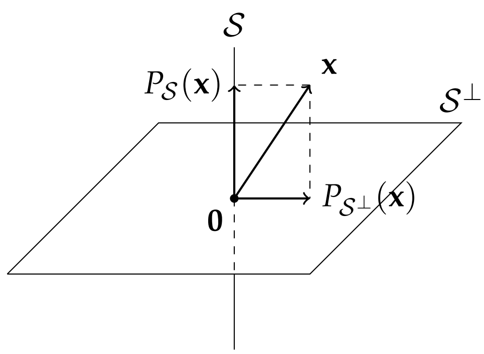
Lições de Álgebra Linear I e II

via Exemplos e Exercícios Resolvidos



Projeções ortogonais do vetor $\mathbf{x}$ sobre os subespaços $\mathcal{S}$ e $\mathcal{S}^{\perp}$.

José Renato Ramos Barbosa

DMAT – UFPR

1a. versão: 2005 – últ. versão: 2024

UNIVERSIDADE FEDERAL DO PARANÁ

DEPARTAMENTO DE MATEMÁTICA

[redacted]

LIÇÕES DE ÁLGEBRA LINEAR
I E II
VIA EXEMPLOS E EXERCÍCIOS RESOLVIDOS

Autor:
Professor José Renato Ramos Barbosa

2024

[redacted]

Dedico essas singelas lições
ao meu filho Theo.

Conteúdo

Capítulo 1

Origem, objetivos e diretrizes

SAUDAÇÕES UNIVERSITÁRIAS!
Nessa era de reducionismo, sumarização e imediatismo, prefácios e bibliografias costumam "passar batidos". Assim, nesse curtíssimo capítulo 1 (com apenas quatro páginas), contextualizaremos o conteúdo dessas *lições* e de suas cinco referências bibliográficas, para que todos possam lê-lo e ter noção do tratamento dado (a partir do capítulo 2) a *álgebra linear* (AL). Por isso, pedimos a atenção de leitores (e leitoras) a descrição das duas partes dessas lições e a abordagem ("quase" sem *determinantes*) que adotamos.[1]
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.[2]
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 essas lições 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 (P_1) se estende até o final da seção 6.2, incluindo alguns exercícios da 6.4. A segunda (P_2) vai da seção 6.3 até o final do livro.
Nossa abordagem é distinta das adotadas na maioria dos livros-texto. Em P_1, 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 $n = 4, 5, \ldots$, o escopo é mais geométrico e alguns resultados mais abstratos, não (ou parcialmente) demonstrados em P_1, são demonstrados em P_2.
Esperando que o(a) leitor(a) tenha absorvido essa generalização,[3] fazemos outra transição sutil, do $\mathbb{R}^n$ para o *espaço* $\mathbb{R}^{m\times n}$ das *matrizes com m linhas e n colunas* e, depois, seguimos em "movimento inercial" para *escalonamento* de matrizes, *sistemas lineares* e um estudo operacional, similar ao de um formulário de algum manual de tabelas e fórmulas, de *determinantes*, com ênfase (apenas) na resolução de exercícios.
Convém ressaltarmos que, a rigor, a utilidade da teoria de determinantes é apenas teórica.

[1] Algumas partes do texto estão escritas em primeira pessoa (do plural).

[2] www.ufpr.br/∼jrrb. A propósito, nesse endereço eletrônico, os leitores poderão encontrar a errata desse livro.

[3] Inclusive, na seção 2.6, apresentamos parte de uma "AL no $\mathbb{R}^3$".

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 de determinantes com o conhecimento adquirido no ensino médio.

Na sequência, ainda em P_1, entramos "matricialmente" - sem nos afastarmos do escopo geométrico supracitado - no estudo de *funções lineares* e suas *diagonalizações* e, só depois dessa abordagem mais aplicada/concreta, apresentamos *escalares* e "espaços" mais gerais.

Em todo esse trajeto, a abstração vai aumentando gradualmente, que é o processo natural para quem está se ambientando com algum conhecimento novo.

Ao longo do texto, as notas de rodapé têm um papel importante e *devem* ser lidas como parte integrante dele, 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 mesmo tamanho das notas de rodapé. Por outro lado, para quem já cursou AL, a leitura pode ser feita em ritmo de revisão.

Por terem um "sabor" semelhante ao de P_1, 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

.

Deliberadamente, não incluí (ou adiei) demonstrações de alguns resultados em P_1, pois ela foi concebida para cursos introdutórios ou mais aplicados (engenharias, por exemplo). Leitores interessados em preencher essas lacunas são convidados a recorrer a outros livros da área (como os supracitados) ou tentar entender P_2 antecipadamente, onde adotamos uma abordagem de um segundo curso (ou de um curso *honors*) de AL e, como dissemos, demonstramos resultados que tinham sido apenas enunciados (ou parcialmente demonstrados) em P_1. Além disso, em P_2, 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

.

Assim, alunos da matemática também podem utilizar essas lições, ficando P_1 como aplicações/exemplos de P_2.

Em linhas gerais, para P_2, seguimos a abordagem do excelente livro:

LINEAR ALGEBRA DONE RIGHT
SHELDON AXLER
3RD EDITION
SPRINGER VERLAG

,

doravante referenciado simplesmente por AXLER.

O capítulo 7 dessas lições é a "jóia da coroa". Nele é feita uma análise pormenorizada de

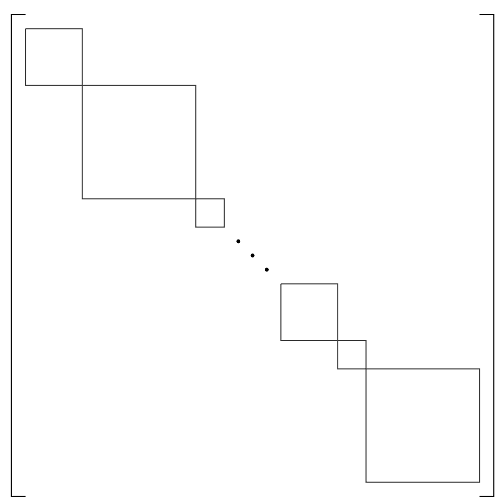
projeções ortogonais e *mínimos quadrados*, fundamentais nas áreas de *otimização* e *estatística*. Contudo, o *teorema espectral real/complexo* e a *forma de Jordan* são os resultados mais importante do capítulo 7 (e talvez da AL), onde serão (rigorosamente) demonstrados. Para darmos uma noção rudimentar desses resultados, consideraremos o seguinte:
Na matemática existem objetos que representam toda uma classe de outros objetos. Por exemplo, a fração irredutível $1/2$ representa qualquer fração da lista

$$\frac{2}{4}, \frac{3}{6}, \frac{4}{8}, \text{etc.}$$

Apresentaremos, para uma matriz quadrada A arbitrária, uma matriz J (de mesmo tamanho) que representa A, onde existem matrizes quadradas (ditas *blocos de Jordan*), em geral de tamanhos variados, cujas *diagonais principais* coincidem com partes da diagonal principal de J. Além disso, só existem zeros (*entradas* nulas) fora dessa "diagonal de blocos". (Confira a figura 1.1 para uma ilustração.)

Figura 1.1: Forma de Jordan

$J =$



Observamos que a figura 1.1 é uma representação qualitativa da forma de Jordan de uma matriz. Assim, os tamanhos dos seus blocos podem diferir dos apresentados nessa figura.[4]
Embora todas as matrizes possam ser representadas por suas respectivas formas de Jordan, algumas delas podem ser representadas por matrizes *diagonais* ainda mais simples, com todas as entradas, fora de suas diagonais principais, nulas.[5] Essa propriedade está estabelecida no teorema espectral supracitado.

[4]A demonstração do teorema da forma de Jordan está baseada (principalmente) no lema 2 do capítulo 7, onde seguimos de perto outro excelente artigo:

A SHORT PROOF OF THE EXISTENCE OF JORDAN NORMAL FORM
MARK WILDON, ROYAL HOLLOWAY, UNIVERSITY OF LONDON

[5]Note que, em termos de *processamento* (e, em particular, *armazenamento*) *de dados*, caso uma matriz não diagonal A seja n por n, tenha $m > n$ entradas não nulas e possa ser representada por uma matriz diagonal D, é mais fácil lidarmos com as n entradas da diagonal principal de D (que são as únicas informações relevantes dessa matriz) do que com as entradas de A, pois $n < m \leq n^2$.

Em tempo:

- A subseção 7.6.3, última do livro, é de autoria do Professor Marcelo Muniz Silva Alves, meu colega do DMAT da UFPR.[6] Nela, é apresentado um método mais eficiente para a obtenção da forma de Jordan, via diagramas de pontos;
- A apresentação das notações de alguns objetos matemáticos apresentados em P_1 foi postergada para P_2;
- Apenas enumeramos fórmulas referenciadas ao longo do texto;
- Em geral, os exercícios foram enumerados e os exemplos, por estarem mais fragmentados, não foram numerados;
- Como utilizamos demonstrações algorítmicas para alguns resultados, partes delas podem ser reescritas como comandos em alguma linguagem de programação;
- O pré-requisito para a leitura de P_1 é um curso de GA e, no início do capítulo 2, fazemos uma revisão de GA em $\mathbb{R}^2$;
- Na produção do texto e de suas figuras, utilizamos o LaTeX, que é uma bem conhecida linguagem científica de editoração eletrônica, e o Ti*k*Z, que é uma ferramenta poderosa para a geração de figuras e gráficos na linguagem supracitada. Assim, esperamos que a leitura seja "agradável aos olhos".

Como conclusão desse capítulo, gostaríamos de expressar que a matemática e algumas áreas do conhecimento podem ser vistas como linguagens, ou seja, do mesmo modo que português, inglês, francês, etc., podemos considerar "matematiquês", "fisiquês", "quimiquês", "informatiquês", "economês", etc. O domínio dessas línguas não ocorre antes de aprendermos o "bê-á-bá" delas e, depois de vencida essa etapa preliminar, é preciso estudá-las e praticá-las para que alguns equívocos não sejam cometidos. Sem esse nível de comprometimento, não é fácil fazermos um estudo avançado dessas línguas, pois, como diz o ditado, "O avançado é fazer o básico bem feito!". Por outro lado, para sermos fluentes, além de estudo e prática, é imprescindível que saibamos utilizar os jargões da língua estudada. Assim, a imersão numa liguagem é fundamental e o hábito de estudarmos (somente) nas vésperas das provas deve ser evitado.

Sugestões para o aprimoramento e/ou a clareza desse livro serão muito bem vindas. Serei, não só grato, mas também todo "ouvidos e olhos".

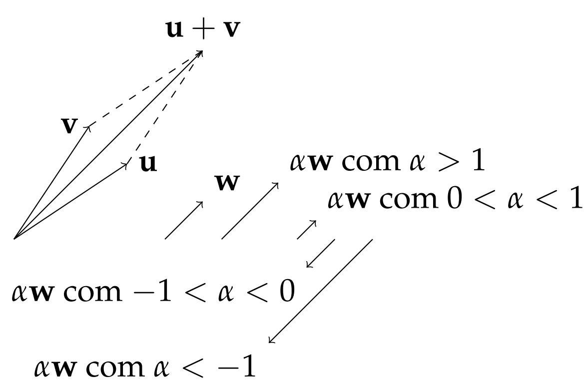
[6] Agradeço ao Professor Marcelo pela gentileza de ter cedido o material dessa subseção e por ter feito várias sugestões na redação do livro. Nesse contexto, também sou grato pelas inúmeras sugestões de formatação e gramática fornecidas por outro colega do DMAT, Professor Adam Luiz de Azevedo.

