\lambda) = (-1)^n (\lambda - \lambda_1) \cdots (\lambda - \lambda_n)$$

Com $\lambda = 0$ vem $\det A = p_A(0) = \lambda_1 \cdots \lambda_n$. Portanto o determinante de A é igual ao produto das suas raízes características, mesmo quando algumas delas são número complexos.

Outro termo de fácil determinação no polinômio $p_A(\lambda)$ é o coeficiente de λ^{n-1} . Na expressão clássica de $\det(A - \lambda I)$ em termos da matriz $\mathbf{a} = [a_{ij}]$ de A numa certa base, as parcelas que contém a potência λ^{n-1} resultam do produto $\prod (a_{ii} - \lambda)$ dos termos da diagonal de $\mathbf{a} - \lambda \mathbf{I}_n$, logo são todas da forma $(-1)^{n-1} a_{ii} \lambda^{n-1}$. Portanto $(-1)^{n-1} \sum a_{ii}$ é o coeficiente de λ^{n-1} no polinômio $p_A(\lambda)$.

A expressão

$$p_A(\lambda) = (-1)^n \prod_{i=1}^n (\lambda - \lambda_i)$$

mostra que o coeficiente de λ^{n-1} é igual a $(-1)^{n-1}$ vezes a soma das raízes do polinômio p_A .

A soma $\sum a_{ii}$ chama-se o *traço* do operador A e é designado por $tr A$.

Teorema de Cayley-Hamilton

Dados um operador linear $A : E \rightarrow E$ e um polinômio

$$p(\lambda) = a_0 + a_1\lambda + \cdots + a_m\lambda^m$$

o símbolo $p(A)$ representará o operador

$$p(A) = a_0I + a_1A + \cdots + a_mA^m$$

obtido de $p(\lambda)$ substituindo-se λ^i por A^i , tendo o cuidado de lembrar que $A^0 = I$, logo $\lambda^0 = 1$ deve ser substituído por I .

Theorem

Se o polinômio característico p_A do operador $A : E \rightarrow E$ é o produto de fatores reais do primeiro grau então $p_A(A) = 0$.

Multiplicidade

Seja λ_0 um autovalor do operador $A : E \rightarrow E$. A *multiplicidade geométrica* de λ_0 é a dimensão do subespaço vetorial $F_{\lambda_0} = \{v \in E; Av = \lambda_0 v\}$. A *multiplicidade algébrica* de λ_0 é sua multiplicidade como raiz do polinômio característico de A , isto é, é o maior inteiro m tal que $p_A(\lambda) = (\lambda_0 - \lambda)^m \cdot q(\lambda)$, onde $q(\lambda)$ é ainda um polinômio.

Theorem

A multiplicidade geométrica de um autovalor é menor ou igual do que a sua multiplicidade algébrica.

Álgebra Linear

Aula 11: Espaços vetoriais complexos

Kelvin Rafael Duarte Machado

dezembro de 2025

Espaço vetorial complexo

Um *espaço vetorial complexo* é um conjunto E , cujos elementos são chamados *vetores*, no qual estão definidas duas operações: a *adição*, que faz corresponder a cada par de vetores $u, v \in E$ um vetor $u + v$, chamado a soma de u e v , e a *multiplicação por um número complexo*, que a cada número complexo ζ e a cada vetor $v \in E$ faz corresponder um vetor $\zeta \cdot v = \zeta v$, chamado de produto de ζ por v . Essas operações devem cumprir as mesmas condições impostas para os espaços vetoriais reais, já apresentadas. Em particular, se $v \in E$ e $\zeta = \alpha + i\beta$ então $\zeta v = \alpha v + i\beta v = \alpha v + \beta(iv)$.

As definições de dependência linear, geradores, subespaço, base, dimensão, etc. se fazem para os espaços vetoriais complexos da mesma maneira como foram feitas para os reais.

Um espaço vetorial complexo E pode ser considerado como um espaço vetorial real: basta que se considere apenas a multiplicação dos vetores de E por números reais. Analogamente, toda transformação $\mathbb{C}$ -linear $A : E \rightarrow F$ entre espaços vetoriais complexos é $\mathbb{R}$ -linear.

Correspondência entre $\mathbb{C}$ e $\mathbb{R}^2$

A correspondência $x + iy \leftrightarrow (x, y)$ é um isomorfismo natural entre os espaços vetoriais reais $\mathbb{C}$ e $\mathbb{R}^2$. Esse isomorfismo faz corresponder a cada operador $\mathbb{C}$ -linear $A : \mathbb{C} \rightarrow \mathbb{C}$, um operador $\mathbb{R}$ -linear $A_r : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, dado por

$$A_r(x, y) = (\alpha x - \beta y, \beta x + \alpha y),$$

cuja matriz na base canônica é

$$\begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}$$

Para o caso trivial $A = 0$. Escrevemos

$$\rho = \sqrt{\alpha^2 + \beta^2}$$

Correspondência entre $\mathbb{C}$ e $\mathbb{R}^2$

Tomando $\theta \in \mathbb{R}$ tal que $\cos \theta = \alpha/\rho$, $\sin \theta = \beta/\rho$. Então a matriz de A_r tem a forma

$$\rho \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

Isto mostra que o operador A_r é uma *semelhança*, composta da rotação de ângulo θ com a homotetia de razão $\rho > 0$. As semelhanças são, portanto, os operadores $\mathbb{R}^2$ que correspondem aos operadores $\mathbb{C}$ -lineares não nulos $A : \mathbb{C} \rightarrow \mathbb{C}$.

