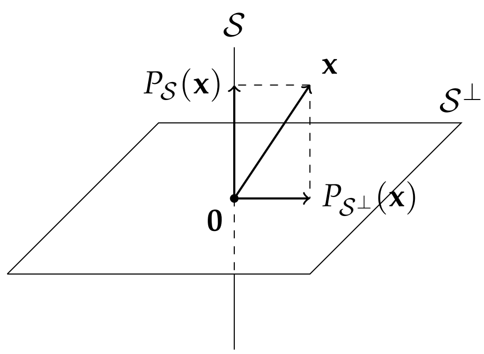
Lições de Álgebra Linear I e II

via Exemplos e Exercícios Resolvidos



Projeções Ortogonais de um Vetor $\mathbf{x}$ sobre um Subespaço $\mathcal{S}$ e sobre seu Complemento Ortogonal $\mathcal{S}^{\perp}$.

José Renato Ramos Barbosa

UFPR - 2020

Universidade Federal do Paraná

Departamento de Matemática

Lições de Álgebra Linear
I e II
Via Exemplos e Exercícios Resolvidos

Autor:
Professor José Renato Ramos Barbosa

2020

Conteúdo

1 Origem, objetivos e diretrizes 7

2 O espaço vetorial $\mathbb{R}^n$ 11

2.1 Geometria analítica do $\mathbb{R}^2$ . . . 11

2.1.1 Exercícios sobre vetores em $\mathbb{R}^2$ . . . 12

2.2 $\mathbb{R}^n$, o espaço euclidiano n-dimensional . . . 17

2.2.1 Notações e definições iniciais . . . 17

2.2.2 Propriedades que caracterizam $\mathbb{R}^n$ como um "espaço vetorial" . . . 18

2.3 Produto interno, módulo, ângulo e ortogonalidade em $\mathbb{R}^n$ . . . 19

2.4 Retas e hiperplanos em $\mathbb{R}^n$ . . . 22

2.5 Subespaços de $\mathbb{R}^n$ . . . 23

2.5.1 Exemplo geral de subespaço: $\mathcal{S}$ gerado por r vetores . . . 26

2.5.2 $\mathcal{S}$ gerado por $r = 3$ vetores em $\mathbb{R}^4$ . . . 26

2.5.3 Bases, LI e LD . . . 27

2.5.4 Dimensão . . . 28

2.5.5 Ortogonalidade e ortonormalidade . . . 29

2.5.6 Subespaços de subespaços . . . 31

2.6 O espaço $\mathbb{R}^3$ . . . 31

2.6.1 Projeção ortogonal de $\mathbf{x}$ sobre $\mathbf{y}$ . . . 31

2.6.2 Produto vetorial de $\mathbf{x}$ e $\mathbf{y}$ . . . 32

2.6.3 Relação entre projeção ortogonal e produto vetorial . . . 32

2.6.4 Dependência linear e produto vetorial . . . 33

2.6.5 Produto misto de $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{z}$ . . . 33

2.6.6 Bases de $\mathbb{R}^3$ via produto misto . . . 34

2.7 Exercícios . . . 39

2.7.1 Resolução do exercício da subseção 2.5.1 . . . 47

3 O espaço vetorial $\mathbb{R}^{m\times n}$ 49

3.1 Adição de matrizes e multiplicação por escalares . . . 49

3.1.1 Matrizes . . . 49

3.1.2 Por que $\mathbb{R}^{m\times n}$ é um espaço vetorial? . . . 51

3.2 Produto e transposição de matrizes . . . 54

3.3 Importância das matrizes quadradas . . . 56

3.4 Determinantes . . . 57

3.5 Sistemas lineares $A\mathbf{x} = \mathbf{b}$ e escalonamento . . . 59

3.5.1 Matriz escalonada reduzida R e escalonamento . . . 61

3.6 Matrizes invertíveis, matrizes elementares e escalonamento . . . 64

3.6.1 Matrizes elementares . . . 66

3.7 Exercícios 68

4 Transformações lineares, autovalores e autovetores 83
4.1 Transformações lineares 83
4.1.1 Núcleo e imagem de A (ou L) 90
4.1.2 Representação de L em outras bases 95
4.2 Autovalores e autovetores 99
4.2.1 Diagonalização 103
4.2.2 Matrizes ortogonais e diagonalização 106
4.3 Exercícios 111
4.3.1 Resoluções de alguns exercícios que precedem a subseção 4.1.1 121

5 Os Espaços vetoriais $\mathbb{K}^n$ e $\mathbb{K}^{m\times n}$ 125
5.1 Definição e propriedades do $(\mathbb{K}^n, +, \cdot)$ 125
5.1.1 Pequena revisão de $\mathbb{C}$, o corpo dos números complexos 125
5.1.2 O corpo $\mathbb{K}$ 126
5.1.3 O espaço $\mathbb{K}^n$ 127
5.2 Exercícios 129
5.3 Informação adicional: diagonalização de $A \in \mathbb{C}^{n\times n}$ 138
5.3.1 Lema de Schur 138
5.3.2 Teorema espectral para $A \in \mathbb{C}^{n\times n}$ 139

6 O espaço vetorial $\mathcal{V}$ sobre o corpo $\mathbb{K}$ 141
6.1 Definição e propriedades de $(\mathcal{V}, +, \cdot)$ 141
6.1.1 Exemplos de $\mathcal{V}$ "diferentes" de $\mathbb{K}^n$ e $\mathbb{K}^{m\times n}$ 142
6.1.2 Subespaços, bases, dimensões, etc. 142
6.2 Isomorfismo entre espaços vetoriais 144
6.2.1 Teorema dos espaços vetoriais isomorfos de dimensões finitas 145
6.2.2 Se $\dim \mathcal{V} = n$, informações sobre $\mathcal{V}$ podem ser obtidas via informações sobre $\mathbb{K}^n$ 146
6.3 Alguns resultados e algumas demonstrações 147
6.3.1 Espaços finitamente gerados, vetores LI e LD, bases 147
6.3.2 O espaço vetorial $\mathcal{L}(\mathcal{V}, \mathcal{W})$ 153
6.3.3 $\dim \mathcal{V} = n$ e $\dim \mathcal{W} = m \Longrightarrow \mathcal{L}(\mathcal{V}, \mathcal{W})$ e $\mathbb{K}^{m\times n}$ são isomorfos 158
6.4 Exercícios 162
6.5 Informação adicional: dimensão da soma direta 175

7 $\mathcal{L}(V)$ 177
7.1 Subespaços invariantes 177
7.2 Autovalores e autovetores 179
7.2.1 Caso $\mathrm{Nu}(L - \lambda I)$ não seja o subespaço nulo de $\mathcal{V}$, seus vetores (não nulos) são os autovetores de L associados ao autovalor λ 179
7.2.2 Polinômios com operadores como variáveis 181
7.2.3 Se $\mathcal{V} \neq \{0_{\mathcal{V}}\}$ é espaço vetorial complexo finitamente gerado, então $T \in \mathcal{L}(\mathcal{V})$ tem autovalor 183
7.2.4 Matrizes triangulares e operadores 184
7.2.5 Diagonalização e operadores 187
7.2.6 Decomposição em somas diretas de autoespaços 188

7.2.7 Complementos e projeções ortogonais. Mínimos quadrados 190
7.3 Funcionais lineares 204
7.3.1 Teorema da Representação de Riesz 204
7.4 Adjuntos, autoadjuntos e normais.
Teorema espectral. Forma de Jordan 205
7.4.1 Construção do operador adjunto 206
7.4.2 $T \in \mathcal{L}(\mathcal{V}) \Longrightarrow T^* \in \mathcal{L}(\mathcal{V})$ 207
7.4.3 Propriedades do adjunto 207
7.4.4 Operador hermiteano, isto é, autoadjunto 208
7.4.5 Operador normal comuta com o seu adjunto 210
7.4.6 Teorema espectral complexo 212
7.4.7 Teorema espectral real 214
7.4.8 Teorema (forma normal de Jordan) 217
7.5 Exercícios 227
7.6 Informações adicionais 237
7.6.1 Matrizes quadradas $\times$ operadores 237
7.6.2 Para $A \in \mathbb{C}^{n\times n}$ não diagonalizável, como podemos obter $\mathsf{P} \in \mathbb{C}^{n\times n}$ invertível com

$$\mathsf{P}^{-1}A\mathsf{P} = J$$

na forma de Jordan? 238

Capítulo 1

Origem, objetivos e diretrizes

SAUDAÇÕES UNIVERSITÁRIAS!
Em geral, quase ninguém presta muita atenção em prefácios e bibliografias. Inicio, assim, chamando o prefácio e a bibliografia de capítulo 1, que é curtíssimo, com apenas três páginas e quatro referências bibliográficas, para que todos possam lê-lo e ter noção do tratamento dado (a partir do capítulo 2) a *álgebra linear* (AL). Peço, portanto, atenção nesse primeiro capítulo e, em particular, na descrição das duas partes nas quais o livro está dividido e no porquê da abordagem "quase" sem *determinantes* que adotamos.
O conteúdo das notas de aulas, que gestaram essas "lições", foi trabalhado e modificado por quase vinte anos. Portanto, é provável que a estrutura (ordem, redação e quantidade de exercícios, definições e resultados) dessas notas tenham variado em muitos dos acessos a minha página institucional.[1]
Durante a quarentena do COVID 19, num momento em que a Terra quase parou e convidou a humanidade para um suspiro coletivo, resolvi tentar organizar as notas supracitadas num formato adequado para a sua publicação. Como o objetivo delas sempre foi o de servir de apoio para cursos de AL ministrados na UFPR, imaginei que o alcance das mesmas pudesse ultrapassar os limites da instituição em que leciono.
Embora possa não parecer claro, tentei escrever esse livro no estilo da renomada coleção *Schaum*, isto é, o conteúdo é, em boa parte, trabalhado via exemplos e exercícios resolvidos. Além disso, esse livro está dividido em duas partes. A primeira delas se estende até o final da seção 6.2, incluindo alguns exercícios da 6.4. A segunda vai da seção 6.3 até o final do livro.
Nossa abordagem é distinta da maioria dos livros-texto comumente adotados. Na primeira parte (supracitada), tentamos fazer uma transição suave da *geometria analítica* (GA) em $\mathbb{R}^2$ e $\mathbb{R}^3$ para a AL do $\mathbb{R}^n$. Nessa "passagem de bastão" para $n = 1, 2, 3, \ldots$, o escopo é mais geométrico, com alguns resultados mais abstratos, embora exaustivamente trabalhados, não demonstrados ou demonstrados apenas nos últimos capítulos. Assim, esperando que o(a) leitor(a) tenha absorvido essa generalização,[2] fazemos depois outra transição suave, do $\mathbb{R}^n$ para o *espaço* $\mathbb{R}^{m\times n}$ das *matrizes com m linhas e n colunas*. Segue, desse ponto, um "movimento inercial" para *escalonamento*, *sistemas lineares* e um estudo operacional, similar ao de um formulário de algum manual de tabelas e fórmulas, de *determinantes*, com ênfase (ape-

[1] www.ufpr.br/∼jrrb. A propósito, nesse endereço eletrônico, os leitores poderão encontrar a errata desse livro.
[2] Inclusive, sendo apresentada, no final do capítulo 2, parte de uma "AL no $\mathbb{R}^3$".

nas) na resolução de exercícios.[3] Na sequência, entramos "matricialmente" no estudo de *função linear* e sua *diagonalização*, sem contudo nos afastarmos do aspecto geométrico. Por fim, só depois dessa abordagem mais aplicada/concreta, é que apresentamos *escalares* e "espaços" mais gerais. Em todo esse trajeto, a abstração vai aumentando gradualmente, como seria natural para aqueles que estão se ambientando com algum conhecimento novo.
Por terem um "sabor" semelhante ao da primeira parte desse livro, recomendo os seguintes livros:

VETORES E MATRIZES UMA INTRODUÇÃO À ÁLGEBRA LINEAR NATHAN MOREIRA DOS SANTOS 4A. EDIÇÃO - 2007 THOMSON

e

INTRODUÇÃO Á ÁLGEBRA LINEAR GILBERT STRANG TRADUÇÃO DA 4A. EDIÇÃO NORTE-AMERICANA LTC

.

Aqui, cada nota de rodapé tem um papel importante (e *deve* ser lida como parte integrante do texto) para quem estiver cursando AL pela primeira vez. O mesmo vale para demonstrações de alguns resultados e resoluções, dicas, sugestões e respostas de alguns exercícios quando o texto estiver escrito no tamanho de uma nota de rodapé. Por outro lado, para quem já cursou AL, a leitura pode ser feita em ritmo de revisão.
Deliberadamente, não incluí, ou adiei, demonstrações de alguns resultados na primeira parte do livro, pois ela foi concebida para cursos mais aplicados (engenharias, por exemplo). Leitores interessados em preencher essas lacunas são convidados a recorrer à outros livros da área (como os supracitados) ou tentar entender antecipadamente a segunda parte desse livro, onde adotamos uma abordagem de um segundo curso, ou de um curso *honors*, de AL e demonstramos resultados que tinham sido apenas enunciados nos capítulos e seções anteriores. Além disso, na segunda parte, resultados (tradicionalmente) demonstrados via determinantes (em outros livros) são obtidos via *determinant-free proofs*, na linha do artigo

DOWN WITH DETERMINANTS SHELDON AXLER https://www.maa.org/sites/default/files/pdf/awards/Axler-Ford-1996.pdf

.

Portanto, essas "lições" também podem ser utilizadas por alunos das exatas, ficando a primeira parte para aplicações/exemplos da segunda.
Em linhas gerais, para essa segunda parte, seguimos a abordagem do excelente livro:

LINEAR ALGEBRA DONE RIGHT SHELDON AXLER 3RD EDITION SPRINGER VERLAG

,

[3] Convém ressaltar que, a rigor, a utilidade da teoria dos determinantes é apenas teórica. Na prática, por exemplo, a obtenção dos *autovalores* de uma *matriz n por n* arbitrária, caso n seja "suficientemente grande", calculando-se as *raízes do polinômio característico* dessa matriz, é uma tarefa (a "tempo polinomial") com quase nenhuma possibilidade de sucesso para a complexidade computacional (binária) atual. Mesmo que n não seja tão grande, o custo computacional desse cálculo é muito alto. Por outro lado, para $n = 2, 3, 4$, por exemplo, é muito provável que os estudantes consigam realizar o cálculo dos determinantes com o conhecimento adquirido no ensino médio.

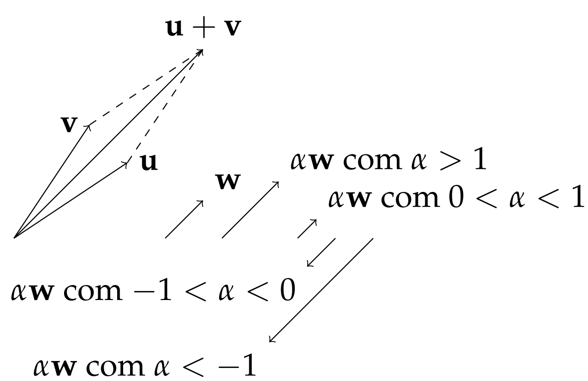
doravante referenciado simplesmente por AXLER.
O pré-requisito para a leitura da primeira parte (supracitada) é um curso de GA. Aliás, inicio o próximo capítulo com uma revisão de GA no $\mathbb{R}^2$. Falando em pré-requisitos, gostaria de expressar que vejo a matemática como uma linguagem tipo português, inglês, francês, etc. Portanto, temos também "matematiquês", "fisiquês", "quimiquês", "informatiquês", etc. O aprendizado de uma língua não ocorre sem antes sermos alfabetizado nela. Já nessa etapa preliminar, é preciso estudá-la e praticá-la, para não cometermos equívocos no uso da mesma. Note que, não é fácil fazermos um estudo avançado da língua sem termos sido alfabetizado nela. Como diz o ditado: "O avançado é fazer o básico bem feito!". Por outro lado, para conseguirmos fluência numa língua é preciso, além do estudo e da prática, conhecermos todo um jargão da área. Apenas estudarmos na proximidade de cada prova é perda de tempo para quase todos que assim procedem.
Sugestões para o aprimoramento e/ou a clareza desse livro serão muito bem vindas. Serei, não só grato, mas também todo "ouvidos e olhos".
Dedico esse pequeno trabalho ao meu filho THEO.

Capítulo 2

O espaço vetorial $\mathbb{R}^n$

2.1 Geometria analítica do $\mathbb{R}^2$

Figura 2.1: Adição de dois vetores e multiplicação de um vetor por um escalar



Em GA, define-se o *espaço* $\mathbb{R}^2$ dos *vetores* $\mathbf{u}, \mathbf{v}, \mathbf{w}$, etc.,[1] dotado de duas operações definidas *coordenada-a-coordenada*: a *adição de* $\mathbf{u}$ *e* $\mathbf{v}$, denotada por

$$\mathbf{u}+\mathbf{v},$$

e a *multiplicação de* $\mathbf{w}$ *pelo escalar* α,[2] denotada por

$$\alpha\mathbf{w},$$

conforme ilustradas na figura 2.1. Além disso, o espaço supracitado é dotado do *produto interno* (ou *escalar*) *de* $\mathbf{u}$ *e* $\mathbf{v}$, denotado por

$$\mathbf{u}\cdot\mathbf{v}$$

e calculado pela soma dos produtos das coordenadas respectivas de $\mathbf{u}$ e $\mathbf{v}$, e do *módulo* (ou *comprimento*) *de* $\mathbf{w}$, definido por

$$\|\mathbf{w}\| := \sqrt{\mathbf{w}\cdot\mathbf{w}} \ \text{ u.c.}^3$$

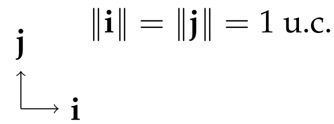
[1] Esses vetores são representados por pares ordenados de números reais.
[2] α é um número real.
[3] Ou seja, unidades de comprimento.

> *Obviamente, podemos calcular* $\mathbf{u}+\mathbf{w}$, $\mathbf{v}+\mathbf{w}$, $\alpha\mathbf{u}$, $\alpha\mathbf{v}$, $\mathbf{u}\cdot\mathbf{w}$, $\mathbf{v}\cdot\mathbf{w}$, $\|\mathbf{u}\|$, $\|\mathbf{v}\|$ *e, no lugar de* $\mathbf{u},\mathbf{v},\mathbf{w}$ *e* α, *utilizar* $\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{a}$, *etc., como vetores e* λ, a, x, t, *etc., como escalares.*

Todas essas operações e suas propriedades,[4] bem como demonstrações ou justificativas que as corroborem, devem ter sido vistas tanto geometricamente,[5] quanto algebricamente e numericamente.
Para os exercícios dessa seção, considere:

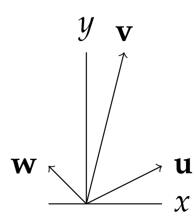
- $\mathbf{i} = (1,0)$ e $\mathbf{j} = (0,1)$, conforme a figura 2.2.

Figura 2.2: "Base canônica" de $\mathbb{R}^2$



- Em geral, $\mathbf{u}$, $\mathbf{v}$ e $\mathbf{w}$ têm a origem do plano cartesiano como ponto inicial, conforme a figura 2.3.

Figura 2.3: Vetores com pontos iniciais na origem



- Ângulos serão medidos em radianos. Contudo, eventuais respostas poderão vir em graus.

2.1.1 Exercícios sobre vetores em $\mathbb{R}^2$

1. Para qual valor de x os vetores $\mathbf{u} = \left(1, x^2-1\right)$ e $\mathbf{v} = (x+2, 0)$ verificam a igualdade $\mathbf{u} = \mathbf{v}$? RESPOSTA $x = -1$.

[4]Comutatividade tanto da adição quanto do produto interno de vetores; o módulo do múltiplo escalar de um vetor iguala o produto do módulo desse escalar pelo módulo desse vetor; $\mathbf{0} = (0,0)$ é o *elemento neutro* aditivo; etc.

[5]Por exemplo, a comutatividade da adição de vetores está ilustrada na figura 2.1.

2. Se o vetor $\mathbf{u}$ tem módulo igual a 3 u.c. e o vetor $\mathbf{v}$ tem módulo igual a 2 u.c., qual é o maior (respectivamente, menor) valor que o módulo da soma $\mathbf{u}+\mathbf{v}$ pode assumir? RESPOSTA 1 u.c. $\leq \|\mathbf{u}+\mathbf{v}\| \leq 5$ u.c.[6]

3. Sejam $\mathbf{v} \neq \mathbf{0}$ e $\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$.

   (a) Verifique que $\mathbf{u}$ é *unitário*, isto é, $\|\mathbf{u}\| = 1$ u.c., e de mesma direção e mesmo sentido que $\mathbf{v}$.[7]

   (b) Determine $\mathbf{u}$ se $\mathbf{v} = (-8, 6)$. RESPOSTA $\mathbf{u} = (-4/5, 3/5)$.

4. Deve ter sido visto em GA que, para quaisquer vetores $\mathbf{u}$, $\mathbf{v}$ e $\mathbf{w}$ e para cada escalar α, as seguintes propriedades são válidas:

   · $(\mathbf{v}+\mathbf{u}) \cdot \mathbf{w} = \mathbf{v} \cdot \mathbf{w} + \mathbf{u} \cdot \mathbf{w}$; (DISTRIBUTIVIDADE DO PRODUTO EM RELAÇÃO À SOMA)

   · $\mathbf{v} \cdot \mathbf{u} = \mathbf{u} \cdot \mathbf{v}$; (COMUTATIVIDADE DO PRODUTO)

   · $\mathbf{v} \cdot (\alpha\mathbf{u}) = (\alpha\mathbf{u}) \cdot \mathbf{v} = \alpha\,(\mathbf{u} \cdot \mathbf{v})$; (ASSOCIATIVIDADE)

   · $\|\mathbf{w}\| = \sqrt{\mathbf{w} \cdot \mathbf{w}}$. (DEFINIÇÃO DE MÓDULO)

   Sejam $\mathbf{u}$ e $\mathbf{v}$ unitários. Utilize as propriedades supracitadas para calcular o produto interno dos vetores dados em cada um dos itens seguintes:

   (a) $\mathbf{u}$ e $-\mathbf{u}$. RESPOSTA -1.

   (b) $\mathbf{v}+\mathbf{u}$ e $\mathbf{v}-\mathbf{u}$. RESPOSTA 0.[8]

---

[6] SUGESTÃO

Por um lado, como deve ser de conhecimento comum, vale a seguinte desigualdade triangular:

$$\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|.$$

Por outro, essa desigualdade aplicada à soma

$$\mathbf{u} = (\mathbf{u}+\mathbf{v}) + (-\mathbf{v})$$

e o uso da igualdade

$$\|-\mathbf{v}\| = \|\mathbf{v}\|$$

acarretam a desigualdade

$$\|\mathbf{u}\| - \|\mathbf{v}\| \leq \|\mathbf{u}+\mathbf{v}\|.$$

[7] SUGESTÕES

Para $\mathbf{v} = (x, y)$, determine $\mathbf{u}$. Então, calcule $\|\mathbf{u}\|$. Para outra resolução, como $\mathbf{v} \neq \mathbf{0}$, considere $\alpha = \frac{1}{\|\mathbf{v}\|}$ e calcule, portanto, o módulo de $\mathbf{u} = \alpha\mathbf{v}$.

[8] RESOLUÇÃO

$$\begin{aligned}
(\mathbf{v}+\mathbf{u}) \cdot (\mathbf{v}-\mathbf{u}) &= (\mathbf{v}+\mathbf{u}) \cdot (\mathbf{v}+(-1)\mathbf{u}) \\
&= \mathbf{v} \cdot (\mathbf{v}+(-1)\mathbf{u}) + \mathbf{u} \cdot (\mathbf{v}+(-1)\mathbf{u}) \quad \text{(DIST.)} \\
&= \mathbf{v} \cdot \mathbf{v} + (-1)\,(\mathbf{u} \cdot \mathbf{v}) + \mathbf{u} \cdot \mathbf{v} + (-1)\,(\mathbf{u} \cdot \mathbf{u}) \quad \text{(DIST., COMUT., ASSOC.)} \\
&= \|\mathbf{v}\|^2 - \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{v} - \|\mathbf{u}\|^2 \quad \text{(DEF. MÓD.)} \\
&= 1 - 0 - 1 \quad (\mathbf{u}, \mathbf{v} \text{ UNITÁRIOS}) \\
&= 0.
\end{aligned}$$

(c) $\mathbf{v} - 2\mathbf{u}$ e $\mathbf{v} + 2\mathbf{u}$. $\boxed{\text{RESPOSTA}}$ -3.

5. O *ângulo* θ entre dois vetores não nulos, $\mathbf{u}$ e $\mathbf{v}$, de medida entre 0 e π radianos, satisfaz a condição

$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}. \tag{2.1}$$

(a) Verifique a validade de (2.1) para $\mathbf{u}$ e $\mathbf{v}$ unitários e tais que:[9]

i. o ângulo entre $\mathbf{u}$ e $\mathbf{i}$ mede $\pi/3$, o ângulo entre $\mathbf{v}$ e $\mathbf{i}$ mede $2\pi/3$ e $\theta = \pi/3$;

ii. o ângulo entre $\mathbf{u}$ e $\mathbf{i}$ mede $\pi/4$, o ângulo entre $\mathbf{v}$ e $\mathbf{i}$ mede $3\pi/4$ e $\theta = \pi/2$.

(b) Verifique a validade de (2.1) em geral, a partir das duas etapas seguintes:

i. Esboce o gráfico da função $f(\theta) = \cos\theta$ para θ entre 0 e π radianos. Observe que, para cada número real r entre -1 e 1 (no eixo das ordenadas), existe um único $\theta = \theta(r)$ entre 0 e π (no eixo das abcissas) com $r = f(\theta)$;

ii. Utilize a desigualdade de Cauchy-Schwarz,[10]

$$|\mathbf{u}\cdot\mathbf{v}| \le \|\mathbf{u}\|\,\|\mathbf{v}\|,$$

para demonstrar que

$$\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}$$

é um número r entre -1 e 1.

6. Considere que $\mathbf{u} = (x, y) \in \mathbb{R}^2$ é unitário e θ é o ângulo entre $\mathbf{u}$ e $\mathbf{i}$.

(a) Verifique:

$$\begin{aligned}\mathbf{u} &= (\cos\theta, \operatorname{sen}\theta)\\ &= \cos\theta\,\mathbf{i} + \operatorname{sen}\theta\,\mathbf{j}.\end{aligned}$$

(b) Determine $\mathbf{u}$ para cada θ dado a seguir:

i. 0;

ii. $\pi/6$;

iii. $\pi/4$;

iv. $\pi/3$;

v. $\pi/2$;

vi. $3\pi/4$.

7. Seja $\mathbf{v} = (x, y)$. Seja θ o ângulo entre $\mathbf{v}$ e $\mathbf{i}$.

(a) Verifique que $\tan\theta = \frac{y}{x}$.[11]

---

[9] $\boxed{\text{DICA}}$
Utilize o exercício 6 dessa subseção, para obter as coordenadas de $\mathbf{u}$ e $\mathbf{v}$.

[10] Possivelmente estudada em GA.

[11] $\boxed{\text{SUGESTÃO}}$
Sem perda de generalidade, suponha que $\mathbf{v}$ é unitário. Agora, aplique o item (a) do exercício 6 dessa subseção.

(b) Determine θ para $\mathbf{v} = -4\mathbf{i} + 3\mathbf{j}$.

(c) Determine θ para $\mathbf{v} = \mathbf{i} - \mathbf{j}$.

8. Calcule

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \, \|\mathbf{v}\| \cos\theta,[12]$$

caso:

(a) $\mathbf{u}$ e $\mathbf{v}$ sejam unitários e representem lados de um triângulo equilátero;

(b) $\mathbf{u}$ seja unitário, esteja na bissetriz do primeiro quadrante e $\mathbf{v} = -\frac{1}{2}\mathbf{i}$.

9. Dizemos que dois vetores não nulos, $\mathbf{u}$ e $\mathbf{v}$, são *ortogonais (entre si)* quando

$$\mathbf{u} \cdot \mathbf{v} = 0.[13]$$

Denotamos a ortogonalidade entre esses vetores por $\mathbf{u} \perp \mathbf{v}$.

(a) Considere $\mathbf{u} = \mathbf{i}$, $\mathbf{v} = \mathbf{j}$ e $\mathbf{u} + \mathbf{v}$. Por um lado, verifique que $\mathbf{u} \perp \mathbf{v}$. Por outro, calcule $\|\mathbf{u}\|$, $\|\mathbf{v}\|$ e $\|\mathbf{u}+\mathbf{v}\|$. Assim, verifique que $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

(b) Faça como no item (a) (dessa questão), considerando, agora, $\mathbf{u} = \left(\sqrt{2}, \sqrt{2}\right)$ e $\mathbf{v} = \left(-\sqrt{2}, \sqrt{2}\right)$.

(c) Utilize o *teorema de Pitágoras*, estudado na *geometria plana*, para demonstrar que, caso $\mathbf{u}$ e $\mathbf{v}$ sejam ortogonais em $\mathbb{R}^2$,

$$\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 . \tag{2.2}$$

(d) Demonstre (2.2) sem utilizar o teorema supracitado.[14]

10. Obtenha a *equação vetorial da reta* r *que passa pelo ponto (final de)* $\mathbf{x}_0$ *com vetor diretor* (ou *na direção do vetor*) $\mathbf{a}$, isto é,

$$\mathrm{r}: \quad \mathbf{x} = \mathbf{x}_0 + t\mathbf{a}, \quad t \in \mathbb{R},$$

para:

(a) $\mathbf{x}_0 = (1,2)$ e $\mathbf{a} = (1,1)$;

(b) $\mathbf{x}_0 = (1,-2)$ e $\mathbf{a} = (-1,2)$;

(c) $\mathbf{x}_0 = (2,2)$ e $\mathbf{a} = \mathbf{i}$;

(d) $\mathbf{x}_0 = (2,2)$ e $\mathbf{a} = \mathbf{j}$.

---

[12] Cf. (2.1), p. 14.

[13] Confira o exercício 8 dessa subseção.

[14] SUGESTÃO

Utilize o seguinte resultado:

$$\|\mathbf{u}+\mathbf{v}\|^2 = (\mathbf{u}+\mathbf{v}) \cdot (\mathbf{u}+\mathbf{v}) .$$

Além disso, em cada item dessa questão, quando possível, determine a equação *afim* $y = ax + b$ da reta r obtida.[15]

11. Caso r e s sejam duas retas com vetores diretores $\mathbf{v}$ e $\mathbf{w}$, respectivamente, elas são:

    · *paralelas* quando seus vetores diretores são *múltiplos escalares* um do outro, ou seja, para algum escalar não nulo α,
    $$\mathbf{w} = \alpha\mathbf{v}.$$

    · *perpendiculares* quando $\mathbf{v} \perp \mathbf{w}$.

    Dê exemplos de retas que sejam paralelas (respectivamente, perpendiculares). Escreva as equações vetoriais dessas retas.

[15] $\boxed{\text{DICA}}$
A equação afim não pode ser obtida no item (d).

2.2 $\mathbb{R}^n$, o espaço euclidiano n-dimensional

Além de vetores em $\mathbb{R}^2$, podemos considerá-los em $\mathbb{R}^3$ e, em geral, em $\mathbb{R}^n$, onde n é um inteiro positivo arbitrário. Passamos, portanto, do plano para o espaço ou hiperespaço. Nesse capítulo, estudaremos esses (hiper)espaços. No capítulo 3, veremos que esses espaços e os espaços das matrizes são "indistinguíveis".

2.2.1 Notações e definições iniciais

- $\mathbf{x} \in \mathbb{R}^n$ representa uma *n-upla ordenada*, cujas *coordenadas* (ou *componentes*) são os números (reais) da lista $x_1, \ldots, x_n$, nessa ordem, isto é,

  $$\mathbf{x} = (x_1, \ldots, x_n).$$

  $\mathbf{x}$ também pode representar uma *matriz* $n \times 1$, cujas *entradas* (da sua única *coluna*) são os números da lista supracitada, dada por

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

  EXEMPLO

  Em $\mathbb{R}^4$, podemos escrever

  $$\mathbf{x} = (1,2,3,4) \quad \text{ou} \quad \mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}.$$

  Caso tenha n coordenadas, diremos que $\mathbf{x}$ é um *vetor* em/do/de $\mathbb{R}^n$.

  Sem perda de generalidade, nesse capítulo, assim como nos capítulos 3 e 4, "vetor" significará "vetor em $\mathbb{R}^n$".

- Do mesmo modo que representamos $\mathbf{x}$, um vetor denotado por outra letra, digamos $\mathbf{y}$, pode ser representado por

  $$\mathbf{y} = (y_1, \ldots, y_n) \;\text{ ou }\; \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

- A palavra "ordenada" supracitada tem relação com a ordem das coordenadas dos vetores, ou seja, caso $\mathbf{x}$ e $\mathbf{y}$ sejam dois vetores tais que $\mathbf{x} = \mathbf{y}$,

  $$x_i = y_i, \;\; i = 1, \ldots, n.$$

  EXEMPLO

  Em $\mathbb{R}^2$, $(1,2) \neq (2,1)$.

- Em $\mathbb{R}^n$, a *soma* dos vetores $\mathbf{x}$ e $\mathbf{y}$ é o vetor $\mathbf{x}+\mathbf{y}$, cuja i-ésima coordenada é dada por

$$x_i + y_i, \quad i = 1, \ldots, n.$$

EXEMPLO

Em $\mathbb{R}^3$, se $\mathbf{x} = (1,2,3)$ e $\mathbf{y} = (-1,1/2,1)$, então $\mathbf{x}+\mathbf{y} = (0,5/2,4)$.

- Dados o vetor $\mathbf{x} \in \mathbb{R}^n$ e o *escalar* $\alpha \in \mathbb{R}$, o vetor $\alpha\mathbf{x}$, cuja i-ésima coordenada é dada por

$$\alpha x_i, \quad i = 1, \ldots, n,$$

é chamado de *produto por escalar*.

EXEMPLO

Em $\mathbb{R}^2$, se $\alpha = \frac{1}{3}$ e $\mathbf{x} = \begin{bmatrix} 3/2 \\ 3 \end{bmatrix}$, então $\alpha\mathbf{x} = \begin{bmatrix} 1/2 \\ 1 \end{bmatrix}$.

2.2.2 Propriedades que caracterizam $\mathbb{R}^n$ como um "espaço vetorial"

Para quaisquer vetores $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{z}$ em $\mathbb{R}^n$ e escalares α e β em $\mathbb{R}$, as seguintes propriedades são válidas:

1. $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$; (COMUTATIVIDADE DA ADIÇÃO)
2. $(\mathbf{x}+\mathbf{y})+\mathbf{z} = \mathbf{x}+(\mathbf{y}+\mathbf{z})$; (ASSOCIATIVIDADE DA ADIÇÃO)
3. $\mathbf{0} = (0,\ldots,0) \in \mathbb{R}^n$ é tal que $\mathbf{x}+\mathbf{0} = \mathbf{x}$; (EXISTÊNCIA DO VETOR NULO)
4. $-\mathbf{x} = (-1)\mathbf{x}$ é tal que $\mathbf{x}+(-\mathbf{x}) = \mathbf{0}$; (EXISTÊNCIA DE VETOR SIMÉTRICO)
5. $\alpha(\mathbf{x}+\mathbf{y}) = \alpha\mathbf{x}+\alpha\mathbf{y}$; (DISTRIBUTIVIDADE DO PRODUTO (POR ESCALAR) EM RELAÇÃO À SOMA DE VETORES)
6. $(\alpha+\beta)\mathbf{x} = \alpha\mathbf{x}+\beta\mathbf{x}$; (DISTRIBUTIVIDADE DO PRODUTO (POR ESCALAR) EM RELAÇÃO À SOMA DOS ESCALARES)
7. $(\alpha\beta)\mathbf{x} = \alpha(\beta\mathbf{x})$; (ASSOCIATIVIDADE DO PRODUTO POR ESCALAR)
8. $1\mathbf{x} = \mathbf{x}$. (ELEMENTO NEUTRO DO PRODUTO POR ESCALAR)

DEMONSTRAÇÃO DA PROPRIEDADE 1

Como visto no início dessa seção, a igualdade dos vetores $\mathbf{x}+\mathbf{y}$ e $\mathbf{y}+\mathbf{x}$ é equivalente a igualdade das suas i-ésimas componentes, $i = 1,\ldots,n$. Essas i-ésimas componentes são dadas por $x_i + y_i$ e $y_i + x_i$, $i = 1,\ldots,n$. Portanto, como a adição em $\mathbb{R}$ é comutativa,[16] temos $x_i + y_i = y_i + x_i$, $i = 1,\ldots,n$. Portanto, $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$.

EXERCÍCIO

Demonstre as propriedades 2–8 supracitadas.

[16] Lembrem-se do "mantra": *a ordem das parcelas não altera a soma.*

2.3 Produto interno, módulo, ângulo e ortogonalidade em $\mathbb{R}^n$

Produto interno. O número real

$$\mathbf{x} \cdot \mathbf{y} := x_1 y_1 + \cdots + x_n y_n$$

é chamado de *produto interno* (ou *escalar*) dos vetores $\mathbf{x}$ e $\mathbf{y}$.

EXEMPLO

Em $\mathbb{R}^3$, se $\mathbf{x} = \begin{bmatrix} \ln 2 \\ 3/\sqrt{2} \\ -3 \end{bmatrix}$ e $\mathbf{y} = \begin{bmatrix} -1 \\ 1 \\ \cos\frac{\pi}{4} \end{bmatrix}$, então $\mathbf{x} \cdot \mathbf{y} = \ln(1/2)$.

OBSERVAÇÃO

Para quaisquer vetores $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{z}$ em $\mathbb{R}^n$, valem as seguintes propriedades:

1. $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$; (COMUTATIVIDADE DO PRODUTO)
2. $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$; (DISTRIBUTIVIDADE DO PRODUTO (INTERNO) EM RELAÇÃO À SOMA DE VETORES)
3. $\alpha \in \mathbb{R} \Longrightarrow \alpha(\mathbf{x} \cdot \mathbf{y}) = (\alpha\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (\alpha\mathbf{y})$; (ASSOCIATIVIDADE)[17]
4. $\mathbf{x} \cdot \mathbf{x} \geq 0$ e $\mathbf{x} \cdot \mathbf{x} = 0$ se, e somente se, $\mathbf{x} = \mathbf{0}$; (NÃO NEGATIVIDADE)
5. $\mathbf{x} \cdot \mathbf{0} = 0$. (VETOR ANULADOR)

EXERCÍCIO

Demonstre as propriedades supracitadas.[18]

Módulo. O número real não negativo

$$\begin{aligned} \|\mathbf{x}\| &:= \sqrt{\mathbf{x} \cdot \mathbf{x}} \\ &= \sqrt{x_1^2 + \cdots + x_n^2} \end{aligned}$$

é chamado de *módulo* (ou *norma* ou *comprimento*) do vetor $\mathbf{x}$.

EXEMPLO

[17] Na segunda igualdade, utilizamos a comutatividade supracitada, antes e depois da associatividade.

[18] A propriedade do vetor anulador pode ser demonstrada diretamente da definição de produto escalar ou da distributividade supracitada. De fato,

$$\begin{aligned} \mathbf{x} \cdot \mathbf{0} &= \mathbf{x} \cdot (\mathbf{0} + \mathbf{0}) \\ &= \mathbf{x} \cdot \mathbf{0} + \mathbf{x} \cdot \mathbf{0} \end{aligned}$$

e, para qualquer $x \in \mathbb{R}$, 0 é o único número que satisfaz a equação $x = x + 0$.

Em $\mathbb{R}^4$, se $\mathbf{x} = \begin{bmatrix} 3 \\ 4 \\ -\sqrt{5} \\ \sqrt{6} \end{bmatrix}$, então

$$\begin{aligned} \|\mathbf{x}\| &= \sqrt{3^2 + 4^2 + \left(-\sqrt{5}\right)^2 + \sqrt{6}^2} \\ &= \sqrt{9 + 16 + 5 + 6} \\ &= 6 \text{ u.c.} \end{aligned}$$

OBSERVAÇÃO

Para quaisquer vetores $\mathbf{x}$ e $\mathbf{y}$ em $\mathbb{R}^n$ e cada escalar $\lambda \in \mathbb{R}$, valem as seguintes propriedades:

1. $\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|$; (O MÓDULO DO PRODUTO IGUALA O PRODUTO DOS MÓDULOS)
2. $\|\mathbf{x}\| \geq 0$ e $\|\mathbf{x}\| = 0$ se, e somente se, $\mathbf{x} = \mathbf{0}$; (NÃO NEGATIVIDADE)
3. $|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|$; (DESIGUALDADE DE CAUCHY-SCHWARZ)
4. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$. (DESIGUALDADE TRIANGULAR)

DEMONSTRAÇÃO DA PROPRIEDADE 3

Se $\mathbf{y} = \mathbf{0}$, ambos os lados da desigualdade se anulam. Assim, seja $\mathbf{y} \neq \mathbf{0}$. Como

$$0 \leq (\mathbf{x} + \lambda \mathbf{y}) \cdot (\mathbf{x} + \lambda \mathbf{y}) = \mathbf{x} \cdot \mathbf{x} + 2\lambda (\mathbf{x} \cdot \mathbf{y}) + \lambda^2 (\mathbf{y} \cdot \mathbf{y}),$$

em particular, se $\lambda = -\frac{\mathbf{x} \cdot \mathbf{y}}{\mathbf{y} \cdot \mathbf{y}}$, temos

$$0 \leq \mathbf{x} \cdot \mathbf{x} - 2\frac{(\mathbf{x} \cdot \mathbf{y})^2}{\mathbf{y} \cdot \mathbf{y}} + \frac{(\mathbf{x} \cdot \mathbf{y})^2}{\mathbf{y} \cdot \mathbf{y}} = \mathbf{x} \cdot \mathbf{x} - \frac{(\mathbf{x} \cdot \mathbf{y})^2}{\mathbf{y} \cdot \mathbf{y}}. \tag{2.3}$$

Portanto, multiplicando-se (2.3) por $\mathbf{y} \cdot \mathbf{y}$, temos

$$0 \leq (\mathbf{x} \cdot \mathbf{x})(\mathbf{y} \cdot \mathbf{y}) - (\mathbf{x} \cdot \mathbf{y})^2,$$

ou seja,

$$|\mathbf{x} \cdot \mathbf{y}|^2 \leq \|\mathbf{x}\|^2 \|\mathbf{y}\|^2.$$

DEMONSTRAÇÃO DA PROPRIEDADE 4

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) \\ &= \mathbf{x} \cdot \mathbf{x} + 2\,\mathbf{x} \cdot \mathbf{y} + \mathbf{y} \cdot \mathbf{y} \\ &= \|\mathbf{x}\|^2 + 2\,\mathbf{x} \cdot \mathbf{y} + \|\mathbf{y}\|^2 \\ &\leq \|\mathbf{x}\|^2 + 2|\mathbf{x} \cdot \mathbf{y}| + \|\mathbf{y}\|^2 \\ &\leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\| \|\mathbf{y}\| + \|\mathbf{y}\|^2, \end{aligned}$$

onde utilizamos:

- na primeira e terceira igualdades, a definição de norma;
- na segunda igualdade, a distributividade do produto interno em relação à soma de vetores e a comutatividade do produto interno;
- na primeira desigualdade, que $t \leq |t|$ para cada real t;
- na última desigualdade, a desigualdade de Cauchy-Schwarz.

Agora, basta aplicarmos raiz quadrada em $\|\mathbf{x}+\mathbf{y}\|^2 \leq (\|\mathbf{x}\|+\|\mathbf{y}\|)^2$.

EXERCÍCIO

Demonstre as propriedades 1 e 2 supracitadas.

Ângulo. Caso $\mathbf{x}$ e $\mathbf{y}$ sejam vetores não nulos em $\mathbb{R}^n$,

$$-1 \leq \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|} \leq 1,$$

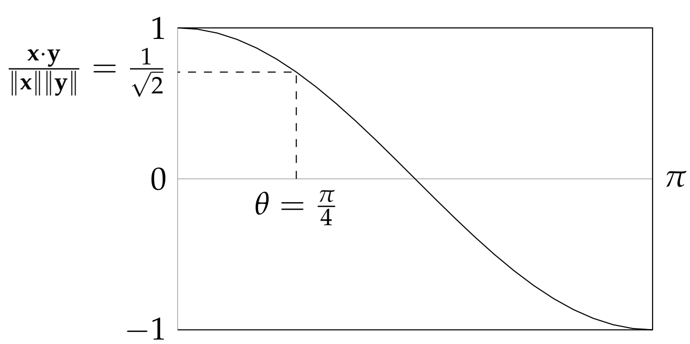
pela desigualdade de Cauchy-Schwarz e pela propriedade da não negatividade de normas.[19] Portanto, como $\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}$ é um número entre -1 e 1, existe um único ângulo $\theta = \theta(\mathbf{x},\mathbf{y})$ (em radianos) entre 0 e π tal que

$$\cos\theta = \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}.$$

Definimos esse θ como o *ângulo* entre os vetores $\mathbf{x}$ e $\mathbf{y}$, conforme ilustrado na figura 2.4 para

$$\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|} = \frac{\sqrt{2}}{2} \ \text{ e } \ \theta = \frac{\pi}{4}.$$

Figura 2.4: θ é o ângulo entre $\mathbf{x}$ e $\mathbf{y}$



EXEMPLO

Em $\mathbb{R}^4$, para $\mathbf{x} = (1,-1,0,2)$ e $\mathbf{y} = \left(-1,1,\frac{1}{2},-2\right)$, $\mathbf{x}\cdot\mathbf{y} = -6$, $\|\mathbf{x}\| = \sqrt{6}$ e $\|\mathbf{y}\| = \frac{5}{2}$ u.c. Então,

$$\begin{aligned}\cos\theta &\approx \frac{-6}{2,45\cdot 2,5}\\ &\approx -1.\end{aligned}$$

Portanto, θ está próximo de π radianos.

[19]Nesse caso, pela positividade das normas!

Note que, agora, podemos calcular o produto interno (de vetores $\mathbf{x}$ *e* $\mathbf{y}$ *arbitrários) por*

$$\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos\theta.$$

Ortogonalidade. $\mathbf{x}$ e $\mathbf{y}$ são ditos *ortogonais (entre si)* quando $\mathbf{x} \cdot \mathbf{y} = 0$, ou seja, um deles é o vetor nulo ou o ângulo entre eles é $\frac{\pi}{2}$ radianos.[20]

EXEMPLO IMPORTANTE DE VETORES ORTOGONAIS

Considere $i \in \{1, \dots, n\}$ e que $\mathbf{e}_i$ seja o vetor do $\mathbb{R}^n$ cuja i-ésima coordenada valha 1 e essa seja a sua única coordenada não nula. Portanto, se $j \in \{1, \dots, n\}$, então

$$\mathbf{e}_i \cdot \mathbf{e}_j = \begin{cases} 0 & \text{se } i \neq j; \\ 1 & \text{se } i = j. \end{cases}$$

EXEMPLO

Em $\mathbb{R}^3$, como $\mathbf{e}_1 = \mathbf{i}$, $\mathbf{e}_2 = \mathbf{j}$ e $\mathbf{e}_3 = \mathbf{k}$ representam, respectivamente, as três primeiras colunas da *matriz identidade* 3×3,

$$\begin{aligned} \mathbf{e}_1 \cdot \mathbf{e}_1 = \mathbf{e}_2 \cdot \mathbf{e}_2 = \mathbf{e}_3 \cdot \mathbf{e}_3 = 1; \\ \mathbf{e}_1 \cdot \mathbf{e}_2 = \mathbf{e}_1 \cdot \mathbf{e}_3 = \mathbf{e}_2 \cdot \mathbf{e}_3 = 0. \end{aligned}$$

Utilize a comutatividade do produto interno, na última linha, para obter os três produtos faltantes.

2.4 Retas e hiperplanos em $\mathbb{R}^n$

Vamos generalizar os conceitos de reta e plano vistos em GA.
Sejam $\mathbf{x}_0$ e $\mathbf{a}$ vetores fixos, $\mathbf{a} \neq \mathbf{0}$, $\mathbf{x}$ um vetor arbitrário e t um escalar que pode assumir qualquer valor em $\mathbb{R}$. Portanto:

1. $\boxed{\mathbf{x} = \mathbf{x}_0 + t\mathbf{a}}$ representa a *reta* r *que passa pelo ponto (final de)* $\mathbf{x}_0$ *na direção do vetor* $\mathbf{a}$;

2. $\boxed{\mathbf{a} \cdot (\mathbf{x} - \mathbf{x}_0) = 0}$ representa o *(hiper)plano* Π *que passa pelo ponto (final de)* $\mathbf{x}_0$ *com normal* $\mathbf{a}$.

Por exemplo, em $\mathbb{R}^3$, se $\mathbf{x}_0 = (x_0, y_0, z_0)$, $\mathbf{a} = (a, b, c)$ e $\mathbf{x} = (x, y, z)$, as *equações paramétricas* de r e a *equação geral de* Π são dadas (respectivamente) por:

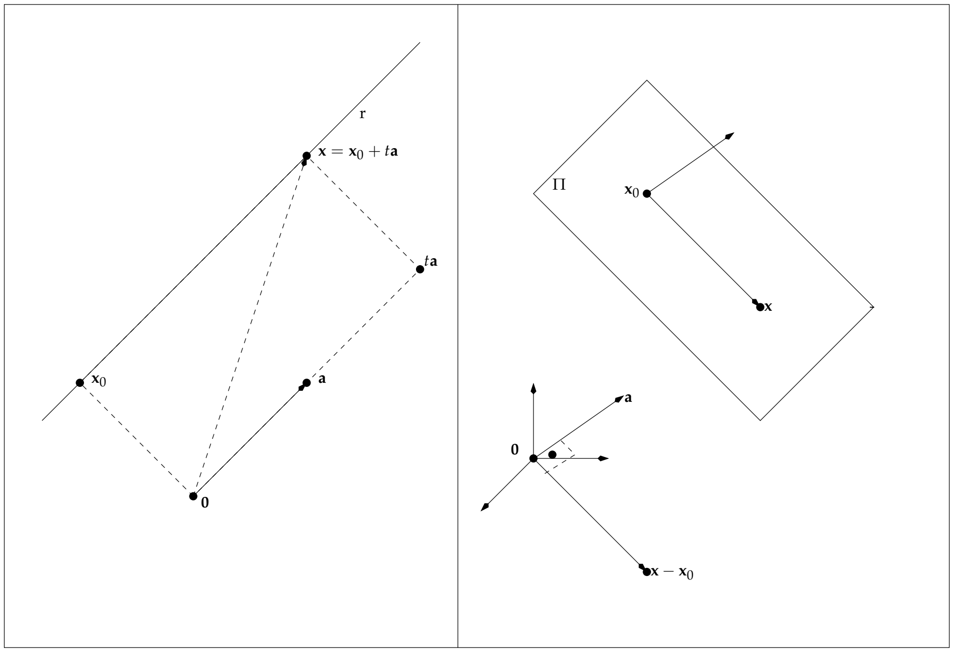
1. $\boxed{x = x_0 + ta,\ y = y_0 + tb,\ z = z_0 + tc}$, onde t representa um escalar arbitrário;

2. $\boxed{ax + by + cz = d}$, onde $d = ax_0 + by_0 + cz_0$.

Para uma ilustração, veja a figura 2.5.

[20] Por exemplo, em $\mathbb{R}^2$, $\mathbf{x} = (1, 1)$ e $\mathbf{y} = (-1, 1)$ são ortogonais.

Figura 2.5: Reta r e Plano Π em $\mathbb{R}^3$.



E quanto às retas e aos planos que "passam" pela origem?

Seja $\mathbf{x}_0 = \mathbf{0}$. Então, r e Π podem ser representados respectivamente por:

1. $\boxed{\mathbf{x} = t\mathbf{a}}$;

2. $\boxed{\mathbf{a} \cdot \mathbf{x} = 0}$.

Para o exemplo em $\mathbb{R}^3$ supracitado, temos:

1. $\boxed{x = ta,\ y = tb,\ z = tc}$;

2. $\boxed{ax + by + cz = 0}$.

2.5 Subespaços de $\mathbb{R}^n$

São subconjuntos $\mathcal{S}$ de $\mathbb{R}^n$ tais que:

1. $\mathbf{0} \in \mathcal{S}$;

2. $\alpha\mathbf{x} \in \mathcal{S}$, para cada escalar $\alpha \in \mathbb{R}$ e qualquer vetor $\mathbf{x} \in \mathcal{S}$;

3. $\mathbf{x} + \mathbf{y} \in \mathcal{S}$, para quaisquer vetores $\mathbf{x}, \mathbf{y} \in \mathcal{S}$.

EXEMPLOS

· A reta (que passa pela origem)

$$\mathcal{S} = \left\{ \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} \,\middle|\, y = x \right\} \tag{2.4}$$

é um subespaço de $\mathbb{R}^2$. De fato, $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \in \mathcal{S}$, pois as coordenadas do vetor nulo satisfazem a equação $y = x$, isto é,

$$x = y = 0 \Longrightarrow y = x.$$

Além disso, se $\alpha \in \mathbb{R}$ e $\mathbf{x} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \in \mathcal{S}$, ou seja,

$$\left\{ \begin{array}{rcl} y_1 & = & x_1, \\ y_2 & = & x_2, \end{array} \right. \tag{2.5}$$

então:

– $\alpha\mathbf{x} = \begin{bmatrix} \alpha x_1 \\ \alpha y_1 \end{bmatrix} \in \mathcal{S}$, pois, ao multiplicarmos a primeira igualdade de (2.5) por α, temos

$$\begin{aligned} y &= \alpha y_1 \\ &= \alpha x_1 \\ &= x; \end{aligned}$$

– $\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix} \in \mathcal{S}$, pois a soma dos membros correspondentes das igualdades de (2.5) é dada por

$$\begin{aligned} y &= y_1 + y_2 \\ &= x_1 + x_2 \\ &= x. \end{aligned}$$

· O plano (que passa pela origem)

$$\mathcal{S} = \left\{ \mathbf{x} = (x, y, z) \,\middle|\, x + y + z = 0 \right\}$$

é um subespaço de $\mathbb{R}^3$. De fato, $\mathbf{0} = (0, 0, 0) \in \mathcal{S}$, pois as coordenadas do vetor nulo satisfazem a equação $x + y + z = 0$, isto é,

$$x = y = z = 0 \Longrightarrow x + y + z = 0.$$

Além disso, se $\alpha \in \mathbb{R}$ e $\mathbf{x} = (x_1, y_1, z_1), \mathbf{y} = (x_2, y_2, z_2) \in \mathcal{S}$, isto é,

$$\left\{ \begin{array}{ccccccc} x_1 & + & y_1 & + & z_1 & = & 0, \\ x_2 & + & y_2 & + & z_2 & = & 0, \end{array} \right.$$

então:

– $\alpha\mathbf{x} = (\alpha x_1, \alpha y_1, \alpha z_1) \in \mathcal{S}$, pois

$$\begin{aligned}\alpha x_1 + \alpha y_1 + \alpha z_1 &= \alpha\,(x_1 + y_1 + z_1)\\ &= \alpha(0)\\ &= 0;\end{aligned}$$

– $\mathbf{x} + \mathbf{y} = (x_1 + x_2, y_1 + y_2, z_1 + z_2) \in \mathcal{S}$, pois

$$\begin{aligned}(x_1 + x_2) + (y_1 + y_2) + (z_1 + z_2) &= x_1 + x_2 + y_1 + y_2 + z_1 + z_2\\ &= x_1 + y_1 + z_1 + x_2 + y_2 + z_2\\ &= 0 + 0\\ &= 0.\end{aligned}$$

· O conjunto

$$\mathcal{S} = \left\{\mathbf{x} = (x, y, z) \;\middle|\; x - y = 0 \text{ e } z = 0\right\}$$

é um subespaço de $\mathbb{R}^3$. De fato, a reta (que passa pela origem)

$$\mathcal{S} = \left\{\mathbf{x} = t(1, 1, 0) \;\middle|\; t \in \mathbb{R}\right\}$$

representa a interseção dos planos (que passam pela origem) $x - y = 0$ e $z = 0$.[21] Além disso, é fácil ver que $\mathcal{S}$ é a reta (2.4), agora representada por um subconjunto do $\mathbb{R}^3$.

Os três exemplos anteriores não são casos isolados, pois, pelo exercício da subseção 2.5.1, quaisquer interseções de hiperplanos, inclusive retas e planos, passando pela origem, são subespaços de $\mathbb{R}^n$.

OBSERVAÇÃO

Para verificarmos que $\mathcal{S} \subset \mathbb{R}^n$ não é um subespaço de $\mathbb{R}^n$, é necessário que alguma das condições seguintes seja válida:

· $\mathbf{0} \notin \mathcal{S}$;

· $\alpha\mathbf{x} \notin \mathcal{S}$ para algum escalar α e algum vetor $\mathbf{x}$ de $\mathcal{S}$;

· $\mathbf{x} + \mathbf{y} \notin \mathcal{S}$ para algum par de vetores $\mathbf{x}$ e $\mathbf{y}$ de $\mathcal{S}$.

EXEMPLO

O plano

$$\mathcal{S} = \left\{\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \;\middle|\; x + y + z = 1\right\}$$

não é um subespaço de $\mathbb{R}^3$, por várias razões. Daremos três:[22]

· $\mathbf{0} \notin \mathcal{S}$, pois suas coordenadas são tais que $0 + 0 + 0 \neq 1$;

· Se $\alpha = -1$ e $\mathbf{x} = \mathbf{e}_1$, então $\alpha\mathbf{x} \notin \mathcal{S}$, pois suas coordenadas são tais que $-1 + 0 + 0 \neq 1$;

· Se $\mathbf{x} = \mathbf{e}_1$ e $\mathbf{y} = \mathbf{e}_2$, então $\mathbf{x} + \mathbf{y} \notin \mathcal{S}$, pois suas coordenadas são tais que $1 + 1 + 0 \neq 1$.

[21] Verifique!

[22] Escolha qualquer uma delas ou estabeleça a sua!

2.5.1 Exemplo geral de subespaço: $\mathcal{S}$ gerado por r vetores

EXERCÍCIO

Em cada um dos itens seguintes, demonstre que $\mathcal{S}$ é um subespaço de $\mathbb{R}^n$, considerando vetores $\mathbf{a}, \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{r-1}$ e $\mathbf{a}_r$ fixados e, para os itens 3, 4 e 5, não nulos.[23]

1. $\mathcal{S} = \{\mathbf{0}\}$; (SUBESPAÇO TRIVIAL)
2. $\mathcal{S} = \mathbb{R}^n$; (SUBESPAÇO TRIVIAL)
3. $\mathcal{S} = \{\mathbf{x} = t\mathbf{a} \,|\, t \in \mathbb{R}\}$; (RETA QUE PASSA PELA ORIGEM NA DIREÇÃO DO VETOR $\mathbf{a}$)
4. $\mathcal{S} = \{\mathbf{x} \,|\, \mathbf{a} \cdot \mathbf{x} = 0\}$; (HIPERPLANO QUE PASSA PELA ORIGEM COM NORMAL $\mathbf{a}$)
5. $\mathcal{S} = \{\mathbf{x} \,|\, \mathbf{a}_1 \cdot \mathbf{x} = 0, \mathbf{a}_2 \cdot \mathbf{x} = 0, \ldots, \mathbf{a}_r \cdot \mathbf{x} = 0\}$;[24] (INTERSEÇÃO DE r HIPERPLANOS QUE PASSAM PELA ORIGEM COM NORMAIS $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$)
6. Para r escalares arbitrários, digamos c_i, $i \in \{1, 2, \ldots, r\}$, a soma

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_r\mathbf{a}_r$$

é chamada de *combinação linear* (CL) de $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{r-1}$ e $\mathbf{a}_r$. Seja $\mathcal{S}$ o conjunto de todas essas CL's, isto é, $\mathcal{S} = \left\{\mathbf{x} \,\middle|\, \mathbf{x} \text{ é CL de } \mathbf{a}_1, \mathbf{a}_2 \ldots, \mathbf{a}_{r-1} \text{ e } \mathbf{a}_r\right\}$. (SUBESPAÇO GERADO POR $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{r-1}$ e $\mathbf{a}_r$)

É importante estabelecer que qualquer subespaço $\mathcal{S}$ de $\mathbb{R}^n$ é gerado por um número finito de vetores, ou seja, é dado como no item 6 do exercício supracitado. A demonstração desse fato encontra-se na seção 6.3.

2.5.2 $\mathcal{S}$ gerado por $r = 3$ vetores em $\mathbb{R}^4$

Se $c_1 = -1$, $\mathbf{a}_1 = (1, -1, 2, 3)$, $c_2 = 2$, $\mathbf{a}_2 = \left(1, 0, -1, \frac{1}{2}\right)$, $c_3 = \frac{3}{4}$ e $\mathbf{a}_3 = \left(\frac{1}{3}, 1, -1, -2\right)$, então

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + c_3\mathbf{a}_3 = \left(\frac{5}{4}, \frac{7}{4}, -\frac{19}{4}, -\frac{7}{2}\right) \in \mathcal{S},$$

[23] As resoluções encontram-se no final desse capítulo!

[24] EXEMPLOS

· Em $\mathbb{R}^3$, seja

$$\mathcal{S} = \left\{(x, y, z) \,\middle|\, \left\{\begin{array}{ccccccc} x & + & y & + & z & = & 0, \\ x & - & 2y & + & z & = & 0, \\ x & + & y & - & 3z & = & 0. \end{array}\right.\right\}.$$

Aqui, $n = r = 3$ e $\mathbf{a}_1 = (1, 1, 1)$, $\mathbf{a}_2 = (1, -2, 1)$ e $\mathbf{a}_3 = (1, 1, -3)$.

· Em $\mathbb{R}^4$, seja

$$\mathcal{S} = \left\{(x, y, z, w) \,\middle|\, \left\{\begin{array}{ccccccccc} x & + & y & & & + & w & = & 0, \\ x & - & y & + & z & & & = & 0. \end{array}\right.\right\}.$$

Aqui, $n = 2r = 4$ e $\mathbf{a}_1 = (1, 1, 0, 1)$ e $\mathbf{a}_2 = (1, -1, 1, 0)$.

ou seja, esse vetor é uma CL de $\mathbf{a}_1, \mathbf{a}_2$ e $\mathbf{a}_3$. Agora, se $c_1 = 1$, $c_2 = -1$ e $c_3 = -3$, então

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + c_3\mathbf{a}_3 = \left(-1, -4, 6, \frac{17}{2}\right) \in \mathcal{S}.$$

Logo, considerando outras possibilidades para c_1, c_2 e c_3, existe uma infinidade de CL's de $\mathbf{a}_1, \mathbf{a}_2$ e $\mathbf{a}_3$. Em particular, além das duas supracitadas, $\mathbf{a}_1 \in \mathcal{S}$.[25] Analogamente, $\mathbf{a}_2$ e $\mathbf{a}_3$ são CL's de $\mathbf{a}_1, \mathbf{a}_2$ e $\mathbf{a}_3$. Do mesmo modo, $\mathbf{0} \in \mathcal{S}$.[26] Portanto, nesse exemplo, os vetores $\mathbf{0}$, $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, $\left(\frac{5}{4}, \frac{7}{4}, -\frac{19}{4}, -\frac{7}{2}\right)$, $\left(-1, -4, 6, \frac{17}{2}\right)$, além das outras possíveis CL's de $\mathbf{a}_1, \mathbf{a}_2$ e $\mathbf{a}_3$,[27] representam os vetores de $\mathcal{S}$.

2.5.3 Bases, LI e LD

Para o subespaço gerado por $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{r-1}$ e $\mathbf{a}_r$, define-se $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r\}$ como *base* de $\mathcal{S}$ quando, além desses r vetores gerarem $\mathcal{S}$, isto é, $\mathcal{S} = \left\{\mathbf{x} \,\middle|\, \mathbf{x} \text{ é CL de } \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{r-1} \text{ e } \mathbf{a}_r\right\}$, eles são *linearmente independentes* (LI), ou seja, a única solução da equação

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_r\mathbf{a}_r = \mathbf{0}$$

é a *trivial*, isto é, dada por

$$x_1 = x_2 = \cdots = x_r = 0.$$

Caso a solução trivial não seja a única solução, dizemos que os r vetores são LD.

EXEMPLOS EM $\mathbb{R}^3$

- Considere que $\mathbf{a}_1 = (1, -1, 1)$, $\mathbf{a}_2 = (-1, 1, 2)$ e $\mathbf{a}_3 = (0, 0, 3)$. Seja

  $$\mathcal{S} = \left\{\mathbf{x} \,\middle|\, \mathbf{x} \text{ é CL de } \mathbf{a}_1, \mathbf{a}_2 \text{ e } \mathbf{a}_3\right\}$$

  o subespaço gerado por $\mathbf{a}_1, \mathbf{a}_2$ e $\mathbf{a}_3$. Assim, para que $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ seja base de $\mathcal{S}$, esses três vetores devem ser LI. Contudo, os vetores $\mathbf{a}_1$, $\mathbf{a}_2$ e $\mathbf{a}_3$ são LD, pois a equação $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = \mathbf{0}$ admite, por exemplo, a solução não trivial $x_1 = 1$, $x_2 = 1$ e $x_3 = -1$.[28] Logo, como $\mathbf{a}_3$ é CL de $\mathbf{a}_1$ e $\mathbf{a}_2$, $\mathbf{a}_3$ pertence ao subespaço $\mathcal{S}$ gerado por $\mathbf{a}_1$ e $\mathbf{a}_2$, ou seja,

  $$\mathbf{a}_3 \in \mathcal{S} = \left\{\mathbf{x} \,\middle|\, \mathbf{x} \text{ é CL de } \mathbf{a}_1 \text{ e } \mathbf{a}_2\right\}.$$

  Por outro lado, $\mathbf{a}_1$ e $\mathbf{a}_2$ são LI, pois

  $$\begin{aligned} x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{0} &\iff (x_1 - x_2, -x_1 + x_2, x_1 + 2x_2) = (0, 0, 0) \\ &\iff \begin{cases} x_1 - x_2 = 0 \\ x_1 + 2x_2 = 0 \end{cases} \\ &\iff x_1 = x_2 = 0. \end{aligned}$$

  Portanto, $\{\mathbf{a}_1, \mathbf{a}_2\}$ é uma base para $\mathcal{S}$.

[25] De fato, considere $c_1 = 1$ e $c_2 = c_3 = 0$.

[26] Basta considerarmos $c_1 = c_2 = c_3 = 0$.

[27] Obtidas ao atribuirmos valores quaisquer aos escalares c_1, c_2 e c_3.

[28] De fato, $\mathbf{a}_3 = \mathbf{a}_1 + \mathbf{a}_2$.

· $\mathbf{e}_1, \mathbf{e}_2$ e $\mathbf{e}_3$ são LI, pois, claramente, $x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 = \mathbf{0}$ admite somente a solução trivial $x_1 = x_2 = x_3 = 0$. Além disso, esses vetores geram $\mathcal{S} = \mathbb{R}^3$. De fato,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \Longleftrightarrow \mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3.$$

Então, $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ é uma base de $\mathcal{S}$.

OBSERVAÇÕES

- Pode ser demonstrado que, $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ é uma base para $\mathbb{R}^n$ se, e somente se, cada $\mathbf{x} \in \mathbb{R}^n$ pode ser escrito de modo único como CL dos vetores dessa base.[29]

- Para n vetores $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ em $\mathbb{R}^n$, demonstra-se a equivalência das seguintes afirmações:

  · Esses vetores formam uma base para $\mathbb{R}^n$.

  · Esses vetores geram $\mathbb{R}^n$.

  · Esses vetores são LI.

  Esse resultado é válido para cada inteiro positivo n. Demonstraremos o caso $n = 3$, na seção dedicada exclusivamente ao espaço vetorial $\mathbb{R}^3$. A demonstração para n arbitrário encontra-se no capítulo 6.[30]

2.5.4 Dimensão

Se $\mathcal{S}$ é um subespaço de $\mathbb{R}^n$, então, pela subseção 6.3.1, demonstra-se que:

- $\mathcal{S}$ tem uma base constituída por r vetores, isto é, $\mathcal{S}$ é gerado por r vetores LI.[31]

- Qualquer base de $\mathcal{S}$ tem o mesmo número de vetores, isto é, para duas bases quaisquer de $\mathcal{S}$, uma com r_1 vetores e a outra com r_2 vetores, temos

$$r_1 = r_2.$$

  Nesse caso, esse número comum de vetores de qualquer uma das bases de $\mathcal{S}$ é chamado de

dimensão de $\mathcal{S}$

  e denotado por

$$\dim \mathcal{S}.$$

EXEMPLOS

[29]Cf. a seção 2.7, exercício 6.

[30]Cf. (R9), p. 152.

[31]Como antecipado na subseção 2.5.1, $\mathcal{S}$ é gerado por um número finito de vetores. Caso esses geradores não sejam LI, podemos eliminar qualquer um deles que seja CL dos outros, restando apenas r vetores que geram $\mathcal{S}$ e são LI.

- No primeiro exemplo da subseção 2.5.3, $\mathcal{S}$ é um subespaço de $\mathbb{R}^3$ com $\dim \mathcal{S} = 2$, ou seja, $\mathcal{S}$ é um plano que passa pela origem do espaço euclidiano tridimensional.

- Para $\mathcal{S} = \mathbb{R}^4$, é fácil ver que $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ é uma base de $\mathcal{S}$ e $\dim \mathcal{S} = 4$.[32]

- Ainda em $\mathbb{R}^4$, considere o subespaço $\mathcal{S}$ gerado por $\mathbf{a}_1 = (1, 1, 0, 0)$, $\mathbf{a}_2 = (0, 1, 1, 0)$ e $\mathbf{a}_3 = (0, 0, 1, 1)$, isto é, cada elemento de $\mathcal{S}$ pode ser escrito da forma $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + c_3\mathbf{a}_3$ com c_1, c_2 e c_3 em $\mathbb{R}$. Note que, $\mathbf{a}_1$, $\mathbf{a}_2$ e $\mathbf{a}_3$ são LI, pois

$$\begin{aligned} x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = \mathbf{0} &\iff (x_1, x_1 + x_2, x_2 + x_3, x_4) = (0, 0, 0, 0) \\ &\iff x_1 = x_2 = x_3 = x_4 = 0. \end{aligned}$$

  Assim, $\mathcal{S}$ é um subespaço de $\mathbb{R}^4$, $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ é uma base de $\mathcal{S}$ e $\dim \mathcal{S} = 3$.

- Se considerarmos $\mathbf{a}_1 = (1, 1, 0, 0)$, $\mathbf{a}_2 = (0, 1, 1, 0)$ e $\mathbf{a}_3 = (1, 2, 1, 0)$, então $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ não é uma base de um subespaço de $\mathbb{R}^4$, pois $\mathbf{a}_1$, $\mathbf{a}_2$ e $\mathbf{a}_3$ são LD. De fato, $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = \mathbf{0}$ admite, além da solução trivial, a solução $x_1 = x_2 = 1$ e $x_3 = -1$, por exemplo. Note que, $\mathbf{a}_3 = \mathbf{a}_1 + \mathbf{a}_2$ pertence ao subespaço $\mathcal{S}$ gerado por $\mathbf{a}_1$ e $\mathbf{a}_2$. Portanto, como $\mathbf{a}_1$ e $\mathbf{a}_2$ são LI,[33] $\{\mathbf{a}_1, \mathbf{a}_2\}$ é base de $\mathcal{S}$ e $\dim \mathcal{S} = 2$.

EXERCÍCIO

Demonstre que $\dim \mathbb{R}^n = n$, pois $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ é uma base de $\mathbb{R}^n$.[34]

2.5.5 Ortogonalidade e ortonormalidade

Caso $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ seja uma lista de vetores não nulos em $\mathbb{R}^n$ e ortogonais entre si, isto é, para quaisquer índices i e j tais que $i \neq j$, $\mathbf{a}_i \cdot \mathbf{a}_j = 0$, dizemos que $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r\}$ é (um conjunto) *ortogonal*.

EXEMPLO

Em $\mathbb{R}^4$,

$$\mathbf{a}_1 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \mathbf{a}_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 1/2 \end{bmatrix} \implies \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\} \text{ é ortogonal.}$$

TEOREMA

Os r vetores de um conjunto ortogonal $\{\mathbf{a}_1, \dots, \mathbf{a}_r\} \subset \mathbb{R}^n$ são LI, *ou seja,*

ortogonalidade $\implies$ independência linear.

---

[32] Analogamente, para $\mathcal{S} = \mathbb{R}^3$, confira o último exemplo da subseção 2.5.3.

[33] Verifique!

[34] Essa base é chamada de *canônica*.

DEMONSTRAÇÃO

Caso $\mathbf{a}_i$ seja um dos vetores da lista $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$, ao multiplicarmos ambos os membros da CL nula

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_r\mathbf{a}_r = \mathbf{0}$$

por $\mathbf{a}_i$, temos

$$c_1\mathbf{a}_1 \cdot \mathbf{a}_i + \cdots + c_{i-1}\mathbf{a}_{i-1} \cdot \mathbf{a}_i + c_i\mathbf{a}_i \cdot \mathbf{a}_i + c_{i+1}\mathbf{a}_{i+1} \cdot \mathbf{a}_i + \cdots + c_r\mathbf{a}_r \cdot \mathbf{a}_i = \mathbf{0} \cdot \mathbf{a}_i. \tag{2.6}$$

Pela hipótese da ortogonalidade entre os r vetores da lista supracitada, temos, para $j \neq i$, $\mathbf{a}_j \cdot \mathbf{a}_i = 0$. Portanto, (2.6) pode ser reescrita como

$$c_i\mathbf{a}_i \cdot \mathbf{a}_i = 0,$$

ou seja, $c_i = 0$.[35] Finalmente, como i é arbitrário,

$$c_1 = c_2 = \cdots = c_r = 0$$

e, assim, os r vetores supracitados são LI.

EXEMPLO

Seja $\mathcal{S}$ o subespaço gerado pelos três vetores dados no primeiro exemplo dessa subseção. Assim, como $\mathbf{a}_1$, $\mathbf{a}_2$ e $\mathbf{a}_3$ são LI, $\dim \mathcal{S} = 3$.

Dizemos que uma base $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r\}$ de um subespaço $\mathcal{S}$ é ortonormal caso seja ortogonal e tenha r vetores unitários, isto é, de comprimento unitário.

Assim, como $\frac{\mathbf{v}}{\|\mathbf{v}\|}$ é unitário para qualquer vetor $\mathbf{v} \neq \mathbf{0}$,[36] temos que, se $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r\}$ é uma base ortogonal, então

$$\{\mathbf{a}_1/\|\mathbf{a}_1\|, \mathbf{a}_2/\|\mathbf{a}_2\|, \ldots, \mathbf{a}_r/\|\mathbf{a}_r\|\}$$

é uma base ortonormal de $\mathcal{S}$.[37]

EXEMPLOS

- O conjunto

$$\left\{ \left(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, 0\right), \left(\frac{1}{\sqrt{6}}, 0, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right), \left(0, \frac{2}{3}, -\frac{2}{3}, \frac{1}{3}\right) \right\}$$

  é uma base ortonormal do subespaço $\mathcal{S}$ dado no segundo exemplo dessa subseção.

- Em $\mathbb{R}^n$, $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ é ortonormal.[38]

---

[35] De fato, $\mathbf{a}_i \neq \mathbf{0}$.

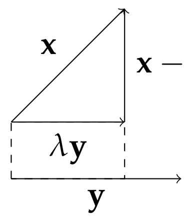
[36] De fato, seja $\alpha = \frac{1}{\|\mathbf{v}\|}$. Então,

$$\begin{aligned} \left\| \frac{\mathbf{v}}{\|\mathbf{v}\|} \right\| &= \|\alpha \mathbf{v}\| \\ &= |\alpha| \|\mathbf{v}\| \\ &= \frac{1}{\|\mathbf{v}\|} \|\mathbf{v}\| \\ &= 1 \text{ u.c.} \end{aligned}$$

[37] Na seção 2.7, veremos que, em geral, essa é a base na qual podemos determinar, mais facilmente, as "coordenadas" de um vetor qualquer de $\mathcal{S}$ e, além disso, existe um método eficiente para calcularmos uma base ortogonal, a partir de qualquer outra base de $\mathcal{S}$.

[38] Confira o exemplo que precede a seção 2.4.

Figura 2.6: Projeção ortogonal de **x** sobre **y**



2.5.6 Subespaços de subespaços

Sejam $\mathcal{S}_1$ e $\mathcal{S}_2$ subespaços de $\mathbb{R}^n$ tais que $\mathcal{S}_1 \subset \mathcal{S}_2$. Nesse caso, dizemos que $\mathcal{S}_1$ *é um subespaço de* $\mathcal{S}_2$.

OBSERVAÇÃO

Via (R8) do capítulo 6,[39] demonstra-se que

$$\dim \mathcal{S}_1 \leq \dim \mathcal{S}_2,$$

onde a igualdade é obtida se, e somente se, $\mathcal{S}_1 = \mathcal{S}_2$.

EXEMPLO

Em $\mathbb{R}^n$, considere que $\mathcal{S}_1$ é uma reta que passa pela origem e $\mathcal{S}_2$ é um plano que contenha essa reta.

2.6 O espaço $\mathbb{R}^3$

Apresentaremos uma nova abordagem para alguns tópicos de GA no espaço $\mathbb{R}^3$ e, simultaneamente, demonstraremos alguns resultados da AL nesse espaço. Assim, considere os vetores $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3$.

2.6.1 Projeção ortogonal de x sobre y

Se $\mathbf{y} \neq \mathbf{0}$ *e* $\lambda = \frac{\mathbf{x}\cdot\mathbf{y}}{\mathbf{y}\cdot\mathbf{y}}$, *então o vetor* $\mathbf{x} - \lambda\mathbf{y}$ *é ortogonal ao vetor* $\mathbf{y}$.

DEMONSTRAÇÃO

Pela linearidade do produto interno e da definição de λ, temos

$$\begin{aligned}(\mathbf{x} - \lambda\mathbf{y}) \cdot \mathbf{y} &= \mathbf{x} \cdot \mathbf{y} - \lambda(\mathbf{y} \cdot \mathbf{y}) \\ &= 0.\end{aligned}$$

$\lambda\mathbf{y}$ é a *projeção ortogonal* de $\mathbf{x}$ sobre $\mathbf{y}$, conforme ilustrada na figura 2.6.

EXEMPLO

$\mathbf{x} = (1,1,0)$, $\mathbf{y} = (0,1,1) \Longrightarrow \lambda = \frac{1}{2}$, $\mathbf{x} - \lambda\mathbf{y} = \left(1, \frac{1}{2}, -\frac{1}{2}\right)$.

[39] Cf. p. 152.

2.6.2 Produto vetorial de x e y

É definido por

$$\mathbf{x} \times \mathbf{y} := (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1)\,.^{40}$$

EXEMPLO

$\mathbf{x} = (1,1,0)$, $\mathbf{y} = (0,1,1) \Longrightarrow \mathbf{x} \times \mathbf{y} = (1,-1,1)$.

OBSERVAÇÃO

O produto vetorial é *antissimétrico* e *linear*, ou seja,

$$\mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x}$$

e

$$(\alpha\mathbf{x} + \beta\mathbf{z}) \times \mathbf{y} = \alpha(\mathbf{x} \times \mathbf{y}) + \beta(\mathbf{z} \times \mathbf{y}),$$

onde α e β são escalares reais quaisquer.

EXERCÍCIOS

1. Demonstre a antissimetria e a linearidade do produto vetorial.

2. Verifique que $\mathbf{i} \times \mathbf{j} = \mathbf{k}$, $\mathbf{k} \times \mathbf{i} = \mathbf{j}$ e $\mathbf{j} \times \mathbf{k} = \mathbf{i}$.

2.6.3 Relação entre projeção ortogonal e produto vetorial

Se $\lambda\mathbf{y}$ é a projeção ortogonal de $\mathbf{x}$ sobre $\mathbf{y}$, então

$$\|\mathbf{x} \times \mathbf{y}\| = \|\mathbf{y}\|\|\mathbf{x} - \lambda\mathbf{y}\|. \tag{2.7}$$

EXEMPLO

Para $\mathbf{x}$ e $\mathbf{y}$ dados nos exemplos das subseções 2.6.1 e 2.6.2, ambos os membros de (2.7) são iguais a $\sqrt{3}$.

DEMONSTRAÇÃO DE (2.7)

[40] Uma mnemônica para essa expressão é dada pelo "determinante"

$$\mathbf{x} \times \mathbf{y} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}.$$

Basta observarmos que

$$\begin{aligned}
\|\mathbf{x}\times\mathbf{y}\|^2 &= (x_2y_3 - x_3y_2)^2 + (x_3y_1 - x_1y_3)^2 + (x_1y_2 - x_2y_1)^2 \\
&= \left(x_1^2 + x_2^2 + x_3^2\right)\left(y_1^2 + y_2^2 + y_3^2\right) - (x_1y_1 + x_2y_2 + x_3y_3)^2 \\
&= \|\mathbf{x}\|^2\|\mathbf{y}\|^2 - (\mathbf{x}\cdot\mathbf{y})^2 \\
&= \|\mathbf{y}\|^2\left(\|\mathbf{x}\|^2 - \frac{(\mathbf{x}\cdot\mathbf{y})^2}{\|\mathbf{y}\|^2}\right) \\
&= \|\mathbf{y}\|^2(\mathbf{x} - \lambda\mathbf{y})\cdot\mathbf{x} \qquad \left(\text{POIS } \lambda = \frac{\mathbf{x}\cdot\mathbf{y}}{\mathbf{y}\cdot\mathbf{y}}\right) \\
&= \|\mathbf{y}\|^2(\mathbf{x} - \lambda\mathbf{y})\cdot(\mathbf{x} - \lambda\mathbf{y}) \quad (\text{POIS } (\mathbf{x} - \lambda\mathbf{y})\cdot\lambda\mathbf{y} = 0) \\
&= \|\mathbf{y}\|^2\|\mathbf{x} - \lambda\mathbf{y}\|^2.
\end{aligned}$$

2.6.4 Dependência linear e produto vetorial

$$\boxed{\mathbf{x} \text{ e } \mathbf{y} \text{ são LD} \iff \mathbf{x}\times\mathbf{y} = \mathbf{0}.}$$

EXEMPLO

Para os exemplos das subseções 2.6.1, 2.6.2 e 2.6.3, $\mathbf{x}$ e $\mathbf{y}$ são LI e tem produto vetorial não nulo.

DEMONSTRAÇÃO

Por um lado, suponha que os vetores sejam LD, ou seja, $\mathbf{y} = \alpha\mathbf{x}$, para algum $\alpha \in \mathbb{R}$. Portanto,

$$\begin{aligned}
\mathbf{x}\times\mathbf{y} &= \mathbf{x}\times\alpha\mathbf{x} \\
&= \alpha(\mathbf{x}\times\mathbf{x}) \quad (\text{PELA LINEARIDADE DO PRODUTO VETORIAL}) \\
&= \alpha\mathbf{0} \qquad\quad (\text{PELA ANTISSIMETRIA DO PRODUTO VETORIAL}) \\
&= \mathbf{0}.
\end{aligned}$$

Por outro, suponha que o produto vetorial seja nulo e um dos vetores, digamos, $\mathbf{y}$, seja não nulo.[41] Como $\|\mathbf{x} - \lambda\mathbf{y}\| = 0$,[42] temos $\mathbf{x} = \lambda\mathbf{y}$, ou seja, $\mathbf{x}$ e $\mathbf{y}$ são LD.

2.6.5 Produto misto de x, y e z

É o escalar definido por

$$(\mathbf{x},\mathbf{y},\mathbf{z}) := \mathbf{x}\cdot(\mathbf{y}\times\mathbf{z}).$$

Note que, $(\mathbf{x},\mathbf{y},\mathbf{z}) = (\mathbf{y}\times\mathbf{z})\cdot\mathbf{x}$, pela comutatividade do produto interno.

EXERCÍCIOS

1. Verifique que o produto misto pode ser obtido pelo determinante

$$(\mathbf{x},\mathbf{y},\mathbf{z}) = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}.$$

[41] O caso $\mathbf{x} = \mathbf{y} = \mathbf{0}$ é trivial.

[42] Por (2.7).

2. Verifique que $(\mathbf{x}, \mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x}, \mathbf{y}) = 0$, isto é,

$$(\mathbf{x} \times \mathbf{y}) \cdot \mathbf{x} = (\mathbf{x} \times \mathbf{y}) \cdot \mathbf{y} = 0,$$

ou seja, $\mathbf{x} \times \mathbf{y}$ é ortogonal aos vetores $\mathbf{x}$ e $\mathbf{y}$.

