Álgebra Linear

Estes *slides* são um resumo do professor para diversas aulas e não devem ser usados separadamente ou como fonte original.

Sugestões são bem-vindas: mande-as para vinicius@ufabc.edu.br

Introdução

Tópicos

- Agradecimentos e fontes principais.
- Corpos e espaços em geral.

Referências

Agradecimentos

A Jerônimo Pellegrini (várias discussões sobre a matéria e ensino e seu material);

Bibliografia central

COELHO, F. U.; LOURENÇO, M. L. *Um curso de álgebra linear.* EDUSP, 2001.

FALEIROS, A. C. *Álgebra linear.* Preliminar.

PELLEGRINI, J. C. *Álgebra linear.* Preliminar.

Contexto matemático

Em muitos aspectos, esta matéria dá continuidade à Geometria Analítica.

Lá, trabalhamos com *vetores* em $\mathbb{R}^2$ ou $\mathbb{R}^3$ e os números reais eram *escalares*.

Aqui, faremos duas inovações:

- usaremos um *corpo de escalares* $\mathbb{K}$;
- trabalharemos em um *espaço euclideano* $\mathbb{K}^n$.

Vejamos o que isso significa...

Entenda $\mathbb{K}$ como um curinga que pode ser tanto $\mathbb{R}$ ou $\mathbb{C}$ ou ocasionalmente $\mathbb{Q}$, mas sempre o mesmo.

O que há de comum entre esses *corpos* é que:

- sabemos somar, subtrair, multiplicar e dividir seus elementos,
- obtendo resultados que também são elementos dos mesmos corpos
- e essas operações têm as propriedades habituais (e.g., $ab = ba$).

Outros corpos podem ser especificados, mas não detalharemos essa teoria aqui.

O *espaço euclideano* $\mathbb{K}^n$ é o conjunto das *n-uplas*

$$x = (x_1, x_2, \ldots, x_n) \text{ com cada } x_i \in \mathbb{K}.$$

Convenções:

- Dado x, suas componentes são $x_1, x_2, \ldots, x_n$.
- Dados escalares $x_1, x_2, \ldots, x_n$, formamos o vetor x.
- Podemos usar também $x = (x, y, z, \ldots)$ etc., especialmente em exemplos e exercícios práticos.
- Para uma matriz A, usamos a_{ij} em vez de A_{ij} para evitar confusão escolar.

n > 3 para quê?

- Novas dimensões em física;
- dados amostrais ou populacionais;
- coeficientes de polinômios ou funções a estimar;
- ? decomposição de ondas, como em $\sum_{k=0}^{\infty} a_k \cos(2\pi kt)$;
- insumos diversos a otimizar...

Cuidados com notação

Vetores também podem ser indexados: será o caso da base canônica dos vetores e_j, que não são "componentes" de um "vetor e".

Usamos letras gregas ou latinas para identificar escalares.

Não usamos flechas, barras ou negrito para identificar vetores. (a em vez de $\vec{a}$, $\bar{a}$ (alguns autores), $\mathbf{a}$.)

Matrizes

Tópicos

- Revisão das operações básicas.
- Relação com sistemas.
- Escalonamento e eliminação gaussiana.

Revisão

Você já sabe:

- o que são matrizes?
- como somar matrizes?
- como multiplicar matrizes?
- por que a soma e o produto são definidos assim?

Uma rede de lojas recolhe, ao final do mês, uma planilha de cada filial, com o número de peças vendidas de cada artigo (por linha) em cada dia do mês (coluna).

Como se obtém a planilha das vendas realizadas por toda a rede?

Uma confecção tem duas planilhas:

Uma contém a quantidade produzida de cada modelo de peça de roupa (por coluna) em cada mês (linha).

Outra contém as quantidades de botões, zíperes, linhas, tecidos (colunas) utilizadas em cada peça (por linha).

Como se obtém a planilha dos totais de materiais utilizados a cada mês?

Notação

O *conjunto de matrizes com m linhas e n colunas cujas entradas estão todas em* $\mathbb{K}$ é indicado

$$M_{m\times n}(\mathbb{K}).$$

Alguns autores escrevem $\mathbb{K}^{m\times n}$, já que essas matrizes consistem de mn escalares que podem ser rearranjados como vetores.

Na matriz A, sua entrada na i-ésima linha e j-ésima coluna é identificada como A_{ij}.

Vetores de $\mathbb{K}^n$ são geralmente vistos como matrizes linha ou coluna.

Na matriz

$$\begin{bmatrix} 5 & & * \\ & * & 1 \\ & & * \end{bmatrix}:$$

- entradas não identificadas são 0;
- entradas identificadas com $*$ são qualquer escalar, inclusive 0, e podem ser distintas entre si.

Comparações entre matrizes

Suponha $A, B \in M_{m\times n}(\mathbb{K})$ (podem ser vetores também). Dizemos:

- $A \leqslant B$ se $A_{ij} \leqslant B_{ij}$ para cada i, j;
- $A > B$ se $A_{ij} > B_{ij}$ para cada i, j;
- $A \geqslant 0$ se $A_{ij} \geqslant 0$ para cada i, j etc.

Esse procedimento "entrada a entrada" será muito utilizado.

Sistemas lineares e representação

Lembre que existem três tipos de sistema:

SPD: "Sistema Possível (ou consistente ou compatível) e Determinado"; tem uma solução e ela é única.

SPI: "Sistema Possível e Indeterminado"; tem mais de uma solução.

SI: "Sistema Impossível (ou inconsistente ou incompatível)"; não tem solução.

Lembre que uma *solução* consiste de um valor para cada incógnita escalar, ou seja, é um vetor.

Considere por exemplo:

$$\begin{cases} 5x - 3y + 2z = 4 \\ 2x \quad\quad - \ z = 1 \\ -x + \ y + \ z = 1 \end{cases} \quad \text{e} \quad \begin{cases} 5x - 3y + 2z = 7 \\ 2x \quad\quad - \ z = 3i \\ -x + \ y + \ z = 0 \end{cases}$$

Os lados esquerdos são as mesmas combinações de x, y, z.

Os dois sistemas são simultaneamente representados pela matriz:

$$\begin{array}{r} 1^{\underline{a}} \text{ equação} \Rightarrow \\ 2^{\underline{a}} \text{ equação} \Rightarrow \\ 3^{\underline{a}} \text{ equação} \Rightarrow \\ \\ \end{array} \begin{array}{c} \left[\begin{array}{rrr|rr} 5 & -3 & 2 & 4 & 7 \\ 2 & 0 & -1 & 1 & 3i \\ -1 & 1 & 1 & 1 & 0 \end{array}\right] \\ \begin{array}{ccc} \Downarrow & \Downarrow & \Downarrow \\ x & y & z \end{array} \end{array}$$

Escalonamento (em linhas)

Diversos tópicos utilizarão matrizes ou poderão ser formulados com o uso de sistemas lineares.

O escalonamento de matrizes será uma ferramenta comum a todos.

Trabalharemos em geral: $A \in M_{m \times n}(\mathbb{K})$ com $m \neq n$ ou $m = n$.

Matrizes (com linhas) escalonadas

Objetivo: usar as operações elementares para pôr uma matriz nas especificações a seguir, que facilitarão cálculos:

(1) As linhas inteiramente nulas estão abaixo de todas as não nulas.

(2) O *pivô* de cada linha — primeiro elemento não nulo a partir da esquerda — está estritamente à direita dos pivôs das linhas acima.

(3) Todas as entradas abaixo dos pivôs são nulas.

(4) Os pivôs valem todos 1.

(5) Todas as entradas acima dos pivôs são nulas.

Exemplos e nomenclatura

$$\begin{bmatrix} 3 & * \\ 0 & 5 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 5 & * & * \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 3 & * & * \\ 0 & 0 & 1 & * \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

todas satisfazem (1)–(3) e são chamadas *escalonadas* (ou "com linhas escalonadas") por outros autores.

$$\begin{bmatrix} 1 & * \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & * & * \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & * & * \\ 0 & 0 & 1 & * \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

todas satisfazem (1)–(4) e são chamadas *escalonadas* (ou "com linhas escalonadas") por alguns autores *e neste curso.*

$$\begin{bmatrix}1 & 0\\0 & 1\\0 & 0\end{bmatrix}, \quad \begin{bmatrix}1 & * & 0\\0 & 0 & 1\\0 & 0 & 0\end{bmatrix}, \quad \begin{bmatrix}0 & 1 & * & *\\0 & 0 & 1 & 0\\0 & 0 & 0 & 1\end{bmatrix}$$

todas satisfazem (1)–(5) e são chamadas *escalonadas reduzidas* (ou "com linhas escalonadas na forma reduzida").

Notas

Uma matriz pode ter várias formas escalonadas.

Uma matriz somente tem uma forma escalonada reduzida.

Há vários modos de escalonar, levando a resultados diferentes, incluindo "truques" com sistemas pequenos (isolar uma variável e substituir etc.) ou fazendo escolhas diferentes na eliminação gaussiana.

Eliminação gaussiana

Algoritmo específico para escalonar.

Para i de 1 a m, faça:

(1º) Determine a primeira coluna não nula (entre linhas i e m) a partir da esquerda. (Interrompa o laço se não houver; acontecerá para $m > n$ e pode parar para $i = m$.)

(2º) Escolha para pivô uma entrada não nula dessa coluna (entre linhas i e m). (Recomenda-se a entrada com maior valor absoluto, para estabilidade numérica.)

(3º) Permute a i-ésima linha com a do pivô.

(4º) Divida a nova i-ésima linha pelo valor do pivô.

(5º) Some (ou subtraia) múltiplos da (nova) i-ésima linha às linhas abaixo, de modo a anular o restante da coluna abaixo do pivô. (Cuidado se fizer programação: armazene o fator ou deixe para anular essa entrada depois de operar as demais.)

Adição reversa (back addition)

Após o laço descrito acima, a partir da linha inferior e subindo, some (ou subtraia) múltiplos de cada linha não nula às linhas acima, de modo a anular o restante da coluna também acima do pivô.

A matriz fica escalonada reduzida.

Espaços vetoriais

Tópicos

- Axiomática e álgebra simbólica.
- Definições e exemplos.
- Subespaços e operações entre eles.
- Geração em espaços vetoriais.
- A combinação linear como operação fundamental.

Espaços vetoriais

Objetivo: identificar propriedades dos espaços euclideanos $\mathbb{R}^n$ que funcionam em outros contextos.

Propriedades são os axiomas e suas consequências.

Novos exemplos são as estruturas em que valem os axiomas.

Funcionamento da axiomática

Demonstramos novos fatos usando apenas os axiomas, não tudo o que sabemos dos espaços $\mathbb{R}^n$.

Esses novos fatos valerão também em todos os outros exemplos.*

*O campo das Probabilidades também é um exemplo, em que propriedades como a aditividade são axiomas para cálculos diversos.

O que são estruturas

$\mathbb{R}$, $\mathbb{C}$ já vêm munidos de "operações naturais".

Distinguiremos, em uma estrutura $(X, +, \times, \ldots)$, o *domínio* X (conjunto-universo não vazio) e as operações $+, \times, \ldots$

Também costumamos chamar a estrutura de X...

... mas há mais de uma estrutura com domínio X!

Nos espaços euclideanos $\mathbb{R}^n$, há duas entidades: *vetores* e *escalares.*

- Não usaremos flecha nem negrito para vetores.
- Usaremos letras gregas e romanas para escalares.

Esteja atento ao contexto!

Sabemos *adicionar* vetores; é uma operação binária:

$$\begin{aligned} +\colon \mathbb{R}^n \times \mathbb{R}^n &\to \mathbb{R}^n \\ (x, y) &\mapsto x + y \end{aligned}$$

Sabemos *multiplicar* um vetor por um escalar; é uma operação binária:

$$\begin{aligned} \cdot\colon \mathbb{R} \times \mathbb{R}^n &\to \mathbb{R}^n \\ (\lambda, x) &\mapsto \lambda \cdot x \end{aligned}$$

Em geral, trabalharemos com:

- um conjunto de vetores $V \neq \emptyset$;
- um corpo de escalares $\mathbb{K}$, leia-se, $\mathbb{R}$ ou $\mathbb{C}$.

Devem ser especificados:

$$+ \colon V \times V \to V$$
$$\cdot \colon \mathbb{K} \times V \to V$$

Para que V seja um *espaço vetorial sobre* $\mathbb{K}$ com tais operações (ou $\mathbb{K}$-*espaço*), os seguintes *axiomas* devem ser verificados (para quaisquer $x, y, z \in V$ e $a, b \in \mathbb{K}$):

- a adição é comutativa: $x + y = y + x$;
- a adição é associativa: $(x + y) + z = x + (y + z)$;
- a adição tem elemento neutro: existe 0 tal que $x + 0 = x$;
- a adição tem opostos: dado x, existe $-x$ tal que $x + (-x) = 0$;

- a multiplicação é associativa com aquela de $\mathbb{K}$: $a(bx) = (ab)x$;
- a multiplicação distribui-se sobre a adição em V: $a(x + y) = ax + ay$;
- a multiplicação distribui-se sobre a adição em $\mathbb{K}$: $(a + b)x = ax + bx$;
- $1 \in \mathbb{K}$ é elemento neutro também desta multiplicação: $1x = x$.

A partir dessas propriedades, podemos mostrar outras (exercícios).

Por exemplo, $\mathbb{K}$ também tem um elemento 0, que não é $0 \in V$.

Mostre que $0x = 0$.

Solução

Note $0x = (0+0)x = 0x + 0x$:

- a primeira igualdade vem de propriedade em $\mathbb{K}$;
- a segunda vem da distributividade.

Agora, $0x$ é um elemento de V, então tem oposto $-(0x)$. Some-o a ambos os lados:

$$0x + [-(0x)] = 0x + 0x + [-(0x)]$$

Então $0 = 0x + 0 = 0x$.

Note ainda:

- os axiomas não mencionavam essa propriedade;
- mas foram “fortes” para prová-la;
- ela indica como um objeto da adição se comporta com respeito à multiplicação (outra operação).

Exercícios similares

Mostre que $\lambda \cdot 0 = 0$ para qualquer escalar λ.

Dados $x \in V$ e $-1 \in \mathbb{K}$, mostre que $(-1)x = -x$, oposto de x.

Dados $x, y, z \in V$ e $a, b \in \mathbb{K}$, valem as “leis do cancelamento”:

- $x + z = y + z \Rightarrow x = y$;
- $ax = ay \;\&\; a \neq 0 \Rightarrow x = y$;
- $ax = bx \;\&\; x \neq 0 \Rightarrow a = b$.

(Note: já conhecemos essas propriedades nos espaços euclideanos; pede-se a demonstração a partir dos axiomas ou outras propriedades já demonstradas.)

Exemplos de espaços vetoriais

Cada $\mathbb{R}^n$ é espaço sobre $\mathbb{R}$ com as operações usuais (coordenada a coordenada).

Analogamente, cada $\mathbb{C}^n$ é espaço sobre $\mathbb{C}$ com operações coordenada a coordenada.

Fixadas as dimensões m e n, o conjunto de matrizes $M_{m\times n}(\mathbb{K})$ é espaço sobre $\mathbb{K}$ com as operações usuais.

O conjunto de todas as sequências

$$\mathrm{Seq}(\mathbb{K}) = \{\, (x_0, x_1, x_2, \ldots) \mid x_i \in \mathbb{K} \,\}$$

é espaço sobre $\mathbb{K}$ com operações entrada a entrada.

Seja D um conjunto não vazio. Então o conjunto de todas as funções $D \to \mathbb{K}$

$$\mathbb{K}^D = \{\, f \colon D \to \mathbb{K} \mid f \text{ qualquer} \,\}$$

é espaço sobre $\mathbb{K}$ com operações ponto a ponto:

- $f + g$ é definida por $(f + g)(x) = f(x) + g(x)$;
- λf é definida por $(\lambda f)(x) = \lambda\big(f(x)\big)$.

Exercício: Determine conjuntos D de modo que os exemplos anteriores sejam vistos na forma $\mathbb{K}^D$.