Se $\mathcal{U} = \{u_1, \dots, u_n\} \in E$ for uma base do espaço vetorial complexo E então o conjunto $\mathcal{U}' = \{u_1, \dots, u_n, iu_1, \dots, iu_n\} \in E$ é uma $\mathbb{R}$ -base, ou seja, é L.I. sobre os reais e, além disso, todo vetor $v \in E$ é uma combinação linear dos u_j e dos $iu_j (j = 1, \dots, n)$ com coeficientes reais.

Se $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \in \mathbb{R}$ então

$$\alpha_1 u_1 + \dots + \alpha_n u_n + \beta_1 iu_1 + \dots + \beta_n iu_n = 0$$

implica

$$(\alpha_1 + i\beta_1)u_1 + \dots + (\alpha_n + i\beta_n)u_n = 0$$

Logo $\alpha_1 + i\beta_1 = \cdots = \alpha_n + i\beta_n = 0$, pois $\mathcal{U}$ é L.I. sobre os complexos. Ao passo que $\alpha_1 = \cdots = \alpha_n = \beta_1 = \cdots = \beta_n = 0$, portanto $\mathcal{U}'$ é L.I. sobre os reais.

Além disso, para qualquer $v \in E$, existem números complexos $\alpha_1 + i\beta_1, \cdots, \alpha_n + i\beta_n$ tais que

$$v = \sum_{j=1}^n (\alpha_j + i\beta_j) u_j = \sum_{j=1}^n \alpha_j u_j + \sum_{j=1}^n \beta_j (iu_j)$$

Assim, $\mathcal{U}'$ é uma $\mathbb{R}$ -base para E .

Podemos, portanto, concluir que se $\dim_{\mathbb{C}} E = n$ então $\dim_{\mathbb{R}} E = 2n$

Produto interno hermitiano

Se o produto interno for bilinear, como definimos para o caso dos espaços vetoriais reais, então $\langle iv, iv \rangle = i^2 \langle v, v \rangle = -\langle v, v \rangle$, logo não pode ser positivo.

Deste modo, precisamos introduzir a noção de *produto interno hermitiano*. Este é, por definição uma função $E \times E \rightarrow \mathbb{C}$ que associa a cada par de vetores u, v no espaço vetorial complexo E um número complexo, representado pela notação $\langle u, v \rangle$, de tal modo que sejam cumpridas as condições seguintes, para quaisquer $u, v, u' \in E$, $\zeta \in \mathbb{C}$, onde uma barra sobre o número complexo $\zeta = \alpha + i\beta$ significa seu conjugado $\bar{\zeta} = \alpha - i\beta$:

- $\langle u, v \rangle = \overline{\langle v, u \rangle}$
- $\langle u + u', v \rangle = \langle u, v \rangle + \langle u', v \rangle$
- $\langle \zeta u, v \rangle = \zeta \langle u, v \rangle$
- $\langle u, u \rangle \geq 0$ se $u \neq 0$

Theorem

Seja E um espaço vetorial complexo de dimensão finita munido de produto interno hermitiano. A correspondência que associa a cada vetor $v \in E$ o funcional linear $\phi(v)v^ : E \rightarrow \mathbb{C}$, tal que $v^*(w) = \langle w, v \rangle$ para todo $w \in E$ é uma bijeção $\phi : E \rightarrow E^*$ tal que $(u + v)^* = u^* + v^*$ e $(\zeta v)^* = \bar{\zeta} \cdot v^*$ para quaisquer $u, v \in E$ e $\zeta \in \mathbb{C}$.*

A correspondência $v \rightarrow v^*$ não é um isomorfismo entre E e E^* , pois falha a condição $(\zeta v)^* = \zeta v^*$, em vez da qual se tem apenas $(\zeta v)^* = \bar{\zeta} v^*$.

Por isso, ela se chama um anti-isomorfismo.

Deste modo, segue-se que A^* é invertível se, e somente se, A é invertível. Já da adjunta, tem-se que $(\zeta A)^* = \bar{\zeta} A^*$.

Um operador $\mathbb{C}$ -linear $A : E \rightarrow E$ chama-se *hermitiano* quando $A = A^*$, ou seja, quando $\langle Au, v \rangle = \langle u, Av \rangle$ para quaisquer $u, v \in E$. Em particular, quando $u = v$, temos $\langle Av, v \rangle = \langle v, Av \rangle$. Portanto, quando A é hermitiano, a forma quadrática $\phi(v) = \langle v, Av \rangle$ só assume valores reais.

Uma matriz $\mathbf{a} = [a_{ij}] \in \mathcal{M}(m \times n; \mathbb{C})$ chama-se *hermitiana* quando $\mathbf{a} = \mathbf{a}^*$, isto é, quando $a_{jk} = \bar{a}_{kj}$ para $k, j = 1, \dots, n$.

Identidade de polarização

Um operador $\mathbb{C}$ -linear $U : E \rightarrow E$ chama-se *unitário* quando $U^* = U^{-1}$. Isto equivale a dizer que $\langle Uv, Uw \rangle = \langle v, w \rangle$ para quaisquer $v, w \in E$.