3. Utilize o exercício 1 dessa subseção, para demonstrar que

$$\begin{aligned}(\mathbf{x}, \mathbf{y}, \mathbf{z}) &= (\mathbf{z}, \mathbf{x}, \mathbf{y}) \\ &= (\mathbf{y}, \mathbf{z}, \mathbf{x}) \\ &= -(\mathbf{y}, \mathbf{x}, \mathbf{z}) \\ &= -(\mathbf{z}, \mathbf{y}, \mathbf{x}),\end{aligned}$$

isto é,

$$\begin{aligned}\mathbf{x} \cdot (\mathbf{y} \times \mathbf{z}) &= \mathbf{z} \cdot (\mathbf{x} \times \mathbf{y}) \\ &= \mathbf{y} \cdot (\mathbf{z} \times \mathbf{x}) \\ &= -[\mathbf{y} \cdot (\mathbf{x} \times \mathbf{z})] \\ &= -[\mathbf{z} \cdot (\mathbf{y} \times \mathbf{x})].\end{aligned}$$

2.6.6 Bases de $\mathbb{R}^3$ via produto misto

Dados $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{z}$ em $\mathbb{R}^3$, vamos estabelecer um critério para determinarmos, via produto misto, se $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ é uma base de $\mathbb{R}^3$.

TEOREMA 1

Caso $\mathbf{x}$ e $\mathbf{y}$ sejam LI *e ortogonais ao vetor $\mathbf{z}$, $\mathbf{z}$ é um múltiplo escalar de $\mathbf{x} \times \mathbf{y}$.*

EXEMPLO

$\mathbf{z} = (1/2, 1, 1/2)$ é ortogonal aos vetores $\mathbf{x} = (1, -1, 1)$ e $\mathbf{y} = (-1, 0, 1)$,[43] que são LI.[44] Além disso, $\mathbf{x} \times \mathbf{y} = -2\mathbf{z}$.

DEMONSTRAÇÃO DO TEO. 1

Pela hipótese da ortogonalidade, temos

$$\begin{cases} x_1 z_1 + x_2 z_2 + x_3 z_3 = 0, \\ y_1 z_1 + y_2 z_2 + y_3 z_3 = 0. \end{cases}$$

Para eliminarmos z_1 desse sistema, basta calcularmos o produto da primeira equação por y_1, da segunda por $-x_1$ e a soma das equações resultantes desses produtos. Analogamente, podemos eliminar z_2 e z_3. Portanto,

$$\begin{cases} z_2\,(x_2 y_1 - x_1 y_2) + z_3\,(x_3 y_1 - x_1 y_3) = 0, \\ z_1\,(x_1 y_2 - x_2 y_1) + z_3\,(x_3 y_2 - x_2 y_3) = 0, \\ z_1\,(x_1 y_3 - x_3 y_1) + z_2\,(x_2 y_3 - x_3 y_2) = 0. \end{cases} \tag{2.8}$$

Por outro lado, pela hipótese da independência linear, temos

$$\mathbf{x} \times \mathbf{y} \neq \mathbf{0}.^{45}$$

[43] De fato, $\mathbf{x} \cdot \mathbf{z} = 0$ e $\mathbf{y} \cdot \mathbf{z} = 0$.
[44] De fato, não podemos escrever um dos vetores, $\mathbf{x}$ ou $\mathbf{y}$, como múltiplo escalar do outro.
[45] Caso contrário, pela equação (2.7), teríamos $\mathbf{x}$ e $\mathbf{y}$ LD!

Assim, alguma componente de $\mathbf{x} \times \mathbf{y}$ é não nula. Sem perda de generalidade, suponha que $x_1y_2 - x_2y_1 \neq 0$. Logo, para

$$\gamma = \frac{z_3}{x_1y_2 - x_2y_1},$$

temos

$$z_3 = \gamma\left(x_1y_2 - x_2y_1\right).$$

Agora, se substituirmos z_3 nas duas primeiras equações de (2.8), então

$$z_2 = \gamma\left(x_3y_1 - x_1y_3\right) \;\text{ e }\; z_1 = \gamma\left(x_2y_3 - x_3y_2\right).$$

Portanto,

$$\mathbf{z} = \gamma(\mathbf{x} \times \mathbf{y}).$$

TEOREMA 2

Caso $\mathbf{x}$ e $\mathbf{y}$ sejam LI, $\{\mathbf{x}, \mathbf{y}, \mathbf{x} \times \mathbf{y}\}$ é uma base de $\mathbb{R}^3$.

DEMONSTRAÇÃO DO TEO. 2

Note que:

- Pela hipótese da independência linear, temos $\mathbf{x} \neq \mathbf{0}$ e $\mathbf{y} \neq \mathbf{0}$;
- Caso $\lambda\mathbf{y}$ seja a projeção ortogonal de $\mathbf{x}$ em $\mathbf{y}$, $\mathbf{x}' = \mathbf{x} - \lambda\mathbf{y}$ é ortogonal a $\mathbf{y}$;[46]
- $\mathbf{x}' \neq \mathbf{0}$, pois $\mathbf{x}$ e $\mathbf{y}$ são LI;
- Para $\mathbf{z} \in \mathbb{R}^3$ arbitrário, podemos considerar

$$\alpha' = \frac{\mathbf{z} \cdot \mathbf{x}'}{\mathbf{x}' \cdot \mathbf{x}'} \;\text{ e }\; \beta' = \frac{\mathbf{z} \cdot \mathbf{y}}{\mathbf{y} \cdot \mathbf{y}}.$$

Demonstraremos que $\mathbf{z}$ pode ser escrito de modo único como uma CL de $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{x} \times \mathbf{y}$.

De fato, da linearidade do produto interno e pela ortogonalidade de $\mathbf{x}'$ e $\mathbf{y}$, verifica-se que $\mathbf{z} - (\alpha'\mathbf{x}' + \beta'\mathbf{y})$ é ortogonal aos vetores $\mathbf{x}'$ e $\mathbf{y}$.[47] Além disso, como $\mathbf{x}'$ e $\mathbf{y}$ são não nulos e ortogonais entre si, também são LI.[48] Portanto, pelo teorema 1 supracitado, existe algum escalar γ' tal que

$$\mathbf{z} - \left(\alpha'\mathbf{x}' + \beta'\mathbf{y}\right) = \gamma'\left(\mathbf{x}' \times \mathbf{y}\right),$$

ou seja,

$$\mathbf{z} = \alpha'\mathbf{x}' + \beta'\mathbf{y} + \gamma'\left(\mathbf{x}' \times \mathbf{y}\right). \tag{2.9}$$

Se substituirmos $\mathbf{x}' = \mathbf{x} - \lambda\mathbf{y}$ em (2.9), utilizarmos a igualdade $\mathbf{y} \times \mathbf{y} = \mathbf{0}$ e reagruparmos os coeficientes, podemos obter escalares α, β e γ tais que

$$\mathbf{z} = \alpha\mathbf{x} + \beta\mathbf{y} + \gamma(\mathbf{x} \times \mathbf{y}).^{49} \tag{2.10}$$

Para concluirmos a demonstração, basta verificarmos que a tripla ordenada (α, β, γ) é única, em relação a CL (2.10). Suponha, por contradição, que não seja única. Assim, existe outra tripla ordenada (a, b, c) tal que

$$\mathbf{z} = a\mathbf{x} + b\mathbf{y} + c(\mathbf{x} \times \mathbf{y}). \tag{2.11}$$

[46] Cf. a subseção 2.6.1.

[47] Verifique!

[48] Cf. a subseção 2.5.5.

[49] Verifique!

Então, da diferença (2.10)−(2.11), obtemos escalares a', b' e c', não todos nulos, tais que

$$\mathbf{0} = a'\mathbf{x} + b'\mathbf{y} + c'(\mathbf{x} \times \mathbf{y}).$$ [50]

Ao multiplicarmos ambos os lados dessa equação por $\mathbf{x} \times \mathbf{y}$, obtemos

$$0 = c'\|\mathbf{x} \times \mathbf{y}\|^2.$$

Logo, como $\|\mathbf{x} \times \mathbf{y}\| \neq 0$, pois $\mathbf{x}$ e $\mathbf{y}$ são LI,[51] temos $c' = 0$. Portanto,

$$\mathbf{0} = a'\mathbf{x} + b'\mathbf{y},$$

com $a' \neq 0$ ou $b' \neq 0$, ou seja, $\mathbf{x}$ e $\mathbf{y}$ são LD, contradizendo a hipótese do teorema 2.

Como observamos na subseção 2.5.3, o próximo resultado também é válido para o espaço $\mathbb{R}^n$, para qualquer inteiro $n \geq 2$, conforme está demonstrado no capítulo 6.

TEOREMA 3

As afirmações seguintes são equivalentes:

1. $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ *é uma base de* $\mathbb{R}^3$.
2. $\mathbf{x}$, $\mathbf{y}$ *e* $\mathbf{z}$ *geram* $\mathbb{R}^3$.
3. $\mathbf{x}$, $\mathbf{y}$ *e* $\mathbf{z}$ *são* LI.

DEMONSTRAÇÃO DO TEO. 3

$1 \Longrightarrow 2$

Essa implicação é uma consequência direta da definição de base.

$2 \Longrightarrow 3$

Se $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{z}$ são geradores do $\mathbb{R}^3$ e a afirmação 3 não é verdadeira, isto é, esses três geradores são LD, então um deles é uma CL dos outros dois. Sem perda de generalidade, suponha que $\mathbf{x}$ e $\mathbf{y}$ gerem $\mathbb{R}^3$. Analisaremos dois casos mutuamente excludentes:

- Caso $\mathbf{x} \times \mathbf{y} = \mathbf{0}$, temos $\mathbf{x}$ e $\mathbf{y}$ LD.[52] Então, como qualquer um deles gera $\mathbb{R}^3$, podemos escrever $\mathbf{i}$ como múltiplo escalar de $\mathbf{j}$, que é uma afirmação falsa.
- Caso $\mathbf{x} \times \mathbf{y} \neq \mathbf{0}$, como esse produto vetorial (não nulo) é ortogonal aos vetores $\mathbf{x}$ e $\mathbf{y}$,[53] ele não pode ser uma CL desses vetores, contradizendo a suposição de que $\mathbf{x}$ e $\mathbf{y}$ geram $\mathbb{R}^3$.

$3 \Longrightarrow 1$

Sejam $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{z}$ LI. Assim, dois deles, digamos, os dois primeiros, também são LI. Logo, pelo teorema 2 supracitado, $\{\mathbf{x}, \mathbf{y}, \mathbf{x} \times \mathbf{y}\}$ é uma base de $\mathbb{R}^3$. Portanto, existem escalares α, β e γ tais que

$$\mathbf{z} = \alpha\mathbf{x} + \beta\mathbf{y} + \gamma(\mathbf{x} \times \mathbf{y}).$$

Note que, pela suposição da independência linear de $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{z}$, γ é não nulo e, assim,

$$\mathbf{x} \times \mathbf{y} = \frac{1}{\gamma}(\mathbf{z} - \alpha\mathbf{x} - \beta\mathbf{y}). \tag{2.12}$$

Analogamente, para $\mathbf{v} \in \mathbb{R}^3$ arbitrário, existem escalares α', β' e γ' tais que

$$\mathbf{v} = \alpha'\mathbf{x} + \beta'\mathbf{y} + \gamma'(\mathbf{x} \times \mathbf{y}). \tag{2.13}$$

[50] $a' = \alpha - a$, $b' = \beta - b$ e $c' = \gamma - c$.

[51] Cf. a subseção 2.6.4.

[52] Idem.

[53] Cf. exercício 2 da subseção 2.6.5.

Ao substituirmos (2.12) em (2.13), obtemos

$$\begin{aligned}\mathbf{v} &= \alpha'\mathbf{x} + \beta'\mathbf{y} + \frac{\gamma'}{\gamma}(\mathbf{z} - \alpha\mathbf{x} - \beta\mathbf{y}) \\ &= \left(\alpha' - \frac{\gamma'}{\gamma}\alpha\right)\mathbf{x} + \left(\beta' - \frac{\gamma'}{\gamma}\beta\right)\mathbf{y} + \frac{\gamma'}{\gamma}\mathbf{z}.\end{aligned}$$

Portanto, $\mathbf{v}$ é uma CL de $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{z}$, que supomos serem LI. Assim, essa CL é única.

Regra de Cramer - I

Demonstra-se que um produto misto arbitrário é não nulo se, e somente se, os três vetores desse produto formam uma base de $\mathbb{R}^3$. Nesse caso, podemos obter as coordenadas de um vetor arbitrário nessa base.

TEOREMA 4

As afirmações seguintes são equivalentes:

1. $(\mathbf{x}, \mathbf{y}, \mathbf{z}) \neq 0$.
2. $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ *é uma base de* $\mathbb{R}^3$.

Além disso, se

$$\mathbf{v} = \alpha\mathbf{x} + \beta\mathbf{y} + \gamma\mathbf{z}, \tag{2.14}$$

então

$$\alpha = \frac{(\mathbf{v}, \mathbf{y}, \mathbf{z})}{(\mathbf{x}, \mathbf{y}, \mathbf{z})}, \quad \beta = \frac{(\mathbf{x}, \mathbf{v}, \mathbf{z})}{(\mathbf{x}, \mathbf{y}, \mathbf{z})} \quad e \quad \gamma = \frac{(\mathbf{x}, \mathbf{y}, \mathbf{v})}{(\mathbf{x}, \mathbf{y}, \mathbf{z})}. \tag{2.15}$$

EXERCÍCIO

Verifique que $\mathbf{x} = (1, 1, 0)$, $\mathbf{y} = (0, 1, 1)$ e $\mathbf{z} = (0, 1, 0)$ formam uma base para o $\mathbb{R}^3$ e calcule as coordenadas de $\mathbf{v} = (1, 1, 1)$ nessa base.

DEMONSTRAÇÃO DO TEO. 4

$1 \Longrightarrow 2$

Suponha que a condição 1 seja válida. Assim,

$$\mathbf{z} \cdot (\mathbf{x} \times \mathbf{y}) = \mathbf{x} \cdot (\mathbf{y} \times \mathbf{z}) \neq 0. \tag{2.16}$$

Então, $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{x} \times \mathbf{y}$ são não nulos e, além disso, $\mathbf{x}$ e $\mathbf{y}$ são LI. Logo, $\{\mathbf{x}, \mathbf{y}, \mathbf{x} \times \mathbf{y}\}$ é base de $\mathbb{R}^3$, pelo teorema 2 supracitado, e, assim, existem escalares α, β e γ tais que

$$\mathbf{z} = \alpha\mathbf{x} + \beta\mathbf{y} + \gamma(\mathbf{x} \times \mathbf{y}). \tag{2.17}$$

Ao calcularmos o produto interno entre cada membro de (2.17) e o vetor $\mathbf{x} \times \mathbf{y}$, obtemos

$$\mathbf{z} \cdot (\mathbf{x} \times \mathbf{y}) = \gamma\|\mathbf{x} \times \mathbf{y}\|^2.$$

Como, por (2.16), $\gamma \neq 0$,

$$\mathbf{x} \times \mathbf{y} = -\frac{\alpha}{\gamma}\mathbf{x} - \frac{\beta}{\gamma}\mathbf{y} + \frac{1}{\gamma}\mathbf{z},$$

pela equação (2.17). Portanto, como cada vetor $\mathbf{v} \in \mathbb{R}^3$ pode ser escrito como CL de $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{x} \times \mathbf{y}$, $\mathbf{v}$ também pode ser escrito como CL de $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{z}$. Assim, para que a condição 2 seja válida, resta provarmos a unicidade dessa CL. De fato, considere escalares α, β e γ para os quais a equação (2.14) seja válida. Multiplicando ambos os membros dessa equação pelos escalares $\mathbf{y} \times \mathbf{z}$, $\mathbf{z} \times \mathbf{x}$ e $\mathbf{x} \times \mathbf{y}$, obtemos

$$\begin{aligned} \mathbf{v} \cdot (\mathbf{y} \times \mathbf{z}) &= \alpha[\mathbf{x} \cdot (\mathbf{y} \times \mathbf{z})], \\ \mathbf{v} \cdot (\mathbf{z} \times \mathbf{x}) &= \beta[\mathbf{y} \cdot (\mathbf{z} \times \mathbf{x})] \text{ e} \\ \mathbf{v} \cdot (\mathbf{x} \times \mathbf{y}) &= \gamma[\mathbf{z} \cdot (\mathbf{x} \times \mathbf{y})], \end{aligned}$$

respectivamente, ou seja, os escalares α, β e γ dados em (2.15) e, concomitantemente, a unicidade da CL (2.14).

$\boxed{2 \Longrightarrow 1}$

Considere a validade da condição 2, ou seja, suponha que $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ seja uma base de $\mathbb{R}^3$. Assim, pelo teorema 3 supracitado, os vetores $\mathbf{x}$ e $\mathbf{y}$ são LI. Portanto, $\mathbf{x} \times \mathbf{y} \neq \mathbf{0}$.[54] Além disso, pelo teorema 2 supracitado, $\{\mathbf{x}, \mathbf{y}, \mathbf{x} \times \mathbf{y}\}$ é uma base de $\mathbb{R}^3$. Logo, existem escalares α, β e γ tais que

$$\mathbf{z} = \alpha\mathbf{x} + \beta\mathbf{y} + \gamma(\mathbf{x} \times \mathbf{y}), \tag{2.18}$$

com $\gamma \neq 0$.[55] Assim, a afirmação 1 é válida, pois,

$$\begin{aligned} \mathbf{x} \cdot (\mathbf{y} \times \mathbf{z}) &= \mathbf{z} \cdot (\mathbf{x} \times \mathbf{y}) \\ &= \gamma\|\mathbf{x} \times \mathbf{y}\|^2 \\ &\neq 0, \end{aligned}$$

onde, na segunda igualdade, de cima para baixo, calculamos o produto interno entre cada membro da equação (2.18) e o vetor $\mathbf{x} \times \mathbf{y}$.

Regra de Cramer - II

Se $\mathbf{x} = (a_{11}, a_{21}, a_{31})$, $\mathbf{y} = (a_{12}, a_{22}, a_{32})$, $\mathbf{z} = (a_{13}, a_{23}, a_{33})$, $\alpha = x$, $\beta = y$ e $\gamma = z$, verifica-se que:

1. O teorema 4 supracitado é equivalente a *regra de Cramer*, estudada no ensino médio;
2. Via independência linear, temos a existência e a unicidade de solução para o "sistema linear" (2.14), pelos teoremas 3 e 4 supracitados.

[54] Cf. a subseção 2.6.4.

[55] Como supusemos que $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ é uma base de $\mathbb{R}^3$, $\mathbf{z}$ não é CL de $\mathbf{x}$ e $\mathbf{y}$.

2.7 Exercícios

1. Determine $\mathbf{x}, \mathbf{y} \in \mathbb{R}^4$ tais que as coordenadas de $\mathbf{x}$ são todas iguais, a última coordenada de $\mathbf{y}$ é igual a 1 e $\mathbf{x} + \mathbf{y} = (2, 3, 4, 5)$.[56]

2. Se $\mathbf{x} = (1, 2, 3)$, $\mathbf{y} = (4, 5, 0)$ e $\mathbf{z} = (6, 0, 0)$, determine escalares x, y e z tais que

$$x\mathbf{x} + y\mathbf{y} + z\mathbf{z} = (41/2, 11, 1/2).$$[57]

3. Em $\mathbb{R}^4$, obtenha a equação geral do hiperplano que passa pelos pontos $\mathcal{P}_1 = (1, -1, 1, 0)$, $\mathcal{P}_2 = (0, -1, 2, 0)$, $\mathcal{P}_3 = (1, 0, -2, 2)$ e $\mathcal{P}_4 = (1, 0, 0, 0)$. Além disso, determine os pontos equidistantes uma u.c. de $\mathcal{P}_1$, pertencentes à reta que passa por $\mathcal{P}_1$ e é perpendicular ao hiperplano supracitado.

RESOLUÇÃO PARCIAL

$ax + by + cz + dw = e$ é a equação geral do hiperplano a ser determinada. Portanto, como $\mathbf{a} = (a, b, c, d)$ e $\mathbf{x}_i$ são ortogonais, $i = 1, 2, 3$, onde $\mathbf{x}_1 = \mathcal{P}_1 - \mathcal{P}_4$, $\mathbf{x}_2 = \mathcal{P}_2 - \mathcal{P}_4$ e $\mathbf{x}_3 = \mathcal{P}_3 - \mathcal{P}_4$, temos $a = b = c = d$.[58] Então, $x + y + z + w = 1$ é a equação do hiperplano, obtida ao substituirmos as coordenadas de $\mathcal{P}_4$ em $a(x + y + z + w) = e$. Além disso, caso $\mathcal{P}_1$ seja o ponto final do vetor $\mathbf{x}_0$, podemos considerar a equação vetorial da reta supracitada, ou seja, $\mathbf{x} = \mathbf{x}_0 + t\mathbf{a}$, $t \in \mathbb{R}$, para determinarmos os pontos $\mathbf{x}$, dessa reta, tais que

$$\|\mathbf{x} - \mathbf{x}_0\| = \|t\mathbf{a}\| = 1 \text{ u.c.}$$

Portanto, $\left(\frac{1}{2}, -\frac{3}{2}, \frac{1}{2}, -\frac{1}{2}\right)$ e $\left(\frac{3}{2}, -\frac{1}{2}, \frac{3}{2}, \frac{1}{2}\right)$ são os pontos supracitados. Note que eles são equidistantes uma u.c. de $\mathcal{P}_1$.

4. Em cada um dos itens seguintes, demonstre que $\mathcal{S}$ é um subespaço de $\mathbb{R}^n$. Além disso, apresente uma base de $\mathcal{S}$ e sua dimensão.

   (a) $n = 3$ e $\mathcal{S} = \{\mathbf{x} = (x_1, x_2, x_3) \mid ax_1 + bx_2 + cx_3 = 0\}$, o plano que passa pela origem com vetor normal $\mathbf{a} = (a, b, c) \neq \mathbf{0}$;

   (b) $n = 4$ e:

   (b.1) $\mathcal{S} = \left\{\mathbf{x} = (x_1, x_2, x_3, x_4) \,\middle|\, x_2 = 2x_1, x_3 = 3x_1, x_4 = 4x_1\right\}$, uma reta que passa pela origem;

   (b.2) $\mathcal{S} = \left\{\mathbf{x} = (x_1, x_2, x_3, x_4) \,\middle|\, x_3 = x_1 + 2x_2, x_4 = x_1 - x_2\right\}$, um plano que passa pela origem;

   (b.3) $\mathcal{S} = \left\{\mathbf{x} = (x_1, x_2, x_3, x_4) \,\middle|\, x_4 = x_1 + 2x_2 + 3x_3\right\}$.

RESOLUÇÃO

(a) Pelo exercício 6 da subseção 2.5.1, caso seja gerado por r vetores, $\mathcal{S}$ é um subespaço de $\mathbb{R}^3$. Verificaremos que $r = 2$ vetores geram $\mathcal{S}$ e são LI. Portanto, esses vetores formam uma base de $\mathcal{S}$ e $\dim \mathcal{S} = 2$.

[56] RESPOSTA $\mathbf{x} = (4, 4, 4, 4)$ e $\mathbf{y} = (-2, -1, 0, 1)$.

[57] RESPOSTA $x = \frac{1}{6}$, $y = \frac{32}{15}$ e $z = \frac{177}{90}$.

[58] Considere o sistema $\mathbf{x}_i \cdot \mathbf{a} = 0$, $i = 1, 2, 3$.

Primeiramente, note que, $(a, b, c) \neq \mathbf{0}$ tem alguma coordenada não nula. Suponha, sem perda de generalidade, que $c \neq 0$. Então, por um lado, como

$$\begin{aligned}\mathcal{S} \ni \mathbf{x} &= \left(x_1, x_2, -\frac{a}{c}x_1 - \frac{b}{c}x_2\right)\\ &= x_1\left(1, 0, -\frac{a}{c}\right) + x_2\left(0, 1, -\frac{b}{c}\right),\end{aligned}$$

temos que $\mathbf{a}_1 = \left(1, 0, -\frac{a}{c}\right)$ e $\mathbf{a}_2 = \left(0, 1, -\frac{b}{c}\right)$ geram $\mathcal{S}$, pois todo $\mathbf{x} \in \mathcal{S}$ é CL de $\mathbf{a}_1$ e $\mathbf{a}_2$. Por outro, $\mathbf{a}_1$ e $\mathbf{a}_2$ são LI, pois $\mathbf{x} = \mathbf{0}$ apenas quando $x_1 = x_2 = 0$.
(b1) Caso seja gerado por r vetores, $\mathcal{S}$ é subespaço de $\mathbb{R}^4$. Verificaremos que $r = 1$ vetor gera $\mathcal{S}$ e, por ser não nulo, é LI. Portanto, esse vetor é o único elemento de uma base de $\mathcal{S}$ e $\dim \mathcal{S} = 1$. De fato, como

$$\begin{aligned}\mathcal{S} \ni \mathbf{x} &= (x_1, 2x_1, 3x_1, 4x_1)\\ &= x_1(1, 2, 3, 4),\end{aligned}$$

temos que $\mathbf{a}_1 = (1, 2, 3, 4) \neq \mathbf{0}$ gera $\mathcal{S}$.
(b.2) Caso seja gerado por r vetores, $\mathcal{S}$ é subespaço de $\mathbb{R}^4$. Verificaremos que $r = 2$ vetores geram $\mathcal{S}$ e são LI. Portanto, esses vetores formam uma base de $\mathcal{S}$ e $\dim \mathcal{S} = 2$. De fato, como

$$\begin{aligned}\mathcal{S} \ni \mathbf{x} &= (x_1, x_2, x_1 + 2x_2, x_1 - x_2)\\ &= x_1(1, 0, 1, 1) + x_2(0, 1, 2, -1),\end{aligned}$$

temos que $\mathbf{a}_1 = (1, 0, 1, 1)$ e $\mathbf{a}_2 = (0, 1, 2, -1)$ geram $\mathcal{S}$ e, como nenhum deles é múltiplo escalar do outro, são LI.
(b.3) $\mathcal{S}$ é subespaço de $\mathbb{R}^4$, caso seja gerado por r vetores. Verificaremos que $r = 3$ vetores geram $\mathcal{S}$ e são LI. Portanto, esses vetores formam uma base de $\mathcal{S}$ e $\dim \mathcal{S} = 3$. De fato, como

$$\begin{aligned}\mathcal{S} \ni \mathbf{x} &= (x_1, x_2, x_3, x_1 + 2x_2 + 3x_3)\\ &= x_1(1, 0, 0, 1) + x_2(0, 1, 0, 2) + x_3(0, 0, 1, 3),\end{aligned}$$

temos que $\mathbf{a}_1 = (1, 0, 0, 1)$, $\mathbf{a}_2 = (0, 1, 0, 2)$ e $\mathbf{a}_3 = (0, 0, 1, 3)$ geram $\mathcal{S}$ e, igualando-se $\mathbf{x}$ ao vetor nulo, obtemos $x_i = 0$, $i = 1, 2, 3$, isto é, esses três vetores são LI.

5. Justifique porque $\mathcal{S}$ não é um subespaço de $\mathbb{R}^2$, para:

   (a) $\mathcal{S} = \left\{\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \,\middle|\, x_2 = x_1^2\right\}$;

   (b) $\mathcal{S} = \left\{\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \,\middle|\, x_2 = 1\right\}$.

RESOLUÇÃO [59]

(a) Note que, por exemplo, $\mathbf{x} = (1, 1) \in \mathcal{S}$, mas $2\mathbf{x} \notin \mathcal{S}$. De fato, a segunda coordenada de $2\mathbf{x}$ não é o quadrado de sua primeira coordenada.
(b) Note que, por exemplo, o vetor nulo não pertence a $\mathcal{S}$.

6. Demonstre as afirmações seguintes:

   (a) Nenhuma base de um subespaço de $\mathbb{R}^n$ pode conter o vetor nulo, pois é LD qualquer conjunto finito de vetores do $\mathbb{R}^n$ que contenha $\mathbf{0}$.[60]

[59] Existem inúmeras resoluções para esse exercício. Tente obter a sua!

[60] RESOLUÇÃO

(b) Nenhum vetor pode ter "coordenadas" distintas numa mesma base, isto é, se $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r\}$ é uma base de um subespaço $\mathcal{S}$ do $\mathbb{R}^n$, $c_1, c_1', c_2, c_2', \ldots, c_r, c_r'$ são escalares e $\mathbf{x} \in \mathcal{S}$ é tal que

$$\begin{aligned} \mathbf{x} &= c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_r\mathbf{a}_r \\ &= c_1'\mathbf{a}_1 + c_2'\mathbf{a}_2 + \cdots + c_r'\mathbf{a}_r, \end{aligned}$$

então $c_1 = c_1'$, $c_2 = c_2' \ldots$, $c_r = c_r'$.[61]

$$\boxed{c_1, \ldots, c_r \textit{ são chamados de coordenadas de } \mathbf{x} \textit{ na base } \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r\}.}$$

(c) Sendo $\mathcal{S}$ um subespaço de $\mathbb{R}^n$ e $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r\}$ uma base ortonormal de $\mathcal{S}$, as coordenadas de $\mathbf{x} = c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_r\mathbf{a}_r$ são dadas por

$$c_1 = \mathbf{x} \cdot \mathbf{a}_1,\ c_2 = \mathbf{x} \cdot \mathbf{a}_2,\ \ldots,\ c_r = \mathbf{x} \cdot \mathbf{a}_r.^{62} \tag{2.19}$$

7. Considere

$$\mathbf{a}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \text{ e } \mathbf{a}_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Demonstre que $\{\mathbf{a}_1, \mathbf{a}_2\}$ é uma base ortonormal de $\mathbb{R}^2$ e determine as coordenadas de

$$\begin{aligned} \mathbf{x} &= \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ &= \mathbf{e}_1 + 2\mathbf{e}_2 \end{aligned}$$

nessa base.[63]

8. Considere

$$\mathbf{a}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{bmatrix},\ \mathbf{a}_2 = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ 0 \\ \frac{2}{\sqrt{6}} \end{bmatrix} \text{ e } \mathbf{a}_3 = \begin{bmatrix} -\frac{1}{2\sqrt{3}} \\ \frac{1}{2\sqrt{3}} \\ \frac{3}{2\sqrt{3}} \\ -\frac{1}{2\sqrt{3}} \end{bmatrix}.$$

Em $\mathbb{R}^n$, considere que $\mathbf{a}_i = \mathbf{0}$ é um entre os vetores $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$. Esses vetores são LD. De fato, para termos uma CL nula

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_r\mathbf{a}_r = \mathbf{0}$$

não trivial, basta considerarmos que c_i é o único coeficiente não nulo. Por exemplo, para $c_i = 1$, temos

$$0 \cdot \mathbf{a}_1 + \cdots + 0 \cdot \mathbf{a}_{i-1} + 1 \cdot \mathbf{a}_i + 0 \cdot \mathbf{a}_{i+1} + \cdots + 0 \cdot \mathbf{a}_r = \mathbf{0}.$$

[61] DICA
Note que, $\left(c_1 - c_1'\right)\mathbf{a}_1 + \cdots + \left(c_r - c_r'\right)\mathbf{a}_r = \mathbf{0}$. Agora, utilize o fato de que os r vetores são LI.

[62] DICA
Multiplique ambos os membros da CL

$$\mathbf{x} = c_1\mathbf{a}_1 + \cdots + c_r\mathbf{a}_r$$

por $\mathbf{a}_i$ e, assim, obtenha o escalar c_i, $i = 1, \ldots, r$, conforme fizemos na subseção 2.5.5.

[63] SUGESTÃO
Calcule c_1 e c_2 via (2.19).

Demonstre que $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ é uma base ortonormal de um subespaço $\mathcal{S}$ (de dimensão 3) do $\mathbb{R}^4$ e, caso seja possível, determine as coordenadas de

$$\begin{aligned}\mathbf{x} &= \begin{bmatrix} 1 \\ 3 \\ 1 \\ 1 \end{bmatrix} \\ &= \mathbf{e}_1 + 3\mathbf{e}_2 + \mathbf{e}_3 + \mathbf{e}_4\end{aligned}$$

nessa base.[64]

9. Sejam r e n dois inteiros positivos com $2 \leq r \leq n$. *(O processo de ortogonalização de) Gram-Schmidt* utiliza uma base $\mathcal{B} = \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r\}$ de um subespaço $\mathcal{S}$ do $\mathbb{R}^n$ para obter uma base ortogonal $\mathcal{B}' = \{\mathbf{a}'_1, \mathbf{a}'_2, \ldots, \mathbf{a}'_r\}$ de $\mathcal{S}$ via:

$$\begin{aligned}
\mathbf{a}'_1 &:= \mathbf{a}_1; \\
\mathbf{a}'_2 &:= \mathbf{a}_2 - \frac{\mathbf{a}_2 \cdot \mathbf{a}'_1}{\mathbf{a}'_1 \cdot \mathbf{a}'_1}\mathbf{a}'_1; \\
\mathbf{a}'_3 &:= \mathbf{a}_3 - \frac{\mathbf{a}_3 \cdot \mathbf{a}'_1}{\mathbf{a}'_1 \cdot \mathbf{a}'_1}\mathbf{a}'_1 - \frac{\mathbf{a}_3 \cdot \mathbf{a}'_2}{\mathbf{a}'_2 \cdot \mathbf{a}'_2}\mathbf{a}'_2; \\
&\vdots \\
\mathbf{a}'_r &:= \mathbf{a}_r - \frac{\mathbf{a}_r \cdot \mathbf{a}'_1}{\mathbf{a}'_1 \cdot \mathbf{a}'_1}\mathbf{a}'_1 - \frac{\mathbf{a}_r \cdot \mathbf{a}'_2}{\mathbf{a}'_2 \cdot \mathbf{a}'_2}\mathbf{a}'_2 - \cdots - \frac{\mathbf{a}_r \cdot \mathbf{a}'_{r-1}}{\mathbf{a}'_{r-1} \cdot \mathbf{a}'_{r-1}}\mathbf{a}'_{r-1}.
\end{aligned}$$

Demonstre que, por exemplo, se $r = 3$, então $\mathcal{B}' = \{\mathbf{a}'_1, \mathbf{a}'_2, \mathbf{a}'_3\}$ é ortogonal.[65]

10. Obtenha uma base ortonormal de $\mathbb{R}^2$ aplicando Gram-Schmidt na base $\{(1,2),(3,4)\}$.

RESOLUÇÃO

Os dois vetores dados são LI, pois nenhum deles é múltiplo escalar do outro.[66] Logo, como esses vetores geram um subespaço $\mathcal{S}$ de dimensão 2 do $\mathbb{R}^2$, $\mathcal{S} = \mathbb{R}^2$. Por outro lado, denotemos $\mathbf{a}_1 = (1,2)$ e $\mathbf{a}_2 = (3,4)$, para podermos utilizar a questão anterior. Assim:

$$\begin{aligned}
\mathbf{a}'_1 &= (1,2); \\
\mathbf{a}'_2 &= (3,4) - \frac{(3,4)\cdot(1,2)}{(1,2)\cdot(1,2)}(1,2) \\
&= \left(3 - \frac{11}{5}, 4 - \frac{22}{5}\right) \\
&= \left(\frac{4}{5}, -\frac{2}{5}\right).
\end{aligned}$$

[64] Se $\mathbf{x}$ é CL dos vetores $\mathbf{a}_1$, $\mathbf{a}_2$ e $\mathbf{a}_3$, então $\mathbf{x} \in \mathcal{S}$. Embora não saibamos, *a priori*, se essa CL pode ser obtida, podemos utilizar o método heurístico, supondo que é possível obtê-la, na tentativa de calcularmos as coordenadas de $\mathbf{x}$ na base supracitada. Caso essas coordenadas sejam obtidas, a suposição feita é, de fato, uma afirmação válida.

[65] SUGESTÃO

Basta verificarmos que $\mathbf{a}'_1 \cdot \mathbf{a}'_2 = 0$, $\mathbf{a}'_1 \cdot \mathbf{a}'_3 = 0$ e $\mathbf{a}'_2 \cdot \mathbf{a}'_3 = 0$. Observe que, ao considerarmos $\mathbf{a}_2 = \mathbf{x}$ e $\mathbf{a}'_1 = \mathbf{y}$ na subseção 2.6.1, $\mathbf{a}'_2$ e $\mathbf{a}'_1$ são ortogonais.

[66] Verifique!

Como $\{\mathbf{a}_1', \mathbf{a}_2'\}$ é ortogonal,[67] podemos obter uma base $\{\mathbf{a}_1'', \mathbf{a}_2''\}$ ortonormal. De fato:

$$\begin{aligned}
\mathbf{a}_1'' &= \frac{\mathbf{a}_1'}{\|\mathbf{a}_1'\|} \\
&= \frac{1}{\sqrt{5}}(1,2); \\
\mathbf{a}_2'' &= \frac{\mathbf{a}_2'}{\|\mathbf{a}_2'\|} \\
&= \frac{1}{2\sqrt{5}/5}(4/5,-2/5).
\end{aligned}$$

11. Aplique Gram-Schmidt na base $\mathcal{B} = \{(0,0,1,1),(0,1,1,0),(1,1,0,0)\}$ de um subespaço $\mathcal{S}$ do $\mathbb{R}^4$.[68]

12. Dê um exemplo de uma base ortonormal do $\mathbb{R}^3$ que contenha o vetor $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$.[69]

13. Dê um exemplo de uma base ortonormal do $\mathbb{R}^4$ que contenha o vetor $\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$. Calcule as coordenadas de $\mathbf{x} = (1,2,3,4)$ na base ortonormal obtida.

RESOLUÇÃO

Como sugerido na "footnote" do exercício anterior, aplicaremos Gram-Schmidt numa base $\mathcal{B} = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4\}$, com $\mathbf{a}_1 = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$ e, por exemplo, $\mathbf{a}_2 = (1,0,0,0)$, $\mathbf{a}_3 = (0,1,0,0)$ e $\mathbf{a}_4 = (0,0,1,0)$.[70] Portanto:

$$\begin{aligned}
\mathbf{a}_1' &= \mathbf{a}_1 \\
&= \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right); \\
\mathbf{a}_2' &= \mathbf{a}_2 - \frac{\mathbf{a}_2 \cdot \mathbf{a}_1'}{\mathbf{a}_1' \cdot \mathbf{a}_1'}\mathbf{a}_1' \\
&= (1,0,0,0) - \frac{1/2}{1}\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) \\
&= (1,0,0,0) - \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right) \\
&= \left(\frac{3}{4}, -\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}\right);
\end{aligned}$$

---

[67] Verifique!

[68] Verifique que os três vetores dados são LI, antes de iniciar o processo de Gram-Schmidt.

[69] SUGESTÃO

Aplique Gram-Schmidt numa base $\mathcal{B} = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ com $\mathbf{a}_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$ e, por exemplo, $\mathbf{a}_2 = (1,0,0)$ e $\mathbf{a}_3 = (0,1,0)$. Outra sugestão é resolver via produto vetorial. Por exemplo, considere $\mathbf{a}_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$, $\mathbf{a}_2 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right)$ e $\mathbf{a}_3 = \frac{\mathbf{a}_1 \times \mathbf{a}_2}{\|\mathbf{a}_1 \times \mathbf{a}_2\|}$.

[70] Antes de iniciar as contas, verifique que os quatro vetores dados são LI.

$$\begin{aligned}
\mathbf{a}_3' &= \mathbf{a}_3 - \frac{\mathbf{a}_3 \cdot \mathbf{a}_1'}{\mathbf{a}_1' \cdot \mathbf{a}_1'}\mathbf{a}_1' - \frac{\mathbf{a}_3 \cdot \mathbf{a}_2'}{\mathbf{a}_2' \cdot \mathbf{a}_2'}\mathbf{a}_2' \\
&= (0,1,0,0) - \frac{1/2}{1}\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right) + \frac{1/4}{3/4}\left(\frac{3}{4},-\frac{1}{4},-\frac{1}{4},-\frac{1}{4}\right) \\
&= (0,1,0,0) - \left(\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4}\right) + \left(\frac{1}{4},-\frac{1}{12},-\frac{1}{12},-\frac{1}{12}\right) \\
&= \left(0,\frac{2}{3},-\frac{1}{3},-\frac{1}{3}\right); \\
\mathbf{a}_4' &= \mathbf{a}_4 - \frac{\mathbf{a}_4 \cdot \mathbf{a}_1'}{\mathbf{a}_1' \cdot \mathbf{a}_1'}\mathbf{a}_1' - \frac{\mathbf{a}_4 \cdot \mathbf{a}_2'}{\mathbf{a}_2' \cdot \mathbf{a}_2'}\mathbf{a}_2' - \frac{\mathbf{a}_4 \cdot \mathbf{a}_3'}{\mathbf{a}_3' \cdot \mathbf{a}_3'}\mathbf{a}_3' \\
&= (0,0,1,0) - \frac{1/2}{1}\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right) + \frac{1/4}{3/4}\left(\frac{3}{4},-\frac{1}{4},-\frac{1}{4},-\frac{1}{4}\right) + \frac{1/3}{2/3}\left(0,\frac{2}{3},-\frac{1}{3},-\frac{1}{3}\right) \\
&= (0,0,1,0) - \left(\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4}\right) + \left(\frac{1}{4},-\frac{1}{12},-\frac{1}{12},-\frac{1}{12}\right) + \left(0,\frac{1}{2},-\frac{1}{6},-\frac{1}{6}\right) \\
&= \left(0,0,\frac{1}{2},-\frac{1}{2}\right).
\end{aligned}$$

Como $\mathcal{B}' = \{\mathbf{a}_1', \mathbf{a}_2', \mathbf{a}_3', \mathbf{a}_4'\}$ é ortogonal, podemos obter $\mathcal{B}'' = \{\mathbf{a}_1'', \mathbf{a}_2'', \mathbf{a}_3'', \mathbf{a}_4''\}$ ortonormal. De fato:

$$\begin{aligned}
\mathbf{a}_1'' &= \frac{\mathbf{a}_1'}{\|\mathbf{a}_1'\|} \\
&= \left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right); \\
\mathbf{a}_2'' &= \frac{\mathbf{a}_2'}{\|\mathbf{a}_2'\|} \\
&= \frac{1}{\sqrt{3}/2}\left(\frac{3}{4},-\frac{1}{4},-\frac{1}{4},-\frac{1}{4}\right) \\
&= \left(\frac{\sqrt{3}}{2},-\frac{\sqrt{3}}{6},-\frac{\sqrt{3}}{6},-\frac{\sqrt{3}}{6}\right); \\
\mathbf{a}_3'' &= \frac{\mathbf{a}_3'}{\|\mathbf{a}_3'\|} \\
&= \frac{1}{\sqrt{2/3}}\left(0,\frac{2}{3},-\frac{1}{3},-\frac{1}{3}\right) \\
&= \left(0,\frac{\sqrt{6}}{3},-\frac{\sqrt{6}}{6},-\frac{\sqrt{6}}{6}\right); \\
\mathbf{a}_4'' &= \frac{\mathbf{a}_4'}{\|\mathbf{a}_4'\|} \\
&= \frac{1}{\sqrt{1/2}}\left(0,0,\frac{1}{2},-\frac{1}{2}\right) \\
&= \left(0,0,\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2}\right).
\end{aligned}$$

Para obtermos as coordenadas de $\mathbf{x} = (1,2,3,4)$ na base $\mathcal{B}''$, escrevemos

$$\mathbf{x} = c_1\mathbf{a}_1'' + c_2\mathbf{a}_2'' + c_3\mathbf{a}_3'' + c_4\mathbf{a}_4''.$$

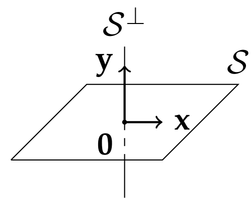
Agora, para evitar a resolução de um sistema linear não trivial de quatro equações nas variáveis c_1, c_2, c_3 e c_4, podemos calcular

$$c_i = \mathbf{x} \cdot \mathbf{a}_i'',\ i = 1,2,3,4,$$

conforme a fórmula (2.19),[71] pois $\mathcal{B}''$ é ortonormal.

[71]Cf. p. 41.

Figura 2.7: Complemento ortogonal de um plano em $\mathbb{R}^3$



14. Em $\mathbb{R}^n$, o *complemento ortogonal* $\mathcal{S}^\perp$ *(do subespaço* $\mathcal{S}$*)* é o conjunto dos vetores ortogonais aos vetores de $\mathcal{S}$, ou seja,

$$\mathcal{S}^\perp = \left\{ \mathbf{y} \in \mathbb{R}^n \,\middle|\, \mathbf{x} \cdot \mathbf{y} = 0 \text{ para cada } \mathbf{x} \in \mathcal{S} \right\},$$

conforme a figura 2.7.

(a) Para todo n, demonstre que $\mathcal{S}^\perp$ é subespaço de $\mathbb{R}^n$.[72]

(b) Determine uma base para $\mathcal{S}^\perp$ se n e a base de $\mathcal{S}$ são dadas, respectivamente, por:

i. 2 e $\{(1,1)\}$; RESPOSTA $\{(1,-1)\}$;

ii. 3 e $\{(1,1,1)\}$; RESPOSTA $\{(1,0,-1),(0,1,-1)\}$;

iii. 3 e $\{(1,0,1),(0,1,1)\}$; RESPOSTA $\{(-1,-1,1)\}$;

iv. 4 e $\{(1,1,1,1)\}$; RESPOSTA $\{(1,0,0,-1),(0,1,0,-1),(0,0,1,-1)\}$;

v. 4 e $\{(1,0,0,1),(0,1,1,0)\}$; RESPOSTA $\{(1,0,0,-1),(0,1,-1,0)\}$;

vi. 4 e $\{(1,1,0,0),(0,1,1,0),(0,0,1,1)\}$; RESPOSTA $\{(1,-1,1,-1)\}$.

OBSERVAÇÕES

· Como, para cada $\mathbf{x} \in \mathcal{S}$, se $\mathbf{y} \in \mathcal{S}^\perp$, então $\mathbf{y}$ e $\mathbf{x}$ são ortogonais, em particular, caso seja possível obtermos alguma base de $\mathcal{S}$, digamos, $\{\mathbf{x}_1, \ldots, \mathbf{x}_r\}$, $\mathbf{y}$ e $\mathbf{x}_i$ são ortogonais, $i = 1, \ldots, r$.

[72] RESOLUÇÃO

As condições 1, 2 e 3, dadas no início da seção 2.5, com $\mathcal{S}^\perp$ no lugar de $\mathcal{S}$, são satisfeitas. De fato:

1. $\mathbf{0} \in \mathcal{S}^\perp$, pois $\mathbf{x} \cdot \mathbf{0} = 0$ para todo $\mathbf{x} \in \mathcal{S}$;

2. Sejam $\alpha \in \mathbb{R}$ e $\mathbf{y} \in \mathcal{S}^\perp$, isto é, $\mathbf{x} \cdot \mathbf{y} = 0$ para todo $\mathbf{x} \in \mathcal{S}$. assim, $\alpha\mathbf{y} \in \mathcal{S}^\perp$, pois, para todo $\mathbf{x} \in \mathcal{S}$, temos

$$\begin{aligned} \mathbf{x} \cdot (\alpha \mathbf{y}) &= \alpha\,(\mathbf{x} \cdot \mathbf{y}) \\ &= \alpha \cdot 0 \\ &= 0; \end{aligned}$$

3. Sejam $\mathbf{y}_1, \mathbf{y}_2 \in \mathcal{S}^\perp$, ou seja, $\mathbf{x} \cdot \mathbf{y}_1 = 0 = \mathbf{x} \cdot \mathbf{y}_2$ para todo $\mathbf{x} \in \mathcal{S}$. Logo, $\mathbf{y}_1 + \mathbf{y}_2 \in \mathcal{S}^\perp$, pois, para todo $\mathbf{x} \in \mathcal{S}$,

$$\begin{aligned} \mathbf{x} \cdot (\mathbf{y}_1 + \mathbf{y}_2) &= \mathbf{x} \cdot \mathbf{y}_1 + \mathbf{x} \cdot \mathbf{y}_2 \\ &= 0 + 0 \\ &= 0. \end{aligned}$$

· Em (R33),[73] demonstraremos que $\dim \mathcal{S} + \dim \mathcal{S}^\perp = n$.

Assim, por exemplo, no item v desse exercício, para escreverermos $\mathbf{y} = (y_1, y_2, y_3, y_4) \in \mathcal{S}^\perp$ como CL dos vetores de uma base de $\mathcal{S}^\perp$, basta resolvermos o sistema linear

$$\begin{cases} \mathbf{x}_1 \cdot \mathbf{y} = 0 \\ \mathbf{x}_2 \cdot \mathbf{y} = 0 \end{cases}$$

onde $\mathbf{x}_1$ e $\mathbf{x}_2$ são os vetores da base de $\mathcal{S}$ dada. A base de $\mathcal{S}^\perp$, obtida do sistema linear supracitado, será composta por dois vetores. De fato,

$$\dim \mathcal{S} = 2 \text{ e } n = 4 \Longrightarrow \dim \mathcal{S}^\perp = 2.$$

15. Considere que $\mathcal{S}$ é o subespaço unidimensional de $\mathbb{R}^4$ gerado por $(1,1,1,1)$. Seja $\mathbf{y} = (1,1,1,-3)$.

    (a) É verdade que
    $$\mathbf{y} \in \mathcal{S}^\perp.$$
    Justifique corretamente essa afirmação.

    (b) Determine uma base ortonormal de $\mathcal{S}^\perp$.

    (c) Calcule as coordenadas de $\mathbf{y}$ na base obtida no item anterior.

RESOLUÇÃO

(a) Como $\mathbf{y}$ e o gerador de $\mathcal{S}$ são ortogonais, $\mathbf{y}$ e qualquer vetor de $\mathcal{S}$ são ortogonais. Portanto, $\mathbf{y} \in \mathcal{S}^\perp$.[74]

(b) Sejam $\mathbf{y}_1$, $\mathbf{y}_2$ e $\mathbf{y}_3$ os vetores de $\mathcal{B}$, na ordem em que aparecem nessa base. Como $\mathcal{B}$ não é ortogonal, usaremos Gram-Schimdt para obtermos uma base $\mathcal{B}' = \{\mathbf{y}_1', \mathbf{y}_2', \mathbf{y}_3'\}$ ortogonal. Portanto:

$$\begin{aligned}
\mathbf{y}_1' &= \mathbf{y}_1 \\
&= (1,0,0,-1); \\
\mathbf{y}_2' &= \mathbf{y}_2 - \frac{\mathbf{y}_2 \cdot \mathbf{y}_1'}{\mathbf{y}_1' \cdot \mathbf{y}_1'} \mathbf{y}_1' \\
&= (0,1,0,-1) - \frac{1}{2}(1,0,0,-1) \\
&= (0,1,0,-1) - \left(\frac{1}{2}, 0, 0, -\frac{1}{2}\right) \\
&= \left(-\frac{1}{2}, 1, 0, -\frac{1}{2}\right); \\
\mathbf{y}_3' &= \mathbf{y}_3 - \frac{\mathbf{y}_3 \cdot \mathbf{y}_1'}{\mathbf{y}_1' \cdot \mathbf{y}_1'} \mathbf{y}_1' - \frac{\mathbf{y}_3 \cdot \mathbf{y}_2'}{\mathbf{y}_2' \cdot \mathbf{y}_2'} \mathbf{y}_2' \\
&= (0,0,1,-1) - \frac{1}{2}(1,0,0,-1) - \frac{1/2}{3/2}\left(-\frac{1}{2}, 1, 0, -\frac{1}{2}\right) \\
&= (0,0,1,-1) - \left(\frac{1}{2}, 0, 0, -\frac{1}{2}\right) - \left(-\frac{1}{6}, \frac{1}{3}, 0, -\frac{1}{6}\right) \\
&= \left(-\frac{1}{3}, -\frac{1}{3}, 1, -\frac{1}{3}\right).
\end{aligned}$$

[73] Cf. p. 193.

[74] Na questão 14 dessa subseção, item iv, determinamos a base $\mathcal{B} = \{(1,0,0,-1), (0,1,0,-1), (0,0,1,-1)\}$ de $\mathcal{S}^\perp$. Assim, outro modo de verificar que $\mathbf{y} \in \mathcal{S}^\perp$ é observarmos que $\mathbf{y}$ é a soma dos três vetores de $\mathcal{B}$.

Agora, $\mathcal{B}'' = \{\mathbf{y}_1'', \mathbf{y}_2'', \mathbf{y}_3''\}$ é ortonormal para:

$$\begin{aligned}
\mathbf{y}_1'' &= \frac{\mathbf{y}_1'}{\|\mathbf{y}_1'\|} \\
&= \left(\frac{1}{\sqrt{2}}, 0, 0, -\frac{1}{\sqrt{2}}\right); \\
\mathbf{y}_2'' &= \frac{\mathbf{y}_2'}{\|\mathbf{y}_2'\|} \\
&= \frac{1}{\sqrt{3/2}}\left(-\frac{1}{2}, 1, 0, -\frac{1}{2}\right) \\
&= \left(-\frac{1}{\sqrt{6}}, \sqrt{\frac{2}{3}}, 0, -\frac{1}{\sqrt{6}}\right); \\
\mathbf{y}_3'' &= \frac{\mathbf{y}_3'}{\|\mathbf{y}_3'\|} \\
&= \frac{1}{2/\sqrt{3}}\left(-\frac{1}{3}, -\frac{1}{3}, 1, -\frac{1}{3}\right) \\
&= \left(\frac{-\sqrt{3}}{6}, -\frac{\sqrt{3}}{6}, \frac{\sqrt{3}}{2}, -\frac{\sqrt{3}}{6}\right).
\end{aligned}$$

(c) Para obtermos as coordenadas de $\mathbf{y}$ na base $\mathcal{B}''$, ao escrevermos

$$\mathbf{y} = c_1\mathbf{y}_1'' + c_2\mathbf{y}_2'' + c_3\mathbf{y}_3'',$$

podemos calcular

$$c_i = \mathbf{y} \cdot \mathbf{y}_i'', \, i = 1, 2, 3,$$

conforme a fórmula (2.19).[75]

2.7.1 Resolução do exercício da subseção 2.5.1

As condições 1, 2 e 3, do início da seção 2.5,[76] devem ser verificadas para cada um dos itens do exercício (supracitado). Essa verificação é trivial para os itens 1 e 2 do exercício. Por exemplo, $\mathcal{S} = \{\mathbf{0}\}$ é subespaço, pois:

- $\mathbf{0} \in \mathcal{S}$;
- $\alpha\mathbf{x} \in \mathcal{S}$, para quaisquer $\alpha \in \mathbb{R}$ e $\mathbf{x} \in \mathcal{S}$, pois

$$\begin{aligned}
\alpha\mathbf{x} &= \alpha\mathbf{0} \\
&= \mathbf{0};
\end{aligned}$$

- $\mathbf{x} + \mathbf{y} \in \mathcal{S}$, para quaisquer $\mathbf{x}, \mathbf{y} \in \mathcal{S}$, pois

$$\begin{aligned}
\mathbf{x} + \mathbf{y} &= \mathbf{0} + \mathbf{0} \\
&= \mathbf{0}.
\end{aligned}$$

Para o item 3 do exercício, isto é, para $\mathcal{S} = \{\mathbf{x} = t\mathbf{a} \mid t \in \mathbb{R}\}$, considere:

- $\mathbf{0} \in \mathcal{S}$, pois $\mathbf{0} = 0\mathbf{a}$;
- $\alpha\mathbf{x} \in \mathcal{S}$, para quaisquer $\alpha \in \mathbb{R}$ e $\mathbf{x} \in \mathcal{S}$, pois $\mathbf{x} = t\mathbf{a}$, onde $t \in \mathbb{R}$, e, portanto,

$$\begin{aligned}
\alpha\mathbf{x} &= \alpha(t\mathbf{a}) \\
&= (\alpha t)\mathbf{a},
\end{aligned}$$

sendo que $\alpha t \in \mathbb{R}$;

[75] Cf. p. 41.
[76] Cf. p. 23.

- $\mathbf{x}+\mathbf{y}\in\mathcal{S}$, para quaisquer $\mathbf{x},\mathbf{y}\in\mathcal{S}$, pois $\mathbf{x}=t_1\mathbf{a}$ e $\mathbf{y}=t_2\mathbf{a}$, onde $t_1,t_2\in\mathbb{R}$, e, assim,

$$\begin{aligned}\mathbf{x}+\mathbf{y}&=t_1\mathbf{a}+t_2\mathbf{a}\\&=(t_1+t_2)\,\mathbf{a},\end{aligned}$$

  sendo que $t_1+t_2\in\mathbb{R}$.

Para o item 5 do exercício,[77] ou seja, $\mathcal{S}=\{\mathbf{x}\,|\,\mathbf{a}_i\cdot\mathbf{x}=0,\ i=1,2,\ldots,r\}$, considere:

- $\mathbf{0}\in\mathcal{S}$, pois $\mathbf{a}_i\cdot\mathbf{0}=0$, $i=1,2,\ldots,r$;
- $\alpha\mathbf{x}\in\mathcal{S}$, para quaisquer $\alpha\in\mathbb{R}$ e $\mathbf{x}\in\mathcal{S}$, pois

$$\begin{aligned}\mathbf{a}_i\cdot(\alpha\mathbf{x})&=\alpha\,(\mathbf{a}_i\cdot\mathbf{x})\\&=\alpha 0\\&=0,\end{aligned}$$

  $i=1,2,\ldots,r$;
- $\mathbf{x}+\mathbf{y}\in\mathcal{S}$, para quaisquer $\mathbf{x},\mathbf{y}\in\mathcal{S}$, pois

$$\begin{aligned}\mathbf{a}_i\cdot(\mathbf{x}+\mathbf{y})&=\mathbf{a}_i\cdot\mathbf{x}+\mathbf{a}_i\cdot\mathbf{y}\\&=0+0\\&=0,\end{aligned}$$

  $i=1,2,\ldots,r$.

Para o item 6 do exercício, isto é, o conjunto $\mathcal{S}$ de todas as CL's dos vetores $\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_r\in\mathbb{R}^n$, considere:

- $\mathbf{0}\in\mathcal{S}$, pois $\mathbf{0}=0\mathbf{a}_1+0\mathbf{a}_2+\cdots+0\mathbf{a}_r$;
- $\alpha\mathbf{x}\in\mathcal{S}$, para quaisquer $\alpha\in\mathbb{R}$ e $\mathbf{x}\in\mathcal{S}$, pois, se $\mathbf{x}=c_1\mathbf{a}_1+c_2\mathbf{a}_2+\cdots+c_r\mathbf{a}_r$, onde $c_i\in\mathbb{R}$, $i=1,2,\ldots,r$, então

$$\begin{aligned}\alpha\mathbf{x}&=\alpha\,(c_1\mathbf{a}_1+c_2\mathbf{a}_2+\cdots+c_r\mathbf{a}_r)\\&=\alpha\,(c_1\mathbf{a}_1)+\alpha\,(c_2\mathbf{a}_2)+\cdots+\alpha\,(c_r\mathbf{a}_r)\\&=(\alpha c_1)\,\mathbf{a}_1+(\alpha c_2)\,\mathbf{a}_2+\cdots+(\alpha c_r)\,\mathbf{a}_r,\end{aligned}$$

  onde $\alpha c_i\in\mathbb{R}$, $i=1,2,\ldots,r$;
- $\mathbf{x}+\mathbf{y}\in\mathcal{S}$, para quaisquer $\mathbf{x},\mathbf{y}\in\mathcal{S}$, pois, se $\mathbf{x}=c_1\mathbf{a}_1+c_2\mathbf{a}_2+\cdots+c_r\mathbf{a}_r$ e $\mathbf{y}=d_1\mathbf{a}_1+d_2\mathbf{a}_2+\cdots+d_r\mathbf{a}_r$, onde $c_i,d_i\in\mathbb{R}$, $i=1,2,\ldots,r$, então

$$\begin{aligned}\mathbf{x}+\mathbf{y}&=c_1\mathbf{a}_1+c_2\mathbf{a}_2+\cdots+c_r\mathbf{a}_r+d_1\mathbf{a}_1+d_2\mathbf{a}_2+\cdots+d_r\mathbf{a}_r\\&=c_1\mathbf{a}_1+d_1\mathbf{a}_1+c_2\mathbf{a}_2+d_2\mathbf{a}_2+\cdots+c_r\mathbf{a}_r+d_r\mathbf{a}_r\\&=(c_1+d_1)\,\mathbf{a}_1+(c_2+d_2)\,\mathbf{a}_2+\cdots+(c_r+d_r)\,\mathbf{a}_r,\end{aligned}$$

  onde $c_i+d_i\in\mathbb{R}$, $i=1,2,\ldots,r$.

[77] Note que, o item 4 é um caso particular do item 5 ($r=1$).

Capítulo 3

O espaço vetorial $\mathbb{R}^{m \times n}$

> *O espaço $\mathbb{R}^n$, do capítulo 2, é um caso particular do espaço $\mathbb{R}^{m\times n}$ das "matrizes" $m \times n$. Esse espaço, além de generalizar o $\mathbb{R}^n$, é fundamental no estudo de "sistemas lineares".*

3.1 Adição de matrizes e multiplicação por escalares

3.1.1 Matrizes

- Doravante, denotaremos *matrizes* por letras maiúsculas em itálico, isto é,

$$A, B, C, D, I, M, R, \text{ etc.}$$

  Contudo, algumas matrizes especiais serão representadas pelas seguintes letras maiúsculas em "sans-serif":

$$\mathsf{O}, \mathsf{D}, \mathsf{I}, \mathsf{R}, \mathsf{E} \text{ e } \mathsf{P}.$$

- As *entradas* (ou os *elementos*) de $A \in \mathbb{R}^{m\times n}$ são representadas(os) por a_{ij}, caso tenham *índices* $i = 1, 2, \ldots, m$ e $j = 1, 2, \ldots, n$, ou seja, se A é $m \times n$, então

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}. \tag{3.1}$$

> *Representações similares valem para matrizes B, C, D, I, M, R, etc.*

- Embora essas entradas sejam números reais em todos os exemplos dos capítulos 3 e 4, elas poderão assumir, a partir do capítulo 5, valores complexos com partes imaginárias não nulas.

- Para i fixo e j variável, a i-ésima *linha* de A é representada por

$$A(i, -) = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix}.$$

- Para i variável e j fixo, a j-ésima *coluna* de A é representada por

$$A(-,j) := \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} = \mathbf{a}_j.$$

$\boxed{\text{EXEMPLO}}$

Para $i = 1,2,3$ e $j = 1,2,3,4,5$, a entrada a_{ij}, a linha $A(i,-)$ e a coluna $A(-,j) = \mathbf{a}_j$ de

$$A = \begin{bmatrix} 1 & -1 & 0 & 2 & \sqrt{2} \\ \pi & \pi/2 & 1 & -2 & 0 \\ 0 & -1 & 1/\sqrt{2} & 1/2 & -\pi \end{bmatrix} \in \mathbb{R}^{3\times 5}$$

são dadas por:

$$\begin{aligned} a_{11} = -a_{12} = a_{23} = -a_{32} = 1; \\ a_{13} = a_{25} = a_{31} = 0; \\ a_{14} = (a_{15})^2 = -a_{24} = (a_{33})^{-2} = (a_{34})^{-1} = 2; \\ a_{21} = 2a_{22} = -a_{35} = \pi; \end{aligned}$$

$$\begin{aligned} A(1,-) &= \begin{bmatrix} 1 & -1 & 0 & 2 & \sqrt{2} \end{bmatrix}; \\ A(2,-) &= \begin{bmatrix} \pi & \pi/2 & 1 & -2 & 0 \end{bmatrix}; \\ A(3,-) &= \begin{bmatrix} 0 & -1 & 1/\sqrt{2} & 1/2 & -\pi \end{bmatrix}; \\ A(-,1) &= \begin{bmatrix} 1 \\ \pi \\ 0 \end{bmatrix} \\ &= \mathbf{a}_1; \\ A(-,2) &= \begin{bmatrix} -1 \\ \pi/2 \\ -1 \end{bmatrix} \\ &= \mathbf{a}_2; \\ A(-,3) &= \begin{bmatrix} 0 \\ 1 \\ 1/\sqrt{2} \end{bmatrix} \\ &= \mathbf{a}_3; \\ A(-,4) &= \begin{bmatrix} 2 \\ -2 \\ 1/2 \end{bmatrix} \\ &= \mathbf{a}_4; \\ A(-,5) &= \begin{bmatrix} \sqrt{2} \\ 0 \\ -\pi \end{bmatrix} \\ &= \mathbf{a}_5. \end{aligned}$$

- A *igualdade* de A e B, ambas $m \times n$, é dada por

$$A = B \Longleftrightarrow a_{ij} = b_{ij},$$

para todos os índices i e j supracitados.

- O denotará a matriz com todas as entradas nulas, isto é,

$$A = \mathsf{O} \Longleftrightarrow a_{ij} = 0,$$

para todos os índices i e j supracitados. Por esse motivo, O é chamada de matriz *nula*.

EXERCÍCIOS

1. A matriz

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 0,0000000001 \end{bmatrix}$$

é igual a matriz nula 2×2. Essa afirmação é verdadeira ou falsa? Justifique!

2. Determine o valor de t para que

$$\begin{bmatrix} t^2-1 & t^2-t \\ t^3-1 & t^2-3t+2 \end{bmatrix}$$

seja igual a matriz nula de $\mathbb{R}^{2\times 2}$. RESPOSTA $t = 1$.

3.1.2 Por que $\mathbb{R}^{m\times n}$ é um espaço vetorial?

- Porque é munido de duas operações "entrada-a-entrada", que satisfazem as oito propriedades que enumeraremos a seguir. Definamos, primeiramente, essas operações da seguinte maneira:

  · A *soma* $A + B$ de A e B, ambas $m \times n$, é a matriz $m \times n$ cuja entrada da linha i e da coluna j é definida por:

$$C = A + B \Longleftrightarrow c_{ij} := a_{ij} + b_{ij}, \quad i = 1, \dots, m, \ j = 1, \dots, n.$$

  EXEMPLO
  Em $\mathbb{R}^{2\times 3}$,

$$\begin{bmatrix} 1 & \sqrt{2} & \ln e \\ 0 & 0,1 & 1/3 \end{bmatrix} + \begin{bmatrix} -1 & -2\sqrt{2} & \ln 1 \\ \pi & 1 & -4/3 \end{bmatrix} = \begin{bmatrix} 0 & -\sqrt{2} & 1 \\ \pi & 1,1 & -1 \end{bmatrix}.$$

  · Para $\lambda \in \mathbb{R}$ e A $m \times n$, o *produto por escalar* λA é a matriz $m \times n$ cuja entrada da linha i e da coluna j é definida por:

$$D = \lambda A \Longleftrightarrow d_{ij} := \lambda a_{ij}, \quad i = 1, \dots, m, \ j = 1, \dots, n.$$

  EXEMPLO
  Em $\mathbb{R}^{2\times 2}$,

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

- Assim como para o $\mathbb{R}^n$,[1] as seguintes propriedades em $\mathbb{R}^{m\times n}$ são sempre válidas:

  1. $A + B = B + A$; (COMUTATIVIDADE DA ADIÇÃO)
  2. $(A + B) + C = A + (B + C)$; (ASSOCIATIVIDADE DA ADIÇÃO)
  3. $A + 0 = A$; (EXISTÊNCIA DA MATRIZ NULA)
  4. $-A := (-1)A \Longrightarrow A + (-A) = 0$; (EXISTÊNCIA DE MATRIZ SIMÉTRICA)
  5. $\lambda(A + B) = \lambda A + \lambda B$; (DISTRIBUTIVIDADE DO PRODUTO (POR ESCALAR) EM RELAÇÃO À SOMA DE MATRIZES)
  6. $(\lambda + \beta)A = \lambda A + \beta A$; (DISTRIBUTIVIDADE DO PRODUTO (POR ESCALAR) EM RELAÇÃO À SOMA DE ESCALARES)
  7. $(\lambda\beta)A = \lambda(\beta A)$; (ASSOCIATIVIDADE DO PRODUTO POR ESCALAR)
  8. $1A = A$. (ELEMENTO NEUTRO DO PRODUTO POR ESCALAR)

DEMONSTRAÇÃO DA PROPRIEDADE 1

Como vimos na subseção 3.1.1, a igualdade das matrizes $A + B$ e $B + A$ é equivalente a igualdade de suas entradas $a_{ij} + b_{ij}$ e $b_{ij} + a_{ij}$, $i = 1, \ldots, m$ e $j = 1, \ldots, n$. Portanto, como a adição em $\mathbb{R}$ é comutativa,[2] temos $a_{ij} + b_{ij} = b_{ij} + a_{ij}$, $i = 1, \ldots, m$ e $j = 1, \ldots, n$. Assim, $A + B = B + A$.

EXERCÍCIO

Demonstre as propriedades 2–8 supracitadas.

- Podemos estabelecer uma

> *Correspondência Biunívoca entre $\mathbb{R}^{m\times n}$ e $\mathbb{R}^{mn}$*
>
> *Em $\mathbb{R}^{m\times n}$, temos conceitos e resultados análogos aos de $\mathbb{R}^n$, tais como subespaço, CL, geradores, LI, LD, base, dimensão, etc, como ilustram alguns exemplos/exercícios desse capítulo. De fato, como pode ser visto na seção 6.2, podemos fazer corresponder biunivocamente matrizes $m \times n$ a vetores em $\mathbb{R}^{mn}$, de modo natural. Por exemplo, a matriz A dada em (3.1), p. 49, e*
>
> $$\mathbf{a} = (a_{11}, \ldots, a_{1n}, a_{21}, \ldots, a_{2n}, \ldots, a_{m1}, \ldots, a_{mn})$$
>
> *onde as mn coordenadas consecutivas de tal vetor são as entradas consecutivas de $A(1,-), A(2,-), \ldots, A(m,-)$, nesta ordem. Essa correspondência preserva combinações lineares, isto é, caso A_i corresponda biunivocamente a $\mathbf{a}_i$ e $\alpha_i \in \mathbb{R}$, $i = 1, \ldots, r$, $\alpha_1 A_1 + \cdots + \alpha_r A_r$ corresponde biunivocamente a $\alpha_1 \mathbf{a}_1 + \cdots + \alpha_r \mathbf{a}_r$.*

EXEMPLOS

· Na tabela seguinte, temos a "mesma" CL nula, escrita matricialmente e vetorialmente:

---

[1] Cf. a subseção 2.2.2, p. 18.

[2] Lembrem-se do mantra: A ORDEM DAS PARCELAS NÃO ALTERA A SOMA.

$\mathbb{R}^{3\times 2}$	$\mathbb{R}^6$
$A = \begin{bmatrix} -2 & 3 \\ 0 & 6 \\ 4 & -12 \end{bmatrix}$	$\mathbf{a} = (-2, 3, 0, 6, 4, -12)$
$B = \begin{bmatrix} 1/6 & -1/4 \\ 0 & -1/2 \\ -1/3 & 1 \end{bmatrix}$	$\mathbf{b} = (1/6, -1/4, 0, -1/2, -1/3, 1)$
$\frac{1}{12}A + B = \mathsf{O}$	$\frac{1}{12}\mathbf{a} + \mathbf{b} = \mathbf{0}$

· Do mesmo modo que, considerando a base canônica $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ de $\mathbb{R}^4$, qualquer vetor

$$\mathbf{a} = (a, b, c, d)$$

desse espaço pode ser escrito da forma

$$\mathbf{a} = a\mathbf{e}_1 + b\mathbf{e}_2 + c\mathbf{e}_3 + d\mathbf{e}_4,$$

temos a base canônica $\{E_1, E_2, E_3, E_4\}$ de $\mathbb{R}^{2\times 2}$, onde qualquer matriz

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

desse espaço pode ser escrita da forma

$$A = aE_1 + bE_2 + cE_3 + dE_4.$$

Aqui, obviamente,

$$E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$
$$E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$
$$E_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \text{ e}$$
$$E_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

EXERCÍCIO

Para números reais C, D, E e F não nulos, seja $\mathcal{S}$ o subespaço de $\mathbb{R}^{2\times 2}$ gerado por

$$A_1 = \begin{pmatrix} C & 0 \\ 0 & D \end{pmatrix} \text{ e } A_2 = \begin{pmatrix} 0 & E \\ F & 0 \end{pmatrix}.$$

(a) Prove que A_1 e A_2 são LI;

(b) Apresente alguma matriz $A_3 \in \mathbb{R}^{2\times 2}$ tal que $A_3 \notin \mathcal{S}$;

(c) Apresente um subespaço $\mathcal{S}'$ de $\mathbb{R}^{2\times 2}$ que contenha $\mathcal{S}$ com $\mathcal{S} \neq \mathcal{S}' \neq \mathbb{R}^{2\times 2}$.

RESOLUÇÃO

(a) Basta verificar que A_1 e A_2 não são múltiplas uma da outra.

(b) Verifique que, por exemplo, se A_3 é uma matriz da base canônica de $\mathbb{R}^{2\times 2}$,[3] A_3 não pode ser escrita CL de A_1 e A_2.

(c) Seja A_3 como no item (b) desse exercício. Considere que $\mathcal{S}'$ seja gerado por A_1, A_2 e A_3. Pelos itens (a) e (b) desse exercício, temos que $\mathcal{S} \subset \mathcal{S}'$, $\mathcal{S} \neq \mathcal{S}'$ e A_1, A_2 e A_3 são LI. Portanto, $\dim \mathcal{S}' = 3$ e $\mathcal{S}' \neq \mathbb{R}^{2\times 2}$, pois $\dim \mathbb{R}^{2\times 2} = 4$.

- Em $\mathbb{R}^{m\times n}$, $A - B := A + (-B)$.

EXEMPLO

Em $\mathbb{R}^{2\times 3}$,

$$\begin{bmatrix} 1 & \sqrt{2} & \ln e \\ 0 & 0,1 & 1/3 \end{bmatrix} - \begin{bmatrix} -1 & -2\sqrt{2} & \ln 1 \\ \pi & 1 & -4/3 \end{bmatrix} = \begin{bmatrix} 2 & 3\sqrt{2} & 1 \\ -\pi & -0,9 & 5/3 \end{bmatrix}.$$

3.2 Produto e transposição de matrizes

- Temos uma multiplicação "linha-por-coluna" análoga ao produto interno visto no capítulo anterior.

O *produto* $AB \in \mathbb{R}^{m\times n}$ de $A \in \mathbb{R}^{m\times p}$ e $B \in \mathbb{R}^{p\times n}$, nessa ordem, está assim definido:

$$C = AB \iff c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj} \quad \text{para } i = 1, 2, \ldots, m \text{ e } j = 1, 2, \ldots, n.$$

Nesse caso, denota-se $c_{ij} := A(i,-)\cdot B(-,j)$.

$$\therefore \quad AB = \begin{bmatrix} A(1,-)\cdot B(-,1) & A(1,-)\cdot B(-,2) & \cdots & A(1,-)\cdot B(-,p) \\ A(2,-)\cdot B(-,1) & A(2,-)\cdot B(-,2) & \cdots & A(2,-)\cdot B(-,p) \\ \vdots & \vdots & \vdots & \vdots \\ A(m,-)\cdot B(-,1) & A(m,-)\cdot B(-,2) & \cdots & A(m,-)\cdot B(-,p) \end{bmatrix}.$$

Seja $\mathbf{b} \in \mathbb{R}^p$ uma coluna qualquer de B. Digamos, $\mathbf{b} = B(-,j)$. Considere, agora, as p entradas consecutivas, da esquerda para a direita, de uma linha arbitrária de A. Digamos que $A(i,-)$ seja a linha considerada. Escreva, então, essas p entradas como as coordenadas consecutivas de um vetor $\mathbf{a} \in \mathbb{R}^p$. Assim, $\mathbf{a}\cdot\mathbf{b}$ é o produto (interno) da linha e da coluna supracitadas, como descrito anteriormente, ou seja,

$$\mathbf{a}\cdot\mathbf{b} = A(i,-)\cdot B(-,j).$$

EXEMPLO

Para $A \in \mathbb{R}^{2\times 2}$ e $B \in \mathbb{R}^{2\times 3}$, digamos

$$A = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \text{ e } B = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix},$$

[3] Lembre-se que a base canônica é composta pelas matrizes que possuem apenas uma entrada não nula e igual a 1, conforme o exemplo anterior.

temos a matriz 2×3

$$AB = \begin{bmatrix} A(1,-)\cdot B(-,1) & A(1,-)\cdot B(-,2) & A(1,-)\cdot B(-,3) \\ A(2,-)\cdot B(-,1) & A(2,-)\cdot B(-,2) & A(2,-)\cdot B(-,3) \end{bmatrix}$$
$$= \begin{bmatrix} \alpha a+\beta d & \alpha b+\beta e & \alpha c+\beta f \\ \gamma a+\delta d & \gamma b+\delta e & \gamma c+\delta f \end{bmatrix}.$$

- Em geral, o produto de matrizes não é comutativo.

EXEMPLOS

- Para A e B dados no exemplo anterior, embora possamos calcular AB, BA não está definido.
- Mesmo que tenhamos $A, B \in \mathbb{R}^{2\times 2}$, em geral, $AB \neq BA$.
  Verifique essa afirmação para
  $$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ e } B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

($*$) *Para $A, B \in \mathbb{R}^{m\times p}$ e $C \in \mathbb{R}^{p\times n}$ arbitrárias, pode ser demonstrado que*

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

EXERCÍCIO

Verifique essa distributividade para A e B dados no último exemplo e $C = I - (A+B)$.

- Temos a operação de *transposição*:

$$A \text{ é } m\times n \Longrightarrow A^t = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}.$$

Note que, A^t é $n \times m$ e

$$T = A^t \Longleftrightarrow t_{ij} = a_{ji} \text{ com } i = 1,2,\ldots,n \text{ e } j = 1,2,\ldots,m.$$

EXEMPLO

$$A = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \text{ e } B = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \Longrightarrow A^t = \begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix} \text{ e } B^t = \begin{bmatrix} a & d \\ b & e \\ c & f \end{bmatrix}.$$

($*$) *Para A e B com AB bem definido, como esse produto de matrizes se comporta sob a ação da transposição?*
Pode ser demonstrado que

$$(AB)^t = B^t A^t.$$

EXERCÍCIO

Verifique essa propriedade de transposição para o exemplo anterior.

3.3 Importância das matrizes quadradas

- Uma matriz A, $m \times n$, é dita *quadrada* quando $m = n$. Nesse caso, dizemos que *A é de ordem n*, sua *diagonal principal* é composta pelas entradas a_{ii}, $i = 1, \ldots, n$, e, caso $i \neq j$, a_{ij} é uma entrada *fora da diagonal principal* (de A).
- Dizemos que uma matriz A é *simétrica* quando

$$A^t = A.$$

Portanto, não existem matrizes simétricas que não sejam quadradas. Além disso, pela definição de A^t, temos igualdade entre as entradas "simétricas" em relação à diagonal principal, ou seja, $a_{ij} = a_{ji}$, caso $i \neq j$.

EXEMPLO DE MATRIZ SIMÉTRICA

$$A = \begin{bmatrix} a & \alpha & \beta \\ \alpha & b & \gamma \\ \beta & \gamma & c \end{bmatrix}.$$

- Uma matriz quadrada D cujas entradas fora da diagonal principal são nulas, ou seja,

$$\mathsf{D} = \begin{bmatrix} d_{11} & & & & \\ & d_{22} & & \text{zeros} & \\ & & \ddots & & \\ & \text{zeros} & & d_{n-1,n-1} & \\ & & & & d_{nn} \end{bmatrix}, \tag{3.2}$$

é dita *matriz diagonal*

- Caso $d_{ii} = 1$, $i = 1, 2, \ldots, n-1, n$, (3.2) é chamada (*matriz*) *identidade* e denotada por I, ou seja,

$$\mathsf{I} = \begin{bmatrix} 1 & & & & \\ & 1 & & \text{zeros} & \\ & & \ddots & & \\ & \text{zeros} & & 1 & \\ & & & & 1 \end{bmatrix}.$$

- Demonstra-se que, caso A seja $m \times p$, I seja $p \times p$ e B seja $p \times n$,

$$A\mathsf{I} = A \quad \text{e} \quad \mathsf{I}B = B, \tag{3.3}$$

isto é, I é um *elemento neutro multiplicativo*.

EXERCÍCIO

Verifique a validade de (3.3) para

$$A = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}, \ B = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \ \text{e} \ \mathsf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

É possível que as matrizes quadradas mais importantes sejam as "invertíveis", isto é, as matrizes que têm "inversas". Estudaremos essas matrizes nas seções 3.6 e 3.7.

3.4 Determinantes

Essa seção e os exercícios da seção 3.7, relacionados ao tema, são operacionais e constituem um formulário reduzido sobre $\det A$, $A \in \mathbb{R}^{n\times n}$. A motivação para essa abordagem encontra-se no capítulo 1.

- Para $n \geq 2$, A_{ij} denota a matriz $(n-1)\times(n-1)$, obtida de $A \in \mathbb{R}^{n\times n}$, eliminando-se a linha $A(i,-)$ e a coluna $A(-,j)$.[4]

  EXEMPLO
  Considere a matriz 3×3 dada por

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

  Obtemos, portanto, as seguintes matrizes 2×2:

  · $A_{11} = \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix}$, eliminando-se $A(1,-)$ e $A(-,1)$;

  · $A_{12} = \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix}$, eliminando-se $A(1,-)$ e $A(-,2)$;

  · $A_{13} = \begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix}$, eliminando-se $A(1,-)$ e $A(-,3)$;

  · $A_{21} = \begin{bmatrix} 2 & 3 \\ 8 & 9 \end{bmatrix}$, eliminando-se $A(2,-)$ e $A(-,1)$;

  · $A_{22} = \begin{bmatrix} 1 & 3 \\ 7 & 9 \end{bmatrix}$, eliminando-se $A(2,-)$ e $A(-,2)$;

  · $A_{23} = \begin{bmatrix} 1 & 2 \\ 7 & 8 \end{bmatrix}$, eliminando-se $A(2,-)$ e $A(-,3)$;

  · $A_{31} = \begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix}$, eliminando-se $A(3,-)$ e $A(-,1)$;

  · $A_{32} = \begin{bmatrix} 1 & 3 \\ 4 & 6 \end{bmatrix}$, eliminando-se $A(3,-)$ e $A(-,2)$;

  · $A_{33} = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}$, eliminando-se $A(3,-)$ e $A(-,3)$.

- Para $A \in \mathbb{R}^{n\times n}$, $\det A$ pode ser calculado ao longo da linha $A(i,-)$, recursivamente, do seguinte modo:

  · Se $n = 1$ e $A = [a]$, então $\det A = a$;

  · Se $n = 2$, então $\det A = a_{11}a_{22} - a_{12}a_{21}$;

[4] Para obtermos A_{ij}, basta eliminarmos a linha e a coluna que se cruzam na entrada a_{ij}.

· Se $n > 2$, então

$$\det A = (-1)^{i+1} a_{i1} \det A_{i1} + (-1)^{i+2} a_{i2} \det A_{i2} + \cdots + (-1)^{i+n} a_{in} \det A_{in}.$$

Obviamente, pode ser demonstrado que, $\det A$ não depende da linha i escolhida para calculá-lo.

EXEMPLO

Ao calcularmos o determinante da matriz A do primeiro exemplo dessa seção, ao longo da primeira linha ($i = 1$), temos

$$\begin{aligned}
\det A &= (-1)^{1+1} a_{11} \det A_{11} + (-1)^{1+2} a_{12} \det A_{12} + (-1)^{1+3} a_{13} \det A_{13} \\
&= (-1)^2 \cdot 1 \cdot \det \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix} + (-1)^3 \cdot 2 \cdot \det \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix} + (-1)^4 \cdot 3 \cdot \det \begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix} \\
&= (5 \cdot 9 - 6 \cdot 8) - 2(4 \cdot 9 - 6 \cdot 7) + 3(4 \cdot 8 - 5 \cdot 7) \\
&= -3 + 12 - 9 \\
&= 0.
\end{aligned}$$

EXERCÍCIO

Para a matriz A do exemplo supracitado, verifique que, ao longo da linha $i \in \{2, 3\}$, $\det A = 0$.

- Para matrizes quadradas, pode ser demonstrado que:

  · o determinante do produto é igual ao produto dos determinantes;

  · a transposição não altera o determinante.

EXERCÍCIO

Resolva o exercício 9 da seção 3.7.

- Como transposição de matrizes não altera seus determinantes, podemos calcular $\det A$ ao longo de qualquer coluna $A(-, j)$, ou seja,

$$\det A = (-1)^{1+j} a_{1j} \det A_{1j} + (-1)^{2+j} a_{2j} \det A_{2j} + \cdots + (-1)^{n+j} a_{nj} \det A_{nj},$$

para $A \in \mathbb{R}^{n\times n}$.

EXERCÍCIO

Ao longo de cada coluna da matriz A do primeiro exemplo dessa seção, calcule $\det A$.