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

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

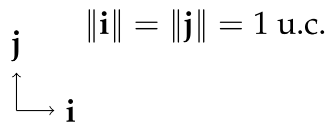
[1]Esses vetores são representados por pares ordenados de números reais.
[2]α é um número real.
[3]Ou seja, unidades de comprimento.

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.*

Essas operações satisfazem algumas propriedades,[4] que podem ser demonstradas ou justificadas geometricamente,[5] algebricamente e numericamente.
Para os exercícios dessa seção, considere:

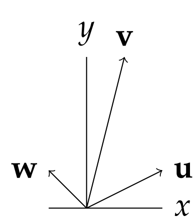
- $\mathbf{i} = (1,0)$ e $\mathbf{j} = (0,1)$, conforme a figura 2.2.

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 $\mathbf{u} = \big(1, x^2-1\big)$ e $\mathbf{v} = (x+2, 0)$ verificam a igualdade $\mathbf{u} = \mathbf{v}$? $\boxed{\text{RESPOSTA}}$ $x = -1$.

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?

[4] Comutatividade tanto da adição quanto do produto interno de vetores; o módulo do múltiplo escalar de um vetor é igual ao 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.

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

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 para cada um dos itens seguintes:

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

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

---

[6] $\boxed{\text{SUGESTÃO}}$
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] $\boxed{\text{SUGESTÕES}}$
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] $\boxed{\text{RESOLUÇÃO}}$

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

(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]
$$|\mathbf{u}\cdot\mathbf{v}| \leq \|\mathbf{u}\|\,\|\mathbf{v}\|,$$
      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]
   (b) Determine θ para $\mathbf{v} = -4\mathbf{i} + 3\mathbf{j}$.

---

[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.

   (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 $\mathbf{a} = \mathbf{j}$.

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

---

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

[13] Confira o exercício 8 dessa subseção.

[14] $\boxed{\text{SUGESTÃO}}$
Utilize o seguinte resultado:
$$\|\mathbf{u} + \mathbf{v}\|^2 = (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}).$$

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

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.

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

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

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

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

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.

OBSERVAÇÃO
Para quaisquer vetores $\mathbf{x}$ e $\mathbf{y}$ em $\mathbb{R}^n$, $\mathbf{x}-\mathbf{y} := \mathbf{x}+(-1)\mathbf{y}$.

[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 $\mathbb{R}^n$,

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

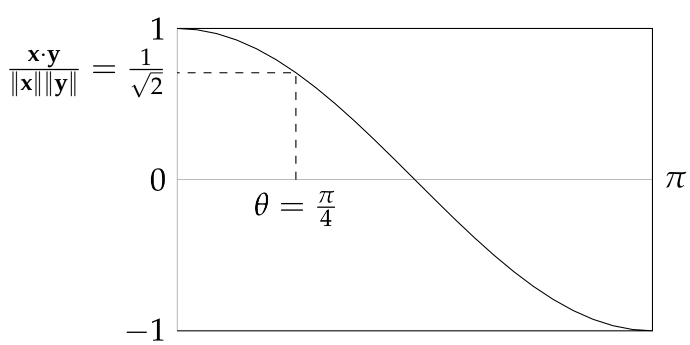
pela desigualdade de Cauchy-Schwarz e pela propriedade da não negatividade de normas.[19] Portanto, como $\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}$ é um número entre -1 e 1, existe um único ângulo $\theta = \theta(\mathbf{x},\mathbf{y})$ (em radianos) entre 0 e π tal que

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

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 $i \in \{1, \ldots, n\}$ 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 $j \in \{1, \ldots, n\}$, então

$$\mathbf{e}_i \cdot \mathbf{e}_j = \begin{cases} 0 & \text{se } i \neq j; \\ 1 & \text{se } i = j. \end{cases} \tag{2.4}$$

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:

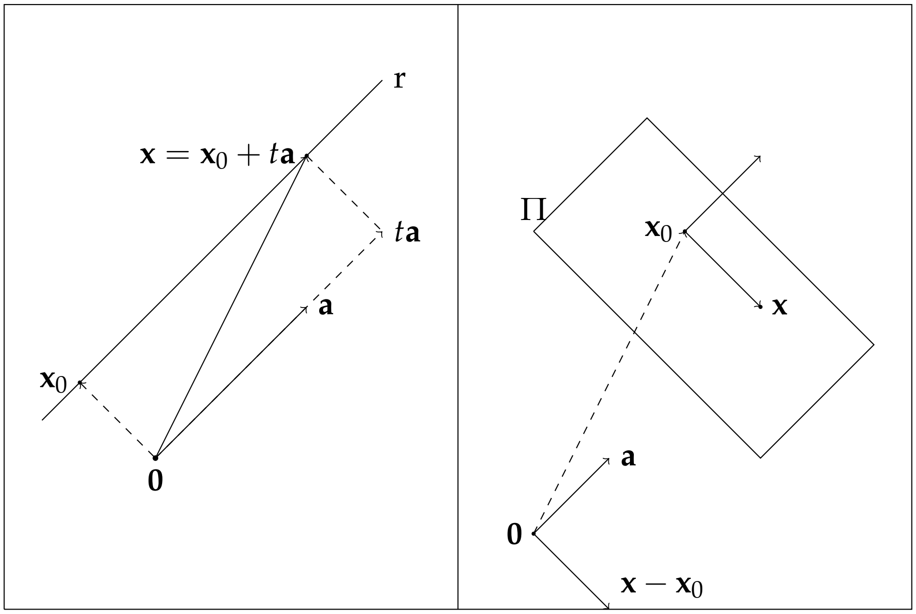
1. $\boxed{x = x_0 + ta,\ y = y_0 + tb,\ z = z_0 + tc}$, onde t representa um escalar arbitrário;

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)

$$\mathcal{S} = \left\{ \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} \;\middle|\; y = x \right\} \tag{2.5}$$

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,

$$\begin{cases} y_1 & = & x_1, \\ y_2 & = & x_2, \end{cases} \tag{2.6}$$

então:

- $\alpha\mathbf{x} = \begin{bmatrix} \alpha x_1 \\ \alpha y_1 \end{bmatrix} \in \mathcal{S}$, pois, ao multiplicarmos a primeira igualdade de (2.6) 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 (2.6) 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,

$$\begin{cases} x_1 & + & y_1 & + & z_1 & = & 0, \\ x_2 & + & y_2 & + & z_2 & = & 0, \end{cases}$$

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 (2.5), 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

Demonstre que $\mathcal{S}$ é um subespaço de $\mathbb{R}^n$. Aqui, tanto o vetor $\mathbf{a}$ quanto os vetores da lista

$$\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r \tag{2.7}$$

estão fixados em $\mathbb{R}^n$.[23]

1. $\mathcal{S} = \{\mathbf{0}\}$; (SUBESPAÇO TRIVIAL)
2. $\mathcal{S} = \mathbb{R}^n$; (SUBESPAÇO TRIVIAL)
3. $\mathcal{S} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = t\mathbf{a} \text{ e } t \in \mathbb{R} \text{ é arbitrário}\}$; (RETA QUE PASSA PELA ORIGEM NA DIREÇÃO DO VETOR $\mathbf{a} \neq \mathbf{0}$)
4. $\mathcal{S} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a} \cdot \mathbf{x} = 0\}$; (HIPERPLANO QUE PASSA PELA ORIGEM COM NORMAL $\mathbf{a} \neq \mathbf{0}$)
5. $\mathcal{S} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_i \cdot \mathbf{x} = 0 \text{ e } \mathbf{a}_i \neq \mathbf{0}, i = 1, 2, \ldots, r\}$;[24] (INTERSEÇÃO DE r HIPERPLANOS QUE PASSAM PELA ORIGEM COM VETORES NORMAIS NA LISTA (2.7))
6. Para r escalares arbitrários, digamos c_i, $i = 1, 2, \ldots, r$, a soma

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_r\mathbf{a}_r \tag{2.8}$$

é chamada de *combinação linear* (CL) dos vetores da lista (2.7). Seja $\mathcal{S}$ o conjunto de todas essas combinações lineares. (SUBESPAÇO GERADO PELOS VETORES DA LISTA (2.7))

Qualquer subespaço $\mathcal{S} \neq \{\mathbf{0}\}$ de $\mathbb{R}^n$ é gerado por um número finito de vetores, ou seja, é dado como no item 6 desse exercício. 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] A resolução desse exercício encontra-se na subseção 2.7.1.

[24] EXEMPLOS

- Para
$$\mathcal{S} = \left\{ (x, y, z) \in \mathbb{R}^3 \,\middle|\, \left\{ \begin{array}{ccccccc} x & + & y & + & z & = & 0 \\ x & - & 2y & + & z & = & 0 \\ x & + & y & - & 3z & = & 0 \end{array} \right. \right\},$$
$n = r = 3$, $\mathbf{a}_1 = (1, 1, 1)$, $\mathbf{a}_2 = (1, -2, 1)$ e $\mathbf{a}_3 = (1, 1, -3)$.
- Para
$$\mathcal{S} = \left\{ (x, y, z, w) \in \mathbb{R}^4 \,\middle|\, \left\{ \begin{array}{ccccccccc} x & + & y & & & + & w & = & 0 \\ x & - & y & + & z & & & = & 0 \end{array} \right. \right\},$$
$n = 2r = 4$, $\mathbf{a}_1 = (1, 1, 0, 1)$ 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

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

Para outros valores de c_1, c_2 e c_3, existe uma infinidade de combinações lineares de $\mathbf{a}_1, \mathbf{a}_2$ e $\mathbf{a}_3$. Em particular, $\mathbf{a}_1 \in \mathcal{S}$.[25] Analogamente, $\mathbf{a}_2$ e $\mathbf{a}_3$ são combinações lineares 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 combinações lineares 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

Caso $\mathcal{S}$ seja o subespaço de $\mathbb{R}^n$ gerado pelos r vetores da lista (2.7), isto é,

$$\mathcal{S} = \left\{\mathbf{x} \in \mathbb{R}^n \,\middle|\, \mathbf{x} \text{ é dado por (2.8), com } c_i \text{ arbitrário, } i = 1, 2, \ldots, r\right\},$$

diremos que $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r\}$ é uma *base* de $\mathcal{S}$ se, e somente se, esses r vetores forem *linearmente independentes* (LI), ou seja, a única solução da equação

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

for *trivial*, isto é, dada por

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

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

EXEMPLOS EM $\mathbb{R}^3$

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

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

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 uma 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 dada por $x_1 = 1$, $x_2 = 1$ e $x_3 = -1$.[28] Logo, como $\mathbf{a}_3$ é uma 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,

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

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 de $\mathcal{S}$.

[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 de $\mathbb{R}^n$ se, e somente se, cada $\mathbf{x} \in \mathbb{R}^n$ pode ser escrito de modo único como uma CL dos vetores dessa base.[29]

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

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

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

  · Esses vetores são LI.

  Essa equivalência é válida para todo inteiro positivo n, conforme o resultado (R9) do capítulo 6.[30] Na seção 2.6, dedicada exclusivamente ao espaço vetorial $\mathbb{R}^3$, demostraremos a equivalência supracitada para o caso $n = 3$.

2.5.4 Dimensão

Seja $\mathcal{S}$ um subespaço de $\mathbb{R}^n$. Na subseção 6.3.1, será demonstrado 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. p. 154.

[31] Como vimos na subseção 2.5.1, $\mathcal{S}$ é gerado por vetores de uma lista finita. Caso esses *geradores* não sejam LI, poderemos descartar todos os vetores dessa lista que forem combinações lineares dos demais, restando apenas vetores LI na lista supracitada.