Se $U : E \rightarrow E$ é um operador unitário então, para todo $v \in E$ tem-se $|Uv| = |v|$. Vale também a recíproca. Para prová-la, usa-se a identidade de polarização

$$\langle u, v \rangle = \frac{1}{4}[|u + v|^2 - |u - v|^2 + i|u + iv|^2 - i|u - iv|^2]$$

Uma matriz $\mathbf{u} \in \mathcal{M}(m \times n; \mathbb{C})$ chama-se unitária quando $\mathbf{u}^* = \mathbf{u}^{-1}$. Para que isto aconteça basta que $\mathbf{u}^* \mathbf{u} = \mathbf{I}_n$ ou que $\mathbf{u} \mathbf{u}^* = \mathbf{I}_m$

Álgebra Linear

Aula 12: Métodos numéricos

Kelvin Rafael Duarte Machado

dezembro de 2025

Matriz de Gram

Seja E um espaço vetorial de dimensão finita, munido de produto interno. A *matriz de Gram* dos vetores $v_1, \dots, v_k \in E$ é a matriz $\mathbf{g} = [g_{ij}] \in \mathcal{M}(k \times k)$, onde $g_{ij} = \langle v_i, v_j \rangle$. Dada uma base $\mathcal{U} = \{u_1, \dots, u_n\} \in E$, seja $\mathbf{a} = [a_{ij}] \in \mathcal{M}(m \times k)$ a matriz das coordenadas dos vetores v_j em relação à base $\mathcal{U}$, isto é

$$v_j = a_{1j}u_1 + \dots + a_{nj}u_n \quad \text{para } j = 1, \dots, k$$

Matriz de Gram

Seja ainda $\mathbf{h} = [h_{ij}] \in \mathcal{M}(m \times n)$ a matriz de Gram na base $\mathcal{U}$, isto é, $h_{ij} = \langle u_i, u_j \rangle$. Então, para $i, j = 1, \dots, k$, temos (escrevendo $\mathbf{m}_{ij}$ para indicar o ij -ésimo elemento de uma matriz $\mathbf{m}$):

$$g_{ij} = \langle v_i, v_j \rangle = \left\langle \sum_{r=1}^n a_{ri} u_r, \sum_{s=1}^n a_{sj} u_s \right\rangle = \sum_{r,s=1}^n a_{ri} a_{sj} h_{rs}$$

$$\sum_{r=1}^n a_{ri} \left(\sum_{s=1}^n h_{rs} a_{sj} \right) = \sum_{r=1}^n (\mathbf{a}^T)_{ir} (\mathbf{h}\mathbf{a})_{rj} = (\mathbf{a}^T \mathbf{h}\mathbf{a})_{ij}$$

portanto $\mathbf{g} = \mathbf{a}^T \mathbf{h}\mathbf{a}$.

Matriz de Gram

Em particular, se tomarmos uma base ortonormal $\{u_1, \dots, u_n\} \subset E$, teremos $\mathbf{h} = \mathbf{I}_n$, portanto a matriz de Gram $\mathbf{g}$ se escreve como

$$\mathbf{g} = \mathbf{g}(v_1, \dots, v_k) = \mathbf{a}^T \cdot \mathbf{a},$$

onde $\mathbf{a}$ é a matriz das coordenadas dos vetores v_j em relação a uma base ortonormal de E . Daí resulta:

- 1 Toda matriz de Gram é não-negativa;
- 2 A matriz de Gram $\mathbf{g} = \mathbf{g}(v_1, \dots, v_k)$ é positiva (isto é invertível) se, e somente se, os vetores $v_1, \dots, v_k$ são L.I.
- 3 Toda matriz não-negativa $\mathbf{g} = [g_{ij}] \in \mathcal{M}(k \times k)$ é a matriz de Gram de uma lista de vetores $v_1, \dots, v_k \in E$.

Matriz de Gram

1. Definição da matriz de Gram

$$G = \begin{bmatrix} v_1 \cdot v_1 & v_1 \cdot v_2 & v_1 \cdot v_3 \\ v_2 \cdot v_1 & v_2 \cdot v_2 & v_2 \cdot v_3 \\ v_3 \cdot v_1 & v_3 \cdot v_2 & v_3 \cdot v_3 \end{bmatrix}$$

2. Produto interno $v_1 \cdot v_j$

$$v_1 \cdot v_1 = 1^2 + 2^2 + 0^2 = 5$$

$$v_1 \cdot v_2 = 1 \cdot 2 + 2 \cdot 3 + 0 \cdot 1 = 8$$

$$v_1 \cdot v_3 = 1 \cdot 0 + 2 \cdot 1 + 0 \cdot 1 = 2$$

$$v_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

3. Produto interno $v_2 \cdot v_j$

$$v_2 \cdot v_1 = 8, \quad v_2 \cdot v_2 = 2^2 + 3^2 + 1^2 = 14$$

$$v_2 \cdot v_3 = 2 \cdot 0 + 3 \cdot 1 + 1 \cdot 1 = 4$$

4. Produto interno $v_3 \cdot v_j$

$$v_3 \cdot v_1 = 2, \quad v_3 \cdot v_2 = 4, \quad v_3 \cdot v_3 = 0^2 + 1^2 + 1^2 = 2$$

5. Matriz de Gram final

$$G = \begin{bmatrix} 5 & 8 & 2 \\ 8 & 14 & 4 \\ 2 & 4 & 2 \end{bmatrix}$$

6. Propriedades

$$G = G^T \text{ (simétrica),}$$

$$G \geq 0 \text{ (semi-definida positiva)}$$

Eliminação de Gauss

Se a primeira coordenada não-nula de cada vetor w_i tem índice menor do que a primeira coordenada não-nula dos vetores subsequentes $w_{i+1}, \dots, w_r$, dizemos que os vetores $w_1, \dots, w_r \in \mathbb{R}^n$ são linearmente independentes.

Se, para cada $i = 1, \dots, r$, escrevemos $w_i = (a_{i1}, \dots, a_{in})$, teremos uma matriz $\mathbf{a} = [a_{ij}] \in \mathcal{M}(r \times n)$, cujos r vetores linha são $w_1, \dots, w_r$. Diremos que essa matriz é *escalorada* quando o primeiro elemento não-nulo de cada uma das linhas está à esquerda do primeiro elemento não-nulo de cada uma das linhas subsequentes. As linhas nulas (se houver) estarão abaixo das demais.