Com suas próprias operações, $\mathbb{K}$ é um espaço vetorial sobre $\mathbb{K}$.

Todo $\mathbb{C}$-espaço é naturalmente um $\mathbb{R}$-espaço, porque "multiplicação por complexo" inclui "multiplicação por real", com as mesmas propriedades.

Do mesmo modo, todo $\mathbb{R}$-espaço é um $\mathbb{Q}$-espaço.

Concluímos que $\mathbb{C}$ é espaço sobre $\mathbb{R}$ e sobre $\mathbb{Q}$; que $\mathbb{R}$ é espaço sobre $\mathbb{Q}$ (este é muito interessante).

Tanto em $\mathbb{C} \cong \mathbb{R}^2$ como em $\mathbb{R}$ sabemos o que é "convergir à origem". Então

$$\mathrm{SeqConv}(\mathbb{K}) = \{\, x \in \mathrm{Seq}(\mathbb{K}) \mid \lim x = 0 \,\}$$

é espaço sobre $\mathbb{K}$ com operações coordenada a coordenada.

De fato, pelas regras de limite, soma e produto de funções que convergem a 0 também convergem a 0.

Também sabemos o que é "tamanho" (módulo) em $\mathbb{C}$ e $\mathbb{R}$. Então

$$\mathrm{SeqLimitadas}(\mathbb{K}) = \{\, x \in \mathrm{Seq}(\mathbb{K}) \mid (\exists M > 0)(\forall i)\, |x_i| < M \,\}$$

é espaço sobre $\mathbb{K}$ etc.

De fato, $|x_i| < M$ e $|y_j| < N$ implicam $|x_i + y_i| \leqslant |x_i| + |y_i| < M + N$ e $|\lambda x_i| \leqslant |\lambda| \cdot |x_i|$.

Também com operações ponto a ponto, para $D \neq \emptyset$:

$$\mathbb{K}^{(D)} = \{\, f\colon D \to \mathbb{K} \mid f(x) \neq 0 \text{ somente para nº finito de } x \in D \,\}$$

Com as operações usuais de polinômios, são espaços sobre $\mathbb{K}$:

- $\mathbb{K}[t] = \{\, P(t) = a_0 + a_1 t + \ldots + a_n t^n \mid n \in \mathbb{N} \text{ e } a_i \in \mathbb{K} \,\}$;
- $\mathbb{K}[t]^{\geqslant n} = \{\, P(t) \in \mathbb{K}[t] \mid \mathrm{grau}(P) \leqslant n \,\}$.

O conjunto vazio $\emptyset$ não pode ser espaço vetorial, porque não contém um elemento neutro.

O conjunto unitário $\{0\}$ é um $\mathbb{K}$-espaço com as operações $0+0=0$ e $\lambda 0 = 0$. Ele é chamado 0.

Também com operações ponto a ponto, para um intervalo qualquer $I \subseteq \mathbb{R}$:

- $C^0(I) = \{\, f\colon I \to \mathbb{R} \mid f \text{ é contínua} \,\}$;
- $C^k(I) = \{\, f\colon I \to \mathbb{R} \mid f \text{ é de classe } C^k \,\}$;
- $L(I) = \{\, f\colon I \to \mathbb{R} \mid f \text{ é integrável sobre } I \,\}$

são todos $\mathbb{R}$-espaços.

Lembrete: uma função é C^k se tem derivada de ordem k e essa derivada é contínua.

Também com as operações usuais, fixada $M \in M_{m \times n}(\mathbb{K})$:

$$\mathrm{SolHomog}(M) = \{\, x \in \mathbb{K}^n \mid Mx = 0 \,\},$$

em que x é visto como vetor coluna, é $\mathbb{K}$-espaço.

Também com operações ponto a ponto, é um $\mathbb{R}$-espaço o conjunto das soluções da EDO linear homogênea

$$p_n(x) \cdot y^{(n)} + \ldots + p_2(x) \cdot y'' + p_1(x) \cdot y' + p_0(x) \cdot y = 0,$$

sendo $p_0, \ldots, p_n$ contínuas e p_n sem zeros.

Há vários outros exemplos!

A álgebra linear permite construir um corpo de conhecimento útil a todos eles simultaneamente!

É preciso verificar que cada exemplo é, de fato, um espaço vetorial.

Isso requer conferir os oito axiomas para cada caso. (Não faremos isso neste curso.)

Alguns exemplos são mais simples, porque podem ser construídos a partir de outros, de vários modos que veremos.

A primeira construção será a de *subespaço.*

Subespaços

Definição

Sejam V espaço vetorial sobre $\mathbb{K}$ (operações implícitas, mas fixadas) e $W \subseteq V$. Dizemos que W é *subespaço* de V se satisfaz três condições:

(i) $W \neq \emptyset$;

(ii) $x, y \in W \Rightarrow x + y \in W$, para quaisquer x, y;

(iii) $\lambda \in \mathbb{K}$ & $x \in W \Rightarrow \lambda x \in W$, para quaisquer λ, x.

Nesse caso, W também é um espaço vetorial sobre $\mathbb{K}$, com as mesmas operações de V restritas a W, porque:

(a) W é "fechado sob as operações", segundo (ii) e (iii), isto é,

$$+ \colon W \times W \to W$$
$$\cdot \colon \mathbb{K} \times W \to W$$

(b) as operações satisfazem os axiomas, porque já satisfazem em V.

Esse é um modo de verificar se uma estrutura é um espaço vetorial.

Exemplo

$C^0(I)$ é subespaço de $\mathbb{R}^I$.

Demonstração: Assumimos que R^I é espaço vetorial e mostramos:

(i) $C^0(I) \neq \emptyset$ porque contém a função constante 0, que é contínua.
Mostrar $0 \in W$ é o modo mais fácil de obter $W \neq \emptyset$; todo subespaço deve conter o elemento neutro.

(ii) Se f, g são funções contínuas, então $f + g$ também é contínua.

(iii) Se f é contínua e $\lambda \in \mathbb{R}$, então λf também é contínua.

Mais exemplos

Os subespaços de $\mathbb{R}^2$ são as retas passando pela origem, *mais* o subespaço 0 (é $\{0\}$) e o espaço todo $\mathbb{R}^2$.

Os subespaços de $\mathbb{R}^3$ são: 0, as retas passando pela origem, os planos pela origem e todo o $\mathbb{R}^3$.

Exercício: Determine, dentre os exemplos de espaços vetoriais já dados, quais são subespaços uns dos outros.

Operações entre subespaços

Dados W_1 e W_2 subespaços de V, então

$$W_1 \cap W_2 = \{\, x \in V \mid x \in W_1 \;\&\; x \in W_2 \,\}$$
$$\text{e } W_1 + W_2 = \{\, x + y \in V \mid x \in W_1 \;\&\; y \in W_2 \,\}$$

também são subespaços de V.

Soma direta

Usa-se $W_1 \oplus W_2$ para $W_1 + W_2$ *quando* $W_1 \cap W_2 = 0$.

Demonstração para $W_1 \cap W_2$

Para (i), como $0 \in W_1$ e $0 \in W_2$, temos $0 \in W_1 \cap W_2$.

Para (ii) e (iii), sejam $x, y \in W_1 \cap W_2$ e $\lambda \in \mathbb{K}$. Então: $x, y \in W_1$ & $x, y \in W_2$, de modo que $x + y, \lambda x \in W_1$, e $x + y, \lambda x \in W_2$, donde $x + y, \lambda x \in W_1 \cap W_2$.

Demonstração para $W_1 + W_2$ (extra)

Para (i), como $0 \in W_1$ e $0 \in W_2$, temos $0 = 0 + 0 \in W_1 + W_2$.

Para (iii), fixe $\lambda \in \mathbb{K}$. Um elemento arbitrário de $W_1 + W_2$ tem a forma $x + y$ com $x \in W_1$ e $y \in W_2$. Então $\lambda(x + y) = \lambda x + \lambda y \in W_1 + W_2$.

Para (ii), dados dois elementos de $W_1 + W_2$, escreva-os como $x + y$ e $u + v$ com $x, u \in W_1$ e $y, v \in W_2$. Então:

$$(x + y) + (u + v) = (x + u) + (y + v) \in W_1 + W_2$$

Inclusões

$$0 \quad \subseteq \quad W_1 \cap W_2 \quad \subseteq \quad W_1, \quad W_2 \quad \subseteq \quad W_1 + W_2 \quad \subseteq \quad V$$

Exercício

Todo elemento de $W_1 + W_2$ pode ser decomposto como $x + y$ com $x \in W_1$ e $y \in W_2$, pela definição de $W_1 + W_2$. Mostre que se $W_1 \cap W_2 = 0$ então a decomposição é única.

Solução

Escreva $x + y = x' + y'$ com $x, x' \in W_1$ e $y, y' \in W_2$.

Então:

$$\underbrace{x - x'}_{\in W_1} = \underbrace{y' - y}_{\in W_2} \in W_1 \cap W_2 = \{0\}$$

Desse modo, $x - x' = 0$ e $y' - y = 0$, donde $x = x'$ e $y = y'$.

Subespaço gerado (definição)

Sejam V espaço vetorial sobre $\mathbb{K}$ e $S \subseteq V$. O subespaço gerado por S é

$$[S] = \{\, \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n \mid n \in \mathbb{N},\ \lambda_i \in \mathbb{K},\ x_i \in S \,\}.$$

Note que a soma é *finita*, mas o nº de elementos (pode repetir) é ilimitado.

Exercícios

- $[S]$ é mesmo subespaço de V.
- Se $S_1 \subseteq S_2$ então $[S_1] \subseteq [S_2]$.

Observação

Se $S = \emptyset$ então $[S] = 0$: é uma *convenção* que uma soma de zero elementos, ou soma vazia, vale 0.

Pode-se mostrar que $[S]$ é, no sentido de inclusão ($\subseteq$), o *menor* subespaço que contém S.

Suponha que $S = \{v_1, v_2, \ldots, v_k\}$ (ou seja, S é finito com k elementos). Então:

$$\begin{aligned}[S] &\overset{\text{resultado}}{=\!=\!=} \{\, a_1 v_1 + a_2 v_2 + \ldots + a_k v_k \mid a_i \in \mathbb{K} \,\} \\ &\overset{\text{notação}}{=\!=\!=} \mathbb{K}v_1 + \mathbb{K}v_2 + \ldots + \mathbb{K}v_k\end{aligned}$$

Dados $x, v_1, \ldots, v_k \in \mathbb{K}^m$, uma questão importante é: $x \in [v_1, \ldots, v_k]$?

$$\begin{aligned}
\text{sim} &\Leftrightarrow x \in \mathbb{K}v_1 + \mathbb{K}v_2 + \ldots + \mathbb{K}v_k \\
&\Leftrightarrow \exists \lambda_i \in \mathbb{K} \text{ tais que } x = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_k v_k \\
&\Leftrightarrow \exists \vec{\lambda} \in \mathbb{K}^k \text{ tal que } \underbrace{\begin{bmatrix} v_{11} & v_{21} & \cdots & v_{k1} \\ v_{12} & v_{22} & \cdots & v_{k2} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}}_{\text{matriz } m \times k \text{ sobre } \mathbb{K}} \vec{\lambda} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}
\end{aligned}$$

Note que as *colunas* da matriz são os vetores $v_1, v_2, \ldots, v_k$.

Por isso, estudaremos sistemas lineares diversos.

Combinações lineares

A combinação linear é uma síntese das operações disponíveis em um espaço vetorial:

- sabemos somar (associativa e comutativa), então podemos somar um nº finito de vezes por repetição;
- sabemos multiplicar por escalar (distributiva etc.).

Qualquer outra operação em um espaço vetorial é uma combinação linear, por exemplo: $\lambda(x+y)-z$ é $\lambda x+\lambda y+(-1)z$.

Então:

Definição

Uma combinação linear de elementos de $S \subseteq V$ é

$$\lambda_1 x_1 + \ldots + \lambda_n x_n \text{ com } n \in \mathbb{N},\ \lambda_i \in \mathbb{K} \text{ e } x_i \in S$$

(os x_i repetidos podem ser unificados).

Note que n é finito, mas ilimitado.

Não sabemos somar um nº infinito de termos, mesmo se S for infinito!

Conclusão

O subespaço gerado $[S]$ é o conjunto de todas as combinações lineares possíveis com elementos de S.

Se S for finito, como observamos para $[S]$, toda combinação linear pode ser escrita listando *todos* os elementos de S — multiplicando alguns por $0 \in \mathbb{K}$ se necessário.

Exercício

Mostre que:

- $y_1, \dots, y_k \in [S] \Rightarrow [y_1, \dots, y_k] \subseteq [S]$;
- $y \in [S] \Rightarrow [S \cup \{y\}] = [S]$.

Em palavras: combinação linear de combinações lineares de S é também combinação linear de S.

Bases e dimensão

Tópicos

- Dependência em espaços vetoriais.
- Bases e dimensão.

Dependência e independência lineares

Objetivo: eliminar redundâncias em conjuntos geradores de subespaços (identificar o conceito de "base").

Definição

Sejam V espaço vetorial sobre $\mathbb{K}$ e $S \subseteq V$. Então S é *linearmente dependente* (LD) se há combinação linear nula, mas não trivial, de elementos de S *sem repetição*, isto é:

$$(\exists n \in \mathbb{N})(\exists \lambda_i \in \mathbb{K} \text{ não todos nulos})(\exists x_i \in S) \sum_{i=1}^{n} \lambda_i x_i = 0$$

Caso contrário, S é *linearmente independente* (LI).

Notas

Sempre há a combinação nula trivial: $0x_1 + 0x_2 + \ldots + 0x_n = 0$.

Na combinação linear não trivial, os *escalares* não podem ser todos nulos, mas os vetores não são especificados. Se $0 \in S$, então S é LD.

S é LI $\Leftrightarrow$ a única combinação linear nula é a trivial.

Se S é LI e $S_1 \subseteq S$, então S_1 é LI.

Quando S é finito, pode-se formular a condição com x_i listando todos os elementos de S *sem repetição.*

S é LI $\Leftrightarrow$ *todo* subconjunto finito de S é LI.

S é LD $\Leftrightarrow$ *algum* subconjunto finito de S é LD.

Assim, podemos verificar a dependência linear de conjuntos arbitrários.

S é LD $\Leftrightarrow$ algum elemento de S é combinação linear dos demais (ou: gerado pelos demais).

Essa é a operação tradicional de "isolar": digamos

$$\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_j x_j + \ldots + \lambda_n x_n = 0 \text{ com } \lambda_j \neq 0.$$

Então

$$x_j = -\frac{\lambda_1}{\lambda_j}\, x_1 - \frac{\lambda_2}{\lambda_j}\, x_2 - \ldots - \frac{\lambda_n}{\lambda_j}\, x_n$$

sem listar o j-ésimo termo.

- $\emptyset$ é LI.
- 0 é LD.
- $\{x\}$ é LI $\Leftrightarrow x \neq 0$.
- $\{x, y\}$ é LD $\Leftrightarrow x = \lambda y$ *ou* $y = \lambda x$ (precisa duas hipóteses caso $\lambda = 0$).

Em $\mathbb{K}^m$, decidir se $\{v_1, \ldots, v_k\}$ é LD ou LI consiste em resolver o sistema homogêneo

$$\begin{bmatrix} v_1 & v_2 & \cdots & v_k \end{bmatrix} \vec{\lambda} = 0$$

(vetores nas colunas) correspondente à questão $0 \in [v_1, \ldots, v_k]$:

- se a única solução for $\vec{\lambda} = 0$, então é LI;
- caso contrário, é LD.

(Cf. discussão sobre subespaços gerados; veremos *determinante* como ferramenta.)

Bases

Objetivo: descrever vetores univocamente e em termos do próprio espaço.

Definição

Sejam V espaço vetorial sobre $\mathbb{K}$ e $B \subseteq V$. Diz-se que B é *base* de V se $[B] = V$ e B é LI.

Proposição

B é base de V se e somente se todo $x \in V$ expressa-se de modo *único* como combinação linear de elementos de B.