- 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 uma 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 $i \neq j$, $\mathbf{a}_i \cdot \mathbf{a}_j = 0$, diremos que $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r\}$ é (um conjunto) *ortogonal*.

EXEMPLO

Em $\mathbb{R}^4$,

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

TEOREMA

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

ortogonalidade $\Longrightarrow$ 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 $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$ e multipliquemos 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 $\mathbf{a}_i$,

$$c_1\mathbf{a}_1 \cdot \mathbf{a}_i + \cdots + c_{i-1}\mathbf{a}_{i-1} \cdot \mathbf{a}_i + c_i\mathbf{a}_i \cdot \mathbf{a}_i + c_{i+1}\mathbf{a}_{i+1} \cdot \mathbf{a}_i + \cdots + c_r\mathbf{a}_r \cdot \mathbf{a}_i = \mathbf{0} \cdot \mathbf{a}_i. \tag{2.9}$$

Como, para $j \neq i$, $\mathbf{a}_j \cdot \mathbf{a}_i = 0$,[35] (2.9) pode ser reescrita como

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

ou seja, $c_i = 0$.[36] 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, $\dim \mathcal{S} = 3$.

Diremos 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 $\mathbf{v} \neq \mathbf{0}$,[37] 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 $\mathcal{S}$.[38]

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.[39]

---

[35] Pela hipótese da ortogonalidade entre os r vetores da lista supracitada.

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

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

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

[38] Pelos exercícios 6 e 9 da seção 2.7, podemos "ortonormalizar" qualquer base de $\mathcal{S}$ e (prontamente) determinar as "coordenadas" de qualquer vetor de $\mathcal{S}$ na base "ortonormalizada".

[39] Cf. (2.4), p. 22.

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,[40] 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

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

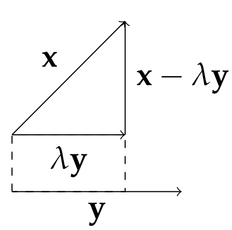
DEMONSTRAÇÃO

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.

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



EXEMPLO

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

[40]Cf. p. 154.

2.6.2 Produto vetorial de x e y

É definido por

$$\mathbf{x} \times \mathbf{y} := (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1)\,.$$[41]

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

$$\|\mathbf{x} \times \mathbf{y}\| = \|\mathbf{y}\|\|\mathbf{x} - \lambda\mathbf{y}\|. \tag{2.10}$$

EXEMPLO

Para $\mathbf{x}$ e $\mathbf{y}$ dados nos exemplos das subseções 2.6.1 e 2.6.2, ambos os membros de (2.10) são iguais a $\sqrt{3}$.

DEMONSTRAÇÃO DE (2.10)

[41]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

$$\boxed{\mathbf{x} \textit{ e } \mathbf{y} \textit{ são } \mathrm{LD} \Longleftrightarrow \mathbf{x}\times\mathbf{y} = \mathbf{0}.}$$

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.[42] Como $\|\mathbf{x}-\lambda\mathbf{y}\| = 0$,[43] 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

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

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

[43] Por (2.10).

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 $\mathbf{y} = (-1, 0, 1)$,[44] que são LI.[45] Além disso, $\mathbf{x} \times \mathbf{y} = -2\mathbf{z}$.

DEMONSTRAÇÃO DO TEO. 1

Pela hipótese da ortogonalidade, temos
$$\begin{cases} x_1z_1 + x_2z_2 + x_3z_3 = 0,\\ y_1z_1 + y_2z_2 + y_3z_3 = 0.\end{cases}$$
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,
$$\begin{cases} z_2\,(x_2y_1 - x_1y_2) + z_3\,(x_3y_1 - x_1y_3) = 0,\\ z_1\,(x_1y_2 - x_2y_1) + z_3\,(x_3y_2 - x_2y_3) = 0,\\ z_1\,(x_1y_3 - x_3y_1) + z_2\,(x_2y_3 - x_3y_2) = 0.\end{cases} \tag{2.11}$$
Por outro lado, pela hipótese da independência linear, temos
$$\mathbf{x} \times \mathbf{y} \neq \mathbf{0}.^{46}$$

[44] De fato, $\mathbf{x} \cdot \mathbf{z} = 0$ e $\mathbf{y} \cdot \mathbf{z} = 0$.
[45] De fato, não podemos escrever um dos vetores, $\mathbf{x}$ ou $\mathbf{y}$, como múltiplo escalar do outro.
[46] Caso contrário, pela equação (2.10), 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 (2.11), 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 $\mathbf{y}$;[47]
- $\mathbf{x}'\neq\mathbf{0}$, 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 $\mathbf{y}$.[48] Além disso, como $\mathbf{x}'$ e $\mathbf{y}$ são não nulos e ortogonais entre si, também são LI.[49] 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,

$$\mathbf{z} = \alpha'\mathbf{x}' + \beta'\mathbf{y} + \gamma'\left(\mathbf{x}'\times\mathbf{y}\right). \tag{2.12}$$

Se substituirmos $\mathbf{x}' = \mathbf{x} - \lambda\mathbf{y}$ em (2.12), utilizarmos a igualdade $\mathbf{y}\times\mathbf{y} = \mathbf{0}$ e reagruparmos os coeficientes, podemos obter escalares α, β e γ tais que

$$\mathbf{z} = \alpha\mathbf{x} + \beta\mathbf{y} + \gamma(\mathbf{x}\times\mathbf{y}).^{50} \tag{2.13}$$

Para concluirmos a demonstração, basta verificarmos que a tripla ordenada (α,β,γ) é única, em relação a CL (2.13). Suponha, por contradição, que não seja única. Assim, existe outra tripla ordenada (a,b,c) tal que

$$\mathbf{z} = a\mathbf{x} + b\mathbf{y} + c(\mathbf{x}\times\mathbf{y}). \tag{2.14}$$

[47] Cf. a subseção 2.6.1.

[48] Verifique!

[49] Cf. a subseção 2.5.5.

[50] Verifique!

Então, da diferença (2.13)−(2.14), obtemos escalares a', b' e c', não todos nulos, tais que

$$\mathbf{0} = a'\mathbf{x} + b'\mathbf{y} + c'(\mathbf{x} \times \mathbf{y}).^{51}$$

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

$$0 = c'\|\mathbf{x} \times \mathbf{y}\|^2.^{52}$$

Logo, como $\|\mathbf{x} \times \mathbf{y}\| \neq 0$, pois $\mathbf{x}$ e $\mathbf{y}$ são LI,[53] 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, isto é, chegamos a uma contradição da 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 positivo n, conforme demonstraremos 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.[54] Então, como qualquer um deles gera $\mathbb{R}^3$, podemos escrever, por exemplo, o vetor $\mathbf{i}$ (da base canônica) como um múltiplo escalar do vetor $\mathbf{j}$ (da base canônica), 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 $\mathbf{y}$,[55] ele não pode ser uma CL desses vetores, ou seja, chegamos a uma contradição da hipótese 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 independência linear de $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{z}$, γ é não nulo e, assim,

$$\mathbf{x} \times \mathbf{y} = \frac{1}{\gamma}(\mathbf{z} - \alpha\mathbf{x} - \beta\mathbf{y}). \tag{2.15}$$

[51] $a' = \alpha - a$, $b' = \beta - b$ e $c' = \gamma - c$.
[52] Cf. o teorema 1 supracitado.
[53] Cf. a subseção 2.6.4.
[54] Idem.
[55] Cf. o exercício 2 da subseção 2.6.5.

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

$$\mathbf{v} = \alpha'\mathbf{x} + \beta'\mathbf{y} + \gamma'(\mathbf{x} \times \mathbf{y}). \tag{2.16}$$

Ao substituirmos (2.15) em (2.16), 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

$$\mathbf{v} = \alpha\mathbf{x} + \beta\mathbf{y} + \gamma\mathbf{z}, \tag{2.17}$$

então

$$\alpha = \frac{(\mathbf{v}, \mathbf{y}, \mathbf{z})}{(\mathbf{x}, \mathbf{y}, \mathbf{z})}, \quad \beta = \frac{(\mathbf{x}, \mathbf{v}, \mathbf{z})}{(\mathbf{x}, \mathbf{y}, \mathbf{z})} \quad e \quad \gamma = \frac{(\mathbf{x}, \mathbf{y}, \mathbf{v})}{(\mathbf{x}, \mathbf{y}, \mathbf{z})}. \tag{2.18}$$

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,

$$\mathbf{z} \cdot (\mathbf{x} \times \mathbf{y}) = \mathbf{x} \cdot (\mathbf{y} \times \mathbf{z}) \neq 0. \tag{2.19}$$

Então, $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{x} \times \mathbf{y}$ são não nulos e, além disso, $\mathbf{x}$ e $\mathbf{y}$ são LI.[56] Logo, $\{\mathbf{x}, \mathbf{y}, \mathbf{x} \times \mathbf{y}\}$ é uma base de $\mathbb{R}^3$, pelo teorema 2 supracitado, e, assim, existem escalares α, β e γ tais que

$$\mathbf{z} = \alpha\mathbf{x} + \beta\mathbf{y} + \gamma(\mathbf{x} \times \mathbf{y}). \tag{2.20}$$

Ao calcularmos o produto interno entre cada membro de (2.20) e o vetor $\mathbf{x} \times \mathbf{y}$, obtemos

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

[56] Cf. a subseção 2.6.4.

Como, por (2.19), $\gamma \neq 0$,

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

pela equação (2.20). Portanto, como cada vetor $\mathbf{v} \in \mathbb{R}^3$ pode ser escrito como uma CL de $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{x} \times \mathbf{y}$, $\mathbf{v}$ também pode ser escrito como uma 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 (2.17) seja válida. Ao multiplicarmos 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 (2.18) e, concomitantemente, a unicidade da CL (2.17).

$\boxed{2 \Longrightarrow 1}$

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, $\mathbf{x} \times \mathbf{y} \neq \mathbf{0}$.[57] 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

$$\mathbf{z} = \alpha\mathbf{x} + \beta\mathbf{y} + \gamma(\mathbf{x} \times \mathbf{y}), \tag{2.21}$$

com $\gamma \neq 0$.[58] 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 (2.21) 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. A existência e a unicidade da solução do "sistema linear" (2.17) são decorrentes da independência linear das colunas do determinante $(\mathbf{x}, \mathbf{y}, \mathbf{z})$.[59]

[57] Cf. a subseção 2.6.4.
[58] 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 $\mathbf{y}$.
[59] Cf. os teoremas 3 e 4 supracitados.

2.7 Exercícios

1. Em $\mathbb{R}^4$, caso o vetor $\mathbf{x}$ tenha quatro coordenadas iguais, a última coordenada do vetor $\mathbf{y}$ seja igual a 1 e $\mathbf{x}+\mathbf{y}=(2,3,4,5)$, determine esses vetores.[60]

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

$$x\mathbf{x}+y\mathbf{y}+z\mathbf{z}=(41/2,11,1/2).$$ [61]

3. Qual é a equação geral do hiperplano do $\mathbb{R}^4$ que contem os pontos $\mathcal{P}_1=(1,-1,1,0)$, $\mathcal{P}_2=(0,-1,2,0)$, $\mathcal{P}_3=(1,0,-2,2)$ e $\mathcal{P}_4=(1,0,0,0)$? Quais são os pontos da reta que passa por $\mathcal{P}_1$, na direção do vetor normal ao hiperplano supracitado, que estão 1 u.c. equidistantes de $\mathcal{P}_1$?