- Para calcularmos o determinante de uma matriz A, podemos escolher entre suas linhas e colunas, alguma que tenha o maior número de zeros. Além disso, caso A tenha alguma linha/coluna nula, $\det A = 0$.

EXERCÍCIO

Resolva o exercício 7 da seção 3.7.

3.5 Sistemas lineares $A\mathbf{x} = \mathbf{b}$ e escalonamento

- Podemos escrever o sistema linear 3×4

$$\left\{\begin{array}{rcrcrcrcr} \frac{\sqrt{2}}{2}x & + & \sqrt{2}y & - & \frac{3\sqrt{2}}{2}z & & & = & 2\sqrt{2} \\ 2x & + & 4y & - & 6z & & & = & 8 \\ & & y & - & z & + & w & = & -1 \end{array}\right. \tag{3.4}$$

da forma

$$\begin{bmatrix} \sqrt{2}/2 & \sqrt{2} & -3\sqrt{2}/2 & 0 \\ 2 & 4 & -6 & 0 \\ 0 & 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 2\sqrt{2} \\ 8 \\ -1 \end{bmatrix}.$$

Em geral, podemos escrever o sistema linear $m \times n$

$$\left\{\begin{array}{rcrcccrcl} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ & & & & \cdots & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}\right.$$

da forma

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

onde

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

- Caso $\mathbf{b} = \mathbf{0}$, o sistema linear supracitado é dito *homogêneo*.

EXEMPLO

$$\left\{\begin{array}{r} x + 2y = 0 \\ 3x + 4y = 0 \end{array}\right.$$

pode ser escrito da forma

$$A\mathbf{x} = \mathbf{0}, \text{ onde } A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{e} \quad \mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

- A é dita *matriz (de coeficientes) do sistema* $A\mathbf{x} = \mathbf{b}$, a_{ij} e b_i são, obviamente, escalares reais, para $i = 1, 2, \ldots, m$ e $j = 1, 2, \ldots, n$, e $\mathbf{x}$ é dito *vetor coluna de n variáveis*.

- Caso as variáveis $x_1, x_2, \ldots, x_n$ possam ser substituídas por escalares fixos, dizemos que o vetor $\mathbf{x} = \mathbf{x}_0$ é uma *solução de* $A\mathbf{x} = \mathbf{b}$ quando $A\mathbf{x}_0 = \mathbf{b}$.[5]

EXEMPLO

Para o sistema do exemplo anterior, como as retas $y = -\frac{1}{2}x$ e $y = -\frac{3}{4}x$ se interceptam na origem de $\mathbb{R}^2$, $\mathbf{x}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ é a única solução.

[5] Note que, o sistema homogêneo $A\mathbf{x} = \mathbf{0}$ admite, pelo menos, a *solução nula* $\mathbf{x}_0 = \mathbf{0}$.

- A *matriz aumentada* de um sistema $A\mathbf{x} = \mathbf{b}$ arbitrário é dada por

$$[A|\mathbf{b}]$$

e seu *conjunto solução* é constituído de todas as soluções desse sistema.

EXEMPLOS

· A matriz aumentada do sistema linear 2×2

$$\begin{cases} -x + 2y = -1 \\ 3x + 4y = 0 \end{cases}$$

é 2×3 e dada por

$$\left[\begin{array}{cc|c} -1 & 2 & -1 \\ 3 & 4 & 0 \end{array}\right].$$

Além disso, $\mathcal{S}$ consiste de todos os pares ordenados (x, y) de números reais que satisfazem, simultaneamente, as duas equações do sistema supracitado. Note que, como a única solução desse sistema é

$$x = \frac{4}{10} \text{ e } y = -\frac{3}{10},$$

temos

$$\mathcal{S} = \left\{ \left(\frac{4}{10}, -\frac{3}{10} \right) \right\}.$$

· A matriz aumentada do sistema (3.4) é 3×5 e dada por

$$\left[\begin{array}{cccc|c} \sqrt{2}/2 & \sqrt{2} & -3\sqrt{2}/2 & 0 & 2\sqrt{2} \\ 2 & 4 & -6 & 0 & 8 \\ 0 & 1 & -1 & 1 & -1 \end{array}\right].$$

Além disso, $\mathcal{S}$ consiste de todas as quádruplas ordenadas (x, y, z, w) que satisfazem, simultaneamente, as três equações do sistema supracitado.

Como podemos determinar essas quádruplas?

Demonstra-se que, para um sistema $m \times n$ arbitrário, ocorre uma das três possibilidades seguintes:

1. $\mathcal{S}$ é *vazio*, isto é, o sistema não admite solução;
2. $\mathcal{S}$ é *unitário*, isto é, o sistema admite uma única solução;
3. $\mathcal{S}$ é *infinito*, isto é, o sistema admite infinitas soluções.

Resolveremos (3.4), ou seja, determinaremos $\mathcal{S}$ para esse sistema, "escalonando" a sua matriz aumentada, até obtermos uma matriz R *adequada. Esse método de resolução, que pode ser utilizado para qualquer sistema linear, será estudado na próxima subseção.*

3.5.1 Matriz escalonada reduzida R e escalonamento

- Uma matriz R é dita *escalonada reduzida* quando as seguintes condições são satisfeitas:

  · A primeira entrada não nula de cada linha não nula de R, chamada de *pivô*, é 1;

  · A partir da segunda linha de R, o pivô de uma linha (não nula) está mais a direita em relação ao pivô da linha anterior;

  · Uma coluna de R que contenha um pivô tem as outras entradas nulas;

  · Possíveis linhas nulas de R estão abaixo das não nulas.

EXEMPLOS

$$\mathsf{R} = \begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix};$$

$$\mathsf{R} = \begin{bmatrix} 1 & 2 & 0 & 3 & 4 & 5 & 0 & 6 & 0 \\ 0 & 0 & 1 & 7 & 8 & 9 & 0 & 10 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 11 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix};$$

$$\mathsf{R} = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

- Sejam $A, B \in \mathbb{R}^{m\times n}$ tais que A é *equivalente (por linhas) à* B, ou seja, B pode ser obtida de A via uma das seguintes *operações elementares sobre as linhas* $A(i,-)$ e $A(j,-)$ fixadas, $i \neq j$:

  · $B(i,-) = A(i,-) + \alpha\, A(j,-)$,[6] com escalar $\alpha \neq 0$;

  · $B(i,-) = \alpha\, A(i,-)$,[7] com escalar $\alpha \neq 0$;

  · $B(i,-) = A(j,-)$, $B(j,-) = A(i,-)$.[8]

  Caso A seja equivalente à B, denota-se

$$A \longrightarrow B.^{9} \tag{3.5}$$

EXEMPLOS

· $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \overbrace{\longrightarrow}^{B(2,-) = A(2,-) + (-4)A(1,-)} B = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 7 & 8 & 9 \end{bmatrix};$

[6] Isto é, a i-ésima linha de B é a soma da i-ésima linha de A e um múltiplo escalar da j-ésima linha de A.
[7] Isto é, a i-ésima linha de B é um múltiplo escalar da i-ésima linha de A.
[8] Isto é, estão trocadas, na matriz B, as linhas i e j de A.
[9] Essa notação pode vir acompanhada de uma das operações elementares supracitadas.

$$\cdot\ A = \begin{bmatrix} 10 & 11 \\ 12 & 13 \end{bmatrix} \underbrace{\overset{B(1,-)=\frac{1}{10}A(1,-)}{}}_{\longrightarrow} B = \begin{bmatrix} 1 & 11/10 \\ 12 & 13 \end{bmatrix};$$

$$\cdot\ A = \begin{bmatrix} 14 & 15 & 16 \\ 17 & 18 & 19 \\ 1 & 2 & 3 \end{bmatrix} \underbrace{\overset{B(1,-)=A(3,-),\ B(3,-)=A(1,-)}{}}_{\longrightarrow} B = \begin{bmatrix} 1 & 2 & 3 \\ 17 & 18 & 19 \\ 14 & 15 & 16 \end{bmatrix}.$$

- Demonstra-se que:

$$A \longrightarrow B \Longrightarrow B \longrightarrow A. \tag{3.6}$$

Nesse caso, A e B são ditas *equivalentes (entre si)*.

- No processo de *escalonamento* de uma matriz $A \in \mathbb{R}^{m\times n}$, *via eliminação de Gauss-Jordan*, obtemos uma sequência finita de matrizes $m \times n$, digamos,

$$A_1, A_2, \ldots, A_r,$$

tal que:

· $A_1 = A$;

· A_{i-1} e A_i são equivalentes, $i = 2, \ldots, r$;

· $A_r = \mathsf{R}$.

Todas as matrizes da sequência supracitada são ditas equivalentes (entre si) e denotadas entre setas horizontais, apontando para à direita. Além disso, caso $r \leq 26$, podemos denotar essas matrizes por letras do nosso alfabeto. Por exemplo, caso o escalonamento termine na décima terceira matriz, denota-se:

$$A \longrightarrow B \longrightarrow C \longrightarrow \cdots \longrightarrow M = \mathsf{R}.$$

EXEMPLO

Seja A, por abuso de notação, a matriz aumentada do sistema (3.4).[10] Portanto,

$$
\begin{aligned}
A &= \left[\begin{array}{cccc|c} \sqrt{2}/2 & \sqrt{2} & -3\sqrt{2}/2 & 0 & 2\sqrt{2} \\ 2 & 4 & -6 & 0 & 8 \\ 0 & 1 & -1 & 1 & -1 \end{array}\right] \underbrace{B(1,-) = \sqrt{2}\,A(1,-)}_{\longrightarrow} \\
B &= \left[\begin{array}{cccc|c} 1 & 2 & -3 & 0 & 4 \\ 2 & 4 & -6 & 0 & 8 \\ 0 & 1 & -1 & 1 & -1 \end{array}\right] \underbrace{C(2,-) = B(2,-) - 2\,B(1,-)}_{\longrightarrow} \\
C &= \left[\begin{array}{cccc|c} 1 & 2 & -3 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 \end{array}\right] \underbrace{D(2,-) = C(3,-),\ D(3,-) = C(2,-)}_{\longrightarrow} \\
D &= \left[\begin{array}{cccc|c} 1 & 2 & -3 & 0 & 4 \\ 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \underbrace{E(1,-) = D(1,-) - 2\,D(2,-)}_{\longrightarrow} \\
E &= \left[\begin{array}{cccc|c} 1 & 0 & -1 & -2 & 6 \\ 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] = \mathsf{R}.
\end{aligned}
$$

Essa R é a matriz aumentada do sistema 2×4

$$
\left\{\begin{array}{ccccccccc} x & & & - & z & - & 2w & = & 6 \\ & & y & - & z & + & w & = & -1 \end{array}\right. . \tag{3.7}
$$

> *Demonstra-se que o escalonamento da matriz aumentada de um sistema $m \times n$ não altera o conjunto solução dele.*

Assim, para o exemplo supracitado, $\mathcal{S}$ é o conjunto solução de (3.4) e (3.7), sendo que, esse último sistema é de fácil resolução. De fato, se $(x, y, z, w) \in \mathcal{S}$, então, para escalares $z = \alpha$ e $w = \beta$ arbitrários,

$$
\begin{aligned}
(x, y, z, w) &= (z + 2w + 6, z - w - 1, z, w) \\
&= (\alpha + 2\beta + 6, \alpha - \beta - 1, \alpha, \beta) \\
&= \alpha(1, 1, 1, 0) + \beta(2, -1, 0, 1) + (6, -1, 0, 0).
\end{aligned} \tag{3.8}
$$

Portanto, $\mathcal{S}$ tem uma infinidade de soluções e, atribuindo um valor real para α e outro para β, (3.8) representa uma dessas infinitas soluções.

- Caso a matriz aumentada $[A|\mathbf{b}]$ seja equivalente à alguma matriz que tenha alguma linha da forma

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

e b seja não nulo, o sistema $A\mathbf{x} = \mathbf{b}$ *não tem solução*. De fato, a linha supracitada representa a equação

$$
0 \cdot x + 0 \cdot y + 0 \cdot z + \text{etc} = b \neq 0,
$$

ou seja,

$$
0 = b \neq 0,
$$

[10] Cf. p. 59.

que não é uma igualdade válida.

EXEMPLO

O sistema $\left\{\begin{array}{rcrcrcr} & x & + & y & - & z & = & 2 \\ - & x & - & 2y & + & 2z & = & -1 \\ & & & y & - & z & = & 0 \end{array}\right.$ não tem solução, pois

$$\left[\begin{array}{rrr|r} 1 & 1 & -1 & 2 \\ -1 & -2 & 2 & -1 \\ 0 & 1 & -1 & 0 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 1 & -1 & 2 \\ 0 & -1 & 1 & 1 \\ 0 & 1 & -1 & 0 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 1 & -1 & 2 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right],$$

cuja terceira linha, ou seja,

$$\left[\begin{array}{rrr|r} 0 & 0 & 0 & 1 \end{array}\right],$$

resulta em $0 = 1$, que não é uma igualdade válida.

- Ao considerarmos o escalonamento de um sistema homogêneo arbitrário, é desnecessário escrevermos a coluna nula $\mathbf{b} = \mathbf{0}$ da matriz aumentada desse sistema, pois as operações elementares não alteram essa coluna.

EXEMPLO

Se

$$\left\{\begin{array}{rcrcrcrcr} \frac{\sqrt{2}}{2}x & + & \sqrt{2}y & - & \frac{3\sqrt{2}}{2}z & & & = & 0 \\ 2x & + & 4y & - & 6z & & & = & 0 \\ & & y & - & z & + & w & = & 0 \end{array}\right. \tag{3.9}$$

e $(x, y, z, w) \in \mathcal{S}$, onde $\mathcal{S}$ é o conjunto solução de (3.9), então, para escalares $z = \alpha$ e $w = \beta$ arbitrários,

$$\begin{aligned} (x, y, z, w) &= (z + 2w, z - w, z, w) \\ &= (\alpha + 2\beta, \alpha - \beta, \alpha, \beta) \\ &= \alpha(1, 1, 1, 0) + \beta(2, -1, 0, 1). \end{aligned} \tag{3.10}$$

Note que, as soluções “gerais” (3.10) e (3.8) dos sistemas (3.9) e (3.4), respectivamente, diferem entre si pela solução “particular” $\mathbf{x}_{\mathrm{P}} = (6, -1, 0, 0)$ de (3.4).

Demonstra-se que, $\mathbf{x}_{\mathrm{NH}}$ é solução “geral” do sistema $A\mathbf{x} = \mathbf{b}$ arbitrário, se, e somente se, $\mathbf{x}_{\mathrm{NH}} = \mathbf{x}_{\mathrm{H}} + \mathbf{x}_{\mathrm{P}}$, onde $\mathbf{x}_{\mathrm{H}}$ é solução “geral’ de $A\mathbf{x} = \mathbf{0}$ e $\mathbf{x}_{\mathrm{P}}$ é uma solução “particular” de $A\mathbf{x} = \mathbf{b}$.

3.6 Matrizes invertíveis, matrizes elementares e escalonamento

Sejam A, B e C matrizes quadradas e de mesma ordem.

- Caso B e C sejam *inversas à direita* e *à esquerda* de A, respectivamente, isto é,

$$AB = \mathsf{I} = CA,$$

temos $B = C$. Além disso, para alguma matriz D tal que $AD = \mathsf{I} = DA$, temos $D = B$.

DEMONSTRAÇÃO

$$\begin{aligned} B &= \mathsf{I}B \\ &= (CA)B \\ &= C(AB) \\ &= C\mathsf{I} \\ &= C. \end{aligned}$$

Além disso,

$$\begin{aligned} D &= D\mathsf{I} \\ &= D(AB) \\ &= (DA)B \\ &= \mathsf{I}B \\ &= B. \end{aligned}$$

Assim, denotamos a matriz B por A^{-1}. Além disso, dizemos que A é *invertível* e A^{-1} é a *inversa* de A.

EXEMPLO

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \Longleftrightarrow A^{-1} = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix}.$$

- Caso A tenha uma linha (respectivamente, coluna) nula, A não é invertível.

DEMONSTRAÇÃO

Para qualquer matriz B (respectivamente, C), AB (respectivamente, CA) tem uma linha (respectivamente, coluna) nula. Portanto, $AB \neq \mathsf{I}$ (respectivamente, $CA \neq \mathsf{I}$).

- Se $\det A \neq 0$, então

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

é invertível e

$$A^{-1} = \begin{bmatrix} \frac{d}{\det A} & \frac{-b}{\det A} \\ \frac{-c}{\det A} & \frac{a}{\det A} \end{bmatrix}$$

é a sua inversa.

DEMONSTRAÇÃO

Basta verificarmos que $AA^{-1} = \mathsf{I}$.[11] Assim,

$$\begin{aligned} AA^{-1} &= \begin{bmatrix} \frac{ad-bc}{\det A} & \frac{-ab+ba}{\det A} \\ \frac{cd-dc}{\det A} & \frac{-cb+da}{\det A} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad (ad - bc = \det A \neq 0) \\ &= \mathsf{I}. \end{aligned}$$

[11] A verificação de $A^{-1}A = \mathsf{I}$ é análoga.

EXEMPLO

Ao calcularmos o produto das matrizes do primeiro exemplo dessa seção, temos $AA^{-1} = \mathsf{I}$.

- Caso A seja invertível, temos as seguintes propriedades:

    · A^{-1} é invertível e $\left(A^{-1}\right)^{-1} = A$;[12]

    · A^t é invertível e $\left(A^t\right)^{-1} = \left(A^{-1}\right)^t$.[13]

EXEMPLO

As propriedades supracitadas são válidas para as matrizes do primeiro exemplo dessa seção.

- Caso A e B sejam invertíveis, AB é invertível e $(AB)^{-1} = B^{-1}A^{-1}$.[14]

EXERCÍCIO

Verifique essa propriedade para

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \text{ e } B = \begin{bmatrix} -3/2 & -1 \\ -1/2 & 0 \end{bmatrix}.$$

3.6.1 Matrizes elementares

Podemos interpretar o processo de escalonamento como um produto de matrizes elementares e calcular a inversa de qualquer matriz (invertível) como um produto dessas matrizes.

Uma matriz E é dita *elememtar* caso seja equivalente à alguma I pela aplicação de *uma* operação elementar, ou seja,

$$\mathsf{I} \longrightarrow \mathsf{E}.^{15} \tag{3.11}$$

EXERCÍCIO

Apresente todas as matrizes elementares 2×2.

RESOLUÇÃO

Se $\mathsf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ e $c \in \mathbb{R}$ é não nulo, então E pode ser dada de uma das seguintes formas:

1. $\begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix}$ ou $\begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix}$;

2. $\begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix}$ ou $\begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix}$;

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

---

[12] Demonstre!

[13] Idem.

[14] Idem.

[15] Cf. 3.5 e 3.6, pp. 61 e 62.

OBSERVAÇÃO

Sejam B e E matrizes obtidas ao aplicarmos a mesma operação elementar em $A \in \mathbb{R}^{m\times n}$ e $\mathsf{I} \in \mathbb{R}^{m\times m}$, respectivamente. Nesse caso, $B = \mathsf{E}A$, ou seja,

$$A \longrightarrow B = \mathsf{E}A, \tag{3.12}$$

onde "$\longrightarrow$" representa a mesma operação elementar dada em (3.11).

EXEMPLO

$$\begin{aligned} A = \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix} \longrightarrow B &= \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix} \\ &= \mathsf{E}A, \end{aligned}$$

com B e E obtidas de $A = \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix}$ e $\mathsf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, respectivamente, pela aplicação da mesma operação elementar, ou seja,

$$B(1,-) = A(1,-) - A(2,-) \quad \text{e} \quad \mathsf{E}(1,-) = \mathsf{I}(1,-) - \mathsf{I}(2,-).$$

- Demonstra-se que *toda matriz elementar é invertível*.

  EXERCÍCIO

  Verifique a invertibilidade de todas as matrizes elementares 2×2 apresentadas no primeiro exercício dessa subseção.[16]

- Como o produto de matrizes invertíveis (de mesma ordem) é invertível, *o produto de matrizes elementares (de mesma ordem) é invertível*.

  EXERCÍCIO

  Multiplique as matrizes elementares 2×2 apresentadas no primeiro exercício dessa subseção, em qualquer ordem, e verifique que o produto delas é invertível.

Na seção 3.7, veremos como escalonamento e invertibilidade estão relacionados pela expressão (3.12).

[16]Esse resultado é válido para matrizes elementares $n \times n$ arbitrárias, conforme demonstrado nas referências sugeridas no primeiro capítulo desse livro.

3.7 Exercícios

1. Considere

$$A = \begin{bmatrix} 1 & -1 & 0 & 2 & \sqrt{2} \\ \pi & \pi/2 & 1 & -2 & 0 \\ 0 & -1 & 1/\sqrt{2} & 1/2 & -\pi \end{bmatrix}$$

e $B = A^t$. Verifique que $C = AB$ e $D = BA$ são simétricas, isto é, $C^t = C$ e $D^t = D$.

DICA

Como A é 3×5 e

$$B = \begin{bmatrix} 1 & \pi & 0 \\ -1 & \pi/2 & -1 \\ 0 & 1 & 1/\sqrt{2} \\ 2 & -2 & 1/2 \\ \sqrt{2} & 0 & -\pi \end{bmatrix}$$

é 5×3, segue que C é 3×3 e D é 5×5. Além disso,

$$\begin{aligned} c_{12} &= A(1,-) \cdot B(-,2) \\ &= 1 \cdot \pi + (-1) \cdot (\pi/2) + 0 \cdot 1 + 2 \cdot (-2) + \sqrt{2} \cdot 0 \\ &= \frac{\pi}{2} - 4 \\ &= \frac{\pi - 8}{2} \end{aligned}$$

e todas as outras entradas (de C e D) são calculadas de modo análogo.

2. Sejam

$$A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 2 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & -1 \end{bmatrix} \text{ e } C = \begin{bmatrix} -2 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

em $\mathbb{R}^{2\times 3}$. Determine:

(a) a CL $A + 2B - C$;

(b) a CL $\alpha A + \beta B - C$ cuja primeira coluna é nula;[17]

(c) se A e B são LI.[18]

3. Demonstre as propriedades $(*)$ da seção 3.2.

RESOLUÇÃO

· Para $M \in \mathbb{R}^{m\times p}$ e $N \in \mathbb{R}^{p\times n}$ arbitrárias, a entrada da linha i e coluna j do produto $MN \in \mathbb{R}^{m\times n}$ é dada por $(MN)_{ij} = M(i,-) \cdot N(-,j)$. Portanto, caso A, B e C sejam matrizes onde o produto

[17] SUGESTÃO

Obtenha escalares α e β tais que $\alpha \mathbf{a}_1 + \beta \mathbf{b}_1 - \mathbf{c}_1$ seja o vetor nulo.

[18] RESOLUÇÃO

Suponha A, B LD. Portanto, existe escalar λ tal que $A = \lambda B$. Logo, ao examinarmos a última entrada da última linha de cada matriz, temos $0 = \lambda \cdot (-1)$, ou seja, $\lambda = 0$. Assim, $A = O$, que é uma igualdade inválida, pois A, dada no enunciado desse exercício, não é nula. Então, a suposição inicial, sobre a dependência linear das matrizes, é falsa.

$(A+B)C$ esteja bem definido, temos

$$\begin{aligned}[(A+B)C]_{ij} &= (A+B)(i,-)\cdot C(-,j)\\ &= (A(i,-)+B(i,-))\cdot C(-,j)\\ &= A(i,-)\cdot C(-,j)+B(i,-)\cdot C(-,j)\\ &= (AC)_{ij}+(BC)_{ij}\\ &= (AC+BC)_{ij}.\end{aligned}$$

Como i e j representam índices arbitrários, $(A+B)C = AC+BC$.

> *Na terceira igualdade anterior, de cima para baixo, utilizamos a distributividade da adição de vetores, em relação ao produto interno.*

· Para qualquer matriz M, a entrada da linha i e coluna j de M^t é dada por $\left(M^t\right)_{ij} = M_{ji}$. Portanto, caso A e B sejam matrizes com produto AB bem definido, temos

$$\begin{aligned}\left[(AB)^t\right]_{ij} &= (AB)_{ji}\\ &= A(j,-)\cdot B(-,i)\\ &= B(-,i)\cdot A(j,-)\\ &= \left(B^t\right)(i,-)\cdot\left(A^t\right)(-,j)\\ &= \left(B^tA^t\right)_{ij}.\end{aligned}$$

Como i e j representam índices arbitrários, $(AB)^t = B^tA^t$.

> *Na terceira igualdade anterior, de cima para baixo, utilizamos a comutatividade do produto interno.*

4. Sejam $\mathbf{a}_1 = A(-,1)$, $\mathbf{a}_2 = A(-,2)$ e $\mathbf{a}_3 = A(-,3)$ os *vetores colunas* e $\mathbf{b}_1 = A(1,-)^t$, $\mathbf{b}_2 = A(2,-)^t$ e $\mathbf{b}_3 = A(3,-)^t$ as transpostas dos *vetores linhas* de

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

(a) Verifique que $\mathbf{a}_3 = 2\mathbf{a}_2 - \mathbf{a}_1$ e $\mathbf{b}_3 = 2\mathbf{b}_2 - \mathbf{b}_1$;

(b) Escreva $\mathbf{a}_1$ e $\mathbf{a}_2$ como combinações lineares de $\mathbf{b}_1$ e $\mathbf{b}_2$;[19]

(c) Escreva $\mathbf{b}_1$ e $\mathbf{b}_2$ como combinações lineares de $\mathbf{a}_1$ e $\mathbf{a}_2$;[20]

(d) A partir dos outros itens desse exercício, conclua que, tanto os vetores linhas de A, quanto seus vetores colunas, geram o mesmo plano de $\mathbb{R}^3$.

RESOLUÇÃO DO ITEM (d)

Denote por $\mathcal{S}_\ell$ o subespaço de $\mathbb{R}^3$ gerado pelos vetores $\mathbf{b}_1$, $\mathbf{b}_2$ e $\mathbf{b}_3$. Denote por $\mathcal{S}_c$ o subespaço de $\mathbb{R}^3$ gerado pelos vetores $\mathbf{a}_1$, $\mathbf{a}_2$ e $\mathbf{a}_3$. Note que, tanto $\mathbf{b}_1$ e $\mathbf{b}_2$, como $\mathbf{a}_1$ e $\mathbf{a}_2$, são não colineares, isto é, são LI. Portanto, $\{\mathbf{b}_1,\mathbf{b}_2\}$ e $\{\mathbf{a}_1,\mathbf{a}_2\}$ são bases de $\mathcal{S}_\ell$ e $\mathcal{S}_c$, respectivamente, pois, pelo item (a), podemos desconsiderar os geradores $\mathbf{b}_3$ e $\mathbf{a}_3$. Além disso, pelo item (b), cada vetor de $\{\mathbf{a}_1,\mathbf{a}_2\}$ pertence à $\mathcal{S}_\ell$. Assim, todos os vetores de $\mathcal{S}_c$ pertencem à $\mathcal{S}_\ell$, ou seja, $\mathcal{S}_c \subset \mathcal{S}_\ell$. Analogamente, pelo item (c), $\mathcal{S}_\ell \subset \mathcal{S}_c$. Então, $\mathcal{S}_\ell = \mathcal{S}_c$ tem dimensão dois.

[19] SOLUÇÃO $\mathbf{a}_1 = \frac{11}{3}\mathbf{b}_1 - \frac{2}{3}\mathbf{b}_2$ e $\mathbf{a}_2 = \frac{10}{3}\mathbf{b}_1 - \frac{1}{3}\mathbf{b}_2$.

[20] SOLUÇÃO $\mathbf{b}_1 = -\frac{1}{3}\mathbf{a}_1 + \frac{2}{3}\mathbf{a}_2$ e $\mathbf{b}_2 = -\frac{10}{3}\mathbf{a}_1 + \frac{11}{3}\mathbf{a}_2$.

5. Verifique que os vetores colunas de

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

geram um subespaço (de dimensão 2) de $\mathbb{R}^3$ diferente daquele gerado por seus vetores linhas.

RESOLUÇÃO

Seja $\mathcal{S}_\ell$ o subespaço de $\mathbb{R}^3$ gerado por $\mathbf{v}_1 = (1,1,0)$ e $\mathbf{v}_2 = (0,1,1)$.[21] Seja $\mathcal{S}_c$ o subespaço de $\mathbb{R}^3$ gerado por $\mathbf{e}_1 = (1,0,0)$ e $\mathbf{e}_2 = (0,1,0)$.[22] Como bases distintas podem gerar o mesmo subespaço, não podemos afirmar que:

$$\{\mathbf{v}_1, \mathbf{v}_2\} \neq \{\mathbf{e}_1, \mathbf{e}_2\} \Longrightarrow \mathcal{S}_\ell \neq \mathcal{S}_c.$$ [23]

Para verificarmos que os subespaços supracitados são distintos, basta apresentarmos algum vetor de um deles que não seja uma CL dos vetores da base do outro. Por exemplo, como $\mathbf{v}_2 \in \mathcal{S}_\ell$ não pode ser escrito como CL dos vetores $\mathbf{e}_1$ e $\mathbf{e}_2$, $\mathbf{v}_2 \notin \mathcal{S}_c$.

6. Seja $\mathcal{M} = \mathbb{R}^{2\times 2}$. A dimensão de $\mathcal{M}$ é 4. De fato, uma base de $\mathcal{M}$, dita *canônica*, é constituída pelas matrizes

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},\ B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},\ C = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \text{ e } D = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

pois toda matriz $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathcal{M}$ é uma CL das matrizes A, B, C e D, ou seja,

$$aA + bB + cC + dD = M,$$

e M é a matriz nula apenas quando $a = b = c = d = 0$, confirmando que A, B, C e D são LI.

(a) Seja $\mathcal{T}_S$ o subconjunto de $\mathcal{M}$ das matrizes triangulares superiores.[24] Verifique que $\mathcal{T}_S$ é um subespaço de dimensão 3 de $\mathcal{M}$, gerado por A, B e D;

(b) Seja $\mathcal{D}$ o subconjunto de $\mathcal{T}_S$ das matrizes diagonais. Verifique que $\mathcal{D}$ é um subespaço de dimensão 2 de $\mathcal{T}_S$, gerado por A e D;

(c) Seja $\mathcal{I}$ o subconjunto de $\mathcal{D}$ das matrizes que são múltiplas da matriz identidade I. Verifique que $\mathcal{I}$ é um subespaço de dimensão 1 de $\mathcal{D}$, gerado por I;

(d) Descreva um subespaço de $\mathcal{M}$ que contém A mas não $-D$.

(e) Caso um subespaço $\mathcal{S}$ de $\mathcal{M}$ contenha A e $-D$, $\mathsf{I} \in \mathcal{S}$?

(f) Como vimos, uma matriz quadrada é dita simétrica quando é igual a sua transposta. Seja $\mathcal{S}$ o subconjunto de $\mathcal{M}$ das matrizes simétricas. Verifique que $\mathcal{S}$ é um subespaço de dimensão 3 de $\mathcal{M}$, gerado por A, D e $B + C$.

---

[21] Note que, embora o vetor nulo represente a terceira linha de A, não podemos incluí-lo na base $\{\mathbf{v}_1, \mathbf{v}_2\}$ de $\mathcal{S}_\ell$.

[22] Note que, embora $\mathbf{v}_1 = (1,1,0)$ represente a segunda coluna de A, como $\mathbf{v}_1 = \mathbf{e}_1 + \mathbf{e}_2$, podemos considerar a base de $\mathcal{S}_c$ composta apenas pelos dois primeiros vetores da base canônica de $\mathbb{R}^3$.

[23] Cf. o exercício anterior.

[24] Cf. o exercício 8 dessa seção.

7. Calcule os seguintes determinantes:

   (a) $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$; RESPOSTA -2;

   (b) $\begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}$; RESPOSTA 3;

   (c) $\begin{bmatrix} 0 & 3 & 1 & 28 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 2 & 71 \\ 2 & 36 & -9 & 3 \end{bmatrix}$; RESPOSTA 36;

   (d) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 5 & 2 & 0 & 0 \\ 8 & 6 & 3 & 0 \\ 0 & 9 & 7 & 4 \end{bmatrix}$; RESPOSTA 24.

8. Caso $A \in \mathbb{R}^{n\times n}$ seja *triangular*, ou seja, as entradas acima ou abaixo da diagonal principal de A sejam nulas, demonstre que $\det A = a_{11} \cdots a_{nn}$.[25]

9. Verifique as regras $\det A^t = \det A$ e $\det(AB) = (\det A)(\det B)$ para

$$A = \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 1 & 1 \\ 5 & -2 \end{bmatrix}.$$

10. A matriz

$$A = \begin{bmatrix} \frac{e^x+e^{-x}}{2} & \frac{e^x-e^{-x}}{2} \\ \frac{e^x-e^{-x}}{2} & \frac{e^x+e^{-x}}{2} \end{bmatrix}$$

é invertível para todo $x \in \mathbb{R}$. Essa afirmação é verdadeira ou falsa? Justifique (corretamente) a sua resposta.

RESOLUÇÃO

Existe A^{-1} pois $\det A \neq 0$. De fato,

$$\begin{aligned} \det A &= \frac{1}{4}\left[e^{2x} + 2 + e^{-2x} - \left(e^{2x} - 2 + e^{-2x}\right)\right] \\ &= 1. \end{aligned}$$

11. Sejam

$$A = \begin{bmatrix} 1 & 1 & 0 & -1 \\ -1 & -1 & 3 & 4 \\ 2 & 2 & 3 & 1 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 0 & 0 & 3 & 3 \\ r & r & 3 & 1 \\ 1 & 1 & 6 & 5 \end{bmatrix}$$

---

[25] SUGESTÃO
Caso as entradas acima da diagonal principal sejam nulas, calcule determinantes, indutivamente, sempre ao longo das primeiras linhas, para obter

$$\det A = a_{11}a_{22} \cdots a_{nn}.$$

Caso as entradas abaixo da diagonal principal sejam nulas, considere A^t no lugar de A, obtenha

$$\det A^t = a_{11} \cdots a_{nn}$$

pelo caso anterior e utilize uma das regras do próximo exercício dessa seção.

equivalentes por linha, isto é, essas matrizes são equivalentes por linha a mesma matriz escalonada reduzida R. Determine r.

RESOLUÇÃO

$r = 2$ pois

$$A \longrightarrow \begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \mathsf{R}$$

e

$$\begin{aligned} B &\longrightarrow \begin{bmatrix} 1 & 1 & 6 & 5 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 3-6r & 1-5r \end{bmatrix} \\ &\longrightarrow \begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & r-2 \end{bmatrix} = \mathsf{R}. \end{aligned}$$

12. Via escalonamento, determine o conjunto solução, $\mathcal{S}$, de cada um dos sistemas seguintes:

(a) $\left\{\begin{array}{rcrcrcr} 2x & + & y & - & 2z & = & 10 \\ 3x & + & 2y & + & 2z & = & 1 \\ 5x & + & 4y & + & 3z & = & 4 \end{array}\right.$ RESPOSTA $\mathcal{S} = \{(1,2,-3)\}$;

(b) $\left\{\begin{array}{rcrcrcr} x & + & 2y & - & z & = & 0 \\ 2x & - & y & + & 3z & = & 0 \\ 4x & + & 3y & + & z & = & 0 \end{array}\right.$ RESPOSTA $\mathcal{S} = \{\alpha(-1,1,1) \mid \alpha \in \mathbb{R}\}$;

(c) $\left\{\begin{array}{rcrcrcr} x & + & y & + & z & = & 4 \\ 2x & + & 5y & - & 2z & = & 3 \\ x & + & 7y & - & 7z & = & 5 \end{array}\right.$ RESPOSTA $\mathcal{S} = \emptyset$;

(d) $\left\{\begin{array}{rrcrcrcrcr} & x & - & y & + & 2z & - & w & = & -1 \\ & 2x & + & y & - & 2z & - & 2w & = & -2 \\ - & x & + & 2y & - & 4z & + & w & = & 1 \\ & 3x & & & & & - & 3w & = & -3 \end{array}\right.$

RESPOSTA $\mathcal{S} = \{\alpha(1,0,0,1) + \beta(0,2,1,0) + (-1,0,0,0) \mid \alpha, \beta \in \mathbb{R}\}$.

13. Caso

$$\left\{\begin{array}{rrcrcrcr} & x_1 & + & 2x_2 & + & x_3 & = & a \\ - & x_1 & - & x_2 & - & 2x_3 & = & b \\ & & - & x_2 & + & x_3 & = & c \end{array}\right. \tag{3.13}$$

represente um sistema linear com três equações e x_1, x_2 e x_3 sejam suas variáveis, determine todos os valores de a, b e c para os quais esse sistema tenha solução. Nesse caso, determine todas as soluções de (3.13).

RESOLUÇÃO

Ao escalonarmos a matriz aumentada de (3.13), temos

$$\begin{aligned} \left[\begin{array}{rrr|c} 1 & 2 & 1 & a \\ -1 & -1 & -2 & b \\ 0 & -1 & 1 & c \end{array}\right] &\longrightarrow \left[\begin{array}{rrr|c} 1 & 2 & 1 & a \\ 0 & 1 & -1 & a+b \\ 0 & -1 & 1 & c \end{array}\right] \\ &\longrightarrow \left[\begin{array}{rrr|c} 1 & 0 & 3 & -a-2b \\ 0 & 1 & -1 & a+b \\ 0 & 0 & 0 & a+b+c \end{array}\right]. \end{aligned}$$

Assim, (3.13) tem solução se, e somente se, $a+b+c=0$ e, nesse caso, (3.13) é equivalente ao sistema

$$\begin{cases} x_1 \quad\quad + 3x_3 = -a-2b; \\ \quad\quad x_2 - x_3 = a+b. \end{cases}$$

Portanto, para $x_3 = t$ arbitrário, a solução de (3.13) é dada por

$$\begin{aligned} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= \begin{bmatrix} -3t-a-2b \\ t+a+b \\ t \end{bmatrix} \\ &= t\begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -a-2b \\ a+b \\ 0 \end{bmatrix}, \end{aligned}$$

com a, b e c constantes e tais que $a+b+c=0$.

14. Determine os valores de a, b e c para os quais o sistema linear com matriz aumentada

$$\left[\begin{array}{cccc|c} 2 & 6 & 0 & 4 & 2a \\ 0 & 0 & 1/2 & 2 & b/2 \\ -1 & -3 & 1 & 2 & c-2a \end{array}\right]$$

tenha solução.[26] Além disso, obtenha a solução do sistema supracitado.[27]

15. Demonstra-se que

> R *é a escalonada reduzida equivalente à* A *se, e somente se,*
>
> $$\mathsf{R} = \mathsf{E}_k \cdots \mathsf{E}_2\mathsf{E}_1 A,$$
>
> *onde* $\mathsf{E}_1, \mathsf{E}_2, \ldots, \mathsf{E}_k$ *são as matrizes elementares obtidas, nessa ordem, nas* k *etapas do escalonamento de* A*, como em (3.12) da página 67.*

Nesse caso, se A é quadrada e $\mathsf{R} = \mathsf{I}$, então A é invertível e

$$A^{-1} = \mathsf{E}_k \cdots \mathsf{E}_2\mathsf{E}_1.$$

EXEMPLO

$$\begin{aligned} A = \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix} &\longrightarrow \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix} = \mathsf{E}_1 A \\ &\longrightarrow \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} = \mathsf{E}_2\mathsf{E}_1 A \\ &\longrightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \mathsf{E}_3\mathsf{E}_2\mathsf{E}_1 A = I; \end{aligned}$$

$$\begin{aligned} \mathsf{E}_3\mathsf{E}_2\mathsf{E}_1 &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} = A^{-1}. \end{aligned}$$

[26] RESPOSTA $a+b=c$.

[27] DICA Cf. a resolução do sistema (3.13) do exercício anterior.

Portanto, para matrizes $n \times n$ arbitrárias, podemos obter A^{-1}(caso A seja invertível) ou afirmar que A^{-1} não existe (caso A não seja invertível) pelo método seguinte:

(a) Escalone $[A|\mathsf{I}]$ até obter $[\mathsf{R}|A']$ após, digamos, k operações elementares, ou seja, efetue o escalonamento

$$[A|\mathsf{I}] \longrightarrow [\mathsf{E}_1A|\mathsf{E}_1\mathsf{I}] \longrightarrow [\mathsf{E}_2\mathsf{E}_1A|\mathsf{E}_2\mathsf{E}_1] \longrightarrow \cdots \longrightarrow [\mathsf{E}_k\cdots\mathsf{E}_1A|\mathsf{E}_k\cdots\mathsf{E}_1] = \left[\mathsf{R}|A'\right]. \qquad (3.14)$$

(b) Para a escalonada reduzida R, obtida de A, temos:

- $\mathsf{R} = \mathsf{I} \Longleftrightarrow A' = A^{-1}$;
- $\mathsf{R} \neq \mathsf{I} \Longleftrightarrow A^{-1}$ não existe.

Utilize o método supracitado nas seguintes matrizes:

$$\begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 2 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix} \quad \text{e}$$

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

Entre as matrizes supracitadas, existe alguma não invertível? Alguma delas é invertível? Caso alguma seja invertível, apresente a sua inversa.[28]

OBSERVAÇÃO

Para as matrizes desse exercício, o objetivo é obtermos suas inversas (caso existam) pelo método supracitado. Contudo, é possível utilizarmos a técnica da "multiplicação de matrizes em blocos" na matriz 6×6 supracitada, ou seja, caso A seja aquela matriz,

$$B = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}, C = \begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix} \text{ e } D = \begin{bmatrix} 5 & 6 \\ 0 & \frac{1}{5} \end{bmatrix},$$

podemos considerar que A é "diagonal por blocos" e as matrizes B, C e D estão representadas ao longo dessa diagonal, isto é,

$$A = \begin{bmatrix} B & & \text{zeros} \\ & C & \\ \text{zeros} & & D \end{bmatrix}.$$

Por outro lado, ao utilizarmos a fórmula para inversas de matrizes 2×2,[29] temos

$$B^{-1} = \begin{bmatrix} 0 & \frac{1}{2} \\ 1 & 0 \end{bmatrix}, C^{-1} = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{4} \end{bmatrix} \text{ e } D^{-1} = \begin{bmatrix} \frac{1}{5} & -6 \\ 0 & 5 \end{bmatrix}.$$

[28] Esse teste de invertibilidade pode ser aplicado mesmo que não seja necessário apresentarmos A^{-1} (caso A seja invertível). De fato, após obtermos a escalonada reduzida R de A, basta observarmos que:

$$A \text{ é invertível} \Longleftrightarrow \mathsf{R} = \mathsf{I}.$$

[29] Cf. a seção 3.6.

Portanto,

$$A^{-1} = \begin{bmatrix} B^{-1} & & \text{zeros} \\ & C^{-1} & \\ \text{zeros} & & D^{-1} \end{bmatrix}$$

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

pois $AA^{-1} = \mathsf{I}$.

16. Determine todos os valores de a para os quais

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

seja não invertível? Justifique (corretamente) a sua resposta.

RESOLUÇÃO

Pelo método supracitado, descrito no exercício anterior, A é invertível se, e somente se, a matriz escalonada reduzida R, equivalente à A, é dada por $\mathsf{R} = \mathsf{I}$. Nesse caso, podemos considerar o seguinte escalonamento:

$$\begin{aligned} A &\longrightarrow \begin{bmatrix} 1 & 2/a & 3/a \\ a & a & 4 \\ a & a & a \end{bmatrix} \\ &\longrightarrow \begin{bmatrix} 1 & 2/a & 3/a \\ 0 & a-2 & 1 \\ 0 & a-2 & a-3 \end{bmatrix} \\ &\longrightarrow \begin{bmatrix} 1 & 2/a & 3/a \\ 0 & 1 & 1/(a-2) \\ 0 & a-2 & a-3 \end{bmatrix} \\ &\longrightarrow \begin{bmatrix} 1 & 0 & (3a-8)/[a(a-2)] \\ 0 & 1 & 1/(a-2) \\ 0 & 0 & a-4 \end{bmatrix} \\ &\longrightarrow \begin{bmatrix} 1 & 0 & (3a-8)/[a(a-2)] \\ 0 & 1 & 1/(a-2) \\ 0 & 0 & 1 \end{bmatrix} \\ &\longrightarrow \mathsf{I}, \end{aligned}$$

onde:

- $a \neq 0$ na primeira etapa, pois, caso contrário, teríamos $A = \begin{bmatrix} 0 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}$ não equivalente à I;
- $a \neq 2$ na terceira etapa, pois, caso contrário, teríamos a matriz $\begin{bmatrix} 1 & 1 & 3/2 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix}$, obtida na segunda etapa, não equivalente à I;

· $a \neq 4$ na quinta etapa, pois, caso contrário, teríamos $\mathsf{R} = \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & 1/2 \\ 0 & 0 & 0 \end{bmatrix}$, obtida na quarta etapa, diferente de I.

Portanto,

$$A \text{ é não invertível } \iff a \in \{0, 2, 4\}.$$

17. Determine todos os valores de a para os quais

$$A = \begin{bmatrix} 2 & a & a \\ a & a & a \\ 8 & 7 & a \end{bmatrix}$$

seja não invertível.[30]

18. Prove que

$$A = \begin{bmatrix} a & b & b \\ a & a & b \\ a & a & a \end{bmatrix}$$

é invertível se $a \neq 0$ e $a \neq b$. Neste caso, determine A^{-1}.

RESOLUÇÃO

Podemos escalonar $[A|\mathsf{I}]$, via Gauss-Jordan, até obtermos $[\mathsf{I}|A^{-1}]$, com

$$A^{-1} = \frac{1}{a(a-b)} \begin{bmatrix} a & 0 & -b \\ -a & a & 0 \\ 0 & -a & a \end{bmatrix}.$$

De fato,

$$\begin{aligned}
\left[\begin{array}{ccc|ccc} a & b & b & 1 & 0 & 0 \\ a & a & b & 0 & 1 & 0 \\ a & a & a & 0 & 0 & 1 \end{array}\right] &\longrightarrow \left[\begin{array}{ccc|ccc} a & b & b & 1 & 0 & 0 \\ 0 & a-b & 0 & -1 & 1 & 0 \\ 0 & a-b & a-b & -1 & 0 & 1 \end{array}\right] \\
&\longrightarrow \left[\begin{array}{ccc|ccc} 1 & b/a & b/a & 1/a & 0 & 0 \\ 0 & 1 & 0 & -1/(a-b) & 1/(a-b) & 0 \\ 0 & 1 & 1 & -1/(a-b) & 0 & 1/(a-b) \end{array}\right] \\
&\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & b/a & 1/(a-b) & -b/[a(a-b)] & 0 \\ 0 & 1 & 0 & -1/(a-b) & 1/(a-b) & 0 \\ 0 & 0 & 1 & 0 & -1/(a-b) & 1/(a-b) \end{array}\right] \\
&\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/(a-b) & 0 & -b/[a(a-b)] \\ 0 & 1 & 0 & -1/(a-b) & 1/(a-b) & 0 \\ 0 & 0 & 1 & 0 & -1/(a-b) & 1/(a-b) \end{array}\right],
\end{aligned}$$

onde, na segunda etapa do escalonamento, consideramos $a \neq 0$ e $a - b \neq 0$.

19. A inversa de qualquer matriz invertível A tal que

$$A^2 - 2A + 5\mathsf{I} = \mathsf{O}$$

é dada por

$$\frac{1}{5}(2\mathsf{I} - A),$$

[30] RESPOSTA $a = 0$; $a = 2$; $a = 7$.

onde $A^2 = AA$. Essa afirmação é verdadeira ou falsa? Justifique (corretamente) a sua resposta.

RESOLUÇÃO

Ao multiplicarmos A^{-1} por cada um dos membros da primeira equação do enunciado desse exercício, temos

$$\left(A^{-1}A\right)A - 2A^{-1}A + 5A^{-1}\mathsf{I} = A^{-1}\mathsf{O},$$

isto é,

$$A - 2\mathsf{I} + 5A^{-1} = \mathsf{O}.$$

Ao adicionarmos $2\mathsf{I} - A$ a cada um dos membros da equação anterior, temos

$$5A^{-1} = 2\mathsf{I} - A.$$

Assim, basta multiplicarmos $1/5$ por cada um dos membros da equação anterior para que a afirmação supracitada seja verdadeira.

20. Demonstra-se que, ao aplicarmos operações elementares sobre as linhas i e j, $i \neq j$, de $A \in \mathbb{R}^{n\times n}$, obtemos $B \in \mathbb{R}^{n\times n}$ tal que:

   · $B(i,-) = A(i,-) + \alpha\, A(j,-) \Longrightarrow \det B = \det A$;
   · $B(i,-) = \alpha\, A(i,-)$, com $\alpha \neq 0 \Longrightarrow \det B = \alpha\, \det A$;
   · $B(i,-) = A(j,-)$ e $B(j,-) = A(i,-) \Longrightarrow \det B = -\det A$.
   · Como $\det A = \det A^t$, os três itens anteriores (desse exercício) permanecem válidos para operações elementares sobre as colunas de A.

   Em cada um dos itens seguintes, utilize esses resultados para calcular $\det A$.

   (a) $A = \begin{bmatrix} 1 & 0 & 0 & 3 \\ 2 & 7 & 0 & 6 \\ 0 & 6 & 3 & 0 \\ 7 & 3 & 1 & -5 \end{bmatrix}$;

   (b) $A = \begin{bmatrix} 1 & -1 & 2 & -3 \\ -1 & 2 & 1 & -2 \\ 0 & -1 & -3 & 5 \\ 0 & 15 & 11 & -1 \end{bmatrix}$;

   (c) $A = \begin{bmatrix} 1 & -1 & 2 & 1 \\ -2 & 1 & 1 & 3 \\ 1 & 1 & 1 & -1 \\ 1 & 3 & 2 & 1 \end{bmatrix}$;

   (d) $A = \begin{bmatrix} 2 & -8 & 6 & 8 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{bmatrix}$;

   (e) $A = \begin{bmatrix} \pi & e & \sqrt{2} & \sqrt{3} \\ -2/\sqrt{2} & -1/\sqrt{2} & \sqrt{2} & \sqrt{2}/2 \\ 1 & 1/2 & -1 & -1/2 \\ \ln 2 & 1/\pi & \sqrt{3} & \pi^2 \end{bmatrix}$;

(f) $A = \begin{bmatrix} 2 & 3 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 2 & 1 & 0 \end{bmatrix}$;

(g) $A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & a & b \\ 0 & 0 & 0 & 0 & c & d \\ 0 & 0 & 0 & 0 & e & f \\ 0 & p & q & r & s & t \\ 0 & v & w & x & y & z \end{bmatrix}$,

onde as letras minúsculas representam números reais não nulos.

RESOLUÇÕES

(a) Via $B(-,4) = A(-,4) - 3A(-,1)$, podemos obter uma matriz triangular cujo determinante é o produto das entradas da diagonal principal de B. De fato,

$$\begin{aligned} \det A &= \det B \\ &= \det \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 7 & 0 & 0 \\ 0 & 6 & 3 & 0 \\ 7 & 3 & 1 & -26 \end{bmatrix} \\ &= 1 \cdot 7 \cdot 3 \cdot (-26) \\ &= -546. \end{aligned}$$

(b)

$$\begin{aligned} \det A &= \det B \\ &= \det \begin{bmatrix} 1 & -1 & 2 & -3 \\ 0 & 1 & 3 & -5 \\ 0 & -1 & -3 & 5 \\ 0 & 15 & 11 & -1 \end{bmatrix} \\ &= \det \begin{bmatrix} 1 & -1 & 2 & -3 \\ 0 & 1 & 3 & -5 \\ 0 & 0 & 0 & 0 \\ 0 & 15 & 11 & -1 \end{bmatrix} \\ &= 0, \end{aligned}$$

onde B foi obtida de A, via $B(2,-) = A(2,-) + A(1,-)$, e o determinante foi calculado ao longo da linha nula de C, obtida de B, via $C(3,-) = B(3,-) + B(2,-)$.

(c) Utilize operações elementares para "zerar" as entradas da primeira coluna de A, abaixo da entrada $a_{11} = 1$. O determinante da matriz, assim obtida, pode ser calculado ao longo de sua primeira coluna. De fato,

$$\begin{aligned}
\det A &= \det \begin{bmatrix} 1 & -1 & 2 & 1 \\ 0 & -1 & 5 & 5 \\ 0 & 2 & -1 & -2 \\ 0 & 4 & 0 & 0 \end{bmatrix} \\
&= \det \begin{bmatrix} -1 & 5 & 5 \\ 2 & -1 & -2 \\ 4 & 0 & 0 \end{bmatrix} \\
&= 4 \det \begin{bmatrix} 5 & 5 \\ -1 & -2 \end{bmatrix} \\
&= 4(-10+5) \\
&= -20,
\end{aligned}$$

onde o determinante da matriz 3×3, na terceira igualdade, de cima para baixo, foi calculado ao longo de sua última linha.

(d)

$$\begin{aligned}
\det A &= 2 \det \begin{bmatrix} 1 & -4 & 3 & 4 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{bmatrix} \\
&= 2 \det \begin{bmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & -12 & 10 & 10 \\ 0 & 0 & -3 & 2 \end{bmatrix} \\
&= 2 \det \begin{bmatrix} 3 & -4 & -2 \\ -12 & 10 & 10 \\ 0 & -3 & 2 \end{bmatrix} \\
&= 2 \det \begin{bmatrix} 3 & -4 & -2 \\ 0 & -6 & 2 \\ 0 & -3 & 2 \end{bmatrix} \\
&= -2 \det \begin{bmatrix} 3 & -4 & -2 \\ 0 & -3 & 2 \\ 0 & -6 & 2 \end{bmatrix} \\
&= -2 \det \begin{bmatrix} 3 & -4 & -2 \\ 0 & -3 & 2 \\ 0 & 0 & -2 \end{bmatrix} \\
&= (-2) \cdot 3 \cdot (-3) \cdot (-2) \\
&= -36,
\end{aligned}$$

onde, na primeira igualdade, de cima para baixo, B foi obtida de A via $B(1,-) = \frac{1}{2}A(1,-)$, ou seja, $\det B = \frac{1}{2}\det A$, isto é, $\det A = 2\det B$. Nas outras igualdades, de cima para baixo, "zeramos" as entradas abaixo do pivô da primeira coluna, duas vezes, trocamos o sinal do determinante, onde permutamos linhas, e utilizamos o exercício 8 dessa seção.

(e)

$$\begin{aligned}
\det A &= \det B \\
&= \det \begin{bmatrix} \pi & e & \sqrt{2} & \sqrt{3} \\ 0 & 0 & 0 & 0 \\ 1 & 1/2 & -1 & -1/2 \\ \ln 2 & 1/\pi & \sqrt{3} & \pi^2 \end{bmatrix} \\
&= 0,
\end{aligned}$$

onde B foi obtida de A via $B(2,-) = A(2,-) + \sqrt{2}A(3,-)$.

(f) Ao considerarmos linhas, podemos permutá-las e adicioná-las a múltiplos escalares de outras linhas, conforme o cálculo seguinte:

$$\begin{aligned}
\det A &= -\det\begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 2 & 3 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 2 & 1 & 0 \end{bmatrix} \\
&= -\det\begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 3 & 0 & 0 & -2 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 2 & 1 & 0 \end{bmatrix} \\
&= \det\begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & -2 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 2 & 1 & 0 \end{bmatrix} \\
&= \det\begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & -1 & -4 \end{bmatrix} \\
&= -\det\begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & -1 & -4 \end{bmatrix} \\
&= \det\begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & -1 & -4 \\ 0 & 0 & 0 & 0 & -2 \end{bmatrix} \\
&= 2.
\end{aligned}$$

(g) Expandindo (sempre) ao longo das primeiras linhas, temos

$$\begin{aligned}
\det A &= 1\cdot\det\begin{bmatrix} 0 & 0 & 0 & a & b \\ 0 & 0 & 0 & c & d \\ 0 & 0 & 0 & e & f \\ p & q & r & s & t \\ v & w & x & y & z \end{bmatrix} \\
&= -a\det\begin{bmatrix} 0 & 0 & 0 & d \\ 0 & 0 & 0 & f \\ p & q & r & t \\ v & w & x & z \end{bmatrix} + b\det\begin{bmatrix} 0 & 0 & 0 & c \\ 0 & 0 & 0 & e \\ p & q & r & s \\ v & w & x & y \end{bmatrix} \\
&= ad\det\begin{bmatrix} 0 & 0 & 0 \\ p & q & r \\ v & w & x \end{bmatrix} - bc\det\begin{bmatrix} 0 & 0 & 0 \\ p & q & r \\ v & w & x \end{bmatrix} \\
&= 0,
\end{aligned}$$

pois linha nula acarreta determinante nulo.

21. Pelas operações elementares sobre $\det A$, dadas no enunciado do exercício anterior, e pelo fato do escalonamento

$$A \longrightarrow \cdots \longrightarrow \mathsf{I}$$

acarretar a existência de A^{-1}, conforme o exercício 15 dessa seção, demonstra-se que

$$\boxed{\textit{A não é invertível} \iff \det A = 0.}$$

Utilize essa equivalência para resolver (novamente) os exercícios 16 e 17 dessa seção.[31]

22. Como estabelecido no final da seção 3.4, se A e B são matrizes $n \times n$, então

$$\det(AB) = (\det A)(\det B).$$

Utilize essa afirmação para demonstrar que, se A é invertível, então

$$\det\left(A^{-1}\right) = (\det A)^{-1}.$$

DICA

$\det A \neq 0$ e $AA^{-1} = I$.

[31] DICA

Resolva

$$\det \begin{bmatrix} a & 2 & 3 \\ a & a & 4 \\ a & a & a \end{bmatrix} = 0 \text{ e } \det \begin{bmatrix} 2 & c & c \\ c & c & c \\ 8 & 7 & c \end{bmatrix} = 0.$$

Capítulo 4

Transformações lineares, autovalores e autovetores

Nesse capítulo, estudaremos *transformações*, isto é, funções,

$$\mathbb{R}^n \ni \mathbf{x} \longmapsto L(\mathbf{x}) \in \mathbb{R}^m,$$

que generalizam a função linear

$$\mathbb{R} \ni x \longmapsto f(x) = ax \in \mathbb{R},$$

onde, para quaisquer reais x, y e α, temos:

$$\begin{aligned} f(x+y) &= a(x+y) \\ &= ax + ay \\ &= f(x) + f(y); \end{aligned} \tag{4.1}$$

$$\begin{aligned} f(\alpha x) &= a(\alpha x) \\ &= \alpha(ax) \\ &= \alpha f(x). \end{aligned} \tag{4.2}$$

Para definirmos *linearidade*, utilizaremos as propriedades (4.1) e (4.2), no contexto da transformação L supracitada, e, como no capítulo 2, assumiremos que o vetor

$$\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$$

pode ser representado pela matriz

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n \times 1}.$$

Portanto, a imagem $L(\mathbf{x}) \in \mathbb{R}^m$ pode ser representada por uma matriz $m \times 1$.

4.1 Transformações lineares

Seja $L_A(\mathbf{x}) := A\mathbf{x}$ o produto das matrizes reais A e $\mathbf{x}$, de tamanhos $m \times n$ e $n \times 1$, respectivamente. Assim, temos a função

$$\begin{array}{rcccl} L_A & : & \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ & & \mathbf{x} & \longmapsto & L_A(\mathbf{x}), \end{array}$$

dita *multiplicação por A*.

EXEMPLO

Se $A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$, então

$$\begin{array}{rccc} L_A : & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ & \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} & \longmapsto & L_A(\mathbf{x}) = \begin{bmatrix} x_1 - x_2 \\ x_2 - x_3 \end{bmatrix} \end{array}.$$

Note que, a lei de correspondência de L_A também é representada por

$$\mathbf{x} = (x_1, x_2, x_3) \longmapsto L_A(\mathbf{x}) = (x_1 - x_2, x_2 - x_3)\,.$$

DEFINIÇÃO

$L : \mathbb{R}^n \to \mathbb{R}^m$ é uma função linear quando as condições

$$L(\mathbf{x} + \mathbf{y}) = L(\mathbf{x}) + L(\mathbf{y}) \quad e \quad L(\alpha\mathbf{x}) = \alpha L(\mathbf{x})$$

são satisfeitas, para quaisquer vetores $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ e qualquer escalar $\alpha \in \mathbb{R}$.

AFIRMAÇÃO 1

$L : \mathbb{R}^n \to \mathbb{R}^m$ é uma função linear se, e somente se, $L = L_A$ para alguma matriz A.

DEMONSTRAÇÃO

Por um lado, é fácil provar que L_A é linear, ou seja, $L_A(\mathbf{x} + \mathbf{y}) = L_A(\mathbf{x}) + L_A(\mathbf{y})$ e $L_A(\alpha\mathbf{x}) = \alpha L_A(\mathbf{x})$ para quaisquer $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$.[1] Por outro, suponha que L seja linear e considere que:

· $j = 1, 2, \ldots, n$;

· $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$;

· $\mathbf{e}_j$ é a matriz $n \times 1$ que representa o j-ésimo vetor da base canônica de $\mathbb{R}^n$, isto é, $\mathbf{e}_j$ é a j-ésima coluna da matriz identidade $n \times n$;

· $\mathbf{a}_j$ é a imagem de $\mathbf{e}_j$ pela função L, isto é, $L\left(\mathbf{e}_j\right) := \mathbf{a}_j$;

· $\mathbf{a}_j$ também é a j-ésima coluna da matriz

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

[1] De fato, faça com L_A o mesmo que fizemos com f em (4.1) e (4.2).

Então, pela linearidade de L, temos

$$\begin{aligned} L(\mathbf{x}) &= L\left(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n\right) \\ &= x_1 L\left(\mathbf{e}_1\right) + x_2 L\left(\mathbf{e}_2\right) + \cdots + x_n L\left(\mathbf{e}_n\right) \\ &= x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n \\ &= x_1\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1}\end{bmatrix} + x_2\begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2}\end{bmatrix} + \cdots + x_n\begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn}\end{bmatrix} \\ &= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} \\ &= A\mathbf{x} \\ &= L_A(\mathbf{x}). \end{aligned} \tag{4.3}$$

Observe que, em (4.3), ao calcularmos $L\left(\mathbf{e}_j\right)$, obtemos a coluna $A(-,j)$, $j = 1, 2, \ldots, n$, tal que $L = L_A$. A é chamada de matriz que representa L (nas bases canônicas).

EXERCÍCIOS

Verifique a linearidade das funções definidas por:

1. $L\left(x_1, x_2, x_3, x_4\right) = \left(x_1 + x_2, x_3 + x_4, x_1 + x_2 + x_3 + x_4\right) \quad \forall\left(x_1, x_2, x_3, x_4\right) \in \mathbb{R}^4$;

   RESOLUÇÃO

   Como

   $$L\left(\mathbf{e}_1\right) = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, L\left(\mathbf{e}_2\right) = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, L\left(\mathbf{e}_3\right) = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \text{ e } L\left(\mathbf{e}_4\right) = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix},$$

   se considerarmos

   $$\begin{aligned} A &= \begin{bmatrix} L\left(\mathbf{e}_1\right) & L\left(\mathbf{e}_2\right) & L\left(\mathbf{e}_3\right) & L\left(\mathbf{e}_4\right) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \end{aligned}$$

   então $L = L_A$.[2] Portanto, L é linear, pela afirmação 1 supracitada.

2. $L\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 2x_1 + x_2 \\ 4x_3 - 3x_4 \\ -6x_5 + 5x_6 \\ -x_1 + 2x_2 + 3x_3 + 4x_4 - 5x_5 + 6x_6 \\ x_1 \\ -x_2 \\ x_3 \\ -x_4 \end{bmatrix} \quad \forall \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} \in \mathbb{R}^6$;[3]

[2] Verifique!

[3] Confira a subseção 4.3.1.

3. ROTAÇÃO (ANTI-HORÁRIA) DE ÂNGULO θ EM TORNO DA ORIGEM

$$R_\theta(x_1, x_2) = (x_1 \cos\theta - x_2 \operatorname{sen}\theta, x_1 \operatorname{sen}\theta + x_2 \cos\theta) \quad \forall (x_1, x_2) \in \mathbb{R}^2; \tag{4.4}$$

RESOLUÇÃO

Como
$$R_\theta(\mathbf{e}_1) = \begin{bmatrix} \cos\theta \\ \operatorname{sen}\theta \end{bmatrix} \text{ e } R_\theta(\mathbf{e}_2) = \begin{bmatrix} -\operatorname{sen}\theta \\ \cos\theta \end{bmatrix},$$
se considerarmos
$$\begin{aligned} A &= \begin{bmatrix} R_\theta(\mathbf{e}_1) & R_\theta(\mathbf{e}_2) \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}, \end{aligned}$$
então $R_\theta = L_A$.[4] Portanto, R_θ é linear, pela afirmação 1 supracitada.

4. SEMELHANÇA (OU CISALHAMENTO) DE RAZÃO $k \in \mathbb{R}$

$$S_k(\mathbf{x}) = k\mathbf{x} \quad \forall \mathbf{x} \in \mathbb{R}^n;^5$$

RESOLUÇÃO

Note que, como $\mathbf{x} = \mathsf{I}\mathbf{x}$, temos
$$\begin{aligned} \mathbf{x} \mapsto S_k(\mathbf{x}) &= k\mathbf{x} \\ &= (k\mathsf{I})\mathbf{x}. \end{aligned}$$
Portanto, se $A = k\mathsf{I}$,[6] então $S_k = L_A$. Assim, S_k é linear, pela afirmação 1 supracitada.

5. PROJEÇÕES ORTOGONAIS SOBRE:

   (a) OS EIXOS x E y

$$P_x(x_1, x_2) = (x_1, 0) \quad \text{e} \quad P_y(x_1, x_2) = (0, x_2) \quad \forall\, (x_1, x_2) \in \mathbb{R}^2;^7$$

   (b) O SUBESPAÇO $\mathcal{S}$ DE $\mathbb{R}^n$ COM BASE ORTONORMAL $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r\}$

$$P_{\mathcal{S}}(\mathbf{x}) = (\mathbf{x} \cdot \mathbf{a}_1)\, \mathbf{a}_1 + (\mathbf{x} \cdot \mathbf{a}_2)\, \mathbf{a}_2 + \cdots + (\mathbf{x} \cdot \mathbf{a}_r)\, \mathbf{a}_r \quad \forall \mathbf{x} \in \mathbb{R}^n; \tag{4.5}$$

Note que, se $\mathbf{x}' := P_{\mathcal{S}}(\mathbf{x})$ e $\mathbf{x}'' := \mathbf{x} - \mathbf{x}'$, então $\mathbf{x}'' \in \mathcal{S}^\perp$. De fato, para $j = 1, 2, \ldots, r$,
$$\begin{aligned} \mathbf{x}'' \cdot \mathbf{a}_j &= [\mathbf{x} - (\mathbf{x} \cdot \mathbf{a}_1)\, \mathbf{a}_1 - \cdots - (\mathbf{x} \cdot \mathbf{a}_r)\, \mathbf{a}_r] \cdot \mathbf{a}_j \\ &= \mathbf{x} \cdot \mathbf{a}_j - (\mathbf{x} \cdot \mathbf{a}_1)(\mathbf{a}_1 \cdot \mathbf{a}_j) - \cdots - (\mathbf{x} \cdot \mathbf{a}_r)(\mathbf{a}_r \cdot \mathbf{a}_j) \\ &= \mathbf{x} \cdot \mathbf{a}_j - \mathbf{x} \cdot \mathbf{a}_j = 0. \end{aligned}$$
Assim, como $\mathbf{x} = \mathbf{x}' + \mathbf{x}'' \in \mathbb{R}^n$ com $\mathbf{x}' \in \mathcal{S}$ e $\mathbf{x}'' \in \mathcal{S}^\perp$, o nome *projeção ortogonal* se justifica. $P_{\mathcal{S}}(\mathbf{x})$ é ainda chamado de *melhor aproximação de* $\mathbf{x}$ *em* $\mathcal{S}$ ou de *vetor de* $\mathcal{S}$ *mais próximo de* $\mathbf{x}$, como veremos no capítulo 7.

---

[4] Verifique!

[5] Podemos assumir $k \neq 0$. De fato, a *função nula*, embora linear, não é uma semelhança. Note que, $S_1 = I$ é a *função identidade* em $\mathbb{R}^n$.

[6] $A = \begin{bmatrix} k & & \\ \text{zeros} & \ddots & \text{zeros} \\ & & k \end{bmatrix}.$

[7] Confira a subseção 4.3.1.

DICA

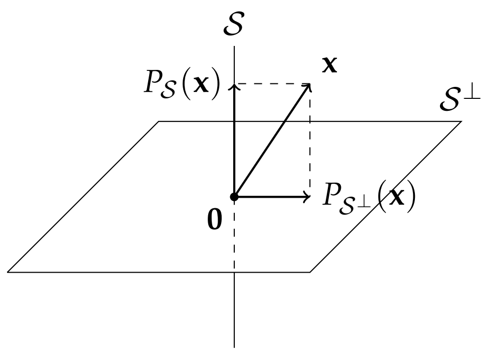
Verifique que

$$P_{\mathcal{S}}(\mathbf{x}+\mathbf{y}) = P_{\mathcal{S}}(\mathbf{x}) + P_{\mathcal{S}}(\mathbf{y}) \ \ e \ \ P_{\mathcal{S}}(\alpha\mathbf{x}) = \alpha P_{\mathcal{S}}(\mathbf{x})$$

para quaisquer $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *e* $\alpha \in \mathbb{R}$.

Ainda em relação ao item (b) do exercício 5, se $n = 3$ e $\mathcal{S}$ é a reta que passa pela origem na direção de $\mathbf{v} = (1,1,1)$, conforme ilustrada na figura 4.1, determine a matriz que representa $P_{\mathcal{S}}$ e a que representa $P_{\mathcal{S}^\perp}$.

Figura 4.1: Representação qualitativa das projeções ortogonais



RESOLUÇÃO PARCIAL

Primeiramente, note que:

- $\{\mathbf{a}\}$ é uma base ortonormal de $\mathcal{S}$ se

$$\begin{aligned} \mathbf{a} &= \frac{\mathbf{v}}{\|\mathbf{v}\|} \\ &= \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right); \end{aligned}$$

- $(1,1,1)\cdot(x,y,z) = 0$, isto é, $x+y+z=0$, se $(x,y,z) \in \mathcal{S}^\perp$. Portanto, uma base para o plano $\mathcal{S}^\perp$ é obtida de

$$\begin{aligned} (x,y,z) &= (x,y,-x-y) \\ &= x(1,0,-1) + y(0,1,-1) \\ &= x\mathbf{u}_1 + y\mathbf{u}_2. \end{aligned}$$

  Assim, uma base ortogonal para esse plano é obtida de

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{u}_1 \\ &= (1,0,-1), \\ \mathbf{v}_2 &= \mathbf{u}_2 - \frac{\mathbf{u}_2\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\mathbf{v}_1 \\ &= \left(-\frac{1}{2}, 1, -\frac{1}{2}\right). \end{aligned}$$

Então, $\{\mathbf{a}_1, \mathbf{a}_2\}$ é uma base ortonormal para $\mathcal{S}^\perp$, com

$$\begin{aligned}
\mathbf{a}_1 &= \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} \\
&= \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right), \\
\mathbf{a}_2 &= \frac{\mathbf{v}_2}{||\mathbf{v}_2||} \\
&= \left(-\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right).
\end{aligned}$$

Finalmente, basta calcularmos:

i. as leis de correspondências

$$\begin{aligned}
\mathbf{x} &\mapsto P_{\mathcal{S}}(\mathbf{x}) = (\mathbf{x} \cdot \mathbf{a})\mathbf{a}, \\
\mathbf{x} &\mapsto P_{\mathcal{S}^\perp}(\mathbf{x}) = (\mathbf{x} \cdot \mathbf{a}_1)\mathbf{a}_1 + (\mathbf{x} \cdot \mathbf{a}_2)\mathbf{a}_2;
\end{aligned}$$

ii. as matrizes 3×3, digamos

$$A_{\mathcal{S}} \quad \text{e} \quad A_{\mathcal{S}^\perp},$$

tais que

$$P_{\mathcal{S}} = L_{A_{\mathcal{S}}} \quad \text{e} \quad P_{\mathcal{S}^\perp} = L_{A_{\mathcal{S}^\perp}}.$$ [8]

6. REFLEXÕES EM TORNO:

(a) DOS EIXOS x E y

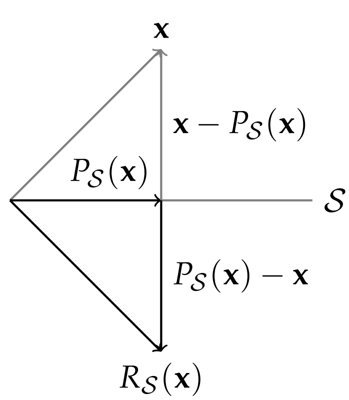
$$R_x(x_1, x_2) = (x_1, -x_2) \quad \text{e} \quad R_y(x_1, x_2) = (-x_1, x_2) \quad \forall (x_1, x_2) \in \mathbb{R}^2;$$ [9]

(b) DO SUBESPAÇO $\mathcal{S}$

$$R_{\mathcal{S}}(\mathbf{x}) = 2P_{\mathcal{S}}(\mathbf{x}) - \mathbf{x} \quad \forall \mathbf{x} \in \mathbb{R}^n,$$

onde $P_{\mathcal{S}}$ é dada como em (4.5) da página 86, conforme a figura 4.2.

Figura 4.2: Relação entre reflexão e projeção



Ainda em relação ao item (b) do exercício 6, se $n = 3$ e $\mathcal{S}$ é a reta que passa pela origem na direção de $\mathbf{v} = (1, 1, 1)$, determine a matriz que representa $R_{\mathcal{S}}$ e a que representa $R_{\mathcal{S}^\perp}$.[10]

[8] Confira a subseção 4.3.1.

[9] Idem!

[10] Idem!

AFIRMAÇÃO 2

Sejam A e B as matrizes que representam as funções lineares $L_1 : \mathbb{R}^n \to \mathbb{R}^m$ e $L_2 : \mathbb{R}^m \to \mathbb{R}^p$, respectivamente, isto é, $L_1 = L_A$ e $L_2 = L_B$. Portanto,

$$L_2 \circ L_1 = L_{BA},$$

ou seja, BA é a matriz que representa a função linear

$$\begin{array}{rccc} L_2 \circ L_1 & : & \mathbb{R}^n & \longrightarrow & \mathbb{R}^p \\ & & x & \longmapsto & L_2\left(L_1(x)\right) \end{array}.$$

DEMONSTRAÇÃO

$$\begin{aligned} \left(L_2 \circ L_1\right)(\mathbf{x}) &= L_2\left(L_1(\mathbf{x})\right) \\ &= L_2\left(L_A(\mathbf{x})\right) \\ &= L_2\left(A\mathbf{x}\right) \\ &= L_B\left(A\mathbf{x}\right) \\ &= B\left(A\mathbf{x}\right) \\ &= \left(BA\right)\mathbf{x} \\ &= L_{BA}\left(\mathbf{x}\right). \end{aligned}$$

EXERCÍCIOS

I. Considere $n = p = 2$, $m = 3$, $\mathbf{x} = (x_1, x_2) \longmapsto L_1(\mathbf{x}) = (x_2, x_1, x_1 + x_2)$ e $\mathbf{y} = (y_1, y_2, y_3) \longmapsto L_2(\mathbf{y}) = (y_1 + y_2, y_2 + y_3)$.

- Determine A e B.[11]
- Sem utilizar a afirmação 2 supracitada, obtenha a matriz C que representa $L_2 \circ L_1$, sendo $\mathbf{x} = (x_1, x_2) \mapsto L_2\left(L_1(\mathbf{x})\right)$, e verifique que, de fato, $C = BA$.[12]

II. Considerando as rotações no plano (4.4),[13] é fácil ver, geometricamente, que

$$R_{\theta_1+\theta_2} = R_{\theta_2} \circ R_{\theta_1},$$

conforme a a figura 4.3. Por outro lado, as matrizes que representam R_{θ_1}, R_{θ_2} e $R_{\theta_1+\theta_2}$ são dadas, respectivamente, por

$$A = \begin{bmatrix} \cos\theta_1 & -\operatorname{sen}\theta_1 \\ \operatorname{sen}\theta_1 & \cos\theta_1 \end{bmatrix}, \quad B = \begin{bmatrix} \cos\theta_2 & -\operatorname{sen}\theta_2 \\ \operatorname{sen}\theta_2 & \cos\theta_2 \end{bmatrix}$$

$$\text{e} \quad C = \begin{bmatrix} \cos(\theta_1+\theta_2) & -\operatorname{sen}(\theta_1+\theta_2) \\ \operatorname{sen}(\theta_1+\theta_2) & \cos(\theta_1+\theta_2) \end{bmatrix}.$$

Sem utilizar a afirmação 2 supracitada, obtenha a matriz que representa $R_{\theta_2} \circ R_{\theta_1}$. Verifique que essa matriz é, de fato, dada por C.[14]

---

[11] Confira a subseção 4.3.1.

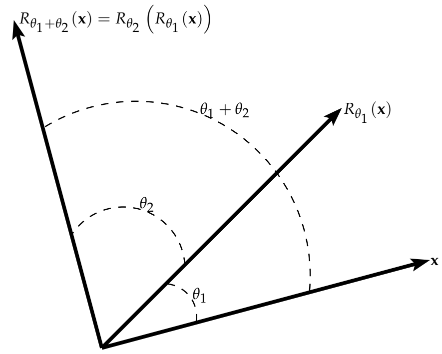
[12] Idem!

[13] Cf. p. 86.

[14] DICA

Utilize o cosseno e o seno da soma de dois ângulos.

Figura 4.3: $R_{\theta_1+\theta_2} = R_{\theta_2} \circ R_{\theta_1}$



4.1.1 Núcleo e imagem de A (ou L)

Considere a função linear $L : \mathbb{R}^n \to \mathbb{R}^m$ e a matriz A, $m \times n$, que a representa (nas bases canônicas). O *núcleo* e a *imagem* de L (ou A) são, respectivamente, os seguintes conjuntos:

(i) $\mathrm{Nu}(L) = \{\mathbf{x} \in \mathbb{R}^n \mid L(\mathbf{x}) = \mathbf{0}\}$, que consiste de cada vetor de $\mathbb{R}^n$ cuja imagem (por L) seja o vetor nulo de $\mathbb{R}^m$;

(ii) $\mathrm{Im}(L) = \{\mathbf{y} \in \mathbb{R}^m \mid \text{existe } \mathbf{x} \in \mathbb{R}^n \text{ com } \mathbf{y} = L(\mathbf{x})\}$, que é constituído de cada vetor de $\mathbb{R}^m$ que seja a imagem (por L) de algum vetor de $\mathbb{R}^n$.

AFIRMAÇÃO 3

$\mathrm{Nu}(L)$ *é um subespaço de* $\mathbb{R}^n$ *e* $\mathrm{Im}(L)$ *é um subespaço de* $\mathbb{R}^m$.

DEMONSTRAÇÃO

Primeiramente, como

$$\begin{aligned} L(\mathbf{0}) &= A\mathbf{0} \\ &= \mathbf{0}, \end{aligned}$$

note que $\mathrm{Nu}(L)$ e $\mathrm{Im}(L)$ são não vazios. Agora, em relação ao item (i), é fácil ver que

$$\mathrm{Nu}(L) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$$

é o subespaço das soluções do sistema homogêneo $A\mathbf{x} = \mathbf{0}$. Por fim, para (ii),

$$\mathrm{Im}(L) = \{\mathbf{y} \in \mathbb{R}^m \mid \text{existe } \mathbf{x} \in \mathbb{R}^n \text{ com } \mathbf{y} = A\mathbf{x}\}$$

é o *subespaço gerado pelas colunas* de A, pois, como vimos em (4.3),[15] se

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \text{e} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

então

$$\begin{aligned} \mathbf{y} &= A\mathbf{x} \\ &= x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n, \end{aligned}$$

[15]Cf. p. 85.

onde $\mathbf{a}_j$ representa a j-ésima coluna de A, $j = 1, 2, \ldots, n$.

EXEMPLO

Seja

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

(i) Note que, $x_4 = t$ acarreta

$$\begin{aligned} \mathrm{Nu}(L) \ni \mathbf{x} &= \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \\ &= t \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \end{aligned}$$

com $t \in \mathbb{R}$. Ainda,

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \right\}$$

é uma base de $\mathrm{Nu}(L)$ e $\dim \mathrm{Nu}(L) = 1$.

(ii) Note que,

$$\begin{aligned} \mathrm{Im}(L) \ni \mathbf{y} &= A\mathbf{x} \\ &= x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \end{aligned}$$

e, como

$$\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} = (-1) \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + (-1) \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 1 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

uma base de $\mathrm{Im}(L) = \mathbb{R}^3$ é a canônica, com $\dim \mathrm{Im}(L) = 3$.

AFIRMAÇÃO 4

Chamando as dimensões de $\mathrm{Nu}(L)$ e $\mathrm{Im}(L)$ de nulidade e posto de L (ou A), respectivamente, demonstra-se, pelo resultado (R14),[16] *que*

$$\textit{nulidade} + \textit{posto} = n.$$

Para o exemplo anterior, temos

$$\text{nulidade} + \text{posto} = 1 + 3 = 4 = n.$$

MÉTODO PRÁTICO PARA DETERMINARMOS $\mathrm{Nu}(L)$ E $\mathrm{Im}(L)$

[16]Cf. p. 157.

- Para determinarmos uma base de $\mathrm{Nu}(L)$, basta obtermos R, a matriz escalonada reduzida equivalente a A. De fato, por um lado, $\mathbf{x}$ é solução de $A\mathbf{x} = \mathbf{0}$ se, e somente se, $\mathbf{x}$ é solução de $\mathsf{R}\mathbf{x} = \mathbf{0}$, isto é, $\mathrm{Nu}(A) = \mathrm{Nu}(\mathsf{R})$. Por outro, a nulidade de A é o número de vetores de uma base do espaço solução do sistema $A\mathbf{x} = \mathbf{0}$.

- Para determinarmos uma base de $\mathrm{Im}(L)$, podemos seguir os passos seguintes:

(P1) Obtemos as colunas de R que contêm os seus pivôs. Assim, sejam

$$\mathbf{r}_{j_1}, \mathbf{r}_{j_2}, \ldots, \mathbf{r}_{j_k}$$

as colunas supracitadas;

(P2) Para obtermos uma base de $\mathrm{Im}(L)$, basta coletarmos as *colunas pivôs*

$$\mathbf{a}_{j_1}, \mathbf{a}_{j_2}, \ldots, \mathbf{a}_{j_k} \tag{4.6}$$

de A.[17]

EXEMPLO

Ao escalonarmos

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

temos

$$\begin{aligned} A &\longrightarrow \begin{bmatrix} 1 & -1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 & 2 \end{bmatrix} \\ &\longrightarrow \begin{bmatrix} 1 & 0 & -1 & 1 & 3 \\ 0 & 1 & -1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \mathsf{R}. \end{aligned}$$

Portanto, para obtermos uma base de $\mathrm{Im}(L)$, basta observarmos que, como $\mathbf{r}_1$ e $\mathbf{r}_2$ são as colunas pivôs de R, as colunas pivôs de A, isto é, $\mathbf{a}_1$ e $\mathbf{a}_2$, formam a base supracitada. De fato, por um lado, as outras colunas de R são CL de $\mathbf{r}_1$ e $\mathbf{r}_2$, pois

$$\begin{aligned} \mathbf{r}_3 &= -\mathbf{r}_1 - \mathbf{r}_2, \\ \mathbf{r}_4 &= \mathbf{r}_1 - \mathbf{r}_2 \text{ e} \\ \mathbf{r}_5 &= 3\mathbf{r}_1 + 2\mathbf{r}_2. \end{aligned} \tag{4.7}$$

Por outro, (4.7) também é válida caso troquemos $\mathbf{r}$ por $\mathbf{a}$, ou seja,

$$\begin{aligned} \mathbf{a}_3 &= -\mathbf{a}_1 - \mathbf{a}_2, \\ \mathbf{a}_4 &= \mathbf{a}_1 - \mathbf{a}_2 \text{ e} \\ \mathbf{a}_5 &= 3\mathbf{a}_1 + 2\mathbf{a}_2. \end{aligned}$$

Assim, $\{\mathbf{a}_1, \mathbf{a}_2\}$ é a base supracitada, pois seus vetores geram $\mathrm{Im}(L)$ e são LI.
Para determinarmos uma base para $\mathrm{Nu}(L)$, basta observarmos que, como $\mathsf{R}\mathbf{x} = \mathbf{0}$ representa o sistema

$$\begin{cases} x_1 \phantom{{}+x_2} - x_3 + x_4 + 3x_5 = 0, \\ \phantom{x_1 +{}} x_2 - x_3 - x_4 + 2x_5 = 0, \end{cases}$$

[17] O posto de A é o número de colunas pivôs de A.

caso $x_3 = \alpha$, $x_4 = \beta$ e $x_5 = \gamma$ sejam números reais quaisquer,

$$\begin{aligned} \mathrm{Nu}(L) \ni \mathbf{x} &= \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} \\ &= \alpha \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} -3 \\ -2 \\ 0 \\ 0 \\ 1 \end{bmatrix}. \end{aligned}$$

Portanto, uma base para $\mathrm{Nu}(L)$ é dada por

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ -2 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Note que,

$$\text{nulidade} + \text{posto} = 3 + 2 = 5 = n.$$

EXERCÍCIO

Seja

$$A = \begin{bmatrix} 1 & 2 & 2 & -1 & 1 & 3 \\ 2 & 4 & 5 & -3 & 3 & 7 \\ 3 & 6 & 7 & -4 & 4 & 10 \end{bmatrix}.$$

- Determine a matriz escalonada reduzida R obtida de A. Que colunas de R não contêm pivôs?[18] Escreva cada uma dessas colunas como CL das colunas pivôs de R.[19]
- Determine as colunas pivôs de A, ou seja, as colunas de A correspondentes às colunas pivôs de R.[20] Como as colunas pivôs de A formam uma base para o espaço gerado por todas as colunas de A, escreva cada uma das outras colunas de A como CL das colunas dessa base.[21]
- Qual a dimensão do núcleo de A, ou seja, qual a nulidade de A?[22]

Bases de subespaços

Como podemos obter uma base de um subespaço do $\mathbb{R}^n$ gerado por m vetores? Por exemplo, dados 4 vetores em $\mathbb{R}^6$, como podemos determinar uma base para o subespaço gerado por

---

[18] RESPOSTA $\mathbf{r}_i$, $i = 2, 4, 5, 6$.