(Entende-se unicidade dos vetores e dos escalares envolvidos, quando os vetores são usados sem repetição.)

Demonstração

($\Leftarrow$) Assuma B base e

$$x = a_1 u_1 + a_2 u_2 + \ldots + a_m u_m$$
$$x = b_1 v_1 + b_2 v_2 + \ldots + b_n v_n$$

com $a_i, b_j \in \mathbb{K}$, $u_i, v_j \in B$ e sem repetições entre os u_i, nem entre os v_j.

Suponha (reindexando) que $u_1 = v_1$, $u_2 = v_2, \ldots,$ $u_k = v_k$ e os demais $u_{k+1}, \ldots, u_m, v_{k+1}, \ldots, v_n$ sejam todos distintos entre si e dos k primeiros.

Então:

$$
\begin{aligned}
0 = x - x = (a_1 - b_1)u_1 + (a_2 - b_2) + \ldots + (a_k - b_k)u_k + \\
+ a_{k+1}u_{k+1} + \ldots + a_m u_m + \\
+ (-b_{k+1})v_{k+1} + \cdots + (-b_n)v_n
\end{aligned}
$$

Como B é LI, temos:

$$a_1 - b_1 = a_2 - b_2 = \ldots = a_k - b_k = 0$$

e

$$a_{k+1} = \ldots = a_m = b_{k+1} = \ldots = b_n = 0$$

Então *ambas* as decomposições de x são iguais a

$$x = a_1 u_1 + \ldots + a_k u_k$$

com cada $a_i = b_i$.

x pôde ser qualquer porque B gera V.

($\Rightarrow$) Assuma que *todo* vetor expressa-se de modo *único* sobre B. Então B gera V (usando "todo") e é LI (usando "único", porque então a única combinação linear nula é a trivial).

Exemplos

A *base canônica* de $\mathbb{K}^n$ é $\{e_1, \ldots, e_n\}$ onde

$$e_i = (0, \ldots, 0, \underset{(i^{\underline{o}})}{1}, 0, \ldots, 0) = (\delta_{ij})_{1 \leqslant j \leqslant n}.$$

(O *delta de Kronecker* vale 1 quando os dois índices são iguais e 0 caso contrário.)

Se B tem número finito de elementos, digamos $B = \{v_1, \ldots, v_n\}$ *sem repetições*, então podemos escrever, para $x = a_1 v_1 + \ldots + a_n v_n$:

$$x_B = (a_1, \ldots, a_n) \in \mathbb{K}^n$$

(são as *coordenadas de* x *na base* B; note o subscrito B).

Isso pressupõe uma *ordem* específica para listar os elementos de B.

Note que

$$(v_i)_B = (0, \ldots, \underset{(i^{\text{o}})}{1}, \ldots, 0) = e_i \in \mathbb{K}^n.$$

Nós faremos uso de três teoremas importantes.

Depois veremos suas demonstrações somente no caso de dimensão finita.

("Dimensão" é definida pelo segundo teorema.)

Como sempre, seja V um espaço vetorial sobre $\mathbb{K}$.

Teorema 1

Sejam $S \subseteq G \subseteq V$ com S LI e $[G] = V$. Então existe base B de V tal que $S \subseteq B \subseteq G$.

Em particular:

- todo conjunto LI estende-se a uma base (tome $G = V$);
- todo espaço vetorial tem base (tome $S = \emptyset$).

Teorema 2

Quaisquer bases de V têm o mesmo nº de elementos. Essa é a *dimensão* de V, indicada $\dim V$.

Por exemplo:

- $\dim \mathbb{K}^n = n$;
- $\dim 0 = 0$ ($\emptyset$ é base do espaço 0).

Usaremos $|S|$ para indicar o nº de elementos (*cardinalidade*) de um conjunto S.

Acima, $\dim V = |B|$.

Teorema 3

Para $S \subseteq V$, temos:

(a) se S é LI então $|S| \leqslant \dim V$;

(b) se $|S| > \dim V$ então S é LD;

(c) se $[S] = V$ então $|S| \geqslant V$;

(d) se $|S| < \dim V$ então S não gera V.

Alguns exemplos...

... de técnicas em demonstrações.

Sejam W_1, W_2 subespaços de V com $W_1 \cap W_2 = 0$. Então

$$\dim(W_1 \oplus W_2) = \dim W_1 + \dim W_2.$$

Faremos isso mostrando que, para B_1, B_2 bases de W_1, W_2 resp., então $B_1 \cup B_2$ é base de $W_1 \oplus W_2$ e $B_1 \cap B_2 = \emptyset$.

...

Mais desafiador é mostrar, para W_1, W_2 subespaços de dimensão finita, que

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2).$$

Aqui, não há garantia *a priori* que $B_1 \cap B_2$ seja base de $W_1 \cap W_2$.

De fato, $\{\vec{\imath}, \vec{\jmath}\}$ e $\{5\vec{\imath}, 2\vec{\imath} - \vec{\jmath}\}$ são bases *disjuntas* do mesmo espaço $\mathbb{R}^2$.

Experimente *começar* com base B_0 de $W_1 \cap W_2$ e *estendê-la* a bases B_1, B_2.

. . .

Completamento de base

Este método permite:

(I) obter uma base para um subespaço de $\mathbb{K}^n$;

(II) completar essa base para obter uma de $\mathbb{K}^n$.

Para um espaço vetorial V com base B de n elementos, podemos então trabalhar com as coordenadas em B, que é equiparada à base canônica de $\mathbb{K}^n$.

Para (I), dados $x_1, \ldots, x_m \in \mathbb{K}^n$:

(1º) Forme a matriz A com os vetores nas linhas:

$$A = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

(2º) Escalone A obtendo uma forma escalonada R_1.
O espaço gerado pelas linhas de R_1 é o mesmo de A, ou seja, $[x_1, \ldots, x_m]$.

(3º) Descarte as linhas nulas de R_1, obtendo R_2.
Como R_2 está escalonada, suas linhas são LI e formam a base desejada.

Para (II), acrescente novas linhas a R_2, assim: em cada coluna originalmente sem pivô, coloque 1, depois 0 nas demais entradas.

Obteve-se uma matriz $n \times n$, cujas linhas são todas LI e formam a base desejada.

Exemplo na lousa

Em $\mathbb{R}^5$, dados

$$\begin{aligned} x_1 &= (0, 3, 5, -1, 2), \\ x_2 &= (0, -2, 0, 8, 0), \\ x_3 &= (0, -4, 1, 0, 0), \end{aligned}$$

temos ...

Exercício

Em $\mathbb{R}^6$, dados

$$x_1 = (2, 0, 3, -1, 0, 0),$$
$$x_2 = (1, 1, 0, 1, 0, 1),$$
$$x_3 = (0, -2, 3, -3, 0, -2),$$
$$x_4 = (1, -1, 0, 0, 0, 0)$$

temos ...

Transformações lineares

Tópicos

- Homomorfismos e seu espaço.
- Subespaços núcleo e imagem.
- Isomorfismos.
- Matrizes associadas, aspecto funtorial e mudança de base.
- Subespaços linha e coluna. Posto e nulidade.
- Funcionais lineares e sua relação com hiperplanos.

Conceito

Objetivo final: permitir a descrição de um espaço vetorial em termos de outro.

Toda vez que se definem estruturas (como espaços vetoriais), verificam-se quais funções entre seus domínios *preservam* suas descrições, funções essas chamadas (*homo*)*morfismos*.

Definição

Sejam V, W espaços vetoriais sobre $\mathbb{K}$ e $T\colon V \to W$ uma função. Diz-se que T é uma *transformação linear* se:

- $T(x +_V y) = T(x) +_W T(y)$ para todos $x, y \in V$ e
- $T(\lambda \cdot_V x) = \lambda \cdot_W T(x)$ para todos $x \in V$ e $\lambda \in \mathbb{K}$.

(Os subscritos V, W ressaltam em quais espaços as operações são feitas.)

Note que, em tal caso, sempre $T(0_V) = 0_W$ — por quê?

Equivalentemente, T é transformação linear se "transporta" combinações lineares:

$$(\forall n \in \mathbb{N})(\forall a_i \in \mathbb{K})(\forall x_i \in V)\ T\left(\sum_{i=1}^{n} a_i x_i\right) = \sum_{i=1}^{n} a_i T(x_i).$$

Costuma-se verificar linearidade calculando $T(ax+by)$ ou $T(ax+y)$.

Exercício: verificar que os próximos exemplos são transformações lineares!

Exemplos

Dada uma matriz $A \in M_{m\times n}(\mathbb{K})$, defina

$$T_A\colon \mathbb{K}^n \to \mathbb{K}^m,\ \ T_A(x) = Ax.$$

Veremos que esse exemplo é arquetípico!

A *derivação* $C^{n+1}(I) \to C^n(I)$, $f \mapsto f'$.

A *valoração* $T_{x_0} \colon \mathbb{K}^D \to \mathbb{K}$ em um ponto fixo $x_0 \in D$ dada por $T_{x_0}(f) = f(x_0)$. (Note que o argumento de T_{x_0} é $f \colon D \to \mathbb{K}$.)

Composição de transformações lineares é linear.

Operadores “construção” e “destruição” T_1, T_2: $\mathrm{Seq}(\mathbb{K}) \to \mathrm{Seq}(\mathbb{K})$:

$$T_1(x_0, x_1, x_2, \ldots) = (0, x_0, x_1, x_2, \ldots)$$
$$\text{e } T_2(x_0, x_1, x_2, \ldots) = (x_1, x_2, \ldots)$$

Proposição

Sejam B base de V e $\varphi\colon B \to W$ função *qualquer.* Então existe única transformação linear $T\colon V \to W$ tal que $(\forall x \in B)\ T(x) = \varphi(x)$.

Ou seja: Para especificar uma transformação linear, basta indicar as imagens dos elementos de uma base do domínio, e isso é totalmente livre.

Demonstração...

Existência

Dado $x \in V$, obtenha $x = \sum \lambda_i x_i$ de modo único com $x_i \in B$. Defina $T(x) = \sum \lambda_i \varphi(x_i)$.

Exercício: mostre que T é linear e extensão de φ.

Unicidade

Suponha também T_1 como no enunciado.

Então $(\forall x \in B)\ T(x) = \varphi(x) = T_1(x)$.

Então também $(\forall x \in V)$

$$T(x) = \sum \lambda_i \varphi(x_i) = \sum \lambda_i T_1(x_i) = T_1\Big(\sum \lambda_i x_i\Big) = T_1(x).$$

Obtemos $T \equiv T_1$.

Proposição

O conjunto $L(V, W) = \{ T\colon V \to W \mid T \text{ é linear} \}$ com as operações de soma e produto por escalar "ponto a ponto" é um espaço vetorial.

Para demonstrar, basta verificar que é subespaço de W^V...

De fato:

$$
\begin{aligned}
(T_1 + T_2)(ax + by) &\overset{\text{definição}}{=\!=\!=} T_1(ax + by) + T_2(ax + by) \overset{\text{lineares}}{=\!=\!=} \\
&= aT_1(x) + bT_1(y) + aT_2(x) + bT_2(y) = \\
&= a(T_1 + T_2)(x) + b(T_1 + T_2)(y),
\end{aligned}
$$

de modo que $T_1 + T_2 \in L(V, W)$;

$$
\begin{aligned}
(\lambda T)(ax + by) &\overset{\text{definição}}{=\!=\!=} \lambda \cdot T(ax + by) \overset{T \text{ linear}}{=\!=\!=} \\
&= \lambda \cdot (aT(x) + bT(y)) = a(\lambda T)(x) + b(\lambda T)(y),
\end{aligned}
$$

de modo que $\lambda T \in L(V, W)$.

Núcleo e imagem

Definições

Dada $T : V \to W$ linear, definimos:

- o *núcleo* ou "*kernel*"

$$\mathrm{Nuc}(T) = \{\, x \in V \mid T(x) = 0 \,\};$$

- a *imagem*

$$\mathrm{Im}(T) = \{\, T(x) \mid x \in V \,\} = \{\, y \in W \mid (\exists x \in V)\, T(x) = y \,\}.$$

Exercícios

$\mathrm{Nuc}(T)$ é subespaço de V.

$\mathrm{Im}(T)$ é subespaço de W.

$\mathrm{Nuc}(T) = 0 \Leftrightarrow T$ é injetora.

$\dim V = \dim \mathrm{Nuc}(T) + \dim \mathrm{Im}(T)$.

Isomorfismo

De um modo geral, duas estruturas isomórficas (com isomorfismo entre elas) são reconhecidas como “exemplos” diferentes de uma *mesma* estrutura, ou seja, mudam-se apenas os *nomes* dos elementos e operações.

Definição

Sejam V, W espaços vetoriais sobre $\mathbb{K}$. Diz-se que $T: V \to W$ é *isomorfismo* se for bijetora e T, T^{-1} forem lineares.

Indica-se $V \cong W$ quando há isomorfismo entre V e W.

Proposição

Basta que T seja bijetora e linear.

Demonstração

Resta mostrar T^{-1} linear: sejam $u = T^{-1}(x)$ e $v = T^{-1}(y)$. Então:

$$T^{-1}(ax+by) = z \Leftrightarrow ax+by = T(z) \Leftrightarrow aT(u)+bT(v) = T(z) \Leftrightarrow$$
$$\Leftrightarrow T(au+bv) = T(z) \Leftrightarrow au+bv = z \Leftrightarrow aT^{-1}(x)+bT^{-1}(y) = z$$

donde $T^{-1}(ax+by) = aT^{-1}(x) + bT^{-1}(y)$.

Exemplo instrumental

Suponha V com base finita B, sendo $n = |B| = \dim V$: então

$$T\colon V \to \mathbb{K}^n,\ \ T(x) = x_B,$$

é isomorfismo ($\mathbb{K}^n$ com operações usuais).

(Esse espaço também é $\mathbb{K}^B$. Para dimensão arbitrária, é preciso considerar $\mathbb{K}^{(B)}$.)

Exercício longo

Uma cadeia de composição de transformações lineares

$$\cdots \xrightarrow{T_{-2}} V_{-1} \xrightarrow{T_{-1}} V_0 \xrightarrow{T_0} V_1 \xrightarrow{T_1} V_2 \xrightarrow{T_2} \cdots$$

é chamada *sequência exata* se $(\forall n)\ \mathrm{Im}(T_n) = \mathrm{Nuc}(T_{n+1})$.

Considere a sequência *curta*

$$0 \to U \xrightarrow{R} V \xrightarrow{T} W \to 0$$

onde U, V, W são espaços vetoriais sobre $\mathbb{K}$ e as quatro funções são lineares.

(a) As funções esquerda e direita não têm nomes porque estão univocamente determinadas. Quais são elas?

(b) Mostre que essa sequência é exata se e somente se três condições são satisfeitas:

  (i) R é injetora;

  (ii) $\mathrm{Im}(R) = \mathrm{Nuc}(T)$;

  (iii) T é sobrejetora.

(c) Assumindo exatidão, mostre $V = \mathrm{Im}(R) \oplus W_1$ onde $\mathrm{Im}(R) \overset{R}{\cong} U$ e $W_1 \overset{T}{\cong} W$.

(d) Identifique as transformações lineares que naturalmente deixam esta sequência exata:

$$0 \to U \to U \oplus W \to W \to 0.$$

Proposição

Para $T\colon V \to W$ linear e V, W de dimensões finitas:

(a) $\dim \operatorname{Im}(T) \leqslant \dim W$ por inclusão e $\dim \operatorname{Im}(T) \leqslant \dim V$ pelo teorema.

(b) T injetora $\Leftrightarrow \dim \operatorname{Im}(T) = \dim V$.

(c) T sobrejetora $\Leftrightarrow \dim \operatorname{Im}(T) = \dim W$.

Para provar, use o teorema $\dim V = \dim \operatorname{Nuc}(T) + \dim \operatorname{Im}(T)$.

Teorema

$V \cong W \Leftrightarrow \dim V = \dim W$.