RESOLUÇÃO PARCIAL

Caso $ax+by+cz+dw=e$ seja a equação geral do hiperplano supracitado e $\mathbf{x}_1=\mathcal{P}_1-\mathcal{P}_4$, $\mathbf{x}_2=\mathcal{P}_2-\mathcal{P}_4$ e $\mathbf{x}_3=\mathcal{P}_3-\mathcal{P}_4$, $\mathbf{a}=(a,b,c,d)$ e $\mathbf{x}_i$ são ortogonais, $i=1,2,3$. Portanto, como $a=b=c=d$,[62] a equação procurada é dada por $x+y+z+w=1$.[63] Além disso, caso $\mathbf{x}=\mathbf{x}_0+t\mathbf{a}$ seja a equação vetorial da reta supracitada e $\mathcal{P}_1$ seja o ponto final do vetor $\mathbf{x}_0$, para obtermos os (dois) pontos dessa reta, equidistantes 1 u.c. de $\mathcal{P}_1$, basta determinarmos cada $\mathbf{x}$ que satisfaça as seguintes equações:

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

Verifique que

$$\mathbf{x}\in\left\{\left(\frac{1}{2},-\frac{3}{2},\frac{1}{2},-\frac{1}{2}\right),\left(\frac{3}{2},-\frac{1}{2},\frac{3}{2},\frac{1}{2}\right)\right\}.$$

4. Para cada um dos itens seguintes, demonstre que $\mathcal{S}$ é um subespaço de $\mathbb{R}^n$, sem utilizar os itens 3, 4 e 5 do exercício da subseção 2.5.1.[64] Além disso, determine uma base de $\mathcal{S}$ e a 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) $\mathcal{S}=\left\{\mathbf{x}=(x_1,x_2,x_3,x_4)\;\middle|\;x_4=x_1+2x_2+3x_3\right\}$.

---

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

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

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

[63] De fato, basta substituirmos as coordenadas de $\mathcal{P}_4$ em $a(x+y+z+w)=e$.

[64] Cf. p. 26.

RESOLUÇÃO

(a) Pelo exercício 6 da subseção supracitada, 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 $\dim \mathcal{S} = 2$. De fato, como $(a, b, c) \neq \mathbf{0}$ tem alguma coordenada não nula, podemos supor (sem perda de generalidade) que $c \neq 0$. Assim, por um lado, como

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

$\mathbf{a}_1 = \left(1, 0, -\frac{a}{c}\right)$ 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

$$\mathbf{x} = \mathbf{0} \iff x_1 = x_2 = 0.$$

(b1) Caso seja gerado por r vetores, $\mathcal{S}$ é um 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

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

$\mathbf{a}_1 = (1, 2, 3, 4) \neq \mathbf{0}$ gera $\mathcal{S}$.
(b.2) Caso seja gerado por r vetores, $\mathcal{S}$ é um 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

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

$\mathbf{a}_1 = (1, 0, 1, 1)$ e $\mathbf{a}_2 = (0, 1, 2, -1)$ geram $\mathcal{S}$ e, como nenhum deles é um múltiplo escalar do outro, são LI.
(b.3) Para que $\mathcal{S}$ seja um subespaço de $\mathbb{R}^4$, basta que 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

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

$\mathbf{a}_1 = (1, 0, 0, 1)$, $\mathbf{a}_2 = (0, 1, 0, 2)$ e $\mathbf{a}_3 = (0, 0, 1, 3)$ geram $\mathcal{S}$ e, como

$$\mathbf{x} = \mathbf{0} \iff x_i = 0, i = 1, 2, 3,$$

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[65]

(a) Considere, por exemplo, $\mathbf{x} = (1, 1)$ e $\alpha = 2$. Assim, a condição

$$\mathbf{x} \in \mathcal{S} \implies \alpha\mathbf{x} \in \mathcal{S}$$

não é satisfeita, pois a segunda coordenada de $2\mathbf{x}$ não é igual ao quadrado da primeira.
(b) O vetor nulo não pertence a $\mathcal{S}$.

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

6. Caso $\mathcal{S}$ seja um subespaço de $\mathbb{R}^n$, demonstre as afirmações seguintes:

   (a) Nenhuma base de $\mathcal{S}$ pode conter o vetor nulo.[66]

   (b) Nenhum vetor pode ter "coordenadas" distintas numa mesma base, isto é, caso $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r\}$ seja uma base de $\mathcal{S}$, $c_1, c_1', c_2, c_2', \ldots, c_r, c_r'$ seja uma lista de escalares e $\mathbf{x}$ seja um vetor de $\mathcal{S}$ tal que

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

   $c_1 = c_1'$, $c_2 = c_2' \ldots$, $c_r = c_r'$.[67]

$$\boxed{c_i \text{ é a } i\text{-ésima coordenada de } \mathbf{x} \text{ na base supracitada, } i = 1, 2, \ldots, r.}$$

   (c) Caso $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r\}$ seja uma base ortonormal de $\mathcal{S}$ e $\mathbf{x}$ seja a CL do item (b) desse exercício, as coordenadas de $\mathbf{x}$ (nessa base) são dadas por

$$c_1 = \mathbf{x}\cdot\mathbf{a}_1,\ c_2 = \mathbf{x}\cdot\mathbf{a}_2,\ \ldots,\ c_r = \mathbf{x}\cdot\mathbf{a}_r. \tag{2.22}$$

[68]

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.[69]

---

[66] RESOLUÇÃO

Em $\mathbb{R}^n$, caso $\mathbf{a}_i = \mathbf{0}$ seja um dos vetores da lista $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$, 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 que c_i seja o único coeficiente não nulo. Por exemplo, se $c_i = 1$, então a CL supracitada é dada por

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

[67] DICA

$(c_1 - c_1')\,\mathbf{a}_1 + \cdots + (c_r - c_r')\,\mathbf{a}_r = \mathbf{0}$ e os r vetores da base supracitada são LI.

[68] 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.

[69] SUGESTÃO

Calcule c_1 e c_2 via (2.22).

8. Considere

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

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

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

na base supracitada.[70]

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

$$\begin{aligned} \mathbf{a}'_1 &:= \mathbf{a}_1; \\ \mathbf{a}'_2 &:= \mathbf{a}_2 - \frac{\mathbf{a}_2 \cdot \mathbf{a}'_1}{\mathbf{a}'_1 \cdot \mathbf{a}'_1}\mathbf{a}'_1; \\ \mathbf{a}'_3 &:= \mathbf{a}_3 - \frac{\mathbf{a}_3 \cdot \mathbf{a}'_1}{\mathbf{a}'_1 \cdot \mathbf{a}'_1}\mathbf{a}'_1 - \frac{\mathbf{a}_3 \cdot \mathbf{a}'_2}{\mathbf{a}'_2 \cdot \mathbf{a}'_2}\mathbf{a}'_2; \\ &\vdots \\ \mathbf{a}'_r &:= \mathbf{a}_r - \frac{\mathbf{a}_r \cdot \mathbf{a}'_1}{\mathbf{a}'_1 \cdot \mathbf{a}'_1}\mathbf{a}'_1 - \frac{\mathbf{a}_r \cdot \mathbf{a}'_2}{\mathbf{a}'_2 \cdot \mathbf{a}'_2}\mathbf{a}'_2 - \cdots - \frac{\mathbf{a}_r \cdot \mathbf{a}'_{r-1}}{\mathbf{a}'_{r-1} \cdot \mathbf{a}'_{r-1}}\mathbf{a}'_{r-1}. \end{aligned} \tag{2.23}$$

Demonstre que, por exemplo, se $r = 3$, então $\mathcal{B}' = \{\mathbf{a}'_1, \mathbf{a}'_2, \mathbf{a}'_3\}$ é ortogonal.[71]

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 um múltiplo escalar do outro.[72] Logo, como esses vetores geram um subespaço $\mathcal{S}$ (de dimensão 2) do $\mathbb{R}^2$, $\mathcal{S} = \mathbb{R}^2$. Portanto, para podermos utilizar o

[70] Caso $\mathbf{x}$ seja uma CL dos vetores $\mathbf{a}_1$, $\mathbf{a}_2$ e $\mathbf{a}_3$, $\mathbf{x} \in \mathcal{S}$ e, embora não saibamos, *a priori*, se essa CL é válida, podemos utilizar o método heurístico, ou seja, supor que seja válida, na tentativa de calcularmos as coordenadas de $\mathbf{x}$ na base supracitada. Caso essas coordenadas sejam obtidas, a suposição feita é, na verdade, um fato.

[71] 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, se $\mathbf{a}_2 = \mathbf{x}$ e $\mathbf{a}'_1 = \mathbf{y}$, então $\mathbf{a}'_2$ e $\mathbf{a}'_1$ são ortogonais, pela subseção 2.6.1.

[72] Verifique!

algoritmo (2.23) da questão anterior, denotemos $\mathbf{a}_1 = (1,2)$ e $\mathbf{a}_2 = (3,4)$. Assim:

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

Como $\{\mathbf{a}_1', \mathbf{a}_2'\}$ é ortogonal,[73] podemos obter uma base $\{\mathbf{a}_1'', \mathbf{a}_2''\}$ ortonormal. De fato:

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

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 $\mathbb{R}^4$.[74]

12. Dê um exemplo de uma base ortonormal do $\mathbb{R}^3$ que contenha o vetor $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$.[75]

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 supracitada.

   RESOLUÇÃO

   Como sugerido na nota de rodapé do exercício anterior, podemos aplicar o processo de Gram-Schmidt numa base $\mathcal{B} = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4\}$, onde $\mathbf{a}_1 = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$ e, por exemplo, $\mathbf{a}_2 = (1,0,0,0)$, $\mathbf{a}_3 = (0,1,0,0)$ e

[73] Verifique!

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

[75] SUGESTÃO

Aplique Gram-Schmidt na base $\mathcal{B} = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$, onde $\mathbf{a}_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$ e, por exemplo, $\mathbf{a}_2 = (1,0,0)$ e $\mathbf{a}_3 = (0,1,0)$. Para outra resolução, 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 $\mathbf{a}_3 = \frac{\mathbf{a}_1 \times \mathbf{a}_2}{\|\mathbf{a}_1 \times \mathbf{a}_2\|}$.

$\mathbf{a}_4 = (0,0,1,0)$.[76] Portanto, por (2.23),

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

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 via:

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

[76] A verificação da independência linear dos quatro vetores de $\mathcal{B}$ fica a cargo do leitor.

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

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

e, para evitarmos a resolução de um sistema linear não trivial de quatro equações nas variáveis c_1, c_2, c_3 e c_4, calcularmos

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

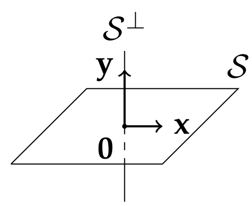
conforme a fórmula (2.22),[77], pois $\mathcal{B}''$ é ortonormal.

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.

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



(a) Para todo n, demonstre que $\mathcal{S}^\perp$ é subespaço de $\mathbb{R}^n$.[78]

(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 $\{(-1,-1,1)\}$;

---

[77] Cf. p. 41.

[78] 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 $\mathbf{x} \in \mathcal{S}$,

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

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,

$$\mathbf{x} \in \mathcal{S} \text{ e } \mathbf{y} \in \mathcal{S}^\perp \Longrightarrow \mathbf{y} \text{ e } \mathbf{x} \text{ são ortogonais},$$

caso seja possível obtermos uma base $\{\mathbf{x}_1, \ldots, \mathbf{x}_r\}$ de $\mathcal{S}$, $\mathbf{y}$ e $\mathbf{x}_i$ são ortogonais, $i = 1, \ldots, r$.