Exemplo:

$$\begin{bmatrix} 1 & 3 & 5 & 2 \\ 0 & 2 & 7 & 1 \\ 0 & 0 & 4 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Eliminação de Gauss

As seguintes modificações, chamadas *operações elementares* levam os vetores $v_1, \dots, v_m \in \mathbb{R}^n$ em vetores $v'_1, \dots, v'_m \in \mathbb{R}^n$ que geram o mesmo subespaço.

- Trocar de posição dois vetores $v_i, v_j (i < j)$ na lista dada.
- Somar a um dos vetores um múltiplo de outro vetor da lista ou seja, substituir v_j por $v'_j = v_j + \alpha v_i, i \neq j$.

Eliminação de Gauss

O *escalonamento* ou *eliminação* consiste em aplicar sucessivamente as operações elementares às linhas de uma matriz de modo a produzir uma matriz escalonada. O procedimento é o seguinte:

- 1 Se $a_{11} \neq 0$, o processo começa deixando a primeira linha intacta e somando a cada linha l_i , com $i \geq 2$, a primeira linha multiplicada por $-a_{i1}/a_{11}$. Com isto se obtém uma matriz cuja a primeira coluna é $(a_{11}, 0, \dots, 0)$.
- 2 Se $a_{11} = 0$, uma troca de linhas fornece uma matriz $a_{11} \neq 0$, desde que a primeira coluna não seja nula. Se, porém todos os elementos da primeira coluna são iguais a zero, passa-se para a segunda coluna ou, mais geralmente, para coluna mais próxima, à direita da primeira, onde haja algum elemento não-nulo e opera-se como antes.

Eliminação de Gauss

O método de eliminação, embora simples e ingênuo, é a maneira mais eficaz de resolver um sistema de m equações lineares, com n incógnitas, apresentado sob a forma matricial $\mathbf{ax} = \mathbf{b}$, onde $\mathbf{a} \in \mathcal{M}(m \times n)$, $\mathbf{x} \in \mathcal{M}(n \times 1)$ e $\mathbf{b} \in \mathcal{M}(m \times 1)$. O sistema $\mathbf{ax} = \mathbf{b}$ possui solução se, e somente se, o vetor $\mathbf{b} \in \mathbb{R}^m$ (correspondente à matriz $\mathbf{b}$) pertence ao subespaço gerado pelas colunas de $\mathbf{a}$. Assim, a *matriz aumentada* $[\mathbf{a}, \mathbf{b}] \in \mathcal{M}(m \times (n + 1))$ tem o mesmo posto que a matriz $\mathbf{a}$ do sistema.

Eliminação de Gauss

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

1. Matriz inicial

$$A_0 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$$

2. Eliminação coluna 1

Multiplicadores:

$$m_{21} = 5, \quad m_{31} = 9$$

Aplicando

$$L_2 \leftarrow L_2 - 5L_1, \quad L_3 \leftarrow L_3 - 9L_1$$

3. Resultado após passo 1

$$A_1 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 0 & -8 & -16 & -24 \end{bmatrix}$$

4. Eliminação coluna 2

Pivô em (2,2): -4 .

Multiplicador:

$$m_{32} = \frac{-8}{-4} = 2$$

Aplicando

$$L_3 \leftarrow L_3 - 2L_2$$

5. Resultado final (escalonada)

$$A_2 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

A_2 é a forma escalonada de A_0

Eliminação de Gauss-Jordan

Outra aplicação do método de eliminação é o cálculo da inversa de uma matriz (invertível) $\mathbf{a} \in \mathcal{M}(n \times n)$. Antes porém devemos advertir que a determinação da inversa não é necessária para resolver o sistema $\mathbf{ax} = \mathbf{b}$. A expressão $\mathbf{x} = \mathbf{a}^{-1} \cdot \mathbf{b}$ para a solução desse sistema é adequada porém, não prática.

Examinando coluna por coluna cada membro da igualdade $\mathbf{aa}^{-1} = \mathbf{I}_n$, vemos que j -ésima coluna de $\mathbf{a}^{-1}$ é a solução do sistema $\mathbf{ax} = \mathbf{e}_j$, portanto o cálculo da inversa equivale a resolver os n sistemas lineares $\mathbf{ax} = \mathbf{e}_j, \dots, \mathbf{ax} = \mathbf{e}_n$.

A este método dá-se o nome de "método de Gauss-Jordan".

Ele continua a eliminação iniciada pelo método de Gauss, chegando no final a uma matriz escalonada, com a propriedade adicional de que, acima e abaixo do primeiro elemento não-nulo de cada linha, todos os elementos são iguais a zero.

Eliminação de Gauss-Jordan

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

1. Formar matriz aumentada

$$[A|I] = \begin{bmatrix} 2 & 4 & 3 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 3 & 5 & 7 & 0 & 0 & 1 \end{bmatrix}$$

2. Eliminar elementos abaixo do pivô (coluna 1)

$$L_3 \leftarrow L_3 - \frac{3}{2}L_1$$

$$\begin{bmatrix} 2 & 4 & 3 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & -1 & \frac{5}{2} & -\frac{3}{2} & 0 & 1 \end{bmatrix}$$

3. Eliminar acima e abaixo do pivô (coluna 2)

$$L_1 \leftarrow L_1 - 4L_2, \quad L_3 \leftarrow L_3 + L_2$$

$$\begin{bmatrix} 2 & 0 & 7 & 1 & -4 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & \frac{3}{2} & -\frac{3}{2} & 1 & 1 \end{bmatrix}$$