Assim, a dimensão é toda a informação necessária para caracterizar o espaço vetorial "a menos de isomorfismo".

(É um "invariante completo" ou "característica de Euler".)

Demonstração

($\Rightarrow$) Segue de (b) e (c).

($\Leftarrow$) Sejam B_1, B_2 bases de V, W respectivamente. Então $|B_1| = |B_2|$ e há uma bijeção entre B_1 e B_2. Tome $T\colon V \to W$ a transformação linear estendendo essa bijeção. Então T também é bijetora:

Sobrejeção: Dada y combinação linear sobre B_2, substitua cada elemento de B_2 pelo correspondente de B_1, obtendo combinação linear x. Por definição, $T(x) = y$.

Injeção: Suponha $x \neq 0$, combinação linear não trivial de B_1, então $T(x)$ é combinação linear não trivial de B_2, então $T(x) \neq 0$.

Matrizes e transformações lineares

Veremos uma conexão profunda entre transformações lineares e matrizes, que respeita suas propriedades e operações.

Nesta seção, todos os espaços vetoriais têm dimensão *finita:*

- V com base $B_1 = \{v_1, \ldots, v_n\}$ e $\dim V = n$;
- W com base $B_2 = \{w_1, \ldots, w_m\}$ e $\dim W = m$.

Seja $T \colon V \to W$ linear.

Vimos que os vetores de V têm *coordenadas* na base B_1:

$$x \in V \Rightarrow x = x_1 v_1 + \ldots + x_n v_n \Rightarrow x_{B_1} = (x_1, \ldots, x_n) \in \mathbb{K}^n$$

O mesmo vale para $T(x)$ em W com respeito a B_2:

$$T(x) = y_1 w_1 + \ldots + y_m w_m \Rightarrow (T(x))_{B_2} = (y_1, \ldots, y_m) \in \mathbb{K}^m$$

Em particular, para cada v_j:

$$T(v_j) = a_{1j}w_1 + \ldots + a_{mj}w_m = \sum_{i=1}^{m} a_{ij}w_i$$

Logo:

$$T(x) = T\left(\sum_{j=1}^{n} x_j v_j\right) = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij}w_i = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}x_j\right)w_i$$

Conclusão:

$$(T(x))_{B_2} = \left(\sum_{j=1}^{n} a_{ij}x_j\right)_{1\leqslant i\leqslant m}$$

Então cada coordenada de $T(x)$ é obtida por uma multiplicação matricial:

$$\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Então $(T(x))_{B_2}$ como vetor coluna é o produto Ax_{B_1}, sendo x_{B_1} coluna e $A = (a_{ij})$.

Notação: $A = [T]_{B_1,B_2}$.

Note que a matriz depende das bases usadas!

Observação

$$[T]_{B_1,B_2} = \begin{bmatrix} & a_{1j} & \\ & a_{2j} & \\ \cdots & \vdots & \cdots \\ & a_{mj} & \end{bmatrix} = \begin{bmatrix} (T(v_1))_{B_2} & (T(v_2))_{B_2} & \cdots & (T(v_n))_{B_2} \end{bmatrix}$$

- A j-ésima coluna é $(T(v_j))_{B_2}$: imagem por T do j-ésimo vetor de B_1, escrita nas coordenadas de B_2.
- São $\dim W$ linhas e $\dim V$ colunas.

Conclusão

$$\big(T(x)\big)_{B_2} = [T]_{B_1,B_2} \cdot (x)_{B_1}$$

Teorema

Defina $\Phi\colon L(V,W) \to M_{m\times n}(\mathbb{K})$, $\Phi(T) = [T]_{B_1,B_2}$. Então, com as operações ponto a ponto em $L(V,W)$ e as usuais em $M_{m\times n}(\mathbb{K})$, a função Φ é um isomorfismo linear.

Além disso, se

$$U \xrightarrow{R} V \xrightarrow{T} W$$

com R linear e B_0 base de U, então

$$[T \circ R]_{B_0,B_2} = [T]_{B_1,B_2} \cdot [R]_{B_0,B_1}$$

com a multiplicação de matrizes.

Finalmente, T é bijetora se e somente se $[T]_{B_1,B_2}$ é invertível e, nesse caso,

$$[T^{-1}]_{B_2,B_1} = \left([T]_{B_1,B_2}\right)^{-1}.$$

Exemplo

$T\colon \mathbb{R}^2 \to \mathbb{R}^3$, $T(x,y) = (3y, 2x, 5y-4x)$ é linear.

Está descrita nas bases canônicas, então:

$$[T]_{\text{can},\text{can}} = \begin{bmatrix} 0 & 3 \\ 2 & 0 \\ -4 & 5 \end{bmatrix}$$

Para conferir, calcule $[T]_{\text{can},\text{can}} \cdot \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ (deverá obter $T(x,y)$).

Agora usaremos as bases

$$B_1 = \{(-3,0),(1,1)\}$$
$$\text{e } B_2 = \{(1,0,0),(0,2,0),(-1,-1,1)\}.$$

Precisamos calcular $(T(v_1))_{B_2}$ e $(T(v_2))_{B_2}$:

$$T(-3,0) = (0,-6,12) = a(1,0,0) + b(0,2,0) + c(-1,-1,1) \Rightarrow$$
$$\Rightarrow \begin{cases} a - c = 0 \\ 2b - c = -6 \\ c = 12 \end{cases} \Rightarrow \begin{cases} a = 12 \\ b = 3 \\ c = 12 \end{cases} \Rightarrow (T(-3,0))_{B_2} = (12,3,12).$$

$$T(1,1) = (3,2,1) = a(1,0,0) + b(0,2,0) + c(-1,-1,1) \Rightarrow$$

$$\Rightarrow \begin{cases} a - c = 3 \\ 2b - c = 2 \\ c = 1 \end{cases} \Rightarrow \begin{cases} a = 4 \\ b = 3/2 \\ c = 1 \end{cases} \Rightarrow (T(1,1))_{B_2} = (4, 3/2, 1).$$

Então:

$$[T]_{B_1,B_2} = \begin{bmatrix} 12 & 4 \\ 3 & 3/2 \\ 12 & 1 \end{bmatrix}$$

Vamos conferir na própria base B_1:

$$[T]_{B_1,B_2}\cdot(v_1)_{B_1} = \begin{bmatrix} 12 & 4 \\ 3 & 3/2 \\ 12 & 1 \end{bmatrix}\cdot\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 12 \\ 3 \\ 12 \end{bmatrix}$$

$$[T]_{B_1,B_2}\cdot(v_2)_{B_1} = \begin{bmatrix} 12 & 4 \\ 3 & 3/2 \\ 12 & 1 \end{bmatrix}\cdot\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 3/2 \\ 1 \end{bmatrix}$$

Note que os vetores de B_1, expressos na própria B_1, são os $\vec{\imath},\vec{\jmath}$ canônicos de $\mathbb{R}^2$.

E se quisermos usá-la para calcular T no vetor $(6,7) \in \mathbb{R}^2$? Ele está na base canônica:

$$T(6,7) = (21,12,11) \text{ também na base canônica}$$

Nas bases dadas:

$$(6,7) = \alpha(-3,0) + \beta(1,1) \Rightarrow \begin{cases} -3\alpha + \beta = 6 \\ \beta = 7 \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} \alpha = 1/3 \\ \beta = 7 \end{cases} \Rightarrow (6,7)_{B_1} = (1/3,7).$$

Então:

$$[T]_{B_1,B_2} \cdot (6,7)_{B_1} = \begin{bmatrix} 12 & 4 \\ 3 & 3/2 \\ 12 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1/3 \\ 7 \end{bmatrix} = \begin{bmatrix} 32 \\ 23/2 \\ 11 \end{bmatrix}$$

Conferindo:

$$(21,12,11) = a(1,0,0) + b(0,2,0) + c(-1,-1,1) \Rightarrow$$

$$\Rightarrow \begin{cases} a - c = 21 \\ 2b - c = 12 \\ c = 11 \end{cases} \Rightarrow \begin{cases} a = 32 \\ b = 23/2 \\ c = 11 \end{cases} \Rightarrow (21,12,11)_{B_2} = (32, 23/2, 11) \text{ ok!}$$

Mudança de base

No exemplo anterior, precisamos calcular coordenadas várias vezes.

Vamos automatizar isso:

- depois, com resolução simultânea de sistemas;
- agora, com matriz de mudança de base.

Suponha V com duas bases: $B_1 = \{v_1, \ldots, v_n\}$ e B_2.

A função identidade $I\colon V \to V$, $I(x) = x$, é uma transformação linear. Portanto:

$$x_{B_2} = [I]_{B_1,B_2} \cdot x_{B_1}$$

Seja

$$M = [I]_{B_1,B_2} = \begin{bmatrix} (v_1)_{B_2} & (v_2)_{B_2} & \cdots & (v_n)_{B_2} \end{bmatrix}.$$

Lembre que M induz $T_M\colon \mathbb{K}^n \to \mathbb{K}^n$, $T_M(x) = Mx$. Temos:

$$T_M((v_i)_{B_1}) = M \begin{bmatrix} 0 \\ \vdots \\ \underset{(i^{\underline{o}})}{1} \\ \vdots \\ 0 \end{bmatrix} = (v_i)_{B_2}$$

Em geral:

$$x = \sum \lambda_i v_i \Rightarrow x_{B_1} = (\lambda_1, \ldots, \lambda_n) \Rightarrow$$

$$\Rightarrow T_M(x_{B_1}) = M \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = \lambda_1 (v_1)_{B_2} + \ldots + \lambda_n (v_n)_{B_2} =$$

$$= \left(\sum \lambda_i v_i \right)_{B_2} = x_{B_2}$$

Isso significa $M x_{B_1} = x_{B_2}$.

Note:

$$I^{-1} = I \Rightarrow [I]_{B_2,B_1} = \left([I]_{B_1,B_2}\right)^{-1} = M^{-1}$$

Ou seja, também:

$$M^{-1} = [I]_{B_2,B_1} = \Big[(1^{\underline{o}} \text{ vetor de } B_2)_{B_1} \quad (2^{\underline{o}} \text{ vetor de } B_2)_{B_1} \quad \cdots\Big]$$

No exemplo anterior

- $T\colon \mathbb{R}^2 \to \mathbb{R}^3$, $T(x,y) = (3y, 2x, 5y - 4x)$;
- $B_1 = \{(-3,0),(1,1)\}$ base de $\mathbb{R}^2$ (na notação original!);
- $B_2 = \{(1,0,0),(0,2,0),(-1,-1,1)\}$ base de $\mathbb{R}^3$ (na notação original!).

Em $\mathbb{R}^2$, tomamos os vetores de B_1 escritos na base canônica:

$$M = [I]_{B_1,\mathsf{can}} = \begin{bmatrix} -3 & 1 \\ 0 & 1 \end{bmatrix}.$$

Então

$$[I]_{\mathsf{can},B_1} = M^{-1} = \begin{bmatrix} -1/3 & 1/3 \\ 0 & 1 \end{bmatrix},$$

cujas colunas são e_1, e_2 na base B_1.

Em $\mathbb{R}^3$, tomamos os vetores de B_2 escritos na base canônica:

$$A = [I]_{B_2,\mathrm{can}} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Então

$$[I]_{\mathrm{can},B_2} = A^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \end{bmatrix},$$

cujas colunas são e_1, e_2, e_3 na base B_2.

Temos:

$$\overbrace{\mathbb{R}^2}^{\text{base } B_1} \xrightarrow[\text{matriz } M]{I} \overbrace{\mathbb{R}^2}^{\text{base canônica}} \xrightarrow[\text{matriz } \begin{bmatrix} 0 & 3 \\ 2 & 0 \\ -4 & 5 \end{bmatrix}]{T} \overbrace{\mathbb{R}^3}^{\text{base canônica}} \xrightarrow[\text{matriz } A^{-1}]{I} \overbrace{\mathbb{R}^3}^{\text{base } B_2}$$

Então:

$$A^{-1}\begin{bmatrix}0 & 3\\ 2 & 0\\ -4 & 5\end{bmatrix}M=\begin{bmatrix}1 & 0 & 1\\ 0 & 1/2 & 1/2\\ 0 & 0 & 1\end{bmatrix}\cdot\begin{bmatrix}0 & 3\\ 2 & 0\\ -4 & 5\end{bmatrix}\cdot\begin{bmatrix}-3 & 1\\ 0 & 1\end{bmatrix}=$$

$$=\begin{bmatrix}-4 & 8\\ -1 & 5/2\\ -4 & 5\end{bmatrix}\cdot\begin{bmatrix}-3 & 1\\ 0 & 1\end{bmatrix}=\begin{bmatrix}12 & 4\\ 3 & 3/2\\ 12 & 1\end{bmatrix}\overset{\text{note!}}{=\!=}[T]_{B_1,B_2}$$

Em suma:

- $[I]_{B_1,\text{can}}$ costuma vir dada (vetores de B_1 na canônica);
- $[I]_{B_2,\text{can}}$ costuma vir dada (vetores de B_2 na canônica);
- inverta-a para obter $[I]_{\text{can},B_2}$.

$$[T]_{\underset{\text{início}}{B_1},\underset{\text{fim}}{B_2}} = [I]_{\underset{(5^{\underline{o}})}{\text{can}},\underset{(6^{\underline{o}})}{B_2}} \cdot [T]_{\underset{(3^{\underline{o}})}{\text{can}},\underset{(4^{\underline{o}})}{\text{can}}} \cdot [I]_{\underset{(1^{\underline{o}})}{B_1},\underset{(2^{\underline{o}})}{\text{can}}}$$

Analogamente:

$$[T]_{\underset{\text{início}}{\text{can}},\underset{\text{fim}}{\text{can}}} = [I]_{\underset{(5^{\underline{o}})}{B_2},\underset{(6^{\underline{o}})}{\text{can}}} \cdot [T]_{\underset{(3^{\underline{o}})}{B_1},\underset{(4^{\underline{o}})}{B_2}} \cdot [I]_{\underset{(1^{\underline{o}})}{\text{can}},\underset{(2^{\underline{o}})}{B_1}}$$

Espaços linha e coluna; posto e nulidade

A matriz $A \in M_{m\times n}(\mathbb{K})$ induz $T_A\colon \mathbb{K}^n \to \mathbb{K}^m$, $T_A(x) = Ax$.

Tomando $x = e_j$ da base canônica, temos $T_A(e_j) = Ae_j = j$-ésima coluna de A.

Então as colunas de A geram $\operatorname{Im} T_A$, chamado *espaço coluna* de A.

Sua dimensão é o *posto* de A: mostraremos que é a mesma dimensão do *espaço linha* de A, gerado pelas linhas de A.

Seja R uma forma escalonada de A: então as operações que levam de A a R correspondem a uma matriz invertível E, produto de matrizes especiais, de modo que $R = EA$.

Como E é invertível, corresponde a um isomorfismo $T_E \colon \mathbb{K}^m \to \mathbb{K}^m$, então os espaços linha e coluna de R têm as mesmas dimensões daqueles de A.

Porém, em R:

- as linhas não nulas são LI entre si (compare posições dos pivôs);
- as colunas com pivôs são LI entre si (*ditto*) e geram aquelas sem pivô;
- portanto, o posto é o número de pivôs e é $\leqslant m, n$.

Nulidade

É a dimensão de $\mathrm{Nuc}(T_A)$ e é $\leqslant n$. Então:

- T_A é injetora $\Leftrightarrow$ A tem nulidade zero.
- posto + nulidade $= n$.

(Prova: $\dim \mathrm{Im}\, T_A + \dim \mathrm{Nuc}\, T_A = \dim \mathbb{K}^n$.)

Cuidado: nulidade só é número de linhas nulas na forma reduzida para matrizes quadradas.

Note:

$$x \in \mathrm{Nuc}(T_A) \Leftrightarrow T_A(x) = 0 \Leftrightarrow Ax = 0 \Leftrightarrow Rx = 0$$

(porque E é invertível).

Então Nuc T_A é o espaço de soluções de $Rx = 0$.