[19] RESPOSTA $\mathbf{r}_2 = 2\mathbf{r}_1$, $\mathbf{r}_4 = \mathbf{r}_1 - \mathbf{r}_3$, $\mathbf{r}_5 = -\mathbf{r}_1 + \mathbf{r}_3$ e $\mathbf{r}_6 = \mathbf{r}_1 + \mathbf{r}_3$.

[20] RESPOSTA $\mathbf{a}_1$ e $\mathbf{a}_3$.

[21] RESPOSTA Na penúltima nota de rodapé, troque $\mathbf{r}$ por $\mathbf{a}$.

[22] RESPOSTA nulidade $= 6 - 2 = 4$.

esses vetores?
Essa questão pode ser respondida de dois modos. O primeiro é baseado no "método prático para determinarmos $\text{Im}(L)$", que acabamos de estudar. Os detalhes do segundo ficam a cargo do leitor.

> PRIMEIRO MODO
> *Considere a matriz A 6×4, cujas colunas são os quatro vetores dados. Determine as colunas pivôs de A. Essas colunas formam uma base do subespaço procurado.*
>
> SEGUNDO MODO
> *Considere a matriz A 4×6, cujas linhas são os quatro vetores dados. Determine as linhas não nulas da escalonada reduzida R, equivalente a A. Essas linhas formam uma base do subespaço procurado, pois elas contêm os pivôs de R e, portanto, nenhuma delas é CL das outras linhas da escalonada reduzida. Além disso, é fácil ver que as linhas de R e A geram o mesmo subespaço.*

EXERCÍCIO

Utilize os dois modos supracitados para determinar bases do subespaço de $\mathbb{R}^6$ gerado por $(1,-1,1,-1,0,1)$, $(-1,1,0,-1,1,-1)$, $(1,-1,2,-3,1,1)$ e $(3,-3,1,1,-2,3)$.

PRIMEIRO MODO

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

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

$$\longrightarrow \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \mathsf{R}.$$

Como $\mathbf{a}_1$ e $\mathbf{a}_2$ são as colunas pivôs de A, $\{\mathbf{a}_1, \mathbf{a}_2\}$ é uma base para o subespaço gerado por $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ e $\mathbf{a}_4$.

$\boxed{\text{SEGUNDO MODO}}$

$$A = \begin{bmatrix} 1 & -1 & 1 & -1 & 0 & 1 \\ -1 & 1 & 0 & -1 & 1 & -1 \\ 1 & -1 & 2 & -3 & 1 & 1 \\ 3 & -3 & 1 & 1 & -2 & 3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -1 & 1 & -1 & 0 & 1 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & -2 & 4 & -2 & 0 \end{bmatrix}$$
$$\longrightarrow \begin{bmatrix} 1 & -1 & 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \mathsf{R}.$$

Como $\mathsf{R}(1,-)$ e $\mathsf{R}(2,-)$, que são as linhas não nulas de R, foram obtidas do escalonamento de A pela aplicação de *operações elementares sobre linhas*,[23] $\{\mathsf{R}(1,-),\mathsf{R}(2,-)\}$ é uma base para o espaço gerado por $A(1,-)$, $A(2,-)$, $A(3,-)$ e $A(4,-)$.

4.1.2 Representação de L em outras bases

Vimos que, caso $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ seja linear, existe uma matriz $A \in \mathbb{R}^{m\times n}$ tal que, para cada $\mathbf{x} \in \mathbb{R}^n$, $L(\mathbf{x}) = A\mathbf{x}$. Como definimos, essa A é a matriz que representa L nas bases canônicas, pois

$$A = [L(\mathbf{e}_1)\ \ L(\mathbf{e}_2)\ \ \cdots\ \ L(\mathbf{e}_n)]\,,$$

onde $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ é a base canônica de $\mathbb{R}^n$, e cada coluna $A(-,j)$ é dada por

$$\begin{aligned} L(\mathbf{e}_j) &= \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \\ &= a_{1j}\mathbf{e}'_1 + a_{2j}\mathbf{e}'_2 + \cdots + a_{mj}\mathbf{e}'_m, \end{aligned}$$

onde $\{\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_m\}$ é a base canônica de $\mathbb{R}^m$. De fato, esse procedimento pode ser generalizado para outras bases, ou seja, podemos representar L por matrizes $m \times n$ em bases quaisquer, digamos $\mathcal{B} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ e $\mathcal{B}' = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_m\}$, de $\mathbb{R}^n$ e $\mathbb{R}^m$:

$$[L]_{\mathcal{B}'}^{\mathcal{B}} := [L(\mathbf{x}_1)\ \ L(\mathbf{x}_2)\ \ \cdots\ \ L(\mathbf{x}_n)]\,,$$

onde cada coluna $[L]_{\mathcal{B}'}^{\mathcal{B}}(-,j)$ é dada por

$$\begin{aligned} L(\mathbf{x}_j) &= \begin{bmatrix} \ell_{1j} \\ \ell_{2j} \\ \vdots \\ \ell_{mj} \end{bmatrix} \\ &= \ell_{1j}\mathbf{y}_1 + \ell_{2j}\mathbf{y}_2 + \cdots + \ell_{mj}\mathbf{y}_m. \end{aligned}$$

[23] Ou seja, cada linha de A é uma CL das linhas não nulas de R. O leitor é convidado a explicitar essas combinações lineares.

EXEMPLO

Considere $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ linear. Sejam $\mathcal{B} = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ e $\mathcal{B}' = \{\mathbf{y}_1, \mathbf{y}_2\}$ bases de $\mathbb{R}^3$ e $\mathbb{R}^2$.[24] Além disso, suponha que, para essas bases,

$$\begin{aligned} L(\mathbf{x}_1) &= 2\mathbf{y}_1 + 3\mathbf{y}_2, \\ L(\mathbf{x}_2) &= 4\mathbf{y}_1 + 5\mathbf{y}_2 \text{ e} \\ L(\mathbf{x}_3) &= 6\mathbf{y}_1 + 7\mathbf{y}_2. \end{aligned}$$

Portanto, as colunas da matriz $[L]_{\mathcal{B}'}^{\mathcal{B}}$ são dadas por

$$\begin{aligned} [L]_{\mathcal{B}'}^{\mathcal{B}}(-,1) &= \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \\ [L]_{\mathcal{B}'}^{\mathcal{B}}(-,2) &= \begin{bmatrix} 4 \\ 5 \end{bmatrix} \text{ e} \\ [L]_{\mathcal{B}'}^{\mathcal{B}}(-,3) &= \begin{bmatrix} 6 \\ 7 \end{bmatrix}, \end{aligned}$$

ou seja,

$$[L]_{\mathcal{B}'}^{\mathcal{B}} = \begin{bmatrix} 2 & 4 & 6 \\ 3 & 5 & 7 \end{bmatrix}.$$

NOTAÇÃO

Caso $n = m$ e $\mathcal{B}' = \mathcal{B}$, a matriz $[L]_{\mathcal{B}'}^{\mathcal{B}}$ é denotada simplesmente por $[L]_{\mathcal{B}}$.

EXERCÍCIO

Se $m = n = 2$, $\mathcal{B}$ é a base canônica,[25] $\mathcal{B}' = \{\mathbf{y}_1 = (1,1), \mathbf{y}_2 = (1,-1)\}$ e $L(\mathbf{x}) = (x_1 - x_2, x_1 + x_2)$ para cada $\mathbf{x} = (x_1, x_2)$, determine as matrizes $[L]_{\mathcal{B}'}^{\mathcal{B}}$, $[L]_{\mathcal{B}}^{\mathcal{B}'}$, $[L]_{\mathcal{B}}$ e $[L]_{\mathcal{B}'}$.

RESOLUÇÃO

Como

$$\begin{cases} L(\mathbf{x}_1) = (1,1) = 1 \cdot \mathbf{y}_1 + 0 \cdot \mathbf{y}_2, \\ L(\mathbf{x}_2) = (-1,1) = 0 \cdot \mathbf{y}_1 - 1 \cdot \mathbf{y}_2, \end{cases}$$

temos

$$[L]_{\mathcal{B}'}^{\mathcal{B}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Como

$$\begin{cases} L(\mathbf{y}_1) = (0,2) = 0 \cdot \mathbf{x}_1 + 2 \cdot \mathbf{x}_2, \\ L(\mathbf{y}_2) = (2,0) = 2 \cdot \mathbf{x}_1 + 0 \cdot \mathbf{x}_2, \end{cases}$$

temos

$$[L]_{\mathcal{B}}^{\mathcal{B}'} = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}.$$

Como

$$\begin{cases} L(\mathbf{x}_1) = (1,1) = 1 \cdot \mathbf{x}_1 + 1 \cdot \mathbf{x}_2, \\ L(\mathbf{x}_2) = (-1,1) = -1 \cdot \mathbf{x}_1 + 1 \cdot \mathbf{x}_2, \end{cases}$$

temos

$$\begin{aligned} A &= [L]_{\mathcal{B}} \\ &= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}. \end{aligned}$$

---

[24] Por exemplo, sejam $\mathbf{x}_1 = (1,1,0)$, $\mathbf{x}_2 = (1,0,1)$, $\mathbf{x}_3 = (0,1,1)$, $\mathbf{y}_1 = (1,1)$ e $\mathbf{y}_2 = (1,-1)$.

[25] Isto é, $\mathbf{x}_1 = \mathbf{e}_1$ e $\mathbf{x}_2 = \mathbf{e}_2$.

Como

$$\begin{cases} L(\mathbf{y}_1) &= (0,2) &= 1 \cdot \mathbf{y}_1 - 1 \cdot \mathbf{y}_2, \\ L(\mathbf{y}_2) &= (2,0) &= 1 \cdot \mathbf{y}_1 + 1 \cdot \mathbf{y}_2, \end{cases}$$

temos

$$\begin{aligned} A' &= [L]_{\mathcal{B}'} \\ &= \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}. \end{aligned}$$

PERGUNTA

L também é uma "multiplicação pela matriz" $[L]_{\mathcal{B}'}^{\mathcal{B}}$?

RESPOSTA

Sim, pelo resultado (R15) da página 159, $\mathbf{x} \mapsto L(\mathbf{x})$ pode ser representada por

$$[\mathbf{x}]_{\mathcal{B}} \mapsto [L(\mathbf{x})]_{\mathcal{B}'} = [L]_{\mathcal{B}'}^{\mathcal{B}} \, [\mathbf{x}]_{\mathcal{B}}, \tag{4.8}$$

onde $[\mathbf{x}]_{\mathcal{B}}$ e $[L(\mathbf{x})]_{\mathcal{B}'}$ são matrizes $n \times 1$ e $m \times 1$, respectivamente, que representam os vetores $\mathbf{x}$ e $L(\mathbf{x})$ nas respectivas bases.

EXERCÍCIO

Com as mesmas hipóteses do exercício anterior, verifique a validade de (4.8) para $\mathbf{x} = (1,2)$.

RESOLUÇÃO

Por um lado, $L(\mathbf{x}) = (-1,3) = 1 \cdot \mathbf{y}_1 - 2\mathbf{y}_2$ implica que $[L(\mathbf{x})]_{\mathcal{B}'} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$. Por outro, temos $[L]_{\mathcal{B}'}^{\mathcal{B}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ e $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Agora, basta verificar o que se pede.

Como as matrizes $A = [L]_{\mathcal{B}}$ e $A' = [L]_{\mathcal{B}'}$ estão relacionadas?

Primeiramente, afirmamos que existe uma única função linear $T : \mathbb{R}^n \to \mathbb{R}^n$ tal que

$$\begin{array}{ccc} \mathbf{x}_1 & \overset{T}{\mapsto} & \mathbf{y}_1; \\ \mathbf{x}_2 & \overset{T}{\mapsto} & \mathbf{y}_2; \\ & \vdots & \\ \mathbf{x}_n & \overset{T}{\mapsto} & \mathbf{y}_n. \end{array} \tag{4.9}$$

A demonstração da existência e da linearidade de T encontra-se em (R10), página 154. Quanto a unicidade de T, seja $S : \mathbb{R}^n \to \mathbb{R}^n$ uma função linear tal que

$$\mathbf{x}_j \overset{S}{\mapsto} \mathbf{y}_j$$

para cada índice $j \in \{1, 2, \ldots, n\}$. Portanto,

$$S\left(\mathbf{x}_j\right) = T\left(\mathbf{x}_j\right)$$

para $j = 1, 2, \ldots, n$. Então, para todo $\mathbf{x} \in \mathbb{R}^n$, pela linearidade de S e T, temos

$$\begin{aligned} S(\mathbf{x}) &= S\left(x_1\mathbf{x}_1 + x_2\mathbf{x}_2 + \cdots + x_n\mathbf{x}_n\right) \\ &= x_1 S\left(\mathbf{x}_1\right) + x_2 S\left(\mathbf{x}_2\right) + \cdots + x_n S\left(\mathbf{x}_n\right) \\ &= x_1 T\left(\mathbf{x}_1\right) + x_2 T\left(\mathbf{x}_2\right) + \cdots + x_n T\left(\mathbf{x}_n\right) \\ &= T\left(x_1\mathbf{x}_1 + x_2\mathbf{x}_2 + \cdots + x_n\mathbf{x}_n\right) \\ &= T(\mathbf{x}), \end{aligned}$$

onde os x_j's são as coordenadas de $\mathbf{x}$ na base $\mathcal{B}$.

Demonstraremos, na seção 7.6, que T é invertível, com inversa T^{-1} dada por

$$T^{-1}\left(\mathbf{y}_i\right) = \mathbf{x}_i, \quad i = 1, 2, \ldots, n,$$

e, se

$$\mathsf{P} := [T]_{\mathcal{B}}, \tag{4.10}$$

então:

$$\mathsf{P}^{-1} = \left[T^{-1}\right]_{\mathcal{B}'}$$

e

$$A' = \mathsf{P}^{-1} A \mathsf{P}. \tag{4.11}$$

Nesse caso, dizemos que as matrizes A e A', que representam a função linear L, são *semelhantes (entre si)*.

EXERCÍCIO

Com as mesmas hipóteses do penúltimo exercício anterior, verifique a validade de (4.11).

RESOLUÇÃO

Como em (4.9), seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ a função linear

$$\begin{aligned} \mathbf{x}_1 &\overset{T}{\mapsto} \mathbf{y}_1; \\ \mathbf{x}_2 &\overset{T}{\mapsto} \mathbf{y}_2. \end{aligned}$$

Portanto,

$$\begin{aligned} (1,0) &\overset{T}{\mapsto} (1,1) = 1(1,0) + 1(0,1); \\ (0,1) &\overset{T}{\mapsto} (1,-1) = 1(1,0) - 1(0,1). \end{aligned}$$

Note que, T é uma multiplicação por P com

$$\begin{aligned} \mathsf{P} &= [T]_{\mathcal{B}} \\ &= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \end{aligned}$$

Por outro lado, temos
$$\mathsf{P}^{-1} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}.$$
Assim, (4.11) é válida para esse exercício, pois
$$\begin{aligned} \mathsf{P}^{-1}A\mathsf{P} &= \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}\begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ &= A'. \end{aligned}$$

EXERCÍCIO

Se $m = n = 2$, $\mathcal{B} = \{\mathbf{x}_1 = (1,0), \mathbf{x}_2 = (1,1)\}$, $\mathcal{B}' = \{\mathbf{y}_1 = (0,1), \mathbf{y}_2 = (-1,1)\}$ e $L(\mathbf{x}) = (x_2, x_1)$ para cada $\mathbf{x} = (x_1, x_2)$, determine $A = [L]_{\mathcal{B}}$, $A' = [L]_{\mathcal{B}'}$, P e P^{-1} tais que $A' = \mathsf{P}^{-1}A\mathsf{P}$.

RESOLUÇÃO

Como
$$\begin{cases} L(\mathbf{x}_1) = (0,1) = -1\cdot\mathbf{x}_1 + 1\cdot\mathbf{x}_2, \\ L(\mathbf{x}_2) = (1,1) = 0\cdot\mathbf{x}_1 + 1\cdot\mathbf{x}_2, \end{cases}$$
temos
$$\begin{aligned} A &= [L]_{\mathcal{B}} \\ &= \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}. \end{aligned}$$
Como
$$\begin{cases} L(\mathbf{y}_1) = (1,0) = 1\cdot\mathbf{y}_1 - 1\cdot\mathbf{y}_2, \\ L(\mathbf{y}_2) = (1,-1) = 0\cdot\mathbf{y}_1 - 1\cdot\mathbf{y}_2, \end{cases}$$
temos
$$\begin{aligned} A' &= [L]_{\mathcal{B}'} \\ &= \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix}. \end{aligned}$$
Sejam, agora, $T : \mathbb{R}^2 \to \mathbb{R}^2$ e $\mathsf{P} = [T]_{\mathcal{B}}$ como dadas em (4.9) e (4.10), isto é, considere
$$\begin{aligned} T(1,0) &= (0,1) \\ &= (-1)\cdot\mathbf{x}_1 + 1\cdot\mathbf{x}_2, \\ T(1,1) &= (-1,1) \\ &= (-2)\cdot\mathbf{x}_1 + 1\cdot\mathbf{x}_2. \end{aligned}$$
Portanto,
$$\mathsf{P} = \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix}$$
e, assim,
$$\mathsf{P}^{-1} = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix}.[26]$$

4.2 Autovalores e autovetores

Para essa seção, considere $L : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ linear e $A = [L]_{\mathcal{B}}$, onde $\mathcal{B}$ é a base canônica, ou seja, $L = L_A$.[27] Um escalar λ é chamado de *autovalor de* A (ou L), caso exista um vetor não

[26] Verifique a validade de (4.11) para esse exercício.

[27] Cf. pp. 85 e 95.

nulo $\mathbf{x}$ tal que

$$\begin{aligned} L(\mathbf{x}) &= A\mathbf{x} \\ &= \lambda\mathbf{x}. \end{aligned} \tag{4.12}$$

Nesse caso, $\mathbf{x}$ é chamado de *autovetor de* A (ou L) *associado à* λ.

EXEMPLO

Se $A = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix}$, então $\lambda_1 = 3$ é autovalor de A associado à $\mathbf{x}_1 = \begin{bmatrix} -4 \\ 1 \end{bmatrix}$. De fato,

$$\begin{aligned} A\mathbf{x}_1 &= \begin{bmatrix} -12 \\ 3 \end{bmatrix} \\ &= \lambda_1\mathbf{x}_1. \end{aligned}$$

Analogamente, $\lambda_2 = -2$ é autovalor de A associado à $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. De fato,

$$\begin{aligned} A\mathbf{x}_2 &= \begin{bmatrix} -2 \\ -2 \end{bmatrix} \\ &= \lambda_2\mathbf{x}_2. \end{aligned}$$

Sendo I *a matriz identidade* $n \times n$*, note que a condição (4.12) pode ser reescrita como*

$$(A - \lambda\mathsf{I})\,\mathbf{x} = \mathbf{0}.$$

De fato,

$$\begin{aligned} A\mathbf{x} = \lambda\mathbf{x} &\iff A\mathbf{x} - \lambda\mathbf{x} = \mathbf{0} \\ &\iff A\mathbf{x} - \lambda\mathsf{I}\mathbf{x} = \mathbf{0} \\ &\iff (A - \lambda\mathsf{I})\,\mathbf{x} = \mathbf{0}. \end{aligned}$$

Calculando autovalores

Por um lado, para que $(A - \lambda\mathsf{I})\,\mathbf{x} = \mathbf{0}$ admita solução $\mathbf{x}$ não nula, $A - \lambda\mathsf{I}$ não pode ser invertível pois, caso contrário, isto é, caso exista $(A - \lambda\mathsf{I})^{-1}$, temos

$$\begin{aligned} \mathbf{x} &= \mathsf{I}\mathbf{x} \\ &= (A - \lambda\mathsf{I})^{-1}(A - \lambda\mathsf{I})\,\mathbf{x} \\ &= (A - \lambda\mathsf{I})^{-1}\,\mathbf{0} \\ &= \mathbf{0}. \end{aligned}$$

Por outro, foi estabelecido (no penúltimo exercício do capítulo anterior) que

uma matriz quadrada não é invertível se, e somente se, seu determinante é nulo.

Portanto, os autovalores de A são as raízes do seu *polinômio característico*, definido por

$$p(\lambda) := \det(A - \lambda \mathsf{I}).$$

EXEMPLO

Seja $A = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix}$.

- CÁLCULO DO(S) AUTOVALOR(ES)

$$\begin{aligned} p(\lambda) &= \det \begin{bmatrix} 2-\lambda & -4 \\ -1 & -1-\lambda \end{bmatrix} \\ &= (2-\lambda)(-1-\lambda) - (-4)(-1) \\ &= \lambda^2 - \lambda - 6 \\ &= (\lambda - 3)(\lambda + 2). \end{aligned}$$

Assim, $\lambda_1 = 3$ e $\lambda_2 = -2$ são os autovalores de A.

Calculando autovetores

Tendo calculado λ, os autovetores de A associados (à λ) podem ser obtidos das soluções não nulas $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ do sistema $(A - \lambda \mathsf{I})\,\mathbf{x} = \mathbf{0}$. O conjunto solução desse sistema, denotado por $\mathcal{S}_\lambda$, é o subespaço de $\mathbb{R}^n$ conhecido como *autoespaço de* A (ou L) *associado à* λ.

EXEMPLO

Considere o exemplo anterior. Seja $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

- CÁLCULO DOS AUTOVETORES $\mathbf{x}$ ASSOCIADOS À $\lambda_1 = 3$

$$\begin{aligned} (A - 3\mathsf{I})\,\mathbf{x} = \mathbf{0} &\iff \begin{bmatrix} -1 & -4 \\ -1 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ &\iff x_1 + 4x_2 = 0. \end{aligned}$$

Assim, $x_2 = t$ e $x_1 = -4x_2 = -4t$ determinam os autovetores

$$\mathbf{x} = t \begin{bmatrix} -4 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}.$$

Portanto, $\left\{ \begin{bmatrix} -4 \\ 1 \end{bmatrix} \right\}$ é uma base para $\mathcal{S}_{\lambda_1}$.

- CÁLCULO DOS AUTOVETORES $\mathbf{x}$ ASSOCIADOS À $\lambda_2 = -2$

$$\begin{aligned} (A + 2\mathsf{I})\,\mathbf{x} = \mathbf{0} &\iff \begin{bmatrix} 4 & -4 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ &\iff x_1 - x_2 = 0. \end{aligned}$$

Assim, $x_1 = x_2 = t$ determina os autovetores

$$\mathbf{x} = t \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}.$$

Portanto, $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ é uma base para $\mathcal{S}_{\lambda_2}$.

Calculando autovalores e autovetores

EXEMPLO

Seja $A = \begin{bmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{bmatrix}$.

· CÁLCULO DO(S) AUTOVALOR(ES)

$$\begin{aligned} p(\lambda) &= \det \begin{bmatrix} 5-\lambda & 8 & 16 \\ 4 & 1-\lambda & 8 \\ -4 & -4 & -11-\lambda \end{bmatrix} \\ &= (\lambda-1)(\lambda+3)^2. \end{aligned}$$

Portanto, $\lambda_1 = 1$ e $\lambda_2 = -3$ são autovalores de A com multiplicidades 1 e 2, respectivamente.

Seja, agora, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$.

· CÁLCULO DOS AUTOVETORES $\mathbf{x}$ ASSOCIADOS À $\lambda_1 = 1$

$$\begin{aligned} (A - \mathsf{I})\,\mathbf{x} = \mathbf{0} &\iff \begin{bmatrix} 4 & 8 & 16 \\ 4 & 0 & 8 \\ -4 & -4 & -12 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ &\iff x_1 + 2x_3 = 0 \quad \text{e} \quad x_2 + x_3 = 0. \end{aligned}$$

Assim, $x_3 = t$, $x_1 = -2t$ e $x_2 = -t$ determinam os autovetores

$$\mathbf{x} = t \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}.$$

Portanto, $\left\{ \mathbf{x}_1 = \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} \right\}$ é uma base para $\mathcal{S}_{\lambda_1}$.

· CÁLCULO DOS AUTOVETORES $\mathbf{x}$ ASSOCIADOS À $\lambda_2 = -3$

$$\begin{aligned} (A + 3\mathsf{I})\,\mathbf{x} = \mathbf{0} &\iff \begin{bmatrix} 8 & 8 & 16 \\ 4 & 4 & 8 \\ -4 & -4 & -8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ &\iff x_1 + x_2 + 2x_3 = 0. \end{aligned}$$

Assim, $x_3 = t$, $x_2 = s$ e $x_1 = -s - 2t$ determinam os autovetores

$$\mathbf{x} = s \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}, \quad s, t \in \mathbb{R}.$$

Portanto, $\left\{ \mathbf{x}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}$ é uma base para $\mathcal{S}_{\lambda_2}$.

AVISOS IMPORTANTES

(1) *Demonstraremos, no final da seção 7.4, que* $\dim(\mathcal{S}_\lambda)$ *não excede a multiplicidade de λ, como raiz de $p(\lambda) = 0$.*

(2) *Cálculos envolvendo determinantes são úteis para matrizes quadradas de ordem n, caso n seja "pequeno". Por exemplo, para $n \in \{2, 3, 4\}$. Contudo, caso n seja "grande", devido ao alto custo computacional desses cálculos, são utilizados outros procedimentos, tais como, os métodos iterativos da álgebra linear numérica.*

(3) *$A \in \mathbb{R}^{n\times n}$ (ou $L = L_A$) pode não ter autovalores (reais). Para um exemplo, confira (5.2), página 133.*

(4) *Além de reais, autovalores e coordenadas de autovetores também podem ser números complexos. Para um exemplo, confira (5.2).*

4.2.1 Diagonalização

Caso $\mathcal{B}$ seja a base canônica de $\mathbb{R}^n$, dizemos que $A = [L]_\mathcal{B}$ é *diagonalizável* quando existe uma base $\mathcal{B}' = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ de $\mathbb{R}^n$ composta de autovetores de A. Nesse caso,

$$\mathbf{x}_i \overset{L}{\mapsto} A\mathbf{x}_i = \lambda_i \mathbf{x}_i, \quad i = 1, 2, \dots, n.$$

Portanto, como

$$\begin{array}{ccllll} \mathbf{x}_1 & \overset{L}{\mapsto} & \lambda_1\mathbf{x}_1, & & & \\ \mathbf{x}_2 & \overset{L}{\mapsto} & & \lambda_2\mathbf{x}_2, & & \\ \vdots & \vdots & & & \ddots & \\ \mathbf{x}_n & \overset{L}{\mapsto} & & & & \lambda_n\mathbf{x}_n, \end{array}$$

temos a seguinte matriz diagonal:

$$[L]_{\mathcal{B}'} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \overbrace{=}^{\text{NOTAÇÃO}} \mathsf{D},$$

onde as entradas de D que não estão na diagonal principal são nulas e foram suprimidas. Note que a diagonal principal de D é composta dos autovalores de A.

Por que A é diagonálizável?
Como vimos, sendo $A = [L]_{\mathcal{B}}$ e $\mathsf{D} = [L]_{\mathcal{B}'}$, existe uma matriz invertível P tal que

$$\mathsf{P}^{-1}A\mathsf{P} = \mathsf{D}.$$

Assim, mesmo que A não seja uma matriz diagonal, A é diagonalizável.

AFIRMAÇÃO 5

$\mathbf{x}_j$ é a j-ésima coluna de P, $j = 1, 2, \cdots, n$, isto é, $\mathsf{P} = \left[\begin{array}{cccc} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{array}\right]$.

De fato, como $\mathsf{P} = [T]_{\mathcal{B}}$ para $\mathbf{e}_j \overset{T}{\mapsto} \mathbf{x}_j$, $j = 1, \ldots, n$,[28] temos

$$\mathsf{P} = \left[\begin{array}{cccc} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{array}\right]$$

onde $T(\mathbf{e}_j) = \mathbf{x}_j$, $j = 1, 2, \ldots, n$.[29]

EXERCÍCIOS

1. Para o exemplo 2×2 anterior,

$$\begin{aligned} A &= \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix}, \\ \mathsf{P} &= \left[\begin{array}{cc} \mathbf{x}_1 & \mathbf{x}_2 \end{array}\right] \\ &= \begin{bmatrix} -4 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{e} \\ \mathsf{D} &= \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}. \end{aligned}$$

   Verifique que

$$\mathsf{P}^{-1} = \begin{bmatrix} -1/5 & 1/5 \\ 1/5 & 4/5 \end{bmatrix} \quad \text{e} \quad \mathsf{P}^{-1}A\mathsf{P} = \mathsf{D}.$$

2. Para o exemplo 3×3 anterior,

$$\begin{aligned} A &= \begin{bmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{bmatrix}, \\ \mathsf{P} &= \left[\begin{array}{ccc} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \end{array}\right] \\ &= \begin{bmatrix} -2 & -1 & -2 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{e} \\ \mathsf{D} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{bmatrix}. \end{aligned}$$

[28] Cf. (4.9), (4.10) e (4.11), pp. 97–98.

[29] Confira o início da subseção 4.1.2.

Verifique que

$$\mathsf{P}^{-1} = \begin{bmatrix} -1 & -1 & -2 \\ -1 & 0 & -2 \\ 1 & 1 & 3 \end{bmatrix} \text{ e } \mathsf{P}^{-1}A\mathsf{P} = \mathsf{D}.$$

3. Diagonalize a matriz

$$A = \begin{bmatrix} 0,8 & 0,3 \\ 0,2 & 0,7 \end{bmatrix}$$

para obter A^{100}.[30] Na resolução, por abuso de notação, considere $\left(\frac{1}{2}\right)^{100} = 0$.

RESOLUÇÃO DO EX. 3

A^{100}, sem diagonalização, pode ser obtida calculando-se A^n, $n = 1, 2, \ldots, 100$, ou seja,

$$\overbrace{\begin{bmatrix} 0,8 & 0,3 \\ 0,2 & 0,7 \end{bmatrix}}^{A}, \overbrace{\begin{bmatrix} 0,70 & 0,45 \\ 0,30 & 0,55 \end{bmatrix}}^{A^2=AA}, \overbrace{\begin{bmatrix} 0,650 & 0,525 \\ 0,350 & 0,475 \end{bmatrix}}^{A^3=AA^2}, \ldots, \overbrace{\begin{bmatrix} 0,6 & 0,6 \\ 0,4 & 0,4 \end{bmatrix}}^{A^{100}=AA^{99}}.$$

Existe um alto custo computacional nesse procedimento. Contudo, caso A seja diagonalizável,

$$A = \mathsf{PDP}^{-1},\ A^2 = AA = \mathsf{PD}^2\mathsf{P}^{-1},\ A^3 = AA^2 = \mathsf{PD}^3\mathsf{P}^{-1},\ \ldots,\ A^{100} = \mathsf{PD}^{100}\mathsf{P}^{-1},$$

onde, sendo λ_1 e λ_2 os autovalores de A,

$$D^{100} = \begin{bmatrix} \lambda_1^{100} & 0 \\ 0 & \lambda_2^{100} \end{bmatrix}.$$

Esses autovalores são as raízes do polinômio característico de A, isto é, os valores de λ para os quais

$$\begin{aligned} \det(A - \lambda \mathsf{I}) &= \begin{vmatrix} \frac{4}{5} - \lambda & \frac{3}{10} \\ \frac{1}{5} & \frac{7}{10} - \lambda \end{vmatrix} \\ &= \lambda^2 - \frac{3}{2}\lambda + \frac{1}{2} \\ &= (\lambda - 1)\left(\lambda - \frac{1}{2}\right) \\ &= 0. \end{aligned}$$

Agora, para obtermos um autovetor $\mathbf{x}_1$ de A associado à $\lambda_1 = 1$, podemos escalonar a matriz do sistema $(A - \mathsf{I})\mathbf{x}_1 = \mathbf{0}$, ou seja,

$$\begin{aligned} \begin{bmatrix} -\frac{1}{5} & \frac{3}{10} \\ \frac{1}{5} & -\frac{3}{10} \end{bmatrix} &\longrightarrow \begin{bmatrix} \frac{1}{5} & -\frac{3}{10} \\ 0 & 0 \end{bmatrix} \\ &\longrightarrow \begin{bmatrix} 1 & -\frac{3}{2} \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

Analogamente, para obtermos um autovetor $\mathbf{x}_2$ de A associado à $\lambda_2 = \frac{1}{2}$, podemos escalonar a matriz do sistema $\left(A - \frac{1}{2}\mathsf{I}\right)\mathbf{x}_2 = \mathbf{0}$, ou seja,

$$\begin{aligned} \begin{bmatrix} \frac{3}{10} & \frac{3}{10} \\ \frac{1}{5} & \frac{1}{5} \end{bmatrix} &\longrightarrow \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\ &\longrightarrow \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

[30] Aqui, A^{100} representa a centésima potência de A.

Portanto, para $\mathbf{x}_1 = \begin{bmatrix} \frac{3}{2} \\ 1 \end{bmatrix}$ e $\mathbf{x}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, como $\{\mathbf{x}_1, \mathbf{x}_2\}$ é uma base de $\mathbb{R}^2$, isto é, A é diagonalizável, temos $\mathsf{P} = \begin{bmatrix} \frac{3}{2} & -1 \\ 1 & 1 \end{bmatrix}$, $\mathsf{P}^{-1} = \begin{bmatrix} \frac{2}{5} & \frac{2}{5} \\ -\frac{2}{5} & \frac{3}{5} \end{bmatrix}$ e

$$
\begin{aligned}
A^{100} &= \mathsf{P} \begin{bmatrix} 1^{100} & 0 \\ 0 & (1/2)^{100} \end{bmatrix} \mathsf{P}^{-1} \\
&= \mathsf{P} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \mathsf{P}^{-1} \\
&= \begin{bmatrix} 3/5 & 3/5 \\ 2/5 & 2/5 \end{bmatrix} \\
&= \begin{bmatrix} 0,6 & 0,6 \\ 0,4 & 0,4 \end{bmatrix}.
\end{aligned}
$$

Observação importantíssima

Pode ser que $A \in \mathbb{R}^{n \times n}$ não seja diagonalizável, isto é, pode não ser possível obter uma base $\mathcal{B}'$ formada por n autovetores de A. Para exemplos, confira a seção 4.3.

4.2.2 Matrizes ortogonais e diagonalização

Uma matriz invertível P com entradas reais e tal que $P^{-1} = P^t$ é chamada de matriz *ortogonal*.

EXEMPLOS

$$
\begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}, \quad \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix} \quad \text{e} \quad \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}
$$

são ortogonais.[31]

AFIRMAÇÃO 6

A ortogonalidade de P é equivalente tanto a ortonormalidade de suas colunas, quanto de suas linhas, ou seja, cada uma das três afirmações seguintes é equivalente as outras duas:

1. *P é ortogonal;*
2. *$\{P(-,1), \ldots, P(-,n)\}$ é uma base ortonormal de $\mathbb{R}^n$;*
3. *$\{P(1,-)^t, \ldots, P(n,-)^t\}$ é uma base ortonormal de $\mathbb{R}^n$.*

EXEMPLOS

Confira os três primeiros exemplos dessa subseção.

DEMONSTRAÇÃO DA AFIRMAÇÃO 6

Como

$$
P^t P = \begin{bmatrix} P^t(1,-) \cdot P(-,1) & \cdots & P^t(1,-) \cdot P(-,n) \\ \vdots & \vdots & \vdots \\ P^t(n,-) \cdot P(-,1) & \cdots & P^t(n,-) \cdot P(-,n) \end{bmatrix}
$$

[31] Para o último exemplo, confira os dois primeiros exercícios relacionados ao processo de Gram-Schmidt da seção 2.7.

e

$$P^t(i,-) = P(-,i)^t, i = 1,\ldots,n,$$

temos

$$\begin{aligned} P^tP = \mathsf{I} &\iff P^t(i,-)\cdot P(-,j) = \begin{cases} 1 \text{ se } i=j \\ \text{e} \\ 0 \text{ se } i \neq j \end{cases} \\ &\iff \big\{P(-,1),\ldots,P(-,n)\big\} \text{ é uma base ortonormal de } \mathbb{R}^n. \end{aligned}$$

Analogamente,

$$PP^t = \mathsf{I} \iff \big\{P(1,-)^t,\ldots,P(n,-)^t\big\} \text{ é uma base ortonormal de } \mathbb{R}^n.$$

Finalmente, note que,

$$PP^t = \mathsf{I} \iff P^tP = \mathsf{I},$$

pois, ao aplicarmos propriedades de transposição de matrizes nessas igualdades, obtemos a validade da equivalência entre elas.

AFIRMAÇÃO 7

Seja $A \in \mathbb{R}^{n\times n}$. Demonstra-se que:

1. *A é ortogonalmente diagonalizável, isto é, existe uma matriz ortogonal P tal que $\mathsf{P}^tA\mathsf{P} = \mathsf{D}$ é diagonal $\iff$ A é simétrica, isto é, $A^t = A$.*[32]

2. *São reais os autovalores de A, caso essa matriz seja simétrica.*

3. *São ortogonais os autovetores associados à eventuais autovalores distintos de A, caso essa matriz seja simétrica.*

DEMONSTRAÇÃO DA AFIRMAÇÃO 7

1. A validade desse resultado, para operadores sobre espaços vetoriais reais, será demonstrada na subseção 7.4.7. Contudo, “para não passar em branco”, considere a seguinte demonstração parcial: A implicação “$\Longrightarrow$” segue de

$$\begin{aligned} A &= \mathsf{PDP}^t \\ &= \mathsf{PD}^t\mathsf{P}^t \\ &= \big(\mathsf{PDP}^t\big)^t \\ &= A^t. \end{aligned}$$

   Adicionalmente, será demonstrada uma forma mais fraca da implicação “$\Longleftarrow$” na subseção 5.3.2.

2. De fato, se $A\mathbf{x} = \lambda\mathbf{x}$, considerando propriedades de transposição de matrizes e conjugação complexa,[33]

[32]Esse resultado é chamado de *teorema espectral* para $A \in \mathbb{R}^{n\times n}$.

[33]Confira o capítulo 5.

temos:

$$\begin{aligned}
\lambda \overline{\mathbf{x}}^t \mathbf{x} &= \overline{\mathbf{x}}^t \lambda \mathbf{x} \\
&= \overline{\mathbf{x}}^t A \mathbf{x} \\
&= \overline{\mathbf{x}}^t A^t \mathbf{x} \\
&= (A \overline{\mathbf{x}})^t \, \mathbf{x} \\
&= \left(\overline{A} \overline{\mathbf{x}}\right)^t \mathbf{x} \\
&= \left(\overline{A \mathbf{x}}\right)^t \mathbf{x} \\
&= \left(\overline{\lambda \mathbf{x}}\right)^t \mathbf{x} \\
&= \left(\overline{\lambda} \overline{\mathbf{x}}\right)^t \mathbf{x} \\
&= \overline{\lambda} \overline{\mathbf{x}}^t \mathbf{x}.
\end{aligned}$$

Portanto, $\left(\lambda - \overline{\lambda}\right) \overline{\mathbf{x}}^t \mathbf{x} = 0$ e, como $\mathbf{x} \neq \mathbf{0}$, temos $\lambda - \overline{\lambda} = 0$, ou seja, $\lambda = \overline{\lambda}$. Assim, $\lambda \in \mathbb{R}$.

3. Sejam $A\mathbf{x}_1 = \lambda_1 \mathbf{x}_1$ e $A\mathbf{x}_2 = \lambda_2 \mathbf{x}_2$. Portanto,

$$\begin{aligned}
\lambda_1 (\mathbf{x}_1 \cdot \mathbf{x}_2) &= (\lambda_1 \mathbf{x}_1) \cdot \mathbf{x}_2 \\
&= (A\mathbf{x}_1) \cdot \mathbf{x}_2 \\
&= \mathbf{x}_2 \cdot (A\mathbf{x}_1) \\
&= \mathbf{x}_2^t (A\mathbf{x}_1) \\
&= \mathbf{x}_2^t \left(A^t \mathbf{x}_1\right) \\
&= \left(\mathbf{x}_2^t A^t\right) \mathbf{x}_1 \\
&= (A\mathbf{x}_2)^t \, \mathbf{x}_1 \\
&= (A\mathbf{x}_2) \cdot \mathbf{x}_1 \\
&= (\lambda_2 \mathbf{x}_2) \cdot \mathbf{x}_1 \\
&= \lambda_2 (\mathbf{x}_2 \cdot \mathbf{x}_1) \\
&= \lambda_2 (\mathbf{x}_1 \cdot \mathbf{x}_2)
\end{aligned}$$

onde alternamos entre produto interno, quando usamos o "$\cdot$", e produto de matrizes, quando não usamos o "$\cdot$". Então, $(\lambda_1 - \lambda_2)(\mathbf{x}_1 \cdot \mathbf{x}_2) = 0$. Assim, $\mathbf{x}_1 \cdot \mathbf{x}_2 = 0$ pois, por hipótese, $\lambda_1 \neq \lambda_2$.

EXEMPLOS

- Considere a matriz $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. Por ser simétrica, A é diagonalizável.[34] Contudo, P não é ortogonal, ou seja, $\mathsf{P}^{-1} \neq \mathsf{P}^t$, caso seja calculada como na subseção 4.2.1.[35] Assim, para podermos utilizar a afirmação 7, P^{-1} não será obtida por escalonamento. Para calcularmos os autovalores de A, observe que

$$\begin{aligned}
\det \begin{bmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{bmatrix} &= \lambda^2 - 2\lambda - 3 \\
&= (\lambda + 1)(\lambda - 3).
\end{aligned}$$

Portanto, esses autovalores são $\lambda_1 = -1$ e $\lambda_2 = 3$, o que acarreta

$$\mathsf{D} = \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}.$$

[34] Cf. item 1 da afirmação 7.

[35] Verifique!

Agora, ao resolvermos o sistema

$$(A+\mathsf{I})\mathbf{x}=\mathbf{0}, \text{ isto é, } \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

obtemos, por exemplo, o autovetor $\mathbf{x}_1=\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ associado à $\lambda_1=-1$. Analogamente, ao resolvermos

$$(A-3\mathsf{I})\mathbf{x}=\mathbf{0}, \text{ ou seja, } \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

obtemos, por exemplo, o autovetor $\mathbf{x}_2=\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ associado à $\lambda_2=3$.[36] Contudo, não escreveremos, como na subseção supracitada, $\mathsf{P}=\begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 \end{bmatrix}$, para depois obtermos P^{-1} via escalonamento.[37] No lugar desse procedimento, normalizando-se $\mathbf{x}_1$ e $\mathbf{x}_2$, temos

$$\mathbf{x}_1'=\frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix} \text{ e } \mathbf{x}_2'=\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Assim, podemos verificar que

$$\mathsf{P}=\begin{bmatrix} \mathbf{x}_1' & \mathbf{x}_2' \end{bmatrix} \Longrightarrow \mathsf{P}^t A\mathsf{P}=\mathsf{D}.$$

· Considere a matriz $A=\begin{bmatrix} 2 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 2 \end{bmatrix}$. Por ser simétrica, A é diagonalizável. Contudo, como no exemplo anterior, P^{-1} não será obtida pelo escalonamento de P.
Para obtermos os autovalores de A, observe que

$$\det\begin{bmatrix} 2-\lambda & 2 & 1 \\ 2 & 5-\lambda & 2 \\ 1 & 2 & 2-\lambda \end{bmatrix}=(\lambda-7)\,(\lambda-1)^2.$$

Logo, esses autovalores são $\lambda_1=7$ e $\lambda_2=\lambda_3=1$, o que acarreta

$$\mathsf{D}=\begin{bmatrix} 7 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Agora, pelo sistema

$$(A-7\mathsf{I})\mathbf{x}=\mathbf{0}, \text{ ou seja, } \begin{bmatrix} -5 & 2 & 1 \\ 2 & -2 & 2 \\ 1 & 2 & -5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

[36] É importante observarmos que, como estabelecido no item 3 da afirmação 7 e pelo fato de A ser simétrica, $\mathbf{x}_1$ e $\mathbf{x}_2$ são ortogonais.

[37] Todavia, a opção do método a ser utilizado é sua!

obtemos, por exemplo, o autovetor $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ associado à $\lambda_1 = 7$. Analogamente, via

$$(A - \mathsf{I})\mathbf{x} = \mathbf{0}, \text{ ou seja, } \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

obtemos, por exemplo, os autovetores $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ e $\mathbf{x}_3 = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$ associados ao autovalor $\lambda_2 = \lambda_3 = 1$.[38] Agora, como no exemplo anterior, não escalonaremos a matriz $[\mathsf{P} \,|\, \mathsf{I}]$, com $\mathsf{P} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \end{bmatrix}$, para determinarmos P^{-1}. No lugar desse procedimento, vamos considerar

$$\mathbf{x}_1' = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix},$$

obtido ao normalizarmos $\mathbf{x}$, e

$$\mathbf{x}_2'' = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad \text{e} \quad \mathbf{x}_3'' = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix},$$

obtidos ao normalizarmos os vetores calculados via Gram-Schmidt em $\mathbf{x}_2$ e $\mathbf{x}_3$. Assim, podemos verificar que

$$\mathsf{P} = \begin{bmatrix} \mathbf{x}_1' & \mathbf{x}_2'' & \mathbf{x}_3'' \end{bmatrix} \Longrightarrow \mathsf{P}^t A \mathsf{P} = \mathsf{D}.$$

[38] É importante observarmos que, como estabelecido no item 3 da afirmação 7, a simetria de A garante que $\mathbf{x}_1$ e $\mathbf{x}_i$ são ortogonais, $i = 2, 3$. Contudo, note que, $\mathbf{x}_2$ e $\mathbf{x}_3$ não são ortogonais entre si.

4.3 Exercícios

1. Seja $L : \mathbb{R}^2 \to \mathbb{R}^2$ linear tal que

$$L(\mathbf{e}_1) = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{e} \quad L(\mathbf{e}_2) = \begin{bmatrix} -3 \\ 4 \end{bmatrix}.$$

Determine $L(\mathbf{x})$ para $\mathbf{x} = \begin{bmatrix} -5 \\ 6 \end{bmatrix}$.

1A. RESOLUÇÃO

$$\begin{aligned} L(\mathbf{x}) &= L(-5\mathbf{e}_1 + 6\mathbf{e}_2) \\ &= (-5)L(\mathbf{e}_1) + 6L(\mathbf{e}_2) \qquad \text{(LINEARIDADE)} \\ &= (-5)\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 6\begin{bmatrix} -3 \\ 4 \end{bmatrix} \\ &= \begin{bmatrix} -23 \\ 14 \end{bmatrix}. \end{aligned}$$

2A. RESOLUÇÃO

$L = L_A$ com

$$\begin{aligned} A &= \begin{bmatrix} L(\mathbf{e}_1) & L(\mathbf{e}_2) \end{bmatrix} \\ &= \begin{bmatrix} 1 & -3 \\ 2 & 4 \end{bmatrix}. \end{aligned}$$

Então,

$$\begin{aligned} L(\mathbf{x}) &= A\mathbf{x} \\ &= \begin{bmatrix} -23 \\ 14 \end{bmatrix}. \end{aligned}$$

2. Seja $L = L_A : \mathbb{R}^2 \to \mathbb{R}^2$ tal que

$$L(\mathbf{u}) = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{e} \quad L(\mathbf{v}) = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

para

$$\mathbf{u} = \begin{bmatrix} 5 \\ 0 \end{bmatrix} \quad \text{e} \quad \mathbf{v} = \begin{bmatrix} 6 \\ 7 \end{bmatrix}.$$

Determine A.

1A. RESOLUÇÃO

Por um lado,

$$\begin{aligned} L(\mathbf{e}_1) &= L\left(\frac{1}{5}\mathbf{u}\right) \\ &= \frac{1}{5}L(\mathbf{u}) \quad \text{(LINEARIDADE)} \\ &= \begin{bmatrix} 1/5 \\ 2/5 \end{bmatrix}. \end{aligned}$$

Por outro, como

$$\begin{aligned} L(\mathbf{v}) &= L\left(6\mathbf{e}_1 + 7\mathbf{e}_2\right) \\ &= 6L\left(\mathbf{e}_1\right) + 7L\left(\mathbf{e}_2\right) \quad (\text{LINEARIDADE}) \\ &= \begin{bmatrix} 6/5 \\ 12/5 \end{bmatrix} + 7L\left(\mathbf{e}_2\right), \end{aligned}$$

$$\begin{aligned} L\left(\mathbf{e}_2\right) &= \frac{1}{7}\left\{\begin{bmatrix} 3 \\ 4 \end{bmatrix} - \begin{bmatrix} 6/5 \\ 12/5 \end{bmatrix}\right\} \\ &= \begin{bmatrix} 9/35 \\ 8/35 \end{bmatrix}. \end{aligned}$$

Então,

$$\begin{aligned} A &= \begin{bmatrix} L\left(\mathbf{e}_1\right) & L\left(\mathbf{e}_2\right) \end{bmatrix} \\ &= \begin{bmatrix} 1/5 & 9/35 \\ 2/5 & 8/35 \end{bmatrix}. \end{aligned}$$

2A. RESOLUÇÃO

Seja

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

Portanto,

$$\begin{aligned} A\mathbf{u} = L(\mathbf{u}) &\Longrightarrow \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 5 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ &\Longrightarrow a = \frac{1}{5} \text{ e } c = \frac{2}{5}. \end{aligned}$$

Analogamente,

$$\begin{aligned} A\mathbf{v} = L(\mathbf{v}) &\Longrightarrow \begin{bmatrix} 1/5 & b \\ 2/5 & d \end{bmatrix}\begin{bmatrix} 6 \\ 7 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \\ &\Longrightarrow b = \frac{9}{35} \text{ e } d = \frac{8}{35}. \end{aligned}$$

Assim, A é dada como na 1a. resolução desse exercício.

3. Para cada vetor $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$, defina

$$L(\mathbf{x}) = x_1 + x_2 + x_3.$$

   (a) Verifique que $L : \mathbb{R}^3 \longrightarrow \mathbb{R}$ é linear obtendo a matriz A tal que $L = L_A$;

   (b) Determine uma base para o núcleo de L.

RESOLUÇÃO

(a) Caso A seja essa matriz, sua i-ésima coluna é a imagem por L do i-ésimo vetor da base canônica de $\mathbb{R}^3$, $i = 1, 2, 3$, ou seja,

$$A = \begin{bmatrix} L(\mathbf{e}_1) & L(\mathbf{e}_2) & L(\mathbf{e}_3) \end{bmatrix},$$

onde $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ é a base canônica de $\mathbb{R}^3$. Portanto, como $L(\mathbf{e}_i) = 1$, $i = 1, 2, 3$, temos

$$A = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}.$$

Além disso, note que, para cada $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$,

$$\begin{aligned} L(\mathbf{x}) &= x_1 + x_2 + x_3 \\ &= \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ &= A\mathbf{x}. \end{aligned}$$

(b) Como $\mathrm{Nu}(L) = \{\mathbf{x} \in \mathbb{R}^3 \, : \, A\mathbf{x} = 0\}$,

$$\mathbf{x} \in \mathrm{Nu}(L) \iff x_1 + x_2 + x_3 = 0.$$

Então, para escalares reais $x_2 = \alpha$ e $x_3 = \beta$ arbitrários,

$$\begin{aligned} \mathrm{Nu}(L) \ni \mathbf{x} &= (x_1, x_2, x_3) \\ &= (-\alpha - \beta, \alpha, \beta) \\ &= \alpha(-1, 1, 0) + \beta(-1, 0, 1). \end{aligned}$$

Assim, $\{(-1, 1, 0), (-1, 0, 1)\}$ é uma base de $\mathrm{Nu}(L)$.

4. Caso a matriz

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & a_{23} & 1 & 1 \\ 0 & 0 & 0 & a_{34} & 1 \end{bmatrix}$$

tenha posto 2, determine todos os valores numéricos que a_{23} e a_{34} podem assumir e, sem recorrer ao método de *tentativa e erro*, justifique os valores obtidos.

RESOLUÇÃO

O posto de A é a dimensão do espaço gerado por suas colunas e é determinado pelo número de colunas pivôs de R, sua escalonada reduzida.[39] Assim, os valores supracitados devem ser ajustados de modo que R tenha duas colunas pivôs. Logo, para que a primeira etapa do escalonamento de A seja dada por

$$A \longrightarrow \begin{bmatrix} 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1/a_{23} & 1/a_{23} \\ 0 & 0 & 0 & a_{34} & 1 \end{bmatrix},$$

a hipótese $a_{23} \neq 0$ deve ser considerada. Nesse caso, independentemente do valor que a_{34} possa assumir, R terá três colunas pivôs. De fato, se $a_{34} \neq 0$, então as colunas 1, 3 e 4 de R serão as pivôs. Por outro lado, se $a_{34} = 0$, as colunas 1, 3 e 5 de R serão as pivôs. Então, apenas $a_{23} = 0$ pode resultar nas duas colunas pivôs supracitadas e, nesse caso,

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & a_{34} & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 - a_{34} \end{bmatrix},$$

onde $1 - a_{34} \neq 0$ resultará em três colunas pivôs: a primeira, a quarta e a quinta. Portanto,

$$A \text{ tem posto } 2 \iff a_{23} = 0 \text{ e } a_{34} = 1.$$

5. Se duas matrizes de mesmo tamanho, digamos $A, B \in \mathbb{R}^{m \times n}$, têm o mesmo posto, então tais matrizes são equivalentes por linha. Essa afirmação é verdadeira ou falsa? Se for verdadeira, demonstre-a; caso contrário, apresente um contraexemplo.

RESOLUÇÃO

Falsa, pois, por exemplo, as escalonadas reduzidas $A = [1 \; 0]$ e $B = [0 \; 1]$, ambas 1×2, têm posto 1 mas não são equivalentes por linha.

[39] Cf. (4.6), p. 92.

6. Obter uma base para o subespaço do $\mathbb{R}^4$ gerado por $(2,1,1,0)$, $(3,0,3,6)$, $(1,0,1,2)$ e $(0,-1,1,4)$.[40]

7. Caso $L : \mathbb{R}^2 \to \mathbb{R}^3$ seja linear,

$$[L]^{\mathcal{B}}_{\mathcal{B}'} = \begin{bmatrix} 2 & 1 \\ -1 & 0 \\ 0 & -2 \end{bmatrix},$$

   $\mathcal{B} = \{(0,1),(1,-1)\}$, $\mathcal{B}' = \{(1,2,3),(1,0,-1),(0,0,2)\}$ e $\mathbf{x} = (-1,2)$ na base canônica, obtenha as coordenadas de $L(\mathbf{x})$ na base $\mathcal{B}'$, isto é, determine $[L(\mathbf{x})]_{\mathcal{B}'}$.[41]

8. Caso $L : \mathbb{R}^3 \to \mathbb{R}^3$ seja linear,

$$\begin{aligned} A &= [L]_{\mathcal{B}} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \end{aligned}$$

   $\mathcal{B} = \left\{ \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{x}_3 = \mathbf{e}_2 \right\}$ e $\mathcal{B}' = \{\mathbf{y}_1 = \mathbf{e}_1, \mathbf{y}_2 = \mathbf{e}_2, \mathbf{y}_3 = \mathbf{e}_3\}$, determine $A' = [L]_{\mathcal{B}'}$.

   RESOLUÇÃO

   Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$ a função linear tal que

$$\begin{aligned} \mathbf{x}_1 &\overset{T}{\mapsto} \mathbf{y}_1; \\ \mathbf{x}_2 &\overset{T}{\mapsto} \mathbf{y}_2; \\ \mathbf{x}_3 &\overset{T}{\mapsto} \mathbf{y}_3. \end{aligned}$$

   Portanto, $A' = \mathsf{P}^{-1} A \mathsf{P}$ com $\mathsf{P} = [T]_{\mathcal{B}}$.[42] Para obtermos P, vamos considerar

$$\begin{cases} \mathbf{x}_1 \overset{T}{\mapsto} t_{11}\mathbf{x}_1 + t_{21}\mathbf{x}_2 + t_{31}\mathbf{x}_3 = \mathbf{y}_1; & (\mathscr{S}_1) \\ \mathbf{x}_2 \overset{T}{\mapsto} t_{12}\mathbf{x}_1 + t_{22}\mathbf{x}_2 + t_{32}\mathbf{x}_3 = \mathbf{y}_2; & (\mathscr{S}_2) \\ \mathbf{x}_3 \overset{T}{\mapsto} t_{13}\mathbf{x}_1 + t_{23}\mathbf{x}_2 + t_{33}\mathbf{x}_3 = \mathbf{y}_3. & (\mathscr{S}_3) \end{cases}$$

   Nesse caso, como

$$\mathsf{P} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix},$$

[40] Faça como no exercício que precede a subseção 4.1.2. Observe que, dos quatro vetores dados, os dois primeiros determinam a base procurada.

[41] Basta utilizarmos a fórmula (4.8), da página 97, e observarmos que

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

e $[L]^{\mathcal{B}}_{\mathcal{B}'}$ é dada no enunciado desse exercício.

[42] Cf. (4.9), (4.10) e (4.11), pp. 97–98.

temos que determinar todos os t_{ij}'s. Assim, como $M = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \end{bmatrix}$ é a matriz do sistema $(\mathscr{S}_k)$, $k = 1, 2, 3$, o escalonamento

$$
\begin{aligned}
[M|\mathsf{I}] = \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] &\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & -1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \\
&\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & -1 & 1 & -1 & 1 & 0 \end{array}\right] \\
&\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right] = [\mathsf{I}|\mathsf{P}]
\end{aligned}
$$

resolve (simultaneamente) os três sistemas supracitados,[43] ou seja, determina P, e (sem qualquer esforço adicional) calcula a inversa de P.[44] Portanto,

$$
\begin{aligned}
A' &= \mathsf{P}^{-1} A \mathsf{P} \\
&= \begin{bmatrix} 0 & 1 & 0 \\ 2 & 0 & -1 \\ -1 & 1 & 1 \end{bmatrix}.
\end{aligned}
$$

9. Para a matriz

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

determine seus autovalores e autovetores associados.

RESOLUÇÃO

Os autovalores $\lambda_1 = 1$ e $\lambda_2 = 5$ de A são as raízes da equação

$$
\begin{vmatrix} 4-\lambda & \sqrt{3} \\ \sqrt{3} & 2-\lambda \end{vmatrix} = \lambda^2 - 6\lambda + 5 = 0.
$$

Os autovetores associados à $\lambda_1 = 1$ são múltiplos de

$$
\mathbf{x}_1 = \begin{bmatrix} -\sqrt{3} \\ 3 \end{bmatrix}.
$$

De fato,

$$
\begin{aligned}
\begin{bmatrix} 4-\lambda_1 & \sqrt{3} \\ \sqrt{3} & 2-\lambda_1 \end{bmatrix} = \begin{bmatrix} 3 & \sqrt{3} \\ \sqrt{3} & 1 \end{bmatrix} &\longrightarrow \begin{bmatrix} 3 & \sqrt{3} \\ 3 & \sqrt{3} \end{bmatrix} \\
&\longrightarrow \begin{bmatrix} 1 & \frac{\sqrt{3}}{3} \\ 0 & 0 \end{bmatrix}.
\end{aligned}
$$

Os autovetores associados à $\lambda_2 = 5$ são múltiplos de

$$
\mathbf{x}_2 = \begin{bmatrix} \sqrt{3} \\ 1 \end{bmatrix}.
$$

[43] As quatro primeiras colunas de cada matriz (do escalonamento supracitado) têm relação com a resolução do sistema $(\mathscr{S}_1)$, cuja solução é $(t_{11}, t_{21}, t_{31}) = (1, 0, -1)$. As três primeiras e a quinta colunas de cada matriz estão relacionadas com a resolução do sistema $(\mathscr{S}_2)$, cuja solução é $(t_{12}, t_{22}, t_{32}) = (0, 0, 1)$. As três primeiras e a última colunas de cada matriz estão relacionadas com a resolução do sistema $(\mathscr{S}_3)$, cuja solução é $(t_{13}, t_{23}, t_{33}) = (-1, 1, 1)$.

[44] $\mathsf{P}^{-1} = M$.

De fato,

$$\begin{bmatrix} 4-\lambda_2 & \sqrt{3} \\ \sqrt{3} & 2-\lambda_2 \end{bmatrix} = \begin{bmatrix} -1 & \sqrt{3} \\ \sqrt{3} & -3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -\sqrt{3} \\ 1 & -\sqrt{3} \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -\sqrt{3} \\ 0 & 0 \end{bmatrix}.$$

10. Quais os autovalores e autovetores da matriz $A = \begin{bmatrix} 2 & 5 \\ 3 & -4 \end{bmatrix}$?

RESOLUÇÃO

Os autovalores $\lambda_{1,2} = -1 \pm 2\sqrt{6}$ de A são as raízes da equação

$$\begin{vmatrix} 2-\lambda & 5 \\ 3 & -4-\lambda \end{vmatrix} = \lambda^2 + 2\lambda - 23 = 0.$$

Os autovetores associados à $\lambda_1 = -1 + 2\sqrt{6}$ são múltiplos de

$$\mathbf{x}_1 = \begin{bmatrix} -\frac{5}{3-2\sqrt{6}} \\ 1 \end{bmatrix}.$$

De fato,

$$\begin{bmatrix} 2-\lambda_1 & 5 \\ 3 & -4-\lambda_1 \end{bmatrix} = \begin{bmatrix} 3-2\sqrt{6} & 5 \\ 3 & -\left(3+2\sqrt{6}\right) \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & \frac{5}{3-2\sqrt{6}} \\ 3 & -\left(3+2\sqrt{6}\right) \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & \frac{5}{3-2\sqrt{6}} \\ 0 & 0 \end{bmatrix}.$$

Os autovetores associados à $\lambda_2 = -1 - 2\sqrt{6}$ são múltiplos de

$$\mathbf{x}_2 = \begin{bmatrix} -\frac{5}{3+2\sqrt{6}} \\ 1 \end{bmatrix}.$$

De fato,

$$\begin{bmatrix} 2-\lambda_2 & 5 \\ 3 & -4-\lambda_2 \end{bmatrix} = \begin{bmatrix} 3+2\sqrt{6} & 5 \\ 3 & -\left(3-2\sqrt{6}\right) \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & \frac{5}{3+2\sqrt{6}} \\ 3 & -\left(3-2\sqrt{6}\right) \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & \frac{5}{3+2\sqrt{6}} \\ 0 & 0 \end{bmatrix}.$$

11. $\mathbf{v}_1 = \left(\frac{1}{2}, 1, 0\right)$, $\mathbf{v}_2 = \left(-\frac{1}{2}, -\frac{3}{2}, 1\right)$ e $\mathbf{v}_3 = \left(-\frac{1}{2}, 0, 1\right)$ são autovetores da matriz

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

Com essa informação, resolva os seguintes itens:

- Determine os autovalores correspondentes aos autovetores supracitados.[45]
- A é diagonalizável? Justifique (corretamente) a sua resposta.

RESOLUÇÃO

Pela equação (4.12),[46] $A\mathbf{v}_1 = \lambda_1\mathbf{v}_1$, $A\mathbf{v}_2 = \lambda_2\mathbf{v}_2$ e $A\mathbf{v}_3 = \lambda_3\mathbf{v}_3$. Assim, se $\mathbf{a}_1 = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right)$,[47] então

$$\begin{aligned}\mathbf{a}_1 \cdot \mathbf{v}_1 &= \lambda_1\mathbf{v}_1,\\ \mathbf{a}_1 \cdot \mathbf{v}_2 &= \lambda_2\mathbf{v}_2 \text{ e}\\ \mathbf{a}_1 \cdot \mathbf{v}_3 &= \lambda_3\mathbf{v}_3.\end{aligned}$$

Portanto, como

$$\begin{aligned}\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{4}\right)(1) &= \lambda_1\left(\frac{1}{2}\right),\\ \left(\frac{1}{2}\right)\left(-\frac{1}{2}\right) + \left(\frac{1}{4}\right)\left(-\frac{3}{2}\right) + \left(\frac{1}{4}\right)(1) &= \lambda_2\left(-\frac{1}{2}\right) \quad \text{e}\\ \left(\frac{1}{2}\right)\left(-\frac{1}{2}\right) + \left(\frac{1}{4}\right)(1) &= \lambda_3\left(-\frac{1}{2}\right),\end{aligned}$$

A tem os seguintes autovalores: $\lambda_1 = 1$, $\lambda_2 = \frac{3}{4}$ e $\lambda_3 = 0$. Por outro lado, para que A seja diagonalizável, basta que o conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ seja uma base de $\mathbb{R}^3$.[48] De fato, os três vetores desse conjunto são LI, pela não nulidade do determinante da matriz com linhas (ou colunas) dadas por esses vetores.[49] Outra maneira de demonstrarmos a independência linear dos vetores supracitados é utilizarmos o resultado (R20) do capítulo 7.[50] De fato, o conjunto supracitado é composto por autovetores associados aos três autovalores distintos de $A \in \mathbb{R}^{3\times 3}$.

12. A matriz

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

não é diagonalizável. Essa afirmação é verdadeira ou falsa? Justifique (corretamente) a sua resposta.

RESOLUÇÃO

Como estamos lidando com uma matriz triangular, temos

$$\det(A - \lambda\mathsf{I}) = (2-\lambda)(1+\lambda)^2,$$

onde I é a matriz identidade. Portanto, $\lambda_1 = 2$ e $\lambda_2 = \lambda_3 = -1$ são os autovalores de A. Por outro lado, A é diagonalizável se, e somente se, existe uma base para o $\mathbb{R}^3$ formada por autovetores de A. Para verificarmos se isso é possível, devemos resolver os sistemas

$$(A - 2\mathsf{I})\mathbf{x} = \mathbf{0} \ \text{ e } \ (A + \mathsf{I})\mathbf{x} = \mathbf{0}.$$

[45] SUGESTÃO
Não obtenha o polinômio característico de A. Utilize a definição de autovalores.

[46] Cf. p. 100.

[47] $\mathbf{a}_i$ representa a i-ésima linha de A, $i = 1, 2, 3$.

[48] Cf. a primeira sentença da subseção 4.2.1.

[49] Cf. o penúltimo exercício do capítulo 3.

[50] Cf. p. 180.

Assim, dos escalonamentos

$$A - 2\mathsf{I} = \begin{bmatrix} 0 & 1 & -1 \\ 0 & -3 & 2 \\ 0 & 0 & -3 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

e

$$A + \mathsf{I} = \begin{bmatrix} 3 & 1 & -1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1/3 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

temos que os autoespaços de A são determinados por

$$\mathcal{S}_2 \ni \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

e

$$\mathcal{S}_{-1} \ni \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3}x_2 \\ x_2 \\ 0 \end{bmatrix} = x_2 \begin{bmatrix} -1/3 \\ 1 \\ 0 \end{bmatrix}.$$

Portanto, como não é possível obtermos três autovetores LI de A, ou seja, uma base do $\mathbb{R}^3$ formada por autovetores de A, a afirmação do enunciado do exercício é verdadeira..

13. Caso seja possível, diagonalize as seguintes matrizes:

(a) $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$;[51]

(b) $A = \begin{bmatrix} -1 & 6 & -12 \\ 0 & -13 & 30 \\ 0 & -9 & 20 \end{bmatrix}$;[52]

---

[51] DICA
Em relação a A, verifique que: $(-1,1,0)$ e $(-1,0,1)$ são autovetores associados ao autovalor -1; $(1,1,1)$ é autovetor associado ao autovalor 2.

[52] DICA
Em relação a A, verifique que -1, 2 e 5 são seus autovalores.

(c) $A = \begin{bmatrix} 5 & -2 & 0 \\ -2 & 6 & 2 \\ 0 & 2 & 7 \end{bmatrix}$.[53]

14. Em (R20),[54] demonstraremos que *autovetores associados à autovalores distintos são* LI. Com essa informação, obtenha algum valor de x tal que

$$A = \begin{bmatrix} D+2 & C+1 & C+1 & C+1 \\ 0 & E+3 & x & C+1 \\ 0 & 0 & E+3 & C+1 \\ 0 & 0 & 0 & C+1 \end{bmatrix}$$

seja diagonalizável. Aqui, C, D e E representam números reais arbitrários com $C \neq -1$ e $C+1$, $D+2$ e $E+3$ distintos entre si.[55]

RESOLUÇÃO

Como estamos lidando com uma matriz triangular, temos

$$\det(A - \lambda \mathsf{I}) = (\lambda - (D+2))(\lambda - (E+3))^2(\lambda - (C+1)),$$

onde I é a matriz identidade. Portanto, $\lambda_1 = D+2$, $\lambda_2 = \lambda_3 = E+3$ e $\lambda_4 = C+1$ são os autovalores de A. Por outro lado, A é diagonalizável se, e somente se, existe uma base $\mathcal{B}$ do $\mathbb{R}^4$ formada por autovetores de A. Bom, pelo fato citado no enunciado da questão, já temos (pelo menos) três autovetores LI dessa $\mathcal{B}$: um associado ao autovalor $D+2$, outro associado ao autovalor $E+3$ e um associado ao autovalor $C+1$. Então, para que A seja diagonalizável, basta que existam dois autovetores LI de A associados ao autovalor $\lambda = E+3$ (cuja multiplicidade é dois).[56] Considere, assim, o sistema

$$(A - (E+3)\mathsf{I})\mathbf{x} = \mathbf{0} \tag{4.13}$$

e escalone a sua matriz para tentar obter os dois autovetores LI procurados. Portanto, como $C+1 \neq 0$,

$$\begin{bmatrix} D-E-1 & C+1 & C+1 & C+1 \\ 0 & 0 & x & C+1 \\ 0 & 0 & 0 & C+1 \\ 0 & 0 & 0 & C-E-2 \end{bmatrix} \longrightarrow \begin{bmatrix} D-E-1 & C+1 & C+1 & C+1 \\ 0 & 0 & x & C+1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & C-E-2 \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} D-E-1 & C+1 & C+1 & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} (D-E-1)/(C+1) & 1 & 1 & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

e a próxima etapa desse escalonamento resulta em uma das seguintes matrizes:

- $\begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ caso $D - E = 1$;

---

[53] DICA

Em relação a A, verifique que 3, 6 e 9 são seus autovalores.

[54] Cf. p. 180.

[55] Confira a nota de rodapé do quarto exemplo dado após o resultado (R34), p. 194.

[56] Esse argumento não vale para $\lambda \in \{C+1, D+2\}$. De fato, confira o item (1) dos avisos importantes que precedem a subseção 4.2.1.

$$\cdot \begin{bmatrix} 1 & (C+1)/(D-E-1) & (C+1)/(D-E-1) & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{caso } D - E \neq 1.$$

Portanto, a matriz do sistema terá posto 2 apenas se $x = 0$.[57]

Em suma, para que o espaço solução do sistema (4.13) tenha uma base formada por dois autovetores LI de A, ou seja, para que a nulidade da matriz $A - (E+3)\mathsf{I}$ seja dois, isto é, para que o posto de $A - (E+3)\mathsf{I}$ seja dois,[58] x deve ser igual a zero.

15. Seja $A \in \mathbb{R}^{n \times n}$ arbitrária e tal que $A = A^2$. Verifique que apenas 0 e 1 podem ser autovalores de A.

SUGESTÕES PARA DUAS RESOLUÇÕES DISTINTAS

1a. Relacione a definição de autovalores/autovetores, dada na página 100, com a condição $A^2 = A$ supracitada.

2a. Reescreva $A = A^2$ da forma $A - A^2 = \mathsf{O}$ e coloque A em evidência. Além disso, utilize o fato de que o determinante do produto de matrizes é o produto dos determinantes dessas matrizes.

1A. RESOLUÇÃO

Seja $\mathbf{x}$ autovetor de A associado ao autovalor λ. Como

$$A\mathbf{x} = \lambda \mathbf{x},$$

$$\begin{aligned} A^2\mathbf{x} &= AA\mathbf{x} \\ &= A(\lambda\mathbf{x}) \\ &= \lambda A\mathbf{x} \\ &= \lambda\lambda\mathbf{x} \\ &= \lambda^2\mathbf{x} \end{aligned}$$

e $A\mathbf{x} = A^2\mathbf{x}$ (pois $A = A^2$), temos que

$$\begin{aligned} \lambda\mathbf{x} = \lambda^2\mathbf{x} &\implies \left(\lambda - \lambda^2\right)\mathbf{x} = \mathbf{0} \\ &\implies \lambda - \lambda^2 = 0 \text{ pois } \mathbf{x} \neq \mathbf{0} \\ &\implies \lambda(1-\lambda) = 0 \\ &\implies \lambda = 0 \text{ ou } \lambda = 1. \end{aligned}$$

2A. RESOLUÇÃO

Primeiramente, lembre-se que, caso λ seja autovalor de A, a equação

$$\det(A - \lambda\mathsf{I}) = 0 \tag{4.14}$$

determina esse autovalor. Assim, como

$$\begin{aligned} \det\left(A - A^2\right) = \det\mathsf{O} &\implies \det(A(A - \mathsf{I})) = 0 \\ &\implies \det A \cdot \det(A - \mathsf{I}) = 0 \\ &\implies \det(A - 0\mathsf{I}) \cdot \det(A - 1\mathsf{I}) = 0 \\ &\implies \det(A - 0\mathsf{I}) = 0 \;\text{ ou }\; \det(A - 1\mathsf{I}) = 0 \\ &\implies \lambda \in \{0, 1\} \;\text{ é autovalor de } A \end{aligned}$$

[57] Lembre-se que o posto de uma matriz é o número de suas colunas pivôs e representa a dimensão do espaço gerado pelas colunas dessa matriz.

[58] Lembre-se que, para qualquer matriz $m \times n$, posto + nulidade $= n$.

com λ como na equação (4.14).

16. Seja $A \in \mathbb{R}^{n\times n}$ arbitrária. Considere que A^2 tem um autovalor não negativo λ^2. Demonstre que, $\pm\lambda$ são os autovalores de A.

RESOLUÇÃO

$$\begin{aligned}\lambda^2 \text{ é autovalor de } A^2 &\Longrightarrow \det\left(A^2-\lambda^2\mathsf{I}\right)=0\\ &\Longrightarrow \det[(A+\lambda\mathsf{I})(A-\lambda\mathsf{I})]=0\\ &\Longrightarrow \det(A+\lambda\mathsf{I})\det(A-\lambda\mathsf{I})=0,\end{aligned}$$

pois o determinante do produto é igual ao produto dos determinantes. Então, ao menos um dos dois fatores do último produto, de cima para baixo, deve ser nulo. Portanto, $\mp\lambda$ são os autovalores de A.

4.3.1 Resoluções de alguns exercícios que precedem a subseção 4.1.1

EX. 2

$$A=\begin{bmatrix}2&1&0&0&0&0\\0&0&4&-3&0&0\\0&0&0&0&-6&5\\-1&2&3&4&-5&6\\1&0&0&0&0&0\\0&-1&0&0&0&0\\0&0&1&0&0&0\\0&0&0&-1&0&0\end{bmatrix}\Longrightarrow L=L_A$$
$$\Longrightarrow L \text{ é linear pela afirmação 1.}$$

EX. 5.(a)

$$A_x=\begin{bmatrix}1&0\\0&0\end{bmatrix} \text{ e } A_y=\begin{bmatrix}0&0\\0&1\end{bmatrix}\Longrightarrow P_x=L_{A_x} \text{ e } P_y=L_{A_y}$$
$$\Longrightarrow P_x \text{ e } P_y \text{ são lineares pela afirmação 1.}$$

EX. 6.(a)

$$A_x=\begin{bmatrix}1&0\\0&-1\end{bmatrix} \text{ e } A_y=\begin{bmatrix}-1&0\\0&1\end{bmatrix}\Longrightarrow R_x=L_{A_x} \text{ e } R_y=L_{A_y}$$
$$\Longrightarrow R_x \text{ e } R_y \text{ são lineares pela afirmação 1.}$$

EX. 5.(b), partes i. e ii., e EX. 6.(b)

Como vetores colunas, temos

$$\mathbf{x}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix},\ \mathbf{a}=\begin{bmatrix}1/\sqrt{3}\\1/\sqrt{3}\\1/\sqrt{3}\end{bmatrix},\ \mathbf{a}_1=\begin{bmatrix}1/\sqrt{2}\\0\\-1/\sqrt{2}\end{bmatrix} \text{ e } \mathbf{a}_2=\begin{bmatrix}-1/\sqrt{6}\\2/\sqrt{6}\\-1/\sqrt{6}\end{bmatrix}.$$

Logo,

$$\begin{aligned}
P_{\mathcal{S}}(\mathbf{x}) &= (\mathbf{x}\cdot\mathbf{a})\mathbf{a} \\
&= \frac{x_1+x_2+x_3}{\sqrt{3}}\begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \\
&= \begin{bmatrix} \frac{x_1+x_2+x_3}{3} \\ \frac{x_1+x_2+x_3}{3} \\ \frac{x_1+x_2+x_3}{3} \end{bmatrix} \\
&= \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\
&= A_{\mathcal{S}}\mathbf{x}
\end{aligned}$$

e

$$\begin{aligned}
P_{\mathcal{S}^\perp}(\mathbf{x}) &= (\mathbf{x}\cdot\mathbf{a}_1)\mathbf{a}_1 + (\mathbf{x}\cdot\mathbf{a}_2)\mathbf{a}_2 \\
&= \frac{x_1-x_3}{\sqrt{2}}\begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix} + \left(\frac{-x_1+2x_2-x_3}{\sqrt{6}}\right)\begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix} \\
&= \begin{bmatrix} \frac{x_1-x_3}{2} \\ 0 \\ \frac{x_3-x_1}{2} \end{bmatrix} + \begin{bmatrix} \frac{x_1-2x_2+x_3}{6} \\ \frac{-x_1+2x_2-x_3}{3} \\ \frac{x_1-2x_2+x_3}{6} \end{bmatrix} \\
&= \begin{bmatrix} \frac{4x_1-2x_2-2x_3}{6} \\ \frac{-x_1+2x_2-x_3}{3} \\ \frac{-2x_1-2x_2+4x_3}{6} \end{bmatrix} \\
&= \begin{bmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\
&= A_{\mathcal{S}^\perp}\mathbf{x}.
\end{aligned}$$

Agora, basta observarmos que

$$\begin{aligned}
R_{\mathcal{S}}(\mathbf{x}) &= 2P_{\mathcal{S}}(\mathbf{x}) - \mathbf{x} \\
&= 2A_{\mathcal{S}}\mathbf{x} - \mathsf{I}\mathbf{x} \\
&= (2A_{\mathcal{S}} - \mathsf{I})\,\mathbf{x}
\end{aligned}$$

e

$$\begin{aligned}
R_{\mathcal{S}^\perp}(\mathbf{x}) &= 2P_{\mathcal{S}^\perp}(\mathbf{x}) - \mathbf{x} \\
&= 2A_{\mathcal{S}^\perp}\mathbf{x} - \mathsf{I}\mathbf{x} \\
&= (2A_{\mathcal{S}^\perp} - \mathsf{I})\,\mathbf{x},
\end{aligned}$$

com I representando a matriz identidade 3×3, e calcularmos as matrizes $2A_{\mathcal{S}} - \mathsf{I}$ e $2A_{\mathcal{S}^\perp} - \mathsf{I}$.

Ex. I

Por um lado,

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

são tais que

$$\mathbb{R}^2 \ni \mathbf{x} \mapsto L_1(\mathbf{x}) = A\mathbf{x} \in \mathbb{R}^3 \quad \text{e} \quad \mathbb{R}^3 \ni \mathbf{y} \mapsto L_2(\mathbf{y}) = B\mathbf{y} \in \mathbb{R}^2.$$

Por outro lado,

$$
\begin{aligned}
\mathbb{R}^2 \ni \mathbf{x} \mapsto L_2\left(L_1(\mathbf{x})\right) &= L_2\left(\begin{bmatrix} x_2 \\ x_1 \\ x_1 + x_2 \end{bmatrix}\right) \\
&= \begin{bmatrix} x_1 + x_2 \\ 2x_1 + x_2 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\
&= C\mathbf{x}.
\end{aligned}
$$

Finalmente, basta verificarmos que

$$BA = C.$$

Capítulo 5

Os Espaços vetoriais $\mathbb{K}^n$ e $\mathbb{K}^{m\times n}$

5.1 Definição e propriedades do $(\mathbb{K}^n, +, \cdot)$

5.1.1 Pequena revisão de $\mathbb{C}$, o corpo dos números complexos

Como vimos no capítulo anterior, existem matrizes sem autovalores (reais). Nesse capítulo,[1] veremos que a matriz de (5.2) não tem autovalores reais, pois não é possível resolvermos, em $\mathbb{R}$, a equação característica

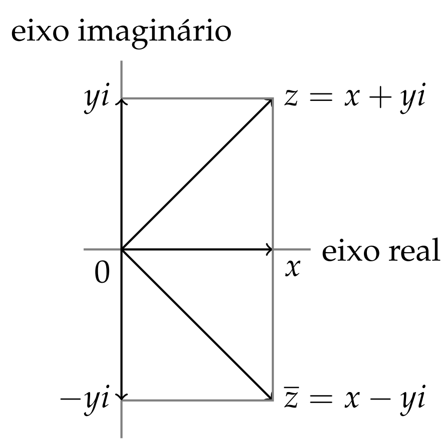
$$\lambda^2 + 1 = 0.$$

Portanto, precisamos de um conjunto que contenha $\mathbb{R}$ e a solução da equação supracitada, de modo que possamos ter uma estrutura algébrica que relacione todos os números reais com essa solução, inexistente em $\mathbb{R}$. Esse conjunto é o conhecido $\mathbb{C}$, que revisaremos sucintamente para aqueles que não tiveram a oportunidade de estudá-lo de maneira apropriada.

1. Para o conjunto $\mathbb{C}$ dos números complexos, as seguintes condições são válidas:

   (a) $z \in \mathbb{C} \iff z = x + yi$ com $x, y \in \mathbb{R}$ e $i^2 = -1$;[2]

Figura 5.1: Plano complexo



Na figura 5.1, podemos associar $z = x + yi \in \mathbb{C}$ ao vetor $\mathbf{z} = (x, y) \in \mathbb{R}^2$.

[1] Cf. p. 133.

[2] Por exemplo, $1 + 2i, 3 + 0i, 0 + 4i, -\frac{1}{\sqrt{2}} + (\ln 3)i \in \mathbb{C}$.