4. Normalizar linha 3 e eliminar acima (coluna 3)

$$L_3 \leftarrow \frac{2}{3}L_3, \quad L_1 \leftarrow L_1 - \frac{7}{3}L_3, \quad L_2 \leftarrow L_2 + \frac{1}{3}L_3$$

$$\begin{bmatrix} 1 & 0 & 0 & 4 & -\frac{13}{3} & -\frac{7}{3} \\ 0 & 1 & 0 & -1 & \frac{5}{3} & \frac{2}{3} \\ 0 & 0 & 1 & -1 & \frac{2}{3} & \frac{2}{3} \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} 4 & -\frac{13}{3} & -\frac{7}{3} \\ -1 & \frac{5}{3} & \frac{2}{3} \\ -1 & \frac{2}{3} & \frac{2}{3} \end{bmatrix}$$

Matrizes triangulares

Uma matriz $\mathbf{t} = [t_{ij}] \in \mathcal{M}(n \times n)$ é dita triangular superior quando $t_{ij} = 0$ para $i > j$ e triangular inferior quando $t_{ij} = 0$ para $i < j$.

Tratando das matrizes triangulares superiores e estendendo para as inferiores, uma matriz triangular superior é a matriz de um operador linear $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tal que $\langle e_i, Te_j \rangle = 0$ para $i > j$. Se chamarmos de $F_i \subset \mathbb{R}^n$ o subespaço vetorial formado pelos vetores $(x_1, \dots, x_i, \dots, 0)$ cujas últimas $n - i$ coordenadas são nulas, a matriz do operador $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ (na base canônica) é triangular superior se, e somente se, todos os subespaços $F_0 \subset \dots \subset F_n$ são invariantes por T .

Matrizes triangulares

Seguem-se algumas propriedades das matrizes triangulares superiores:

- 1 O produto de duas matrizes triangulares superiores é ainda uma matriz triangular superior.
- 2 Uma matriz triangular superior $\mathbf{t} = [t_{ij}]$ é invertível se, e somente se, os elementos t_{ii} da sua diagonal são todos diferentes de zero. No caso afirmativo, a inversa $\mathbf{t}^{-1}$ é triangular superior e os elementos de sua diagonal são t_{ii}^{-1} .
- 3 Os autovalores de uma matriz triangular superior $\mathbf{t} = [t_{ij}] \in \mathcal{M}(m \times n)$ são os elementos t_{ii} da sua diagonal.
- 4 Seja $\mathcal{M} = \{u_1, \dots, u_n\} \subset E$ uma base ortonormal obtida pelo processo de Gram-Schmidt a partir da base $\mathcal{V} = \{v_1, \dots, v_n\} \in E$. A matriz de passagem de $\mathcal{V}$ para $\mathcal{U}$ é triangular superior e seus autovalores são todos positivos.

Decomposição LU

Cada uma das três operações elementares sobre as linhas de uma matriz, pode ser interpretada como a multiplicação à esquerda por uma matriz invertível de tipo especial, chamada uma matriz elementar.

Mais precisamente, uma *matriz elementar* $m \times m$ é uma matriz que resulta da aplicação de uma operação à matriz identidade $\mathbf{I}_m$. Há portanto 3 tipos de matrizes elementares. Vejamos alguns exemplos:

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

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

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

Ao realizar sucessivas operações elementares em uma matriz, pode-se escrever a matriz $\mathbf{a}$ como um produto:

$$\mathbf{a} = \mathbf{l}\mathbf{u}$$

onde $\mathbf{l} \in \mathcal{M}(m \times m)$ é uma matriz triangular inferior (lower triangular) com elementos diagonais todos iguais a 1 e $\mathbf{u}$ é uma matriz escalonada. Se $\mathbf{a}$ for uma matriz quadrada $m \times m$ então $\mathbf{u} \in \mathcal{M}(m \times m)$ será uma matriz triangular superior. Esta é chamada *decomposição lu* da matriz $\mathbf{a}$.

Decomposição LU

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 4 & 7 & 2 \\ 6 & 18 & -1 \end{bmatrix}$$

1. Elimin. coluna 1

Multiplicadores:

$$m_{21} = \frac{4}{2} = 2, \quad m_{31} = \frac{6}{2} = 3$$

Aplicando

$$L_2 \leftarrow L_2 - 2L_1, \quad L_3 \leftarrow L_3 - 3L_1$$

Obtemos

$$\begin{bmatrix} 2 & 3 & 1 \\ 0 & 1 & 0 \\ 0 & 9 & -4 \end{bmatrix}$$

2. Elimin. coluna 2

Pivô na posição (2,2): 1.

Multiplicador:

$$m_{32} = \frac{9}{1} = 9$$

Aplicando

$$L_3 \leftarrow L_3 - 9L_2$$

Resulta

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

3. Construção de L e verificação

Os multiplicadores formam L:

$$L = \begin{bmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & m_{32} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 9 & 1 \end{bmatrix}$$

Verificação:

$$LU = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 9 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -4 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 1 \\ 4 & 7 & 2 \\ 6 & 18 & -1 \end{bmatrix}$$

$$\boxed{A = LU}$$

Decomposição QR

Esta é uma interpretação matricial do processo de Gram-Schmidt. Segundo ela, toda matriz invertível $\mathbf{a} = [a_{ij}] \in \mathcal{M}(n \times n)$ admite uma decomposição do tipo $\mathbf{a} = \mathbf{q}\mathbf{r}$ onde $\mathbf{q}$ é ortogonal e $\mathbf{r}$ é triangular superior, com elementos positivos na diagonal.