Essa forma já facilita descrever a solução por adição reversa (e achar base).

Corolário

Para A quadrada ($m = n$), equivalem:

(a) As linhas de A são LI.

(b) As colunas de A são LI.

(c) O posto de A é n.

(d) A é invertível (não singular).

(e) A nulidade de A é zero.

Demonstração: (c) $\Leftrightarrow$ (e) $\Leftrightarrow$ T_A injetora $\Leftrightarrow$ T_A bijetora $\Leftrightarrow$ (d).

(a) $\Leftrightarrow$ (c) $\Leftrightarrow$ (b) pelas definições de espaço linha e coluna.

Hiperplanos e funcionais lineares

Há uma conexão profunda entre hiperplanos e funcionais lineares, que exploraremos aqui.

Os funcionais lineares também têm uma teoria bem desenvolvida e aplicações, que não veremos no todo.

Seja V um espaço vetorial sobre $\mathbb{K}$.

Definições

Um *hiperplano* é um subespaço próprio maximal de V, isto é, é um subespaço H tal que não existe subespaço entre H e V:

$$H \subseteq W \subseteq V \Rightarrow W = H \text{ ou } H = V$$

Um *funcional linear* é uma transformação linear $f : V \to \mathbb{K}$.

Exemplos

Os hiperplanos de $\mathbb{R}^2$ e $\mathbb{R}^3$ são as retas e planos, resp., passando pela origem.

Todo núcleo de funcional linear não trivial é um hiperplano e reciprocamente.

No espaço $L(I)$ das funções integráveis sobre o intervalo I, temos o funcional

$$\varphi \mapsto \int_I \varphi(t)\,dt.$$

A valoração

$$f_x\colon \mathbb{K}^D \to \mathbb{K},\ \ f_x(\varphi) = \varphi(x),$$

que calcula $\varphi\colon D \to \mathbb{K}$ em $x \in D$ fixo, é funcional linear.

Justificaremos esses exemplos com as próximas proposições.

Proposição

Para $\dim V = n \geqslant 1$, finita: Os hiperplanos são os subespaços de dimensão $n-1$.

Argumento: Seja W subespaço de V e B_1 base de W. Estenda B_1 a uma base B de V.

Se $|B_1| \leqslant n-2$ então existem $v, u \in B \setminus B_1$ distintos e $W = [B_1] \subset [B_1, v] \subset [B_1, v, u] \subseteq V$.

Se $|B_1| \geqslant n-1$ então $(\forall v \notin W)\ [B_1, v] = V$.

Proposição

Para $V \neq 0$: Os hiperplanos são os núcleos dos funcionais lineares não nulos.

(1) Suponha H hiperplano e $v \in V \setminus H$. Tome B_1 base de H: então $B_1 \cup \{v\}$ é base de V pela definição: $H \subseteq [B_1, v] \subseteq V$. Defina $f\colon V \to \mathbb{K}$ linear via $f|_{B_1} = 0$ e $f(v)$ sendo escalar não nulo de sua escolha. Então

$$\begin{aligned} \mathrm{Nuc}(f) &= \{\, x \in V \mid f(x) = 0 \,\} = \\ &= \{\, x \in V \mid x \text{ não tem componente em } v \,\} = H. \end{aligned}$$

(2) Dado funcional linear $f \neq 0$, então $\mathrm{Nuc}(f) \neq V$. Se $\mathrm{Nuc}(f) \subset W \subseteq V$ com W subespaço, fixe $v \in W \smallsetminus \mathrm{Nuc}(f)$. Dado $x \in V$, mostraremos que $x \in W$: temos

$$f\left(x - \frac{f(x)}{f(v)} \cdot v\right) = f(x) - \frac{f(x)}{f(v)} \cdot f(v) = 0 \Rightarrow$$

$$\Rightarrow x - \frac{f(x)}{f(v)} \cdot v \in \mathrm{Nuc}(f) \subset W,$$

mas $\dfrac{f(x)}{f(v)} \cdot v \in W$, então a soma $x \in W$.

Espaço dual

É o conjunto V^* dos funcionais lineares com as operações usuais, ou seja, $L(V, \mathbb{K})$ com adição e multiplicação por escalar ponto a ponto.

A notação varia entre autores e com V'.

Base dual

Fixe uma base B de V e defina, para cada $v \in B$, o funcional

$$v^*\colon V \to \mathbb{K},\ v^*|_B(x) = \begin{cases} 0 & \text{se } x \neq v, \\ 1 & \text{se } x = v. \end{cases}$$

(Lembre que transformação linear pode ser definida só pela base.)

Forme o conjunto $B^* = \{\, v^* \mid v \in B \,\}$.

Proposição

B^* é linearmente independente.

Para tanto, suponha $a_1 v_1^* + \ldots + a_n v_n^* \equiv 0$ para $v_1, \ldots, v_n \in B$ distintos.

Então $(\forall j)$

$$0 = \sum_{i=1}^{n} a_i v_i^*(v_j) = \sum_{i=1}^{n} a_i \delta_{ij} = a_j,$$

de modo que todos $a_j = 0$.

Proposição

Para $\dim V$ finita: B^* é base de V^*, donde $\dim V^* = \dim V$.

Basta gerar V^*: dada $f \in V^*$, para cada $v \in B$ temos $f(v) \in \mathbb{K}$. Como B é finita, podemos calcular $f_1 = \sum_{v \in B} f(v) \cdot v^* \in [B^*]$. Vamos mostrar $f = f_1$.

Para $x \in B$: $f_1(x) = \sum_{v \in B} f(v) \cdot \overset{\text{0 ou 1}}{v^*(x)} = f(x)$. Note: $f(v)$ é a coordenada de f no vetor v^* de B^*.

Escólio

Para dim V qualquer: B^* é base de $\{ f \in V^* \mid f(v) = 0$ para todos, exceto nº finito, $v \in B \}$.

Exemplo clássico

Tome $V = \mathbb{K}[t]$ e a valoração $f(p(t)) = p(1)$.

Com a base canônica $B = \{1, t, t^2, \ldots\}$ temos $(\forall p \in B)\ f(p) = 1$, logo, $f\left(\sum_{i=0}^{n} a_i t^i\right) = \sum_{i=0}^{n} a_i$.

Se $f \in [B^*]$, digamos $f = \sum \lambda_i (t^i)^*$, suponha N o maior grau: então $1 = f(t^{N+1}) = \sum \lambda_i (t^i)^* (t^{N+1}) = \sum \lambda_i 0 = 0$, absurdo.

Transformação transposta

Para V, W espaços vetoriais sobre $\mathbb{K}$ e $T\colon V \to W$ linear.

Note:

$$g \in W^* \Rightarrow V \xrightarrow{T} W \xrightarrow{g} \mathbb{K} \Rightarrow (g \circ T) \in V^*.$$

Defina $T^t\colon W^* \to V^*$, $T^t(g) = g \circ T$.

Proposição

T^t é linear.

Suponha $a, b \in \mathbb{K}$ e $g, h \in W^*$: queremos mostrar

$$T^t(ag + bh) = aT^t(g) + bT^t(h),$$

ou seja, queremos mostrar

$$(ag + bh) \circ T = a(g \circ T) + b(h \circ T).$$

Seja $x \in V$ arbitrário: então

$$
\begin{aligned}
[(ag + bh) \circ T](x) = (ag + bh)\big(T(x)\big) = \\
= ag(T(x)) + bh(T(x)) = \\
= a(g \circ T)(x) + b(h \circ T)(x) = [a(g \circ T) + b(h \circ T)](x).
\end{aligned}
$$

Proposição

Para $\dim V$, $\dim W$ finitas: Sejam B, C bases de V, W resp. Então

$$\left[T^t\right]_{C^*,B^*} = \left([T]_{B,C}\right)^t,$$

no lado direito sendo transposta de matriz.

Para demonstrar, escreva: $B = \{v_i\}$, $B^* = \{v_i^*\}$, $C = \{w_j\}$, $C^* = \{w_j^*\}$ e $[T^t]_{C^*,B^*} = [a_{ij}]$.

A j-ésima coluna de $[T^t]_{C^*,B^*}$ é $\big(T^t(w_j^*)\big)_{B^*}$, cujo i-ésimo elemento a_{ij} é a coordenada de $T^t(w_j^*)$ no vetor v_i^* de B^*.

Como B^* é a base dual a B, temos essa coordenada:

$$a_{ij} = (T^t(w_j^*))(v_i) = (w_j^* \circ T)(v_i) = w_j^*(T(v_i)).$$

Escrevendo $T(v_i) = \sum b_k w_k$, vem:

$$w_j^*(T(v_i)) = \sum b_k \overbrace{w_j^*(w_k)}^{0 \text{ ou } 1} = b_j,$$

então a_{ij} é a j-ésima coordenada de $T(v_i)$, donde a_{ij} é a j-ésima linha da i-ésima coluna de $[T]_{B,C}$.

Determinantes

Tópicos

- Concepção como volume orientado.
- Propriedades e métodos de cálculo.
- O determinante de operador linear.
- Bases orientadas.

Definição

São funções $M_{n\times n}(\mathbb{K}) \to \mathbb{K}$ com diversas propriedades e aplicações.

Introduzi-las-emos para calcular volumes orientados em $\mathbb{R}^n$.

Volume orientado (ou "com sinal")

(Figura na lousa.)

Paralelepípedo gerado

Seja V espaço vetorial sobre $\mathbb{R}$ com $\dim V = n$ e fixe $x_1, \ldots, x_m \in V$.

O *paralelepípedo gerado* por esses vetores é

$$P(x_1, \ldots, x_m) = \{\, a_1 x_1 + \ldots + a_m x_m \in V \mid (\forall i)\, 0 \leqslant a_i \leqslant 1 \,\}.$$

(Figuras na lousa.)

Fixe uma base $B = \{v_1, \ldots, v_n\}$ de V e uma ordem entre os elementos da base.

Queremos definir o volume de $P(x_1, \ldots, x_m)$ em termos do volume de $P(v_1, \ldots, v_n)$ fixado como uma *unidade*.

Para tanto, desejamos:

- Se $m < n$ então $\operatorname{vol} P(x_1, \ldots, x_m) = 0$ porque P é uma lâmina sem altura.
- Se $\dim[x_1, \ldots, x_m] < n$ então $\operatorname{vol} P(x_1, \ldots, x_m) = 0$, pelo mesmo motivo.

Portanto, fixamos $m = n$.

Então, desejamos:

- $x_1, \ldots, x_n$ são LD $\Leftrightarrow$ $\operatorname{vol} P(x_1, \ldots, x_n) = 0$, pelo mesmo motivo.
- Multilinearidade (figuras na lousa):

$$\begin{aligned}\operatorname{vol} P(x_1, \ldots, x_k + x_k', \ldots, x_n) &= \operatorname{vol} P(x_1, \ldots, x_k, \ldots, x_n) \\ &\quad + \operatorname{vol} P(x_1, \ldots, x_k', \ldots, x_n)\end{aligned}$$

  e também

$$\operatorname{vol} P(x_1, \ldots, \lambda x_k, \ldots, x_n) = \lambda \operatorname{vol} P(x_1, \ldots, x_k, \ldots, x_n)$$

  (se não quiséssemos sinal, λ sairia em módulo).
- $\operatorname{vol} P(v_1, \ldots, v_n) = 1$.

Teorema

Existe uma única função vol (da família de paralelepípedos a $\mathbb{R}$) que satisfaz essas três condições. Além disso, existe uma única função $\det\colon M_{n\times n}(\mathbb{K}) \to \mathbb{K}$ tal que (com $\mathbb{K} = \mathbb{R}$)

$$\operatorname{vol} P(x_1, x_2, \ldots, x_n) = \det[(x_1)_B|(x_2)_B|\ldots|(x_n)_B].$$

Não demonstraremos esse teorema.

Corolário

det é a função volume para $\mathbb{R}^n$ com a base canônica ordenada $\{e_1, \ldots, e_n\}$ como usual e vetores listados em colunas. Assim, det também satisfaz:

- para $A \in M_{n\times n}(\mathbb{K})$, as colunas de A são LD (então A não é invertível e as linhas de A são LD) $\Leftrightarrow \det A = 0$;
- det é multilinear em suas colunas (nas linhas também, cf. abaixo para transposta);
- $\det I = 1$, onde I é a matriz identidade (colunas e_i em sequência).

Vejamos outras propriedades do determinante e maneiras de calculá-lo...

Não justificaremos várias delas.

Regras de Sarrus

. . .

Cuidado com produtos de diversos sinais!

Fórmula de Leibniz?

$$\det[a_{ij}]_{1\leqslant i,j\leqslant n} = \sum_{\sigma} \varepsilon(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}$$

onde a soma é feita sobre todas as bijeções $\sigma\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$ e $\varepsilon(\sigma) = (-1)^N$ onde N é o número de pares $i < j$ tais que $\sigma(i) > \sigma(j)$.

Essa fórmula é a definição algébrica de determinante, para qualquer $\mathbb{K}$, e permite muitas demonstrações.

Algumas consequências:

- $\det A^t = \det A$;
- $\det(AB) = \det A \cdot \det B$ (para ambas A, B quadradas de mesma dimensão);
- $\det(A^{-1}) = (\det A)^{-1}$ (para A invertível).

Por recursão ()

$$\begin{aligned}
\det[\lambda]_{1\times 1} &= \lambda; \\
\det A &= \sum_{j=1}^{n} (-1)^{i_0+j} a_{i_0 j} \det A_{\widehat{i_0 j}} \text{ ou} \\
&= \sum_{i=1}^{n} (-1)^{i+j_0} a_{i j_0} \det A_{\widehat{i j_0}}
\end{aligned}$$

onde i_0 ou j_0 é fixo ("ao longo de uma linha ou coluna") e $A_{\widehat{ij}}$ é a matriz $(n-1)\times(n-1)$ obtida de A eliminando-se a linha i e a coluna j.

Usa-se, então, a linha ou coluna com mais zeros.

Consequência: Se A é triangular superior ou inferior, basta multiplicar os elementos da diagonal principal (por indução em n):

$$\begin{vmatrix} a_{11} & * & * & \cdots \\ & a_{22} & * & \cdots \\ & & a_{33} & \cdots \\ & & & \ddots \end{vmatrix} = \begin{vmatrix} a_{11} & & & \\ * & a_{22} & & \\ * & * & a_{33} & \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix} = \prod_{i=1}^{n} a_{ii}$$

Mais geralmente, vale a decomposição em blocos:

$$\det \begin{bmatrix} A_{p\times p} & * & \cdots \\ & B_{q\times q} & \cdots \\ & & \ddots \end{bmatrix} = \det A \times \det B \times \ldots$$

Por escalonamento

Toda matriz escalonada é triangular (potencialmente com várias entradas nulas na diagonal).

Dada A, seja A_1 obtida por eliminação gaussiana com N permutações de linha e *sem normalização dos pivôs* (4º passo).

Então $\det A = (-1)^N \times$ produto da diagonal de A_1. (Dos métodos listados, este é mais eficiente em geral e pode ser mais simplificado.)

Isso porque, das matrizes elementares, $\det P_{ij} = -1$ e $\det {M_i^\lambda}^{-1} S_{ij} M_i^\lambda = \det S_{ij} = 1$, enquanto $A_1 = EA$ com E produto dessas matrizes.

De fato. . .

Efeito das operações elementares

- Permutação de linhas (ou de colunas) inverte o sinal do determinante.
- Multiplicação de uma linha (ou coluna) pelo escalar λ multiplica o det. por λ.
- Multiplicação de toda a matriz $n \times n$ por λ multiplica o det. por λ^n.
- Adição, a uma linha (ou coluna), de uma combinação linear de *outras* linhas (ou colunas) não altera o det.

Determinante de um operador linear

Sejam $T : V \to V$ transformação linear e B_1, B_2 bases de V.

Então $[T]_{B_i,B_j}$ são quadradas e

$$[T]_{B_2,B_2} = [I]_{B_1,B_2} \cdot [T]_{B_1,B_1} \cdot [I]_{B_2,B_1}.$$

(Note mesma base em cada $[T]$ e a transformação identidade I.)