· Pelo resultado (R33),[79] $\dim \mathcal{S} + \dim \mathcal{S}^\perp = n$.

Logo, no item v desse exercício, por exemplo, para escrevermos

$$\mathbf{y} = (y_1, y_2, y_3, y_4) \in \mathcal{S}^\perp$$

como uma CL dos vetores de uma base de $\mathcal{S}^\perp$, basta resolvermos o sistema linear

$$\begin{cases} \mathbf{x}_1 \cdot \mathbf{y} = 0, \\ \mathbf{x}_2 \cdot \mathbf{y} = 0, \end{cases} \tag{2.24}$$

onde $\mathbf{x}_1$ e $\mathbf{x}_2$ são os vetores da base de $\mathcal{S}$ supracitada. A base de $\mathcal{S}^\perp$, obtida de (2.24), será composta por dois vetores. De fato,

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

15. Caso $\mathcal{S}$ seja o subespaço unidimensional de $\mathbb{R}^4$ gerado por $(1,1,1,1)$ e consideremos $\mathbf{y} = (1,1,1,-3)$, resolva os seguintes itens:

(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 supracitada.

RESOLUÇÃO

(a) Verificaremos que $\mathbf{y} \in \mathcal{S}^\perp$ de dois modos:

I. Como $\mathbf{y}$ e o gerador de $\mathcal{S}$ são ortogonais, $\mathbf{y}$ e qualquer vetor de $\mathcal{S}$ são ortogonais;

II Na questão 14 dessa seção, item iv, determinamos a base

$$\mathcal{B} = \{(1,0,0,-1),(0,1,0,-1),(0,0,1,-1)\} \tag{2.25}$$

de $\mathcal{S}^\perp$. Observe que $\mathbf{y}$ é a soma dos três vetores de (2.25).

[79]Cf. p. 195.

(b) Sejam $\mathbf{y}_1$, $\mathbf{y}_2$ e $\mathbf{y}_3$ os vetores de (2.25), 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:

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

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 (2.22).[80]

[80] Cf. p. 41.

2.7.1 Resolução do exercício da subseção 2.5.1

As condições 1, 2 e 3, do início da seção 2.5,[81] 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 $\alpha t \in \mathbb{R}$;
- $\mathbf{x}+\mathbf{y} \in \mathcal{S}$, 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,[82] ou seja, $\mathcal{S} = \{\mathbf{x} \mid \mathbf{a}_i \cdot \mathbf{x} = 0, i = 1,2,\ldots,r\}$, 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 combinações lineares dos vetores da lista $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$, considere:

- $\mathbf{0} \in \mathcal{S}$, pois $\mathbf{0} = 0\mathbf{a}_1 + 0\mathbf{a}_2 + \cdots + 0\mathbf{a}_r$;

[81] Cf. p. 23.

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

- $\alpha \mathbf{x} \in \mathcal{S}$, 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

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

  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

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

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

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

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

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

$$A = B \iff a_{ij} = b_{ij}, i = 1, \ldots, m, j = 1, \ldots, n.$$

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

$$A = \mathsf{O} \iff a_{ij} = 0, i = 1, \ldots, m, j = 1, \ldots, n.$$

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:

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

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:

$$D = \lambda A \iff d_{ij} := \lambda a_{ij}, i = 1, \ldots, m, j = 1, \ldots, n.$$

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, gerador, LI, LD, base, dimensão, etc., conforme ilustram alguns exemplos/exercícios desse capítulo. De fato, como veremos na seção 6.2, podemos fazer, de modo natural, as matrizes $m \times n$ corresponderem biunivocamente aos vetores de $\mathbb{R}^{mn}$. Basta considerarmos, por exemplo, a matriz A dada em (3.1), p. 51, e

 $$\mathbf{a} = (a_{11}, \ldots, a_{1n}, a_{21}, \ldots, a_{2n}, \ldots, a_{m1}, \ldots, a_{mn}),$$

 onde as mn coordenadas consecutivas desse vetor são as entradas consecutivas de $A(1,-), A(2,-), \ldots, A(m,-)$, nessa ordem. Essa correspondência preserva combinações lineares, isto é, caso a matriz A_i corresponda biunivocamente ao vetor $\mathbf{a}_i$ e α_i seja um escalar, $i = 1, \ldots, r$, a CL $\alpha_1 A_1 + \cdots + \alpha_r A_r$ corresponderá biunivocamente à CL $\alpha_1\mathbf{a}_1 + \cdots + \alpha_r\mathbf{a}_r$.

 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, ao considerarmos 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

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

podemos considerar a base canônica $\{E_1, E_2, E_3, E_4\}$ de $\mathbb{R}^{2\times 2}$ e, assim, qualquer matriz

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

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 $\mathcal{S}$ e satisfaça a seguinte condição:

$$\mathcal{S} \neq \mathcal{S}' \neq \mathbb{R}^{2\times 2}.$$

RESOLUÇÃO

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

(b) Seja A_3 uma matriz da base canônica de $\mathbb{R}^{2\times 2}$.[3] Assim, A_3 não pode ser escrita como uma CL de A_1 e A_2.[4]

(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 $\mathcal{S} \subset \mathcal{S}'$, $\mathcal{S} \neq \mathcal{S}'$ e A_1, A_2 e A_3 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

- Existe uma multiplicação “linha-por-coluna”, análoga ao produto interno visto no capítulo 2. De fato, 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, é definido por:

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

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.

[4] Verifique!

temos a matriz 2×3

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

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

EXEMPLOS

- Para as matrizes A e B do exemplo anterior, embora possamos calcular AB, o produto BA não está definido.

- Mesmo que tenhamos $A, B \in \mathbb{R}^{n\times n}$, em geral, $AB \neq BA$. De fato, considere, por exemplo,

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ e } B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

(∗) *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 as matrizes A e B do último exemplo e $C = A + B$.

- Caso A seja dada por (3.1),[5] a sua *transposta* é definida por

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

Note que, A^t é $n \times m$ e

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

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 tais que AB esteja bem definido, como esse produto de matrizes se comporta sob a ação da transposição? Pode ser demonstrado que*

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

[5] Cf. p. 51.

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, diremos que *A é de ordem n*, sua *diagonal principal* é composta pelas entradas a_{ii}, $i = 1, \ldots, n$, e, para $i \neq j$, a_{ij} é uma entrada *fora da diagonal principal* (de A).
- Uma matriz A é dita *simétrica* quando
$$A^t = A.$$
Portanto, não existem matrizes simétricas que não sejam quadradas e, pela definição de A^t, duas entradas de A "simétricas" em relação à diagonal principal são iguais, ou seja, para $i \neq j$, $a_{ij} = a_{ji}$.

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 sejam nulas, isto é,
$$\mathsf{D} = \begin{bmatrix} d_{11} & & & & \\ & d_{22} & & \text{zeros} & \\ & & \ddots & & \\ & \text{zeros} & & d_{n-1,n-1} & \\ & & & & d_{nn} \end{bmatrix}, \tag{3.2}$$
é dita *matriz diagonal*
- Se $d_{ii} = 1$, $i = 1, 2, \ldots, n-1, n$, então (3.2) é chamada de (*matriz*) *identidade* e denotada por I, ou seja,
$$\mathsf{I} = \begin{bmatrix} 1 & & & & \\ & 1 & & \text{zeros} & \\ & & \ddots & & \\ & \text{zeros} & & 1 & \\ & & & & 1 \end{bmatrix}.$$
- Caso A seja $m \times p$, I seja $p \times p$ e B seja $p \times n$, demonstra-se que
$$A\mathsf{I} = A \quad \text{e} \quad \mathsf{I}B = B, \tag{3.3}$$
isto é, I é um *elemento neutro multiplicativo*.

EXERCÍCIO

Verifique a validade de (3.3) para
$$A = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}, \ B = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \ \text{e} \ \mathsf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

As matrizes que têm "inversas", ditas "invertíveis", figuram entre as matrizes quadradas mais importantes. 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 representam 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 $A(-,j)$.[6]

  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:

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

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

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

· Se $n > 2$, então

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

OBSERVAÇÃO

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.

- Para $A \in \mathbb{R}^{n\times n}$, como $\det A = \det A^t$, $\det A$ pode ser calculado ao longo de qualquer coluna $A(-, j)$, ou seja,

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

EXERCÍCIO

Calcule $\det A$ ao longo de cada coluna da matriz A do primeiro exemplo dessa seção.

- Para simplificarmos o cálculo de $\det A$, podemos escolher alguma linha/coluna de A que tenha o maior número de zeros. 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

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

- Quando $\mathbf{b} = \mathbf{0}$, 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*.

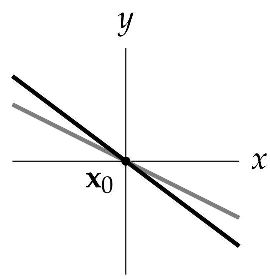
- Caso a variável x_j possa ser substituída pelo escalar fixo x_{0_j}, $j = 1, 2, \ldots, n$, diremos que o vetor $\mathbf{x} = \mathbf{x}_0$ *é uma solução de* $A\mathbf{x} = \mathbf{b}$ quando $A\mathbf{x}_0 = \mathbf{b}$.[7]

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, conforme a figura 3.1.

[7] O sistema homogêneo $A\mathbf{x} = \mathbf{0}$ admite, no mínimo, a *solução nula* $\mathbf{x}_0 = \mathbf{0}$.

Figura 3.1: Gráficos de $x + 2y = 0$ (em cinza) e $3x + 4y = 0$ (em negrito)



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

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

e o seu *conjunto solução* $\mathcal{S}$ 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 dada por

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

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

· 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, $i \neq j$:
  - $B(i,-) = A(i,-) + \alpha\, A(j,-)$,[8] com escalar $\alpha \neq 0$;
  - $B(i,-) = \alpha\, A(i,-)$,[9] com escalar $\alpha \neq 0$;
  - $B(i,-) = A(j,-)$, $B(j,-) = A(i,-)$.[10]

[8] 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.
[9] Isto é, a i-ésima linha de B é um múltiplo escalar da i-ésima linha de A.
[10] Isto é, estão trocadas, na matriz B, as linhas i e j de A.

Caso A seja equivalente à B, denota-se

$$A \longrightarrow B.^{11} \tag{3.5}$$

EXEMPLOS

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

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

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

- Demonstra-se que:

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

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

- Pelo *(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:

$\cdot$ $A_1 = A$;

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

$\cdot$ $A_r = \mathsf{R}$.

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:

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

[11] Essa notação pode vir acompanhada de uma das operações elementares supracitadas.

EXEMPLO

Seja A, por abuso de notação, a matriz aumentada do sistema (3.4).[12] Portanto,

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

$$B = \left[\begin{array}{cccc|c} 1 & 2 & -3 & 0 & 4 \\ 2 & 4 & -6 & 0 & 8 \\ 0 & 1 & -1 & 1 & -1 \end{array}\right] \underbrace{C(2,-) = B(2,-) - 2\,B(1,-)}_{\longrightarrow}$$

$$C = \left[\begin{array}{cccc|c} 1 & 2 & -3 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 \end{array}\right] \underbrace{D(2,-) = C(3,-),\, D(3,-) = C(2,-)}_{\longrightarrow}$$

$$D = \left[\begin{array}{cccc|c} 1 & 2 & -3 & 0 & 4 \\ 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \underbrace{E(1,-) = D(1,-) - 2\,D(2,-)}_{\longrightarrow}$$

$$E = \left[\begin{array}{cccc|c} 1 & 0 & -1 & -2 & 6 \\ 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] = \mathrm{R}.$$

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, caso $\mathcal{S}$ seja o conjunto solução dos sistemas (3.4) e (3.7), 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, que podem ser obtidas ao atribuirmos valores reais para α e β.[13]

- 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

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

[12] Cf. p. 61.