Para chegar a este resultado, chamemos de $v_1, \dots, v_n$ as colunas da matriz $\mathbf{a}$ e de $\mathcal{U} = \{u_1, \dots, u_n\} \subset \mathbb{R}^n$ a base ortonormal obtida dos v_i pelo processo de Gram-Schmidt. Como sabemos, a matriz $\mathbf{p} = [p_{ij}]$ de passagem da base $\mathcal{V} = \{v_1, \dots, v_n\}$ para a base $\mathcal{U}$ é triangular superior, com elementos positivos na diagonal. Além disso, a matriz $\mathbf{q} = [q_{ij}]$ cujas colunas são os vetores $u_j = (q_{1j}, q_{2j}, \dots, q_{nj})$, é ortogonal.

Decomposição QR

Tomando a i -ésima coordenada de ambos os membros da igualdade vetorial $u_j = \sum_k p_{kj} v_k$, obtemos

$$q_{ij} = \sum_{k=1}^n p_{kj} a_{ik} = \sum_{k=1}^n a_{ik} p_{kj} = (\mathbf{ap})_{ij}$$

Para quaisquer $i, j = 1, \dots, n$. Logo $\mathbf{q} = \mathbf{ap}$. A matriz $\mathbf{r} = \mathbf{p}^{-1}$ é, como vimos acima, triangular superior com elementos positivos na diagonal. De $\mathbf{ap} = \mathbf{q}$ resulta imediatamente $\mathbf{a} = \mathbf{qr}$.

Decomposição QR

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

1. Vetor a_1

$$a_1 = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

2. Normalizar: q_1

$$\|a_1\| = \sqrt{4^2 + 2^2} = 2\sqrt{5}$$

$$q_1 = \frac{1}{2\sqrt{5}} \begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$$

3. Vetor a_2

$$a_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

4. Projeção de a_2 em q_1

$$u_2 = a_2 - (q_1^T a_2) q_1$$

$$q_1^T a_2 = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \frac{7}{\sqrt{5}}$$

$$u_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix} - \frac{7}{\sqrt{5}} \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} = \begin{bmatrix} -4/5 \\ 8/5 \end{bmatrix}$$

5. Normalizar u_2

$$q_2 = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

6. Montar Q

$$Q = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

7. Montar $R = Q^T A$

$$R = \begin{bmatrix} 2\sqrt{5} & 7/\sqrt{5} \\ 0 & 4\sqrt{5}/5 \end{bmatrix}$$

8. Verificação

$$QR = A$$

$$\begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 2\sqrt{5} & 7/\sqrt{5} \\ 0 & 4\sqrt{5}/5 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix}$$

Decomposição de Cholesky

Toda a matriz positiva $\mathbf{a} = [a_{ij}] \in \mathcal{M}(n \times n)$ pode ser expressa como o produto $\mathbf{a} = \mathbf{t}^T \cdot \mathbf{t}$, onde $\mathbf{t} \in \mathcal{M}(n \times n)$ é uma matriz triangular superior cujos elementos da diagonal são todos positivos. A expressão $\mathbf{a} = \mathbf{t}^T \cdot \mathbf{t}$ chama-se *decomposição de Cholesky* da matriz $\mathbf{a}$.

Existem vetores $v_1, \dots, v_n \in \mathbb{R}^n$ tais que $a_{ij} = \langle v_i, v_j \rangle$, ou seja, $\mathbf{a} = \mathbf{g}(v_1, \dots, v_n)$ é a matriz de Gram dos vetores $v_1, \dots, v_n$, os quais formam uma base de $\mathbb{R}^n$, pois $\mathbf{a} > 0$.

Decomposição de Cholesky

Pelo processo de Gram-Schmidt, obtemos uma base ortonormal $\{u_1, \dots, u_n\} \subset \mathbb{R}^n$ a partir de $v_1, \dots, v_n$. Para $i, j = 1, \dots, n$, temos

$$u_i = \sum_r p_{ri} v_r, \quad u_j = \sum_s p_{sj} v_s$$

onde a matriz de passagem $\mathbf{p} = [p_{ij}]$ é triangular superior, com $p_{ii} > 0$ para todo i . Usando o símbolo de Kronecker δ_{ij} , temos

$$\delta_{ij} = \langle u_i, u_j \rangle = \sum_{r,s=1}^n p_{ri} p_{sj} \langle v_r, v_s \rangle = \sum_{r,s=1}^n p_{ri} a_{rs} p_{sj}$$

logo $\mathbf{p}^T \mathbf{a} \mathbf{p} = \mathbf{I}_n$. Pondo $\mathbf{t} = \mathbf{p}^{-1}$, obtemos $\mathbf{a} = \mathbf{t}^T \cdot \mathbf{t}$.

Decomposição de Cholesky

Com efeito, de $\mathbf{s}^T \cdot \mathbf{s} = \mathbf{t}^T \cdot \mathbf{t}$ resulta $\mathbf{s}\mathbf{t}^{-1} = (\mathbf{s}^T)^{-1} \cdot \mathbf{t}^T$. Como o primeiro membro desta última igualdade é uma matriz triangular superior e o segundo é triangular inferior, concluímos que $\mathbf{d} = \mathbf{s}\mathbf{t}^{-1} = (\mathbf{s}^T)^{-1} \cdot \mathbf{t}^T$ é uma matriz diagonal, com $d_{ii} > 0$ (e $d_{ij} = 0$ se $i \neq j$). Segue-se imediatamente das igualdades acima que $\mathbf{s} = \mathbf{d}\mathbf{t}$ e $\mathbf{t} = \mathbf{d}\mathbf{s}$. Olhando para os elementos da diagonal, temos $s_{ii} = d_{ii}t_{ii}$ e $t_{ii} = d_{ii}s_{ii}$. Como $s_{ii} > 0$ e $t_{ii} > 0$, isto implica $d_{ii} = 1$, logo $\mathbf{d} = \mathbf{I}_n$ e $\mathbf{s} = \mathbf{t}$.