Mas $[I]_{B_2,B_1} = \left([I]_{B_1,B_2}\right)^{-1}$, então:

$$\det[T]_{B_2,B_2} = \frac{1}{\det[I]_{B_2,B_1}} \cdot \det[T]_{B_1,B_1} \cdot \det[I]_{B_2,B_1} = \det[T]_{B_1,B_1}.$$

Portanto, o determinante da matriz do operador independe da base (contanto que seja a mesma para argumento e imagem) e é chamado "determinante do operador".

Orientação de base (sobre $\mathbb{R}$)

Em GA, diz-se que B_1, B_2 *têm a mesma orientação* se $\det[I]_{B_1,B_2} > 0$.

Exercício

Mostre que "ter a mesma orientação" é uma relação de equivalência.

Produto interno, norma e ortogonalização

Tópicos

- Abstração de ângulo.
- Abstração de magnitude.
- Construção de bases ortonormais.
- Transformações que preservam ângulo e magnitude.
- Noções de distância e continuidade.
- Otimização por mínimos quadrados.

Produto interno

O produto interno é um modo de definir e medir *ângulos* entre vetores. Mais uma vez nos inspiramos na Geometria Analítica.

Há várias notações: $\langle x|y\rangle$, $\langle x, y\rangle$, $x \bullet y$ etc.

Para números complexos, indicaremos o *conjugado* com $*$:

$$\begin{aligned} \lambda &= \Re(\lambda) + i\Im(\lambda) \\ \lambda^* &= \Re(\lambda) - i\Im(\lambda) \end{aligned}$$

Trabalharemos com V espaço vetorial sobre $\mathbb{K}$.

Definição

Um *produto interno* sobre V é uma função $V^2 \to \mathbb{K}$ satisfazendo:

- $\langle\, x + u \mid y \,\rangle = \langle x|y\rangle + \langle u|y\rangle$ e
  $\langle\, y \mid x + u \,\rangle = \langle y|x\rangle + \langle y|u\rangle$;
- $\langle \lambda x|y\rangle = \lambda^* \langle x|y\rangle$ (cuidado!) e
  $\langle x|\lambda y\rangle = \lambda\langle x|y\rangle$ (vários autores invertem a conjugação!);

- $\langle x|y\rangle = \langle y|x\rangle^*$ (esta propriedade permute deduzir as 2ªˢ linhas acima);
- $\langle x|x\rangle$ é real estritamente positivo quando $x \neq 0$.

Exemplos

No espaço euclideano $\mathbb{K}^n$, dados $\lambda_1, \ldots, \lambda_n > 0$, pomos:

$$\langle x|y\rangle = \sum_{i=1}^{n} \lambda_i x_i^* y_i$$

onde x_i, y_i são as coordenadas de x, y.

Com todos $\lambda_i = 1$, é o produto *canônico.*

No espaço $L(I, \mathbb{K})$ das funções integráveis sobre um intervalo I, com valores em $\mathbb{K}$, pomos:

$$\langle f|g\rangle = \int_I \big(f(t)\big)^* g(t)\, dt$$

canônico desse espaço.

Propriedades

- $\langle 0|x\rangle = 0$ e $\langle x|0\rangle = 0$.
- $\langle x|x\rangle = 0 \Leftrightarrow x = 0$.
- Para vetores $x_i, y_j \in V$ (não coordenadas!) e $a_i, b_j \in \mathbb{K}$,

$$\left\langle \sum_{i=1}^{m} a_i x_i \,\middle|\, \sum_{j=1}^{n} b_j y_j \right\rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} a_i^* b_j \langle x_i|y_j\rangle.$$

Exercício: Deduza essas propriedades a partir dos axiomas.

Espaço normado

É V sobre $\mathbb{K}$ munido de uma função $V \to \mathbb{R}$ chamada *norma* e satisfazendo:

- $\|x\| \geqslant 0$ para todo x e
  $\|x\| = 0 \Leftrightarrow x = 0$;
- $\|x+y\| \leqslant \|x\|+\|y\|$ (desigualdade triangular: figura na lousa);
- $\|\lambda x\| = |\lambda| \cdot \|x\|$.

Nos espaços normados, emerge uma medida natural de *distância:*

$$d(x, y) = \|x - y\|$$

Mostre que, então:

- $d(x, z) \leqslant d(x, y) + d(y, z)$ (desigualdade triangular para distâncias);
- $\big||x| - |y|\big| \leqslant \|x + y\| \leqslant \|x\| + \|y\|$.

Vamos relacionar a norma com o produto interno. . .

Proposição

Todo produto interno origina uma norma via $\|x\| = \sqrt{\langle x|x\rangle}$.

Exercício

Demonstre a "regra do paralelogramo":

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Desigualdade de Cauchy–Schwarz

$$\big|\langle x|y\rangle\big| \leqslant \|x\| \cdot \|y\|,$$

com igualdade $\Leftrightarrow \{x, y\}$ é LD.

Ângulo e ortogonalidade

No caso do produto real, como

$$-1 \leqslant \frac{\langle x|y\rangle}{\|x\| \cdot \|y\|} \leqslant 1$$

para $x, y \neq 0$, definimos o *ângulo* entre eles como o arco cosseno desse número.

Em geral ($\mathbb{K} = \mathbb{R}, \mathbb{C}$), chamamos x, y *ortogonais* ($x \perp y$) quando $\langle x|y\rangle = 0$.

Exercícios

Prove o Teorema de Pitágoras ($\mathbb{K} = \mathbb{R}$):

- $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| \cos\theta$, ou também
- $\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\|\|y\| \cos\theta$.

Mostre que um conjunto de vetores não nulos ortogonais entre si é LI.

Projeção

(Figura na lousa.) Sendo θ o ângulo entre x e y, a *projeção* (escalar) de y na direção de x é

$$\operatorname{proj}_x y = \|y\| \cos\theta = \frac{\langle x|y\rangle}{\|x\|}.$$

Note que

$$\operatorname{proj}_{\lambda x} y = \frac{\langle \lambda x | y \rangle}{\|\lambda x\|} = \frac{\lambda^* \langle x | y \rangle}{|\lambda| \cdot \|x\|} = \frac{\lambda^*}{|\lambda|} \operatorname{proj}_x y.$$

Em particular, quando $\lambda > 0$, as projeções são iguais (a projeção depende apenas da orientação de x).

Usareмos isso com $\lambda = \|x\|^{-1}$.

Normalização

Um vetor x é *unitário* se $\|x\| = 1$.

Para normalizar um vetor $\neq 0$, divida-o pela norma: mostre que $\dfrac{x}{\|x\|}$ é unitário.

Bases ortogonais e ortonormais

Suponha que $B = \{v_1, v_2, \ldots\}$ é base de V (finita ou infinita).

B é *ortogonal* se $\langle v_i | v_j \rangle = 0$ para $i \neq j$.

B é *ortonormal* se é ortogonal e cada $\|v_i\| = 1$.

Exercício

Se B é base ortogonal, mostre que

$$\left\{ \frac{v}{\|v\|} \;\middle|\; v \in B \right\}$$

é base (LI, geradora) ortonormal.

Vamos calcular as coordenadas de um vetor em B ortogonal:

$$x = \sum a_j v_j \Rightarrow$$
$$\langle v_i | x \rangle = \left\langle v_i \;\middle|\; \sum a_j v_j \right\rangle = \sum a_j \langle v_i | v_j \rangle = a_i \|v_i\|^2$$
$$\Rightarrow x = \sum \frac{\langle v_j | x \rangle}{\|v_j\|^2} v_j = \sum \operatorname{proj}_{v_j} x \cdot \frac{v_j}{\|v_j\|}$$

No caso de B ortonormal, então, é só omitir $\|v_j\|$.

No caso de B ortonormal, para $x = \sum a_i v_i$ e $y = \sum b_i v_i$, valem:

- $\langle x|y\rangle = \sum a_i^* b_i$;
- $\|x\| = \sqrt{\sum |a_i|^2}$;
- $d(x, y) = \sqrt{\sum |a_i - b_i|^2}$.

Conclusão: com B ortonormal finita, o isomorfismo $V \cong \mathbb{K}^n$, $x \mapsto (x)_B$, preserva também o produto interno, ou seja, $(V, \langle|\rangle_{\text{dado}}) \cong (\mathbb{K}^n, \langle|\rangle_{\text{can}})$.

Exercícios

Pratique as propriedades de produto e norma demonstrando essas fórmulas.

Suponha $S \subseteq V$ e defina $S^\perp = \{\, x \in V \mid (\forall y \in S)\, x \perp y \,\}$. Mostre que:

- $S^\perp$ é subespaço de V;
- $x \perp y$ para cada $x \in [S]$ e $y \in S^\perp$;
- $[S] \oplus S^\perp = V$ no caso de dim V finita.

Para obter $[S] \oplus S^\perp = V$, comece com uma base B_1 de $[S]$; ortogonalize-a para obter B_2; estenda B_2 a uma base B_3 de V; ortogonalize B_3 *começando pelos vetores de* B_2 para obter uma base ortogonal B_4 de V.

Desse modo, B_2 é preservada em B_4 e $B_4 \setminus B_2$ é base ortogonal de $S^\perp$.

Isso pode ser mostrado fazendo produto interno na decomposição de um vetor sobre B_4.

Ortogonalização de Gram–Schmidt

Matrizes ortogonais–unitárias e simétricas–hermitianas

Objetivo: identificar transformações lineares que preservam produto interno (como no teorema a seguir).

Trabalharemos com $\mathbb{K}^n$ com o produto canônico $\langle x|y\rangle = \sum x_i^* y_i$.

Dada $A = [a_{ij}]$, defina $A^* = [a_{ji}^*]$, ou seja, A^* é a *transposta conjugada* (*operador adjunto*) de A.

Note $(A^*)^* = A$.

Lema

Para todos $x, y \in \mathbb{K}^n$, valem $\langle Ax|y\rangle = \langle x|A^*y\rangle$ e $\langle A^*x|y\rangle = \langle x|Ay\rangle$.

Para demonstrar, calculamos:

$$\langle Ax|y\rangle = \left\langle \left(\sum_{j=1}^{n} a_{ij}x_j\right)_{i=1}^{n} \middle| (y_i)_{i=1}^{n} \right\rangle = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}x_j\right)^{*} y_i =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^{*}x_j^{*}y_i = \sum_{j=1}^{n} x_j^{*}\sum_{i=1}^{n} a_{ij}^{*}y_i \text{ etc.}$$

Teorema + Definição

Suponha $A \in M_{n\times n}(\mathbb{K})$ (quadrada). Equivalem:

(a) $A^{-1} = A^*$ (chama-se A *ortogonal* no caso real e *unitária* no caso complexo);

(b) $\langle Ax|Ay\rangle = \langle x|y\rangle$ para quaisquer $x, y \in \mathbb{K}^n$;

(c) $\|Ax\| = \|x\|$ para qualquer $x \in \mathbb{K}^n$;

(d) se $\|x\| = 1$ então $\|Ax\| = 1$;

(e) as linhas de A são vetores ortonormais entre si;

(f) as colunas de A são vetores ortonormais entre si.

Vamos demonstrar...

(a⇒b) $\langle Ax|Ay\rangle = \langle x|A^*Ay\rangle = \langle x|A^{-1}Ay\rangle = \langle x|y\rangle$.

(b⇒a) $\langle x|(A^*A - I)y\rangle = \langle x|A^*Ay\rangle - \langle x|y\rangle = \langle Ax|Ay\rangle - \langle x|y\rangle = 0$.

Com $x = (A^*A - I)y$, vem $(A^*A - I)y = 0$.

Como y é qualquer, vem $A^*A - I = 0$.

Analogamente podemos mostrar $AA^* = I$.

(b⇒c) Por definição: $\|v\| = \sqrt{\langle v|v\rangle}$.

(c⇒b) Pelas identidades de polarização.

(c⇔d) Calcule $\|A\frac{x}{\|x\|}\|$.

(b⇒f) A j-ésima coluna de A é Ae_j, com e_j vetor da base canônica. Então $\langle Ae_j|Ae_k\rangle = \langle e_j|e_k\rangle = \delta_{jk}$.

(f⇒b) Com $x = \sum x_j e_j$ e $y = \sum y_k e_k$, vem

$$\langle Ax|Ay\rangle = \sum_{j,k} x_j^* y_k \langle Ae_j|Ae_k\rangle = \sum_j x_j^* y_j = \langle x|y\rangle.$$

(e⇔a) As linhas de A são as colunas de A^t, mas:

$$\begin{aligned} A^t \text{ é unitária} &\Leftrightarrow (A^t)^{-1} = (A^t)^* \\ &\Leftrightarrow (A^{-1})^t = (A^*)^t \Leftrightarrow A^{-1} = A^*. \end{aligned}$$

Exemplos de matrizes ortogonais

- Matrizes de mudança de bases ortonormais.
- Rotações no plano e no espaço euclideanos.
- Produtos de matrizes ortogonais.
- Uso em simulador de vôo.

Definição

A é *hermitiana* (*autoadjunta*) se $A = A^*$ (chama-se *simétrica* no caso real).

(A não precisa ser unitária.)

Proposição

Para $\mathbb{K} = \mathbb{C}$: A é hermitiana $\Leftrightarrow (\forall x \in \mathbb{C}^n)\, \langle Ax|x \rangle \in \mathbb{R}$.

De fato, se A é hermitiana então $\langle Ax|x\rangle = \langle x|Ax\rangle = \langle Ax|x\rangle^*$, de modo que $\langle Ax|x\rangle$ é real.

Usaremos matrizes hermitianas no tópico de diagonalização.

Continuidade

Mínimos quadrados

Objetivo: dados $A = [a_{ij}] \in M_{m\times n}(\mathbb{R})$ e $b \in \mathbb{R}^m$, achar $x \in \mathbb{R}^n$ que minimize a distância entre Ax e b (conforme exemplos).

Minimizar $d(Ax, b) = \|Ax - b\|$ equivale a minimizar

$$f(x) = \|Ax - b\|^2 = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}x_j - b_i\right)^2.$$

Achamos os pontos críticos:

$$\frac{\partial f}{\partial x_k}(x) = \sum_{i=1}^{m} 2\left(\sum_{j=1}^{n} a_{ij}x_j - b_i\right)a_{ik} =$$

$$= 2\left\langle \underbrace{(a_{1k}, \ldots, a_{mk})}_{k\text{-ésima linha de } A^t} \;\middle|\; \underbrace{\left(\sum_{j=1}^{n} a_{1j}x_j - b_1, \ldots, \sum_{j=1}^{n} a_{mj}x_j - b_m\right)}_{\text{vetor coluna } Ax-b} \right\rangle.$$

Então x_0 é crítico $\Leftrightarrow$ $(\forall k)\,\frac{\partial f}{\partial x_k}(x_0) = 0 \Leftrightarrow A^t(Ax_0 - b) = 0 \Leftrightarrow A^tAx_0 = A^tb$.

Sistema a ser resolvido é $(A^tA)x = A^tb$.

Veremos depois que o ponto é de mínimo e que existe de fato, mas não é único.

Exemplo de ajuste linear

Determinar a reta $s = mt + k$ que passa mais próxima aos pontos (p_i, q_i) com $1 \leqslant i \leqslant N$.

Temos duas incógnitas m, k que formam $x = (m, k)$ na notação acima.