[13] Verifique que (3.8) é uma solução arbitrária de (3.4).

ou seja,

$$0 = b \neq 0,$$

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

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

cuja terceira linha, ou seja,

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

representa $0 = 1$, 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 uma 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 uma 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.

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

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

Nesse caso,

$$B = C.$$

DEMONSTRAÇÃO

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

Portanto, denotaremos B por A^{-1} e diremos 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 $AA^{-1} = \mathsf{I}$.[14] Assim,

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

EXEMPLO
Pelo produto das matrizes do primeiro exemplo dessa seção, $AA^{-1} = \mathsf{I}$.

- Caso A seja invertível, as seguintes propriedades são válidas:

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

· A^{-1} é invertível e $\left(A^{-1}\right)^{-1} = A$;[15]

· A^t é invertível e $\left(A^t\right)^{-1} = \left(A^{-1}\right)^t$.[16]

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

- Caso A e B sejam invertíveis, AB é invertível e $(AB)^{-1} = B^{-1}A^{-1}$.[17]

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 *elementar* caso seja equivalente à alguma I pela aplicação de *uma* operação elementar, ou seja,

$$\mathsf{I} \longrightarrow \mathsf{E}.^{18} \tag{3.11}$$

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

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,

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

[15]Demonstre!
[16]Idem.
[17]Idem.
[18]Cf. (3.5) e (3.6), pp. 64 e 64.

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

EXEMPLO

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

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.[19]

- 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).

[19] 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;[20]

   (c) se A e B são LI.[21]

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

---

[20] SUGESTÃO

Obtenha escalares α e β tais que $\alpha\mathbf{a}_1 + \beta\mathbf{b}_1 - \mathbf{c}_1$ seja o vetor nulo.

[21] RESOLUÇÃO

Caso A e B sejam LD, existe algum escalar λ tal que $A = \lambda B$. Logo, ao examinarmos a última entrada da última linha da matriz de cada membro dessa igualdade, temos $0 = \lambda \cdot (-1)$, ou seja, $\lambda = 0$. Assim, $A = \mathrm{O}$, que não é uma igualdade válida, 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,

$$\begin{aligned}[(A+B)C]_{ij} &= (A+B)(i,-)\cdot C(-,j)\\ &= (A(i,-)+B(i,-))\cdot C(-,j)\\ &= A(i,-)\cdot C(-,j)+B(i,-)\cdot C(-,j)\\ &= (AC)_{ij}+(BC)_{ij}\\ &= (AC+BC)_{ij}.\end{aligned}$$

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,

$$\begin{aligned}\left[(AB)^t\right]_{ij} &= (AB)_{ji}\\ &= A(j,-)\cdot B(-,i)\\ &= B(-,i)\cdot A(j,-)\\ &= \left(B^t\right)(i,-)\cdot\left(A^t\right)(-,j)\\ &= \left(B^tA^t\right)_{ij}.\end{aligned}$$

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 $\mathbf{b}_2$;[22]

(c) Escreva $\mathbf{b}_1$ e $\mathbf{b}_2$ como combinações lineares de $\mathbf{a}_1$ e $\mathbf{a}_2$;[23]

(d) Utilize os outros itens desse exercício para verificar que o mesmo plano de $\mathbb{R}^3$ é gerado pelos vetores linhas ou colunas de A.

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 $\mathbf{b}_2$ quanto $\mathbf{a}_1$ 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.

[22] SOLUÇÃO $\mathbf{a}_1 = \frac{11}{3}\mathbf{b}_1 - \frac{2}{3}\mathbf{b}_2$ e $\mathbf{a}_2 = \frac{10}{3}\mathbf{b}_1 - \frac{1}{3}\mathbf{b}_2$.

[23] 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 $\mathbf{v}_2 = (0,1,1)$.[24] Seja $\mathcal{S}_c$ o subespaço de $\mathbb{R}^3$ gerado por $\mathbf{e}_1 = (1,0,0)$ e $\mathbf{e}_2 = (0,1,0)$.[25] Como bases distintas podem gerar o mesmo subespaço, não podemos afirmar que:

$$\{\mathbf{v}_1, \mathbf{v}_2\} \neq \{\mathbf{e}_1, \mathbf{e}_2\} \Longrightarrow \mathcal{S}_\ell \neq \mathcal{S}_c.^{26}$$

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 uma 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.[27] 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 $B + C$.

[24] 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 $\mathcal{S}_\ell$.

[25] 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 $\mathbb{R}^3$.

[26] Cf. o exercício anterior.

[27] Cf. o exercício 8 dessa seção.

7. Calcule os seguintes determinantes:

(a) $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$; $\boxed{\text{RESPOSTA}}$ -2;

(b) $\begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}$; $\boxed{\text{RESPOSTA}}$ 3;

(c) $\begin{bmatrix} 0 & 3 & 1 & 28 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 2 & 71 \\ 2 & 36 & -9 & 3 \end{bmatrix}$; $\boxed{\text{RESPOSTA}}$ 36;

(d) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 5 & 2 & 0 & 0 \\ 8 & 6 & 3 & 0 \\ 0 & 9 & 7 & 4 \end{bmatrix}$; $\boxed{\text{RESPOSTA}}$ 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 $\det A = a_{11} \cdots a_{nn}$.[28]

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

Como $\det A \neq 0$, existe A^{-1}. 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. Caso

$$A = \begin{bmatrix} 1 & 1 & 0 & -1 \\ -1 & -1 & 3 & 4 \\ 2 & 2 & 3 & 1 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 0 & 0 & 3 & 3 \\ r & r & 3 & 1 \\ 1 & 1 & 6 & 5 \end{bmatrix}$$

[28] $\boxed{\text{SUGESTÃO}}$

Caso as entradas acima da diagonal principal sejam nulas, calcule determinantes, indutivamente, sempre ao longo das primeiras linhas, até que

$$\det A = a_{11}a_{22} \cdots a_{nn}$$

seja obtido. 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.

sejam equivalentes (por linha), isto é, elas sejam equivalentes à mesma escalonada reduzida R, determine r.

$\boxed{\text{Resolução}}$

Como

$$A \longrightarrow \begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \mathsf{R}$$

e

$$\begin{aligned} B \longrightarrow & \begin{bmatrix} 1 & 1 & 6 & 5 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 3-6r & 1-5r \end{bmatrix} \\ \longrightarrow & \begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & r-2 \end{bmatrix} = \mathsf{R}, \end{aligned}$$

$r = 2$.

12. Via escalonamento, determine o conjunto solução ($\mathcal{S}$) de cada um dos sistemas seguintes:

(a) $\begin{cases} 2x + y - 2z = 10 \\ 3x + 2y + 2z = 1 \\ 5x + 4y + 3z = 4 \end{cases}$ $\boxed{\text{Resposta}}$ $\mathcal{S} = \{(1, 2, -3)\}$;

(b) $\begin{cases} x + 2y - z = 0 \\ 2x - y + 3z = 0 \\ 4x + 3y + z = 0 \end{cases}$ $\boxed{\text{Resposta}}$ $\mathcal{S} = \{\alpha(-1, 1, 1) \mid \alpha \in \mathbb{R}\}$;

(c) $\begin{cases} x + y + z = 4 \\ 2x + 5y - 2z = 3 \\ x + 7y - 7z = 5 \end{cases}$ $\boxed{\text{Resposta}}$ $\mathcal{S} = \emptyset$;

(d) $\begin{cases} x - y + 2z - w = -1 \\ 2x + y - 2z - 2w = -2 \\ -x + 2y - 4z + w = 1 \\ 3x - 3w = -3 \end{cases}$

$\boxed{\text{Resposta}}$ $\mathcal{S} = \{\alpha(1, 0, 0, 1) + \beta(0, 2, 1, 0) + (-1, 0, 0, 0) \mid \alpha, \beta \in \mathbb{R}\}$.

13. Caso

$$\begin{cases} x_1 + 2x_2 + x_3 = a \\ -x_1 - x_2 - 2x_3 = b \\ -x_2 + x_3 = c \end{cases} \tag{3.13}$$

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).

$\boxed{\text{Resolução}}$

Ao escalonarmos a matriz aumentada de (3.13), temos

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & a \\ -1 & -1 & -2 & b \\ 0 & -1 & 1 & c \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & 1 & a \\ 0 & 1 & -1 & a+b \\ 0 & -1 & 1 & c \end{array}\right]$$
$$\longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 3 & -a-2b \\ 0 & 1 & -1 & a+b \\ 0 & 0 & 0 & a+b+c \end{array}\right].$$

Assim, (3.13) tem solução se, e somente se, $a+b+c=0$ e, nesse caso, (3.13) é equivalente ao sistema

$$\begin{cases} x_1 \phantom{{}+x_2} + 3x_3 = -a-2b; \\ \phantom{x_1+{}} x_2 - x_3 = a+b. \end{cases}$$

Portanto, para $x_3 = t$ arbitrário, a solução de (3.13) é dada por

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -3t-a-2b \\ t+a+b \\ t \end{bmatrix} = t\begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -a-2b \\ a+b \\ 0 \end{bmatrix},$$

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.[29] Além disso, obtenha a solução do sistema supracitado.[30]

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 as matrizes elementares da lista* $\mathsf{E}_1, \mathsf{E}_2, \ldots, \mathsf{E}_k$ *são obtidas, nessa ordem, nas* k *etapas do escalonamento de* A*, como em (3.12) da página 68.*

Nesse caso, para A quadrada e $\mathsf{R} = \mathsf{I}$, A é invertível e

$$A^{-1} = \mathsf{E}_k \cdots \mathsf{E}_2\mathsf{E}_1.$$

EXEMPLO

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix} = \mathsf{E}_1 A$$
$$\longrightarrow \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} = \mathsf{E}_2\mathsf{E}_1 A$$
$$\longrightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \mathsf{E}_3\mathsf{E}_2\mathsf{E}_1 A = I;$$

[29] RESPOSTA $a+b=c$.

[30] DICA Cf. a resolução do sistema (3.13) do exercício anterior.

$$\begin{aligned}
\mathsf{E}_3\mathsf{E}_2\mathsf{E}_1 &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} = A^{-1}.
\end{aligned}$$

Portanto, para $A \in \mathbb{R}^{n\times n}$ arbitrária, 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 que $[\mathsf{R}|A']$ seja obtida, digamos, após k operações elementares, ou seja, efetue o escalonamento

$$[A|\mathsf{I}] \longrightarrow [\mathsf{E}_1A|\mathsf{E}_1\mathsf{I}] \longrightarrow [\mathsf{E}_2\mathsf{E}_1A|\mathsf{E}_2\mathsf{E}_1] \longrightarrow \cdots \longrightarrow [\mathsf{E}_k\cdots\mathsf{E}_1A|\mathsf{E}_k\cdots\mathsf{E}_1] = [\mathsf{R}|A'] \,. \tag{3.14}$$

(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.[31]

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,

$$A = \begin{bmatrix} B & & \text{zeros} \\ & C & \\ \text{zeros} & & D \end{bmatrix}.$$

[31] 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:

$$A \text{ é invertível} \Longleftrightarrow \mathsf{R} = \mathsf{I}.$$

Por outro lado, ao utilizarmos a fórmula para inversas de matrizes 2×2,[32] temos

$$B^{-1} = \begin{bmatrix} 0 & \frac{1}{2} \\ 1 & 0 \end{bmatrix}, C^{-1} = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{4} \end{bmatrix} \text{ e } D^{-1} = \begin{bmatrix} \frac{1}{5} & -6 \\ 0 & 5 \end{bmatrix}.$$

Portanto,

$$\begin{aligned} A^{-1} &= \begin{bmatrix} B^{-1} & & \text{zeros} \\ & C^{-1} & \\ \text{zeros} & & D^{-1} \end{bmatrix} \\ &= \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{5} & -6 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix}, \end{aligned}$$

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:

[32] Cf. a seção 3.6.