(b) $z = w$ com $z = x + yi$ e $w = u + vi \iff x = u$ e $y = v$.[3]

2. Se $z = x + yi$ e $y = 0$, então, por abuso de notação, denotamos $z = x$ e $\mathbb{R} \subset \mathbb{C}$.[4]

3. Se $z = x + yi$ e $w = u + vi$, define-se as operações de adição e multiplicação em $\mathbb{C}$ por:

   (a) $z + w = (x + u) + (y + v)i$;[5]

   (b) $zw = (xu - yv) + (xv + yu)i$.[6]

4. As operações supracitadas são comutativas e associativas; 0 é o elemento neutro aditivo e 1 é o multiplicativo; a adição é distributiva em relação a multiplicação; caso $z = x + yi$, $-z = -x - yi$ é seu inverso aditivo e, se $z \neq 0$, $z^{-1} = \frac{x-yi}{x^2+y^2}$ é seu inverso multiplicativo.[7]

5. $\overline{z} = x - yi$ e $|z| = \sqrt{x^2 + y^2}$ são o *conjugado* e o *módulo* de z, respectivamente. Note que, $\overline{z}$ é a imagem espelhada de z, em relação ao eixo real, conforme a figura 5.1, $z\overline{z} = |z|^2$, $z^{-1} = \dfrac{\overline{z}}{|z|^2}$ e, caso $\mathbf{z}$ seja o vetor supracitado, $|z| = \|\mathbf{z}\|$.[8]

6. Para a potenciação e a radiciação em $\mathbb{C}$, bem como para interpretar geometricamente a adição e a multiplicação complexas, confira qualquer bom livro sobre números complexos.[9]

5.1.2 O corpo $\mathbb{K}$

É possível generalizarmos o conceito de escalar e o escopo das coordenadas dos vetores e das entradas das matrizes. Assim, além de $\mathbb{R}$, escalares e coordenadas de vetores podem assumir valores em outros subconjuntos de $\mathbb{C}$. De fato, um subconjunto $\mathbb{K}$ de $\mathbb{C}$ é chamado de *corpo* (*de escalares*) quando:

1. $0, 1 \in \mathbb{K}$;

2. $k_1, k_2 \in \mathbb{K} \Longrightarrow k_1 + k_2, k_1k_2 \in \mathbb{K}$;

3. $k \in \mathbb{K} \Longrightarrow -k \in \mathbb{K}$ e, se $k \neq 0$, então $k^{-1} \in \mathbb{K}$.

$\boxed{\text{EXEMPLOS DE CORPOS}}$

$\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$.

---

[3] Por exemplo, $2 + 3i \neq 3 + 2i$.

[4] Por exemplo, $3 + 0i = 3$.

[5] Por exemplo, $(1 + 2i) + (3 + 4i) = 4 + 6i$. Note que a adição em $\mathbb{C}$ funciona como a soma dos vetores $\mathbf{z} = (x, y)$ e $\mathbf{w} = (u, v)$ em $\mathbb{R}^2$, isto é, $\mathbf{z} + \mathbf{w} = (x + u, y + v)$.

[6] Por exemplo, $(1 + i)(1 + 2i) = -1 + 3i$, $(1 + i)(1 + i) = 2i$ e $(1 + i)(1 - i) = 2$. Note que $\mathbb{C}$ pode ser interpretado como o $\mathbb{R}^2$, munido de uma operação de multiplicação entre seus elementos, não definida nos capítulos anteriores.

[7] Por exemplo, se $z = 1 + i$, então $-z = -1 - i$ e $z^{-1} = \frac{1-i}{2}$.

[8] Por exemplo, se $z = 1 + i$, então $\overline{z} = 1 - i$ e $|z| = \sqrt{2}$.

[9] Para os nossos propósitos, essas operações não são necessárias!

Embora existam outros exemplos de corpos,[10] a partir do próximo capítulo, sem perda de generalidade, consideraremos apenas $\mathbb{K} = \mathbb{R}, \mathbb{C}$.

CONTRAEXEMPLOS

$\mathbb{N}$ e $\mathbb{Z}$ não são corpos. De fato, se $k = 2$, por exemplo, então $-k \notin \mathbb{N}$ e $k^{-1} \notin \mathbb{Z}$.

5.1.3 O espaço $\mathbb{K}^n$

As definições de vetor, multiplicação de escalar por vetor e soma de vetores em $\mathbb{R}^n$, bem como as propriedades decorrentes delas, podem ser generalizadas se, no lugar de $\mathbb{R}$, considerarmos outro corpo $\mathbb{K}$ qualquer e, concomitantemente, trocarmos $\mathbb{R}^n$ por $\mathbb{K}^n$.

Portanto, ao considerarmos $\mathbf{x} \in \mathbb{K}^n$, ou seja, a *n-upla* ordenada $\mathbf{x} = (x_1, \ldots, x_n)$ com *coordenadas/componentes* em $\mathbb{K}$,[11] chamaremos $\mathbf{x}$ de *vetor* em/do/de $\mathbb{K}^n$, onde a palavra "ordenada" foi utilizada no seguinte sentido: se $\mathbf{x} = (x_1, \ldots, x_n)$ e $\mathbf{y} = (y_1, \ldots, y_n)$, então

$$\mathbf{x} = \mathbf{y} \iff x_i = y_i,\ i = 1, \ldots, n.$$ [12]

Em $\mathbb{K}^n$, o vetor *soma* de $\mathbf{x} = (x_1, \ldots, x_n)$ e $\mathbf{y} = (y_1, \ldots, y_n)$ é definido por

$$\mathbf{x} + \mathbf{y} := (x_1 + y_1, \ldots, x_n + y_n).$$ [13]

Em $\mathbb{K}^n$, se $\lambda \in \mathbb{K}$ e $\mathbf{x} = (x_1, \ldots, x_n)$, então o vetor

$$\lambda\mathbf{x} := (\lambda x_1, \ldots, \lambda x_n)$$

é chamado de *produto por escalar*.[14]
Assim como em $\mathbb{R}^n$,[15] as seguintes propriedades são sempre válidas em $\mathbb{K}^n$:

1. $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$;
2. $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$;
3. $\mathbf{0} = (0, \ldots, 0) \in \mathbb{K}^n$ é tal que $\mathbf{x} + \mathbf{0} = \mathbf{x}$;
4. $-\mathbf{x} := (-1)\mathbf{x}$ é tal que $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$;
5. $\lambda(\mathbf{x} + \mathbf{y}) = \lambda\mathbf{x} + \lambda\mathbf{y}$;
6. $(\lambda + \beta)\mathbf{x} = \lambda\mathbf{x} + \beta\mathbf{x}$;

---

[10] Por exemplo, verifique que

$$\mathbb{K} = \left\{ x + y\sqrt{2} \mid x, y \in \mathbb{Q} \right\}$$

é um corpo, denotado por $\mathbb{K} = \mathbb{Q}\left[\sqrt{2}\right]$.

[11] Por exemplo, $\mathbf{x} = (1, i, 1+i, 1-i) \in \mathbb{C}^4$.

[12] Por exemplo, em $\mathbb{C}^2$, $(1, i) \neq (i, 1)$.

[13] Por exemplo, se $\mathbf{x} = (1, i, 1+i)$, $\mathbf{y} = (-1, 1, 1-i) \in \mathbb{C}^3$, então $\mathbf{x} + \mathbf{y} = (0, 1+i, 2) \in \mathbb{C}^3$.

[14] Por exemplo, se $\lambda = \frac{1}{i}$ e $\mathbf{x} = (i/2, -1/i) \in \mathbb{C}^2$, então $\lambda\mathbf{x} = (1/2, 1) \in \mathbb{C}^2$.

[15] Cf. a subseção 2.2.2, p. 18.

7. $(\lambda\beta)\mathbf{x} = \lambda(\beta\mathbf{x})$;

8. $1\mathbf{x} = \mathbf{x}$.

Como veremos nos capítulos 6 e 7, os conceitos e resultados relacionados ao $\mathbb{R}^n$, tais como, produto interno, norma, CL, gerador, subespaço, LI, LD, base, dimensão, sistema linear, operador linear, autovalor, autovetor, diagonalização, etc., podem ser generalizados para o espaço $\mathbb{K}^n$. Contudo, algumas adaptações são necessárias, como ilustram alguns dos exercícios da seção 5.2. Além disso, a generalização supracitada também ocorre de $\mathbb{R}^{m\times n}$ para $\mathbb{K}^{m\times n}$.

5.2 Exercícios

1. Sejam:

   · $\mathbb{K}$ um corpo;

   · $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{K}^n$;

   · $\mathbf{x} \cdot \mathbf{y} := x_1 y_1 + \cdots + x_n y_n$.

   Como no capítulo 2, para quaisquer $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{K}^n$ e $c \in \mathbb{K}$, as seguintes propriedades são satisfeitas:

   (a) $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$;

   (b) $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$;

   (c) $c(\mathbf{x} \cdot \mathbf{y}) = (c\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (c\mathbf{y})$;

   (d) $\mathbb{K} \subset \mathbb{R}$ acarreta:

   - $\mathbf{x} \cdot \mathbf{x} \geq 0$;
   - $\mathbf{x} \cdot \mathbf{x} = 0 \iff \mathbf{x} = \mathbf{0}$.

   Agora, caso $\mathbb{K}$ contenha algum número complexo com parte imaginária não nula, resolva as seguintes questões:

   1.1. Apresente algum exemplo em que a propriedade (d) supracitada não seja satisfeita.[16]

   1.2. Para $\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$ ser *bem definido* em $\mathbb{K}^n$, considere

   $$\mathbf{x} \cdot \mathbf{y} := x_1 \overline{y_1} + x_2 \overline{y_2} + \cdots + x_n \overline{y_n}. \tag{5.1}$$

   Então, verifique que (d) é satisfeita para qualquer corpo $\mathbb{K}$.[17]

   1.3. Como ficam as propriedades (a) e (c) supracitadas, considerando (5.1)?[18]

   OBSERVAÇÃO

   Note que, como a conjugação complexa não altera números reais, caso as coordenadas de $\mathbf{x}, \mathbf{y} \in \mathbb{K}^n$ sejam reais, temos

   $$x_1 \overline{y_1} + x_2 \overline{y_2} + \cdots + x_n \overline{y_n} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

   Isso significa que, o produto interno estudado até o capítulo 4 é, exatamente, o definido em (5.1), isto é, independente do $\mathbb{K}^n$ considerado, $\mathbf{x} \cdot \mathbf{y}$ conjuga as coordenadas de $\mathbf{y}$.

---

[16] RESOLUÇÃO

Considere os exemplos

$$\begin{aligned} \mathbf{x} = (2i, 1) &\implies \mathbf{x} \cdot \mathbf{x} = -3 < 0; \\ \mathbf{x} = (i, 1) &\implies \mathbf{x} \cdot \mathbf{x} = 0. \end{aligned}$$

[17] DICA

Note que, pelo item 5 da subseção 5.1.1, $\mathbf{x} \cdot \mathbf{x} = |x_1|^2 + \cdots + |x_n|^2$.

[18] RESPOSTAS

(a) $\mathbf{x} \cdot \mathbf{y} = \overline{\mathbf{y} \cdot \mathbf{x}}$; (c) $c(\mathbf{x} \cdot \mathbf{y}) = (c\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (\bar{c}\mathbf{y})$.

2. Utilize Gram-Schmidt para transformar $\{(i,i,i),(0,i,i),(0,0,i)\}$ numa base ortonormal de $\mathbb{C}^3$.[19]

3. Obtenha uma base ortonormal do subespaço de $\mathbb{C}^3$ gerado por $(1,i,0)$ e $(1,2,1-i)$.[20]

RESOLUÇÃO

Sejam $\mathbf{a}_1 = (1,i,0)$ e $\mathbf{a}_2 = (1,2,1-i)$. É fácil verificar que, nenhum desses vetores pode ser escrito como múltiplo escalar do outro. Logo, $\mathcal{B} = \{\mathbf{a}_1, \mathbf{a}_2\}$ é uma base de um subespaço $\mathcal{S}$ (de dimensão 2) de $\mathbb{C}^3$. Contudo, essa base não é ortonormal, pois, não temos ortogonalidade entre $\mathbf{a}_1$ e $\mathbf{a}_2$ e, além disso, esses vetores não são unitários. De fato,

$$\begin{aligned}
\mathbf{a}_1 \cdot \mathbf{a}_2 &= 1\cdot 1 + i\cdot 2 + 0\cdot(1+i)\\
&= 1+2i\\
&\neq 0,\\
\mathbf{a}_1 \cdot \mathbf{a}_1 &= 1\cdot 1 + i\cdot(-i) + 0\cdot 0\\
&= 2\\
&\neq 1 \ \text{ e}\\
\mathbf{a}_2 \cdot \mathbf{a}_2 &= 1\cdot 1 + 2\cdot 2 + (1+i)\cdot(1-i)\\
&= 7\\
&\neq 1.
\end{aligned}$$

Portanto, para obter uma base ortogonal $\mathcal{B}' = \left\{\mathbf{a}_1', \mathbf{a}_2'\right\}$ de $\mathcal{S}$, considere

$$\begin{aligned}
\mathbf{a}_1' &= \mathbf{a}_1\\
&= (1,i,0) \ \text{ e}\\
\mathbf{a}_2' &= \mathbf{a}_2 - \frac{\mathbf{a}_2\cdot\mathbf{a}_1'}{\mathbf{a}_1'\cdot\mathbf{a}_1'}\mathbf{a}_1'\\
&= (1,2,1-i) - \frac{1-2i}{2}(1,i,0)\\
&= \left(1-\frac{1-2i}{2}, 2-\frac{2+i}{2}, 1-i\right)\\
&= \left(\frac{1+2i}{2}, \frac{2-i}{2}, 1-i\right).
\end{aligned}$$

De fato,

$$\begin{aligned}
\mathbf{a}_1'\cdot\mathbf{a}_2' &= 1\cdot\frac{1-2i}{2} + i\cdot\frac{2+i}{2} + 0\cdot(1+i)\\
&= \frac{1-2i}{2} + \frac{-1+2i}{2}\\
&= 0.
\end{aligned}$$

[19] DICA
Faça como no exercício 9 da seção 2.7, mas agora utilize a definição (5.1).

[20] Idem.

Finalmente, para obter uma base ortonormal $\mathcal{B}'' = \{\mathbf{a}_1'', \mathbf{a}_2''\}$ de $\mathcal{S}$, considere

$$\begin{aligned}
\mathbf{a}_1'' &= \frac{\mathbf{a}_1'}{\|\mathbf{a}_1'\|} \\
&= \frac{1}{\sqrt{2}}(1, i, 0) \ \text{ e} \\
\mathbf{a}_2'' &= \frac{\mathbf{a}_2'}{\|\mathbf{a}_2'\|} \\
&= \frac{\sqrt{2}}{3}\left(\frac{1+2i}{2}, \frac{2-i}{2}, 1-i\right) \\
&= \frac{\sqrt{2}}{6}(1+2i, 2-i, 2-2i).
\end{aligned}$$

4. Sejam C, D, E e F números reais arbitrários e diferentes de -1.[21] Dê um exemplo de uma base ortonormal de $\mathbb{C}^3$ que contenha o vetor $\mathbf{x}/\|\mathbf{x}\|$, onde

$$\mathbf{x} = ((C+1) + (D+1)i, E+1, F+1).$$[22]

5. Em $\mathbb{C}^{3\times 4}$, seja

$$A = \begin{bmatrix} 1 & -1 & 0 & 2 \\ -2 & 1 & 1-i & 0 \\ -1+i & 0 & 1 & 2 \end{bmatrix}.$$

Determine a entrada a_{ij}, a i-ésima linha $A(i,-)$ e a j-ésima coluna $A(-,j)$ de A, para $i = 1,2,3$ e $j = 1,2,3,4$.

6. Esse exercício exemplifica a seguinte propriedade distributiva para matrizes com entradas num corpo $\mathbb{K}$:

$$\boxed{(A+B)C = AC + BC \quad \forall A, B \in \mathbb{K}^{m\times n} \text{ e } \forall C \in \mathbb{K}^{n\times p}.}$$

Assim, para $m = p = 2$, $n = 3$, $\mathbb{K} = \mathbb{C}$,

$$A = \begin{bmatrix} 1+i & 1 & -i \\ 2i & 2+3i & 2 \end{bmatrix}, B = \begin{bmatrix} 1-i & i & -1 \\ 3 & 1 & 2 \end{bmatrix} \text{ e } C = \begin{bmatrix} 1+i & i \\ -i & 2-3i \\ 1 & -1 \end{bmatrix},$$

calcule:

(a) $A+B$;
(b) $(A+B)C$;
(c) AC;
(d) BC;
(e) $AC+BC$.

[21] Confira a *footnote* do quarto exemplo dado após o resultado (R34), p. 194.

[22] $\boxed{\text{SUGESTÃO}}$
Resolva como na seção 2.7. Por exemplo, considere $\mathbf{a}_1 = \frac{\mathbf{x}}{\|\mathbf{x}\|}$, $\mathbf{a}_2 = \mathbf{e}_1$ e $\mathbf{a}_3 = \mathbf{e}_2$. Verifique que esses vetores são LI. Então, obtenha $\mathbf{a}_1'$, $\mathbf{a}_2'$ e $\mathbf{a}_3'$ via Gram-Schmidt. Por último, obtenha $\{\mathbf{a}_1'', \mathbf{a}_2'', \mathbf{a}_3''\}$ ortonormal.

7. Esse exercício exemplifica a seguinte propriedade de transposição para matrizes com entradas num corpo $\mathbb{K}$:

$$\boxed{(AB)^t = B^t A^t \quad \forall A \in \mathbb{K}^{m\times n} \text{ e } \forall B \in \mathbb{K}^{n\times p}.}$$

Assim, para $m = p = 2$, $n = 3$, $\mathbb{K} = C$,

$$A = \begin{bmatrix} 1+i & 1 & -i \\ 2i & 2+3i & 2 \end{bmatrix} \text{ e } B = \begin{bmatrix} 1+i & i \\ -i & 2-3i \\ 1 & -1 \end{bmatrix},$$

calcule:

(a) AB;
(b) $(AB)^t$;
(c) B^t;
(d) A^t;
(e) $B^t A^t$.

8. Sejam

$$A = \begin{bmatrix} 2 & i \\ -i & 1 \end{bmatrix}, \text{ isto é, } A^t = \begin{bmatrix} 2 & -i \\ i & 1 \end{bmatrix}, \text{ e } B = \begin{bmatrix} 1+i & 1 \\ 1 & 1-i \end{bmatrix}.$$

Verifique que:

(a) $AB = \begin{bmatrix} 2+3i & 3+i \\ 2-i & 1-2i \end{bmatrix}$ e $(AB)^{-1} = \begin{bmatrix} 1-2i & -3-i \\ -2+i & 2+3i \end{bmatrix}$;
(b) $A^{-1} = \begin{bmatrix} 1 & -i \\ i & 2 \end{bmatrix}$, $(A^t)^{-1} = \begin{bmatrix} 1 & i \\ -i & 2 \end{bmatrix}$ e $B^{-1} = \begin{bmatrix} 1-i & -1 \\ -1 & 1+i \end{bmatrix}$;
(c) $(A^t)^{-1} = (A^{-1})^t$ e $(AB)^{-1} = B^{-1}A^{-1}$.

9. Caso

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

seja invertível, determine A^{-1}.

RESOLUÇÃO

Podemos adaptar, para qualquer matriz quadrada com entradas complexas, o método descrito no exercício 15 da seção 3.7. De fato, se considerarmos (7.12),[23] então

$$\begin{aligned}
[A|\mathrm{I}] = \left[\begin{array}{ccc|ccc} 1 & 1+i & 0 & 1 & 0 & 0 \\ 0 & 1 & i & 0 & 1 & 0 \\ -i & 1-2i & 2 & 0 & 0 & 1 \end{array}\right] &\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 1+i & 0 & 1 & 0 & 0 \\ 0 & 1 & i & 0 & 1 & 0 \\ 0 & -i & 2 & i & 0 & 1 \end{array}\right] \\
&\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 1-i & 1 & -(1+i) & 0 \\ 0 & 1 & i & 0 & 1 & 0 \\ 0 & 0 & 1 & i & i & 1 \end{array}\right] \\
&\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -i & -2(1+i) & -(1-i) \\ 0 & 1 & 0 & 1 & 2 & -i \\ 0 & 0 & 1 & i & i & 1 \end{array}\right] = \left[\mathrm{I}|A'\right].
\end{aligned}$$

[23] Cf. p. 186.

Portanto, $A' = A^{-1}$.[24]

10. Sejam C, D, E e F números reais não negativos arbitrários, com C e D diferentes de -1.[25] Escalone a matriz

$$A = \begin{pmatrix} (C+1)+(D+1)i & E+1 & F+1 \\ (D+1)i & E & F \\ (D+1)i & E & F-i \end{pmatrix} \in \mathbb{C}^3$$

e obtenha a sua escalonada reduzida R. Com base na R obtida, o que podemos concluir sobre a invertibilidade de A?

RESOLUÇÃO

Como

$$A = \begin{pmatrix} (C+1)+(D+1)i & E+1 & F+1 \\ (D+1)i & E & F \\ (D+1)i & E & F-i \end{pmatrix} \longrightarrow \begin{pmatrix} C+1 & 1 & 1 \\ (D+1)i & E & F \\ 0 & 0 & -i \end{pmatrix}$$
$$\longrightarrow \begin{pmatrix} C+1 & 1 & 1 \\ (D+1)i & E & F \\ 0 & 0 & 1 \end{pmatrix}$$
$$\longrightarrow \begin{pmatrix} C+1 & 1 & 0 \\ (D+1)i & E & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$\longrightarrow \begin{pmatrix} 1 & 1/(C+1) & 0 \\ (D+1)i & E & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$\longrightarrow \begin{pmatrix} 1 & 1/(C+1) & 0 \\ 0 & E - \frac{D+1}{C+1}i & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$\longrightarrow \begin{pmatrix} 1 & 1/(C+1) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$\longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathsf{R},$$

A é invertível.

11. Obtenha os autovalores e autovetores do operador linear

$$\begin{array}{rccc} T : & \mathbb{K}^2 & \longrightarrow & \mathbb{K}^2 \\ & (z,w) & \longmapsto & (-w,z) \end{array} \tag{5.2}$$

para:

(a) $\mathbb{K} = \mathbb{R}$;

(b) $\mathbb{K} = \mathbb{C}$.

RESOLUÇÃO

(a) Note que, nesse caso, $T = R_{90^\circ}$, a rotação de 90 graus no sentido anti-horário.[26] Assim, por um lado, podemos analisar o problema respondendo as seguintes perguntas:

---

[24] De fato, verifique que $AA' = \mathsf{I} = A'A$.

[25] Confira a primeira nota de rodapé do exercício 2, dado após o resultado (R34) da página 194.

[26] Cf. (4.4), p. 86.

· O que acontece quando aplicamos uma rotação de 90 graus num vetor não nulo?

· Essa rotação pode levar esse vetor a um múltiplo escalar dele mesmo?[27]

Por outro, podemos calcular os autovalores (diretamente) do operador

$$R_{90^\circ}\left(\begin{bmatrix} z \\ w \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} z \\ w \end{bmatrix},$$

estudado no capítulo 4. Portanto,

$$\lambda \text{ é autovalor de } T \iff \lambda^2 + 1 = 0.$$

Então, nesse caso, T não tem autovalores nem autovetores.[28]
(b) Nesse caso, como

$$\begin{aligned} T(z,w) = \lambda(z,w) &\Longrightarrow (-w,z) = (\lambda z, \lambda w) \\ &\Longrightarrow -w = \lambda z, z = \lambda w \\ &\Longrightarrow -w = \lambda^2 w \\ &\Longrightarrow \lambda^2 = -1, \end{aligned}$$

pois $w \neq 0$,[29] $\lambda = \pm i$ são os autovalores com autovetores associados $w(\pm i, 1)$, $w \in \mathbb{C}$, respectivamente.

12. Seja $A \in \mathbb{K}^{n\times n}$. A matriz $B = \overline{A} \in \mathbb{K}^{n\times n}$ é definida da seguinte maneira:

$$b_{ij} = \overline{a_{ij}}, i,j = 1,\ldots,n.$$

EXEMPLO

$A = \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} \Longrightarrow \overline{A} = \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}.$

Em geral, denota-se $A^* := \overline{A}^t$.

EXEMPLO

Para a matriz A do exemplo anterior, temos

$$A^* = A. \tag{5.3}$$

Uma matriz é dita *hermiteana* caso satisfaça a equação (5.3).

EXEMPLO

Matrizes simétricas são hermiteanas.

---

[27] Pense um pouco!

[28] Note que, com esse exemplo, podemos responder a pergunta

"Existe operador (ou matriz quadrada) que não tenha autovalores (reais)?"

com um simples "Sim!".

[29] De fato,

$$\begin{aligned} w = 0 &\Longrightarrow z = 0 \\ &\Longrightarrow (z,w) = (0,0) \text{ não é autovetor.} \end{aligned}$$

$P \in \mathbb{K}^{n\times n}$ é chamada de *unitária* quando

$$P^{-1} = P^*. \tag{5.4}$$

EXEMPLO

Matrizes ortogonais são unitárias.

A é dita *normal* quando

$$AA^* = A^*A. \tag{5.5}$$

12.1. Demonstre que matrizes hermiteanas e unitárias são normais.

12.2. Apresente matrizes normais que não sejam hermiteanas nem unitárias.[30]

Considere a seguinte proposição:[31]

$$A \text{ é normal} \iff A \text{ é unitariamente diagonalizável}, \tag{5.6}$$

ou seja, *a normalidade de A é equivalente a existência de uma matriz unitária* P *tal que* $\mathsf{P}^*A\mathsf{P} = \mathsf{D}$ *é diagonal.*[32]

> *Demonstraremos, na seção 5.3, uma forma mais fraca do resultado (5.6), para matrizes hermiteanas, e, na seção 7.4, uma outra forma desse resultado, para matrizes/operadores normais.*

12.3. Considerando a matriz hermiteana A dada no primeiro exemplo desse exercício, resolva as seguintes questões:

(a) Determine os autovalores de A.

(b) Para cada autovalor de A, determine os autovetores associados.

(c) Diagonalize A por uma matriz unitária P.

RESOLUÇÃO

(a) Como

$$\begin{aligned}\det(A - \lambda \mathsf{I}) &= \begin{vmatrix} 1-\lambda & i \\ -i & 1-\lambda \end{vmatrix} \\ &= \lambda^2 - 2\lambda + 1 - 1 \\ &= \lambda(\lambda - 2),\end{aligned}$$

---

[30] RESOLUÇÃO

Considere, por exemplo,

$$\begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \text{ e } \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

[31] Essa proposição é uma generalização da afirmação 7 do capítulo 4.

[32] Nesse caso, P e D são obtidas como no capítulo 4.

$\lambda_1 = 0$ e $\lambda_2 = 2$ são os autovalores de A.
(b) Ao resolvermos o sistema $(A-\lambda_1\mathsf{I})\mathbf{x} = \mathbf{0}$, podemos obter os autovetores de A associados à $\lambda_1 = 0$. De fato, como

$$A - 0\mathsf{I} = \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} \longrightarrow \mathsf{R} = \begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix},$$

os autovetores supracitados são da forma

$$c\begin{bmatrix} 1 \\ i \end{bmatrix},$$

para qualquer escalar complexo c. Agora, da resolução do sistema $(A-\lambda_2\mathsf{I})\mathbf{x} = \mathbf{0}$, podemos obter os autovetores de A associados à $\lambda_2 = 2$. De fato, como

$$A - 2\mathsf{I} = \begin{bmatrix} -1 & i \\ -i & -1 \end{bmatrix} \longrightarrow \mathsf{R} = \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix},$$

os autovetores supracitados são da forma

$$c\begin{bmatrix} 1 \\ -i \end{bmatrix},$$

para qualquer escalar complexo c.
(c) Considere os autovetores

$$\mathbf{x}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ i \end{bmatrix} \text{ e } \mathbf{x}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -i \end{bmatrix}.$$

Além disso, considere a matriz unitária $\mathsf{P} = [\mathbf{x}_1\ \mathbf{x}_2]$ e a matriz diagonal D cuja diagonal principal consiste dos autovalores λ_1 e λ_2, nessa ordem. Agora, basta verificarmos que $\mathsf{P}^*A\mathsf{P} = \mathsf{D}$.

13. Diagonalize unitariamente a seguinte matriz unitária

$$A = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}.$$ [33]

RESOLUÇÃO

Como

$$\det(A-\lambda\mathsf{I}) = \begin{vmatrix} \frac{1}{\sqrt{2}} - \lambda & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} - \lambda \end{vmatrix} = \lambda^2 - \sqrt{2}\lambda + 1,$$

$\lambda_{1,2} = (1 \pm i)\sqrt{2}$ são os autovalores de A. Assim, por um lado, da resolução do sistema $(A-\lambda_1\mathsf{I})\mathbf{x} = \mathbf{0}$, podemos obter os autovetores de A associados à λ_1. De fato, como

$$A - \lambda_1\mathsf{I} = \frac{1}{\sqrt{2}}\begin{bmatrix} -i & i \\ i & -i \end{bmatrix} \longrightarrow \mathsf{R} = \frac{1}{\sqrt{2}}\begin{bmatrix} i & -i \\ 0 & 0 \end{bmatrix},$$ [34]

os autovetores supracitados são da forma

$$c\begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

para qualquer escalar complexo c. Por outro lado, da resolução do sistema $(A-\lambda_2\mathsf{I})\mathbf{x} = \mathbf{0}$, podemos obter os autovetores de A associados à λ_2. De fato, como

$$A - \lambda_2\mathsf{I} = \frac{1}{\sqrt{2}}\begin{bmatrix} i & i \\ i & i \end{bmatrix} \longrightarrow \mathsf{R} = \frac{1}{\sqrt{2}}\begin{bmatrix} i & i \\ 0 & 0 \end{bmatrix},$$ [35]

[33] Por ser unitária, A é normal!
[34] Note que R, aqui, não é a escalonada reduzida!
[35] Idem!

os autovetores supracitados são da forma

$$c\begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

para qualquer escalar complexo c. Finalmente, para

$$\mathsf{P} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix},$$

temos

$$\mathsf{P}^* A \mathsf{P} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1+i & 0 \\ 0 & 1-i \end{bmatrix}.$$

14. Diagonalize unitariamente a seguinte matriz hermiteana

$$A = \begin{bmatrix} 2 & 3-3i \\ 3+3i & 5 \end{bmatrix}.^{36}$$

RESOLUÇÃO

Como

$$\begin{aligned} \det(A - \lambda \mathsf{I}) &= \begin{vmatrix} 2-\lambda & 3-3i \\ 3+3i & 5-\lambda \end{vmatrix} \\ &= \lambda^2 - 7\lambda - 8, \end{aligned}$$

$\lambda_1 = 8$ e $\lambda_2 = -1$ são os autovalores de A. Portanto, por um lado, os autovetores de A associados à $\lambda_1 = 8$ são obtidos ao resolvermos o sistema $(A - 8\mathsf{I})\mathbf{x} = \mathbf{0}$. Assim, do escalonamento

$$\begin{aligned} A - 8\mathsf{I} = \begin{bmatrix} -6 & 3(1-i) \\ 3(1+i) & -3 \end{bmatrix} &\longrightarrow \begin{bmatrix} -2 & 1-i \\ 1+i & -1 \end{bmatrix} \\ &\longrightarrow \begin{bmatrix} -2 & 1-i \\ 2 & -(1-i) \end{bmatrix} \\ &\longrightarrow \begin{bmatrix} 1 & -(1-i)/2 \\ 0 & 0 \end{bmatrix} = \mathsf{R}, \end{aligned}$$

segue que os autovetores supracitados são da forma

$$c\begin{bmatrix} (1-i)/2 \\ 1 \end{bmatrix},$$

para qualquer escalar complexo c. Por outro lado, os autovetores de A associados à $\lambda_2 = -1$ são obtidos ao resolvermos o sistema $(A + \mathsf{I})\mathbf{x} = \mathbf{0}$. De fato, do escalonamento

$$\begin{aligned} A + \mathsf{I} = \begin{bmatrix} 3 & 3(1-i) \\ 3(1+i) & 6 \end{bmatrix} &\longrightarrow \begin{bmatrix} 1 & 1-i \\ 1+i & 2 \end{bmatrix} \\ &\longrightarrow \begin{bmatrix} 1 & 1-i \\ 0 & 0 \end{bmatrix} = \mathsf{R}, \end{aligned}$$

segue que os autovetores supracitados são da forma

$$c\begin{bmatrix} -(1-i) \\ 1 \end{bmatrix},$$

para qualquer escalar complexo c. Finalmente, para

$$\begin{aligned} \mathbf{x}_1 &= \frac{1}{\sqrt{3/2}}\begin{bmatrix} (1-i)/2 \\ 1 \end{bmatrix}, \\ \mathbf{x}_2 &= \frac{1}{\sqrt{3}}\begin{bmatrix} -(1-i) \\ 1 \end{bmatrix} \text{ e} \\ \mathsf{P} &= \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 \end{bmatrix}, \end{aligned}$$

[36] Por ser hermiteana, A é normal!

temos

$$\mathsf{P}^*A\mathsf{P} = \begin{bmatrix} 8 & 0 \\ 0 & -1 \end{bmatrix}.$$

15. Diagonalize unitariamente a seguinte matriz hermiteana

$$A = \begin{bmatrix} 0 & i & 1 \\ -i & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$[37]

RESOLUÇÃO PARCIAL

Os autovalores de A são

$$0,\ \sqrt{2}\ \text{e}\ -\sqrt{2},$$

correspondendo aos autovetores

$$\begin{bmatrix} 0 \\ i \\ 1 \end{bmatrix},\ \begin{bmatrix} \sqrt{2} \\ -i \\ 1 \end{bmatrix}\ \text{e}\ \begin{bmatrix} -\sqrt{2} \\ -i \\ 1 \end{bmatrix},$$

respectivamente. Observe que esses vetores são ortogonais em relação ao produto interno canônico em $\mathbb{C}^3$. Agora, ao normalizarmos os autovetores, obtemos a matriz unitária

$$\mathsf{P} = \begin{bmatrix} 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ i/\sqrt{2} & -i/2 & -i/2 \\ 1/\sqrt{2} & 1/2 & 1/2 \end{bmatrix}.$$

Portanto, $\mathsf{P}^*A\mathsf{P} = D$ é a matriz diagonal cujos elementos da diagonal principal são os autovalores dados acima, na ordem em que aparecem.

5.3 Informação adicional: diagonalização de $A \in \mathbb{C}^{n\times n}$

5.3.1 Lema de Schur

A é unitariamente semelhante a alguma matriz triangular superior.

DEMONSTRAÇÃO

Usaremos indução matemática sobre n. Portanto, como o caso $n = 1$ é trivial, suponha que a afirmação (em itálico) supracitada seja verdadeira para $n = r - 1 \geq 1$, $A \in \mathbb{C}^{r\times r}$ e λ_1 seja um autovalor de A.[38] Seja $\mathbf{x}$ um autovetor unitário de A associado à λ_1. Então, por Gram-Schmidt, podemos obter uma base ortonormal $\{\mathbf{x}, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$ de $\mathbb{C}^r$. Assim, ao considerarmos a matriz unitária $\mathsf{P}_1 = [\mathbf{x}\,\mathbf{u}_2\,\cdots\,\mathbf{u}_r]$, temos

$$\begin{aligned} \mathsf{P}_1^{-1}A\mathsf{P}_1 &= \mathsf{P}_1^*A\mathsf{P}_1 \\ &= \mathsf{P}_1^*\,[A\mathbf{x}\,A\mathbf{u}_2\,\cdots\,A\mathbf{u}_r] \\ &= \begin{bmatrix} \overline{\mathbf{x}}^T \\ \overline{\mathbf{u}}_2^T \\ \vdots \\ \overline{\mathbf{u}}_r^T \end{bmatrix} [A\mathbf{x}\,A\mathbf{u}_2\,\cdots\,A\mathbf{u}_r] \\ &= \begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & B \end{bmatrix}, \end{aligned}$$

[37]Por ser hermiteana, A é normal.

[38]Para a existência desse autovalor, confira a subseção 7.2.3.

onde, na matriz da última igualdade, de cima para baixo, $*$ representa todas as entradas de sua primeira linha, excetuando-se λ_1. Então, pela hipótese de indução, existe uma matriz unitária $\mathsf{P}_2 \in \mathbb{C}^{(r-1)\times(r-1)}$ tal que $\mathsf{P}_2^{-1}B\mathsf{P}_2$ é triangular superior. Se denotarmos as entradas da diagonal principal dessa matriz triangular por $\lambda_2, \ldots, \lambda_r$, então

$$\mathsf{P} = \mathsf{P}_1 \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathsf{P}_2 \end{bmatrix}$$

também é unitária e

$$\begin{aligned} \mathsf{P}^{-1}A\mathsf{P} &= \mathsf{P}^*A\mathsf{P} \\ &= \begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & \mathsf{P}_2^*B\mathsf{P}_2 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_r \end{bmatrix}, \end{aligned}$$

onde a matriz da última igualdade, de cima para baixo, é triangular superior, com 0 e $*$ representando, respectivamente, todas as entradas abaixo e acima da diagonal principal dessa matriz.

5.3.2 Teorema espectral para $A \in \mathbb{C}^{n\times n}$

A é hermiteana $\Longrightarrow$ A é unitariamente diagonalizável.

DEMONSTRAÇÃO

Primeiramente, note que, pelo lema de Schur, existe uma matriz P unitária tal que $\mathsf{P}^*A\mathsf{P}$ é triangular superior. Contudo, como A é hermiteana, $\mathsf{P}^*A\mathsf{P}$ também o é, pois

$$\begin{aligned} (\mathsf{P}^*A\mathsf{P})^* &= \mathsf{P}^*A^*\mathsf{P} \\ &= \mathsf{P}^*A\mathsf{P}. \end{aligned}$$

Assim, como $\mathsf{P}^*A\mathsf{P}$ é triangular superior e hermiteana, também é diagonal.

Capítulo 6

O espaço vetorial $\mathcal{V}$ sobre o corpo $\mathbb{K}$

6.1 Definição e propriedades de $(\mathcal{V}, +, \cdot)$

> *A mesma estrutura algébrica que estudamos, para as n-uplas ordenadas e matrizes $m \times n$, pode ser considerada para outras construções matemáticas, ou seja, $\mathbb{K}^n$ (e.g., $\mathbb{R}^n$) e $\mathbb{K}^{m\times n}$ (e.g., $\mathbb{R}^{m\times n}$) não são os únicos exemplos de espaços vetoriais sobre o corpo de escalares $\mathbb{K}$.*

Como nos capítulos 2–5, um conjunto não vazio $\mathcal{V}$ é chamado de *espaço vetorial* (*sobre o corpo* $\mathbb{K}$), caso seja dotado das operações (bem definidas em $\mathcal{V}$) de adição $u + v$ e multiplicação por escalar $\lambda \cdot v = \lambda v$, com $u, v \in \mathcal{V}$ e $\lambda \in \mathbb{K}$ arbitrários,[1] e sejam válidas, para quaisquer $u, v, w \in \mathcal{V}$ e $\lambda, \beta \in \mathbb{K}$, as seguintes propriedades:

1. $u + v = v + u$; (COMUTATIVIDADE DA ADIÇÃO)
2. $(u + v) + w = u + (v + w)$; (ASSOCIATIVIDADE DA ADIÇÃO)
3. $0 \in \mathcal{V}$ satisfaz a equação $u + 0 = u$;[2] (EXISTÊNCIA DO NEUTRO ADITIVO)
4. $-u := (-1)u$ satisfaz a equação $u + (-u) = 0$; (EXISTÊNCIA DO SIMÉTRICO ADITIVO)
5. $\lambda(u + v) = \lambda u + \lambda v$; (DISTRIBUTIVIDADE DA MULTIPLICAÇÃO POR ESCALAR EM RELAÇÃO À ADIÇÃO DE VETORES)
6. $(\lambda + \beta)u = \lambda u + \beta u$; (DISTRIBUTIVIDADE DA MULTIPLICAÇÃO POR ESCALAR EM RELAÇÃO À ADIÇÃO DE ESCALARES)
7. $(\lambda\beta)u = \lambda(\beta u)$; (ASSOCIATIVIDADE DA MULTIPLICAÇÃO POR ESCALAR)
8. $1u = u$. (EXISTÊNCIA DO NEUTRO MULTIPLICATIVO)

[1] Caso não seja necessário que especifiquemos o espaço vetorial $\mathcal{V}$, denotaremos seus elementos, isto é, seus *vetores*, por letras minúsculas, em itálico.

[2] Podemos denotar esse neutro por $0 = 0_{\mathcal{V}}$.

6.1.1 Exemplos de $\mathcal{V}$ "diferentes" de $\mathbb{K}^n$ e $\mathbb{K}^{m\times n}$

- Como vimos no capítulo 5, $\mathcal{V} = \mathbb{K}^n$ é um espaço vetorial sobre o corpo $\mathbb{K}$. Por exemplo, $\mathcal{V} = \mathbb{C}^n$ é um espaço vetorial sobre $\mathbb{K} = \mathbb{C}$. Além disso, $\mathcal{V} = \mathbb{C}^n$ também é um espaço vetorial sobre $\mathbb{K} = \mathbb{R}$.[3]

- Podemos generalizar o espaço $\mathbb{K}^n$ das sequências finitas em $\mathbb{K}$. De fato, seja $\mathcal{V} = \mathbb{K}^{\mathbb{N}}$ o espaço das sequências infinitas em $\mathbb{K}$. Assim, $x \in \mathbb{K}^{\mathbb{N}}$ se, e somente se, x é uma função definida por
$$\begin{array}{rccc} x: & \mathbb{N} & \to & \mathbb{K} \\ & n & \mapsto & x_n \end{array}.$$
Nesse caso, denota-se
$$x = (x_1, x_2, \ldots)\,.$$
Aqui, a adição de vetores e a multiplicação de escalares por vetores são definidas como em $\mathbb{K}^n$, ou seja, componente a componente. Isso significa que,
$$x + y : n \mapsto x_n + y_n \text{ e } \lambda x : n \mapsto \lambda x_n,$$
para quaisquer $x, y \in \mathbb{K}^{\mathbb{N}}$ e qualquer escalar $\lambda \in \mathbb{K}$. Além disso, $0 = (0, 0, \ldots)$ é o vetor nulo e, se x é uma sequência infinita, então $-x = (-x_1, -x_2, \ldots)$.

- Seja $\mathcal{V} = \mathcal{P}_n(\mathbb{K})$ o espaço cujos vetores são os polinômios com coeficientes em $\mathbb{K}$ e de graus menores do que n. Assim, $p \in \mathcal{P}_n(\mathbb{K})$ se, e somente se, para cada $t \in \mathbb{K}$,
$$p(t) = c_0 + c_1 t + c_2 t^2 + \cdots + c_{n-1} t^{n-1},$$
com coeficientes $c_0, c_1, c_2, \ldots, c_{n-1} \in \mathbb{K}$.[4] Aqui, a adição de vetores é a adição de polinômios, a multiplicação de escalares por vetores é a multiplicação de escalares por polinômios, 0 é o polinômio nulo e $-p$ é definido por
$$-p(t) = -c_0 - c_1 t - c_2 t^2 - \cdots - c_{n-1} t^{n-1},$$
para todo $t \in \mathbb{K}$.

$\boxed{\text{EXERCÍCIO}}$

Para cada um dos exemplos dessa subseção, demonstre que o $\mathcal{V}$ considerado é um espaço vetorial sobre $\mathbb{K}$.

6.1.2 Subespaços, bases, dimensões, etc.

Conceitos e resultados análogos aos dos capítulos anteriores permanecem válidos para outros espaços vetoriais.

$\boxed{\text{EXEMPLOS}}$

---

[3] Verifique!

[4] Alguns desses coeficientes podem ser nulos. Por exemplo, $-2 + t^2 + \frac{1}{2}t^3 + t^5 \in \mathcal{P}_6(\mathbb{K})$.
Em alguns livros, $\mathcal{P}_n(\mathbb{K})$ representa o espaço dos polinômios com coeficientes em $\mathbb{K}$ de graus menores do que *ou iguais a* n. Nesse caso, temos um espaço $(n+1)$-dimensional. Como veremos, com a nossa representação, esse espaço tem dimensão n.

· Em relação ao primeiro exemplo da subseção 6.1.1, considere $\mathbb{C}^n$ como espaço vetorial sobre $\mathbb{C}$. Portanto, $\dim(\mathbb{C}^n) = n$ e a base canônica é uma de suas bases. De fato,

$$\mathbb{C}^n \ni (a_1 + b_1 i, \ldots, a_n + b_n i) = (a_1 + b_1 i)\,\mathbf{e}_1 + \cdots + (a_n + b_n i)\,\mathbf{e}_n$$

e $\mathbf{e}_1$, ..., $\mathbf{e}_{n-1}$ e $\mathbf{e}_n$ são LI. Contudo, caso $\mathbb{C}^n$ seja espaço vetorial sobre $\mathbb{R}$, $\dim(\mathbb{C}^n) = 2n$. De fato,

$$\mathbb{C}^n \ni (a_1 + b_1 i, \ldots, a_n + b_n i) = a_1\mathbf{e}_1 + b_1\,(i\mathbf{e}_1) + \cdots + a_n\mathbf{e}_n + b_n\,(i\mathbf{e}_n)$$

e $\mathbf{e}_1$, $i\mathbf{e}_1$, ..., $\mathbf{e}_{n-1}$, $i\mathbf{e}_{n-1}$, $\mathbf{e}_n$ e $i\mathbf{e}_n$ são LI.

· Uma base para o segundo exemplo da subseção 6.1.1 é obtida ao considerarmos os vetores

$$e_1 = (1, 0, 0, 0, \ldots),\ e_2 = (0, 1, 0, 0, \ldots),\ e_3 = (0, 0, 1, 0, \ldots), \text{ etc.},$$

ou seja, para cada inteiro positivo n, a n-ésima componente de e_n é igual a 1 e todas as outras componentes são nulas.[5] Aqui, $x \in \mathbb{K}^{\mathbb{N}}$ é tal que $x = x_1 e_1 + x_2 e_2 + x_3 e_3 + \cdots$. Sem nos aprofundarmos na questão, dizemos que esse espaço tem dimensão *infinita*.

· O terceiro exemplo da subseção 6.1.1 tem dimensão n e uma de suas bases é constituída pelos polinômios $1, t, t^2, \ldots, t^{n-2}$ e t^{n-1}. De fato, esses n polinômios geram $\mathcal{P}_n(\mathbb{K})$, pois qualquer polinômio p de grau menor que n pode ser escrito como CL dos n polinômios supracitados, ou seja, da forma $p(t) = c_0 + c_1 t + c_2 t^2 + \cdots + c_n t^{n-1}$. Além disso, esses geradores são LI. De fato, se $p(t) = 0$, então $c_0 = c_1 = c_2 = \cdots = c_{n-1} = 0$.[6]

· Seja $\mathcal{V} = \mathcal{P}(\mathbb{K})$ o espaço vetorial de todos os polinômios com coeficientes em $\mathbb{K}$.[7] A adição de polinômios e a multiplicação de escalares por polinômios são definidas como no terceiro exemplo da subseção 6.1.1.[8] $\mathcal{S} = \mathcal{P}_n(\mathbb{K})$ é um subespaço de dimensão finita (n) de $\mathcal{V}$ e os vetores $1, t, t^2, \ldots, t^{n-1}, t^n, t^{n+1}, \ldots$ formam uma das bases de $\mathcal{V}$.[9] Por esse motivo, dizemos que $\mathcal{V}$ tem dimensão *infinita*.

· Considere, novamente, o segundo exemplo da subseção 6.1.1. É fácil ver que, o conjunto

$$\mathcal{S} = \left\{ x \in \mathbb{K}^{\mathbb{N}} \mid x \text{ tem um número finito de componentes não nulas} \right\}$$

é um subespaço de $\mathbb{K}^{\mathbb{N}}$. De fato, a sequência nula pertence à $\mathcal{S}$. Além disso, se $\lambda \in \mathbb{K}$ e $x, y \in \mathcal{S}$, isto é, x tem m componentes não nulas e y tem n componentes não nulas, então $\lambda x \in \mathcal{S}$, pois esse produto tem (no máximo) m componentes não nulas, e $x + y \in \mathcal{S}$, pois essa soma tem (no máximo) $m + n$ componentes não nulas.

[5] Note a semelhança com a base canônica de $\mathbb{K}^m$!
[6] Iguale os coeficientes de $p(t)$ aos do polinômio nulo $0 + 0 \cdot t + 0 \cdot t^2 + \cdots + 0 \cdot t^{n-1}$.
[7] Agora, os graus dos polinômios pertencentes à $\mathcal{V}$ são dados por inteiros não negativos quaisquer.
[8] Verifique que esse $\mathcal{V}$ é um espaço vetorial sobre $\mathbb{K}$.
[9] Verifique!

$\boxed{\text{OBSERVAÇÃO}}$

Se $\mathcal{S}$ é um subespaço de $\mathcal{V}$, então $\mathcal{S}$ também é um espaço vetorial (sobre $\mathbb{K}$).[10]

Note que, dessa observação, podemos obter uma quantidade infinita de exemplos de espaços vetoriais.

6.2 Isomorfismo entre espaços vetoriais

Os espaços $\mathbb{R}^4$, $\mathbb{R}^{2\times 2}$ e $\mathcal{P}_4(\mathbb{R})$ são indistinguíveis, diferindo apenas em como seus vetores são representados. De fato, podemos escrever um "vetor de quatro coordenadas" de uma das seguintes formas:

· (a, b, c, d);

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

· $a + bt + ct^2 + dt^3$.

Além disso, cada vetor do espaço $\mathcal{V}$ pode ser escrito, de modo único, como uma CL dos vetores da base constituída por:

· $(1, 0, 0, 0)$, $(0, 1, 0, 0)$, $(0, 0, 1, 0)$ e $(0, 0, 0, 1)$, caso $\mathcal{V} = \mathbb{R}^4$;

· $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ e $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, caso $\mathcal{V} = \mathbb{R}^{2\times 2}$;

· 1, t, t^2 e t^3, caso $\mathcal{V} = \mathcal{P}_4(\mathbb{R})$.

Logo, independentemente da representação escolhida para um vetor de quatro coordenadas, tudo funciona da mesma forma para $\mathcal{V} = \mathbb{R}^4, \mathbb{R}^{2\times 2}, \mathcal{P}_4(\mathbb{R})$.

$\boxed{\text{EXEMPLO}}$

Nos três casos supracitados, considere a mesma CL, representada nas três formas seguintes:

· $2(1, 1, 0, 0) - (0, 1, 1, 0) = (2, 1, -1, 0)$;

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

· $2(1 + t) - (t + t^2) = 2 + t - t^2$.

---

[10] De fato, $0 \in \mathcal{S}$ e, como, para cada $\lambda \in \mathbb{K}$ e $u, v \in \mathcal{S}$, temos $\lambda u, u + v \in \mathcal{S}$, podemos obter, em $\mathcal{S}$, uma operação de adição de vetores e uma operação de multiplicação de escalares por vetores, herdadas das respectivas operações em $\mathcal{V}$, e o simétrico aditivo de cada $u \in \mathcal{S}$ (pois $-u = (-1)u \in \mathcal{S}$). Por outro lado, as outras propriedades, que devem ser satifeitas para que $\mathcal{S}$ seja um espaço vetorial, são (trivialmente) válidas, ou seja, como $\mathcal{S} \subset \mathcal{V}$ e essas outras propriedades valem para todos os vetores de $\mathcal{V}$, elas também valem para todos os vetores de $\mathcal{S}$.

Dizemos que esses espaços são *isomorfos*. Em geral, dizemos que dois espaços vetoriais, digamos $\mathcal{V}_1$ e $\mathcal{V}_2$, sobre o mesmo corpo $\mathbb{K}$, são *isomorfos*, caso exista uma função linear $L : \mathcal{V}_1 \to \mathcal{V}_2$ invertível, chamada de *isomorfismo (entre $\mathcal{V}_1$ e $\mathcal{V}_2$)*. Como no capítulo 4, da linearidade de L, temos que, a imagem da soma de vetores, pela L, é igual a soma das imagens desses vetores, pela L, e a imagem do produto de um escalar qualquer por um vetor arbitrário, pela L, é igual ao produto desse escalar pela imagem desse vetor, pela L. A invertibilidade de L, aqui, significa que L tem inversa $L^{-1} : \mathcal{V}_2 \to \mathcal{V}_1$, cuja composição com L resulta numa função identidade.[11]

EXEMPLOS

- É fácil ver que,
$$\begin{array}{rccc} L : & \mathbb{R}^{2\times 2} & \to & \mathbb{R}^4 \\ & \begin{bmatrix} a & b \\ c & d \end{bmatrix} & \mapsto & (a,b,c,d) \end{array}$$
é um isomorfismo entre $\mathbb{R}^{2\times 2}$ e $\mathbb{R}^4$.[12]

- É fácil ver que,
$$\begin{array}{rccc} L : & \mathcal{P}_4(\mathbb{R}) & \to & \mathbb{R}^4 \\ & a + bt + ct^2 + dt^3 & \mapsto & (a,b,c,d) \end{array} \tag{6.1}$$
é um isomorfismo entre $\mathcal{P}_4(\mathbb{R})$ e $\mathbb{R}^4$.[13]

- A correspondência biunívoca entre $\mathbb{K}^{m\times n}$ e $\mathbb{K}^{mn}$, dada na subseção 3.1.2 (para $\mathbb{K} = \mathbb{R}$) é, claramente, um isomorfismo entre esses espaços.

Note que, os isomorfismos dos três exemplos supracitados associam a base canônica dos respectivos domínios à base canônica das respectivas imagens. Além disso, as dimensões dos espaços isomorfos, via L, são iguais. Na verdade, vale o seguinte resutado mais geral:

6.2.1 Teorema dos espaços vetoriais isomorfos de dimensões finitas

Dois espaços vetoriais de dimensões finitas, $\mathcal{V}_1$ e $\mathcal{V}_2$, sobre $\mathbb{K}$, são isomorfos (entre si) se, e somente se, $\dim \mathcal{V}_1 = \dim \mathcal{V}_2$. *Em particular, esse isomorfismo associa bases de $\mathcal{V}_1$ à bases de $\mathcal{V}_2$.*

DEMONSTRAÇÃO

Se $\dim \mathcal{V}_1 = \dim \mathcal{V}_2 = n$, então $\mathcal{V}_1$ tem uma base $\{v_1, v_2, \ldots, v_n\}$ e $\mathcal{V}_2$ tem uma base $\{w_1, w_2, \ldots, w_n\}$. Assim, pelo resultado (R10) da seção 6.3 vindoura,[14] existe um isomorfismo $L : \mathcal{V}_1 \to \mathcal{V}_2$ tal que
$$v_1 \mapsto L(v_1) = w_1,\ v_2 \mapsto L(v_2) = w_2,\ \ldots,\ v_n \mapsto L(v_n) = w_n.$$
Para a recíproca, caso $L : \mathcal{V}_1 \to \mathcal{V}_2$ seja um isomorfismo e $\dim \mathcal{V}_1 = n$, $\mathcal{V}_1$ tem uma base $\{v_1, v_2, \ldots, v_n\}$ e (é fácil ver que)
$$\{L(v_1), L(v_2), \ldots, L(v_n)\}$$
é uma base de $\mathrm{Im}(L) = \mathcal{V}_2$.

[11] Por exemplo, se $\mathcal{V}_1 = \mathbb{R}^n$, então $L^{-1} \circ L = I = S_1$ é a semelhança de razão $k = 1$, vista no capítulo 4.

[12] Verifique!

[13] Idem!

[14] Cf. p. 154.

6.2.2 Se $\dim \mathcal{V} = n$, informações sobre $\mathcal{V}$ podem ser obtidas via informações sobre $\mathbb{K}^n$

A idéia é simples. Por um lado, temos algum dado inicial associado à $\mathcal{V}$, aqui chamado de *entrada* ou *input*. Por outro, procuramos alguma informação associada à $\mathcal{V}$, aqui chamada de *saída* ou *output*. Assim, via algum isomorfismo $L : \mathcal{V} \to \mathbb{K}^n$ adequado, transformamos a entrada associada à $\mathcal{V}$ numa entrada associada à $\mathbb{K}^n$ e, após alguns cálculos, obtemos a saída associada à $\mathbb{K}^n$. Finalmente, essa saída é transformada, via L^{-1}, naquela associada à $\mathcal{V}$, inicialmente procurada. Simples assim!

EXEMPLOS

- Considere que queremos determinar duas bases do subespaço de $\mathbb{R}^{2\times 3}$ gerado por

$$\begin{bmatrix} 1 & -1 & 1 \\ -1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 1 & 0 \\ -1 & 1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & -1 & 2 \\ -3 & 1 & 1 \end{bmatrix} \text{ e } \begin{bmatrix} 3 & -3 & 1 \\ 1 & -2 & 3 \end{bmatrix}.$$

Note que, o isomorfismo

$$\begin{array}{rccc} L : & \mathbb{R}^{2\times 3} & \to & \mathbb{R}^6 \\ & \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} & \mapsto & (a,b,c,d,e,f) \end{array}$$

reduz o problema a determinação de duas bases do subespaço de $\mathbb{R}^6$, gerado por

$$(1,-1,1,-1,0,1),\ (-1,1,0,-1,1,-1),\ (1,-1,2,-3,1,1) \text{ e } (3,-3,1,1,-2,3),$$

cuja solução encontra-se no exercício que precede a subseção 4.1.2. Assim, ao utilizarmos as bases obtidas naquele exercício, podemos obter as bases procuradas nesse exemplo. Portanto, uma das bases é constituída pelas matrizes

$$\begin{bmatrix} 1 & -1 & 1 \\ -1 & 0 & 1 \end{bmatrix} \text{ e } \begin{bmatrix} -1 & 1 & 0 \\ -1 & 1 & -1 \end{bmatrix},$$

e a outra é constituída por

$$\begin{bmatrix} 1 & -1 & 0 \\ 1 & -1 & 1 \end{bmatrix} \text{ e } \begin{bmatrix} 0 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix}.$$

- Seja $L : \mathbb{K}^{m\times n} \to \mathbb{K}^{mn}$ o isomorfismo do último exemplo que antecede essa subseção. Podemos definir uma norma em $\mathbb{K}^{m\times n}$ do modo seguinte: Para cada $A \in \mathbb{K}^{m\times n}$, dada como em (3.1),[15] defina $\|A\| := \|L(A)\|$ (onde a segunda norma é a norma em $\mathbb{K}^{mn}$ definida no exercício 1 da seção 5.2).[16] Assim, nesse caso,

$$\begin{aligned} \|A\| &= \sqrt{|a_{11}|^2 + \cdots + |a_{1n}|^2 + |a_{21}|^2 + \cdots + |a_{2n}|^2 + \cdots + |a_{m1}|^2 + \cdots + |a_{mn}|^2} \\ &= \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2} \end{aligned}$$

calcula o "comprimento" de A.

[15] Cf. p. 49.

[16] Confira o exercício 4 da seção 6.4, para um melhor entendimento de *norma*.

Final da parte I do livro

Essencialmente, a parte II desse livro, que iniciaremos na próxima seção, é um curso mais avançado de AL, sobre o que estudamos desde o capítulo 2. No lugar da abordagem matricial empregada até agora, o enfoque das transformações/operadores lineares será considerado.

Para o material já estudado, recomendo que sejam trabalhados apenas os exercícios 1–8 e 13–14 da seção 6.4. Os outros exercícios, mais conceituais, são recomendados apenas para aqueles que estudarem a parte II.

Alguns dos exercícios que podem ser resolvidos com o que estudamos até agora, utilizam alguns conceitos e resultados que serão aprofundados a partir da próxima seção. Especificamente, em analogia ao capítulo 4, caso $\mathcal{V}$ e $\mathcal{W}$ sejam espaços vetoriais arbitrários sobre o corpo $\mathbb{K}$, definimos o *núcleo* e a *imagem* da transformação linear $L : \mathcal{V} \to \mathcal{W}$ como os conjuntos

$$\mathrm{Nu}(L) := \{v \in \mathcal{V} \,:\, L(v) = 0_{\mathcal{W}}\} \;\text{ e }\; \mathrm{Im}(L) := \{w \in \mathcal{W} \,:\, w = L(v) \text{ com } v \in \mathcal{V}\}\,,$$

respectivamente. L é *sobrejetiva* quando $\mathrm{Im}(L) = \mathcal{W}$. L é *injetiva* quando

$$u, v \in \mathcal{V} \text{ com } L(u) = L(v) \Longrightarrow u = v.$$

Esses conceitos já foram estudados para funções reais do ensino médio e, na subseção 6.3.2, demonstraremos que:

- $\mathrm{Nu}(L)$ é um subespaço de $\mathcal{V}$;
- $\mathrm{Im}(L)$ é um subespaço de $\mathcal{W}$;
- L é injetiva $\Longleftrightarrow \mathrm{Nu}(L) = \{0_{\mathcal{V}}\}$.

Agora, "mãos à obra", ou seja, tentem resolver os exercícios que acabamos de recomendar.

6.3 Alguns resultados e algumas demonstrações

Aqui começa a parte II das NA. A parte I será utilizada para exemplificarmos alguns conceitos e resultados que estudaremos na II.

6.3.1 Espaços finitamente gerados, vetores LI e LD, bases

O subespaço de $\mathcal{V}$ gerado pelos m vetores

$$v_j \in \mathcal{V},\, j = 1, \ldots, m,$$

é denotado por

$$\begin{aligned}[v_1, \ldots, v_m] &= \{v \in \mathcal{V} \,:\, v \text{ é uma CL de } v_1, \ldots, v_m\} \\ &= \left\{ \sum_{j=1}^{n} a_j v_j \,:\, a_1, \ldots, a_m \in \mathbb{K} \right\}.\end{aligned}$$

EXERCÍCIO

Verifique que, de fato,

$$0_{\mathcal{V}}, \lambda v, u + v \in [v_1, \ldots, v_m]\,,$$

para quaisquer $\lambda \in \mathbb{K}$ e $u, v \in [v_1, \ldots, v_m]$.

EXEMPLOS

Confira as seções 2.5.1 e 2.5.2 do capítulo 2.

DEFINIÇÃO

Caso $\mathcal{V} = \{0\}$ ou

$$[v_1, \ldots, v_m] = \mathcal{V}, \tag{6.2}$$

dizemos que $\mathcal{V}$ é *finitamente gerado*.

EXEMPLOS

$\mathbb{K}^m = [\mathbf{e}_1, \ldots, \mathbf{e}_m]$ e $\mathcal{P}_m(\mathbb{K}) = \left[1, t, t^2, \ldots, t^{m-1}\right]$ são finitamente gerados.

OBSERVAÇÃO SOBRE (IN)DEPENDÊNCIA LINEAR

É possível que os m geradores de $\mathcal{V}$ em (6.2) sejam LI, ou seja, a única CL nula desses geradores seja a trivial, isto é, a condição

$$\sum_{j=1}^{m} a_j v_j = 0_{\mathcal{V}}, \text{ com } a_j \in \mathbb{K}, \ j = 1, \ldots, m,$$

seja válida apenas para

$$a_1 = \cdots = a_m = 0.$$

Nesse caso, $\{v_1, \ldots, v_m\}$ é uma *base* de $\mathcal{V}$. Portanto, caso sejam LD, os geradores supracitados não formam uma base de $\mathcal{V}$.

EXEMPLO

$$\begin{aligned} \mathbb{K}^3 &= [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3] \\ &= [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_3, \mathbf{e}_2 + \mathbf{e}_3]\,. \end{aligned}$$

Note que, o segundo conjunto de geradores, de cima para baixo, não é uma base de $\mathbb{K}^3$.

(R1) Lema da dependência linear

Se $v_1, \ldots, v_m$ são LD *e $v_1 \neq 0$, então, para algum índice $j \in \{2, \ldots, m\}$, temos:*

1. $v_j \in \left[v_1, \ldots, v_{j-1}\right]$;
2. $\left[v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m\right] = [v_1, \ldots, v_m]$.

EXERCÍCIO

Analise o exemplo (com $\mathbb{K}^3$) supracitado, utilizando (R1).

DEMONSTRAÇÃO DO ITEM 1 DE (R1)

Por hipótese, existem $a_1, \ldots, a_m \in \mathbb{K}$, não todos nulos, tais que

$$\sum_{j=1}^{m} a_j v_j = 0.$$

Na verdade, como $v_1 \neq 0$, algum escalar não nulo deve existir entre $a_2, \ldots, a_m$. Assim, seja j o maior índice em $\{2, \ldots, m\}$ tal que $a_j \neq 0$. Portanto, como

$$v_j = -\frac{a_1}{a_j} v_1 - \cdots - \frac{a_{j-1}}{a_j} v_{j-1}, \tag{6.3}$$

o item 1 é válido.

DEMONSTRAÇÃO DO ITEM 2 DE (R1)

Seja $v \in [v_1, \ldots, v_m]$. Então, existem $b_1, \ldots, b_m \in \mathbb{K}$ tais que

$$v = b_1 v_1 + \cdots + b_j v_j + \cdots + b_m v_m. \tag{6.4}$$

Assim, para provarmos o item 2, basta substituirmos o lado direito de (6.3) no lugar do v_j da equação (6.4).

(R2) Em (6.2),[17] os m geradores majoram n vetores LI

Se $\mathcal{V}$ é finitamente gerado por m vetores e tem n vetores LI, *então $n \leq m$.*

EXERCÍCIO

Via (R2), analise o exemplo (com $\mathbb{K}^3$) supracitado.

DEMONSTRAÇÃO DE (R2)

Considere os m geradores listados em (6.2) e que os vetores da lista $u_1, \ldots, u_n$ são LI em $\mathcal{V}$.

Passo 1. Os $m+1$ vetores da lista ordenada $u_1, v_1, \ldots, v_m$ são LD.[18] Assim, podemos excluir algum vetor, diferente de u_1, da lista supracitada.[19] Portanto, podemos obter uma lista ordenada ℓ_1, começando por u_1, formada por m geradores de $\mathcal{V}$;

Passo 2. Ao incluirmos u_2 em ℓ_1, podemos obter uma lista ordenada com $m+1$ vetores LD,[20] começando por u_1, u_2. Assim, podemos excluir algum vetor, diferente de u_1 e u_2, da lista supracitada.[21] Portanto, podemos obter uma lista ordenada ℓ_2, começando por u_1, u_2, formada por m geradores de $\mathcal{V}$;

Agora, para $j \in \{2, \ldots, m\}$, considere o seguinte:

Passo j. A lista ℓ_{j-1}, com m vetores, obtida no passo $j-1$, gera $\mathcal{V}$. Ao incluirmos o vetor u_j em ℓ_{j-1}, podemos obter uma lista ordenada com $m+1$ vetores LD,[22] começando por $u_1, u_2, \ldots, u_{j-1}, u_j$. Assim, podemos excluir algum vetor, diferente de u_k, $k = 1, \ldots, j$, da lista ordenada supracitada.[23] Portanto, podemos obter uma lista ordenada ℓ_j, começando por $u_1, \ldots, u_j$, formada por m geradores de $\mathcal{V}$.

---

[17] Cf. p. 148.

[18] De fato, $u_1 \in [v_1, \ldots, v_m]$.

[19] Cf. (R1), p. 148.

[20] De fato, como $u_2 \in \mathcal{V}$ e $\mathcal{V}$ é gerado pelos vetores de ℓ_1, u_2 é CL dos vetores de ℓ_1.

[21] Cf. (R1). Além disso, como u_1 e u_2 são LI, nenhum dos dois pode ser excluído.

[22] De fato, como $u_j \in \mathcal{V}$ e $\mathcal{V}$ é gerado pelos vetores de ℓ_{j-1}, u_j é CL dos vetores de ℓ_{j-1}.

[23] Cf. (R1). Além disso, como os vetores da lista $u_1, \ldots, u_j$ são LI, nenhum deles pode ser excluído.

Na lista ℓ_j, à esquerda dos $m-j$ geradores que não foram excluídos da lista dada em (6.2), foi incluída a lista $u_1, \ldots, u_j$. Assim, o número máximo de passos é m e não pode haver mais u's do que v's. De fato, caso o passo $m+1$ ocorresse, como todos os geradores listados em (6.2) foram excluídos no passo m, os vetores da lista

$$u_1, \ldots, u_m, u_{m+1}$$

seriam LD, contradizendo a hipótese da independência linear dos vetores $u_1, \ldots, u_n$.

(R3) Subespaço de espaço finitamente gerado é finitamente gerado

Se $\mathcal{V}$ é finitamente gerado e $\mathcal{W}$ é subespaço de $\mathcal{V}$, então $\mathcal{W}$ é finitamente gerado.

OBSERVAÇÃO

Esse resultado cumpre o prometido no final da subseção 2.5.1.

DEMONSTRAÇÃO DE (R3)

Passo 1. Caso $\mathcal{W} = \{0_\mathcal{V}\}$, $\mathcal{W}$ é finitamente gerado. Caso $\mathcal{W} \neq \{0_\mathcal{V}\}$, ao considerarmos

$$w_1 \in \mathcal{W} - \{0_\mathcal{V}\},$$

w_1 é LI.

Passo 2. Caso $\mathcal{W} = [w_1]$, $\mathcal{W}$ é finitamente gerado. Caso $\mathcal{W} \neq [w_1]$, ao considerarmos

$$w_2 \in \mathcal{W} - [w_1],$$

w_1 e w_2 são LI, por (R1).

Passo 3. Caso $\mathcal{W} = [w_1, w_2]$, $\mathcal{W}$ é finitamente gerado. Caso $\mathcal{W} \neq [w_1, w_2]$, ao considerarmos

$$w_3 \in \mathcal{W} - [w_1, w_2],$$

w_1, w_2 e w_3 são LI, por (R1).

Em geral, para algum índice inteiro $j \geq 2$, considere o seguinte:

Passo j. Caso $\mathcal{W} = \left[w_1, \ldots, w_{j-1}\right]$, $\mathcal{W}$ é finitamente gerado e os vetores da lista ordenada $w_1, \ldots, w_{j-1}$ são LI, pelo passo $j-1$. Caso $\mathcal{W} \neq \left[w_1, \ldots, w_{j-1}\right]$, ao considerarmos

$$w_j \in \mathcal{W} - \left[w_1, \ldots, w_{j-1}\right],$$

os vetores da lista $w_1, \ldots, w_j$ são LI, por (R1).

Assim, caso $\mathcal{W}$ não seja finitamente gerado no passo $j-1$, devemos prosseguir para o passo j. Contudo, como $\mathcal{V}$ é gerado por m vetores, o maior valor que o índice supracitado pode assumir é $m+1$ e, caso o passo $m+1$ ocorra, $\mathcal{W} = [w_1, \ldots, w_m]$ e os vetores da lista $w_1, \ldots, w_m$ são LI, pelo passo m. Portanto, não precisamos prosseguir para o passo $m+2$. De fato, caso prosseguíssemos para o passo $m+2$, os vetores da lista $w_1, \ldots, w_{m+1}$ seriam LI em $\mathcal{V}$, contradizendo (R2).

(R4) Geradores reduzidos a bases

Como uma base de $\mathcal{V}$ é um conjunto finito de geradores LI de $\mathcal{V}$, temos:

Caso seja necessário, para obtermos uma base de $\mathcal{V} = [v_1, \ldots, v_m]$, podemos reduzir a lista de geradores $v_1, \ldots, v_m$.

EXEMPLOS

Confira a subseção 2.5.3.

DEMONSTRAÇÃO DE (R4)

Passo 1. Se $v_1 = 0$, exclua v_1 da lista supracitada. Caso contrário, mantenha-o nessa lista.

Passo 2. Se $v_2 \in [v_1]$, exclua v_2 da lista obtida no passo 1. Caso contrário, mantenha-o nessa lista.

Para $j \in \{2, \ldots, m\}$, considere o seguinte:

Passo *j*. Se $v_j \in [v_1, \ldots, v_{j-1}]$, exclua v_j da lista obtida no passo $j - 1$. Caso contrário, mantenha-o nessa lista.

No passo m, obtemos uma lista ordenada, constituída de $n \leq m$ geradores de $\mathcal{V}$, onde nenhum deles é CL dos vetores que o antecedem (nessa lista). Assim, os n geradores de $\mathcal{V}$ obtidos no passo m são LI, por (R1).

EXERCÍCIO

Para $\mathbb{K}^2 = [(1,2), (3,6), (4,7), (5,9)]$, verifique que o passos supracitados produzem a base $\{(1,2), (4,7)\}$.

(R5) $\mathcal{V}$ finitamente gerado tem base

Esse resultado é uma consequência imediata de (R4).

OBSERVAÇÃO

Até agora, todos os espaços vetoriais apresentados nessas NA tiveram bases. Por (R5), qualquer espaço da forma (6.2) tem base (finita).[24]

(R6) LI estendidos a bases

Caso seja necessário, para obtermos uma base de $\mathcal{V} = [v_1, \ldots, v_m]$, podemos estender uma lista, digamos, $w_1, \ldots, w_n$, constituída de vetores LI *em $\mathcal{V}$.*

EXEMPLO

Considere $w_1 = \mathbf{e}_1$, $w_2 = \mathbf{e}_2 + \mathbf{e}_3$ e $\mathbb{R}^3 = [\mathbf{e}_1, \mathbf{e}_2 + \mathbf{e}_3, \mathbf{e}_1 \times (\mathbf{e}_2 + \mathbf{e}_3)]$.

DEMONSTRAÇÃO DE (R6)

Passo 1. Se v_1 é CL dos vetores da lista supracitada, considere $\mathcal{B}_1 = \{w_1, \ldots, w_n\}$.[25] Caso contrário, considere $\mathcal{B}_1 = \{w_1, \ldots, w_n, v_1\}$.[26]

Passo 2. Se v_2 é CL dos vetores de $\mathcal{B}_1$, considere $\mathcal{B}_2 = \mathcal{B}_1$. Caso contrário, considere $\mathcal{B}_2 = \mathcal{B}_1 \cup \{v_2\}$.[27]

Em geral, para algum índice $j \geq 2$, considere o seguinte:

Passo *j*. Se v_j é CL dos vetores de $\mathcal{B}_{j-1}$, considere $\mathcal{B}_j = \mathcal{B}_{j-1}$. Caso contrário, considere $\mathcal{B}_j = \mathcal{B}_{j-1} \cup \{v_j\}$.[28]

---

[24] Cf. p. 148.

[25] Note que, $n \leq m$, por (R2).

[26] Note que, $n + 1 \leq m$, por (R2).

[27] Note que, o número de vetores de $\mathcal{B}_2$ não pode ser maior que m, por (R2).

[28] Note que, o número de vetores de $\mathcal{B}_j$ não pode ser maior que m, por (R2).

Pelos passos supracitados, os vetores de $\mathcal{B}_j$ são LI e o subespaço gerado por eles inclui os vetores da lista $v_1, \dots, v_j$. Caso $\mathcal{B}_j$ tenha m vetores, o subespaço supracitado inclui os vetores da lista $v_1, \dots, v_m$.[29] Portanto, para algum índice $j \leq m$, os vetores de $\mathcal{B}_j$ geram $\mathcal{V}$ e são LI, ou seja, $\mathcal{B}_j$ é uma base de $\mathcal{V}$.

(R7) Dimensão é o cardinal comum a todas as bases

Caso $\mathcal{V}$ seja finitamente gerado, quaisquer duas bases de $\mathcal{V}$ têm o mesmo número de vetores

DEMONSTRAÇÃO DE (R7)

Sejam $\mathcal{B}_1$ e $\mathcal{B}_2$ bases de $\mathcal{V}$ com m e n vetores, respectivamente. Assim, como temos m vetores que geram $\mathcal{V}$ e n vetores LI em $\mathcal{V}$, $n \leq m$, por (R2). Analogamente, como temos m vetores LI em $\mathcal{V}$ e n vetores que geram $\mathcal{V}$, $m \leq n$, por (R2). Portanto, $m = n$.

NOTAÇÃO

$\dim \mathcal{V} := m = n$ da demonstração de (R7).

EXEMPLO

$\dim \mathbb{K}^n = n$.

(R8) A dimensão do subespaço $\mathcal{W}$ do finitamente gerado $\mathcal{V}$ não excede a dimensão de $\mathcal{V}$

DEMONSTRAÇÃO DE (R8)

Por (R3), (R5) e (R7), temos:

- $\mathcal{W}$ tem $\dim \mathcal{W}$ vetores LI em V;
- $\mathcal{V}$ tem $\dim \mathcal{V}$ vetores que geram V.

Portanto, $\dim \mathcal{W} \leq \dim \mathcal{V}$, por (R2).[30]

OBSERVAÇÃO

Note que,

$$\dim \mathcal{W} = \dim \mathcal{V} \iff \mathcal{W} = \mathcal{V},$$[31]

por (R8).

(R9) Verificação para que $\{v_1, \dots, v_{\dim \mathcal{V}}\}$ seja uma base de $\mathcal{V}$

Para essa verificação, é suficiente que os vetores da lista $v_1, \dots, v_{\dim \mathcal{V}}$ sejam LI *ou gerem $\mathcal{V}$.*

EXEMPLO

$\{(1,2), (3,4)\}$ é base de $\mathbb{R}^2$ pois seus dois vetores são LI.

[29] De fato, caso o subespaço gerado pelos m vetores de $\mathcal{B}_j$ não incluísse algum vetor da lista supracitada, digamos, v_k, com $j < k \leq m$, os $m + 1$ vetores de $\mathcal{B}_j \cup \{v_k\}$ seriam LI em $\mathcal{V}$, contradizendo (R2).

[30] Cf. p. 149.

[31] Cf. a subseção 2.5.6.

DEMONSTRAÇÃO DE (R9)

Ao considerarmos que os vetores da lista supracitada são LI ou que $m = \dim \mathcal{V}$ em (6.2),[32] essa lista pode ser reduzida ou estendida a uma base de $\mathcal{V}$, por (R4) ou (R6). Contudo, é desnecessário estendê-la ou reduzi-la, pois $\dim \mathcal{V}$ é o número de vetores de qualquer base de $\mathcal{V}$, por (R7).

6.3.2 O espaço vetorial $\mathcal{L}(\mathcal{V}, \mathcal{W})$

Sejam $\mathcal{V}$ e $\mathcal{W}$ espaços vetoriais sobre o mesmo corpo $\mathbb{K}$.

NOTAÇÃO

$L : \mathcal{V} \to \mathcal{W}$ é uma transformação linear $\iff L \in \mathcal{L}(\mathcal{V}, \mathcal{W})$.

Para exemplos de $L \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$, confira os capítulos 4 e 5 anteriores. Para exemplos de $L \in \mathcal{L}(\mathcal{P}(\mathbb{K}), \mathcal{P}(\mathbb{K}))$, $L \in \mathcal{L}(\mathcal{P}(\mathbb{K}), \mathbb{K})$ e $L \in \mathcal{L}\big(\mathbb{K}^{\mathbb{N}}, \mathbb{K}^{\mathbb{N}}\big)$, confira os exercícios 5–8 da seção 6.4. Além disso, note que, os isomorfismos L apresentados anteriormente nesse capítulo, são elementos dos conjuntos $\mathcal{L}\big(\mathbb{R}^{2\times 2}, \mathbb{R}^4\big)$ e $\mathcal{L}\big(\mathcal{P}_4(\mathbb{R}), \mathbb{R}^4\big)$.

OBSERVAÇÃO

Considere $\lambda \in \mathbb{K}$ e $T, L \in \mathcal{L}(\mathcal{V}, \mathcal{W})$. Defina, para qualquer $v \in \mathcal{V}$,

$$(\lambda T)(v) = \lambda(T(v)) \text{ e } (T + L)(v) = T(v) + L(v).$$

Demonstra-se que:

- $\lambda T, T + L \in \mathcal{L}(\mathcal{V}, \mathcal{W})$;
- $\mathcal{L}(\mathcal{V}, \mathcal{W})$ é um espaço vetorial sobre $\mathbb{K}$.

De fato, para verificarmos que λT e $T + L$ são transformações lineares, note que, para quaisquer $\alpha \in \mathbb{K}$ e $u, v \in \mathcal{V}$, como T e L são lineares,

$$\begin{aligned}(\lambda T)(\alpha v) &= \lambda(T(\alpha v)) \\ &= \lambda(\alpha(T(v))) \\ &= \lambda\alpha(T(v)) \\ &= \alpha\lambda(T(v)) \\ &= \alpha((\lambda T)(v))\end{aligned}$$

e

$$\begin{aligned}(T + L)(u + v) &= T(u + v) + L(u + v) \\ &= T(u) + T(v) + L(u) + L(v) \\ &= T(u) + L(u) + T(v) + L(v) \\ &= (T + L)(u) + (T + L)(v).\end{aligned}$$

Para verificarmos que $\mathcal{L}(\mathcal{V}, \mathcal{W})$ é um espaço vetorial, note que, seu vetor nulo é a função nula $\mathcal{V} \ni v \mapsto 0_{\mathcal{W}}$, o simétrico aditivo de $L \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ é a função $\mathcal{V} \ni v \mapsto -L(v)$ e as demonstrações das outras propriedades

[32] Cf. p. 148.

são triviais e seguem, basicamente, do fato de que $\mathcal{W}$ é um espaço vetorial. Por exemplo, a comutatividade da adição segue de

$$\begin{aligned}(T+L)(v) &= T(v)+L(v)\\ &= L(v)+T(v)\\ &= (L+T)(v),\end{aligned}$$

para quaisquer $T, L \in \mathcal{L}(\mathcal{V},\mathcal{W})$ e $v \in \mathcal{V}$.

NOTAÇÃO/DEFINIÇÃO

Caso $\mathcal{V}=\mathcal{W}$ em $\mathcal{L}(\mathcal{V},\mathcal{W})$, denota-se

$$\mathcal{L}(\mathcal{V}) := \mathcal{L}(\mathcal{V},\mathcal{V})$$

e define-se

$$L \in \mathcal{L}(\mathcal{V}) \iff L \text{ é um } \textit{operador (linear)} \text{ em } \mathcal{V}.$$

EXEMPLO

Para $A \in \mathbb{K}^{n\times n}$, $\mathbf{x} \overset{L}{\mapsto} A\mathbf{x}$ é um operador em $\mathcal{L}(\mathbb{K}^n)$.

(R10) Construção de funções $L \in \mathcal{L}(\mathcal{V},\mathcal{W})$ via bases de $\mathcal{V}$

Se $\{v_1,\ldots,v_n\}$ é base de $\mathcal{V}$, $w_i \in \mathcal{W}$ está fixado e $a_i \in \mathbb{K}$ é arbitrário, $i=1,\ldots,n$, então

$$a_1v_1+\cdots+a_nv_n \overset{L}{\mapsto} a_1w_1+\cdots+a_nw_n$$

define uma única $L \in \mathcal{L}(\mathcal{V},\mathcal{W})$ tal que

$$L(v_j) = w_j,\ j=1,\ldots,n.$$

EXEMPLO

Em (R10), considere $\mathcal{V}=\mathbb{R}^2$, $\mathcal{W}=\mathbb{R}^3$, $v_j=\mathbf{e}_j$, $j=1,2$, $w_1=(1,1,0)$ e $w_2=(1,0,1)$. Como vimos no capítulo 4, temos a função linear "multiplicação por matriz", dada por

$$\begin{pmatrix} x_1\\ x_2\end{pmatrix} \overset{L}{\mapsto} \begin{pmatrix}1&1\\1&0\\0&1\end{pmatrix}\begin{pmatrix} x_1\\ x_2\end{pmatrix}.$$

DEMONSTRAÇÃO DE (R10)

L está bem definida. De fato, qualquer $v \in \mathcal{V}$ pode ser escrito, de modo único, como CL dos vetores da base, ou seja, existem únicos $a_1,\ldots,a_n \in \mathbb{K}$, dependentes de v, tais que

$$v=\sum_{j=1}^{n} a_jv_j. \tag{6.5}$$

Por outro lado, é fácil ver que, se $a_j=1$ e os a's com outros índices são nulos, temos

$$v_j \overset{L}{\mapsto} w_j.$$

Em relação à linearidade de L, sejam $\lambda \in \mathbb{K}$ e $u, v \in \mathcal{V}$ como em (6.5), onde os escalares de u tem c's nos lugares dos a's de v. Então,

$$\begin{aligned} L(\lambda v) &= L\left(\sum_{j=1}^{n} (\lambda a_j)\, v_j\right) \\ &= \sum_{j=1}^{n} (\lambda a_j)\, w_j \\ &= \lambda \sum_{j=1}^{n} a_j w_j \\ &= \lambda L(v) \end{aligned}$$

e

$$\begin{aligned} L(u+v) &= L\left(\sum_{j=1}^{n} (c_j + a_j)\, v_j\right) \\ &= \sum_{j=1}^{n} (c_j + a_j)\, w_j \\ &= \sum_{j=1}^{n} c_j w_j + \sum_{j=1}^{n} a_j w_j \\ &= L(u) + L(v). \end{aligned}$$

Quanto a unicidade de L, seja $S \in \mathcal{L}(\mathcal{V}, \mathcal{U})$ tal que $v_j \overset{S}{\mapsto} w_j$, $j = 1, \ldots, n$. Assim,

$$S(v_j) = L(v_j)\,,$$

$j = 1, , \ldots, n$. Portanto, por (6.5) e pela linearidade de S e L,

$$\begin{aligned} S(v) &= S(a_1 v_1 + \cdots + a_n v_n) \\ &= a_1 S(v_1) + \cdots + a_n S(v_n) \\ &= a_1 L(v_1) + \cdots + a_n L(v_n) \\ &= L(a_1 v_1 + \cdots + a_n v_n) \\ &= L(v). \end{aligned}$$

(R11) $\mathrm{Nu}(L)$ **é subespaço de** $\mathcal{V}$ **e** $\mathrm{Im}(L)$ **é subespaço de** $\mathcal{W}$

Em analogia ao capítulo 4, para $L \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, temos:

(I) *O núcleo de L, isto é, o conjunto*

$$\mathrm{Nu}(L) = \{v \in \mathcal{V} \,:\, L(v) = 0_{\mathcal{W}}\}\,,$$

é subespaço de $\mathcal{V}$*;*

(II) *A imagem de L, isto é, o conjunto*

$$\mathrm{Im}(L) = \{L(v) \,:\, v \in \mathcal{V}\}\,,$$

é subespaço de W.