Então queremos minimizar a distância entre

$$\underbrace{\begin{bmatrix} p_1 & 1 \\ p_2 & 1 \\ \vdots & \vdots \\ p_N & 1 \end{bmatrix}}_{A} \cdot \underbrace{\begin{bmatrix} m \\ k \end{bmatrix}}_{x} \quad \text{e} \quad \underbrace{\begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_N \end{bmatrix}}_{b}.$$

Temos

$$A^t A = \begin{bmatrix} p_1 & p_2 & \cdots & p_N \\ 1 & 1 & \cdots & 1 \end{bmatrix} \cdot \begin{bmatrix} p_1 & 1 \\ p_2 & 1 \\ \vdots & \vdots \\ p_N & 1 \end{bmatrix} = \begin{bmatrix} \sum p_i^2 & \sum p_i \\ \sum p_i & \cancelto{N}{\sum 1} \end{bmatrix}$$

e

$$A^t b = \begin{bmatrix} p_1 & p_2 & \cdots & p_N \\ 1 & 1 & \cdots & 1 \end{bmatrix} \cdot \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_N \end{bmatrix} = \begin{bmatrix} \sum p_i q_i \\ \sum q_i \end{bmatrix},$$

então

$$\begin{bmatrix} \sum p_i^2 & \sum p_i \\ \sum p_i & N \end{bmatrix} \cdot \begin{bmatrix} m \\ k \end{bmatrix} = \begin{bmatrix} \sum p_i q_i \\ \sum q_i \end{bmatrix}.$$

Se $A^t A$ for invertível, podemos usar Cramer:

$$m = \frac{\begin{vmatrix} \sum p_i q_i & \sum p_i \\ \sum q_i & N \end{vmatrix}}{\begin{vmatrix} \sum p_i^2 & \sum p_i \\ \sum p_i & N \end{vmatrix}} = \frac{N \sum p_i q_i - (\sum p_i)(\sum q_i)}{N \sum p_i^2 - (\sum p_i)^2}$$

e

$$k = \frac{\begin{vmatrix} \sum p_i^2 & \sum p_i q_i \\ \sum p_i & \sum q_i \end{vmatrix}}{\begin{vmatrix} \sum p_i^2 & \sum p_i \\ \sum p_i & N \end{vmatrix}} = \frac{(\sum p_i)^2 (\sum q_i) - (\sum p_i)(\sum p_i q_i)}{N \sum p_i^2 - (\sum p_i)^2}.$$

Essas são as fórmulas usuais dos livros-texto.

Exemplo de ajuste parabólico

Determinar a parábola $s = at^2 + bt + c$ que passa mais próxima aos pontos (p_i, q_i) com $1 \leqslant i \leqslant N$.

Formamos $x = (a, b, c)$,

$$A = \begin{bmatrix} p_1^2 & p_1 & 1 \\ p_2^2 & p_2 & 1 \\ \vdots & \vdots & \vdots \\ p_N^2 & p_N & 1 \end{bmatrix} \quad \text{e} \quad b = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_N \end{bmatrix}.$$

Portanto,

$$A^t A = \begin{bmatrix} p_1^2 & p_2^2 & \cdots & p_N^2 \\ p_1 & p_2 & \cdots & p_N \\ 1 & 1 & \cdots & 1 \end{bmatrix} \cdot \begin{bmatrix} p_1^2 & p_1 & 1 \\ p_2^2 & p_2 & 1 \\ \vdots & \vdots & \\ p_N^2 & p_N & 1 \end{bmatrix} = \begin{bmatrix} \sum p_i^4 & \sum p_i^3 & \sum p_i^2 \\ \sum p_i^3 & \sum p_i^2 & \sum p_i \\ \sum p_i^2 & \sum p_i & \sum 1 \end{bmatrix}^{N}$$

e

$$A^t b = \begin{bmatrix} p_1^2 & p_2^2 & \cdots & p_N^2 \\ p_1 & p_2 & \cdots & p_N \\ 1 & 1 & \cdots & 1 \end{bmatrix} \cdot \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_N \end{bmatrix} = \begin{bmatrix} \sum p_i^2 q_i \\ \sum p_i q_i \\ \sum q_i \end{bmatrix}.$$

Note que $A^t A$ é 3×3 e $A^t b$ é 3×1 (dimensões independem de N).

Notamos, nos exemplos, que $A^t A$ é matriz quadrada simétrica, o que facilita seu cálculo:

$$(A^t A)^t = A^t A^{tt} = A^t A.$$

Discussão teórica

Determinamos um ponto crítico x_0.

É ponto de mínimo porque função distância é ilimitada (não há ponto de máximo, nem de sela): de fato, $A(\lambda x) = \lambda x$ para λ arbitrariamente grande.

Não é único se $A^t A$ for singular (não invertível).

Existe?

(Figura na lousa.)

Para todo $v \in \mathbb{R}^n$, temos

$$\langle Av | b - y_0 \rangle_{\text{em } \mathbb{R}^m} = \langle Av | b - Ax_0 \rangle_{\text{em } \mathbb{R}^m} =$$
$$= \langle v | A^t(b - Ax_0) \rangle_{\text{em } \mathbb{R}^n} = \langle v | \vec{0} \rangle_{\text{em } \mathbb{R}^n} = 0.$$

Então $b - y_0$ é ortogonal a todo o subespaço Im T_A. (Note que v terá a forma $x - x_0$.)

y_0 é o *vetor projeção ortogonal* de b em Im T_A.

Se $B = \{v_1, \ldots, v_m\}$ é base ortonormal de $\operatorname{Im} T_A$, então

$$y_0 = \sum_{k=1}^{m} \langle v_k | b \rangle v_k.$$

Então y_0 é único determinado por b.

x_0 existe porque $y_0 \in \operatorname{Im} T_A$.

x_0 é ponto de mínimo também porque (Pitágoras):

$$\|y-b\|^2 = \|y-y_0\|^2+\|y_0-b\|^2 \geqslant \|y_0-b\|^2 \Rightarrow \|Ax-b\| \geqslant \|Ax_0-b\|.$$

Quando A^tA é invertível? O que acontece?

Nesse caso, $A^{\mathrm{ig}} = (A^tA)^{-1}A^t$ é chamada *inversa generalizada* de A e temos:

- $x_0 = A^{\mathrm{ig}}b$;
- $A^{\mathrm{ig}}A = (A^tA)^{-1}(A^tA) = I$;
- $y_0 = Ax_0 = AA^{\mathrm{ig}}b = A(A^tA)^{-1}A^tb$, sendo a matriz $A(A^tA)^{-1}A^t$ dita *de projeção em* $\mathrm{Im}\, T_A$;
- se A já é quadrada e invertível, então $A^{\mathrm{ig}} = A^{-1}(A^t)^{-1}A^t = A^{-1}$ e $x_0 = A^{-1}b$ como esperado.

Proposição

A^tA é invertível se e somente se as *colunas* de A são LI.

($\Rightarrow$) Se as colunas de A são LD, existe combinação linear nula não trivial delas: os escalares formam um vetor $x \in \mathbb{R}^n$, $x \neq 0$, de modo que $Ax = 0$ em $\mathbb{R}^m$. Então $A^tAx = 0$, de modo que as colunas de A^tA são LD e A^tA é singular.

($\Rightarrow$) Assuma que as colunas de A são LI, então (como acima) $(\forall x \neq 0)\ Ax \neq 0$.

Suponha, porém, algum x tal que $(A^t A)x = 0$, ou seja, $A^t(Ax) = 0$. Então Ax é ortogonal a cada linha de A^t (definições dos produtos matricial e interno canônico), ou seja, Ax é ortogonal a cada coluna de A.

Assim, $Ax \in (\text{esp. coluna de } A)^\perp$, embora Ax seja combinação linear das colunas de A, ou seja, $Ax \in (\text{esp. coluna de } A)$.

Portanto $Ax \perp Ax$, o que impõe $Ax = 0$, contradição.

Autovalores

Tópicos

- Autovalores e autovetores.
- Diagonalização.
- Cadeias de Markov.

Definição

Trabalharemos com $A \in M_{n\times n}(\mathbb{C})$ (quadrada) e a base canônica de $\mathbb{C}^n$.

Neste curso, a principal utilidade dos autovalores será a diagonalização de matrizes.

Definição

$\lambda \in \mathbb{C}$ é *autovalor* (*valor próprio*) de A se existe $x \in \mathbb{C}^n$, $x \neq 0$, tal que $Ax = \lambda x$ (ou seja, a multiplicação por A é simplificada pela multiplicação por λ).

Note que $A0 = \lambda 0$ sempre, então requeremos a existência de $x \neq 0$.

Podemos ter $\lambda = 0$.

Definição

Para um autovalor λ, todo x satisfazendo $Ax = \lambda x$ é chamado *autovetor (vetor próprio) de A correspondente a λ.*

Note que o vetor 0 sempre é autovetor.

Proposição

Os autovalores de A são as raízes de $P(t) = \det(A - tI)$, onde I é a matriz identidade $n \times n$.

Esse polinômio é $(-1)^n \det(tI - A)$. O *polinômio característico* $\det(tI - A)$ é sempre *mônico*, isto é, tem coeficiente de maior grau 1.

Prefira o que for mais fácil de calcular, *com cuidado com os sinais.*

Demonstração

Note a equivalência entre afirmações consecutivas:

λ é autovalor

$\Leftrightarrow$ $(\exists x \neq 0)\ Ax = \lambda x$

$\Leftrightarrow$ $(\exists x \neq 0)\ (A - \lambda I)x = 0$ (em que é preciso ter I para operar com A)

$\Leftrightarrow$ $(\exists x \neq 0)$ as coordenadas de x são escalares de uma combinação linear nula (não trivial) das colunas de $A - \lambda I$

$\Leftrightarrow$ $A - \lambda I$ é singular ou não invertível (porque é quadrada)

Exemplo do Prof. Jerônimo

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

$$\Rightarrow P(t) = \begin{vmatrix} 1-t & -2 & 0 \\ -2 & 1-t & 0 \\ 0 & 0 & 3-t \end{vmatrix} =$$

$$= (1-t)^2(3-t) + 0 + 0 - 0 - 0 - 4(3-t) = (3-t)(t^2 - 2t - 3) =$$

$$= -t^3 = 5t^2 - 3t - 9$$

com raízes 3 dupla e -1 simples.

Esses são os autovalores (guarde suas multiplicidades).

Encontrar os autovetores usando $Ax = \lambda x$:

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

$$\Rightarrow \begin{cases} x - 2y = 3x \\ -2x + y = 3y \\ \cancel{3z} = \cancel{3z} \end{cases} \Rightarrow \begin{cases} -2x - 2y = 0 \\ \cancel{-2x - 2y = 0} \end{cases} \Rightarrow y = -x \Rightarrow$$

$$\Rightarrow \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ -x \\ z \end{bmatrix} = x \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + z \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

(Note a dimensão 2 igual à multiplicidade: *não vale sempre.*)

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

$$\Rightarrow \begin{cases} x - 2y = -x \\ -2x + y = -y \\ 3z = -z \end{cases} \Rightarrow \begin{cases} 2x - 2y = 0 \\ \cancel{-2x + 2y = 0} \\ 2z = 0 \end{cases} \Rightarrow \begin{cases} x = y \\ z = 0 \end{cases} \Rightarrow$$

$$\Rightarrow \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ x \\ 0 \end{bmatrix} = x \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

(Note a dimensão 1 igual à multiplicidade: *não vale sempre.*)

Exercícios de EDO na Graduação

Calcule os autovalores e autovetores destas matrizes:

$$\begin{bmatrix} 3 & 5 \\ -5 & -3 \end{bmatrix}, \quad \begin{bmatrix} 0 & 2 \\ 3 & -1 \end{bmatrix}, \quad \begin{bmatrix} 2 & 1 \\ 6 & 3 \end{bmatrix}, \quad \begin{bmatrix} 3 & -1 \\ 4 & -1 \end{bmatrix}.$$

Exercício

Mostre que A e A^t têm os mesmos autovalores com as mesmas multiplicidades e, de fato, o mesmo polinômio característico.

Construa um exemplo em que os autovetores, porém, são diferentes.

O que acontece com A^* ?

Observações

$\det(A) = P(0)$ é o coeficiente constante de $P(t)$.

É $(-1)^n$ o coef. constante do polinômio característico (mônico).

É $(-1)^n$ o produto dos autovalores (repetidos com multiplicidades) (relação de Girard).

Também se mostra relação entre a soma dos autovalores e o *traço* de A.

Se A é triangular (inclui diagonal), então as entradas na diagonal são os autovalores (pelo cálculo de $\det(A - tI)$).

Para uma R quadrada $n \times n$, tanto A como $R^{-1}AR$ têm os mesmos autovalores com as mesmas multiplicidades. De fato:

$$\det(R^{-1}AR - tI) = \det(R^{-1}AR - R^{-1}tIR) = \det(R^{-1}(A - tI)R) = \dots$$

Definição–Exercício

Para λ autovalor de A, mostre que

$$\mathrm{Aut}_A(\lambda) = \{\, x \in \mathbb{C}^n \mid Ax = \lambda x \,\}$$

é subespaço de $\mathbb{C}^n$, chamado *espaço próprio* ou *autoespaço* de λ.

No exemplo

- $\mathrm{Aut}(3) = [\,(1,-1,0),\,(0,0,1)\,]$.
- $\mathrm{Aut}(-1) = [\,(1,1,0)\,]$.

Proposição

Sejam $\lambda_1, \ldots, \lambda_k$ autovalores *distintos* de A e $x_i \in \mathrm{Aut}(\lambda_i)$. Se $\sum x_i = 0$ então cada $x_i = 0$.

Assim, escrevemos $\mathrm{Aut}(\lambda_1) \oplus \ldots \oplus \mathrm{Aut}(\lambda_k)$.

Demonstraremos por indução em $k \geqslant 2$.

Caso $k = 2$: $x_1 + x_2 = 0 \Rightarrow A(x_1 + x_2) = 0 \Rightarrow \lambda_1 x_1 + \lambda_2 x_2 = 0$, mas $x_2 = -x_1 \Rightarrow \lambda_1 x_1 - \lambda_2 x_1 = 0 \Rightarrow \overbrace{(\lambda_1 - \lambda_2)}^{\neq 0} x_1 = 0 \Rightarrow x_1 = 0 \Rightarrow x_2 = 0$.

Caso $k > 2$:

$\sum x_i = 0 \Rightarrow A(\sum x_i) = 0 \Rightarrow \sum Ax_i = 0 \Rightarrow \sum \lambda_i x_i = 0 \Rightarrow \lambda_1 x_1 + \sum_{i \neq 1} \lambda_i x_i = 0 \Rightarrow \lambda_1(-x_2 - x_3 - \ldots - x_k) + \sum_{i \neq 1} \lambda_i x_i = 0 \Rightarrow \sum_{i \neq 1} (\lambda_i - \lambda_1) x_i = 0$, mas cada $(\lambda_i - \lambda_1) x_i \in \mathrm{Aut}(\lambda_i)$ que é subespaço.

Pela hipótese de indução para $k - 1$, cada $\cancelto{\neq 0}{(\lambda_i - \lambda_1)} x_i = 0 \Rightarrow x_i \neq 0$ para cada $i \neq 1$. Também $x_1 = 0$ pela soma inicial nula.

Corolário

Autovetores não nulos de autovalores distintos são LI.

Para tanto, dada uma combinação linear nula deles, aplique a proposição aos *múltiplos* (x_i) deles, que também são autovetores e cuja soma é 0.

Conclusão

Se cada B_i é base de $\mathrm{Aut}(\lambda_i)$, então $B_1 \cup \ldots \cup B_k$ é base de $\mathrm{Aut}(\lambda_1) \oplus \ldots \oplus \mathrm{Aut}(\lambda_k)$.

Diagonalização

Trabalharemos com $A \in M_{n\times n}(\mathbb{C})$ (quadrada) e a base canônica de $\mathbb{C}^n$ sobre $\mathbb{C}$.

Objetivo é anular o máximo nº elementos fora da diagonal de A, para facilitar cálculos e interpretações. Por exemplo:

- Se $A = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$, então $A^k = \begin{bmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{bmatrix}$.
- Se A é matriz de um sistema de EDO de 1ª ordem e é diagonal, então os eixos representam equações independentes.

Definição

A é *diagonalizável* se existe R invertível tal que $D = R^{-1}AR$ é diagonal.

Teorema

A é diagonalizável se e somente se A tem n autovetores LI.

A demonstração se constituirá em método para diagonalizar.