Decomposição de Cholesky

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

1. Verificar se A é simétrica

$$A^T = \begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix} = A$$

2. Determinar l_{11}

$$l_{11} = \sqrt{4} = 2$$

$$L = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

3. Determinar l_{21}

$$l_{21} = \frac{2}{l_{11}} = \frac{2}{2} = 1$$

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

4. Determinar l_{22}

$$l_{22} = \sqrt{3 - l_{21}^2} = \sqrt{3 - 1^2} = \sqrt{2}$$

5. Matriz final L

$$L = \begin{bmatrix} 2 & 0 \\ 1 & \sqrt{2} \end{bmatrix}$$

6. Verificação

$$LL^T = \begin{bmatrix} 2 & 0 \\ 1 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & \sqrt{2} \end{bmatrix}$$

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

$$A = LL^T \quad \checkmark$$

Diagonalização, decomposição polar e valores singulares

Se a matriz quadrada $\mathbf{a} \in \mathcal{M}(n \times n)$ possui n autovalores diferentes então existe uma matriz invertível $\mathbf{p} \in \mathcal{M}(n \times n)$ tal que $\mathbf{p}^{-1}\mathbf{a}\mathbf{p} = \mathbf{d}$ é uma matriz diagonal. Os elementos da diagonal de $\mathbf{d}$ são os autovalores de $\mathbf{a}$.

Se a matriz $\mathbf{a} \in \mathcal{M}(n \times n)$ é simétrica então existe uma matriz ortogonal $\mathbf{q} \in \mathcal{M}(n \times n)$ tal que $\mathbf{q}^T \mathbf{a} \mathbf{q} = \mathbf{d}$ é uma matriz diagonal. Os elementos da diagonal de $\mathbf{d}$ são os autovalores de $\mathbf{a}$.

Toda matriz quadrada $\mathbf{a} \in \mathcal{M}(m \times n)$ se exprime como produto $\mathbf{a} = \mathbf{p}\mathbf{u}$, onde $\mathbf{p} \in \mathcal{M}(m \times m)$ é uma matriz não-negativa e $\mathbf{u}$ é ortogonal.

Para toda matriz $\mathbf{a} \in \mathcal{M}(m \times n)$ existem matrizes ortogonais $\mathbf{p} \in \mathcal{M}(m \times m)$, $\mathbf{q} \in \mathcal{M}(n \times n)$, tais que $\mathbf{p}\mathbf{a}\mathbf{q} = \mathbf{d}$, onde $\mathbf{d} \in \mathcal{M}(m \times n)$ é uma matriz diagonal, $\mathbf{d} = [d_{ij}]$, isto é, $d_{ij} = 0$ se $i \neq j$. Para $i = 1, \dots, r = \text{posto de } \mathbf{a}$, tem-se $d_{ii} = \sigma_i > 0$, σ_i^2 é um autovalor de $\mathbf{a}^T \mathbf{a}$ (e de $\mathbf{a}\mathbf{a}^T$) e, para $i > r$, $d_{ii} = 0$.

Diagonalização

1. Autovalores

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

$$(4 - \lambda)(3 - \lambda) - 2$$

$$\lambda^2 - 7\lambda + 10 = 0$$

$$\lambda_1 = 5, \quad \lambda_2 = 2$$

2. Autovetores

Para $\lambda_1 = 5$:

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

$$(A - 5I)v = 0$$

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

$$v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Para $\lambda_2 = 2$:

$$(A - 2I)v = 0$$

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

$$v_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

3. Diagonalização

$$P = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}$$

$$D = \begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix}$$

$$P^{-1}AP = D$$

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

Decomposição polar

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

1. Calcular $A^T A$

$$A^T A = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 9 & 3 \\ 3 & 5 \end{bmatrix}$$

2. Autovalores

$$\lambda_1 = \lambda_2 = 6$$

3. Raiz positiva

$$S = \sqrt{6} I$$

$$S = \begin{bmatrix} \sqrt{6} & 0 \\ 0 & \sqrt{6} \end{bmatrix}$$

4. Inversa de S

$$S^{-1} = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

5. Cálculo de Q

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

$$Q = \frac{1}{\sqrt{6}} \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$

$$Q = \begin{bmatrix} \frac{3}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{2}{\sqrt{6}} \end{bmatrix}$$

6. Verificação

$$QS = \begin{bmatrix} \frac{3}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{2}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} \sqrt{6} & 0 \\ 0 & \sqrt{6} \end{bmatrix}$$

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

$$A = QS \quad \checkmark$$

Polar de A :

$$A = QS$$

Decomposição em valores singulares

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

1. Calcular $A^T A$

$$A^T A = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 9 & 3 \\ 3 & 5 \end{bmatrix}$$

2. Autovalores de $A^T A$

$$\lambda_1 = 10, \quad \lambda_2 = 4$$

3. Valores singulares

$$\sigma_1 = \sqrt{10}, \quad \sigma_2 = 2$$

$$\Sigma = \begin{bmatrix} \sqrt{10} & 0 \\ 0 & 2 \end{bmatrix}$$

4. Cálculo de V

Autovetores de $A^T A$:

$$V = \begin{bmatrix} 0.9487 & -0.3162 \\ 0.3162 & 0.9487 \end{bmatrix}$$

$$V^T = \begin{bmatrix} 0.9487 & 0.3162 \\ -0.3162 & 0.9487 \end{bmatrix}$$

5. Matriz Σ^{-1}

$$\Sigma^{-1} = \begin{bmatrix} \frac{1}{\sqrt{10}} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$$