OBSERVAÇÃO

Para exemplos, confira o capítulo 4 e a seção 6.4.

DEMONSTRAÇÃO DO ITEM (I) DE (R11)[33]

Note que, $0_{\mathcal{V}} \in \mathrm{Nu}(L)$. De fato, como

$$\begin{aligned} L(0_{\mathcal{V}}) &= L(0_{\mathcal{V}} + 0_{\mathcal{V}}) \\ &= L(0_{\mathcal{V}}) + L(0_{\mathcal{V}}) \,, \end{aligned}$$

$L(0_{\mathcal{V}}) = 0_{\mathcal{W}}$. Agora, para $\lambda \in \mathbb{K}$ e $u, v \in \mathrm{Nu}(L)$, temos $\lambda u, u + v \in \mathrm{Nu}(L)$. De fato,

$$\begin{aligned} L(\lambda u) &= \lambda L(u) \\ &= \lambda \cdot 0_{\mathcal{W}} \\ &= 0_{\mathcal{W}} \end{aligned}$$

e

$$\begin{aligned} L(u+v) &= L(u) + L(v) \\ &= 0_{\mathcal{W}} + 0_{\mathcal{W}} \\ &= 0_{\mathcal{W}}. \end{aligned}$$

(R12) Injetividade e núcleo

$$\boxed{L \in \mathcal{L}(\mathcal{V}, \mathcal{W}) \textit{ é injetiva} \Longleftrightarrow \mathrm{Nu}(L) = \{0_{\mathcal{V}}\}.}$$

OBSERVAÇÃO

A “injetividade” de L significa o seguinte:

$$u, v \in \mathcal{V} \textit{ com } L(u) = L(v) \Longrightarrow u = v.$$

EXEMPLOS

Confira exercícios 5–8 da seção 6.4.

DEMONSTRAÇÃO DE (R12)

Para a implicação “$\Longrightarrow$”, suponha que L seja injetiva. Logo, como $\{0_{\mathcal{V}}\} \subset \mathrm{Nu}(L)$, resta provarmos que $\mathrm{Nu}(L) \subset \{0_{\mathcal{V}}\}$. Assim, ao considerarmos $v \in \mathrm{Nu}(L)$, ou seja,

$$L(v) = 0_{\mathcal{W}},$$

e a igualdade

$$L(0_{\mathcal{V}}) = 0_{\mathcal{W}},$$

obtida na demonstração do item (I) de (R11), temos

$$L(v) = L(0_{\mathcal{V}}) \,.$$

Portanto, pela suposição supracitada,

$$v = 0_{\mathcal{V}},$$

ou seja, $v \in \{0_{\mathcal{V}}\}$.
Para a implicação “$\Longleftarrow$”, suponha que

$$\mathrm{Nu}(L) = \{0_{\mathcal{V}}\} \,.$$

[33] A demonstração do item (II) fica como exercício.

Portanto, como

$$\begin{aligned} L(u) = L(v) &\Longrightarrow L(u) - L(v) = 0_{\mathcal{W}} \\ &\Longrightarrow L(u-v) = 0_{\mathcal{W}} \\ &\Longrightarrow u - v \in \mathrm{Nu}(L) \\ &\Longrightarrow u - v = 0_{\mathcal{V}} \\ &\Longrightarrow u = v, \end{aligned}$$

para $u, v \in \mathcal{V}$, L é injetiva.[34]

Além de injetividade, precisaremos do conceito de "sobrejetividade", ou seja:

DEFINIÇÃO

$L \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ é *sobrejetiva* $\Longleftrightarrow \mathrm{Im}(L) = \mathcal{W}$.

EXEMPLOS

Confira exercícios 5–8 da seção 6.4.

(R13) Isomorfismo e subespaços triviais

$$\boxed{L \in \mathcal{L}(\mathcal{V}, \mathcal{W}) \text{ é isomorfismo} \Longleftrightarrow \mathrm{Nu}(L) = \{0_{\mathcal{V}}\} \text{ e } \mathrm{Im}(L) = \mathcal{W}.}$$

DEMONSTRAÇÃO

Considere (R12), a definição supracitada e que, para $L \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, conforme o livro do AXLER citado no capítulo 1, vale o seguinte resultado:

$$\boxed{L \text{ é isomorfismo} \Longleftrightarrow L \text{ é injetiva e sobrejetiva.}}$$

EXEMPLOS

Verifique a validade de (R13) para os isomorfismos apresentados na seção 6.2 e para o isomorfismo do exercício 13 da seção 6.4.

(R14) Nulidade mais posto

$$\boxed{\mathcal{V} \text{ é finitamente gerado e } L \in \mathcal{L}(\mathcal{V}, \mathcal{W}) \Longrightarrow \dim \mathrm{Nu}(L) + \dim \mathrm{Im}(L) = \dim \mathcal{V}.}$$

EXEMPLOS

Confira a subseção 4.1.1.

DEMONSTRAÇÃO DE (R14)

Por (I) de (R11), (R3) e (R5),[35] podemos considerar uma base

$$\{v_1, \ldots, v_m\} \tag{6.6}$$

de $\mathrm{Nu}(L)$. Assim, por (R6),[36] podemos estender a lista $v_1, \ldots, v_m$, caso necessário, para obtermos uma base

$$\{v_1, \ldots, v_m, u_1, \ldots, u_n\} \tag{6.7}$$

[34] Note que, na segunda implicação, de cima para baixo, usamos a linearidade de L. Na penúltima, de cima para baixo, usamos a suposição supracitada.

[35] Cf. pp. 155, 150 e 151, respectivamente.

[36] Cf. p. 151.

de $\mathcal{V}$.[37] Para concluirmos a demonstração, basta provarmos que

$$\{L(u_1), \ldots, L(u_n)\}$$

é base de $\mathrm{Im}(L)$. Assim, por um lado, essas n imagens por L são LI. De fato, como, para $a_1, \ldots, a_n \in \mathbb{K}$,

$$\begin{aligned}\sum_{i=1}^{n} a_i L(u_i) = 0_{\mathcal{W}} &\Longrightarrow L\left(\sum_{i=1}^{n} a_i u_i\right) = 0_{\mathcal{W}}\\ &\Longrightarrow \sum_{i=1}^{n} a_i u_i \in \mathrm{Nu}(L)\\ &\Longrightarrow \sum_{i=1}^{n} a_i u_i = \sum_{j=1}^{m} c_j v_j,\end{aligned}$$

com $c_1, \ldots, c_m \in \mathbb{K}$ (onde, na primeira implicação, de cima para baixo, utilizamos a linearidade de L e, na última, a consideração inicial de que (6.6) é base de $\mathrm{Nu}(L)$), ou seja,

$$\sum_{j=1}^{m} c_j v_j + \sum_{i=1}^{n} (-a_i)\, u_i = 0_V, \tag{6.8}$$

e os $m+n$ vetores de (6.7) são LI, todos os escalares de (6.8) são nulos e, em particular,

$$a_i = 0,\, i = 1, \ldots, n.$$

Por outro,

$$\mathrm{Im}(L) = [L(u_1), \ldots, L(u_n)]\,.$$

De fato, para $w \in \mathrm{Im}(L)$, existe uma CL

$$v = \sum_{j=1}^{m} \alpha_j v_j + \sum_{i=1}^{n} \beta_j u_i \in \mathcal{V},$$

com escalares em $\mathbb{K}$, tal que

$$\begin{aligned} w &= L(v)\\ &= L\left(\sum_{j=1}^{m} \alpha_j v_j + \sum_{i=1}^{n} \beta_j u_i\right)\\ &= \sum_{j=1}^{m} \alpha_j L(v_j) + \sum_{i=1}^{n} \beta_j L(u_i)\\ &= \sum_{i=1}^{n} \beta_j L(u_i)\,,\end{aligned}$$

onde, na terceira igualdade, de cima para baixo, utilizamos a linearidade de L e, na última, a consideração inicial de que (6.6) é base de $\mathrm{Nu}(L)$.

6.3.3 $\dim \mathcal{V} = n$ e $\dim \mathcal{W} = m \Longrightarrow \mathcal{L}(\mathcal{V}, \mathcal{W})$ e $\mathbb{K}^{m\times n}$ são isomorfos

Sejam $\mathcal{B} = \{v_1, \ldots, v_n\}$ e $\mathcal{B}' = \{w_1, \ldots, w_m\}$ bases ordenadas de $\mathcal{V}$ e $\mathcal{W}$, respectivamente.[38] Considere a transformação linear

$$\begin{array}{rcccc} \mathcal{M}(-, \mathcal{B}, \mathcal{B}') & : & \mathcal{L}(\mathcal{V}, \mathcal{W}) & \longrightarrow & \mathbb{K}^{m\times n}\\ & & L & \longmapsto & \mathcal{M}(L, \mathcal{B}, \mathcal{B}') \end{array} \tag{6.9}$$

[37] Caso não seja necessário estendermos a lista supracitada, é fácil ver que $\mathrm{Nu}(L) = \mathcal{V}$ e $\mathrm{Im}(L) = \{0_{\mathcal{W}}\}$. Portanto, (R14) é um resultado válido, nesse caso.

[38] Ao escrevermos qualquer vetor de $\mathcal{V}$ (respectivamente, $\mathcal{W}$) como CL dos vetores de sua base ordenada, a ordem em que esses vetores são apresentados nessa base - ordem crescente de seus índices - é mantida na CL.

tal que, para cada $j \in \{1, \ldots, n\}$, se

$$L(v_j) = a_{1j}w_1 + \cdots + a_{mj}w_m,$$

com $a_{1j}, \ldots, a_{mj} \in \mathbb{K}$, então a j-ésima coluna de $\mathcal{M}(L, \mathcal{B}, \mathcal{B}')$ é dada por

$$\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}.$$

EXEMPLOS

Confira o capítulo 4, onde $\mathcal{V} = \mathbb{R}^n$, $\mathcal{W} = \mathbb{R}^m$ e

$$\mathcal{M}(L, \mathcal{B}, \mathcal{B}') = [L]^{\mathcal{B}}_{\mathcal{B}'}.$$

OBSERVAÇÕES

- Caso, num dado contexto, as bases tenham sido previamente determinadas, podemos denotar $\mathcal{M}(-) := \mathcal{M}(-, \mathcal{B}, \mathcal{B}')$.[39]

- $\mathcal{M}(-)$ é linear, ou seja,

$$\mathcal{M}(\lambda T) = \lambda\, \mathcal{M}(T) \ \text{ e } \ \mathcal{M}(T + L) = \mathcal{M}(T) + \mathcal{M}(L),$$

  para quaisquer $\lambda \in \mathbb{K}$ e $T, L \in \mathcal{L}(\mathcal{V}, \mathcal{W})$.[40]

- Para que $\mathcal{M}(-)$ seja um isomorfismo, basta que seja invertível.[41]

- Como no capítulo 4, podemos demonstrar que

$$\mathcal{M}(TL, \mathcal{B}, \mathcal{B}'') = \mathcal{M}(T, \mathcal{B}', \mathcal{B}'')\, \mathcal{M}(L, \mathcal{B}, \mathcal{B}'), \tag{6.10}$$

  para quaisquer transformações $L \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ e $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ e quaisquer bases ordenadas $\mathcal{B}$ de $\mathcal{U}$, $\mathcal{B}'$ de $\mathcal{V}$ e $\mathcal{B}''$ de $\mathcal{W}$, onde o primeiro produto é a composição de funções e o segundo é a multiplicação de matrizes.

(R15) Transformações lineares e produtos de matrizes por vetores

Toda transformação linear é uma "multiplicação por matriz".

DEMONSTRAÇÃO DE (R15)

Sejam $v \in \mathcal{V}$ e $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ tais que

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n.$$

A *matriz de v na base $\mathcal{B}$* é definida por

$$\mathcal{M}(v, \mathcal{B}) = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}.^{42} \tag{6.11}$$

[39] Contudo, nesse caso, não podemos esquecer que $\mathcal{M}(-)$ depende das bases $\mathcal{B}$ e $\mathcal{B}'$.

[40] Verifique!

[41] A invertibilidade de $\mathcal{M}(-)$ será demonstrada em (R16), p. 160. A linearidade de $\mathcal{M}(-)^{-1}$ é consequência da propriedade simétrica do exercício 12 da seção 6.4.

[42] Caso, num dado contexto, $\mathcal{B}$ tenha sido previamente determinada, podemos denotar $\mathcal{M}(v) := \mathcal{M}(v, \mathcal{B})$.

Assim,

$$\mathcal{M}\big(L(v),\mathcal{B}'\big) = \mathcal{M}\big(L,\mathcal{B},\mathcal{B}'\big)\,\mathcal{M}(v,\mathcal{B}).^{43} \tag{6.12}$$

De fato, seja

$$\mathcal{M}(L) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}. \tag{6.13}$$

Logo, para cada índice j,

$$L(v_j) = \sum_{i=1}^{m} a_{ij} w_i.$$

Então,

$$\begin{aligned} L(v) &= \alpha_1 L(v_1) + \cdots + \alpha_n L(v_n) \\ &= \alpha_1 \sum_{i=1}^{m} a_{i1} w_i + \cdots + \alpha_n \sum_{i=1}^{m} a_{in} w_i \\ &= \sum_{i=1}^{m} \left(a_{i1}\alpha_1 + \cdots + a_{in}\alpha_n\right) w_i. \end{aligned}$$

Portanto,

$$\mathcal{M}(L(v)) = \begin{bmatrix} a_{11}\alpha_1 + \cdots + a_{1n}\alpha_n \\ \vdots \\ a_{m1}\alpha_1 + \cdots + a_{mn}\alpha_n \end{bmatrix}. \tag{6.14}$$

Assim, (6.12) é demonstrada ao verificarmos que a matriz coluna (6.14) é igual ao produto da matriz (6.13) pela matriz coluna (6.11).

(R16) $\mathcal{M}(-)$ é invertível

DEMONSTRAÇÃO DE (R16)

Por (R12),[44] para que $\mathcal{M}(-)$ seja injetiva, basta que

$$\mathrm{Nu}\left(\mathcal{M}(-)\right) = \left\{0_{\mathcal{L}(\mathcal{V},\mathcal{W})}\right\}.$$

Assim, considere $L \in \mathcal{L}(\mathcal{V},\mathcal{W})$ tal que

$$\mathcal{M}(L) = 0_{\mathbb{K}^{m\times n}}.$$

Então, para cada índice j,

$$L(v_j) = 0_{\mathcal{W}}.$$

Portanto, como $\mathcal{B}$ é uma base de $\mathcal{V}$, temos, para cada $v \in \mathcal{V}$, $L(v) = 0_{\mathcal{W}}$, isto é,

$$L = 0_{\mathcal{L}(\mathcal{V},\mathcal{W})}.$$

Agora, para que $\mathcal{M}(-)$ seja sobrejetiva, basta que

$$\mathrm{Im}\left(\mathcal{M}(-)\right) = \mathbb{K}^{m\times n}.$$

Então, para

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{K}^{m\times n}$$

e $L \in \mathcal{L}(\mathcal{V},\mathcal{W})$ tal que, para cada índice j,

$$L(v_j) = \sum_{i=1}^{m} a_{ij} w_i,$$

temos, claramente,

$$\mathcal{M}(L) = A.$$

[43]Cf. (4.8), p. 97.

[44]Cf. p. 156.

(R17) $\dim \mathcal{L}(\mathcal{V}, \mathcal{W}) = (\dim \mathcal{V})(\dim \mathcal{W})$

DEMONSTRAÇÃO DE (R17)

Como $\dim \mathcal{V} = n$, $\dim \mathcal{W} = m$, $\dim \mathbb{K}^{m\times n} = mn$ e $\mathcal{L}(\mathcal{V}, \mathcal{W}) \overset{\mathcal{M}(-)}{\longrightarrow} \mathbb{K}^{m\times n}$ é um isomorfismo, temos

$$\begin{aligned} \dim \mathcal{L}(\mathcal{V}, \mathcal{W}) &= \dim \mathbb{K}^{m\times n} \\ &= mn \\ &= (\dim \mathcal{V})(\dim \mathcal{W}), \end{aligned}$$

pelo teorema 6.2.1.[45]

[45]Cf. p. 145.

6.4 Exercícios

1. Considere que 0 é o polinômio nulo e m é um inteiro positivo. O conjunto

$$\mathcal{S} = \{0\} \cup \{p(t) \in \mathcal{P}(\mathbb{K}) \ : \ \text{o grau de } p(t) \text{ é } m\}$$

é um subespaço de $\mathcal{P}(\mathbb{K})$. Essa afirmação é verdadeira ou falsa? Justifique sua resposta.

RESOLUÇÃO

Falsa pois, por exemplo, para $m = 2$, $p(t) = t^2 + t, q(t) = -t^2 + 1 \in \mathcal{S}$ mas $p(t) + q(t) = t + 1 \notin \mathcal{S}$.

2. Seja $\mathbb{R}^{[a,b]}$ o conjunto de todas as funções reais definidas no intervalo real $[a, b]$. Sejam $f, g : [a, b] \to \mathbb{R}$ dois elementos quaisquer de tal conjunto e $\lambda \in \mathbb{R}$. Defina

$$(f + g)(x) := f(x) + g(x) \quad \text{e} \quad (\lambda f)(x) := \lambda f(x)$$

para qualquer $x \in [a, b]$. Portanto, demonstre que, munido dessas operações de adição e multiplicação por escalar, $\mathbb{R}^{[a,b]}$ é um espaço vetorial sobre $\mathbb{R}$.

RESOLUÇÃO

Sejam $f, g, h \in \mathbb{R}^{[a,b]}$, $\lambda, \beta \in \mathbb{R}$ e $x \in [a, b]$ arbitrários. Para verificarmos as propriedades 1–8 que precedem a subseção 6.1.1, basta observarmos que

$$\begin{aligned}(f+g)(x) &= f(x) + g(x)\\ &= g(x) + f(x)\\ &= (g+f)(x);\end{aligned}$$

$$\begin{aligned}[(f+g)+h](x) &= (f+g)(x) + h(x)\\ &= f(x) + g(x) + h(x)\\ &= f(x) + (g+h)(x)\\ &= [f + (g+h)](x);\end{aligned}$$

o vetor nulo é a função real nula definida em $[a, b]$;

$$(-f)(x) = -f(x);$$

$$\begin{aligned}[\lambda(f+g)](x) &= \lambda(f+g)(x)\\ &= \lambda(f(x) + g(x))\\ &= \lambda f(x) + \lambda g(x)\\ &= (\lambda f)(x) + (\lambda g)(x)\\ &= (\lambda f + \lambda g)(x);\end{aligned}$$

$$\begin{aligned}[(\lambda+\beta)f](x) &= (\lambda+\beta)f(x)\\ &= \lambda f(x) + \beta f(x)\\ &= (\lambda f)(x) + (\beta f)(x)\\ &= (\lambda f + \beta f)(x);\end{aligned}$$

$$\begin{aligned}[(\lambda\beta)f](x) &= (\lambda\beta)f(x)\\ &= \lambda(\beta f(x))\\ &= \lambda(\beta f)(x)\\ &= [\lambda(\beta f)](x)\end{aligned}$$

e

$$(1f)(x) = f(x).$$

3. Seja $C[a,b]$ o conjunto das funções reais contínuas definidas no intervalo real $[a,b]$. Sejam $f, g \in C[a,b]$ e $\lambda \in \mathbb{R}$. Defina $f+g$ e λg como na questão 2 dessa seção. Então, verifique que $C[a,b]$ é um espaço vetorial real.

   RESOLUÇÃO

   Basta verificarmos que $C[a,b]$ é subespaço de $\mathbb{R}^{[a,b]}$. De fato, pelo "cálculo das funções reais de uma variável real", a função nula é contínua, somas de funções contínuas são contínuas e produtos de números reais por funções contínuas são funções contínuas.

4. Seja $\mathcal{V}$ um espaço vetorial sobre $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Uma função

$$\begin{array}{rccc}\langle\ ,\ \rangle: & \mathcal{V}\times\mathcal{V} & \to & \mathbb{K}\\ & (u,v) & \mapsto & \langle u,v\rangle\end{array}$$

   é chamada de *produto interno real* (respectivamente, *complexo*) *em* $\mathcal{V}$ quando $\mathbb{K} = \mathbb{R}$ (respectivamente, $\mathbb{K} = \mathbb{C}$) e, para quaisquer $u, v, w \in \mathcal{V}$ e qualquer $\lambda \in \mathbb{K}$, as seguintes condições são satisfeitas:

   · $\langle u,v\rangle = \overline{\langle v,u\rangle}$; (SIMETRIA CONJUGADA)
   · $\langle \lambda u,v\rangle = \lambda\langle u,v\rangle$; $\langle u+v,w\rangle = \langle u,w\rangle + \langle v,w\rangle$; (LINEARIDADE NO PRIMEIRO FATOR)[46]
   · $\langle u,u\rangle \geq 0$; $\langle u,u\rangle = 0 \iff u = 0$. (NÃO NEGATIVIDADE)

   EXEMPLO

   Como vimos nos capítulos 2 e 5,

$$\mathcal{V}\times\mathcal{V} \ni (\mathbf{x},\mathbf{y}) \mapsto \langle\mathbf{x},\mathbf{y}\rangle = \mathbf{x}\cdot\mathbf{y} \in \mathbb{K}$$

   é um produto interno real (respectivamente, complexo) para $\mathcal{V} = \mathbb{R}^n$ e $\mathbb{K} = \mathbb{R}$ (respectivamente, $\mathcal{V} = \mathbb{C}^n$ e $\mathbb{K} = \mathbb{C}$).

   DEFINIÇÃO

   Caso $\mathcal{V}$ seja munido de um produto interno $\langle\ ,\ \rangle$, dizemos que a função

$$\mathcal{V} \ni v \xmapsto{\|\ \|} \|v\| = \sqrt{\langle v,v\rangle} \in \mathbb{R}$$

   é uma *norma em* $\mathcal{V}$.

   EXEMPLOS

[46] Verifique que a *linearidade no segundo fator* é dada por

$$\langle u,\lambda v\rangle = \overline{\lambda}\langle u,v\rangle;\ \langle u,v+w\rangle = \langle u,v\rangle + \langle u,w\rangle.$$

· Como vimos nos capítulos 2 e 5,

$$\mathcal{V} \ni \mathbf{x} \longmapsto \|\mathbf{x}\| = \sqrt{|x_1|^2 + \cdots + |x_n|^2} \in \mathbb{R} \tag{6.15}$$

é uma norma em $\mathcal{V} = \mathbb{K}^n$.

· Pelo isomorfismo $L : \mathbb{K}^{m\times n} \to \mathbb{K}^{mn}$, dado no último exemplo da subseção 6.2.2, podemos obter a norma

$$\mathcal{V} \ni A \longmapsto \|A\| := \|L(A)\| \in \mathbb{R}$$

em $\mathcal{V} = \mathbb{K}^{m\times n}$, onde $\|L(A)\|$ é calculada como em (6.15), ou seja,

$$\|A\| = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2}.$$

Note que, $\|A\| = \sqrt{\langle A, A\rangle}$, onde o produto interno $\langle\ ,\ \rangle$ é definido por

$$\mathcal{V} \times \mathcal{V} \ni (A, B) \longmapsto \langle A, B\rangle := \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\overline{b_{ij}} \in \mathbb{K}.$$

OBSERVAÇÕES

- Pela definição de produto interno, dada no início desse exercício, podemos verificar que uma norma em $\mathcal{V}$ satisfaz as condições seguintes:[47]
  - (NÃO NEGATIVIDADE) Para cada $v \in \mathcal{V}$, $\|v\| \geq 0$ e
  $$\|v\| = 0 \Longleftrightarrow v = 0_{\mathcal{V}};$$
  - Para cada $\lambda \in \mathbb{K}$ e cada $v \in \mathcal{V}$,
  $$\|\lambda v\| = |\lambda|\|v\|;$$
  - (DESIGUALDADE TRIANGULAR) Para todos $u, v \in \mathcal{V}$,
  $$\|u + v\| \leq \|u\| + \|v\|.$$
- Analogamente aos espaços $\mathbb{R}^n$ e $\mathbb{C}^n$, munidos, respectivamente, dos produtos internos dos capítulos 2 e 5,[48] caso $\langle\ ,\ \rangle$ seja um produto interno real (respectivamente, complexo) definido num espaço vetorial $\mathcal{V}$ sobre $\mathbb{R}$ (respectivamente, $\mathbb{C}$), temos:
  - $u, v \in \mathcal{V}$ são ditos *ortogonais* (*entre si*) quando $\langle u, v\rangle = 0$;
  - $\mathcal{V} \supset \{e_1, \ldots, e_n\}$ é uma base *ortonormal* de $\mathcal{V}$ quando seus vetores são *unitários*, isto é, $\|e_i\| = 1$, $i = 1, \ldots, n$, e quaisquer dois de seus vetores são ortogonais;
  - (*O processo de*) *Gram-Schmidt* também é válido para $\mathcal{V}$, dotado do produto interno supracitado.[49]

[47] Caso seja de vosso interesse, verifique-as!

[48] $\mathbb{K}^n \times \mathbb{K}^n \ni (\mathbf{x}, \mathbf{y}) \longmapsto \mathbf{x} \cdot \mathbf{y} \in \mathbb{K}$.

[49] Confira as seções 2.7 e 5.2. Além disso, para uma demonstração do caso geral, indico o livro do AXLER, referenciado no capítulo 1.

Resolva as seguintes questões:

4.1. Seja $\mathcal{V}$ um espaço vetorial com produto interno $\langle\ ,\ \rangle$. Demonstre a validade da *desigualdade de Cauchy-Schwarz* para vetores de $\mathcal{V}$, ou seja,

$$u, v \in \mathcal{V} \Longrightarrow |\langle u, v\rangle| \leq \|u\|\|v\|.$$ [50]

4.2. Considere o espaço da questão 3 dessa seção.

(a) Demonstre que $C[a, b]$ é munido do produto interno $\langle\ ,\ \rangle$, definido por:

$$f, g \in C[a, b] \Longrightarrow \langle f, g\rangle = \int_a^b f(t)g(t)dt.$$

(b) Determine uma base ortonormal para o subespaço $\mathcal{S}$ de $C[0, 1]$ gerado pelos polinômios $p_1(t) = 1 + t$ e $p_2(t) = t$.

(c) Considere que $\mathcal{S}$ é um subespaço de $C[a, b]$ e $\{f_1, f_2, \ldots, f_r\}$ é uma base ortonormal de $\mathcal{S}$. Então, a projeção ortogonal, ou seja, a melhor aproximação, [51] de $f \in C[a, b]$ em $\mathcal{S}$ é dada por

$$P_{\mathcal{S}}(f) = \langle f, f_1\rangle f_1 + \langle f, f_2\rangle f_2 + \cdots + \langle f, f_r\rangle f_r.$$

Assim, em relação ao item (b) desse exercício, determine o vetor de $\mathcal{S}$ mais próximo da função $f : [0, 1] \to \mathbb{R}$, tal que, $f(t) = e^t$ para todo $t \in [0, 1]$.

RESOLUÇÃO DE 4.1

Caso $v = 0_{\mathcal{V}}$, ambos os lados da desigualdade de Cauchy-Schwarz são nulos. Assim, seja $v \neq 0_{\mathcal{V}}$ e considere $\lambda \in \mathbb{K}$. Note que,

$$\begin{aligned} 0 \leq \langle u + \lambda v, u + \lambda v\rangle &= \langle u, u\rangle + \overline{\lambda}\langle u, v\rangle + \lambda\langle v, u\rangle + \lambda\overline{\lambda}\langle v, v\rangle \\ &= \|u\|^2 + \overline{\lambda\langle v, u\rangle} + \lambda\langle v, u\rangle + |\lambda|^2\|v\|^2 \\ &= \|u\|^2 + 2\operatorname{Re}(\lambda\langle v, u\rangle) + |\lambda|^2\|v\|^2 \end{aligned}$$

com $\operatorname{Re}(z)$ representando a parte real de $z \in \mathbb{K}$. Considere, agora,

$$\begin{aligned} \lambda &= -\frac{\langle u, v\rangle}{\|v\|^2} \\ &= -\frac{\overline{\langle v, u\rangle}}{\|v\|^2}. \end{aligned}$$

Então, como $|z| = |\overline{z}|$, para cada $z \in \mathbb{K}$,

$$\begin{aligned} 0 \leq \|u\|^2 - 2\operatorname{Re}\left(\frac{|\langle v, u\rangle|^2}{\|v\|^2}\right) + \frac{|\langle u, v\rangle|^2}{\|v\|^2} &= \|u\|^2 - \frac{2|\langle u, v\rangle|^2}{\|v\|^2} + \frac{|\langle u, v\rangle|^2}{\|v\|^2} \\ &= \|u\|^2 - \frac{|\langle u, v\rangle|^2}{\|v\|^2}. \end{aligned}$$

RESOLUÇÃO PARCIAL DE 4.2

Primeiramente, note que, como estamos considerando um produto interno real, a propriedade de simetria conjugada é, simplesmente, a propriedade comutativa.

---

[50] SUGESTÃO

Adapte a demonstração dessa desigualdade para o caso $\mathcal{V} = \mathbb{R}^n$, com produto interno usual, da seção 2.3, para demonstrar o caso geral.

[51] Na subseção 7.2.7, estudaremos esses conceitos de modo mais aprofundado!

(a) Sejam $f, g, h \in C[a, b]$ e $\lambda \in \mathbb{R}$ arbitrários. Assim:

$$\begin{aligned}\langle f, g\rangle &= \int_a^b f(t)g(t)dt \\ &= \int_a^b g(t)f(t)dt \\ &= \langle g, f\rangle;\end{aligned}$$

$$\begin{aligned}\langle \lambda f, g\rangle &= \int_a^b (\lambda f(t))g(t)dt \\ &= \int_a^b \lambda(f(t)g(t))dt \\ &= \lambda \int_a^b f(t)g(t)dt \\ &= \lambda\langle f, g\rangle;\end{aligned}$$

$$\begin{aligned}\langle f+g, h\rangle &= \int_a^b (f(t)+g(t))h(t)dt \\ &= \int_a^b (f(t)h(t)+g(t)h(t))dt \\ &= \int_a^b f(t)h(t)dt + \int_a^b g(t)h(t)dt \\ &= \langle f, h\rangle + \langle g, h\rangle\end{aligned}$$

e, como

$$\begin{aligned}\langle f, f\rangle &= \int_a^b f(t)f(t)dt \\ &= \int_a^b f(t)^2 dt\end{aligned}$$

representa a área de uma região do plano delimitada pelo gráfico de uma função contínua não negativa, o eixo das abcissas e as retas $x = a$ e $x = b$, temos

$$\langle f, f\rangle \geq 0 \;\text{ e }\; \langle f, f\rangle = 0 \Longleftrightarrow f = 0.$$

(b) Via Gram-Schmidt,

$$\begin{aligned}q_1(t) &= p_1(t) \\ &= 1+t\end{aligned}$$

e

$$\begin{aligned}q_2(t) &= p_2(t) - \frac{\langle p_2, q_1\rangle}{\langle q_1, q_1\rangle}q_1(t) \\ &= t - \frac{\int_0^1 t(1+t)dt}{\int_0^1 (1+t)^2 dt}(1+t),\end{aligned}$$

que, normalizados, são dados por

$$f_1(t) = \frac{q_1(t)}{\|q_1(t)\|} \quad \text{e} \quad f_2(t) = \frac{q_2(t)}{\|q_2(t)\|},$$

onde

$$\|q_i\| = \sqrt{\langle q_i, q_i\rangle}, i = 1, 2.$$ [52]

[52] Deixaremos o cálculo de $q_2(t)$, $f_1(t)$ e $f_2(t)$ para o leitor!

(c) Basta calcularmos

$$P_{\mathcal{S}}(f) = \langle f, f_1 \rangle f_1 + \langle f, f_2 \rangle f_2,$$

para f_1 e f_2 calculados no item (b) desse exercício.[53]

5. Demonstre que o operador

$$\begin{array}{rccc} L: & \mathcal{P}(\mathbb{R}) & \to & \mathcal{P}(\mathbb{R}) \\ & p(t) & \mapsto & \frac{dp}{dt} \end{array}$$

é linear.[54] Além disso, determine o núcleo de L, isto é, o conjunto

$$\mathrm{Nu}(L) = \{p(t) \in \mathcal{P}(\mathbb{R}) \,:\, L(p(t)) = 0\}.$$

L é injetivo? É sobrejetivo?

DICA

Pelo "cálculo das funções reais de uma variável real", temos:

- a derivada da soma de funções diferenciáveis é igual a soma das derivadas dessas funções;
- a derivada do múltiplo escalar de uma função diferenciável é igual ao produto desse escalar pela derivada dessa função.

Para o núcleo, responda a seguinte questão:

Que funções, quando derivadas, se anulam?

Para a injetividade, responda a seguinte questão:

$$\mathrm{Nu}(L) = \{0\}\,?$$

Para a sobrejetividade, responda a seguinte questão:

$\mathrm{Im}(L) = \mathcal{P}(\mathbb{R})$?[55]

6. Demonstre que a função

$$\begin{array}{rccc} L: & \mathcal{P}(\mathbb{R}) & \to & \mathbb{R} \\ & p(t) & \mapsto & \int_0^1 p(t)dt \end{array}$$

é linear. L é sobrejetiva? É injetiva?

DICA

Pelo "cálculo das funções reais de uma variável real", temos:

- a integral da soma de funções contínuas é igual a soma das integrais dessas funções;
- a integral do múltiplo escalar de uma função contínua é igual ao produto desse escalar pela integral dessa função.

---

[53] Deixamos o cálculo dessa projeção para o leitor!

[54] dp/dt é a derivada do polinômio $p(t)$ em relação à variável t.

[55] Todo polinômio $a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$, com coeficientes em $\mathbb{R}$, é a derivada do polinômio

$$a_0 t + \frac{a_1}{2} t^2 + \frac{a_2}{3} t^3 + \cdots + \frac{a_n}{n+1} t^{n+1}.$$

Para a sobrejetividade, responda: $\text{Im}(L) = \mathbb{R}$?[56]
Para a injetividade, responda: $\text{Nu}(L) = \{0\}$?[57]

7. Demonstre que o operador

$$\begin{array}{rccc} L: & \mathcal{P}(\mathbb{R}) & \to & \mathcal{P}(\mathbb{R}) \\ & p(t) & \mapsto & t^2 p(t) \end{array}$$

é linear. L é injetivo? É sobrejetivo?

RESOLUÇÃO PARCIAL

A linearidade de L é trivial.[58] Para a injetividade, devemos responder a pergunta: $\text{Nu}(L) = \{0\}$?
Agora, note que, $L(p(t)) = t^2 p(t)$ tem grau no mínimo 2 se $p(t) \neq 0$. Logo, $L(p(t)) = 0$ apenas se $p(t) = 0$. Então, L é injetiva.
Quanto a sobrejetividade, devemos responder a pergunta: $\text{Im}(L) = \mathcal{P}(\mathbb{R})$?
A resposta é "não!", pois polinômios de grau 1 ou polinômios não nulos de grau 0 não podem ser imagens, por essa L, de polinômios em $\mathcal{P}(\mathbb{R})$!

8. Demonstre que o operador

$$\begin{array}{rccc} L: & \mathbb{K}^{\mathbb{N}} & \longrightarrow & \mathbb{K}^{\mathbb{N}} \\ & (x_1, x_2, \dots) & \longmapsto & (x_2, x_3, \dots) \end{array}$$

é linear. L é sobrejetivo? É injetivo?

RESOLUÇÃO

L é linear pois, para quaisquer $\lambda \in \mathbb{K}$ e $x, y \in \mathbb{K}^{\mathbb{N}}$, temos

$$\begin{aligned} L(\lambda x) &= L(\lambda x_1, \lambda x_2, \dots) \\ &= (\lambda x_2, \lambda x_3, \dots) \\ &= \lambda (x_2, x_3, \dots) \\ &= \lambda L(x_1, x_2, \dots) \\ &= \lambda L(x) \end{aligned}$$

e

$$\begin{aligned} L(x+y) &= L(x_1 + y_1, x_2 + y_2, \dots) \\ &= (x_2 + y_2, x_3 + y_3, \dots) \\ &= (x_2, x_3, \dots) + (y_2, y_3, \dots) \\ &= L(x_1, x_2, \dots) + L(y_1, y_2, \dots) \\ &= L(x) + L(y). \end{aligned}$$

L é sobrejetivo pois $\text{Im}(L) = \mathbb{K}^{\mathbb{N}}$, ou seja, para cada $y \in \mathbb{K}^{\mathbb{N}}$, existe $x \in \mathbb{K}^{\mathbb{N}}$ tal que $L(x) = y$. De fato, para pré-imagem de y por L, considere a sequência $x = (x_1, x_2, \dots)$ cujo primeiro termo seja dado por algum escalar real ($x_1 = 0$, por exemplo) e cujos termos restantes sejam dados por $x_2 = y_1$, $x_3 = y_2$,

[56] Todo número real r pode ser escrito da forma $\int_0^1 r\, dt$.
[57] $\int_0^1 (2t-1) dt = 0$.
[58] Verifique!

etc.[59] Assim,

$$\begin{aligned} L(x) &= L(x_1, x_2, \ldots) \\ &= (x_2, x_3, \ldots) \\ &= (y_1, y_2, \ldots) \\ &= y. \end{aligned}$$

L não é injetivo pois $\mathrm{Nu}(L) \neq \{0\}$. De fato, note que,

$$\mathrm{Nu}(L) = \{a(1,0,0,\ldots) \, : \, a \in \mathbb{K}\}.$$

9. Dê exemplo de alguma $L \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ que não seja sobrejetiva nem injetiva.

RESOLUÇÃO

Seja L o vetor nulo de $\mathcal{L}(\mathcal{V}, \mathcal{W})$, ou seja,

$$L(v) = 0_{\mathcal{W}} \;\; \forall v \in \mathcal{V}.$$

Portanto, como $\mathrm{Nu}(L) = \mathcal{V}$ e $\mathrm{Im}(L) = \{0_{\mathcal{W}}\}$, L não é injetiva (caso $\mathcal{V} \neq \{0_{\mathcal{V}}\}$) nem sobrejetiva (caso $\mathcal{W} \neq \{0_{\mathcal{W}}\}$).

10. O operador

$$\begin{array}{rccc} T & : & \mathbb{C}^{2\times 2} & \longrightarrow & \mathbb{C}^{2\times 2} \\ & & A & \longmapsto & \overline{A} \end{array}$$

não é linear. Essa afirmação é verdadeira ou falsa? Justifique corretamente a sua resposta.

RESOLUÇÃO

A afirmação é verdadeira. De fato, se $\lambda \in \mathbb{C}$ e $A \in \mathbb{C}^{2\times 2}$, então

$$\begin{aligned} T(\lambda A) &= \overline{\lambda A} \\ &= \overline{\lambda}\,\overline{A} \\ &= \overline{\lambda} T(A) \\ &\neq \lambda T(A), \end{aligned}$$

a menos que λ seja real.

11. Sejam $\mathcal{V}_1$, $\mathcal{V}_2$ e $\mathcal{V}_3$ espaços vetoriais sobre o corpo $\mathbb{K}$. Demonstre que se $L_1 : \mathcal{V}_1 \to \mathcal{V}_2$ e $L_2 : \mathcal{V}_2 \to \mathcal{V}_3$ são funções lineares, então a composição

$$\begin{array}{ccccccc} L_2 L_1 & := & L_2 \circ L_1 & : & \mathcal{V}_1 & \longrightarrow & \mathcal{V}_3 \\ & & & & v & \longmapsto & L_2(L_1(v)) \end{array}$$

é linear.[60]

RESOLUÇÃO

Sejam $u, v \in \mathcal{V}_1$ e $\lambda \in \mathbb{K}$ arbitrários. Portanto,

$$\begin{aligned} (L_2 L_1)(u + v) &= L_2(L_1(u + v)) \\ &= L_2(L_1(u) + L_1(v)) && (L_1 \text{ É LINEAR}) \\ &= L_2(L_1(u)) + L_2(L_1(v)) && (L_2 \text{ É LINEAR}) \\ &= (L_2 L_1)(u) + (L_2 L_1)(v) \end{aligned}$$

[59] Isto é, $x_{n+1} = y_n$ para cada inteiro positivo n.

[60] Cf. a afirmação 2 que precede a subseção 4.1.1 e os dois exercícios que acompanham aquela afirmação.

e

$$\begin{aligned}(L_2L_1)(\lambda v) &= L_2(L_1(\lambda v)) \\ &= L_2(\lambda L_1(v)) \quad (L_1 \text{ É LINEAR}) \\ &= \lambda\,(L_2(L_1(v)))\ (L_2 \text{ É LINEAR}) \\ &= \lambda\,(L_2L_1)(v).\end{aligned}$$

12. Demonstre que isomorfismo entre espaços vetoriais é uma relação de equivalência, ou seja, é, simultaneamente:

REFLEXIVA. Todo espaço vetorial $\mathcal{V}$ é isomorfo a si mesmo;

SIMÉTRICA. Se $\mathcal{V}_1$ é isomorfo a $\mathcal{V}_2$, então $\mathcal{V}_2$ é isomorfo a $\mathcal{V}_1$;[61]

TRANSITIVA. Se $\mathcal{V}_1$ é isomorfo a $\mathcal{V}_2$ e $\mathcal{V}_2$ é isomorfo a $\mathcal{V}_3$, então $\mathcal{V}_1$ é isomorfo a $\mathcal{V}_3$.[62]

RESOLUÇÃO PARCIAL

REFLEXIVA. Basta provarmos que a função identidade em $\mathcal{V}$ é um isomorfismo;

SIMÉTRICA. Devemos provar que L^{-1} é linear e tem inversa linear. Como $\left(L^{-1}\right)^{-1} = L$, resta demonstrar que L^{-1} é linear. Para isso, sejam $u, v \in \mathcal{V}_2$ arbitrários. Logo, existem únicos $u', v' \in \mathcal{V}_1$ tais que

$$L(u') = u \text{ e } L(v') = v.$$

Então,

$$\begin{aligned}L^{-1}(u+v) &= L^{-1}\big(L(u') + L(v')\big) \\ &= L^{-1}\big(L(u'+v')\big) \qquad (L \text{ É LINEAR}) \\ &= \left(L^{-1} \circ L\right)(u'+v') \\ &= u' + v',\end{aligned}$$

pois essa composição resulta no operador identidade em $\mathcal{V}_2$. Assim, como

$$u' = L^{-1}(u) \text{ e } v' = L^{-1}(v),$$

temos

$$L^{-1}(u+v) = L^{-1}(u) + L^{-1}(v).$$

Seja, agora, $\lambda \in \mathbb{K}$ arbitrário. Logo,

$$\begin{aligned}L^{-1}(\lambda u) &= L^{-1}\big(\lambda L(u')\big) \\ &= L^{-1}\big(L(\lambda u')\big)\ (L \text{ É LINEAR}) \\ &= \lambda u' \\ &= \lambda L^{-1}(u);\end{aligned}$$

TRANSITIVA. Utilize o exercício 11 dessa seção,[63] a identidade

$$(L_2L_1)^{-1} = L_1^{-1}L_2^{-1}$$

e que, como acabamos de demonstrar, isomorfismo é uma relação simétrica.

[61]Ou seja, se $L : \mathcal{V}_1 \to \mathcal{V}_2$ é um isomorfismo, então $L^{-1} : \mathcal{V}_2 \to \mathcal{V}_1$ também o é.
[62]Ou seja, se $L_1 : \mathcal{V}_1 \to \mathcal{V}_2$ e $L_2 : \mathcal{V}_2 \to \mathcal{V}_3$ são isomorfismos, então $L_2L_1 : \mathcal{V}_1 \to \mathcal{V}_3$ também o é.
[63]6.4.

13. Podemos considerar $\mathbb{K}^n$ como subespaço de $\mathbb{K}^{\mathbb{N}}$, isto é, existe subespaço $\mathcal{S}$ de $\mathbb{K}^{\mathbb{N}}$ isomorfo a $\mathbb{K}^n$. Essa afirmação é verdadeira ou falsa? Justifique corretamente a sua resposta.

RESOLUÇÃO PARCIAL

Verdadeira!
De fato, seja

$$\mathcal{S} = \left\{ x \in \mathbb{K}^{\mathbb{N}} \mid x_i = 0 \text{ para } i > n \right\}.$$

O leitor fica encarregado de verificar que

$$\begin{array}{rccc} L: & \mathbb{K}^n & \to & \mathcal{S} \\ & (x_1, \ldots, x_n) & \mapsto & (x_1, \ldots, x_n, 0, 0, \ldots) \end{array}$$

é um isomorfismo.

14. Determine uma base para o subespaço de $\mathcal{P}_4(\mathbb{R})$ gerado pelos polinômios

$$p(t) = 2 + t + t^2,\ q(t) = 3 + 3t^2 + 6t^3,\ r(t) = 1 + t^2 + 2t^3 \text{ e } s(t) = -t + t^2 + 4t^3.$$

SUGESTÃO

Considere os vetores $(2,1,1,0)$, $(3,0,3,6)$, $(1,0,1,2)$ e $(0,-1,1,4)$. Determine uma base para o subespaço do $\mathbb{R}^4$ gerado por esses vetores, como no exercício 6 da seção 4.3. Converta essa base numa base de $\mathcal{P}_4(\mathbb{R})$ via (6.1).[64]

15. Suponha que $\mathcal{V}$ e $\mathcal{W}$ são espaços vetoriais (sobre $\mathbb{K}$) finitamente gerados. Demonstre que:

   (a) $\dim \mathcal{V} > \dim \mathcal{W} \Longrightarrow \nexists\, L \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ injetiva;[65]

   (b) $\dim \mathcal{V} < \dim \mathcal{W} \Longrightarrow \nexists\, L \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ sobrejetiva.[66]

RESOLUÇÃO

Seja $L \in \mathcal{L}(\mathcal{V}, \mathcal{W})$. Por (R14),[67] temos

$$\dim \mathrm{Nu}(L) + \dim \mathrm{Im}(L) = \dim \mathcal{V}.$$

DEMONSTRAÇÃO DE (a)

Como, por (II) de (R11),[68] $\mathrm{Im}(L)$ é subespaço de $\mathcal{W}$, temos

$$\dim \mathrm{Im}(L) \le \dim \mathcal{W},$$

---

[64] Cf. p. 145.

[65]

$$\mathbb{K}^3 \ni (x_1, x_2, x_3) \overset{L}{\longmapsto} (x_1, x_2) \in \mathbb{K}^2$$

não pode ser injetiva pois, por exemplo, $(0,0,1) \in \mathrm{Nu}(L)$.

[66]

$$\mathbb{K}^2 \ni (x_1, x_2) \overset{L}{\longmapsto} (x_1, x_2, x_1 + x_2) \in \mathbb{K}^3$$

não pode ser sobrejetiva pois, por exemplo, $(0,0,1) \notin \mathrm{Im}(L)$.

[67] Cf. p. 157.

[68] Cf. p. 155.

por (R8).[69] Portanto,

$$\begin{aligned}\dim \mathrm{Nu}(L) &= \dim \mathcal{V} - \dim \mathrm{Im}(L) \\ &\geq \dim \mathcal{V} - \dim \mathcal{W} \\ &> 0.\end{aligned}$$

Assim, $\mathrm{Nu}(L) \neq \{0\}$. Logo, por (R12),[70] L não é injetiva.

DEMONSTRAÇÃO DE (b)

Como, por (I) de (R11), $\mathrm{Nu}(L)$ é subespaço de $\mathcal{V}$, temos

$$\dim \mathrm{Nu}(L) \leq \dim \mathcal{V},$$

por (R8). Logo,

$$\begin{aligned}\dim \mathrm{Im}(L) &= \dim \mathcal{V} - \dim \mathrm{Nu}(L) \\ &\leq \dim \mathcal{V} \\ &< \dim \mathcal{W}.\end{aligned}$$

Portanto, $\mathrm{Im}(L) \subsetneqq \mathcal{W}$. Então, L não é sobrejetiva.

16. Considere uma lista formada por m subespaços de $\mathcal{V}$, digamos $\mathcal{U}_1, \ldots, \mathcal{U}_m$, e o conjunto $\mathcal{U}_1 + \cdots + \mathcal{U}_m$ formado pelas somas de m vetores cujas j-ésimas parcelas pertençam à $\mathcal{U}_j$, $j = 1, \ldots, m$, ou seja,

$$\mathcal{U}_1 + \cdots + \mathcal{U}_m := \left\{ u_1 + \cdots + u_m \mid u_j \in \mathcal{U}_j, j = 1, \ldots, m \right\}. \tag{6.16}$$

O conjunto (6.16) é chamado de *soma de* $\mathcal{U}_1, \ldots, \mathcal{U}_m$.[71]
Por exemplo, se $\mathcal{V} = \mathbb{K}^4, \mathcal{U}_1 = \{x\mathbf{e}_1 \mid x \in \mathbb{K}\}, \mathcal{U}_2 = \{y\mathbf{e}_2 \mid y \in \mathbb{K}\}$ e $\mathcal{U}_3 = \{z\mathbf{e}_3 \mid z \in \mathbb{K}\}$, então $\mathcal{U}_1 + \mathcal{U}_2 + \mathcal{U}_3 = \{(x, y, z, 0) \mid x, y, z \in \mathbb{K}\}$.

(a) Demonstre que $\sum_{j=1}^{m} \mathcal{U}_j$ é o menor subespaço de $\mathcal{V}$ que contém $\mathcal{U}_j$, $j = 1, \ldots, m$.

(b) $\mathcal{U}_1 + \cdots + \mathcal{U}_m$ é uma *soma direta de* $\mathcal{U}_1, \ldots, \mathcal{U}_m$ quando

$$u_1 + \cdots + u_m = u'_1 + \cdots + u'_m \Longrightarrow u_1 = u'_1, \ldots, u_m = u'_m.$$

Nesse caso, denotamos (6.16) por

$$\mathcal{U}_1 \oplus \cdots \oplus \mathcal{U}_m.$$

Por exemplo, se $\mathcal{V} = \mathbb{K}^3, \mathcal{U}_1 = \{x\mathbf{e}_1 + y\mathbf{e}_2 \mid x, y \in \mathbb{K}\}$ e $\mathcal{U}_2 = \{z\mathbf{e}_3 \mid z \in \mathbb{K}\}$, então

$$\mathcal{V} = \mathcal{U}_1 \oplus \mathcal{U}_2.$$

Agora, ainda para esse exemplo, se $\mathcal{U}_3 = \{\lambda\,(\mathbf{e}_2 + \mathbf{e}_3) \mid \lambda \in \mathbb{K}\}$, verifique que $\mathbb{K}^3$ não é soma direta de $\mathcal{U}_1, \mathcal{U}_2$ e $\mathcal{U}_3$.

(c) Demonstre que a soma $\sum_{j=1}^{n} \mathcal{U}_j$ é direta se, e somente se, a CL $0_{\mathcal{V}} = u_1 + \cdots + u_m$ é válida apenas para $u_1 = \cdots = u_m = 0_{\mathcal{V}}$.

(d) Sejam $\mathcal{U}$ e $\mathcal{W}$ subespaços de $\mathcal{V}$. Então, $\mathcal{U} + \mathcal{W}$ é uma soma direta se, e somente se, $\mathcal{U} \cap \mathcal{W} = \{0_{\mathcal{V}}\}$.

[69] Cf. p. 152.

[70] Cf. p. 156.

[71] Nesse exercício, ao considerarmos algum vetor $u_1 + \cdots + u_m$, estará implícito que $u_i \in \mathcal{U}_i$, $i = 1, \ldots, m$.

(e) Sejam $\mathcal{U}$ e $\mathcal{W}$ subespaços de $\mathcal{V}$ tais que $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$. Sejam $\{u_1, \dots, u_m\}$ e $\{w_1, \dots, w_n\}$ bases de $\mathcal{U}$ e $\mathcal{W}$, respectivamente. Demonstre que

$$\{u_1, \dots, u_m, w_1, \dots, w_n\}$$

é uma base de $\mathcal{V}$.

(f) Para $D, E, F \in \mathbb{K}$ com E e F não nulos, considere o seguinte subespaço de $\mathbb{K}^{2\times 2}$:

$$\mathcal{U} = \left\{ \begin{pmatrix} Ex & Dx \\ Dy & Fy \end{pmatrix} \,:\, x, y \in \mathbb{K} \right\}.$$

i. Determine uma base de $\mathcal{U}$.
ii. Determine um subespaço $\mathcal{W}$ de $\mathbb{K}^{2\times 2}$ tal que $U \bigoplus \mathcal{W} = \mathbb{K}^{2\times 2}$.

RESOLUÇÃO

(a) É fácil verificar que $\sum_{j=1}^{n} \mathcal{U}_j$ é subespaço. De fato:

- $0_{\mathcal{V}} \in \sum_{j=1}^{m} \mathcal{U}_j$;[72]
- $\lambda \in \mathbb{K}$ e $u, v \in \sum_{j=1}^{m} \mathcal{U}_j \Longrightarrow \lambda u, u + v \in \sum_{j=1}^{m} \mathcal{U}_j$.[73]

Agora, claramente, $\mathcal{U}_i \subset \sum_{j=1}^{m} \mathcal{U}_j$, para cada $i \in \{1, \dots, m\}$.[74] Por outro lado, seja $\mathcal{U}$ um subespaço de $\mathcal{V}$ que contenha $\mathcal{U}_j$, $j = 1, \dots, m$. Portanto,

$$\begin{aligned} u_j \in \mathcal{U}_j, j = 1, \dots, m &\Longrightarrow u_j \in \mathcal{U}, j = 1, \dots, m \\ &\Longrightarrow u_1 + \cdots + u_m \in \mathcal{U} \text{ (pois } \mathcal{U} \text{ é subespaço)} \\ &\Longrightarrow \sum_{j=1}^{m} \mathcal{U}_j \subset \mathcal{U}. \end{aligned}$$

(b) Claramente, $\mathbb{K}^3 = \mathcal{U}_1 + \mathcal{U}_2 + \mathcal{U}_3$, pois cada vetor $(x, y, z) \in \mathbb{K}^3$ pode ser escrito da forma

$$(x, y, z) = (x, y, 0) + (0, 0, z) + (0, 0, 0).$$

Contudo, a soma não é direta, pois a forma supracitada não é única para, por exemplo, o vetor nulo de $\mathbb{K}^3$. De fato,

$$\begin{aligned} (0, 0, 0) &= (0, 0, 0) + (0, 0, 0) + (0, 0, 0) \\ &= (0, 1, 0) + (0, 0, 1) + (0, -1, -1). \end{aligned}$$

(c) Suponha, primeiramente, que a soma seja direta. Portanto, pela definição de somas diretas dada no item (b) supracitado, caso

$$\begin{aligned} u_1 + \cdots + u_m &= 0_{\mathcal{V}} \\ &= 0_{\mathcal{V}} + \cdots + 0_{\mathcal{V}}, \end{aligned}$$

onde estamos considerando m parcelas nulas na última igualdade (de cima para baixo),

$$u_j = 0_{\mathcal{V}}, j = 1, \dots, m.$$

Para a recíproca, considere a seguinte hipótese:

$$0_{\mathcal{V}} = u_1 + \cdots + u_m \Longrightarrow u_j = 0_{\mathcal{V}}, j = 1, \dots, m.$$

[72] Verifique!
[73] Idem!
[74] De fato, basta considerarmos somas $u_1 + \cdots + u_m$ onde todos os u's sejam nulos, exceto um deles.

Verificaremos que a soma é direta pela definição dada no item (b) supracitado. Assim, caso

$$u_1 + \cdots + u_m = u'_1 + \cdots + u'_m,$$

$$0_{\mathcal{V}} = u_1 - u'_1 + \ldots + u_m - u'_m.$$

Portanto, pela hipótese supracitada, $u_j - u'_j = 0_{\mathcal{V}}$, $j = 1, \ldots, m$, ou seja, $u_1 = u'_1, \ldots, u_m = u'_m$.

(d) Suponha, primeiramente, que $\mathcal{U} + \mathcal{W}$ é uma soma direta. Como $0_{\mathcal{V}} \in \mathcal{U} \cap \mathcal{W}$, $\{0_{\mathcal{V}}\} \subset \mathcal{U} \cap \mathcal{W}$. Agora, para demonstrarmos que $\mathcal{U} \cap \mathcal{W} \subset \{0_{\mathcal{V}}\}$, seja $v \in \mathcal{U} \cap \mathcal{W}$. Então,

$$0_{\mathcal{V}} = v + (-v),$$

com $v \in \mathcal{U}$ e $-v \in \mathcal{W}$. Portanto, como essa representação de $0_{\mathcal{V}}$ é única,[75] $v = 0_{\mathcal{V}}$. Logo, $v \in \{0_{\mathcal{V}}\}$.
Para a recíproca, suponha que $\mathcal{U} \cap \mathcal{W} = \{0_{\mathcal{V}}\}$. Assim, para que a soma $\mathcal{U} + \mathcal{W}$ seja direta, pelo item (c) supracitado, devemos demonstrar que o único modo de escrevermos

$$0_{\mathcal{V}} = u + w \text{ com } u \in \mathcal{U} \text{ e } w \in \mathcal{W} \tag{6.17}$$

é aquele em que $u = w = 0_{\mathcal{V}}$. De fato, caso (6.17) ocorra, $u = -w \in \mathcal{W}$ e, então, $u \in \mathcal{U} \cap \mathcal{W}$. Logo, $u = 0_{\mathcal{V}}$ e, portanto, $w = 0_{\mathcal{V}}$.

(e) Primeiramente, vamos demonstrar que os $m + n$ vetores dados geram $\mathcal{V}$. Seja, então, $v \in \mathcal{V}$. Como $\mathcal{V} = \mathcal{U} + \mathcal{W}$, existem $u \in \mathcal{U}$ e $w \in \mathcal{W}$ tais que

$$v = u + w. \tag{6.18}$$

Por outro lado, considerando as bases de $\mathcal{U}$ e $\mathcal{W}$ dadas, existem escalares $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n$ em $\mathbb{K}$ tais que

$$u = \alpha_1 u_1 + \cdots + \alpha_m u_m \quad \text{e} \quad w = \beta_1 w_1 + \cdots + \beta_n w_n. \tag{6.19}$$

Substituindo (6.19) em (6.18), temos que v é uma CL dos $m + n$ vetores dados.
Vamos demonstrar, agora, que estes $m + n$ vetores são LI. Assim, seja

$$\alpha_1 u_1 + \cdots \alpha_m u_m + \beta_1 w_1 + \cdots + \beta_n w_n = 0_{\mathcal{V}}, \tag{6.20}$$

onde os $m + n$ escalares, como antes, estão em $\mathbb{K}$. Denote u e w como em (6.19). Portanto, $0_{\mathcal{V}} = u + w$ e, como a soma é direta, pelo item (c) supracitado,

$$u = w = 0_{\mathcal{V}}. \tag{6.21}$$

Então, substituindo (6.21) em (6.19) e observando que tanto os m vetores de $\mathcal{U}$ dados quanto os n vetores de $\mathcal{W}$ dados são LI, todos os $m + n$ escalares da CL nula (6.20) devem ser nulos.

(f)
i. Note que toda matriz pertencente a $\mathcal{U}$ é uma CL da forma

$$x \begin{pmatrix} E & D \\ 0 & 0 \end{pmatrix} + y \begin{pmatrix} 0 & 0 \\ D & F \end{pmatrix}$$

com $x, y \in \mathbb{K}$. Assim, como as duas matrizes dessa CL são LI,[76] uma base de $\mathcal{U}$ é composta por elas.
ii. Considere que $\mathcal{W}$ seja, por exemplo, gerado por

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ e } \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Essas matrizes são LI.[77] Considere, então, o conjunto $\mathcal{B}$ formado por elas e pelas duas matrizes da base obtida no item i. supracitado. Então, por uma lado, a identidade

$$\begin{pmatrix} Ex & Dx \\ Dy & Fy \end{pmatrix} = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}$$

[75] Cf. o item (c) supracitado!
[76] De fato, nenhuma delas é um múltiplo escalar da outra.
[77] Idem.

acarreta $x = y = a = b = 0$. Logo, $\mathcal{U} \cap W = \{0\}$ com 0 representando a matriz nula de ordem 2. Por outro, $\mathcal{B}$ é uma base de $\mathcal{U} \bigoplus \mathcal{W}$. De fato, considere o determinante da matriz de ordem 4 cujas colunas sejam os vetores-coluna das coordenadas das matrizes de $\mathcal{B}$ na base canônica de $\mathbb{K}^{2\times 2}$:

$$\begin{aligned}
\det \begin{pmatrix} E & 0 & 0 & 0 \\ D & 0 & 1 & 0 \\ 0 & D & 0 & 1 \\ 0 & F & 0 & 0 \end{pmatrix} &= -\det \begin{pmatrix} E & 0 & 0 & 0 \\ 0 & F & 0 & 0 \\ 0 & D & 0 & 1 \\ D & 0 & 1 & 0 \end{pmatrix} \\
&= \det \begin{pmatrix} E & 0 & 0 & 0 \\ 0 & F & 0 & 0 \\ D & 0 & 1 & 0 \\ 0 & D & 0 & 1 \end{pmatrix} \\
&= EF \\
&\neq 0.
\end{aligned}$$

Portanto, $\mathcal{U} \bigoplus \mathcal{W} = \mathbb{K}^{2\times 2}$ pois $\dim(\mathcal{U} \bigoplus \mathcal{W}) = 4 = \dim \mathbb{K}^{2\times 2}$.

6.5 Informação adicional: dimensão da soma direta

Vamos generalizar o item (e) do último exercício da seção 6.4. Assim, caso $\mathcal{U}_1, \ldots, \mathcal{U}_m$ sejam subespaços finitamente gerados de $\mathcal{V}$ e a soma $\mathcal{U}_1 + \cdots + \mathcal{U}_m$ seja direta,

> *$\mathcal{U}_1 \oplus \cdots \oplus \mathcal{U}_m$ é finitamente gerado e*
>
> $$\dim \oplus_{i=1}^m \mathcal{U}_i = \sum_{i=1}^m \dim \mathcal{U}_i.$$

Esse resultado é uma consequência direta desse outro:

> *Se*
> $$\mathcal{B}_i = \left\{u_{i_1}, \ldots, u_{i_{n_i}}\right\}$$
> *é uma base de $\mathcal{U}_i$, $i = 1, \ldots, m$, então*
> $$\mathcal{B} = \cup_{i=1}^m \mathcal{B}_i$$
> *é uma base de $\mathcal{U} = \oplus_{i=1}^m \mathcal{U}_i$.*

De fato, seja $u \in \mathcal{U}$. Então, $u = \sum_{i=1}^m u_i$, com $u_i \in \mathcal{U}_i$, $i = 1, \ldots, m$. Assim, como u_i é uma CL dos vetores de $\mathcal{B}_i$, $i = 1, \ldots, m$, é fácil ver que os vetores de $\mathcal{B}$ geram $\mathcal{U}$. Portanto, resta verificarmos que os vetores de $\mathcal{B}$ são LI. Para isso, considere a seguinte CL nula:

$$c_{1_1} u_{1_1} + \cdots + c_{1_{n_1}} u_{1_{n_1}} + \cdots + c_{m_1} u_{m_1} + \cdots + c_{m_{n_m}} u_{m_{n_m}} = 0,$$

com $c_{i_1} u_{i_1} + \cdots + c_{i_{n_i}} u_{i_{n_i}} \in \mathcal{U}_i$ e $c_{i_j} \in \mathbb{K}$, $i = 1, \ldots, m$ e $j = 1, \ldots, n_i$. Então, $c_{i_1} u_{i_1} + \cdots + c_{i_{n_i}} u_{i_{n_i}} = 0$, $i = 1, \ldots, m$, pelo item (c) do último exercício da seção 6.4. Portanto, como $\mathcal{B}_i$ é uma base de $\mathcal{U}_i$, $i = 1, \ldots, m$, cada escalar c_{i_j} da CL supracitada é nulo .

Capítulo 7

$\mathcal{L}(V)$

Nesse capítulo, excetuando-se a seção 7.3, apenas operadores lineares sobre $\mathcal{V}$ serão considerados e, caso L seja qualquer um desses operadores, somente matrizes $\mathcal{M}(L, \mathcal{B}) \in \mathbb{K}^{n \times n}$ serão consideradas.

Dando prosseguimento aos resultados da seção 6.3, temos:

(R18) Isomorfismo sobre finitamente gerados

Se $\mathcal{V}$ é finitamente gerado, as seguintes condições são equivalentes:

1. *L é isomorfismo;*
2. *L é injetiva;*
3. *L é sobrejetiva.*

DEMONSTRAÇÃO

$\boxed{1 \Longrightarrow 2}$ Segue de (R13).[1]

$\boxed{2 \Longrightarrow 3}$ Caso L seja injetiva, $\mathrm{Nu}(L) = \{0\}$, por (R12).[2] Assim, $\dim \mathrm{Im}(L) = \dim \mathcal{V}$, por (R14).[3] Então, pela observação que segue (R8),[4] $\mathrm{Im}(L) = \mathcal{V}$, ou seja, L é sobrejetiva.

$\boxed{3 \Longrightarrow 1}$ Caso L seja sobrejetiva, isto é, $\mathrm{Im}(L) = \mathcal{V}$, temos $\dim \mathrm{Nu}(L) = 0$, por (R14). Assim, L é injetiva, por (R12). Portanto, por ser injetiva e sobrejetiva, L é isomorfismo, por (R13).

OBSERVAÇÃO

Que fique claro: *para operadores, qualquer uma das condições de* (R18) *acarreta as outras duas*.

7.1 Subespaços invariantes

Seja $\mathcal{U}$ subespaço de $\mathcal{V}$. Define-se a *restrição de L em $\mathcal{U}$* como a transformação linear $L|_{\mathcal{U}}$ dada por

$$\mathcal{U} \ni u \longmapsto L|_{\mathcal{U}}(u) := L(u) \in \mathcal{V}.$$

[1] Cf. p. 157.
[2] Cf. p. 156.
[3] Cf. p. 157.
[4] Cf. p. 152.

Em outras palavras, essa restrição funciona como a L, mas com o domínio restrito ao subespaço $\mathcal{U}$.

Exemplo

A restrição da função identidade

$$\mathcal{V} \ni v \longmapsto I(v) = v \tag{7.1}$$

em $\mathcal{U}$ é a *inclusão*

$$\mathcal{U} \ni u \longmapsto I|_{\mathcal{U}}(u) = u.$$

Observação/Definição

Caso a imagem da restrição $L|_{\mathcal{U}}$ esteja contida em $\mathcal{U}$, isto é,

$$u \in \mathcal{U} \Longrightarrow L(u) \in \mathcal{U},$$

essa restrição é um operador linear em $\mathcal{U}$, ou seja,

$$L|_{\mathcal{U}} \in \mathcal{L}(\mathcal{U}).$$

Nesse caso, dizemos que $\mathcal{U}$ é *invariante por L*.

Exemplos

- Confira a inclusão supracitada.

- Para $S, T \in \mathcal{L}(\mathcal{V})$ tais que $ST = TS$, $\mathrm{Nu}(S)$ e $\mathrm{Im}(S)$ são invariantes por T.

  De fato, basta observarmos que:

$$\begin{aligned} v \in \mathrm{Nu}(S) &\Longrightarrow S(v) = 0 \\ &\Longrightarrow T(S(v)) = T(0) \\ &\Longrightarrow TS(v) = 0 \\ &\Longrightarrow ST(v) = 0 \\ &\Longrightarrow S(T(v)) = 0 \\ &\Longrightarrow T(v) \in \mathrm{Nu}(S); \end{aligned}$$

$$\begin{aligned} v \in \mathrm{Im}(S) &\Longrightarrow \exists w \in V \text{ tal que } v = S(w) \\ &\Longrightarrow T(v) = T(S(w)) \\ &\Longrightarrow T(v) = TS(w) \\ &\Longrightarrow T(v) = ST(w) \\ &\Longrightarrow T(v) = S(T(w)) \\ &\Longrightarrow T(v) = S(u) \text{ com } u = T(w) \\ &\Longrightarrow \exists u \in V \text{ tal que } T(v) = S(u) \\ &\Longrightarrow T(v) \in \mathrm{Im}(S). \end{aligned}$$

- Se $T \in \mathcal{L}(\mathcal{V})$ e $\mathcal{U}_i$ é subespaço de $\mathcal{V}$ invariante por T, $i = 1, \ldots, m$, então $\sum_{i=1}^{m} \mathcal{U}_i$ é invariante por T.

De fato,

$$\begin{aligned} v \in \sum_{i=1}^{m} \mathcal{U}_i &\Longrightarrow v = \sum_{i=1}^{m} u_i \text{ com } u_i \in \mathcal{U}_i,\, i = 1, \ldots, m \\ &\Longrightarrow T(v) = T\left(\sum_{i=1}^{m} u_i\right) \\ &\Longrightarrow T(v) = \sum_{i=1}^{m} T(u_i) \text{ com } T(u_i) \in \mathcal{U}_i,\, i = 1, \ldots, m, \text{ pois cada } \mathcal{U}_i \text{ é invariante por } T \\ &\Longrightarrow T(v) \in \sum_{i=1}^{m} \mathcal{U}_i. \end{aligned}$$

CONTRA EXEMPLO

Se $\mathcal{V} = \mathbb{R}^2$, $L = R_{\pi/2}$ é a rotação de $\pi/2$ em torno da origem (como no exercício 3 da seção 4.1) e $\mathcal{U}$ é uma reta que passa pela origem, então $\mathcal{U}$ não é invariante por L.

7.2 Autovalores e autovetores

Como no capítulo 4, dizemos que $\lambda \in \mathbb{K}$ é *autovalor de L*, caso exista algum vetor não nulo $v \in \mathcal{V}$, dito *autovetor de L associado à λ*, tal que

$$L(v) = \lambda v.$$

EXEMPLOS

- Confira os exemplos da seção 4.2.
- Caso $v \in \mathcal{V}$ seja não nulo e $\mathcal{U} = [v]$ seja invariante por L, qualquer vetor não nulo de $\mathcal{U}$ é um autovetor de L.

  De fato, por um lado, caso $u \in \mathcal{U}$ seja não nulo, existe $\alpha \in \mathbb{K}$ não nulo tal que

  $$u = \alpha v.$$

  Por outro, caso $\mathcal{U}$ seja invariante por L, $L(u) \in \mathcal{U}$. Portanto, existe $\beta \in \mathbb{K}$ tal que

  $$L(u) = \beta v.$$

  Assim, se $\lambda = \beta\alpha^{-1}$, então

  $$L(u) = \lambda u.$$

7.2.1 Caso $\mathrm{Nu}(L - \lambda I)$ não seja o subespaço nulo de $\mathcal{V}$, seus vetores (não nulos) são os autovetores de L associados ao autovalor λ

De fato, para I dado em (7.1),

$$\begin{aligned} v \text{ é autovetor de } L \text{ associado à } \lambda &\Longleftrightarrow Lv = \lambda v \\ &\Longleftrightarrow Lv = \lambda I v \\ &\Longleftrightarrow Lv - \lambda I v = 0_{\mathcal{V}} \\ &\Longleftrightarrow (L - \lambda I)v = 0_{\mathcal{V}}. \end{aligned}$$

EXEMPLOS

Confira os exemplos da seção 4.2.[5]

[5] Note que, na seção supracitada, denotamos $\mathcal{S}_\lambda = \mathrm{Nu}(A - \lambda I)$. Em geral, $\mathrm{Nu}(L - \lambda I)$ é dito *autoespaço de L associado à λ*.

(R19) Autovalores via não isomorfismos

Se $\mathcal{V}$ é finitamente gerado, as condições seguintes são equivalentes:

1. *λ é autovalor de L;*
2. $\mathrm{Nu}(L - \lambda I) \neq \{0_{\mathcal{V}}\}$;
3. *$L - \lambda I$ não é injetiva;*
4. *$L - \lambda I$ não é sobrejetiva;*
5. *$L - \lambda I$ não é isomorfismo.*

DEMONSTRAÇÃO

A equivalência $1 \iff 2$ segue do enunciado da subseção 7.2.1. A equivalência $2 \iff 3$ segue de (R12).[6] As equivalências $3 \iff 4 \iff 5$ seguem de (R18).[7]

(R20) Autovetores de L associados à autovalores distintos entre si

Caso v_i e λ_i, $i = 1, \ldots, n$, sejam, respectivamente, os autovetores e autovalores supracitados, esses autovetores são LI.

EXEMPLO

Confira o exercício 14 da seção 4.3.

DEMONSTRAÇÃO DE (R20)

Suponha que os n autovetores supracitados sejam LD. Portanto, por (R1),[8] podemos considerar o menor índice j tal que

$$v_j \in \left[v_1, \ldots, v_{j-1}\right]. \tag{7.2}$$

Assim, existem escalares $\alpha_1, \ldots, \alpha_{j-1}$ tais que

$$v_j = \alpha_1 v_1 + \cdots + \alpha_{j-1} v_{j-1}. \tag{7.3}$$

Então, ao aplicarmos L à ambos os membros da equação (7.3), obtemos

$$\lambda_j v_j = \alpha_1 \lambda_1 v_1 + \cdots + \alpha_{j-1} \lambda_{j-1} v_{j-1}. \tag{7.4}$$

Por outro lado, ao multiplicarmos ambos os membros de (7.3) por λ_j, obtemos

$$\lambda_j v_j = \alpha_1 \lambda_j v_1 + \cdots + \alpha_{j-1} \lambda_j v_{j-1}. \tag{7.5}$$

Agora, ao subtrairmos (membro a membro) as equações (7.4) e (7.5), temos

$$\alpha_1 \left(\lambda_j - \lambda_1\right) v_1 + \cdots + \alpha_{j-1} \left(\lambda_j - \lambda_{j-1}\right) v_{j-1} = 0_{\mathcal{V}}.$$

Logo, como os autovalores supracitados são distintos e j é o menor índice para o qual (7.2) é válida, temos

$$\alpha_1 = \cdots = \alpha_{j-1} = 0.$$

Portanto, $v_j = 0_{\mathcal{V}}$, por (7.3). Isto é um absurdo, pois v_j é autovetor.

[6] Cf. p. 156.
[7] Cf. p. 177.
[8] Cf. p. 148.

(R21) Número máximo de autovalores distintos

Caso $\mathcal{V}$ seja finitamente gerado, $L \in \mathcal{L}(\mathcal{V})$ tem, no máximo, $\dim \mathcal{V}$ *autovalores distintos.*

DEMONSTRAÇÃO

O resultado (R21) é uma consequência (quase imediata) de (R20) e (R2).[9]

OBSERVAÇÃO

Em geral, noutros livros de AL, a demonstração de (R21) é baseada em determinantes, com a seguinte conclusão: *os autovalores de L são (exatamente) as raízes do polinômio característico de L.*[10] Na demonstração que adotamos, utilizamos uma abordagem "livre" de determinantes!

7.2.2 Polinômios com operadores como variáveis

Sejam i e j inteiros não negativos e $T \in \mathcal{L}(\mathcal{V})$. Considere as seguintes definições:

- Caso $j = 0$ e I seja o operador identidade em $\mathcal{V}$,

$$T^j := I;$$

- Caso $j \neq 0$,

$$T^j := \underbrace{T \cdots T}_{j \text{ fatores}};^{11}$$

Assim, $T^0 = I$, $T^1 = T$, $T^2 = TT$, $T^3 = TTT$, etc.

- Caso

$$p(t) = \sum_{j=0}^{m} a_j t^j \in \mathcal{P}(\mathbb{K}),$$

$$p(T) := \sum_{j=0}^{m} a_j T^j.$$

EXEMPLO

Se $p(t) = 3 - 2t + t^2$, então $p(T) = 3I - 2T + T^2$.

EXERCÍCIOS

1. Demonstre que

$$T^i T^j = T^{i+j} \text{ e } \left(T^i\right)^j = T^{ij}.^{12}$$

[9] Cf. p. 149.

[10] Cf. a seção 4.2.

[11] Cf. exercício 11 da seção 6.4.

[12] Caso T seja invertível, essas duas igualdades também são válidas para potências negativas. De fato, para cada inteiro k, podemos definir

$$T^{-k} := \left(T^{-1}\right)^k.$$

2. Fixado T, demonstre a linearidade da função

$$\mathcal{P}(\mathbb{K}) \ni p(t) \longmapsto p(T) \in \mathcal{L}(\mathcal{V}). \tag{7.6}$$

3. Demonstre que (7.6) preserva produto de polinômios, ou seja, para $p(t), q(t) \in \mathcal{P}(\mathbb{K})$ arbitrários, se

$$(pq)(t) := p(t)q(t),$$

então

$$(pq)(T) = p(T)q(T).$$

OBSERVAÇÃO

Quaisquer dois polinômios em T comutam.[13]

EXEMPLOS

- Se $T^2 = I$ e -1 não é autovalor de T, então $T = I$.[14]

- 9 é autovalor de $T^2 \iff$ 3 é autovalor de T ou -3 é autovalor de T.[15]

[13]De fato,

$$\begin{aligned} p(T)q(T) &= (pq)(T) \\ &= (qp)(T) \\ &= q(T)p(T). \end{aligned}$$

[14]De fato, suponha que $T - I \neq 0$. Assim, existe um vetor não nulo $v \in V$ tal que

$$\begin{aligned} w &:= (T-I)(v) \\ &\neq 0. \end{aligned}$$

Portanto, como

$$\begin{aligned} (T+I)(w) &= (T+I)(T-I)(v) \\ &= (T^2 - I)(v) \\ &= (I - I)(v) \\ &= 0, \end{aligned}$$

$T(w) = -w$, isto é, -1 é autovalor de T, que é uma contradição da hipótese inicial!

[15]De fato, como

$$\begin{aligned} T^2 - 9I &= (T+3I)(T-3I) \\ &= (T-3I)(T+3I), \end{aligned}$$

$$\begin{aligned} 9 \text{ é autovalor de } T^2 &\iff T^2 - 9I \text{ não é injetiva} \\ &\iff (T-3I) \text{ não é injetiva ou } (T+3I) \text{ não é injetiva} \\ &\iff 3 \text{ é autovalor de } T \text{ ou } -3 \text{ é autovalor de } T, \end{aligned}$$

por (R19).

7.2.3 Se $\mathcal{V} \neq \{0_{\mathcal{V}}\}$ é espaço vetorial complexo finitamente gerado, então $T \in \mathcal{L}(\mathcal{V})$ tem autovalor

Embora esse resultado não seja válido para espaços vetoriais reais,[16] podemos demonstrá-lo utilizando o seguinte fato algébrico:

> *(*) Se $p(t) \in \mathcal{P}(\mathbb{C})$ é não constante, então esse polinômio pode ser fatorado de modo único, a menos da ordem de seus fatores, na forma*
>
> $$p(t) = c\,(t-\lambda_1)\cdots(t-\lambda_m)\,,$$
>
> *com $c, \lambda_1, \ldots, \lambda_m \in \mathbb{C}$.*

A demonstração de (*), que é um corolário do *teorema fundamental da álgebra*, pode ser encontrada, por exemplo, no livro do AXLER, mencionado no primeiro capítulo.

DEMONSTRAÇÃO DE 7.2.3

Sejam $v \in \mathcal{V} - \{0_{\mathcal{V}}\}$ e $n = \dim \mathcal{V}$. Logo, os $n+1$ vetores

$$v, T(v), \ldots, T^n(v)$$

são LD, por (R2).[17] Assim, existem escalares $a_0, a_1, \ldots, a_n \in \mathbb{C}$, não todos nulos, tais que

$$\sum_{j=0}^{n} a_j T^j(v) = 0_{\mathcal{V}}. \tag{7.7}$$

Caso m seja o máximo do conjunto

$$\left\{ j \in \{0, 1, \ldots, n\} \,:\, a_j \neq 0 \right\},$$

$m \neq 0$. De fato, $m = 0$ reduz (7.7) a

$$a_0 v = 0_{\mathcal{V}}, \text{ com } a_0 \neq 0 \text{ e } v \neq 0_{\mathcal{V}},$$

que não é uma igualdade válida. Portanto, por (*),

$$p(t) = \sum_{j=0}^{n} a_j t^j$$

pode ser fatorado como

$$p(t) = c\Pi_{i=1}^{m}\,(t-\lambda_i)\,.$$

Assim,

$$\begin{aligned} c\left(\Pi_{i=1}^{m}\,(T-\lambda_i I)\right)(v) &= p(T)(v) \\ &= \sum_{j=0}^{n} a_j T^j(v) \\ &= 0_{\mathcal{V}}, \end{aligned}$$

por (7.6) e (7.7). Então, como $c \neq 0$,

$$\Pi_{i=1}^{m}\,(T-\lambda_i I)\,(v) = 0_{\mathcal{V}}. \tag{7.8}$$

Portanto, algum operador $T - \lambda_i I$ é não injetivo pois, caso contrário, (7.8) não pode ser nulo, por (R19).[18] Assim, também por (R19), algum λ_i é autovalor de T.

---

[16] Cf. (5.2), p. 133.

[17] Cf. p. 149.

[18] Cf. p 180.

7.2.4 Matrizes triangulares e operadores

(R22) Matrizes triangulares e invariância de bases

Se $T \in \mathcal{L}(\mathcal{V})$ e $\mathcal{B} = \{v_1, \ldots, v_n\}$ é base de $\mathcal{V}$, então as condições seguintes são equivalentes:

1. *$\mathcal{M}(T, \mathcal{B})$ é triangular superior;*
2. *$T(v_j) \in [v_1, \ldots, v_j]$, $j = 1, \ldots, n$;*
3. *$[v_1, \ldots, v_j]$ é invariante por T, $j = 1, \ldots, n$.*

EXEMPLO

Considere $\mathcal{V} = \mathbb{K}^2$ e

$$\mathcal{M}(T, \mathcal{B}) = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}.$$

Note que,

$$\begin{array}{rcl} T(v_1) & = & v_1; \\ T(v_2) & = & 2v_1 + 3v_2. \end{array}$$

Então, a condição 2 supracitada é satisfeita. Logo, como

$$\begin{aligned} T(\alpha_1 v_1 + \alpha_2 v_2) &= \alpha_1 T(v_1) + \alpha_2 T(v_2) \\ &= (\alpha_1 + 2\alpha_2)\, v_1 + 3\alpha_2 v_2, \end{aligned}$$

com $\alpha_1, \alpha_2 \in \mathbb{K}$, a condição 3 supracitada também é satisfeita.

DEMONSTRAÇÃO DE (R22)

$1 \Longleftrightarrow 2$ Segue de

$$\begin{array}{l} T(v_1) = a_{11}v_1; \\ T(v_2) = a_{12}v_1 + a_{22}v_2; \\ T(v_3) = a_{13}v_1 + a_{23}v_2 + a_{33}v_3; \\ \cdots \\ T(v_n) = a_{1n}v_1 + a_{2n}v_2 + a_{3n}v_3 + \cdots + a_{nn}v_n. \end{array}$$

$3 \Longrightarrow 2$ Óbvia.

$2 \Longrightarrow 3$ Fixe $j \in \{1, \ldots, n\}$. Assim, pela condição 2,

$$\begin{array}{l} T(v_1) \in [v_1] \subset [v_1, \ldots, v_j]\,; \\ T(v_2) \in [v_1, v_2] \subset [v_1, \ldots, v_j]\,; \\ \cdots \\ T(v_j) \in [v_1, \ldots, v_j]\,. \end{array}$$

Então, se $v \in [v_1, \ldots, v_j]$,

$$T(v) \in [v_1, \ldots, v_j]\,.$$

(R23) Existência de matrizes triangulares

Caso $\mathcal{V}$ seja um espaço vetorial finitamente gerado sobre $\mathbb{C}$ e $T \in \mathcal{L}(\mathcal{V})$, $\mathcal{M}(T, \mathcal{B})$ é triangular superior para alguma base $\mathcal{B}$ de $\mathcal{V}$.

EXEMPLO

Confira o primeiro exemplo da subseção 7.2.4.