Sejam $v_1, \dots, v_n$ autovetores LI de A e seja λ_i o autovalor correspondente a v_i. Forme $R = \left[\begin{smallmatrix} v_1 & \cdots & v_n \end{smallmatrix}\right]$ com os vetores em colunas: então R é invertível e

$$AR = \begin{bmatrix} Av_1 & \cdots & Av_n \end{bmatrix} = \begin{bmatrix} \lambda_1 v_1 & \cdots & \lambda_n v_n \end{bmatrix} =$$

$$= \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}.$$

(Compare com $\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$ que bagunça tudo.)

Reciprocamente, se

$$AR = R\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

e v_i é a i-ésima coluna de R, então $Av_i = \lambda_i v_i$ etc.

Corolário

Se A tem n autovalores distintos, então A é diagonalizável.

Exemplo

Para

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

calculamos autovetores $(1, -1, 0)$, $(0, 0, 1)$ (do autovalor 3) e $(1, 1, 0)$ (do autovalor -1), que são LI entre si.

Então

$$R = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad R^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Vamos aplicar calculando A^{500}:

$$D = R^{-1}AR \Rightarrow A = RDR^{-1} \Rightarrow A^{500} = (RD\cancel{R^{-1}})(\cancel{R}D\cancel{R^{-1}})\ldots =$$

$$= RD^{500}R^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 3^{500} & 0 & 0 \\ 0 & 3^{500} & 0 \\ 0 & 0 & (-1)^{500} \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{1+3^{500}}{2} & \frac{1-3^{500}}{2} & 0 \\ \frac{1-3^{500}}{2} & \frac{1+3^{500}}{2} & 0 \\ 0 & 0 & 3^{500} \end{bmatrix}.$$

Notas

R é uma matriz de mudança de base: R^{-1} leva coordenadas na base canônica à base em que T_A tem matriz diagonal (simples contração/dilatação/reflexão independente em cada eixo).

A pode ser real, mas somente diagonalizável com R e/ou D complexa!

Definição

Seja λ autovalor de A:

- A *multiplicidade algébrica* $\mathrm{ma}(\lambda)$ é sua multiplicidade como raiz de $P(t)$.
- A *multiplicidade geométrica* $\mathrm{mg}(\lambda)$ é $\dim \mathrm{Aut}_A(\lambda)$.

Estenda uma base B de $\mathrm{Aut}(\lambda)$ a uma base B_1 de $\mathbb{C}^n$ listando B no início. Então

$$[T_A]_{B_1} = \left[\begin{array}{ccc|c} \lambda & & & \\ & \ddots & & * \\ & & \lambda & \\ \hline & 0 & & * \end{array}\right]$$

(onde o bloco superior esquerdo é $|B| \times |B|$).

Portanto $P(t)$ tem um fator $(\lambda - t)^{|B|}$, donde $\mathrm{mg}(\lambda) = |B| \leqslant \mathrm{ma}(\lambda)$.

Corolário

A é diagonalizável

$\Leftrightarrow$ sendo λ_i seus autovalores distintos para $1 \leqslant i \leqslant k$, vale

$$\mathbb{C}^n = \bigoplus_{i=1}^{k} \mathrm{Aut}(\lambda_i)$$

$\Leftrightarrow$ $\mathrm{mg}(\lambda_i) = \mathrm{ma}(\lambda_i)$ para todo $1 \leqslant i \leqslant k$.

Exercícios de EDO na Graduação

Quais destas matrizes são diagonalizáveis? Diagonalize as possíveis:

$$\begin{bmatrix} 3 & 5 \\ -5 & -3 \end{bmatrix}, \quad \begin{bmatrix} 0 & 2 \\ 3 & -1 \end{bmatrix}, \quad \begin{bmatrix} 2 & 1 \\ 6 & 3 \end{bmatrix}, \quad \begin{bmatrix} 3 & -1 \\ 4 & -1 \end{bmatrix}.$$

Diagonalização de matrizes hermitianas

Assumiremos A hermitiana, ou seja, $A^* = A$. É o caso de matrizes reais simétricas.

Lema A

Os autovalores de A são reais.

Para demonstrar, escreva $Ax = \lambda x$ com $x \neq 0$. Então:

- $\langle Ax | x \rangle = \langle \lambda x | x \rangle = \lambda^* \|x\|^2$ e também
- $\langle Ax | x \rangle = \langle x | Ax \rangle = \langle x | \lambda x \rangle = \lambda \|x\|^2$,

donde $\lambda^* = \lambda$.

Lema B

Se $x \neq 0$ é autovetor de A, então $(\forall y \perp x)\ Ay \perp x$.

Escreva $Ax = \lambda x$, então: $\langle Ay|x\rangle = \langle y|Ax\rangle = \langle y|\lambda x\rangle = \lambda\langle y|x\rangle = 0$.

Lema C

Se λ, μ são autovalores distintos de A, então $\mathrm{Aut}(\lambda) \perp \mathrm{Aut}(\mu)$.

Suponha x, y autovetores de λ, μ resp.: queremos $\langle x|y\rangle = 0$. Calcule $\langle Ax|y\rangle$ de dois modos:

- $\langle Ax|y\rangle = \langle \lambda x|y\rangle = \lambda\langle x|y\rangle$ (porque $\lambda \in \mathbb{R}$) e
- $\langle Ax|y\rangle = \lambda x|Ay\rangle = \langle x|\mu y\rangle = \mu\langle x|y\rangle$.

Como $\lambda \neq \mu$, é preciso $\langle x|y\rangle = 0$.

Proposição

Existe uma base ortonormal de $\mathbb{C}^n$ formada por autovetores de A.

Por indução em n: para $n = 1$, tome x qualquer autovetor não nulo e normalize.

Para $n > 1$, tome x qualquer autovetor não nulo unitário e note que o subespaço $x^\perp = \{\, y \in \mathbb{C}^n \mid y \perp x \,\}$ tem dimensão $n - 1$. Pelo lema B, A induz $T_A \colon x^\perp \to x^\perp$.

Veja método prático depois.

Teorema Espectral

Existe matriz unitária U (isto é, $U^{-1} = U^*$) tal que U^*AU é diagonal.

Este resultado é usado em FVV, juntamente com a próx. prop.

Para demonstrar, a prop. anterior mostra que A é diagonalizável.

Sendo B a base normalizada, como antes forme U com os autovetores em colunas, então $U^{-1}AU$ é diagonal.

Porém, U^*U é a matriz de produtos respectivos de linhas de U^* com colunas de U, isto é, produtos *internos* de colunas de U (conjugado é usado no 1º fator) por colunas de U.

Como B é ortonormal, $U^*U = I$ e então $U^* = U^{-1}$.

Método para obter U

Como provamos A diagonalizável, temos $\mathbb{C}^n = \mathrm{Aut}(\lambda_1) \oplus \ldots \oplus \mathrm{Aut}(\lambda_k)$.

Tome B_i base de $\mathrm{Aut}(\lambda_i)$ e ortogonalize-a com Gram–Schmidt, depois normalize-a obtendo $\widehat{B}_i$.

Como $\mathrm{Aut}(\lambda_i)$ é subespaço, o processo resulta em novos autovetores. Pelo Lema C, todos os autovetores são ortogonais entre si.

Assim $\widehat{B}_1 \cup \ldots \widehat{B}_k$ é base ortonormal de $\mathbb{C}^n$.

Exemplo

$$A = \begin{bmatrix} -1 & 0 & 3 \\ 0 & -4 & 0 \\ 3 & 0 & -1 \end{bmatrix} \text{ real simétrica}$$

Autovalores:

$$P(t) = \begin{vmatrix} -1-t & 0 & 3 \\ 0 & -4-t & 0 \\ 3 & 0 & -1-t \end{vmatrix} = (1+t)^2(-4-t) - 9(-4-t) =$$

$$= (4+t)(9-(1+t)^2) = (4+t)(2-t)(4+t) \Rightarrow$$

raízes -4 dupla e 2 simples.

Autovetores para -4:

$$\begin{bmatrix} -1 & 0 & 3 \\ 0 & -4 & 0 \\ 3 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = -4 \begin{bmatrix} x \\ y \\ z \end{bmatrix} \Rightarrow$$

$$\Rightarrow \begin{cases} -x + 3z = -4x \\ \cancel{-4y} = \cancel{-4y} \\ 3x - z = -4z \end{cases} \Rightarrow \begin{cases} 3x + 3z = 0 \\ \cancel{3x + 3z = 0} \end{cases} \Rightarrow z = -x \Rightarrow$$

$$\Rightarrow \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ -x \end{bmatrix} = x \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + y \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Os autovetores $(1,0,-1)$ e $(0,1,0)$ já são ortogonais (coincidência).

Caso não fossem, usaríamos Gram–Schmidt.

Normalizando: $(\frac{1}{\sqrt{2}},0,-\frac{1}{\sqrt{2}})$ e $(0,1,0)$.

Autovetores para 2:

$$\begin{bmatrix} -1 & 0 & 3 \\ 0 & -4 & 0 \\ 3 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 2 \begin{bmatrix} x \\ y \\ z \end{bmatrix} \Rightarrow$$

$$\Rightarrow \begin{cases} -x + 3z = 2x \\ -4y = 2y \\ 3x - z = 2z \end{cases} \Rightarrow \begin{cases} -3x + 3z = 0 \\ -6y = 0 \\ \cancel{3x - 3z = 0} \end{cases} \Rightarrow \begin{cases} z = x \\ y = 0 \end{cases} \Rightarrow$$

$$\Rightarrow \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ 0 \\ x \end{bmatrix} = x \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

O autovetor $(1, 0, 1)$ é ortogonal com os anteriores, pela 1ª proposição.

Caso houvesse mais, usaríamos Gram–Schmidt para obter ortogonais entre si.

Normalizando: $(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$.

Então

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}, \quad U^{-1} = U^t = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\text{e} \quad D = \begin{bmatrix} -4 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Temos $D = U^t A U$ e $A = U D U^t$.

Nesse exemplo, pudemos conduzir todas as contas em $\mathbb{R}$.

Proposição

No caso de A real simétrica, cada autovalor tem um autovetor não nulo real.

Para tanto, escreva $Ax = \lambda x$ com $x = u + iv \neq 0$ sendo $u, v \in \mathbb{R}^n$.

Então $Ax = Au + iAv$ com $Au, Av \in \mathbb{R}^n$ e $\lambda x = \lambda u + i\lambda v$ com $\lambda u, \lambda v \in \mathbb{R}^n$.

Concluímos que $Au = \lambda u$ e $Av = \lambda v$. Mas $u \neq 0$ ou $v \neq 0$ porque $x \neq 0$.

Trabalhar diretamente com os autovetores assim obtidos pode não funcionar.

Porém, todas as demonstrações (Lemas A, B, C, Proposição e Teorema Espectral) passam a funcionar em $\mathbb{R}^n$ sobre $\mathbb{R}$, assim como Gram–Schmidt e normalização.

Consequência

Desse modo, obtém-se U real ortogonal e $U^t A U$ é diagonal.

Cadeias de Markov

Objetivo é estudar situações de *market share*, por exemplo:

- Empresas detêm fatias do mercado e consumidores migram de um produto a outro.
- Uma doença se espalha na população e pessoas adoecem, saram ou morrem.

Uma cadeia de Markov consiste de:

- *estados* ou *vértices* $1, 2, \ldots n$;
- uma unidade de tempo, período ou ciclo;
- para cada $i, j \in \{1, \ldots, n\}$, a *proporção* P_{ij} do estado i que migra para o estado j.

Por exemplo, P_{ij} é a probabilidade do consumidor trocar o produto i pelo produto j ao fim de cada mês.

Podemos ter $P_{ii} > 0$ (o produto retém alguns consumidores) ou não.

Porém, é exigido sempre que todos $P_{ij} \geqslant 0$.

Também se exige $\sum_{j=1}^{n} P_{ij} = 1$ para cada i.

Vetor de distribuição ou probabilidades

É $x = (x_1, \ldots, x_n)$ indicando que há x_i indivíduos no estado i, no início de um ciclo ou período.

Não há razão, nem garantia que cada x_i seja inteiro, então se usa medida contínua (ex.: biomassa), proporção ou probabilidade.

Requeremos $\sum_{i=1}^{n} x_i = N$, onde N é constante (geralmente 1) e cada $x_i \geqslant 0$.

Matriz de transição (matriz estocástica)

Começando com x, após um período teremos no estado k:

$$\sum_{j=1}^{n} x_j P_{jk}$$

(soma dos produtos: total que havia em j, vezes proporção que migrou de j para k).

Essa é a k-ésima linha do vetor coluna Ax, onde

$$A = [P_{jk}]^t$$

(note transposição).

Ou seja, para formar A, em cada *coluna* liste as proporções de *saída* do estado correspondente:

$$A = \begin{bmatrix} \cdots & P_{r1} & \cdots \\ \cdots & P_{r2} & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & P_{rn} & \cdots \end{bmatrix} \begin{array}{l} \Leftarrow \text{prop. de } r \text{ a } 1 \\ \Leftarrow \text{prop. de } r \text{ a } 2 \\ \qquad\vdots \\ \Leftarrow \text{prop. de } r \text{ a } n \end{array}$$

- Todas as entradas são $\geqslant 0$ e A é quadrada.
- Cada coluna tem soma 1.
- Cada linha lista as proporções de *entrada* no estado correspondente (não soma 1).

Nota

Muitos autores trabalham com x linha e xA^t; tudo aqui ficará transposto!

Lema de preservação

Se x satisfaz $\sum x_i = N$ e cada $x_i \geqslant 0$, então $y = Ax$ também satisfaz $\sum y_i = N$ e cada $y_i \geqslant 0$.

De fato, $\sum_i y_i = \sum_{i,j} x_j P_{ji} = \sum_j x_j \cancelto{1}{\sum_i P_{ji}} = \sum_j x_j = N$ e $y_i = \sum_j A_{ij} x_j = \sum_j P_{ji} x_j \geqslant 0$.

Corolário

Por indução, $A^k x$ é vetor de distribuição para o mesmo total, para qualquer potência k.

Corolário

Produtos e potências de matrizes de transição também são matrizes de transição.

Para tanto, sejam A, B matrizes de transição. Cada coluna B_j de B é um vetor de distribuição com total 1. Então AB_j é um vetor de distribuição com total 1. Mas essa é a j-ésima coluna de AB. Então AB é matriz de transição.

Vetor estável ou estacionário

É x tal que $Ax = x$ (autovetor com autovalor 1).

Descreve ponto fixo da cadeia, em que os totais nos estados são constantes, apesar de trocas de indivíduos.

(Reveja "equilíbrio" em sistemas de EDO como um tema afim.)

Para existir x estável, é preciso que 1 seja autovalor de A, ou seja, $P(1) = 0$, isto é, $\det(A - I) = 0$.

Nesse caso, para achar x, resolva o SPI $(A - I)x = 0$ com condições adicionais $x_1 + \ldots + x_n = N$ e $x_1 \geqslant 0, \ldots, x_n \geqslant 0$.

Porém, temos informação adicional:

Teorema

Toda matriz de transição tem autovalor 1 e este é seu maior autovalor em módulo.

A seguir, demonstração devida a Mike Spivey...

As somas das colunas de A são 1, então

$$\begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} \cdot A = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix},$$

donde

$$A^t \cdot \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

e vê-se que A^t tem autovetor com autovalor 1, de modo que A também.

Suponha λ autovalor de A, sendo $A^t x = \lambda x$ com $x \neq 0$.

Cada linha de A^t soma 1, então a i-ésima coordenada de $A^t x$ tem a forma $\sum_j P_{ij} x_j = \lambda x_i$ com $\sum_j P_{ij} = 1$.

Desse modo, $|\lambda x_i| \leqslant \sum_j P_{ij} |x_j|$.

Escolha i de modo que $|x_i|$ é máximo dentre $|x_1|, \ldots, |x_n|$, então $x_i \neq 0$ e

$$|\lambda| \cdot |x_i| \leqslant \sum_j P_{ij} |x_j| \leqslant \sum_j P_{ij} |x_i| = |x_i|,$$

donde $|\lambda| \leqslant 1$.

. . .

2017 VCL