6. Cálculo de U

$$U = A V \Sigma^{-1}$$

$$U = \begin{bmatrix} 0.9487 & -0.3162 \\ 0.3162 & 0.9487 \end{bmatrix}$$

Verificação

$$U \Sigma V^T = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$

$$A = U \Sigma V^T \quad \checkmark$$

SVD de A :

$$A = U \Sigma V^T$$

Tópicos Computacionais em Álgebra Linear

- Em problemas reais de engenharia e ciência de dados, resolver sistemas lineares e decompor matrizes vai muito além da teoria.
- Decomposições como **LU**, **QR**, **Cholesky** e **SVD** são a base para métodos eficientes, estáveis e robustos.
- É essencial entender o **custo computacional**, a **estabilidade numérica** e o **condicionamento** para aplicar corretamente esses métodos.
- Softwares modernos (NumPy, MATLAB, Julia, LAPACK) utilizam essas técnicas de forma otimizada, evitando, por exemplo, o cálculo direto da inversa.

Custo Computacional e Decomposições

- **LU:** $\mathcal{O}(n^3)$
 - Método padrão para resolver sistemas lineares.
 - Eficiente, mas estabilidade depende de pivotamento.
- **QR:** maior que LU
 - Estável numericamente, usado em mínimos quadrados e autovalores.
- **Cholesky:** aproximadamente metade do custo do LU
 - Apenas para matrizes simétricas definidas positivas.
 - Rápido e estável quando aplicável.
- **SVD:** muito mais caro
 - Alta robustez numérica.
 - Útil em aproximação de posto reduzido e matrizes mal condicionadas.

Estabilidade Numérica em Decomposições Matriciais

- **LU com pivotamento parcial**

- Evita divisão por elementos muito pequenos que poderiam amplificar erros de arredondamento.
- Pivotamento parcial: troca linhas para colocar o maior pivô possível na diagonal.

- **QR via Householder**

- Muito estável numericamente, mesmo para matrizes quase singulares.
- Preferível ao GramSchmidt clássico, que pode gerar perdas de ortogonalidade.
- Muito usado em problemas de mínimos quadrados e cálculos de autovalores.

- **Cholesky**

- Aplicável apenas a matrizes **simétricas e definidas positivas**.
- Se a matriz não for SPD, a fatoração falha imediatamente, prevenindo resultados incorretos.
- Estável e eficiente quando a condição é atendida.

- **SVD (Singular Value Decomposition)**

- Fornece decomposição robusta mesmo para matrizes mal condicionadas.
- Permite encontrar aproximações de posto reduzido (Teorema de Eckart-Young).
- Alto custo computacional, mas máxima robustez numérica.

Condicionalmento e Erro Relativo

- Matrizes **mal condicionadas** amplificam erros de arredondamento durante cálculos numéricos.
- **Número de condição** de uma matriz A :

$$\kappa(A) = \|A\| \|A^{-1}\|$$

- Relação entre erro relativo na solução e no vetor de termos independentes:

$$\frac{\|\delta x\|}{\|x\|} \lesssim \kappa(A) \frac{\|\delta b\|}{\|b\|}$$

- Interpretação:
 - $\kappa(A) \approx 1$ matriz bem condicionada; erros pequenos.
 - $\kappa(A) \gg 1$ matriz mal condicionada; erros amplificados.
- Implicação prática: mesmo métodos estáveis podem gerar soluções imprecisas se a matriz for mal condicionada.

Decomposições na Prática

- Softwares modernos (**NumPy, MATLAB, Julia, LAPACK**) não calculam a inversa para resolver sistemas lineares.
- Em vez disso, utilizam **fatorações diretas** ou métodos robustos:
 - `solve(A,b)` LU (ou Cholesky se a matriz for SPD)
 - `lstsq(A,b)` QR (para problemas de mínimos quadrados)
 - `svd(A)` SVD completa ou truncada (para aproximação de posto reduzido)
- Benefícios:
 - Evita instabilidade de calcular A^{-1}
 - Explora propriedades da matriz (simetria, positividade definida)
 - Otimiza desempenho computacional

Armazenamento, Sparsidade e Desempenho

- Em problemas grandes (engenharia civil, FEM, CFD), as matrizes geralmente são **esparsas**.
- Armazenamento eficiente:
 - Comprimido por linhas ou colunas (CSR, CSC)
 - Reduz uso de memória e acesso desnecessário a zeros
- Métodos diretos esparsos:
 - Variantes de LU e Cholesky controlando **fill-in**
 - Evitam crescimento excessivo de elementos não nulos
- Métodos iterativos dominam na prática:
 - **CG, GMRES, BiCGSTAB**
 - Precondicionadores essenciais: ILU, Jacobi, Incomplete Cholesky
- Estratégia prática: combinar **método apropriado** com **armazenamento eficiente** para desempenho ótimo.

Paralelização e Computação em Blocos

- Algoritmos de álgebra linear modernos exploram **paralelização** e **blocos** para desempenho.
- **BLAS / LAPACK:**
 - Operações orientadas a blocos melhor uso de cache e memória.
 - Permite escalabilidade em CPUs multi-core.
- **GPUs:**
 - Implementações de QR, Cholesky e SVD usando **kernels CUDA / MAGMA**.
 - Ganhos significativos em matrizes grandes.
- Princípio prático: dividir a matriz em blocos permite combinar eficiência de memória com paralelismo, maximizando velocidade sem comprometer estabilidade.