DEMONSTRAÇÃO DE (R23)

Usaremos indução sobre $\dim \mathcal{V}$. Assim, o primeiro passo da indução consiste em provarmos a validade de (R23) para $\dim \mathcal{V} = 1$. De fato, como $\mathcal{M}(T, \mathcal{B})$ é de ordem 1, independentemente da base $\mathcal{B}$ considerada, o resultado (R23) é (trivialmente) válido. Agora, para o segundo passo da indução, vamos supor que (R23) seja válido para qualquer espaço vetorial $\mathcal{U}$ (sobre $\mathbb{C}$) tal que $\dim \mathcal{U} < \dim \mathcal{V} \neq 1$. Logo, caso λ seja um autovalor de T,[19] considere

$$\mathcal{U} := \operatorname{Im}(T - \lambda I).$$

Note que:

- $T - \lambda I$ não é sobrejetiva, por (R19).[20] Assim, $\dim \mathcal{U} < \dim \mathcal{V}$;
- $\mathcal{U}$ é invariante por T. De fato, se $u \in \mathcal{U}$, então

$$T(u) = (T - \lambda I)(u) + \lambda u,$$

  com $(T - \lambda I)(u) \in \mathcal{U}$ (pela definição de $\mathcal{U}$) e $\lambda u \in \mathcal{U}$. Logo, $T|_{\mathcal{U}} \in \mathcal{L}(\mathcal{U})$.

Assim, $\mathcal{M}\left(T|_{\mathcal{U}}, \mathcal{B}'\right)$ é triangular superior para alguma base $\mathcal{B}' = \{u_1, \ldots, u_n\}$ de $\mathcal{U}$. Note que,

$$T(u_i) \in [u_1, \ldots, u_i]\,, i = 1, \ldots, n, \tag{7.9}$$

por (R22). Além disso, ao estendermos $\mathcal{B}'$ à uma base $\mathcal{B} = \mathcal{B}' \cup \{v_1, \ldots, v_m\}$ de $\mathcal{V}$,[21] temos

$$T(v_j) = (T - \lambda I)(v_j) + \lambda v_j,$$

com $(T - \lambda I)(v_j) \in \mathcal{U} = [u_1, \ldots, u_n]$,[22] $j = 1, \ldots, m$. Logo,

$$T(v_j) \in [u_1, \ldots, u_n, v_1, \ldots, v_j]\,, j = 1, \ldots, m. \tag{7.10}$$

Portanto, (R23) é válido para $\mathcal{V}$, por (7.9), (7.10) e (R22).

OBSERVAÇÕES

- Que fique claro:

  Operadores sobre espaços vetoriais complexos podem ser representados por matrizes triangulares superiores.

- (R23) é a versão do lema de Schur para operadores.[23] De fato, sem entrarmos em detalhes técnicos, podemos assumir a ortonormalidade da base $\mathcal{B}$ supracitada.

(R24) Matrizes triangulares invertíveis

Se $T \in \mathcal{L}(\mathcal{V})$, $\mathcal{B} = \{v_1, \ldots, v_n\}$ é uma base de $\mathcal{V}$, $\mathsf{T} = \mathcal{M}(T, \mathcal{B})$ é triangular superior e $\mathsf{T}_{ii} = \lambda_i$, $i = 1, \ldots, n$, então:

$$\boxed{T \text{ é invertível} \iff \lambda_i \neq 0,\, i = 1, \ldots, n.}$$

[19] Cf. 7.2.3, p. 183.
[20] Cf. p. 180.
[21] Cf. (R6), p. 151.
[22] Pela definição de $\mathcal{U}$.
[23] Cf. a subseção 5.3.1.

EXEMPLO

O operador do primeiro exemplo da subseção 7.2.4 é invertível.

DEMONSTRAÇÃO DE (R24)

Primeiramente, note que,

$$\mathsf{T} = \begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}, \tag{7.11}$$

onde:

- $*$ representa todas as entradas acima da diagonal principal;
- 0 representa todas as entradas abaixo da diagonal principal.[24]

$\boxed{\Longleftarrow}$ Suponha que $\lambda_j \neq 0$, $j = 1, \ldots, n$. Assim:

- Como $T(v_1) = \lambda_1 v_1$, por (7.11), e $\lambda_1 \neq 0$,

  $$T\left(\frac{1}{\lambda_1} v_1\right) = v_1.$$

  Portanto, $v_1 \in \mathrm{Im}(T)$.
- Como, para algum escalar α, $T(v_2) = \alpha v_1 + \lambda_2 v_2$, por (7.11), e $\lambda_2 \neq 0$,

  $$T\left(\frac{1}{\lambda_2} v_2\right) = \frac{\alpha}{\lambda_2} v_1 + v_2. \tag{7.12}$$

  Então, como (7.12) e $-\frac{\alpha}{\lambda_2} v_1$ pertencem à $\mathrm{Im}(T)$, $v_2 \in \mathrm{Im}(T)$.
- Como $\lambda_3 \neq 0$, para escalares β e γ adequados,

  $$T\left(\frac{1}{\lambda_3} v_3\right) = \frac{\beta}{\lambda_3} v_1 + \frac{\gamma}{\lambda_3} v_2 + v_3.$$[25]

  Portanto, $v_3 \in \mathrm{Im}(T)$.

Ao repetirmos esse raciocínio mais $n-3$ vezes, concluímos que $v_j \in \mathrm{Im}(T)$, $j = 1, \ldots, n$, e, como $\mathcal{B}$ é uma base de $\mathcal{V}$, $\mathrm{Im}(T) = \mathcal{V}$, isto é, T é sobrejetivo, ou seja, T é invertível, por (R18).[26]

$\boxed{\Longrightarrow}$ Suponha que T seja invertível. Assim, como T é triangular superior e $\mathsf{T}_{11} = \lambda_1$, $\lambda_1 \neq 0$. De fato, seja $\lambda_1 = 0$. Logo, $T(v_1) = 0$, ou seja, $\mathrm{Nu}(T) \neq \{0\}$, isto é, T não é injetivo,[27] ou seja, T não é invertível,[28] contradizendo a suposição supracitada. Agora, considere que, para algum índice $j > 1$, $\lambda_j = 0$. Logo,

$$\mathrm{Im}\left(T|_{[v_1,\ldots,v_j]}\right) \subset [v_1, \ldots, v_{j-1}],$$

por (7.11). Assim, como

$$\dim [v_1, \ldots, v_j] = j \ \text{ e } \ \dim [v_1, \ldots, v_{j-1}] = j-1,$$

a restrição $T|_{[v_1,\ldots,v_j]}$ não é injetiva,[29] ou seja, existe um vetor não nulo $v \in [v_1, \ldots, v_j]$ tal que $T(v) = 0 = T(0)$, isto é, $T : \mathcal{V} \to \mathcal{V}$ não é injetivo, ou seja, T não é invertivel,[30] contradizendo a suposição supracitada. Portanto, $\lambda_i \neq 0$, $i = 1, \ldots, n$.

[24] Obviamente, essas entradas são nulas.
[25] Por (7.11).
[26] Cf. p. 177.
[27] Cf. (R12), p. 156.
[28] Cf. (R18), p. 177.
[29] Cf. o penúltimo exercício da seção 6.4.
[30] Cf. (R18).

(R25) Autovalores de uma matriz triangular

Se $T \in \mathcal{L}(\mathcal{V})$, $\mathcal{B} = \{v_1, \ldots, v_n\}$ é uma base de $\mathcal{V}$ e $\mathsf{T} = \mathcal{M}(T, \mathcal{B})$ é triangular superior, então:

$$\boxed{\lambda \text{ é autovalor de } T \iff \lambda \in \{\mathsf{T}_{11}, \ldots, \mathsf{T}_{nn}\}.}$$

EXEMPLOS

- 1 e 3 são os autovalores de $T \in \mathcal{L}(\mathbb{K}^2)$ do primeiro exemplo da subseção 7.2.4.
- Se $T \in \mathcal{L}(\mathbb{K}^3)$ é dada por $(x, y, z) \longmapsto T(x, y, z) = (x + 2y + 3z, 4y + 5z, 6z)$ e $\mathcal{B}$ é a base canônica de $\mathbb{K}^3$, então
$$\mathsf{T} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}.$$
Portanto, 1, 4 e 6 são os autovalores de T.

DEMONSTRAÇÃO DE (R25)

Como, ao considerarmos (7.11),[31]

$$\mathcal{M}(T - \lambda I) = \begin{bmatrix} \lambda_1 - \lambda & & * \\ & \ddots & \\ 0 & & \lambda_n - \lambda \end{bmatrix},$$

$$\begin{aligned} \lambda \text{ é autovalor de } T &\iff T - \lambda I \text{ não é invertível} \\ &\iff \lambda_i = \lambda, \text{ para algum } i \in \{1, \ldots, n\}, \end{aligned}$$

por (R19) e (R24).[32]

OBSERVAÇÃO

Noutros livros, (R25) é (geralmente) demonstrado via determinantes.

7.2.5 Diagonalização e operadores

Considere o seguinte corolário de (R25):

(R26) Autovalores de matrizes diagonais

Caso $T \in \mathcal{L}(\mathcal{V})$ seja diagonalizável, ou seja, caso exista alguma base $\mathcal{B} = \{v_1, \ldots, v_n\}$ de $\mathcal{V}$ tal que $\mathcal{M}(T, \mathcal{B}) = \mathsf{D}$,[33] e $d_{ii} = \lambda_i$, $i = 1, \ldots, n$,

$$\boxed{\lambda \text{ é autovalor de } T \iff \lambda \in \{\lambda_1, \ldots, \lambda_n\}}$$

[31] Cf. p. 186.
[32] Cf. pp. 180 e 185.
[33] Cf. (3.2), p. 56.

e $\mathcal{B}$ é constituída de autovetores de T.[34]

EXEMPLO

Para $T \in \mathcal{L}(\mathbb{K}^n)$ diagonalizável, confira as subseções 4.2.1–2 e os exercícios 12 – 15 da seção 5.2.

(R27) Autovalores distintos e diagonalização

$\lambda_1, \ldots, \lambda_n$ é uma lista de autovalores distintos de $T \in \mathcal{L}(\mathcal{V})$ e $\dim \mathcal{V} = n \Longrightarrow T$ é diagonalizável.

DEMONSTRAÇÃO

Seja v_j um autovetor de T associado à λ_j, $j = 1, \ldots, n$. Assim, $v_1, \ldots, v_n$ é uma lista de vetores LI, por (R20).[35] Portanto, $\mathcal{B} = \{v_1, \ldots, v_n\}$ é uma base de $\mathcal{V}$, por (R9).[36] Logo, $\mathcal{M}(T, \mathcal{B}) = \mathsf{D}$.

EXEMPLO

Para $T \in \mathcal{L}(\mathbb{K}^n)$ diagonalizável com n autovalores distintos, confira as subseções 4.2.1–2 e os exercícios 12 – 15 da seção 5.2.

7.2.6 Decomposição em somas diretas de autoespaços

OBSERVAÇÃO

Antes de prosseguir, confira o último exercício da seção 6.4.

(R28) Somas diretas e autovalores

Caso $\mathcal{V}$ seja finitamente gerado, $T \in \mathcal{L}(\mathcal{V})$ e $\lambda_1, \ldots, \lambda_m$ seja uma lista de autovalores distintos de T, a soma dos autoespaços

$$\sum_{i=1}^{m} \mathrm{Nu}(T - \lambda_i I)$$

é direta e

$$\sum_{i=1}^{m} \dim \mathrm{Nu}(T - \lambda_i I) \leq \dim \mathcal{V}.$$

[34]De fato, $T(v_i) = \lambda_i v_i$, $i = 1, \ldots, n$, pois

$$\mathsf{D} = \begin{bmatrix} \lambda_1 & & & & \\ & \lambda_2 & & \text{zeros} & \\ & & \ddots & & \\ & \text{zeros} & & \lambda_{n-1} & \\ & & & & \lambda_n \end{bmatrix}.$$

[35]Cf. p. 180.
[36]Cf. p. 152.

EXEMPLO

Confira o exercício 2 da seção 7.5.

DEMONSTRAÇÃO DE (R28)

Para demonstrarmos que a soma supracitada é direta, consideremos

$$\sum_{i=1}^{m} u_i = 0_{\mathcal{V}}, \tag{7.13}$$

com

$$u_i \in \mathrm{Nu}(T - \lambda_i I)\,, i = 1, \ldots, m. \tag{7.14}$$

Note que, caso alguns dos vetores listados em (7.14) sejam não nulos, eles são LI, por (R20).[37] Assim, um desses vetores não nulos pode ser escrito como CL dos outros, por (7.13). Como isso contradiz a independência linear dos vetores não nulos supracitados, $u_i = 0_{\mathcal{V}}$, $i = 1, \ldots, m$. Portanto, a soma dos autoespaços listados em (7.14) é direta, pelo item (c) do último exercício da seção 6.4.

Para concluirmos a demonstração de (R28), basta observaramos que, pela seção 6.5,

$$\begin{aligned} \sum_{i=1}^{n} \dim \mathrm{Nu}(T - \lambda_i I) &= \dim\left(\oplus_{i=1}^{m} \mathrm{Nu}(T - \lambda_i)\right) \\ &\leq \dim \mathcal{V}, \end{aligned}$$

pois a soma direta é subespaço de $\mathcal{V}$, pelo item (a) do último exercício da seção 6.4.

(R29) Diagonalização e decomposição em somas diretas

Caso $\mathcal{V}$ seja finitamente gerado, $T \in \mathcal{L}(\mathcal{V})$ e a lista $\lambda_1, \ldots, \lambda_m$ seja composta pelos autovalores distintos de T, as seguintes condições são equivalentes:

1. *T é diagonalizável;*
2. $\mathcal{V} = \oplus_{i=1}^{m} \mathrm{Nu}(T - \lambda_i I)$;
3. $\dim \mathcal{V} = \sum_{i=1}^{m} \dim \mathrm{Nu}(T - \lambda_i I)$.

EXEMPLO

Pelos cálculos da seção 4.2, caso

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

$$\mathbb{R}^2 \ni \mathbf{x} \longmapsto L(\mathbf{x}) = A\mathbf{x} \in \mathbb{R}^2$$

é diagonalizável e

$$\mathbb{R}^2 = \mathcal{S}_{-2} \oplus \mathcal{S}_3.$$

DEMONSTRAÇÃO DE (R29)

Seja

$$\mathcal{B} = \cup_{i=1}^{m} \mathcal{B}_i, \tag{7.15}$$

onde $\mathcal{B}_i$ é uma base de $\mathrm{Nu}(T - \lambda_i I)$, $i = 1, \ldots, m$.[38]

$\boxed{1 \Longrightarrow 3}$ Caso T seja diagonalizável, podemos considerar que a base $\mathcal{B}$ de (R26) é da forma (7.15). Portanto, a

[37] Cf. p. 180.

[38] A soma desses núcleos é direta, por (R28).

condição 3 supracitada é válida.

$\boxed{3 \Longrightarrow 2}$ Como a soma direta supracitada é um subespaço de $\mathcal{V}$,[39] a condição 2 supracitada é válida, pela seção 6.5.

$\boxed{2 \Longrightarrow 1}$ De fato, basta considerarmos (7.15).

7.2.7 Complementos e projeções ortogonais. Mínimos quadrados

OBSERVAÇÃO

Confira os exercícios 4 e 16 da seção 6.4, antes de prosseguir.

CONVENÇÃO

Até o final da subseção 7.2.7, considere que $\mathcal{V}$ é um espaço vetorial munido do produto interno $\langle\ ,\ \rangle$ e $\mathcal{B} = \{u_1, \ldots, u_r\}$ é uma base ortonormal de um subespaço de $\mathcal{V}$.[40]

(R30) Coordenadas na base ortonormal

Os escalares $c_1, \ldots, c_r$ da CL $u = \sum_{i=1}^{r} c_i u_i$ *são dados por* $c_i = \langle u, u_i \rangle$, $i = 1 \ldots, r$.

EXEMPLO

Confira os exercícios 6.(c), 7 e 8 da seção 2.7.

DEMONSTRAÇÃO DE (R30)

Para cada índice $j \in \{1, \ldots, r\}$, ao multiplicarmos a CL supracitada por u_j e utilizarmos a ortonormalidade de $\mathcal{B}$, obtemos

$$\begin{aligned}
\langle u, u_j \rangle &= \left\langle \sum_{i=1}^{r} c_i u_i, , u_j \right\rangle \\
&= \sum_{i=1}^{r} c_i \langle u_i, u_j \rangle \\
&= c_j.
\end{aligned}$$

(R31) Complemento ortogonal

Se $\emptyset \neq \mathcal{U} \subset \mathcal{V}$, então o complemento ortogonal de $\mathcal{U}$, isto é, o conjunto

$$\mathcal{U}^{\perp} := \{v \in \mathcal{V} \,:\, \langle v, u \rangle = 0 \text{ para cada } u \in \mathcal{U}\}\,, \tag{7.16}$$

satisfaz as seguintes condições:

1. $\mathcal{U}^{\perp}$ *é subespaço de* $\mathcal{V}$;
2. $\emptyset \neq \mathcal{W} \subset \mathcal{U} \Longrightarrow \mathcal{U}^{\perp} \subset \mathcal{W}^{\perp}$;
3. $\mathcal{U} = \mathcal{B} \Longrightarrow \mathcal{U}^{\perp} = [u_1, \ldots, u_r]^{\perp}$;
4. $v \in \mathcal{U} \cap \mathcal{U}^{\perp} \Longrightarrow v = 0_{\mathcal{V}}$.

[39] Cf. o item (a) do último exercício da seção 6.4.

[40] A existência dessa base é uma consequência do processo de Gram-Schmidt!

EXEMPLO

Confira o último exercício da seção 2.7.

DEMONSTRAÇÃO DO ITEM 1 DE (R31)

Seja $u \in \mathcal{U}$ arbitrário. $0_{\mathcal{V}} \in \mathcal{U}^{\perp}$, isto é, $\langle 0_{\mathcal{V}}, u\rangle = 0$, pois

$$\begin{aligned} \langle 0_{\mathcal{V}}, u\rangle &= \langle 0_{\mathcal{V}} + 0_{\mathcal{V}}, u\rangle \\ &= \langle 0_{\mathcal{V}}, u\rangle + \langle 0_{\mathcal{V}}, u\rangle. \end{aligned}$$

Agora, para quaisquer $v, w \in \mathcal{U}^{\perp}$ e $\lambda \in \mathbb{K}$, temos

$$\begin{aligned} \langle v + w, u\rangle &= \langle v, u\rangle + \langle w, u\rangle \\ &= 0 + 0 \\ &= 0 \end{aligned}$$

e

$$\begin{aligned} \langle \lambda v, u\rangle &= \lambda\langle v, u\rangle \\ &= \lambda \cdot 0 \\ &= 0, \end{aligned}$$

isto é, $v + w, \lambda v \in \mathcal{U}^{\perp}$.

DEMONSTRAÇÃO DO ITEM 2 DE (R31)

$$\begin{aligned} v \in \mathcal{U}^{\perp} &\Longrightarrow \langle v, u\rangle \text{ para todo } u \in \mathcal{U} \\ &\Longrightarrow \langle v, u\rangle \text{ para todo } u \in \mathcal{W}, \text{ pois } \mathcal{W} \subset \mathcal{U} \\ &\Longrightarrow v \in \mathcal{W}^{\perp}. \end{aligned}$$

DEMONSTRAÇÃO DO ITEM 3 DE (R31)

Por um lado, pelo item 2 supracitado, como $\mathcal{U} \subset [u_1, \ldots, u_r]$,

$$[u_1, \ldots, u_r]^{\perp} \subset \mathcal{U}^{\perp}.$$

Por outro, considere $v \in \mathcal{U}^{\perp}$ e seja u uma CL dos r vetores de $\mathcal{U}$, ou seja, considere $c_1, \ldots, c_r \in \mathbb{K}$ tais que $u = c_1 u_1 + \ldots + c_r u_r$. Então,

$$\begin{aligned} \langle v, u\rangle &= \overline{c_1}\langle v, u_1\rangle + \cdots + \overline{c_r}\langle v, u_r\rangle \\ &= \overline{c_1} \cdot 0 + \cdots + \overline{c_r} \cdot 0 \\ &= 0. \end{aligned}$$

Assim, $v \in [u_1, \ldots, u_r]^{\perp}$. Logo,

$$\mathcal{U}^{\perp} \subset [u_1, \ldots, u_r]^{\perp}.$$ [41]

DEMONSTRAÇÃO DO ITEM 4 DE (R31)

$$\begin{aligned} v \in \mathcal{U} \cap \mathcal{U}^{\perp} &\Longrightarrow \langle v, v\rangle = 0 \\ &\Longrightarrow v = 0_{\mathcal{V}}. \end{aligned}$$

[41] Essa demonstração também é válida para qualquer subconjunto finito $\mathcal{U} \neq \mathcal{B}$ de $\mathcal{V}$.

(R32) Decomposição numa soma de subespaços ortogonais

Na convenção dada no início da subseção 7.2.7, caso $\mathcal{U}$ seja o subespaço gerado pelos vetores de $\mathcal{B}$,

$$\mathcal{V} = \mathcal{U} \oplus \mathcal{U}^{\perp}.$$

EXEMPLOS

Confira o exercício 14 de 2.7.

DEMONSTRAÇÃO DE (R32)

Na primeira metade da demonstração, provaremos que

$$\mathcal{V} = \mathcal{U} + \mathcal{U}^{\perp}. \tag{7.17}$$

Assim, para $v \in \mathcal{V}$,

$$u = \sum_{i=1}^{r} \langle v, u_i \rangle u_i \in \mathcal{U} \tag{7.18}$$

e, para que (7.17) seja válida, basta que

$$w = v - u \in \mathcal{U}^{\perp}.[42]$$

Portanto, basta demonstrarmos que

$$w \in \mathcal{B}^{\perp},$$

pelo item 3 de (R31). De fato, para cada $j \in \{1, \ldots, r\}$,

$$\begin{aligned}
\langle w, u_j \rangle &= \langle v, u_j \rangle - \langle u, u_j \rangle \\
&= \langle v, u_j \rangle - \left\langle \sum_{i=1}^{r} \langle v, u_i \rangle u_i, u_j \right\rangle \\
&= \langle v, u_j \rangle - \sum_{i=1}^{r} \langle v, u_i \rangle \langle u_i, , u_j \rangle \\
&= \langle v, u_j \rangle - \langle v, u_j \rangle \\
&= 0,
\end{aligned}$$

onde, na segunda igualdade, de cima para baixo, utilizamos (7.17).
Na segunda parte da demonstração, provaremos que

$$\mathcal{U} \cap \mathcal{U}^{\perp} = \{0_{\mathcal{V}}\}.$$

De fato, por um lado,

$$\mathcal{U} \cap \mathcal{U}^{\perp} \subset \{0_{\mathcal{V}}\},$$

pelo item 4 de (R31). Por outro, como $\mathcal{U}$ e $\mathcal{U}^{\perp}$ são subespaços de $\mathcal{V}$, $0_{\mathcal{V}} \in \mathcal{U} \cap \mathcal{U}^{\perp}$. Portanto,

$$\{0_{\mathcal{V}}\} \subset \mathcal{U} \cap \mathcal{U}^{\perp}.$$

[42]De fato, nesse caso,

$$v = u + w, \tag{7.19}$$

com $u \in \mathcal{U}$ e $w \in \mathcal{U}^{\perp}$.

(R33) Corolário de (R32)

$$\boxed{\mathcal{V} \text{ é finitamente gerado e } \mathcal{U} \text{ é subespaço de } \mathcal{V} \Longrightarrow \dim \mathcal{V} = \dim \mathcal{U} + \dim \mathcal{U}^{\perp}.}$$

De fato, (R33) é uma consequência de (R3),[43] do item 1 de (R31), de (R32) e da seção 6.5.

EXEMPLO

Ao considerarmos $\mathcal{V} = \mathbb{R}^4$, podemos obter, via Gram-Schmidt, bases ortornomais para os subespaços $\mathcal{U} = [(1,2,3,-4),(-5,4,3,2)]$ e $\mathcal{U}^{\perp}$. De fato, caso

$$\begin{aligned}
\mathbf{x}_1 &= (1,2,3,-4),\\
\mathbf{x}_2 &= (-5,4,3,2),\\
\mathbf{x}_1' &= \mathbf{x}_1,\\
\mathbf{x}_2' &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{x}_1'}{\|\mathbf{x}_1'\|^2}\mathbf{x}_1',\\
\mathbf{x}_1'' &= \mathbf{x}_1'/\|\mathbf{x}_1'\| \text{ e}\\
\mathbf{x}_2'' &= \mathbf{x}_2'/\|\mathbf{x}_2'\|,
\end{aligned}$$

$\{\mathbf{x}_1'', \mathbf{x}_2''\}$ é uma base ortonormal de $\mathcal{U}$. Por outro lado, como

$$\mathcal{U}^{\perp} = \{\mathbf{x}_1, \mathbf{x}_2\}^{\perp},$$

pelo item 3 de (R31), e $\dim \mathcal{U}^{\perp} = 2$, por (R33), caso $\mathbf{y}_1$ e $\mathbf{y}_2$ sejam LI e satisfaçam as equações

$$\mathbf{y}_i \cdot \mathbf{x}_j = 0,\, i,j \in \{1,2\},$$

$\{\mathbf{y}_1, \mathbf{y}_2\}$ é uma base de $\mathcal{U}^{\perp}$. Portanto, caso $\mathbf{y} = (a,b,c,d) \in \mathbb{R}^4$ satisfaça as equações

$$\mathbf{y} \cdot \mathbf{x}_j = 0,\, j = 1,2,$$

ou seja,

$$\begin{cases} a + 2b + 3c - 4d = 0,\\ -5a + 4b + 3c + 2d = 0, \end{cases}$$

isto é,

$$\begin{cases} a = 15c - 14d,\\ b = -9c + 9d, \end{cases}$$

$$\begin{aligned}
\mathbf{y} &= (15c - 14d, -9c + 9d, c, d)\\
&= c(15,-9,1,0) + d(-14,9,0,1),
\end{aligned}$$

para quaisquer números reais c e d. Assim, caso

$$\begin{aligned}
\mathbf{y}_1 &= (15,-9,1,0),\\
\mathbf{y}_2 &= (-14,9,0,1),\\
\mathbf{y}_1' &= \mathbf{y}_1,\\
\mathbf{y}_2' &= \mathbf{y}_2 - \frac{\mathbf{y}_2 \cdot \mathbf{y}_1'}{\|\mathbf{y}_1'\|^2}\mathbf{y}_1',\\
\mathbf{y}_1'' &= \mathbf{y}_1'/\|\mathbf{y}_1'\| \text{ e}\\
\mathbf{y}_2'' &= \mathbf{y}_2'/\|\mathbf{y}_2'\|,
\end{aligned}$$

[43] Cf. p. 150.

$\{\mathbf{y}_1'', \mathbf{y}_2''\}$ é uma base ortonormal de $\mathcal{U}^\perp$.

DEFINIÇÃO

Seja $\mathcal{U}$ o subespaço dado em (R32). Para $v \in \mathcal{V}$, seja u o vetor definido em (7.18).[44] Dizemos que

$$P_{\mathcal{U}}(v) := u$$

é a *projeção ortogonal de v sobre $\mathcal{U}$* e

$$P_{\mathcal{U}} : \mathcal{V} \longrightarrow \mathcal{V}$$

é a *projeção ortogonal de $\mathcal{V}$ sobre $\mathcal{U}$*.

Para uma representação geométrica de $P_{\mathcal{U}}(v)$, confira a figura 7.1.

EXEMPLO

Confira o item (b) do exercício 5 que antecede a subseção 4.1.1.

(R34) Propriedades das projeções ortogonais

Considere a definição supracitada. Assim:

1. $P_{\mathcal{U}}(v) = \sum_{i=1}^{r} \langle v, u_i \rangle u_i$;
2. $v - P_{\mathcal{U}}(v) \in \mathcal{U}^\perp$;
3. $P_{\mathcal{U}} \in \mathcal{L}(\mathcal{V})$;
4. *$P_{\mathcal{U}}(v)$ é o vetor de $\mathcal{U}$ mais próximo de v, ou seja, para todo $x \in \mathcal{U}$,*

$$\|v - P_{\mathcal{U}}(v)\| \leq \|v - x\|.$$

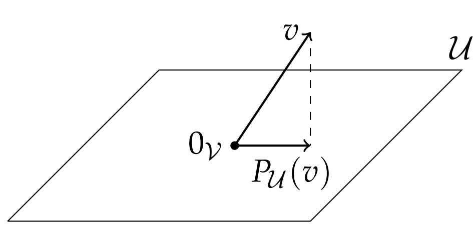
Para uma representação geométrica dos itens 2 e 4 de (R34), confira a figura 7.1.

EXEMPLO

Confira o item (c) do exercício 4.2 da seção 6.4.

[44] Cf. p. 192.

Figura 7.1: Projeção ortogonal de $v \in \mathcal{V}$ sobre $\mathcal{U}$



DEMONSTRAÇÃO DO ITEM 1 DE (R34)

Pela definição supracitada, u é dado por (7.18).

DEMONSTRAÇÃO DO ITEM 2 DE (R34)

$$\begin{aligned} v - P_{\mathcal{U}}(v) &= v - u \\ &= w \in \mathcal{U}^{\perp}, \end{aligned}$$

por (7.19).[45]

DEMONSTRAÇÃO DO ITEM 3 DE (R34)

$$\begin{aligned} v_1, v_2 \in \mathcal{V} &\Longrightarrow v_1 - P_{\mathcal{U}}(v_1)\,,\, v_2 - P_{\mathcal{U}}(v_2) \in \mathcal{U}^{\perp}, \text{ pelo item 2 de (R34)} \\ &\Longrightarrow (v_1 + v_2) - [P_{\mathcal{U}}(v_1) + P_{\mathcal{U}}(v_2)] \in \mathcal{U}^{\perp} \\ &\Longrightarrow v_1 + v_2 = [P_{\mathcal{U}}(v_1) + P_{\mathcal{U}}(v_2)] + w_{v_1+v_2}, \\ &\qquad \text{onde a parcela entre colchetes pertence à } \mathcal{U} \\ &\qquad \text{(pela definição supracitada) e } w_{v_1+v_2} \in \mathcal{U}^{\perp} \\ &\Longrightarrow P_{\mathcal{U}}(v_1 + v_2) = P_{\mathcal{U}}(v_1) + P_{\mathcal{U}}(v_2)\,; \end{aligned}$$

$$\begin{aligned} v \in \mathcal{V} \text{ e } \lambda \in \mathbb{K} &\Longrightarrow \lambda\,(v - P_{\mathcal{U}}(v)) \in \mathcal{U}^{\perp}, \text{ pelo item 2 de (R34)} \\ &\Longrightarrow \lambda v = [\lambda P_{\mathcal{U}}(v)] + w_{\lambda v}, \\ &\qquad \text{onde a parcela entre colchetes pertence à } \mathcal{U} \\ &\qquad \text{(pela definição supracitada) e } w_{\lambda v} \in \mathcal{U}^{\perp} \\ &\Longrightarrow P_{\mathcal{U}}(\lambda v) = \lambda P_{\mathcal{U}}(v). \end{aligned}$$

DEMONSTRAÇÃO DO ITEM 4 DE (R34)

TEOREMA DE PITÁGORAS

Se

$v_1, v_2 \in \mathcal{V}$, *com* $\langle v_1, v_2 \rangle = 0$,

então

$\|v_1 + v_2\|^2 = \|v_1\|^2 + \|v_2\|^2$.

De fato, basta observarmos que

$$\begin{aligned} \|v_1 + v_2\|^2 &= \langle v_1 + v_2, v_1 + v_2 \rangle \\ &= \langle v_1, v_1 + v_2 \rangle + \langle v_2, v_1 + v_2 \rangle \\ &= \overline{\langle v_1 + v_2, v_1 \rangle} + \overline{\langle v_1 + v_2, v_2 \rangle} \\ &= \overline{\langle v_1, v_1 \rangle} + \overline{\langle v_2, v_1 \rangle} + \overline{\langle v_1, v_2 \rangle} + \overline{\langle v_2, v_2 \rangle} \\ &= \overline{\|v_1\|^2} + \langle v_1, v_2 \rangle + \overline{\langle v_1, v_2 \rangle} + \overline{\|v_1\|^2}. \end{aligned}$$

Assim, como demonstramos o teorema de Pitágoras, podemos usá-lo na segunda igualdade dada a seguir, pois $v - P_{\mathcal{U}}(v) \in \mathcal{U}^{\perp}$, pelo item 2 de (R34), e $P_{\mathcal{U}}(v) - x \in \mathcal{U}$, por ser a diferença entre dois vetores do mesmo subespaço. Portanto,

$$\begin{aligned} \|v - x\|^2 &= \|(v - P_{\mathcal{U}}(v)) + (P_{\mathcal{U}}(v) - x)\|^2 \\ &= \|v - P_{\mathcal{U}}(v)\|^2 + \|P_{\mathcal{U}}(v) - x\|^2 \\ &\geq \|v - P_{\mathcal{U}}(v)\|^2 \end{aligned}$$

[45] Cf. p. 192.

e, assim,

$$\begin{aligned} \|v - x\| = \|v - P_{\mathcal{U}}(v)\| &\iff \|P_{\mathcal{U}}(v) - x\| = 0 \\ &\iff x = P_{\mathcal{U}}(v). \end{aligned}$$

OBSERVAÇÃO

Em muitos problemas de otimização, os itens 1 e 4 de (R34) podem ser utilizados, simultaneamente, para obtermos minimizadores globais. Em particular, em $\mathcal{V} = \mathbb{K}^n$, dotado do produto canônico, caso

$$\mathcal{B} = \{\mathbf{a}_1, \ldots, \mathbf{a}_r\}$$

seja uma base ortonormal do subespaço $\mathcal{U}$ de $\mathcal{V}$,

$$P_{\mathcal{U}}(\mathbf{x}) = \sum_{i=1}^{r} (\mathbf{x} \cdot \mathbf{a}_i)\, \mathbf{a}_i$$

é o vetor de $\mathcal{U}$ mais próximo de $\mathbf{x} \in \mathbb{K}^n$. A expressão *mínimos quadrados* vem da minimalidade de $\|\mathbf{x} - P_{\mathcal{U}}(\mathbf{x})\|^2$.

EXEMPLO

Reveja o item (c) do exercício 4.2 da seção 6.4.

EXERCÍCIOS

1. Para $\mathcal{V} = \mathbb{R}^4$ e $\mathcal{U} = [(1,1,0,0), (1,1,1,2)]$, obtenha $\mathbf{u} \in \mathcal{U}$ com $\|\mathbf{u} - (1,2,3,4)\|$ mínimo.

   RESOLUÇÃO

   Se

$$\begin{aligned} \mathbf{x}_1 &= (1,1,0,0), \\ \mathbf{x}_2 &= (1,1,1,2), \\ \mathbf{x}_1' &= \mathbf{x}_1, \\ \mathbf{x}_2' &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{x}_1'}{\|\mathbf{x}_1'\|^2}\mathbf{x}_1', \\ \mathbf{x}_1'' &= \mathbf{x}_1' / \|\mathbf{x}_1'\| \text{ e} \\ \mathbf{x}_2'' &= \mathbf{x}_2' / \|\mathbf{x}_2'\|, \end{aligned}$$

   então $\{\mathbf{x}_1'', \mathbf{x}_2''\}$ é uma base ortonormal de $\mathcal{U}$, por Gram-Schmidt. Portanto,

$$\begin{aligned} \mathbf{u} &= P_{\mathcal{U}}(1,2,3,4) \\ &= ((1,2,3,4) \cdot \mathbf{x}_1'')\, \mathbf{x}_1'' + ((1,2,3,4) \cdot \mathbf{x}_2'')\, \mathbf{x}_2''. \end{aligned}$$

2. Para $C, D, E, F \in \mathbb{R}$, considere o subespaço

$$\mathcal{U} = \left[\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1/2 \\ C - D & E - F \end{bmatrix} \right]$$

   de $\mathbb{R}^{2\times 2}$, munido do produto interno

$$\left\langle \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \right\rangle = a_{11}b_{11} + a_{12}b_{12} + a_{21}b_{21} + a_{22}b_{22}.$$

- Obtenha uma base de $\mathcal{U}^\perp$.
- Considere $C = 3$, $D = 4$, $E = 5$ e $F = 6$.[46]
    - Obtenha uma base ortonormal de $\mathcal{U}^\perp$.
    - Para $N = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, determine a matriz $M \in \mathcal{U}^\perp$ com $\|M - N\|$ mínimo.

RESOLUÇÃO

Sejam X_1 e X_2 as matrizes geradoras de $\mathcal{U}$, dadas no enunciado desse exercício, na ordem em que aparecem. Assim, caso Y_1 e Y_2 sejam duas matrizes LI de $\mathcal{U}^\perp$, $\{Y_1, Y_2\}$ é uma base de $\mathcal{U}^\perp$ e $\langle X_i, Y_j \rangle = 0$, $i, j \in \{1, 2\}$. Portanto, se

$$Y = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathcal{U}^\perp,$$

então, para cada $i \in \{1, 2\}$,

$$\langle Y, X_i \rangle = 0,$$

isto é,

$$a + b + c = 0 \;\text{ e }\; a + \frac{b}{2} + (C - D)c + (E - F)d = 0,$$

ou seja,

$$a = (2(D - C) + 1)c + 2(F - E)d \;\text{ e }\; b = 2(C - D - 1)c + 2(E - F)d,$$

isto é,

$$Y = c \begin{bmatrix} 2(D - C) + 1 & 2(C - D - 1) \\ 1 & 0 \end{bmatrix} + d \begin{bmatrix} 2(F - E) & 2(E - F) \\ 0 & 1 \end{bmatrix}. \tag{7.20}$$

Assim, a escolha mais simples para Y_1 e Y_2 é a seguinte:

- $c = 1$ e $d = 0 \Longrightarrow Y = Y_1$;
- $c = 0$ e $d = 1 \Longrightarrow Y = Y_2$. [47]

---

[46] A escolha desses valores numéricos é arbitrária. Inclusive, numa eventual resolução alternativa, o leitor pode utilizar outros valores reais para C, D, E e F.
Durante a pandemia de *covid-19*, o departamento de matemática, onde leciono, aplicou algumas provas *online*. Para otimizarem o trabalho e, simultaneamente, coibirem tentativas de burlar o sistema de avaliações, alguns dos professores preparavam questões individuais, que dependiam de números relacionados as matrículas dos alunos. Assim, embora as questões fossem alfanuméricas e iguais para todos os estudantes, no lugar de letras, como C, D, E e F, eles tinham de utilizar algarismos, de 0 a 9, correspondentes à determinadas posições, como em $1234CDEF$, de suas respectivas matrículas.

[47] Como Y_1 e Y_2 não são múltiplas uma da outra, elas são LI.

Agora, ao considerarmos $C = 3$, $D = 4$, $E = 5$ e $F = 6$, temos

$$\begin{aligned}
Y_1' &= Y_1 \\
&= \begin{bmatrix} 3 & -4 \\ 1 & 0 \end{bmatrix}, \\
Y_2' &= Y_2 - \frac{\langle Y_2, Y_1' \rangle}{\|Y_1'\|^2} Y_1' \\
&= \begin{bmatrix} 2 & -2 \\ 0 & 1 \end{bmatrix} - \frac{7}{13} \begin{bmatrix} 3 & -4 \\ 1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 5/13 & 2/13 \\ -7/13 & 1 \end{bmatrix}, \\
Y_1'' &= \frac{Y_1'}{\|Y_1'\|} \\
&= \frac{1}{\sqrt{26}} \begin{bmatrix} 3 & -4 \\ 1 & 0 \end{bmatrix} \quad \text{e} \\
Y_2'' &= \frac{Y_2'}{\|Y_2'\|} \\
&= \frac{13}{\sqrt{247}} \begin{bmatrix} 5/13 & 2/13 \\ -7/13 & 1 \end{bmatrix},
\end{aligned}$$

com $\{Y_1'', Y_2''\}$ ortonormal, por Gram-Schmidt. Portanto, como M é dado pela projeção de N sobre $\mathcal{U}^\perp$ via

$$M = \langle N, Y_1'' \rangle Y_1'' + \langle N, Y_2'' \rangle Y_2'',$$

com $\langle N, Y_1'' \rangle = -\frac{2}{\sqrt{26}}$ e $\langle N, Y_2'' \rangle = \frac{40}{\sqrt{247}}$,

$$M = -\frac{1}{13} \begin{bmatrix} 3 & -4 \\ 1 & 0 \end{bmatrix} + \frac{40}{19} \begin{bmatrix} 5/13 & 2/13 \\ -7/13 & 1 \end{bmatrix}. \tag{7.21}$$

$\boxed{\text{OBSERVAÇÃO}}$

Alternativamente, podemos obter a matriz (7.21) pelo *cálculo de várias variáveis*. De fato, para Y dada em (7.20), podemos calcular o(s) ponto(s) crítico(s) (c_0, d_0) da função $f(c, d) = \|Y - N\|^2$, via $\frac{\partial f}{\partial c} = 0$ e $\frac{\partial f}{\partial d} = 0$.[48] Em seguida, para garantirmos que o ponto (c_0, d_0) calculado seja um minimizador local de f,[49] podemos utilizar, nesse ponto, o *teste da derivada (parcial) de segunda ordem*. Finalmente, para obtermos M, basta substituirmos $c = c_0$ e $d = d_0$ em (7.20).

3. Em $\mathcal{P}_3(\mathbb{R})$, munido do produto interno

$$(p_1, p_2) \mapsto \int_0^1 p_1(x) p_2(x)\, dx,$$

determine o polinômio $p(x) = c + bx + ax^2$ tal que $p(0) = 0$ e a integral

$$\int_0^1 |(C + D + 1) + (E + F + 1)x - p(x)|^2 dx$$

seja mínima, atribuindo valores numéricos aos escalares $C, D, E, F \in \mathbb{R}$ tais que $C + D$ e $E + F$ sejam diferentes de -1.[50]

[48] Para o exercício 2 supracitado, existe (apenas) um ponto crítico.
[49] Esse ponto é, de fato, o mínimo global!
[50] Confira a primeira nota de rodapé do exercício 2 supracitado.

RESOLUÇÃO

Como $p(0) = 0$, $c = 0$. Assim, claramente, $p \in \mathcal{U} = [q, r]$, onde $q(x) = x$ e $r(x) = x^2$.[51] Portanto, se

$$u(x) = (C + D + 1) + (E + F + 1)x \text{ e } p(x) = p_{\mathcal{U}}(u(x)),$$

então

$$\|u(x) - p(x)\|^2 = \int_0^1 |u(x) - p(x)|^2 dx$$

é mínima e

$$p = \langle u, q_2 \rangle q_2 + \langle u, r_2 \rangle r_2, \tag{7.22}$$

onde $\{q_2, r_2\}$ é uma base ortonormal de $\mathcal{U}$, que pode ser obtida por Gram-Schmidt. Assim, ao considerarmos, por exemplo, $C = 3$, $D = 4$, $E = 5$ e $F = 6$, ou seja, $u(x) = 8 + 12x$, temos

$$\begin{aligned}
q_1(x) &= q(x) \\
&= x, \\
r_1(x) &= r(x) - \frac{\int_0^1 r(x) q_1(x)\, dx}{\int_0^1 q_1(x)^2\, dx} q_1(x) \\
&= x^2 - \frac{\int_0^1 x^3\, dx}{\int_0^1 x^2\, dx} x \\
&= x^2 - \frac{1/4}{1/3} x \\
&= x^2 - \frac{3}{4} x.
\end{aligned}$$

Portanto, ao normalizarmos q_1 e r_1, temos

$$\begin{aligned}
q_2(x) &= \frac{1}{\sqrt{\int_0^1 q_1(x)^2\, dx}} q_1(x) \\
&= \sqrt{3} x, \\
r_2(x) &= \frac{1}{\sqrt{\int_0^1 r_1(x)^2\, dx}} r_1(x) \\
&= \frac{1}{\sqrt{\int_0^1 x^4 - \frac{3}{2}x^3 + \frac{9}{16}x^2\, dx}} \left(x^2 - \frac{3}{4} x \right) \\
&= \frac{1}{\sqrt{1/80}} \left(x^2 - \frac{3}{4} x \right) \\
&= \sqrt{5}(4x^2 - 3x).
\end{aligned}$$

Assim, por (7.22),

$$\begin{aligned}
p(x) &= \left(\int_0^1 (8 + 12x)(\sqrt{3}x)\, dx \right) \sqrt{3}x + \left(\int_0^1 (8 + 12x)(\sqrt{5}(4x^2 - 3x))\, dx \right) (\sqrt{5}(4x^2 - 3x)) \\
&= 3 \left(8 \int_0^1 x\, dx + 12 \int_0^1 x^2\, dx \right) x \\
&\quad + 5 \left(32 \int_0^1 x^2\, dx - 24 \int_0^1 x\, dx + 48 \int_0^1 x^3\, dx - 36 \int_0^1 x^2\, dx \right) (4x^2 - 3x) \\
&= 3(4 + 4)x + 5((32/3) - 12 + 12 - 12)(4x^2 - 3x) \\
&= 24x - (20/3)(4x^2 - 3x) \\
&= 44x - (80/3)x^2.
\end{aligned}$$

[51] Verifique que q e r são LI.

4. Para $C, D, E \in \mathbb{R}$, considere $\mathcal{V} = \mathcal{C}[1, C+2]$,[52] munido do produto interno

$$\mathcal{V} \times \mathcal{V} \ni (g,h) \mapsto \langle g,h \rangle = \int_1^{C+2} g(x)h(x)\,dx \in \mathbb{R},$$

e

$$[1, C+1] \ni x \mapsto f(x) = (D+1)x^{\frac{1}{E+1}} \in \mathbb{R}.^{53}$$

Determine o polinômio $p \in \mathcal{U} = \mathcal{P}_2(\mathbb{R})$ mais próximo de f.[54,55]

RESOLUÇÃO

p é a projeção ortogonal de f sobre $\mathcal{U}$ e podemos calculá-lo pela expressão

$$p = \langle f, p_1 \rangle p_1 + \langle f, p_2 \rangle p_2, \tag{7.23}$$

onde $\{p_1, p_2\}$ é uma base ortonormal de $\mathcal{U}$, que pode ser obtida pela aplicação de Gram-Schmidt na base canônica de $\mathcal{U}$, ou seja, $\{1, x\}$. Assim, como $p_i = q_i / \|q_i\|$, $i = 1, 2$, com

$$\begin{aligned}
q_1(x) &= 1 \qquad \text{e} \\
q_2(x) &= x - \frac{\langle x, 1 \rangle}{\langle 1, 1 \rangle} 1 \\
&= x - \frac{\int_1^{C+2} x\,dx}{\int_1^{C+2} dx} \\
&= x - \frac{\frac{(C+2)^2 - 1}{2}}{C+1} \\
&= x - \frac{C+3}{2},
\end{aligned}$$

$$\begin{aligned}
p_1(x) &= \frac{1}{\sqrt{\langle 1, 1 \rangle}} \\
&= \frac{1}{\sqrt{C+1}} \qquad \text{e} \\
p_2(x) &= \frac{1}{\sqrt{\langle x - \frac{C+3}{2}, x - \frac{C+3}{2} \rangle}} \left(x - \frac{C+3}{2} \right) \\
&= \frac{1}{\sqrt{\int_1^{C+2} \left(x - \frac{C+3}{2} \right)^2 dx}} \left(x - \frac{C+3}{2} \right) \\
&= \frac{1}{\sqrt{\frac{u^3}{3}\Big|_{-\frac{C+1}{2}}^{\frac{C+1}{2}}}} \left(x - \frac{C+3}{2} \right) \qquad \left(\text{onde } u = x - \frac{C+3}{2} \right) \\
&= \frac{1}{\sqrt{\frac{(C+1)^3}{12}}} \left(x - \frac{C+3}{2} \right).
\end{aligned}$$

[52] $f \in \mathcal{V} \iff f : [1, C+2] \to \mathbb{R}$ é contínua.
[53] Note que, $f \in \mathcal{V}$.
[54] $\mathcal{U}$ é o espaço dos polinômios constantes ou lineares.
[55] Confira a primeira nota de rodapé do exercício 2 supracitado.

Portanto, ao substituirmos f, p_1 e p_2 na expressão (7.23), obtemos

$$p(x) = (D+1)\Bigg(\langle x^{1/(E+1)}, 1/\sqrt{C+1}\rangle \frac{1}{\sqrt{C+1}} + \langle x^{1/(E+1)}, (1/\sqrt{(C+1)^3/12})(x-(C+3)/2)\rangle \frac{1}{\sqrt{\frac{(C+1)^3}{12}}}\left(x - \frac{C+3}{2}\right)\Bigg),$$

isto é,

$$p(x) = (D+1)\Bigg(\frac{1}{C+1}\int_1^{C+2} x^{1/(E+1)}dx + \frac{12}{(C+1)^3}\int_1^{C+2} x^{1/(E+1)}(x-(C+3)/2)\,dx \left(x - \frac{C+3}{2}\right)\Bigg).$$

Assim, para $C=3$, $D=4$ e $E=5$, por exemplo, temos

$$\begin{aligned} p(x) &= 5\left(\frac{1}{4}\int_1^5 x^{1/6}dx + \frac{12}{4^3}\left(\int_1^5 x^{1/6}(x-3)\,dx\right)(x-3)\right) \\ &= 5\left(\frac{1}{4}\left[\frac{x^{7/6}}{7/6}\right]_1^5 + \frac{3}{16}\left[\frac{x^{13/6}}{13/6} - \frac{3x^{7/6}}{7/6}\right]_1^5 (x-3)\right) \\ &= 5\Bigg(\frac{3}{14}\left(5^{7/6}-1\right) + \frac{3}{16}\left(\frac{6}{13}\left(5^{13/6}-1\right) - \frac{18}{7}\left(5^{7/6}-1\right)\right)(x-3)\Bigg). \end{aligned}$$

5. Considere $D, E, F \in \mathbb{R} - \{-1\}$, $\mathbf{x}_0 = (D+1, E+1, F+1)$, $\mathcal{U} = [\mathbf{x}_0]$ e $\mathcal{V} = \mathbb{R}^3$, munido do produto interno canônico.
   (a) Determine uma base ortonormal para $\mathcal{U}$;
   (b) Determine uma base ortonormal para $\mathcal{U}^\perp$;
   (c) Para $D=4$, $E=5$ e $F=6$,[56] determine:
      i. $P_{\mathcal{U}}$;
      ii. $P_{\mathcal{U}^\perp}$;
      iii. as matrizes de $P_{\mathcal{U}}$ e $P_{\mathcal{U}^\perp}$ em alguma base de $\mathbb{R}^3$.

RESOLUÇÃO

(a) Se

$$\begin{aligned} \mathbf{x}' &= \frac{\mathbf{x}_0}{\|\mathbf{x}_0\|} \\ &= \frac{1}{\sqrt{(D+1)^2+(E+1)^2+(F+1)^2}}(D+1, E+1, F+1), \end{aligned}$$

então $\{\mathbf{x}'\}$ é uma base ortonormal de $\mathcal{U}$.[57]

[56] Confira a primeira nota de rodapé do exercício 2 supracitado.

[57] Para $D=4$, $E=5$ e $F=6$,

$$\mathbf{x}' = \frac{1}{\sqrt{5^2+6^2+7^2}}(5,6,7). \tag{7.24}$$

(b) Seja $\{\mathbf{y}_1, \mathbf{y}_2\}$ uma base de $\mathcal{U}^\perp$, isto é, considere $\mathbf{y}_1$ e $\mathbf{y}_2$ LI e $\langle \mathbf{x}_0, \mathbf{y}_j \rangle = 0$, $j = 1, 2$. Portanto, caso $\mathbf{y} = (a, b, c) \in \mathcal{U}^\perp$, ou seja, $\langle \mathbf{x}_0, \mathbf{y} \rangle = 0$, isto é,

$$(D+1)a + (E+1)b + (F+1)c = 0,$$

ou seja,

$$a = -\frac{E+1}{D+1}b - \frac{F+1}{D+1}c,$$

temos

$$\mathbf{y} = b\left(-\frac{E+1}{D+1}, 1, 0\right) + c\left(-\frac{F+1}{D+1}, 0, 1\right),$$

onde $\mathbf{y} = \mathbf{y}_1$, para $b = 1$ e $c = 0$, e $\mathbf{y} = \mathbf{y}_2$, para $b = 0$ e $c = 1$.[58,59]
Agora, se

$$\begin{aligned}
\mathbf{y}_1' &= \mathbf{y}_1 \\
&= \left(-\frac{E+1}{D+1}, 1, 0\right), \\
\mathbf{y}_2' &= \mathbf{y}_2 - \frac{\langle \mathbf{y}_2, \mathbf{y}_1' \rangle}{\|\mathbf{y}_1'\|^2}\mathbf{y}_1' \\
&= \left(-\frac{F+1}{D+1}, 0, 1\right) - \frac{\frac{(E+1)(F+1)}{(D+1)^2}}{\left(\frac{E+1}{D+1}\right)^2 + 1}\left(-\frac{E+1}{D+1}, 1, 0\right) \\
&= \left(-\frac{F+1}{D+1}, 0, 1\right) + \left(\frac{(E+1)^2(F+1)}{((E+1)^2 + (D+1)^2)(D+1)}, -\frac{(E+1)(F+1)}{(E+1)^2 + (D+1)^2}, 0\right) \\
&= \left(-\frac{(D+1)(F+1)}{(E+1)^2 + (D+1)^2}, -\frac{(E+1)(F+1)}{(E+1)^2 + (D+1)^2}, 1\right), \\
\mathbf{y}_1'' &= \frac{\mathbf{y}_1'}{\|\mathbf{y}_1'\|} \\
&= \frac{1}{\sqrt{\left(\frac{E+1}{D+1}\right)^2 + 1}}\left(-\frac{E+1}{D+1}, 1, 0\right) \text{ e} \\
\mathbf{y}_2'' &= \frac{\mathbf{y}_2'}{\|\mathbf{y}_2'\|} \\
&= \frac{1}{\sqrt{\left(\frac{(D+1)(F+1)}{(E+1)^2+(D+1)^2}\right)^2 + \left(\frac{(E+1)(F+1)}{(E+1)^2+(D+1)^2}\right)^2 + 1}}\left(-\frac{(D+1)(F+1)}{(E+1)^2 + (D+1)^2}, -\frac{(E+1)(F+1)}{(E+1)^2 + (D+1)^2}, 1\right),
\end{aligned}$$

$\{\mathbf{y}_1'', \mathbf{y}_2''\}$ é ortonormal, por Gram-Schmidt.[60]

(c) Sejam $D = 4$, $E = 5$, $F = 6$ e $\mathbf{x} = (x, y, z)$.

---

[58]Como $\mathbf{y}_1$ e $\mathbf{y}_2$ não são múltiplos um do outro, eles são LI.

[59]Para $D = 4$, $E = 5$ e $F = 6$,

$$\mathbf{y}_1 = \left(-\frac{6}{5}, 1, 0\right) \text{ e } \mathbf{y}_2 = \left(-\frac{7}{5}, 0, 1\right).$$

[60]Para $D = 4$, $E = 5$ e $F = 6$,

$$\mathbf{y}_1'' = \frac{1}{\sqrt{(6/5)^2 + 1}}\left(-\frac{6}{5}, 1, 0\right) \text{ e } \mathbf{y}_2'' = \frac{1}{\sqrt{(35/61)^2 + (42/61)^2 + 1}}\left(-\frac{35}{61}, -\frac{42}{61}, 1\right). \tag{7.25}$$

i. Por (7.24),

$$\begin{aligned}P_{\mathcal{U}}(\mathbf{x}) &= \langle \mathbf{x}, \mathbf{x}' \rangle \mathbf{x}' \\ &= \frac{1}{110}(5x+6y+7z)(5,6,7) \\ &= \frac{1}{110}(25x+30y+35z, 30x+36y+42z, 35x+42y+49z).\end{aligned}$$

ii. Em vez de calcularmos

$$P_{\mathcal{U}^\perp}(\mathbf{x}) = \langle \mathbf{x}, \mathbf{y}_1'' \rangle \mathbf{y}_1'' + \langle \mathbf{x}, \mathbf{y}_2'' \rangle \mathbf{y}_2''$$

via (7.25), utilizaremos a fórmula $P_{\mathcal{U}} + P_{\mathcal{U}^\perp} = I$, onde $I : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ é a função identidade.[61] Assim,

$$\begin{aligned}P_{\mathcal{U}^\perp}(\mathbf{x}) &= \mathbf{x} - P_{U}(\mathbf{x}) \\ &= (x,y,z) - \frac{1}{110}(25x+30y+35z, 30x+36y+42z, 35x+42y+49z) \\ &= \left(x - \frac{1}{110}(25x+30y+35z), y - \frac{1}{110}(30x+36y+42z), z - \frac{1}{110}(35x+42y+49z)\right) \\ &= \left(\frac{85}{110}x - \frac{30}{110}y - \frac{35}{110}z, -\frac{30}{110}x + \frac{74}{110}y - \frac{42}{110}z, -\frac{35}{110}x - \frac{42}{110}y + \frac{61}{110}z\right).\end{aligned}$$

iii. Para obtermos as matrizes pedidas, na base canônica, basta lembrarmos que a i-ésima coluna de cada uma delas é a imagem do i-ésimo vetor da base canônica pela projeção correspondente. Por exemplo, a segunda coluna da matriz de $P_{\mathcal{U}}$ na base canônica é dada por

$$\begin{bmatrix} 30/110 \\ 36/110 \\ 42/110 \end{bmatrix},$$

pois

$$P_{\mathcal{U}}(0,1,0) = \left(\frac{30}{110}, \frac{36}{110}, \frac{42}{110}\right).$$

Analogamente, a primeira coluna da matriz de $P_{\mathcal{U}^\perp}$ na base canônica é dada por

$$\begin{bmatrix} 85/110 \\ -30/110 \\ -35/110 \end{bmatrix},$$

pois

$$P_{\mathcal{U}^\perp}(1,0,0) = \left(\frac{85}{110}, -\frac{30}{110}, -\frac{35}{110}\right).$$

O cálculo das outras colunas fica a cargo do(a) leitor(a).

6. Considere o subespaço $\mathcal{U} = [(1/2,1/2,1/2,1/2)]$ de $\mathbb{R}^4$, munido do produto interno canônico. Caso $C, D, E, F \in \mathbb{R}$ sejam diferentes de -1 e $\mathbf{u} = (C+1, D+1, E+1, F+1)$, determine $P_{\mathcal{U}^\perp}(\mathbf{u})$.

RESOLUÇÃO

Como $(1/2,1/2,1/2,1/2) \in \mathcal{U}$ e, para cada $\mathbf{y} \in \mathcal{U}^\perp$, $\langle (1/2,1/2,1/2,1/2), \mathbf{y} \rangle = 0$, podemos obter uma base ortonormal $\{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3\}$ de $\mathcal{U}^\perp$, por Gram-Schmidt, e calcularmos

$$P_{\mathcal{U}^\perp}(\mathbf{u}) = \langle \mathbf{u}, \mathbf{y}_1 \rangle \mathbf{y}_1 + \langle \mathbf{u}, \mathbf{y}_2 \rangle \mathbf{y}_2 + \langle \mathbf{u}, \mathbf{y}_3 \rangle \mathbf{y}_3.$$

[61] Confira o item (e) do exercício 4 da seção 7.5.

Contudo, em vez disso, aplicaremos a fórmula $P_{\mathcal{U}^\perp} + P_{\mathcal{U}} = I$, onde $I : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ é a função identidade.[62] Portanto, basta obtermos $P_{\mathcal{U}}(\mathbf{u})$ e calcularmos $P_{\mathcal{U}^\perp}(\mathbf{u}) = \mathbf{u} - P_{\mathcal{U}}(\mathbf{u})$. Assim, como $\mathbf{x}$ é unitário,

$$\begin{aligned} P_{\mathcal{U}}(\mathbf{u}) &= \langle \mathbf{u}, \mathbf{x} \rangle \mathbf{x} \\ &= \left(\frac{C+1}{2} + \frac{D+1}{2} + \frac{E+1}{2} + \frac{F+1}{2} \right) (1/2, 1/2, 1/2, 1/2) \\ &= \frac{C+D+E+F+4}{2} (1/2, 1/2, 1/2, 1/2) \end{aligned}$$

e, portanto,

$$P_{\mathcal{U}^\perp}(\mathbf{u}) = (C+1, D+1, E+1, F+1) - \frac{C+D+E+F+4}{2} (1/2, 1/2, 1/2, 1/2).$$

Por exemplo, para $C = 3$, $D = 4$, $E = 5$ e $F = 6$,

$$\begin{aligned} P_{\mathcal{U}^\perp}(\mathbf{u}) &= (4, 5, 6, 7) - 11(1/2, 1/2, 1/2, 1/2) \\ &= (-3/2, -1/2, 1/2, 3/2). \end{aligned}$$

7.3 Funcionais lineares

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

DEFINIÇÃO

$L \in \mathcal{L}(\mathcal{V}, \mathbb{K}) \iff L$ é um *funcional linear em* $\mathcal{V}$.

EXEMPLOS

- Caso $\langle\ ,\ \rangle$ seja um produto interno em $\mathcal{V}$, real ou complexo,[63] e $u \in \mathcal{V}$ esteja fixado,

$$\mathcal{V} \ni v \overset{\langle -, u \rangle}{\longmapsto} \langle v, u \rangle \in \mathbb{K} \tag{7.26}$$

  é um funcional linear em $\mathcal{V}$, pela linearidade desse produto no primeiro fator.[64]

- Caso $\mathbf{y} \in \mathbb{R}^n$ esteja fixado,

$$\mathbb{R}^n \ni \mathbf{x} \longmapsto \mathbf{x} \cdot \mathbf{y} \in \mathbb{R}$$

  é um funcional linear em $\mathcal{V} = \mathbb{R}^n$, por (7.26).[65]

OBSERVAÇÃO

Caso $\mathcal{V}$ seja finitamente gerado, (7.26) é o exemplo geral de funcionais lineares, conforme o seguinte resultado:

7.3.1 Teorema da Representação de Riesz

Caso $\mathcal{V}$ seja finitamente gerado e $\varphi \in \mathcal{L}(\mathcal{V}, \mathbb{K})$, existe um único $u \in \mathcal{V}$ tal que

$$\varphi = \langle -, u \rangle.$$

[62] Confira o item (e) do exercício 4 da seção 7.5.

[63] Confira o exercício 4 de 6.4.

[64] Idem.

[65] Aqui, "$\cdot$" representa o produto interno canônico em $\mathbb{R}^n$.

DEMONSTRAÇÃO

Caso $v \in \mathcal{V}$ seja arbitrário e $\mathcal{B} = \{v_1, \ldots, v_r\}$ seja uma base ortonormal arbitrária de $\mathcal{V}$,

$$v = \sum_{i=1}^{r} \langle v, v_i \rangle v_i,$$

por (R30).[66] Assim, como φ é linear,

$$\begin{aligned}
\varphi(v) &= \varphi\left(\sum_{i=1}^{r} \langle v, v_i \rangle v_i\right) \\
&= \sum_{i=1}^{r} \langle v, v_i \rangle \varphi(v_i) \\
&= \left\langle v, \sum_{i=1}^{r} \overline{\varphi(v_i)} v_i \right\rangle, \\
&= \langle v, u \rangle,
\end{aligned}$$

onde

$$u := \sum_{i=1}^{r} \overline{\varphi(v_i)} v_i.$$

A unicidade de u segue do lema (R35), que veremos a seguir.

OBSERVAÇÕES

- Pela unicidade supracitada, φ não depende da base ortonormal $\mathcal{B}$ considerada;
- Pela demonstração supracitada, temos um método para calcularmos φ.

(R35) Lema para 7.3.1

Se $u, w \in \mathcal{V}$ e, para todo $v \in \mathcal{V}$, $\langle v, u \rangle = \langle v, w \rangle$, então $u = w$.

DEMONSTRAÇÃO

Como, para cada $v \in \mathcal{V}$,

$$\begin{aligned}
\langle v, u \rangle = \langle v, w \rangle &\Longrightarrow \overline{\langle u, v \rangle} = \overline{\langle w, v \rangle} \quad (\text{SIMETRIA CONJUGADA DE } \langle\ ,\ \rangle) \\
&\Longrightarrow \langle u, v \rangle = \langle w, v \rangle \quad (\overline{\overline{z}} = z \text{ PARA TODO } z \in \mathbb{C}) \\
&\Longrightarrow \langle u - w, v \rangle = 0 \quad (\langle\ ,\ \rangle \text{ É LINEAR NO PRIMEIRO FATOR}),
\end{aligned}$$

em particular, caso $v = u - w$,

$$\langle u - w, u - w \rangle = 0,$$

isto é, $u - w = 0$.

7.4 Adjuntos, autoadjuntos e normais. Teorema espectral. Forma de Jordan

CONVENÇÃO

Nessa seção, considere:

[66] Cf. p. 190.

· $\mathcal{V}$ finitamente gerado e dotado do produto interno $\langle\ ,\ \rangle$;[67]

· $T \in \mathcal{L}(\mathcal{V})$.

7.4.1 Construção do operador adjunto

Como vimos no primeiro exemplo da seção 7.3, para $w \in \mathcal{V}$ fixo, $\langle -, w\rangle \in \mathcal{L}(\mathcal{V}, \mathbb{K})$. Assim,

$$\varphi = \langle -, w\rangle \circ T \in \mathcal{L}(\mathcal{V}, \mathbb{K}). \tag{7.27}$$

Por outro lado, por 7.3.1, existe um único $u \in \mathcal{V}$ tal que

$$\varphi = \langle -, u\rangle. \tag{7.28}$$

Portanto, se

$$u := T^*(w),$$

temos uma função $T^* : \mathcal{V} \to \mathcal{V}$ tal que, para quaisquer $v, w \in \mathcal{V}$,

$$\langle T(v), w\rangle = \langle v, T^*(w)\rangle,$$

por (7.27) e (7.28)

DEFINIÇÃO T^* é dito *adjunto* de T.

EXEMPLO
Determinaremos T^* para

$$\mathbb{R}^2 \ni (x_1, x_2) \overset{T}{\longmapsto} (x_1 - x_2, x_1 + x_2) \in \mathbb{R}^2,$$

caso $\langle\ ,\ \rangle$ seja o produto interno usual em $\mathbb{R}^2$. Assim, para $(y_1, y_2) \in \mathbb{R}^2$ fixo e $(x_1, x_2) \in \mathbb{R}^2$ arbitrário,

$$\begin{aligned}
\langle (x_1, x_2), T^*(y_1, y_2)\rangle &= \langle T(x_1, x_2), (y_1, y_2)\rangle \\
&= \langle (x_1 - x_2, x_1 + x_2), (y_1, y_2)\rangle \\
&= x_1 y_1 - x_2 y_1 + x_1 y_2 + x_2 y_2 \\
&= x_1(y_1 + y_2) + x_2(y_2 - y_1) \\
&= \langle (x_1, x_2), (y_1 + y_2, y_2 - y_1)\rangle.
\end{aligned}$$

Portanto,

$$\mathbb{R}^2 \ni (y_1, y_2) \overset{T^*}{\longmapsto} (y_1 + y_2, y_2 - y_1) \in \mathbb{R}^2,$$

por (R35).

[67] Como no exercício 4 da seção 6.4.

7.4.2 $T \in \mathcal{L}(\mathcal{V}) \Longrightarrow T^* \in \mathcal{L}(\mathcal{V})$

DEMONSTRAÇÃO

Para $w_1, w_2 \in \mathcal{V}$ fixos e $v \in \mathcal{V}$ arbitrário, como

$$\begin{aligned}\langle v, T^*(w_1 + w_2)\rangle &= \langle T(v), w_1 + w_2\rangle \\ &= \langle T(v), w_1\rangle + \langle T(v), w_2\rangle \\ &= \langle v, T^*(w_1)\rangle + \langle v, T^*(w_2)\rangle \\ &= \langle v, T^*(w_1) + T^*(w_2)\rangle,\end{aligned}$$

temos

$$T^*(w_1 + w_2) = T^*(w_1) + T^*(w_2).$$ [68]

Para $\lambda \in \mathbb{K}$ e $w \in \mathcal{V}$ fixos e $v \in \mathcal{V}$ arbitrário, como

$$\begin{aligned}\langle v, T^*(\lambda w)\rangle &= \langle T(v), \lambda w\rangle \\ &= \overline{\lambda}\langle T(v), w\rangle \\ &= \overline{\lambda}\langle v, T^*(w)\rangle \\ &= \langle v, \lambda T^*(w)\rangle,\end{aligned}$$

temos

$$T^*(\lambda w) = \lambda T^*(w).$$ [69]

7.4.3 Propriedades do adjunto

Se $T, L \in \mathcal{L}(V)$, então:

1. $(T + L)^* = T^* + L^*$;
2. $(\lambda T)^* = \overline{\lambda} T^*$;
3. $(T^*)^* = T$;
4. *I é o operador identidade em $\mathcal{V} \Longrightarrow I^* = I$;*
5. $(TL)^* = L^* T^*$;
6. *Caso $\mathcal{B} = \{v_1, \ldots, v_n\}$ seja uma base ortonormal de $\mathcal{V}$,*

   $$A = \mathcal{M}(T, \mathcal{B}) \Longrightarrow \mathcal{M}(T^*, \mathcal{B}) = A^*,$$

   onde A^* é a conjugada transposta de A, como no exercício 12 da seção 5.2;
7. *Se T é invertível, então T^* é invertível e*

   $$(T^*)^{-1} = \left(T^{-1}\right)^*;$$

8. *λ é autovalor de $T \Longleftrightarrow \overline{\lambda}$ é autovalor de T^*.*

[68]Cf. (R35).
[69]Idem.

Demonstraremos as propriedades 3 e 6.[70]

DEMONSTRAÇÃO DE 7.4.2.3

Considere $v \in \mathcal{V}$ fixo e $w \in \mathcal{V}$ arbitrário. Logo,

$$\begin{aligned} \langle w, (T^*)^*(v) \rangle &= \langle T^*(w), v \rangle \\ &= \overline{\langle v, T^*(w) \rangle} \\ &= \overline{\langle T(v), w \rangle} \\ &= \langle w, T(v) \rangle. \end{aligned}$$

Então, $(T^*)^*(v) = T(v)$, por (R35).[71]

DEMONSTRAÇÃO DE 7.4.2.6

Como vimos na subseção 6.3.3, para $j \in \{1, \ldots, n\}$ fixo, a j-ésima coluna de A é obtida ao escrevermos

$$T(v_j) = \sum_{i=1}^{n} a_{ij} v_i.$$

Nesse caso, essa coluna é dada por

$$\begin{bmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{bmatrix},$$

isto é, a_{ij} é a entrada da linha i e coluna j de A, $i = 1, \ldots, n$. Por outro lado,

$$a_{ij} = \langle T(v_j), v_i \rangle, \; i = 1, \ldots, n,$$

por (R30).[72] Portanto, ao considerarmos T^* no lugar de T, a entrada da linha j e coluna i de $B = \mathcal{M}(T^*, \mathcal{B})$ é dada por

$$\begin{aligned} b_{ji} &= \langle T^*(v_i), v_j \rangle \\ &= \langle v_i, T(v_j) \rangle \\ &= \overline{\langle T(v_j), v_i \rangle} \\ &= \overline{a_{ij}}. \end{aligned}$$

7.4.4 Operador hermiteano, isto é, autoadjunto

DEFINIÇÃO

Para que T seja o operador supracitado, a seguinte condição deve ser satisfeita:

$$T^* = T,$$

ou seja,

$$v, w \in \mathcal{V} \Longrightarrow \langle T(v), w \rangle = \langle v, T(w) \rangle.$$

EXEMPLOS

- Para $A \in \mathbb{R}^{n \times n}$ simétrica (como na subseção 4.2.2),

  $$\mathbb{R}^n \ni \mathbf{x} \stackrel{T}{\longmapsto} A\mathbf{x} \in \mathbb{R}^n$$

  é autoadjunto;

[70] As demonstrações das outras propriedades foram deixadas como exercícios da seção 7.5.

[71] Cf. p. 205.

[72] Cf. p. 190.

· Para $A \in \mathbb{C}^{n\times n}$ hermiteana (como nos exercícios 12.3, 14 e 15 da seção 5.2),

$$\mathbb{C}^n \ni \mathbf{x} \stackrel{T}{\longmapsto} A\mathbf{x} \in \mathbb{C}^n$$

é autoadjunto.

De fato, para $\mathbb{K} = \mathbb{R}, \mathbb{C}$, caso $\langle\, ,\, \rangle$ seja o produto interno usual em $\mathbb{K}^n$, temos, para quaisquer $\mathbf{x}, \mathbf{y} \in \mathbb{K}^n$,

$$\begin{aligned}\langle T(\mathbf{x}), \mathbf{y}\rangle &= \langle \mathbf{x}, T^*(\mathbf{y})\,\rangle \text{ (pela definição de } T^*) \\ &= \langle \mathbf{x}, A^*\mathbf{y}\rangle \quad \text{(via 7.4.2.6)} \\ &= \langle \mathbf{x}, A\mathbf{y}\rangle \quad \text{(pois } A^* = A) \\ &= \langle \mathbf{x}, T(\mathbf{y})\rangle.\end{aligned}$$

OBSERVAÇÃO

Os autovalores para os operadores dos dois exemplos supracitados são reais. Na verdade, vale o seguinte resultado mais geral:

(R36) Autovalor real

$$\boxed{\lambda \textit{ é autovalor do operador hermiteano } T \Longrightarrow \lambda \in \mathbb{R}.}$$

DEMONSTRAÇÃO

Seja v autovetor de T associado a λ.[73] Assim, como

$$\begin{aligned}\lambda\|v\|^2 &= \lambda\langle v, v\rangle \\ &= \langle \lambda v, v\rangle \\ &= \langle T(v), v\rangle \\ &= \langle v, T(v)\rangle \\ &= \langle v, \lambda v\rangle \\ &= \overline{\lambda}\langle v, v\rangle \\ &= \overline{\lambda}\|v\|^2,\end{aligned}$$

temos $\lambda = \overline{\lambda}$ e, portanto, $\lambda \in \mathbb{R}$.

OBSERVAÇÃO

À guisa de comparação, confira a demonstração do segundo item da afirmação 7 da subseção 4.2.2, referente aos autovalores das matrizes simétricas.

(R37) Autoadjuntos nulos

$$\boxed{T \textit{ é autoadjunto e } \langle T(v), v\rangle = 0 \textit{ para todo } v \in \mathcal{V} \Longrightarrow T = 0.}$$

DEMONSTRAÇÃO

Sejam $u, w \in \mathcal{V}$ arbitrários. Portanto, como $\langle T(v), v\rangle = 0$ para cada $v \in \mathcal{V}$, temos

$$\langle T(u + w), u + w\rangle - \langle T(u), u\rangle - \langle T(w), w\rangle = 0.$$

[73] Lembre-se que autovetores são não nulos!

Assim,

$$\langle T(u), w\rangle + \langle T(w), u\rangle = 0. \tag{7.29}$$

Então,

$$\begin{aligned} 0 &= \langle T(u), w\rangle + \langle w, T(u)\rangle \text{ (T é hermiteano)} \\ &= \langle T(u), w\rangle + \overline{\langle T(u), w\rangle} \\ &= \text{ duas vezes a parte real de } \langle T(u), w\rangle. \end{aligned}$$

Logo,

$$\text{a parte real de } \langle T(u), w\rangle \text{ é nula.} \tag{7.30}$$

Agora, ao substituirmos w por iw em (7.29), temos a equação

$$\begin{aligned} 0 &= \langle T(u), iw\rangle + \langle iT(w), u\rangle \\ &= \bar{i}\langle T(u), w\rangle + i\langle T(w), u\rangle. \end{aligned}$$

Então, ao multiplicarmos cada membro da equação supracitada por i, temos

$$\begin{aligned} 0 &= \langle T(u), w\rangle - \langle T(w), u\rangle \\ &= \langle T(u), w\rangle - \langle w, T(u)\rangle \text{ (T é hermiteano)} \\ &= \langle T(u), w\rangle - \overline{\langle T(u), w\rangle} \\ &= \text{ duas vezes a parte imaginária de } \langle T(u), w\rangle. \end{aligned}$$

Logo,

$$\text{a parte imaginária de } \langle T(u), w\rangle \text{ é nula.} \tag{7.31}$$

Portanto,

$$\langle T(u), w\rangle = 0,$$

por (7.30) e (7.31). Então, como w é arbitrário, ao considerarmos $w = T(u)$, temos

$$\langle T(u), T(u)\rangle = 0,$$

isto é,

$$T(u) = 0.$$

Assim, como u é arbitrário, temos

$$T = 0.$$

7.4.5 Operador normal comuta com o seu adjunto

DEFINIÇÃO

Para que T seja um operador *normal*, a seguinte condição deve ser satisfeita:

$$TT^* = T^*T.$$

EXEMPLO

Todo operador autoadjunto é, claramente, normal. Contudo, a recíproca não é verdadeira.[74]

[74] Cf. o exercício 12 da seção 5.2.

(R38) Propriedades do operador normal T

1. *T é normal se, e somente se,*

$$\|T(v)\| = \|T^*(v)\| \ \forall v \in \mathcal{V};$$

2. *T é normal $\Longleftrightarrow$ T^* é normal;*
3. *T é normal $\Longrightarrow \mathrm{Nu}(T) = \mathrm{Nu}(T^*)$;*
4. *T é normal $\Longrightarrow T - \lambda I$ é normal $\forall \lambda \in \mathbb{K}$;*
5. *T é normal $\Longrightarrow T$ e T^* têm os mesmos autovetores;*
6. *Se λ_1 e λ_2 são autovalores distintos do operador normal T, com autovetores associados v_1 e v_2, respectivamente, então esses autovetores são ortogonais.*

DEMONSTRAÇÃO DE (R38.1)

$$\begin{aligned}
\|T(v)\|^2 = \|T^*(v)\|^2 \ \forall v \in \mathcal{V} &\Longleftrightarrow \langle T(v), T(v)\rangle = \left\langle T^*(v), T^*(v)\right\rangle \forall v \in \mathcal{V} \\
&\Longleftrightarrow \left\langle\, (T^*T)(v), v\right\rangle = \left\langle\, (TT^*)(v), v\right\rangle \forall v \in \mathcal{V} \\
&\Longleftrightarrow \left\langle\, (T^*T - TT^*)(v), v\right\rangle = 0\, \forall v \in \mathcal{V} \\
&\Longleftrightarrow T^*T - TT^* = 0,
\end{aligned}$$

pois, como $T^*T - TT^*$ é autoadjunto,[75] podemos usar (R37) na implicação "$\Longrightarrow$" da última equivalência, de cima para baixo.

DEMONSTRAÇÃO DE (R38.2)

É uma consequência direta de (R38.1).[76]

DEMONSTRAÇÃO DE (R38.3)

É uma consequência direta de (R38.1).

DEMONSTRAÇÃO DE (R38.4)

Via 7.4.3,

$$\begin{aligned}
(T - \lambda I)(T - \lambda I)^* &= (T - \lambda I)\left(T^* - \overline{\lambda} I\right) \\
&= TT^* - T\left(\overline{\lambda} I\right) - (\lambda I)T^* + (\lambda I)\left(\overline{\lambda} I\right) \\
&= T^*T - T^*(\lambda I) - \left(\overline{\lambda} I\right)T + \left(\overline{\lambda} I\right)(\lambda I) \\
&= \left(T^* - \overline{\lambda} I\right)(T - \lambda I) \\
&= (T - \lambda I)^*(T - \lambda I).
\end{aligned}$$

DEMONSTRAÇÃO DE (R38.5)

Caso λ seja um autovalor de T, isto é, $\overline{\lambda}$ seja um autovalor de T^*,[77]

$$\mathrm{Nu}(T - \lambda I) = \mathrm{Nu}\left(T^* - \overline{\lambda} I\right), \tag{7.32}$$

[75] Verifique!

[76] Cf. a propriedade 3 de 7.4.3.

[77] Cf. a propriedade 8 de 7.4.3.

por (R38.4), (R38.3) e 7.4.3.[78]

DEMONSTRAÇÃO DE (R38.6)

Como

$$\begin{aligned}(\lambda_1-\lambda_2)\langle v_1,v_2\rangle &= \lambda_1\langle v_1,v_2\rangle-\lambda_2\langle v_1,v_2\rangle\\&=\langle \lambda_1 v_1,v_2\rangle-\langle v_1,\overline{\lambda_2}v_2\rangle\\&=\langle T(v_1),v_2\rangle-\langle v_1,T^*(v_2)\rangle\\&=\langle T(v_1),v_2\rangle-\langle T(v_1),v_2\rangle\\&=0,\end{aligned}$$

onde usamos (7.32) na terceira igualdade, e $\lambda_1-\lambda_2\neq 0$, temos

$$\langle v_1,v_2\rangle=0.$$

OBSERVAÇÃO

À guisa de comparação com a demonstração supracitada, confira a demonstração do terceiro item da afirmação 7 da subseção 4.2.2, referente aos autovalores distintos das matrizes simétricas.

7.4.6 Teorema espectral complexo

Ao considerarmos a convenção inicial de 7.4, o teorema supracitado estabelece o seguinte:

Para $\mathbb{K}=\mathbb{C}$,

existe alguma base ortonormal $\mathcal{B}$ de $\mathcal{V}$ com $\mathcal{M}(T,\mathcal{B})$ diagonal

$$\Longleftrightarrow$$

T é normal.

EXEMPLO

Caso A seja a matriz do exercício 12.3 da seção 5.2, $T\in\mathcal{L}(\mathbb{C}^2)$ seja a multiplicação por A, isto é,

$$\mathbf{x}\overset{T}{\longmapsto}A\mathbf{x},$$

onde A é a matriz de T na base canônica, e $\mathbf{x}_1$ e $\mathbf{x}_2$ sejam os autovetores ortonormalizados calculados no exercício supracitado, temos

$$\mathcal{M}(T,\mathcal{B})=\begin{bmatrix}0&0\\0&2\end{bmatrix},$$

onde $\mathcal{B}=\{\mathbf{x}_1,\mathbf{x}_2\}$. Além disso, note que, como T é autoadjunto,[79] T é normal.

DEMONSTRAÇÃO DE 7.4.6

Para a implicação "$\Longrightarrow$", seja $\mathsf{D}=\mathcal{M}(T,\mathcal{B})$, isto é, $\mathsf{D}^*=\mathcal{M}(T^*,\mathcal{B})$,[80] que, obviamente, também é diagonal.

[78] Cf. 7.2.1, p. 179.
[79] Cf. o segundo item dos exemplos de 7.4.4.
[80] Cf. a propriedade 6 de 7.4.3.

Então, como D e D* comutam, T e T^* também comutam.[81]
Para a implicação "$\Longleftarrow$", note que, pelo lema de Schur,[82] existe uma base ortonormal $\mathcal{B} = \{v_1, \ldots, v_n\}$ de $\mathcal{V}$ tal que

$$\mathcal{M}(T, \mathcal{B}) = \begin{bmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{bmatrix} \tag{7.33}$$

é triangular superior, onde $*$ representa as entradas a_{ij} acima da diagonal principal e 0 representa a nulidade das entradas abaixo da diagonal principal. Além disso, pela propriedade 6 de 7.4.3,

$$\mathcal{M}(T^*, \mathcal{B}) = \begin{bmatrix} \overline{a_{11}} & & 0 \\ & \ddots & \\ \overline{*} & & \overline{a_{nn}} \end{bmatrix} \tag{7.34}$$

é triangular inferior, onde $\overline{*}$ representa as entradas $\overline{a_{ij}}$ abaixo da diagonal principal e 0 representa a nulidade das entradas acima da diagonal principal. Portanto,

$$T(v_1) = a_{11} v_1$$

e

$$T^*(v_1) = \overline{a_{11}} v_1 + \overline{a_{12}} v_2 + \cdots + \overline{a_{1n}} v_n,$$

por (7.33) e (7.34), respectivamente. Assim,

$$\|T(v_1)\|^2 = |a_{11}|^2$$

e

$$\|T^*(v_1)\|^2 = |a_{11}|^2 + |a_{12}|^2 + \cdots + |a_{1n}|^2.$$

Logo, como T é normal,

$$|a_{12}|^2 + \cdots + |a_{1n}|^2 = 0,$$

pela propriedade 1 de (R38). Portanto, como $a_{12} = \cdots = a_{1n} = 0$,

$$T(v_2) = a_{22} v_2$$

e

$$T^*(v_2) = \overline{a_{22}} v_2 + \overline{a_{23}} v_3 + \cdots + \overline{a_{2n}} v_n,$$

por (7.33) e (7.34), respectivamente. Assim,

$$\|T(v_2)\|^2 = |a_{22}|^2$$

e

$$\|T^*(v_2)\|^2 = |a_{22}|^2 + |a_{23}|^2 + \cdots + |a_{2n}|^2.$$

Então, como T é normal,

$$|a_{23}|^2 + \cdots + |a_{2n}|^2 = 0,$$

pela propriedade 1 de (R38). Portanto, $a_{23} = \cdots = a_{2n} = 0$ e, ao prosseguirmos com esse raciocínio, $* = 0$, ou seja, (7.33) é diagonal.