· $a \neq 0$ 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.[33]

18. Prove que

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

é invertível se $a \neq 0$ e $a \neq b$. Nesse caso, determine A^{-1}.

RESOLUÇÃO

Podemos escalonar $[A|\mathsf{I}]$, via Gauss-Jordan, até obtermos $\left[\mathsf{I}|A^{-1}\right]$, com

$$A^{-1} = \frac{1}{a(a-b)} \begin{bmatrix} a & 0 & -b \\ -a & a & 0 \\ 0 & -a & a \end{bmatrix}.$$

De fato,

$$\left[\begin{array}{ccc|ccc} a & b & b & 1 & 0 & 0 \\ a & a & b & 0 & 1 & 0 \\ a & a & a & 0 & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|ccc} a & b & b & 1 & 0 & 0 \\ 0 & a-b & 0 & -1 & 1 & 0 \\ 0 & a-b & a-b & -1 & 0 & 1 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & b/a & b/a & 1/a & 0 & 0 \\ 0 & 1 & 0 & -1/(a-b) & 1/(a-b) & 0 \\ 0 & 1 & 1 & -1/(a-b) & 0 & 1/(a-b) \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & b/a & 1/(a-b) & -b/[a(a-b)] & 0 \\ 0 & 1 & 0 & -1/(a-b) & 1/(a-b) & 0 \\ 0 & 0 & 1 & 0 & -1/(a-b) & 1/(a-b) \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/(a-b) & 0 & -b/[a(a-b)] \\ 0 & 1 & 0 & -1/(a-b) & 1/(a-b) & 0 \\ 0 & 0 & 1 & 0 & -1/(a-b) & 1/(a-b) \end{array}\right],$$

onde, na segunda etapa do escalonamento, consideramos $a \neq 0$ e $a - b \neq 0$.

[33] RESPOSTA $a = 0$; $a = 2$; $a = 7$.

19. A inversa de qualquer matriz invertível A tal que
$$A^2 - 2A + 5\mathsf{I} = \mathsf{O}$$
é dada por
$$\frac{1}{5}(2\mathsf{I} - A),$$
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.

Utilize os itens anteriores (desse exercício) para calcular $\det A$, onde A é dada por:

(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

$$\boxed{\textit{A não é invertível} \Longleftrightarrow \det A = 0.}$$

Utilize essa equivalência para resolver (novamente) os exercícios 16 e 17 dessa seção.[34]

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 $AA^{-1} = I$.

[34] 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

$$\begin{array}{rccc} L_A & : & \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ & & \mathbf{x} & \longmapsto & L_A(\mathbf{x}), \end{array}$$

dita *multiplicação por A*.

EXEMPLO

Se $A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$, então

$$\begin{array}{rcccc} L_A & : & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ & & \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} & \longmapsto & L_A(\mathbf{x}) = \begin{bmatrix} x_1 - x_2 \\ x_2 - x_3 \end{bmatrix} \end{array}.$$

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,

$$\begin{aligned}
L(\mathbf{x}) &= L\left(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n\right) \\
&= x_1 L\left(\mathbf{e}_1\right) + x_2 L\left(\mathbf{e}_2\right) + \cdots + x_n L\left(\mathbf{e}_n\right) \\
&= x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n \\
&= x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \\
&= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} \\
&= A\mathbf{x} \\
&= L_A(\mathbf{x}).
\end{aligned} \tag{4.3}$$

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 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!

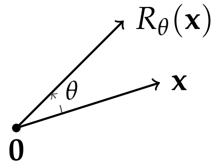
[3] Confira a subseção 4.3.1.

3. ROTAÇÃO (ANTI-HORÁRIA) DE ÂNGULO θ EM TORNO DA ORIGEM

$$R_\theta(x_1, x_2) = (x_1 \cos\theta - x_2 \operatorname{sen}\theta, x_1 \operatorname{sen}\theta + x_2 \cos\theta) \quad \forall (x_1, x_2) \in \mathbb{R}^2; \tag{4.4}$$

Para $\mathbf{x} = (x_1, x_2)$, podemos considerar uma ilustração de R_θ como a da figura 4.1.

Figura 4.1: Rotação de θ radianos do vetor $\mathbf{x}$ em torno de $\mathbf{0}$



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

$$\begin{aligned} A &= \begin{bmatrix} R_\theta(\mathbf{e}_1) & R_\theta(\mathbf{e}_2) \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}, \end{aligned}$$

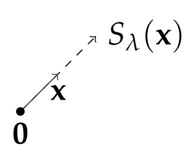
então $R_\theta = L_A$.[4] Portanto, R_θ é linear, pela afirmação 1 supracitada.

4. SEMELHANÇA (OU CISALHAMENTO) DE RAZÃO $\lambda \in \mathbb{R}$

$$S_\lambda(\mathbf{x}) = \lambda\mathbf{x} \quad \forall \mathbf{x} \in \mathbb{R}^n;^5$$

Para uma ilustração de S_λ, considere a figura 4.2.

Figura 4.2: Cisalhamento do vetor $\mathbf{x}$, $\lambda > 1$



RESOLUÇÃO

Como $\mathbf{x} = \mathsf{I}\mathbf{x}$, $S_\lambda(\mathbf{x}) = (\lambda\mathsf{I})\mathbf{x}$. Portanto, para $A = \lambda\mathsf{I}$,[6] $S_\lambda = L_A$ é linear, pela afirmação 1 supracitada.

[4] Verifique!

[5] Podemos assumir $\lambda \neq 0$. De fato, a *função nula*, embora linear (verifique!), não é uma semelhança. Note que $S_1 = I$ é a *função identidade* em $\mathbb{R}^n$.

[6] $A = \begin{bmatrix} k & & \\ \text{zeros} & \ddots & \text{zeros} \\ & & k \end{bmatrix}$.

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 $j = 1, 2, \ldots, r$,

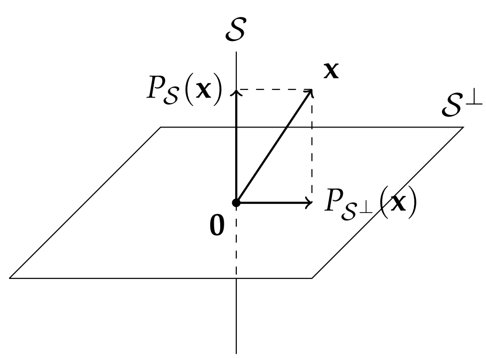
$$\begin{aligned}
\mathbf{x}''\cdot\mathbf{a}_j &= [\mathbf{x} - (\mathbf{x}\cdot\mathbf{a}_1)\,\mathbf{a}_1 - \cdots - (\mathbf{x}\cdot\mathbf{a}_r)\,\mathbf{a}_r]\cdot\mathbf{a}_j \\
&= \mathbf{x}\cdot\mathbf{a}_j - (\mathbf{x}\cdot\mathbf{a}_1)(\mathbf{a}_1\cdot\mathbf{a}_j) - \cdots - (\mathbf{x}\cdot\mathbf{a}_r)(\mathbf{a}_r\cdot\mathbf{a}_j) \\
&= \mathbf{x}\cdot\mathbf{a}_j - \mathbf{x}\cdot\mathbf{a}_j = 0.
\end{aligned}$$

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 *melhor aproximação de* $\mathbf{x}$ *em* $\mathcal{S}$ ou *vetor de* $\mathcal{S}$ *mais próximo de* $\mathbf{x}$, conforme veremos no capítulo 7.)

> 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 4.3, determine a matriz que representa $P_{\mathcal{S}}$ e a que representa $P_{\mathcal{S}^{\perp}}$.

Figura 4.3: Representação qualitativa das projeções ortogonais



Resolução parcial

Primeiramente, note que:

- $\{\mathbf{a}\}$ é uma base ortonormal de $\mathcal{S}$ se

$$\begin{aligned}
\mathbf{a} &= \frac{\mathbf{v}}{\|\mathbf{v}\|} \\
&= \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right);
\end{aligned}$$

[7] Confira a subseção 4.3.1.

· $(1,1,1)\cdot(x,y,z)=0$, 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 $\mathcal{S}$

$$R_{\mathcal{S}}(\mathbf{x})=2P_{\mathcal{S}}(\mathbf{x})-\mathbf{x} \quad \forall\mathbf{x}\in\mathbb{R}^n;$$ [10]

Para uma ilustração de $R_{\mathcal{S}}$ e $P_{\mathcal{S}}$, confira a figura 4.4.

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 $R_{\mathcal{S}^{\perp}}$.[11]

---

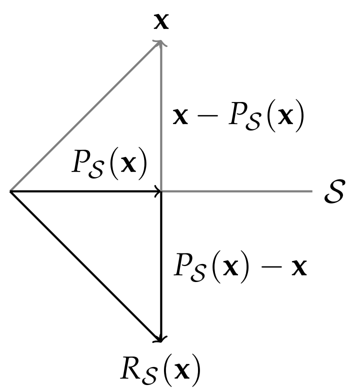
[8] Confira a subseção 4.3.1.

[9] Idem!

[10] Cf. (4.5), p. 89.

[11] Confira a subseção 4.3.1.

Figura 4.4: Relação entre reflexão e projeção



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.[12]
- Sem utilizar a afirmação 2 supracitada, obtenha a matriz C que representa $L_2 \circ L_1$, onde $\mathbf{x} = (x_1, x_2) \mapsto L_2\left(L_1(\mathbf{x})\right)$, e verifique que, de fato, $C = BA$.[13]

[12] Cf. a subseção 4.3.1.

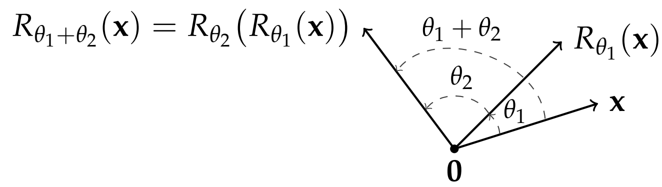
[13] Idem!

II. Para as rotações planares (4.4),[14] é fácil ver (geometricamente) que

$$R_{\theta_1+\theta_2} = R_{\theta_2} \circ R_{\theta_1},$$

conforme a a figura 4.5. Por outro lado, as matrizes que representam R_{θ_1}, R_{θ_2} e $R_{\theta_1+\theta_2}$

Figura 4.5: $R_{\theta_1+\theta_2} = R_{\theta_2} \circ R_{\theta_1}$



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.[15]

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

$$\mathrm{Nu}(L) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$$

[14] Cf. p. 88.

[15] DICA

Utilize o cosseno e o seno da soma de dois ângulos.

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),[16] 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

$$\begin{aligned} \mathbf{y} &= A\mathbf{x} \\ &= x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n, \end{aligned}$$

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,

$$\begin{aligned} \mathrm{Im}(L) \ni \mathbf{y} &= A\mathbf{x} \\ &= x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \end{aligned}$$

[16] Cf. p. 87.