---

[81] De fato, considere $\mathcal{V} = \mathcal{W}$ e $\mathcal{B} = \mathcal{B}' = \mathcal{B}''$ na subseção 6.3.3. Assim, o isomorfismo (6.9) satisfaz a condição (6.10) e, portanto,

$$\begin{aligned} \mathcal{M}(TT^*, \mathcal{B}) &= \mathcal{M}(T, \mathcal{B}) \mathcal{M}(T^*, \mathcal{B}) \\ &= \mathsf{D}\mathsf{D}^* \\ &= \mathsf{D}^*\mathsf{D} \\ &= \mathcal{M}(T^*, \mathcal{B}) \, \mathcal{M}(T, \mathcal{B}) \\ &= \mathcal{M}(T^* T, \mathcal{B}), \end{aligned}$$

ou seja, $TT^* = T^*T$.

[82] Cf. a observação que precede (R24), p. 185.

7.4.7 Teorema espectral real

Ao considerarmos a convenção inicial de 7.4, o teorema supracitado estabelece o seguinte:

Para $\mathbb{K} = \mathbb{R}$,

existe alguma base ortonormal $\mathcal{B}$ *de* $\mathcal{V}$ *com* $\mathcal{M}(T, \mathcal{B})$ *diagonal*

$$\iff$$

T *é autoadjunto.*

EXEMPLOS

Confira os dois exemplos do final da subseção 4.2.2.

OBSERVAÇÃO

Antes da demonstração do teorema supracitado, alguns resultados preliminares são necessários. O primeiro deles é um fato algébrico decorrente de (*) da subseção 7.2.3:

(**) *Caso* $p(t) \in \mathcal{P}(\mathbb{R})$ *não seja um polinômio constante, podemos fatorá-lo de modo único, a menos da ordem de seus fatores, da seguinte maneira:*

$$p(t) = c\,(t - \lambda_1) \cdots (t - \lambda_m) \left(t^2 + b_1 t + c_1\right) \cdots \left(t^2 + b_M t + c_M\right),$$

onde $c, \lambda_1, \ldots, \lambda_m, b_1, \ldots, b_M, c_1, \ldots, c_M \in \mathbb{R}$, $c \neq 0$, $b_1^2 < 4c_1, \ldots, b_M^2 < 4c_M$, $m + M > 0$ *e podem existir apenas fatores lineares (caso* $M = 0$*) ou quadráticos (caso* $m = 0$*).*

Para uma demonstração de (**), confira o livro do AXLER, referenciado no capítulo 1. Além disso, as demonstrações de alguns dos próximos resultados foram adaptadas de demonstrações do livro supracitado.

(R39) Invertibilidade via autoadjuntos

Se T *é autoadjunto e* $b, c \in \mathbb{R}$ *são tais que* $b^2 < 4c$, *então*

$$T^2 + bT + cI$$

é invertível.

DEMONSTRAÇÃO DE (R39)

Se $v \in \mathcal{V}$ é não nulo, então

$$\begin{aligned}
\left\langle \left(T^2 + bT + cI\right)(v), v \right\rangle &= \left\langle T^2(v), v \right\rangle + b\langle T(v), v\rangle + c\langle v, v\rangle \\
&= \langle T(v), T(v)\rangle + b\langle T(v), v\rangle + c\|v\|^2 \\
&\geq \|T(v)\|^2 - |b||\langle T(v), v\rangle| + c\|v\|^2 \\
&\geq \|T(v)\|^2 - |b|\|T(v)\|\|v\| + c\|v\|^2,
\end{aligned}$$

onde utilizamos, de cima para baixo, os fatos seguintes:

· $\langle \ , \ \rangle$ é linear na primeira coordenada;

· T é autoadjunto;

· $-|b|$ é o mínimo entre $-b$ e b;

· Desigualdade de Cauchy-Schwarz.[83]

Portanto, como $v \neq 0_{\mathcal{V}}$,

$$\|T(v)\|^2 - |b|\|T(v)\|\|v\| + c\|v\|^2 = \left(\|T(v)\| - \frac{|b|\|v\|}{2}\right)^2 + \left(c - \frac{b^2}{4}\right)\|v\|^2 > 0,^{84}$$

ou seja,

$$\left\langle \left(T^2 + bT + cI\right)(v), v \right\rangle > 0,$$

isto é,

$$\left(T^2 + bT + cI\right)(v) \neq 0_{\mathcal{V}}.$$

Portanto, $\mathrm{Nu}\big(T^2 + bT + cI\big) = \{0_{\mathcal{V}}\}$, ou seja, $T^2 + bT + cI$ é invertível.[85]

(R40) Existência de autovalores para autoadjuntos

Se T é autoadjunto e $\mathcal{V} \neq \{0_{\mathcal{V}}\}$, então T tem autovalor (em $\mathbb{K}$).

DEMONSTRAÇÃO DE (R40)

Por 7.2.3, resta demonstrarmos o resultado para $\mathbb{K} = \mathbb{R}$. Assim, para $v \in \mathcal{V}$ não nulo, como a lista

$$v, T(v), T^2(v), \ldots, T^{\dim \mathcal{V}}(v)$$

é formada por $\dim \mathcal{V} + 1$ vetores LD em $\mathcal{V}$, existe alguma lista de escalares reais, não todos nulos, digamos

$$a_0, a_1, a_2, \ldots, a_{\dim \mathcal{V}},$$

tal que

$$\sum_{j=0}^{\dim \mathcal{V}} a_j T^j(v) = 0_{\mathcal{V}}. \tag{7.35}$$

Por outro lado, se

$$p(t) = \sum_{j=0}^{\dim \mathcal{V}} a_j t^j \in \mathcal{P}(\mathbb{R}), \tag{7.36}$$

então, pelo fato algébrico (**) supracitado,

$$p(t) = c\left(t^2 + b_1 t + c_1\right) \cdots \left(t^2 + b_M t + c_M\right)(t - \lambda_1) \cdots (t - \lambda_m), \tag{7.37}$$

onde $c, \lambda_1, \ldots, \lambda_m, b_1, \ldots, b_M, c_1, \ldots, c_M \in \mathbb{R}$, $c \neq 0$, $b_i^2 < 4c_i$, $i = 1, \ldots, M$, e $M + m > 0$, com $M \geq 0$ ou $m \geq 0$. Note que

$$\left(c\left(T^2 + b_1 T + c_1 I\right) \cdots \left(T^2 + b_M T + c_M I\right)(T - \lambda_1 I) \cdots (T - \lambda_m I)\right)(v) = 0_{\mathcal{V}}, \tag{7.38}$$

por 7.2.2, (7.35), (7.36) e (7.37). Portanto:

· Caso $M = 0$, temos $m > 0$ e (7.38) reduzida a equação

$$((T - \lambda_1 I) \cdots (T - \lambda_m I))(v) = 0_{\mathcal{V}}; \tag{7.39}$$

[83] Cf. o ex. 4.1, seção 6.4.

[84] Completando quadrados, obtemos o segundo membro da igualdade, que é positivo, pois $b^2 < 4c$.

[85] Cf. (R12) e (R18), pp. 156 e 177, respectivamente.

· Caso $M > 0$, como $T^2 + b_i T + c_i I$ é invertível, $i = 1, \ldots, M$,[86] temos

$$m > 0,$$

pois, como $c \neq 0$, se $m = 0$, o operador invertível

$$c\left(T^2 + b_1 T + c_1 I\right) \cdots \left(T^2 + b_M T + c_M I\right) \tag{7.40}$$

leva $v \neq 0_{\mathcal{V}}$ em $0_{\mathcal{V}}$, por (7.38), que é um absurdo.[87] Portanto, para que (7.38) seja válida, (7.39) deve ser verdadeira.[88]

Para concluirmos a demonstração, observe que, caso $T - \lambda_k I$ seja invertível, $k = 1, \ldots, m$,

$$\prod_{k=1}^{m} (T - \lambda_k I)$$

é invertível. Logo, (7.39) é válida apenas para $v = 0_{\mathcal{V}}$,[89] ou seja, chegamos a uma contradição da hipótese inicial sobre a não nulidade de v. Assim, como algum $T - \lambda_k I$ é não invertível, algum λ_k é autovalor de T. [90]

DEMONSTRAÇÃO DE 7.4.7

Para a implicação "$\Longrightarrow$", seja $\mathsf{D} = \mathcal{M}(T, \mathcal{B})$, isto é, $\mathsf{D}^* = \mathcal{M}(T^*, \mathcal{B})$,[91] que, obviamente, também é diagonal. Logo, como

$$\begin{aligned}
\mathsf{D}^* &= \overline{\mathsf{D}}^t \text{ (D^* é a transposta conjugada de D)} \\
&= \mathsf{D}^t \text{ (D tem apenas entradas reais)} \\
&= \mathsf{D} \text{ (a transposta de uma matriz diagonal é a própria matriz),}
\end{aligned}$$

$T^* = T$, pois, se $\dim \mathcal{V} = n$, então

$$\begin{array}{rcccc}
\mathcal{M}(-, \mathcal{B}) & : & \mathcal{L}(\mathcal{V}) & \longrightarrow & \mathbb{K}^{n \times n} \\
& & L & \longmapsto & \mathcal{M}(L, \mathcal{B})
\end{array}$$

é um isomorfismo.[92]

Demonstraremos a implicação "$\Longleftarrow$" por indução sobre $n = \dim \mathcal{V}$. Na verdade,[93] basta demonstrarmos que

$$T \text{ é autoadjunto} \Longrightarrow \mathcal{V} \text{ admite base ortonormal } \mathcal{B} \text{ consistindo de autovetores de } T. \tag{7.41}$$

Obviamente, (7.41) é verdadeira para $n = 1$. Suponha, então, $n > 1$ e (7.41) verdadeira para qualquer espaço com produto interno real cuja dimensão pertença ao conjunto $\{1, 2 \ldots, n-1\}$.[94] Agora, por (R40), T admite algum autovetor u que, sem perda de generalidade, pode ser considerado unitário.[95] Portanto, $\mathcal{U} = [u]$ é invariante por T e

$$T|_{\mathcal{U}^\perp} \in \mathcal{L}\left(\mathcal{U}^\perp\right)$$

é autoadjunto, pelo exercício 8 da seção 7.5. Logo, podemos aplicar (7.41) ao operador $T|_{\mathcal{U}^\perp}$, pela hipótese de indução supracitada. Portanto, para concluirmos a demonstração, basta considerarmos o resultado (R32), a seção 6.5 e $\mathcal{B} = \{u, u_1, \ldots, u_{n-1}\}$, onde $\{u_1, \ldots, u_{n-1}\}$ representa uma base ortonormal de $\mathcal{U}^\perp$, composta de autovetores de $T|_{\mathcal{U}^\perp}$.

---

[86] Cf. (R39).

[87] Cf. (R12) e (R18), pp. 156 e 177.

[88] De fato, a única pré-imagem de $0_{\mathcal{V}}$ pelo operador invertível (7.40) é o próprio $0_{\mathcal{V}}$.

[89] Cf. (R19), p. 180.

[90] Idem.

[91] Cf. a propriedade 6 de 7.4.3.

[92] Cf. a subseção 6.3.3.

[93] Cf. (R26), p. 187.

[94] Essa é a nossa *hipótese de indução*.

[95] Caso u não seja unitário, considere o versor de u.

OBSERVAÇÃO

Os operadores das subseções 7.4.6 e 7.4.7 são diagonalizáveis. O próximo resultado estabelece que operadores não diagonalizáveis sobre espaços vetoriais complexos são sempre *diagonalizáveis por blocos de Jordan*.

7.4.8 Teorema (forma normal de Jordan)

Se $T \in \mathcal{L}(\mathcal{V})$ e $\mathcal{V}$ é finitamente gerado sobre $\mathbb{C}$, existe alguma base $\mathcal{B}$ de $\mathcal{V}$ tal que

$$\mathcal{M}(T, \mathcal{B}) = J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{pmatrix},$$

onde, para cada $j \in \{1, \ldots, k\}$,

$$J_j = \begin{pmatrix} \lambda_j & 1 & & & \\ & \lambda_j & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_j & 1 \\ & & & & \lambda_j \end{pmatrix}$$

e λ_j é um autovalor de T.

OBSERVAÇÕES

- As entradas de J diferentes de 1 e λ_j, $j = 1, \ldots, k$, são nulas.
- J é dita *diagonal por blocos de Jordan*.
- Alguns dos autovalores supracitados podem ser iguais.[96]
- A demonstração do teorema supracitado será precedida por dois lemas, juntamente com suas respectivas demonstrações.

Lema 1

Seja T como no teorema supracitado e considere a lista $\lambda_1, \ldots, \lambda_r$ dos autovalores distintos de T. Então,

$$\mathcal{V} = \bigoplus_{j=1}^{r} \mathrm{Nu}\Big((T - \lambda_j I)^{\dim \mathcal{V}} \Big).$$

DEMONSTRAÇÃO DO LEMA 1

Pela subseção 7.2.3, podemos considerar algum autovalor λ de T na lista supracitada.[97]

[96] Cf. a seção 7.6.

[97] $r = 1$ é uma possibilidade!

Dividiremos a demonstração em 5 partes:

PARTE 1

Para cada inteiro positivo i, considere

$$\mathcal{U}_i(\lambda) := \mathrm{Nu}\Big((T-\lambda I)^i\Big)$$

e, caso o autovalor considerado esteja implícito no contexto,

$$\mathcal{U}_i(\lambda) = \mathcal{U}_i.$$ [98]

Portanto,

$$\mathcal{U}_1 \subset \mathcal{U}_2 \subset \ldots \subset \mathcal{U}_i \subset \ldots \tag{7.42}$$ [99]

Logo, como $\mathcal{V}$ é finitamente gerado, a sequência (7.42) não é estritamente crescente, ou seja, a condição

$$\dim \mathcal{U}_i < \dim \mathcal{U}_{i+1}$$

não é válida para todo inteiro positivo i. Assim, existe o menor inteiro ℓ satisfazendo à condição

$$\mathcal{U}_\ell = \mathcal{U}_{\ell+1}. \tag{7.43}$$ [100]

Portanto,

$$\mathcal{U}_{\ell+1} = \mathcal{U}_{\ell+2} = \mathcal{U}_{\ell+3} = \ldots \tag{7.44}$$ [101]

Informação adicional

Chamamos $0 \neq v \in \mathcal{U}_{\dim \mathcal{V}}(\lambda)$ de *autovetor generalizado de T* e $\mathcal{U}_{\dim \mathcal{V}}$ de *autoespaço generalizado de T*, ambos *associados a λ*.

Note que:

I. $\ell \leq \dim \mathcal{V}$;

II. v é autovetor de $T \Longrightarrow v$ é autovetor generalizado de T.

De fato, para I, suponha $\ell > \dim \mathcal{V}$. Logo, por (7.42) e (7.43), a sequência

$$\mathcal{U}_1 \subset \mathcal{U}_2 \subset \ldots \subset \mathcal{U}_{\dim \mathcal{V}} \subset \mathcal{U}_{\dim \mathcal{V}+1}$$

é estritamente crescente, que é uma afirmação absurda, pois $\mathcal{U}_{\dim \mathcal{V}+1}$ não pode ter dimensão maior que $\dim \mathcal{V}$. Agora, para II, basta observarmos que

$$\mathcal{U}_1 \subset \mathcal{U}_\ell = \mathcal{U}_{\dim \mathcal{V}},$$

por (7.42), (7.43), (7.44) e I.

PARTE 2

Para provarmos que

$$\mathrm{Nu}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big) \cap \mathrm{Im}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big) = \{0_{\mathcal{V}}\},$$

consideremos $v \in \mathcal{V}$ pertencente à essa interseção. Então, $(T-\lambda I)^{\dim \mathcal{V}}(v) = 0_{\mathcal{V}}$ e existe $u \in \mathcal{V}$ tal que $(T-\lambda I)^{\dim \mathcal{V}}(u) = v$. Logo,

$$(T-\lambda I)^{2(\dim \mathcal{V})}(u) = (T-\lambda I)^{\dim \mathcal{V}}(v) = 0_{\mathcal{V}}.$$

Assim, $u \in \mathrm{Nu}\Big((T-\lambda I)^{2(\dim \mathcal{V})}\Big)$. Portanto, como $\mathcal{U}_{\dim \mathcal{V}} = \mathcal{U}_{2(\dim \mathcal{V})}$ (pela parte 1 supracitada), $u \in \mathcal{U}_{\dim \mathcal{V}}$, isto é,

$$v = (T-\lambda I)^{\dim \mathcal{V}}(u) = 0_{\mathcal{V}}.$$

[98] Note que,

$$v \in \mathcal{U}_i \Longleftrightarrow (T-\lambda I)^i(v) = 0.$$

[99] Cf. o antepenúltimo exercício da seção 7.5.

[100] Observe que ℓ depende do autovalor λ considerado.

[101] Cf. o antepenúltimo exercício da seção 7.5.

PARTE 3
Pela parte 2 supracitada e por (R14),[102]

$$\mathcal{V} = \mathrm{Nu}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big) \bigoplus \mathrm{Im}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big).$$

PARTE 4
Vamos demonstrar que $\mathrm{Nu}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big)$ e $\mathrm{Im}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big)$ são subespaços invariantes por T.[103] Para isso, note que

$$\begin{aligned} T(T-\lambda I) &= TT - \lambda(TI) \\ &= TT - \lambda(IT) \\ &= (T-\lambda I)T. \end{aligned}$$

Assim, por um lado,

$$v \in \mathrm{Nu}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big)$$

$$\Downarrow$$

$$\begin{aligned} (T-\lambda I)^{\dim \mathcal{V}} T(v) &= T(T-\lambda I)^{\dim \mathcal{V}}(v) \\ &= T(0) \\ &= 0. \end{aligned}$$

$$\Downarrow$$

$$T(v) \in \mathrm{Nu}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big).$$

Por outro,

$$v \in \mathrm{Im}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big)$$

$$\Downarrow$$

$$\text{existe } u \in \mathcal{V} \text{ tal que } (T-\lambda I)^{\dim \mathcal{V}}(u) = v$$

$$\Downarrow$$

$$\begin{aligned} (T-\lambda I)^{\dim \mathcal{V}} T(u) &= T(T-\lambda I)^{\dim \mathcal{V}}(u) \\ &= T(v) \end{aligned}$$

$$\Downarrow$$

$$T(v) \in \mathrm{Im}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big).$$

PARTE 5
O lema 1 pode ser demonstrado por indução sobre r, o número de autovalores distintos de T. De fato, seja $\lambda_1 = \lambda$. Assim, como o caso $r = 1$ é trivial, suponha $r > 1$ e que o lema seja válido para operadores que possuam $r-1$ autovalores distintos. Portanto, como os autovalores da restrição de T a $\mathrm{Im}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big)$, que é um operador sobre esse subespaço (pela parte 4 supracitada), são $\lambda_2, \ldots, \lambda_r$,[104]

$$\mathrm{Im}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big) = \bigoplus_{j=2}^{r} \mathrm{Nu}\Big((T-\lambda_j I)^{\dim \mathcal{V}}\Big), \tag{7.45}$$

[102]Cf. p. 157.
[103]Logo, a restrição de T a cada um desses subespaços é um operador.
[104]De fato, pela parte 2 supracitada, nenhum dos autovetores generalizados de $T|_{\mathrm{Im}\left((T-\lambda I)^{\dim \mathcal{V}}\right)}$ pode estar associado a λ, pois os autovetores generalizados de T associados a λ pertencem a $\mathrm{Nu}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big)$.

pela hipótese de indução. Logo, para concluirmos a demonstração, basta substituirmos (7.45) na equação da parte 3 supracitada.

OBSERVAÇÃO

O lema 1 supracitado estabelece que

> *$\mathcal{V}$ pode ser escrito como soma direta dos autoespaços generalizados associados aos autovalores distintos de T.*

EXERCÍCIO

Defina $T \in \mathcal{L}(\mathbb{C}^3)$ por

$$(z_1, z_2, z_3) \stackrel{T}{\longmapsto} (z_2, 0, z_3)\,.$$

1. Determine:
   (a) os autovalores de T;
   (b) os autoespaços de T;
   (c) os autoespaços generalizados de T.
2. Mostre que $\mathbb{C}^3$ é soma direta dos autoespaços generalizados de T associados aos autovalores distintos de T.

RESOLUÇÃO

(a)

$$\begin{aligned} T(z_1,z_2,z_3) = \lambda(z_1,z_2,z_3) &\Longrightarrow (z_2,0,z_3) = (\lambda z_1, \lambda z_2, \lambda z_3) \\ &\Longrightarrow \lambda \in \{0,1\} \text{ é autovalor de } T. \end{aligned}$$

(b)

$$\begin{aligned} \lambda = 0 \text{ e } T(z_1,z_2,z_3) = \lambda(z_1,z_2,z_3) &\Longrightarrow (z_2,0,z_3) = (0,0,0) \\ &\Longrightarrow z_2 = z_3 = 0 \\ &\Longrightarrow \mathrm{Nu}\,(T - 0I) = \{(z_1,0,0) \,:\, z_1 \in \mathbb{C}\}\,; \end{aligned}$$

$$\begin{aligned} \lambda = 1 \text{ e } T(z_1,z_2,z_3) = \lambda(z_1,z_2,z_3) &\Longrightarrow (z_2,0,z_3) = (z_1,z_2,z_3) \\ &\Longrightarrow z_1 = z_2 = 0 \\ &\Longrightarrow \mathrm{Nu}\,(T - 1I) = \{(0,0,z_3) \,:\, z_3 \in \mathbb{C}\}\,. \end{aligned}$$

> *Note que $\mathbb{C}^3$ não possui base de autovetores de T!*

(c) Primeiramente, para $\lambda = 0$, como

$$\begin{aligned} (T - \lambda I)^3 (z_1,z_2,z_3) &= T^3(z_1,z_2,z_3) \\ &= T^2(z_2,0,z_3) \\ &= T(0,0,z_3) \\ &= (0,0,z_3)\,, \end{aligned}$$

$$\mathrm{Nu}\Big((T - 0I)^3\Big) = \{(z_1,z_2,0) \,:\, z_1,z_2 \in \mathbb{C}\}\,.$$

Agora, para $\lambda = 1$, como

$$\begin{aligned}(T-\lambda I)^3(z_1,z_2,z_3) &= (T-I)^2((T-I)(z_1,z_2,z_3))\\ &= (T-I)^2(T(z_1,z_2,z_3)-(z_1,z_2,z_3))\\ &= (T-I)^2((z_2,0,z_3)-(z_1,z_2,z_3))\\ &= (T-I)^2(z_2-z_1,-z_2,0)\\ &= (T-I)((T-I)(z_2-z_1,-z_2,0))\\ &= (T-I)(T(z_2-z_1,-z_2,0)-(z_2-z_1,-z_2,0))\\ &= (T-I)((-z_2,0,0)-(z_2-z_1,-z_2,0))\\ &= (T-I)(-2z_2+z_1,z_2,0)\\ &= T(-2z_2+z_1,z_2,0)-(-2z_2+z_1,z_2,0)\\ &= (z_2,0,0)-(-2z_2+z_1,z_2,0)\\ &= (3z_2-z_1,-z_2,0),\end{aligned}$$

$$\mathrm{Nu}\Big((T-1I)^3\Big) = \{(0,0,z_3)\,:\,z_3\in\mathbb{C}\}.$$

2. Pelo item (c) do exercício 1 supracitado,

$$\mathbb{C}^3 = \mathrm{Nu}\Big((T-0I)^3\Big)\oplus\mathrm{Nu}\Big((T-1I)^3\Big).$$

Lema 2

Considere $L\in\mathcal{L}(\mathcal{V})$, $\mathcal{V}$ finitamente gerado sobre $\mathbb{C}$ e L nilpotente, isto é, existe algum inteiro positivo p tal que $L^p = 0$. Então, existe uma lista de vetores $v_1,\ldots,v_s$ de $\mathcal{V}$ e existe uma lista de inteiros positivos $p_1,\ldots,p_s$ tais que

$$L^{p_k}(v_k) = 0_{\mathcal{V}},\ k=1,\ldots,s,$$

e

$$\Big\{L^{p_1-1}(v_1),\ldots,L(v_1),v_1,\ldots,L^{p_s-1}(v_s),\ldots,L(v_s),v_s\Big\}$$

é uma base de V.

Exemplo de nilpotência

Caso $\lambda,\mathcal{U}_i$ e $i\geq\ell$ sejam como na demonstração do lema 1, a restrição

$$(T-\lambda I)|_{\mathcal{U}_i}\in\mathcal{L}(\mathcal{U}_i)$$

é nilpotente. De fato, essa restrição é um operador em $\mathcal{U}_i$, pois, como

$$\begin{aligned}\mathcal{U}_i &= \mathcal{U}_\ell\\ &= \mathcal{U}_{\dim\mathcal{V}}\end{aligned}\qquad(7.46)$$

é invariante por T,[105] $\mathcal{U}_i$ é invariante por $T-\lambda I$,[106] e sua nilpotência é trivial, pois, para cada $v\in\mathcal{U}_i$,

$$\begin{aligned}\big((T-\lambda I)|_{\mathcal{U}_i}\big)^\ell(v) &= (T-\lambda I)^\ell(v)\\ &= 0_{\mathcal{V}}.\end{aligned}$$

[105]Cf. (7.43), (7.44), a informação adicional e a parte 4 da demonstração do lema 1.
[106]Qualquer subespaço $\mathcal{U}$ de $\mathcal{V}$ é invariante por λI.

DEMONSTRAÇÃO DO LEMA 2 [107]

Note que,

$$\text{o lema 2 é válido para } L = 0. \tag{7.47}$$

De fato, basta considerarmos uma base $\{v_1, \ldots, v_s\}$ arbitrária de $\mathcal{V}$ e $p_1 = \cdots = p_s = 1$.[108] Por outro lado, para $L \neq 0$, podemos aplicar indução sobre $\dim \mathcal{V}$. Assim, se $\dim \mathcal{V} = 1$, então

$$L^p = 0 \iff L = 0,^{109} \tag{7.48}$$

ou seja, o lema 2 é válido para $\mathcal{V}$ unidimensional, por (7.47). Agora, se $\dim \mathcal{V} > 1$, suponha que

$$\text{o lema 2 é válido para qualquer espaço que tenha dimensão inferior a } \dim \mathcal{V} \tag{7.49}$$

e considere

$$\mathcal{U} = \operatorname{Im}(L). \tag{7.50}$$

- Caso $\mathcal{U}$ seja nulo, $L = 0$ e, para tal operador, o lema 2 é válido.[110]
- Caso $\mathcal{U} = \mathcal{V}$, L é injetivo.[111] Logo, L^p é injetivo, que é uma contradição da nilpotência de L.
- Caso $\mathcal{U}$ seja um subespaço não trivial de $\mathcal{V}$, isto é,

$$0 < \dim \mathcal{U} < \dim \mathcal{V},$$

a hipótese de indução (7.49) é aplicável a $\mathcal{U}$. Então, existe uma lista de vetores $u_1 \ldots, u_r$ de $\mathcal{U}$ e uma lista de inteiros positivos $q_1, \ldots, q_r$ tais que

$$L^{q_j}(u_j) = 0_{\mathcal{V}},\ j = 1, \ldots, r, \tag{7.51}$$

e

$$\left\{L^{q_1-1}(u_1), \ldots, L(u_1), u_1, \ldots, L^{q_r-1}(u_r), \ldots, L(u_r), u_r\right\} \tag{7.52}$$

é uma base de $\mathcal{U}$. Portanto, como existe $v_j \in \mathcal{V}$, $j = 1, \ldots, r$, tal que $L(v_j) = u_j$, por (7.50), e

$$L^i(v_j) = L^{i-1}(u_j), \tag{7.53}$$

para todo inteiro positivo i, os vetores da lista $L^{q_1}(v_1), \ldots, L^{q_r}(v_r)$ são LI em $\operatorname{Nu}(L)$, por (7.51) e (7.52). Logo, podemos estender a lista supracitada, caso seja necessário, a uma base

$$\left\{L^{q_1}(v_1), \ldots, L^{q_r}(v_r), w_1, \ldots, w_q\right\} \tag{7.54}$$

de $\operatorname{Nu}(L)$.[112] Assim,

$$\left\{L^{q_1}(v_1), \ldots, L(v_1), v_1, \ldots, L^{q_r}(v_r), \ldots, L(v_r), v_r, w_1, \ldots, w_q\right\} \tag{7.55}$$

é uma base de $\mathcal{V}$.[113]

---

[107] Confira

MARK WILDON, ROYAL HOLLOWAY, UNIVERSITY OF LONDON
A SHORT PROOF OF THE EXISTENCE OF JORDAN NORMAL FORM

[108] Nesse caso, $s = \dim \mathcal{V}$.
[109] Cf. a resolução do penúltimo exercício da seção 7.5.
[110] Cf. (7.47).
[111] Cf. (R12) e (R14), pp. 156 e 157.
[112] Cf. (R6), p. 151.
[113] Cf. a resolução do último exercício da seção 7.5.

DEMONSTRAÇÃO DO TEOREMA 7.4.7

Considere λ_i como no lema 1,

$$\mathcal{U}_{\lambda_i} = \mathrm{Nu}\Big((T-\lambda_i I)^{\dim \mathcal{V}}\Big) \tag{7.56}$$

e

$$\mathcal{B} = \bigcup_{i=1}^{r} \mathcal{B}_i,$$

onde $\mathcal{B}_i$ é uma base arbitrária de $\mathcal{U}_{\lambda_i}$, $i = 1, \ldots, r$. Então, como

$$T|_{\mathcal{U}_{\lambda_i}} \in \mathcal{L}(\mathcal{U}_{\lambda_i}),\ i = 1, \ldots, r,$$ [114]

temos

$$\mathcal{M}(T,\mathcal{B}) = \begin{pmatrix} \mathcal{M}\Big(T|_{\mathcal{U}_{\lambda_1}}, \mathcal{B}_1\Big) & & \text{zeros} \\ & \ddots & \\ \text{zeros} & & \mathcal{M}\Big(T|_{\mathcal{U}_{\lambda_r}}, \mathcal{B}_r\Big) \end{pmatrix}.$$

Portanto, para $i = 1, \ldots, r$, como

$$L_i := (T - \lambda_i I)\,|_{\mathcal{U}_{\lambda_i}}$$

é nilpotente,[115] caso $\mathcal{B}_i$ seja uma base como a do lema 2,

$$\begin{aligned} \mathcal{M}\Big(T|_{\mathcal{U}_{\lambda_i}}, \mathcal{B}_i\Big) &= \mathcal{M}\Big(\lambda_i I|_{\mathcal{U}_{\lambda_i}}, \mathcal{B}_i\Big) + \mathcal{M}(L_i, \mathcal{B}_i) \\ &= \begin{pmatrix} J_{i,1} & & \text{zeros} \\ & \ddots & \\ \text{zeros} & & J_{i,s_i} \end{pmatrix}, \end{aligned}$$

onde, para $j = 1, \ldots, s_i$, as colunas do bloco

$$J_{i,j} = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \text{zeros} \\ & & \ddots & \ddots & \\ \text{zeros} & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix}$$

são obtidas ao aplicarmos

$$T|_{\mathcal{U}_{\lambda_i}} = L_i + \lambda_i I|_{\mathcal{U}_{\lambda_i}}$$

à cada vetor da base

$$\mathcal{B}_{i,j} := \Big\{L_i^{p_j - 1}(v_j), \ldots, L_i(v_j), v_j\Big\}, \tag{7.57}$$

com

$$\mathcal{B}_i = \bigcup_{j=1}^{s_i} \mathcal{B}_{i,j}. \tag{7.58}$$

OBSERVAÇÃO

O primeiro vetor da base (7.57) é um autovetor de T associado ao autovalor λ_i.[116]

DEFINIÇÃO

A lista de vetores de (7.57) é um *ciclo* da base $\mathcal{B}_i$, $i = 1, \ldots, r$, de *comprimento* p_j.

---

[114] Cf. a parte 4 da demonstração do lema 1.

[115] Cf. o exemplo de nilpotência que precede a demonstração do lema 2.

[116] Cf. a primeira coluna de $J_{i,j}$.

EXEMPLO.1[117]

Para $\mathbb{C}^3 = \mathcal{U}_\lambda$ e

$$\begin{aligned} A &= \mathcal{M}(T, \mathcal{B}) \\ &= \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}, \end{aligned}$$

$\mathcal{B}$ consiste de um único ciclo. De fato, como $\mathbf{e}_1$ é um autovetor associado ao (único) autovalor λ de T e $(A - \lambda\mathsf{I})\mathbf{e}_3 = \mathbf{e}_2$, $(A - \lambda\mathsf{I})^2\mathbf{e}_3 = \mathbf{e}_1$ e $(A - \lambda\mathsf{I})^3\mathbf{e}_3 = \mathbf{0}$,

$$\begin{aligned} \mathcal{B} &= \left\{(A - \lambda\mathsf{I})^2\mathbf{e}_3, (A - \lambda\mathsf{I})\mathbf{e}_3, \mathbf{e}_3\right\} \\ &= \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}. \end{aligned}$$

DEFINIÇÃO

$m_i := \dim\big(\mathcal{U}_{\lambda_i}\big)$ é a *multiplicidade algébrica* de λ_i, $i = 1, \ldots, r$.[118]

OBSERVAÇÕES

- Para $i = 1, \ldots, r$,

$$m_i = \sum_{j=1}^{s_i} p_j,$$

  ou seja, m_i é a soma dos comprimentos dos ciclos de $\mathcal{B}_i$, por (7.58).

- Pelo lema 1,

$$\dim \mathcal{V} = \sum_{i=1}^{r} m_i. \tag{7.59}$$

DEFINIÇÃO

$\mu_i := \dim(\mathrm{Nu}\,(T - \lambda_i I))$ é a *multiplicidade geométrica* de λ_i, $i = 1, \ldots, r$.[119]

OBSERVAÇÃO

Pela parte 1 da demonstração do lema 1,

$$\mu_i \leq m_i, \, i = 1, \ldots, r,$$

isto é, a multiplicidade algébrica de um autovalor majora a sua multiplicidade geométrica.

EXEMPLO.2

Na exemplo.1, as multiplicidades algébrica e geométrica de λ são dadas, respectivamente, por $m = 3$ e $\mu = 1$.[120]

[117] Para mais exemplos, confira a subseção 7.6.2.

[118] Que fique claro:

a multiplicidade algébrica de λ_i é a dimensão do autoespaço generalizado de T associado a λ_i, $i = 1, \ldots, r$.

[119] Que fique claro:

a multiplicidade geométrica de λ_i é a dimensão do autoespaço de T associado a λ_i, $i = 1, \ldots, r$.

[120] Verifique!

DEFINIÇÃO

Para um espaço vetorial $\mathcal{V}$ sobre $\mathbb{C}$, caso m_i seja a multiplicidade algébrica do autovalor λ_i de $T \in \mathcal{L}(\mathcal{V})$, $i = 1, \dots, r$, definimos o *polinômio característico* de T por

$$p(t) := \prod_{i=1}^{r} (t - \lambda_i)^{m_i} \,. \tag{7.60}$$

OBSERVAÇÃO

Por (7.60) e (7.59),

$$\begin{aligned} \text{grau de } p(t) &= \sum_{i=1}^{r} m_i \\ &= \dim \mathcal{V}. \end{aligned}$$

EXEMPLO.3

Em relação ao exemplo.1, $p(t) = (t - \lambda)^3$.[121]

Teorema de Cayley-Hamilton

$$\boxed{p(t) \text{ é o polinômio característico de } T \Longrightarrow p(T) = 0.}$$

DEMONSTRAÇÃO

Caso λ_i seja um dos r autovalores distintos de T e m_i seja a sua multiplicidade algébrica, ou seja, a dimensão do autoespaço generalizado $\mathcal{U}_{\lambda_i}$,[122]

$$(T - \lambda_i I)\,|_{\mathcal{U}_{\lambda_i}}$$

é nilpotente, pelo exemplo de nilpotência que antecede à demonstração ao lema 2, e, portanto,

$$(T - \lambda_i I)^{m_i}\,|_{\mathcal{U}_{\lambda_i}} = 0, \tag{7.61}$$

pelo item (a) do exercício 12 da seção 7.5. Por outro lado,

$$p(T) := \prod_{i=1}^{r} (T - \lambda_i I)^{m_i} \,,$$

por (7.60). Assim, como

$$p(T)|_{\mathcal{U}_{\lambda_i}} = 0,\text{[123]}$$

$$p(T) = 0,$$

pois $p(T)$ é linear e $\mathcal{V} = \bigoplus_{j=1}^{r} \mathcal{U}_{\lambda_j}$.[124]

[121] Para mais exemplos de polinômios característicos, confira os capítulos 4 e 5 e a subseção 7.6.2.
[122] Cf. (7.56), p. 223.
[123] Cf. (7.61).
[124] Cf. a observação que segue à demonstração do lema 1.

OBSERVAÇÃO

Seja $\mathcal{V}_{\mathbb{K}}$ um espaço vetorial sobre o corpo $\mathbb{K}$. No livro do AXLER, referenciado no capítulo 1, alguns resultados sobre operadores definidos em $\mathcal{V}_{\mathbb{R}}$ são obtidos via resultados sobre operadores definidos em $\mathcal{V}_{\mathbb{C}}$, similares aos estudados nesse capítulo. De fato, para $T \in \mathcal{L}(\mathcal{V}_{\mathbb{R}})$, basta considerarmos a "complexificação" $\mathcal{V}_{\mathbb{C}}$ e $T_{\mathbb{C}}$ de $\mathcal{V}_{\mathbb{R}}$ e T, respectivamente, dadas por:

1. $\mathcal{V}_{\mathbb{C}} := \mathcal{V}_{\mathbb{R}} \times \mathcal{V}_{\mathbb{R}}$;
2. $\mathcal{V}_{\mathbb{C}} \ni (u, v) := u + iv$;
3. $u_1, u_2, v_1, v_2 \in \mathcal{V}_{\mathbb{R}} \Longrightarrow (u_1 + iv_1) + (u_2 + iv_2) := (u_1 + u_2) + i\,(v_1 + v_2)$;
4. $a, b \in \mathbb{C}$, $u, v \in \mathcal{V}_{\mathbb{R}} \Longrightarrow (a + ib)(u + iv) := (au - bv) + i(av + bu)$;
5. $u, v \in \mathcal{V}_{\mathbb{R}} \Longrightarrow T_{\mathbb{C}}(u + iv) := T(u) + iT(v)$.

Com essa complexificação, em particular, o polinômio característico de T pode ser definido (sem determinantes), o teorema de Cayley-Hamilton (para T) pode ser demonstrado e podemos provar que a definição usual de polinômio característico (via determinantes), estabelecida na maioria dos livros de álgebra linear, é equivalente a definição dada nesse capítulo.

7.5 Exercícios

1. Se $L \in \mathcal{L}(\mathcal{V})$, $\mathcal{B} = \{v_1, v_2, v_3\}$ é uma base de $\mathcal{V}$ e

$$\begin{aligned} A &= \mathcal{M}(L, \mathcal{B}) \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \end{aligned}$$

obtenha os autoespaços de L associados aos seus autovalores.

RESOLUÇÃO

Como A é diagonal, $\lambda_1 = 1$ e $\lambda_2 = 2$ são os autovalores de L, por (R26).[125] Além disso,

$$L(v_1) = v_1,\ L(v_2) = 2v_2 \text{ e } L(v_3) = 2v_3,$$

por A. Portanto, pela subseção 7.2.1,

$$\mathrm{Nu}(L - I) = [v_1] \ \text{ e } \mathrm{Nu}(L - 2I) = [v_2, v_3]$$

são os autoespaços de L associados aos autovalores λ_1 e λ_2, respectivamente.

2. Para $L \in \mathcal{L}(\mathbb{K}^5)$ e $\dim \mathrm{Nu}(L - 8I) = 4$, demonstre a invertibilidade de $L - 2I$ ou $L - 6I$.

RESOLUÇÃO

Demonstraremos por contradição. Assim, caso os operadores $L - 2I$ e $L - 6I$ não sejam invertíveis, 2 e 6 são autovalores de L, por (R19).[126] Como $\mathrm{Nu}(L - 8I) \neq \{0\}$, 8 também é autovalor de L, por (R19). Logo,

$$\dim \mathrm{Nu}(L - 2I) + \dim \mathrm{Nu}(L - 6I) + \dim \mathrm{Nu}(L - 8I) \leq 5,$$

por (R28).[127] Portanto, como

$$\mathrm{Nu}(L - 2I) = \{0\} \text{ ou } \mathrm{Nu}(L - 6I) = \{0\},$$

um dos operadores supracitados é invertível, por (R19). Contudo, esses operadores foram considerados não invertíveis no início da demonstração.

3. Seja $\mathcal{U}$ um subespaço finitamente gerado do espaço vetorial $\mathcal{V}$, munido do produto interno $\langle\ ,\ \rangle$. Demonstre que

$$\left(\mathcal{U}^{\perp}\right)^{\perp} = \mathcal{U}.$$

RESOLUÇÃO

Para a inclusão

$$\mathcal{U} \subset \left(\mathcal{U}^{\perp}\right)^{\perp},$$

note que,[128] para cada $v \in \mathcal{U}^{\perp}$,

$$\begin{aligned} u \in \mathcal{U} &\Longrightarrow \langle v, u \rangle = 0 \\ &\Longrightarrow \langle u, v \rangle = 0 \\ &\Longrightarrow u \in \left(\mathcal{U}^{\perp}\right)^{\perp}, \end{aligned}$$

[125] Cf. p. 187.
[126] Cf. p. 180.
[127] Cf. p. 188.
[128] Cf. (7.16), p. 190.

pois

$$\left(\mathcal{U}^{\perp}\right)^{\perp} = \left\{w \in \mathcal{V} : \langle w, v\rangle = 0 \text{ para cada } v \in \mathcal{U}^{\perp}\right\}.$$

Para a inclusão

$$\left(\mathcal{U}^{\perp}\right)^{\perp} \subset \mathcal{U},$$

considere $w \in \left(U^{\perp}\right)^{\perp}$. Assim, como $\mathcal{U}$ admite alguma base ortonormal,[129] existe um (único) par de vetores $(u, v) \in \mathcal{U} \times \mathcal{U}^{\perp}$ tal que

$$w = u + v,$$

por (R32).[130] Portanto, como

$$\begin{aligned} 0 &= \langle w, v\rangle \\ &= \langle u + v, v\rangle \\ &= \langle u, v\rangle + \langle v, v\rangle \\ &= 0 + \langle v, v\rangle \\ &= \langle v, v\rangle, \end{aligned}$$

$v = 0_{\mathcal{V}}$ e, assim, $w = u \in \mathcal{U}$.

4. Seja $\mathcal{U}$ um subespaço finitamente gerado do espaço $\mathcal{V}$, dotado do produto interno $\langle\ ,\ \rangle$. Demonstre que:

(a) $u \in \mathcal{U} \Longrightarrow P_{\mathcal{U}}(u) = u$;

(b) $w \in \mathcal{U}^{\perp} \Longrightarrow P_{\mathcal{U}}(w) = 0$;

(c) $\mathrm{Im}(P_{\mathcal{U}}) = \mathcal{U}$;

(d) $\mathrm{Nu}(P_{\mathcal{U}}) = \mathcal{U}^{\perp}$;

(e) $\mathcal{U}^{\perp}$ é finitamente gerado $\Longrightarrow P_{\mathcal{U}^{\perp}} = I - P_{\mathcal{U}}$;[131]

(f) $P_{\mathcal{U}}$ é idempotente, ou seja, $(P_{\mathcal{U}})^2 = P_{\mathcal{U}}$;[132]

(g) $v \in \mathcal{U} \Longrightarrow \|P_{\mathcal{U}}(v)\| \leq \|v\|$.

RESOLUÇÃO DE 4.(a)

$$\begin{aligned} u \in \mathcal{U} &\Longrightarrow \mathcal{V} \ni u = u + 0_{\mathcal{V}}, \text{ com } u \in \mathcal{U} \text{ e } 0_{\mathcal{V}} \in \mathcal{U}^{\perp} \\ &\Longrightarrow P_{\mathcal{U}}(u) = u, \text{ pela definição de } P_{\mathcal{U}}. \end{aligned}$$

RESOLUÇÃO DE 4.(b)

$$\begin{aligned} w \in \mathcal{U}^{\perp} &\Longrightarrow \mathcal{V} \ni w = 0_{\mathcal{V}} + w, \text{ com } 0_{\mathcal{V}} \in \mathcal{U} \text{ e } w \in \mathcal{U}^{\perp} \\ &\Longrightarrow P_{\mathcal{U}}(w) = 0_{\mathcal{V}}, \text{ pela definição de } P_{\mathcal{U}}. \end{aligned}$$

---

[129] De fato, pelo exercício 4 da seção 6.4, podemos aplicar Gram-Schmidt a uma base qualquer de $\mathcal{U}$.

[130] Cf. p. 192.

[131] Cf. as matrizes das projeções $P_{\mathcal{S}}$ e $P_{\mathcal{S}^{\perp}}$ da subseção 4.3.1.

[132] Pelo item (e) desse exercício, temos

$$\begin{aligned} P_{\mathcal{U}} P_{\mathcal{U}^{\perp}} &= P_{\mathcal{U}}\left(I - P_{\mathcal{U}}\right) \\ &= 0. \end{aligned}$$

RESOLUÇÃO DE 4.(c)

Por um lado,

$$\mathrm{Im}(P_{\mathcal{U}}) \subset \mathcal{U},$$

pela definição de $P_{\mathcal{U}}$. Por outro,

$$U \subset \mathrm{Im}(P_{\mathcal{U}}),$$

pelo item (a) desse exercício.

RESOLUÇÃO DE 4.(d)

Por um lado,

$$\mathcal{U}^{\perp} \subset \mathrm{Nu}(P_{\mathcal{U}}),$$

pelo item (b) desse exercício. Por outro, se $v \in \mathcal{V}$ e $P_{\mathcal{U}}(v) = 0_{\mathcal{V}}$, como $v = 0_{\mathcal{V}} + v$ e $0_{\mathcal{V}} \in \mathcal{U}$, então $v \in \mathcal{U}^{\perp}$, pela definição de $P_{\mathcal{U}}$. Portanto,

$$\mathrm{Nu}(P_{\mathcal{U}}) \subset \mathcal{U}^{\perp}.$$

RESOLUÇÃO DE 4.(e)

Basta observarmos que, para $v \in \mathcal{V}$ arbitrário,

$$v = (v - P_{\mathcal{U}}(v)) + P_{\mathcal{U}}(v) \overset{P_{\mathcal{U}^{\perp}}}{\longmapsto} v - P_{\mathcal{U}}(v)$$

define a projeção ortogonal de $\mathcal{V}$ sobre $\mathcal{U}^{\perp}$, onde $v - P_{\mathcal{U}}(v) \in \mathcal{U}^{\perp}$, pelo item 2 de (R34),[133] e $P_{\mathcal{U}}(v) \in \mathcal{U} = \left(\mathcal{U}^{\perp}\right)^{\perp}$, pelo exercício 3 dessa seção.[134]

RESOLUÇÃO DE 4.(f)

Caso u e v sejam como na demonstração de (R32),[135]

$$\begin{aligned} P_{\mathcal{U}}(P_{\mathcal{U}}(v)) &= P_{\mathcal{U}}(u) \text{ (pois } P_{\mathcal{U}}(v) = u, \text{ pela definição de } P_{\mathcal{U}}) \\ &= u \qquad \text{(pelo item (a) desse exercício)} \\ &= P_{\mathcal{U}}(v) \text{ (pois } P_{\mathcal{U}}(v) = u, \text{ pela definição de } P_{\mathcal{U}}). \end{aligned}$$

RESOLUÇÃO DE 4.(g)

Sejam u, v e w como na demonstração de (R32). Logo, pelo teorema de Pitágoras,[136]

$$\begin{aligned} \|v\|^2 &= \|u\|^2 + \|w\|^2 \\ &\geq \|u\|^2 = \|P_{\mathcal{U}}(v)\|^2, \end{aligned}$$

pois $P_{\mathcal{U}}(v) = u$, pela definição de $P_{\mathcal{U}}$.

5. Demonstre os seis itens (não demonstrados) da subseção 7.4.3.

RESOLUÇÃO

1. Considere $w \in \mathcal{V}$ fixo e $v \in \mathcal{V}$ arbitrário. Logo,

$$\begin{aligned} \langle v, (T+L)^*(w) \rangle &= \langle (T+L)(v), w \rangle \\ &= \langle T(v), w \rangle + \langle L(v), w \rangle \\ &= \langle v, T^*(w) \rangle + \langle v, L^*(w) \rangle \\ &= \langle v, (T^* + L^*)(w) \rangle. \end{aligned}$$

Então, $(T+L)^*(w) = (T^* + L^*)(w)$, por (R35).[137]

[133] Cf. p. 194.
[134] Seção 7.5.
[135] Cf. p. 192.
[136] Cf. a demonstração do item 4 de (R34).
[137] Cf. p. 205.

2. Sejam $\lambda \in \mathbb{K}$ e $w \in \mathcal{V}$ fixos. Logo, para $v \in \mathcal{V}$ arbitrário,

$$\begin{aligned}\langle v, (\lambda T)^*(w)\rangle &= \langle (\lambda T)(v), w\rangle \\ &= \lambda \langle T(v), w\rangle \\ &= \lambda \langle v, T^*(w)\rangle \\ &= \langle v, (\overline{\lambda} T^*)(w)\rangle.\end{aligned}$$

Então, $(\lambda T)^*(w) = \left(\overline{\lambda} T^*\right)(w)$, por (R35).

4. Considere $w \in \mathcal{V}$ fixo e $v \in \mathcal{V}$ arbitrário. Logo,

$$\begin{aligned}\langle v, I^*(w)\rangle &= \langle I(v), w\rangle \\ &= \langle v, w\rangle \\ &= \langle v, I(w)\rangle.\end{aligned}$$

Então, $I^*(w) = I(w)$, por (R35).

5. Considere $w \in \mathcal{V}$ fixo e $v \in \mathcal{V}$ arbitrário. Logo,

$$\begin{aligned}\langle v, (TL)^*(w)\rangle &= \langle (TL)(v), w\rangle \\ &= \langle T(L(v)), w\rangle \\ &= \langle L(v), T^*(w)\rangle \\ &= \langle v, L^*\left(T^*(w)\right)\rangle \\ &= \langle v, (L^* T^*)(w)\rangle.\end{aligned}$$

Então, $(TL)^*(w) = (L^*T^*)(w)$, por (R35).

7. Pelos itens 4 e 5 de 7.4.3,[138] se

$$\begin{aligned}TT^{-1} &= I \\ &= T^{-1}T,\end{aligned}$$

então

$$\begin{aligned}\left(T^{-1}\right)^* T^* &= I \\ &= T^* \left(T^{-1}\right)^*.\end{aligned}$$

8. Por (R19) e pelos itens 1, 2, 3 e 7 de 7.4.3,[139]

$$\begin{aligned}\lambda \text{ é autovalor de } T &\iff T - \lambda I \text{ não é invertível} \\ &\iff (T - \lambda I)^* = T^* - \overline{\lambda} I \text{ não é invertível} \\ &\iff \overline{\lambda} \text{ é autovalor de } T^*.\end{aligned}$$

6. Para $L \in \mathcal{L}(\mathcal{V})$, onde $\mathcal{V}$ é finitamente gerado e dotado do produto interno $\langle\ ,\ \rangle$,

$$\mathrm{Nu}(L), \mathrm{Im}(L), \mathrm{Nu}(L^*) \text{ e } \mathrm{Im}(L^*)$$

são conhecidos como os *quatro subespaços fundamentais associados a L*.[140] Demonstre que:

(a) $\mathrm{Nu}(L^*) = (\mathrm{Im}(L))^\perp$;

[138] Cf. p. 207.

[139] Cf. pp. 180 e 207.

[140] Cf. o livro do STRANG, referenciado no capítulo 1.

(b) $\mathrm{Nu}(L) = (\mathrm{Im}(L^*))^{\perp}$;

(c) $\mathrm{Im}(L^*) = (\mathrm{Nu}(L))^{\perp}$;

(d) $\mathrm{Im}(L) = (\mathrm{Nu}(L^*))^{\perp}$.

RESOLUÇÃO

Para o item (a), note que

$$\begin{aligned} w \in \mathrm{Nu}(L^*) &\iff L^*(w) = 0_{\mathcal{V}} \\ &\iff \langle v, L^*(w)\rangle = 0 \;\; \forall v \in \mathcal{V} \\ &\iff \langle L(v), w\rangle = 0 \;\; \forall v \in \mathcal{V} \\ &\iff \langle w, L(v)\rangle = 0 \;\; \forall v \in \mathcal{V} \\ &\iff \langle w, u\rangle = 0 \;\; \forall u \in \mathrm{Im}(L) \\ &\iff w \in (\mathrm{Im}(L))^{\perp}. \end{aligned}$$

Para o item (b), temos

$$\begin{aligned} \mathrm{Nu}(L) &= \mathrm{Nu}\big((L^*)^*\big) \\ &= (\mathrm{Im}(L^*))^{\perp}, \end{aligned}$$

pelo item 3 de 7.4.3, p. 207, e pelo item (a) desse exercício.
O item (c) segue do exercício 3 dessa subseção (7.5) e do item (b) desse exercício.
Para o item (d), temos

$$\begin{aligned} \mathrm{Im}(L) &= \mathrm{Im}\big((L^*)^*\big) \\ &= (\mathrm{Nu}(L^*))^{\perp}, \end{aligned}$$

pelo item 3 de 7.4.3 e pelo item (c) desse exercício.

7. Considere que a matriz de $T \in \mathcal{L}\big(\mathbb{C}^2\big)$ na base canônica é dada por

$$\begin{bmatrix} 1 & \alpha \\ 1+i & 2 \end{bmatrix}.$$

Determine todos os valores de α para os quais T é autoadjunto.

RESOLUÇÃO

Caso A e A^* sejam como no item 6 de 7.4.3,

$$\begin{aligned} T \text{ é autoadjunto} &\iff T = T^* \\ &\iff A = A^*, \end{aligned}$$

pelo isomorfismo (6.9) de 6.3.3, com $\mathcal{V} = \mathcal{W}$ e $\mathcal{B} = \mathcal{B}'$. Portanto, $\alpha = 1 - i$.

8. Caso $T \in \mathcal{L}(\mathcal{V})$ seja hermiteano e $\mathcal{U}$ seja um subespaço de $\mathcal{V}$ invariante por T, isto é, $T|_{\mathcal{U}} \in \mathcal{L}(\mathcal{U})$, demonstre que:

(a) $\mathcal{U}^{\perp}$ é invariante por T, ou seja, $T|_{\mathcal{U}^{\perp}} \in \mathcal{L}\big(\mathcal{U}^{\perp}\big)$;

(b) $T|_{\mathcal{U}}$ e $T|_{\mathcal{U}^{\perp}}$ são hermiteanos.

RESOLUÇÃO

(a) Para $v \in \mathcal{U}^{\perp}$ e $u \in \mathcal{U}$ arbitrários,

$$\begin{aligned} \langle T(v), u\rangle &= \langle v, T(u)\rangle \;\; (T \text{ é autoadjunto}) \\ &= 0 \qquad\qquad (v \in \mathcal{U}^{\perp} \text{ e } T(u) \in \mathcal{U}) \end{aligned}$$

acarreta $T(v) \in \mathcal{U}^{\perp}$.
(b) Para $u, w \in \mathcal{U}$ arbitrários,

$$\begin{aligned}\big\langle (T|_{\mathcal{U}})(u), w\big\rangle &= \langle T(u), w\rangle \\ &= \langle u, T(w)\rangle \qquad (T \text{ é autoadjunto}) \\ &= \big\langle u, (T|_{\mathcal{U}})(w)\big\rangle\end{aligned}$$

acarreta $T|_{\mathcal{U}}$ autoadjunto. Portanto, ao considerarmos $\mathcal{U}^{\perp}$ no lugar de $\mathcal{U}$ e o item (a) desse exercício, $T|_{\mathcal{U}^{\perp}}$ também é autoadjunto.

9. O teorema espectral complexo (subseção 7.4.6) caracteriza todo operador $T \in \mathcal{L}(\mathcal{V})$ normal, caso $\mathcal{V}$ seja finitamente gerado e munido de algum produto interno complexo. Utilize o teorema supracitado para demonstrar que, caso T seja normal e λ seja um dos seus autovalores,

$$T \text{ é autoadjunto } \iff \lambda \in \mathbb{R}.$$

OBSERVAÇÃO

Na demonstração da implicação "$\Longrightarrow$", não utilize o resultado (R36).[141]

RESOLUÇÃO

Para a implicação "$\Longrightarrow$", suponha

$$T^* = T$$

e que λ seja um autovalor de T. Logo,

$$\begin{aligned}\mathrm{Nu}(T - \lambda I) &= \mathrm{Nu}\big(T^* - \overline{\lambda} I\big) \\ &= \mathrm{Nu}\big(T - \overline{\lambda} I\big)\,,\end{aligned} \tag{7.62}$$

pelos itens 3 e 4 de (R38),[142] e, assim, $\overline{\lambda}$ também é autovalor de T. Além disso, como T é normal,[143] T é diagonalizável, por 7.4.6.[144] Portanto, os autovalores λ e $\overline{\lambda}$ não podem ser distintos, por (7.62) e pelo item 2 de (R29).[145] Assim, como $\lambda = \overline{\lambda}$,

$$\lambda \in \mathbb{R}.$$

Para a implicação "$\Longleftarrow$", suponha que

$$\lambda \text{ é autovalor de } T \Longrightarrow \lambda \in \mathbb{R}. \tag{7.63}$$

Logo, por 7.4.6,[146] existe alguma base ortonormal $\mathcal{B}$ de $\mathcal{V}$ tal que

$$D = \mathcal{M}(T, \mathcal{B})$$

é diagonal, onde todos os seus autovalores são reais, por (7.63), e são representados pelas entradas da diagonal principal de D. Portanto,

$$\begin{aligned}\mathcal{M}(T^*, \mathcal{B}) &= D^* \\ &= D,\end{aligned}$$

pelo item 6 de 7.4.3.[147] Assim,

$$T^* = T,$$

pelo isomorfismo (6.9).[148]

[141] Cf. p. 209.
[142] Cf. p. 211.
[143] Por hipótese ou por ser autoadjunto!
[144] Cf. p. 212.
[145] Cf. p. 189.
[146] T é normal, por hipótese.
[147] Cf. p. 207.
[148] Cf. p. 158.

10. Se $T \in \mathcal{L}(\mathcal{V})$, $v \in \mathcal{V}$, $T^{m-1}(v) \neq 0$ e $T^m(v) = 0$, demonstre a independência linear dos vetores da lista
$$v, T(v), T^2(v), \dots, T^{m-1}(v).$$

RESOLUÇÃO

Note que:

- $v \neq 0$, pois, caso contrário, para todo inteiro não negativo j, $T^j(v) = 0$ e, portanto, chegaríamos a uma contradição da hipótese
$$T^{m-1}(v) \neq 0; \tag{7.64}$$
- $T^m(v) = 0$ acarreta
$$T^{m+j}(v) = 0,\ j = 0, 1, 2, \dots \tag{7.65}$$

Assim, para
$$\sum_{j=0}^{m-1} a_{j+1} T^j(v) = 0, \tag{7.66}$$
onde $a_{j+1} \in \mathbb{K}$, $j = 0, 1, \dots, m-1$, temos
$$\begin{aligned} T^{m-1}\left(\sum_{j=0}^{m-1} a_{j+1} T^j v\right) = T^{m-1}(0) &\Longrightarrow \sum_{j=0}^{m-1} a_{j+1} T^{m+j-1}(v) = 0 \\ &\Longrightarrow a_1 T^{m-1}(v) = 0 \\ &\Longrightarrow a_1 = 0, \end{aligned}$$
por (7.65) e (7.64). Portanto, (7.66) pode ser escrita como
$$\sum_{j=1}^{m-1} a_{j+1} T^j(v) = 0.$$
Assim,
$$\begin{aligned} T^{m-2}\left(\sum_{j=1}^{m-1} a_{j+1} T^j v\right) = T^{m-2}(0) &\Longrightarrow \sum_{j=1}^{m-1} a_{j+1} T^{m+j-2}(v) = 0 \\ &\Longrightarrow a_2 T^{m-1}(v) = 0 \\ &\Longrightarrow a_2 = 0, \end{aligned}$$
por (7.65) e (7.64). Obviamente, ao repetirmos o mesmo raciocínio $m-2$ vezes, temos
$$a_3 = \cdots = a_m = 0.$$

11. Para um espaço vetorial complexo finitamente gerado $\mathcal{V}$, o autoespaço generalizado de $T \in \mathcal{L}(\mathcal{V})$ associado a um de seus autovalores contém todos os autovetores associados ao autovalor supracitado,[149] pela informação adicional da parte 1 da demonstração do lema 1.[150]

    (a) Caso T tenha autovalores distintos, demonstre que nenhum autovetor de T associado a um deles pertence ao autoespaço generalizado de T associado a outro desses autovalores.

    (b) Demonstre que a interseção de autoespaços generalizados de T associados a autovalores distintos contém apenas o vetor nulo.

[149] A existência desse autovalor segue de (R23) e (R25), pp. 184 e 187.

[150] O lema 1 encontra-se na página 217.

RESOLUÇÃO

Sejam α e β autovalores distintos de T.
(a) Para demonstrarmos que

$$\operatorname{Nu}(T-\alpha I)\cap\operatorname{Nu}\left((T-\beta I)^{\dim\mathcal{V}}\right)=\{0\}\,, \tag{7.67}$$

consideremos v pertencente à interseção do primeiro membro de (7.67). Logo,

$$\begin{aligned}0&=(T-\beta I)^{\dim\mathcal{V}}(v)\\&=(T-\beta I)^{\dim\mathcal{V}-1}(T-\beta I)(v)\\&=(\alpha-\beta)(T-\beta I)^{\dim\mathcal{V}-1}(v).\end{aligned}$$

Ao repetirmos esse procedimento mais $\dim\mathcal{V}-1$ vezes, temos

$$0=(\alpha-\beta)^{\dim V}v.$$

Assim, $v=0$, pois $\alpha\neq\beta$. Portanto, (7.67) é válida.
(b) Caso

$$\mathcal{W}=\operatorname{Nu}\left((T-\alpha I)^{\dim\mathcal{V}}\right)\cap\operatorname{Nu}\left((T-\beta I)^{\dim\mathcal{V}}\right)$$

seja não nulo, como $T|_{\mathcal{W}}$ possui algum autovalor,[151] $\operatorname{Nu}\left((T-\alpha I)^{\dim\mathcal{V}}\right)$ e $\operatorname{Nu}\left((T-\beta I)^{\dim\mathcal{V}}\right)$ contem autovetores de T associados ao autovalor supracitado, que é diferente de α ou β, pois $\alpha\neq\beta$. Portanto, como temos uma contradição do item (a) desse exercício, $\mathcal{W}=\{0\}$.

12. Seja $L\in\mathcal{L}(\mathcal{V})$ nilpotente.[152] Demonstre que:
   (a) $L^{\dim\mathcal{V}}=0$;
   (b) 0 é o único autovalor de L;
   (c) $\mathbb{K}=\mathbb{C}\Longrightarrow$ existe alguma base $\mathcal{B}$ de $\mathcal{V}$ tal que

$$\mathcal{M}(L,\mathcal{B})=\begin{bmatrix}0 & & *\\ & \ddots & \\ \text{zeros} & & 0\end{bmatrix}$$

   é triangular superior. [153]

RESOLUÇÕES

(a) Existe algum inteiro positivo p tal que

$$\begin{aligned}(L-0I)^p&=L^p\\&=0.\end{aligned}$$

Então, como, para $\lambda=0$,

$$\operatorname{Nu}\left((L-\lambda I)^p\right)=\mathcal{V}$$

é um dos subespaços da sequência (7.42) da demonstração do lema 1,[154]

$$\operatorname{Nu}\left(L^{\dim\mathcal{V}}\right)=\mathcal{V},$$

[151] Cf. (R23) e (R25).
[152] Cf. lema 2, p. 221.
[153] Esse resultado também é válido para $\mathbb{K}=\mathbb{R}$, conforme demonstrado no livro do AXLER, citado no capítulo 1.
[154] Cf. p. 218.

por (7.43), (7.44) e I (da informação adicional daquela demonstração).
(b) Seja $\lambda \in \mathbb{K}$, $0 \neq v \in \mathcal{V}$ e $L^p = 0$ para algum inteiro positivo p. Portanto,

$$\begin{aligned} L(v) = \lambda v &\Longrightarrow L^p(v) = \lambda^p v \\ &\Longrightarrow \lambda^p v = 0_{\mathcal{V}} \\ &\Longrightarrow \lambda = 0. \end{aligned}$$

(c) Esse resultado é uma consequência direta de (R23), (R25) e do item (b) desse exercício.

13. Prove (7.42) e (7.44) da demonstração do lema 1.[155]

RESOLUÇÃO

Para cada inteiro positivo i,

$$\mathcal{U}_i \subset \mathcal{U}_{i+1},$$

pois

$$\begin{aligned} v \in \mathcal{U}_i &\Longrightarrow v \in \mathrm{Nu}\Big((T-\lambda I)^i\Big) \\ &\Longrightarrow (T-\lambda I)^i(v) = 0_{\mathcal{V}} \\ &\Longrightarrow (T-\lambda I)\Big((T-\lambda I)^i(v)\Big) = (T-\lambda I)(0_{\mathcal{V}}) \\ &\Longrightarrow (T-\lambda I)^{i+1}(v) = 0_{\mathcal{V}} \\ &\Longrightarrow v \in \mathrm{Nu}\Big((T-\lambda I)^{i+1}\Big) \\ &\Longrightarrow v \in \mathcal{U}_{i+1}. \end{aligned}$$

Portanto, como a sequência de inclusões (7.42) é, de fato, crescente, em relação a (7.44), basta demonstrarmos a validade da seguinte sequência de inclusões reversas:

$$\mathcal{U}_{\ell+1} \supset \mathcal{U}_{\ell+2} \supset \mathcal{U}_{\ell+3} \supset \dots$$

Na verdade, basta demonstrarmos a primeira inclusão dessa sequência, pois as outras são análogas. Assim,

$$\begin{aligned} v \in \mathcal{U}_{\ell+2} &\Longrightarrow v \in \mathrm{Nu}\Big((T-\lambda I)^{\ell+2}\Big) \\ &\Longrightarrow (T-\lambda I)^{\ell+2}(v) = 0_{\mathcal{V}} \\ &\Longrightarrow (T-\lambda I)^{\ell+1}((T-\lambda I)(v)) = 0_{\mathcal{V}} \\ &\Longrightarrow (T-\lambda I)(v) \in \mathrm{Nu}\Big((T-\lambda I)^{\ell+1}\Big) \\ &\Longrightarrow (T-\lambda I)(v) \in \mathcal{U}_{\ell+1} \\ &\Longrightarrow (T-\lambda I)(v) \in \mathcal{U}_{\ell} \text{ (via (7.43))} \\ &\Longrightarrow (T-\lambda I)(v) \in \mathrm{Nu}\Big((T-\lambda I)^{\ell}\Big) \\ &\Longrightarrow (T-\lambda I)^{\ell}((T-\lambda I)(v)) = 0_{\mathcal{V}} \\ &\Longrightarrow (T-\lambda I)^{\ell+1}(v) = 0_{\mathcal{V}} \\ &\Longrightarrow v \in \mathrm{Nu}\Big((T-\lambda I)^{\ell+1}\Big) \\ &\Longrightarrow v \in \mathcal{U}_{\ell+1}. \end{aligned}$$

14. Prove a equivalência (7.48) da demonstração do lema 2.

RESOLUÇÃO

[155] Cf. p. 218.

Como $a \in \mathbb{K}$ é uma "matriz" 1×1 e toda transformação linear é do tipo "multiplicação por matriz",[156] se $v \in \mathcal{V}$ é arbitrário, então $L(v) = av$ e, caso p seja um inteiro positivo, como $L^p(v) = a^p v$,

$$\begin{aligned} L^p(v) = 0_{\mathcal{V}}\ \forall v \in \mathcal{V} &\iff a^p v = 0_{\mathcal{V}}\ \forall v \in \mathcal{V} \\ &\iff a^p = 0 \\ &\iff a = 0 \\ &\iff av = 0_{\mathcal{V}}\ \forall v \in \mathcal{V} \\ &\iff L(v) = 0_{\mathcal{V}}\ \forall v \in \mathcal{V}. \end{aligned}$$

15. Prove que (7.55), da demonstração do lema 2, é base de $\mathcal{V}$.

RESOLUÇÃO

Por (R9),[157] basta demonstrarmos os seguintes fatos:

I. Os vetores de (7.55) são LI;

II. $\dim \mathcal{V}$ é igual ao número de vetores de (7.55).

DEMONSTRAÇÃO DO FATO I

Ao aplicarmos L a ambos os membros da CL nula

$$\sum_{i=1}^{r} \left(\sum_{j=0}^{q_i} a_{i,j} L^j(v_i) \right) + \sum_{k=1}^{q} b_k w_k = 0_{\mathcal{V}},^{158} \tag{7.68}$$

podemos reduzi-la a

$$\sum_{i=1}^{r} \left(\sum_{j=0}^{q_i-1} a_{i,j} L^{j+1}(v_i) \right) = 0_{\mathcal{V}},^{159}$$

isto é,[160]

$$\sum_{i=1}^{r} \left(\sum_{j=0}^{q_i-1} a_{i,j} L^{j}(u_i) \right) = 0_{\mathcal{V}},$$

que é uma CL nula dos vetores da base (7.52) de $\mathrm{Im}(T)$,[161] ou seja,

$$a_{1,0} = a_{1,1} = \ldots = a_{1,q_1-1} = \ldots = a_{r,0} = a_{r,1} = \cdots = a_{r,q_r-1} = 0. \tag{7.69}$$

Portanto, ao substituirmos (7.69) em (7.68), obtemos a CL nula

$$\sum_{i=1}^{r} a_{i,q_i} L^{q_i}(v_i) + \sum_{k=1}^{q} b_k w_k = 0_{\mathcal{V}}. \tag{7.70}$$

Logo, como (7.54) é base de $\mathrm{Nu}(L)$,[162] os escalares de (7.70) são nulos e, portanto, todos os escalares de (7.68) também são nulos.

DEMONSTRAÇÃO DO FATO II

$$\begin{aligned} \dim \mathcal{V} &= \dim(\mathrm{Im}(L)) + \dim(\mathrm{Nu}(L)) && \text{(por (R14) do final da subseção 6.3.2)} \\ &= (q_1 + \cdots + q_r) + (r + q) && \text{(por (7.52) e (7.54) da demonstração do lema 2)} \\ &= (q_1 + 1) + \cdots + (q_r + 1) + q, \end{aligned}$$

que é exatamente o número de vetores de (7.55).

---

[156] Cf. (R15), p. 159.
[157] Cf. p. 152.
[158] Para todos os índices i, j, k admissíveis, $a_{i,j}, b_k \in \mathbb{C}$.
[159] De fato, os vetores de (7.54) pertencem a $\mathrm{Nu}(L)$, conforme a p. 222.
[160] Cf. (7.53), p. 222.
[161] Cf. p. 222.
[162] Idem.

7.6 Informações adicionais

7.6.1 Matrizes quadradas × operadores

Caso $\mathcal{B} = \{v_1, \ldots, v_n\}$ e $\mathcal{B}' = \{w_1, \ldots, w_n\}$ sejam bases ordenadas de $\mathcal{V}$ e I seja o operador identidade em $\mathcal{V}$,[163]

$$\begin{aligned}\mathcal{M}(I,\mathcal{B}',\mathcal{B})\,\mathcal{M}(I,\mathcal{B},\mathcal{B}') &= \mathcal{M}(I,\mathcal{B}) \\ &= \begin{bmatrix} 1 & & \text{zeros} \\ & \ddots & \\ \text{zeros} & & 1 \end{bmatrix},\end{aligned}$$

por (6.10).[164] Assim, $\mathsf{P} := \mathcal{M}(I,\mathcal{B}',\mathcal{B})$ e $\mathsf{P}' := \mathcal{M}(I,\mathcal{B},\mathcal{B}')$ são invertíveis, com $\mathsf{P}' = \mathsf{P}^{-1}$, e ditas *matrizes de mudança de bases*, pois, para cada $L \in \mathcal{L}(\mathcal{V})$,

$$\mathcal{M}(L,\mathcal{B}') = \mathsf{P}^{-1}\mathcal{M}(L,\mathcal{B})\mathsf{P}, \tag{7.71}$$

por (6.10).[165]

EXEMPLO

Podemos conciliar o conceito de diagonalização de matrizes com o de diagonalização de operadores. De fato, qualquer matriz $A \in \mathbb{R}^{n\times n}$ pode ser vista como a representação do operador "multiplicação por A" em $\mathbb{R}^n$, ou seja, se $\mathcal{B}$ é a base canônica de $\mathbb{R}^n$ e, para cada $\mathbf{x} \in \mathbb{R}^n$, $L(\mathbf{x}) = A\mathbf{x}$, então $A = \mathcal{M}(L,\mathcal{B})$. Caso o operador L seja diagonalizável, existe uma base $\mathcal{B}'$ de $\mathbb{R}^n$, formada por autovetores de L, tal que $A' = \mathcal{M}(L,\mathcal{B}')$ é diagonal, pela subseção 7.2.5. Portanto, A é diagonalizável (num sentido semelhante ao da subseção 4.2.1), pela equação (7.71).[166]

OBSERVAÇÃO

O operador (4.9) (da página 97) não é o operador identidade. Caso fosse, $[T]_{\mathcal{B}}$ e $\left[T^{-1}\right]_{\mathcal{B}'}$ seriam a matriz identidade em $\mathbb{R}^{n\times n}$ e, portanto, o uso de (4.11) (da página 98) estaria comprometido.[167] Contudo, T e o operador identidade I, no que se refere a matriz P desses operadores, desempenham um papel semelhante nas fórmulas (4.11) e (7.71), pois:

Caso $\mathcal{B}$ e $\mathcal{B}'$ sejam as bases ordenadas supracitadas e o operador $T \in \mathcal{L}(\mathcal{V})$ satisfaça a condição

$$T(v_k) = w_k,\, k = 1, \ldots, n,$$

$$\mathcal{M}(T,\mathcal{B}) = \mathcal{M}(I,\mathcal{B}',\mathcal{B})\,.$$

[163] Para cada $v \in \mathcal{V}$, $I(v) = v$.

[164] Cf. p. 159.

[165] Matrizes que representam L (em bases distintas) são ditas *semelhantes (entre si)*.

[166] Essa abordagem também é válida para a forma de Jordan de uma matriz quadrada A, a ser estudada na subseção 7.6.2.

[167] Além disso, exemplos (nas subseções 4.1.2 e 4.2.1) comprovam que essas matrizes não podem ser a identidade.

De fato,

$$\begin{aligned}\mathcal{M}(T,\mathcal{B}) &= \mathcal{M}(IT,\mathcal{B},\mathcal{B})\\ &= \mathcal{M}(I,\mathcal{B}',\mathcal{B})\,\mathcal{M}(T,\mathcal{B},\mathcal{B}') \text{ (por (6.10))}\\ &= \mathcal{M}(I,\mathcal{B}',\mathcal{B})\,\mathsf{I},\end{aligned}$$

pela condição supracitada, onde I denota a matriz identidade de tamanho compatível.

7.6.2 Para $A \in \mathbb{C}^{n\times n}$ não diagonalizável, como podemos obter $\mathsf{P} \in \mathbb{C}^{n\times n}$ invertível com $$\mathsf{P}^{-1}A\mathsf{P} = J$$ na forma de Jordan?

Para os três exemplos seguintes, considere o estudo feito na subseção 7.4.8 e as seguintes etapas:

I. Cálculo dos autovalores de A;

II. Cálculo dos autoespaços de A;

III. Cálculo de J;

IV. Cálculo da base $\mathcal{B}$ **de** $\mathcal{U}_\lambda$ **consistindo de ciclos**;

V. Cálculo de P.

EXEMPLO 1

Seja

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

I. Como

$$\begin{aligned}\det(A-\lambda\mathsf{I}) &= \begin{vmatrix} 2-\lambda & 2 & 3\\ 1 & 3-\lambda & 3\\ -1 & -2 & -2-\lambda \end{vmatrix}\\ &= -(\lambda-1)^3,\end{aligned}$$

$\lambda = 1$ é o único autovalor de A.

II. Como

$$\begin{aligned}\mathrm{Nu}(A-\mathsf{I}) &= \mathrm{Nu}\begin{pmatrix} 1 & 2 & 3\\ 1 & 2 & 3\\ -1 & -2 & -3 \end{pmatrix}\\ &= \left[\begin{pmatrix} -2\\ 1\\ 0 \end{pmatrix}, \begin{pmatrix} -3\\ 0\\ 1 \end{pmatrix}\right],\end{aligned}$$

a multiplicidade geométrica de $\lambda = 1$ é dada por

$$\begin{aligned}\mu &= \dim(\mathrm{Nu}(A - \mathsf{I})) \\ &= 2 \\ &< 3.\end{aligned}$$

Portanto, $\mathbb{C}^3$ não tem base formada por autovetores de A, isto é, A não é diagonalizável.

III. Pelas etapas I e II, temos (apenas) duas possibilidades para J:

1. $J = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, composta de um único bloco de Jordan;
2. $J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, a menos da ordem dos dois blocos.

Na verdade, 1 não é uma possibilidade viável, pelo exemplo que precede a o teorema de Cayley-Hamilton.[168] De fato, como

$$\begin{aligned}(A - \mathsf{I})^2 &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ -1 & -2 & -3 \end{pmatrix}^2 \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\end{aligned}$$

não tem como $\mathbb{C}^3$ ter uma base que consista de um único ciclo de comprimento 3. Assim, devemos obter a forma dada em 2.

IV. Pela etapa III,

$$\mathcal{B} = \{v_1\} \cup \{(A - \mathsf{I})v_2, v_2\},$$

onde v_1 pode ser qualquer vetor de $\mathrm{Nu}(A - \mathsf{I})$. Por exemplo, seja $v_1 = (-2, 1, 0)$. Por outro lado, para obtermos um ciclo de comprimento 2, o autovetor generalizado v_2 deve satisfazer a condição

$$v_2 \in \mathrm{Nu}\left((A - \mathsf{I})^2\right) \text{ mas } v_2 \notin \mathrm{Nu}(A - \mathsf{I}).$$

Logo, como $(A - \mathsf{I})^2 = \mathsf{O}$, podemos escolher qualquer $v_2 \notin \mathrm{Nu}(A - \mathsf{I})$. Por exemplo, considere $v_2 = \mathbf{e}_1$. Logo,

$$\{(A - \mathsf{I})v_2, v_2\} = \left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

[168] Esse teorema encontra-se na página 225.

V. Considere as colunas de P como sendo v_1, $(A-\mathsf{I})v_2$ e v_2, nessa ordem. Verifique, então, que

$$\mathsf{P}^{-1}A\mathsf{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

EXEMPLO 2

Seja

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

I. Por (R25),[169] $\lambda = -1$ é o único autovalor da matriz triangular superior A.

II. Como

$$\begin{aligned} \mathrm{Nu}(A+\mathsf{I}) &= \mathrm{Nu}\begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \left[\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right], \end{aligned}$$

a multiplicidade geométrica de $\lambda = -1$ é dada por

$$\begin{aligned} \mu &= \dim(\mathrm{Nu}(A+\mathsf{I})) \\ &= 1 \\ &< 3. \end{aligned}$$

Portanto, $\mathbb{C}^3$ não tem base formada por autovetores de A, isto é, A não é diagonalizável.

III. Pelas etapas I e II, temos (apenas) duas possibilidades para J:

1. $J = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$, composta de um único bloco de Jordan;
2. $J = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$, a menos da ordem dos dois blocos.

Note que 2 é uma possibilidade viável, caso $\mathbb{C}^3$ tenha alguma base $\mathcal{B}$ consistindo de dois ciclos, um de comprimento 1 e o outro de comprimento 2, digamos

$$\mathcal{B} = \{v_1\} \cup \{(A+\mathsf{I})v_2, v_2\},$$

onde v_1 e $(A+\mathsf{I})v_2$ são autovetores LI associados a $\lambda = -1$. Contudo, isso não pode acontecer, pois $\mu = 1$. Portanto, devemos obter a forma dada em 1.

[169] Cf. p. 187.

IV. $\mathcal{B}$ é composta de apenas um ciclo de comprimento 3, pela etapa III, dado por

$$\mathcal{B} = \left\{(A+\mathsf{I})^2 v, (A+\mathsf{I})v, v\right\},$$

onde o autovetor generalizado v deve satisfazer a condição

$$v \in \mathrm{Nu}\left((A+\mathsf{I})^3\right) \text{ mas } v \notin \mathrm{Nu}(A+\mathsf{I}) \cup \mathrm{Nu}\left((A+\mathsf{I})^2\right).$$

Pelo que obtivemos na etapa II e como

$$\begin{aligned}\mathrm{Nu}\left((A+\mathsf{I})^2\right) &= \mathrm{Nu}\begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \left[\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right],\end{aligned}$$

podemos considerar, por exemplo, $v = \mathbf{e}_3$. Logo,

$$\mathcal{B} = \left\{\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}.$$

V. Considere as colunas de P como sendo os vetores da base ordenada $\mathcal{B}$, na ordem em que esses vetores aparecem nessa base. Então, verifique que

$$\mathsf{P}^{-1} A \mathsf{P} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

$\boxed{\text{EXEMPLO 3}}$

Seja

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

I. Como

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

(onde calculamos os dois primeiros determinantes ao longo da primeira e da última colunas, respectivamente), $\lambda_1 = 2$ e $\lambda_2 = 3$ são os autovalores de A com multiplicidades algébricas dadas respectivamente por $m_1 = 3$ e $m_2 = 1$.

II. Como

$$\begin{aligned}\mathrm{Nu}(A - 2\mathrm{I}) &= \mathrm{Nu}\begin{pmatrix} 0 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} \\ &= \left[\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}\right],\end{aligned}$$

com multiplicidade geométrica de $\lambda_1 = 2$ dada por

$$\begin{aligned}\mu_1 &= \dim(\mathrm{Nu}(A - 2\mathrm{I})) \\ &= 2 \\ &< m_1,\end{aligned}$$

e

$$\begin{aligned}\mathrm{Nu}(A - 3\mathrm{I}) &= \mathrm{Nu}\begin{pmatrix} -1 & -1 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \\ &= \left[\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}\right],\end{aligned}$$

com multiplicidade geométrica de $\lambda_2 = 3$ dada por

$$\begin{aligned}\mu_2 &= \dim(\mathrm{Nu}(A - 2\mathrm{I})) \\ &= 1 \\ &= m_2,\end{aligned}$$

temos que $\mathbb{C}^4$ não tem base formada por autovetores de A, isto é, A não é diagonalizável.

III. Pelas etapas I e II, temos (apenas) duas possibilidades para J:

1. $J = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$, a menos da ordem dos dois blocos, com um dos blocos associado a um ciclo de comprimento 3 e o outro associado a um ciclo de comprimento 1;

2. $J = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$, a menos da ordem dos três blocos, com dois blocos de comprimento 1 e o terceiro de comprimento 2.

Como $\mu_1 = 2$, a base de $\mathcal{U}_2$, composta de ciclos, deve conter dois autovetores LI. Como não conseguimos isso na possibilidade 1, devemos obter a J dada em 2.

IV. Pelas etapas II e III, para

$$v_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

$\{v_3\}$ é uma base de $\mathcal{U}_3$ e uma base de $\mathcal{U}_2$ é da forma

$$\{v_1\} \cup \{(A - 2\mathsf{I})v_2, v_2\},$$

onde v_1 e $(A - 2\mathsf{I})v_2$ são autovetores de A e v_2 é um autovetor generalizado de A tal que $(A - 2\mathsf{I})v_2 \neq \mathbf{0}$, todos associados ao autovalor $\lambda_1 = 2$. Portanto, para obtermos esses ciclos, considere

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

por exemplo, e que v_2 satisfaça a condição

$$v_2 \in \mathrm{Nu}\left((A - 2\mathsf{I})^2\right) \text{ mas } v_2 \notin \mathrm{Nu}(A - 2\mathsf{I}).$$

Pelo que obtivemos na etapa II e como

$$\begin{aligned} \mathrm{Nu}\left((A - 2\mathsf{I})^2\right) &= \mathrm{Nu}\begin{pmatrix} 0 & -2 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 1 \end{pmatrix} \\ &= \left[\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \end{pmatrix}\right], \end{aligned}$$

podemos considerar, por exemplo,

$$v_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}.$$

Nesse caso,

$$(A - 2I)v_2 = \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix}.$$

V. Caso as colunas de P sejam v_1, $(A - 2\mathsf{I})v_2$, v_2 e v_3, nessa ordem, verifique que

$$\mathsf{P}^{-1}A\mathsf{P} = J$$

é dada como no item 2 da etapa III.