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 $\operatorname{Im}(L) = \mathbb{R}^3$ é a canônica, com $\dim \operatorname{Im}(L) = 3$.

AFIRMAÇÃO 4

Chamando as dimensões de $\operatorname{Nu}(L)$ *e* $\operatorname{Im}(L)$ *de nulidade e posto de* L *(ou* A*), respectivamente, demonstra-se, pelo resultado (R14),*[17] *que*

$$\textit{nulidade} + \textit{posto} = n.$$

Para o exemplo anterior, temos

$$\text{nulidade} + \text{posto} = 1 + 3 = 4 = n.$$

MÉTODO PRÁTICO PARA DETERMINARMOS $\operatorname{Nu}(L)$ E $\operatorname{Im}(L)$

- Para determinarmos uma base de $\operatorname{Nu}(L)$, 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, $\operatorname{Nu}(A) = \operatorname{Nu}(\mathsf{R})$. 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 $\operatorname{Im}(L)$, 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 $\operatorname{Im}(L)$, basta coletarmos as *colunas pivôs*

$$\mathbf{a}_{j_1}, \mathbf{a}_{j_2}, \ldots, \mathbf{a}_{j_k} \tag{4.6}$$

de A.[18]

EXEMPLO

Ao escalonarmos

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

temos

$$\begin{aligned} A &\longrightarrow \begin{bmatrix} 1 & -1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 & 2 \end{bmatrix} \\ &\longrightarrow \begin{bmatrix} 1 & 0 & -1 & 1 & 3 \\ 0 & 1 & -1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \mathsf{R}. \end{aligned}$$

---

[17]Cf. p. 159.

[18]O posto de A é o número de colunas pivôs de A.

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

$$\begin{cases} x_1 \quad\quad - x_3 + x_4 + 3x_5 = 0, \\ \quad\quad x_2 - x_3 - x_4 + 2x_5 = 0, \end{cases}$$

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,

$$\text{nulidade} + \text{posto} = 3 + 2 = 5 = n.$$

$\boxed{\text{EXERCÍCIO}}$
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?[19] Escreva cada uma dessas colunas como CL das colunas pivôs de R.[20]

· Determine as colunas pivôs de A, ou seja, as colunas de A correspondentes às colunas pivôs de R.[21] 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.[22]

· Qual é a dimensão do núcleo de A, ou seja, a nulidade de A?[23]

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 esses vetores?
Essa questão pode ser respondida de dois modos. O primeiro é baseado no "método prático para determinarmos $\mathrm{Im}(L)$", 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 $(3,-3,1,1,-2,3)$.

[19] RESPOSTA $\mathbf{r}_i$, $i = 2,4,5,6$.

[20] 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 $\mathbf{r}_6 = \mathbf{r}_1 + \mathbf{r}_3$.

[21] RESPOSTA $\mathbf{a}_1$ e $\mathbf{a}_3$.

[22] RESPOSTA Na penúltima nota de rodapé, troque $\mathbf{r}$ por $\mathbf{a}$.

[23] RESPOSTA nulidade $= 6 - 2 = 4$.

$\boxed{\text{PRIMEIRO MODO}}$

$$A = \begin{bmatrix} 1 & -1 & 1 & 3 \\ -1 & 1 & -1 & -3 \\ 1 & 0 & 2 & 1 \\ -1 & -1 & -3 & 1 \\ 0 & 1 & 1 & -2 \\ 1 & -1 & 1 & 3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -1 & 1 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & -2 \\ 0 & -2 & -2 & 4 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
$$\longrightarrow \begin{bmatrix} 1 & -1 & 1 & 3 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
$$\longrightarrow \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \mathsf{R}.$$

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*,[24] $\{\mathsf{R}(1, -), \mathsf{R}(2, -)\}$ 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

$$A = [L(\mathbf{e}_1) \;\; L(\mathbf{e}_2) \;\; \cdots \;\; L(\mathbf{e}_n)]\,,$$

[24] Ou seja, cada linha de A é uma CL das linhas não nulas de R. O leitor é convidado a explicitar essas combinações lineares.

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 $\mathbb{R}^m$:

$$[L]^{\mathcal{B}}_{\mathcal{B}'} := \left[L(\mathbf{x}_1) \;\; L(\mathbf{x}_2) \;\; \cdots \;\; L(\mathbf{x}_n)\right],$$

onde cada coluna $[L]^{\mathcal{B}}_{\mathcal{B}'}(-,j)$ é dada por

$$\begin{aligned} L(\mathbf{x}_j) &= \begin{bmatrix} \ell_{1j} \\ \ell_{2j} \\ \vdots \\ \ell_{mj} \end{bmatrix} \\ &= \ell_{1j}\mathbf{y}_1 + \ell_{2j}\mathbf{y}_2 + \cdots + \ell_{mj}\mathbf{y}_m. \end{aligned}$$

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 $\mathbb{R}^2$.[25] 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 $[L]_{\mathcal{B}}$.

[25] 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 $\mathbf{y}_2 = (1,-1)$.

EXERCÍCIO

Se $m = n = 2$, $\mathcal{B}$ é a base canônica,[26] $\mathcal{B}' = \{\mathbf{y}_1 = (1,1), \mathbf{y}_2 = (1,-1)\}$ 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

$$\begin{aligned} A &= [L]_{\mathcal{B}} \\ &= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}. \end{aligned}$$

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, pois, pelo resultado (R15) da página 162, $\mathbf{x} \mapsto L(\mathbf{x})$ 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 $\mathbf{x} = (1,2)$.

[26]Isto é, $\mathbf{x}_1 = \mathbf{e}_1$ e $\mathbf{x}_2 = \mathbf{e}_2$.

RESOLUÇÃO

Como $L(\mathbf{x}) = (-1,3) = 1 \cdot \mathbf{y}_1 - 2\mathbf{y}_2$, $[L(\mathbf{x})]_{\mathcal{B}'} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$. Por outro lado, $[L]_{\mathcal{B}'}^{\mathcal{B}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ e $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Agora, basta verificarmos 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 156. 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,

$$S(\mathbf{x}_j) = T(\mathbf{x}_j)$$

para $j = 1, 2, \ldots, n$. Então, para todo $\mathbf{x} \in \mathbb{R}^n$, pela linearidade de S e T, temos

$$\begin{aligned} S(\mathbf{x}) &= S(x_1\mathbf{x}_1 + x_2\mathbf{x}_2 + \cdots + x_n\mathbf{x}_n) \\ &= x_1 S(\mathbf{x}_1) + x_2 S(\mathbf{x}_2) + \cdots + x_n S(\mathbf{x}_n) \\ &= x_1 T(\mathbf{x}_1) + x_2 T(\mathbf{x}_2) + \cdots + x_n T(\mathbf{x}_n) \\ &= T(x_1\mathbf{x}_1 + x_2\mathbf{x}_2 + \cdots + x_n\mathbf{x}_n) \\ &= T(\mathbf{x}), \end{aligned}$$

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

$$T^{-1}(\mathbf{y}_i) = \mathbf{x}_i, \quad i = 1, 2, \ldots, n,$$

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 dos dois exercícios anteriores, verifique a validade de (4.11).

RESOLUÇÃO

Como em (4.9), seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ a função linear

$$\begin{array}{ccc} \mathbf{x}_1 & \overset{T}{\mapsto} & \mathbf{y}_1; \\ \mathbf{x}_2 & \overset{T}{\mapsto} & \mathbf{y}_2. \end{array}$$

Portanto,

$$\begin{array}{ccc} (1,0) & \overset{T}{\mapsto} & (1,1) = 1(1,0) + 1(0,1); \\ (0,1) & \overset{T}{\mapsto} & (1,-1) = 1(1,0) - 1(0,1). \end{array}$$

Note que T é uma multiplicação por P, onde

$$\begin{aligned} \mathsf{P} &= [T]_{\mathcal{B}} \\ &= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \end{aligned}$$

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

Considere $m = n = 2$, $\mathcal{B} = \{\mathbf{x}_1 = (1,0), \mathbf{x}_2 = (1,1)\}$, $\mathcal{B}' = \{\mathbf{y}_1 = (0,1), \mathbf{y}_2 = (-1,1)\}$ 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

$$\left\{ \begin{array}{lclcl} L(\mathbf{x}_1) & = & (0,1) & = & -1 \cdot \mathbf{x}_1 + 1 \cdot \mathbf{x}_2, \\ L(\mathbf{x}_2) & = & (1,1) & = & 0 \cdot \mathbf{x}_1 + 1 \cdot \mathbf{x}_2, \end{array} \right.$$

temos

$$\begin{aligned} A &= [L]_{\mathcal{B}} \\ &= \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}. \end{aligned}$$

Como

$$\left\{ \begin{array}{lclcl} L(\mathbf{y}_1) & = & (1,0) & = & 1 \cdot \mathbf{y}_1 - 1 \cdot \mathbf{y}_2, \\ L(\mathbf{y}_2) & = & (1,-1) & = & 0 \cdot \mathbf{y}_1 - 1 \cdot \mathbf{y}_2, \end{array} \right.$$

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,

$$\mathsf{P}^{-1} = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix}.$$ [27]

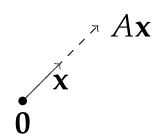
4.2 Autovalores e autovetores

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, $L = L_A$.[28] Um escalar λ é chamado *autovalor de A* (ou L), caso exista um vetor não 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 *autovetor de A* (ou L) *associado à* (ou *ao autovalor*) λ. (Cf. a figura 4.6.)

Figura 4.6: Autovetor $\mathbf{x}$ de A associado à $\lambda > 1$



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,

$$\begin{aligned} A\mathbf{x}_2 &= \begin{bmatrix} -2 \\ -2 \end{bmatrix} \\ &= \lambda_2\mathbf{x}_2. \end{aligned}$$

[27] Verifique a validade de (4.11) para esse exercício.
[28] Cf. pp. 87 e 97.

> *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 $\lambda_1 = 1$

$$(A - \mathsf{I})\,\mathbf{x} = \mathbf{0} \Longleftrightarrow \begin{bmatrix} 4 & 8 & 16 \\ 4 & 0 & 8 \\ -4 & -4 & -12 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$
$$\Longleftrightarrow x_1 + 2x_3 = 0 \quad \text{e} \quad x_2 + x_3 = 0.$$

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 $\lambda_2 = -3$

$$(A + 3\mathsf{I})\,\mathbf{x} = \mathbf{0} \Longleftrightarrow \begin{bmatrix} 8 & 8 & 16 \\ 4 & 4 & 8 \\ -4 & -4 & -8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$
$$\Longleftrightarrow x_1 + x_2 + 2x_3 = 0.$$

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 134.*

(4) *Além de reais, autovalores e coordenadas de autovetores também podem ser números complexos. Para um exemplo, confira (5.2), p. 134.*

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 $\mathcal{B}' = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ de $\mathbb{R}^n$ composta de autovetores de A. Nesse caso,

$$\mathbf{x}_i \overset{L}{\mapsto} A\mathbf{x}_i = \lambda_i \mathbf{x}_i, \quad i = 1, 2, \ldots, n.$$

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*
>
> $$\mathsf{P}^{-1}A\mathsf{P} = \mathsf{D}.$$
>
> *Nesse caso, ainda que 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, $\mathsf{P} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix}$.

De fato, como $\mathsf{P} = [T]_{\mathcal{B}}$ para $\mathbf{e}_j \overset{T}{\mapsto} \mathbf{x}_j$, $j = 1, \ldots, n$,[29] temos

$$\mathsf{P} = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix}$$

onde $T(\mathbf{e}_j) = \mathbf{x}_j$, $j = 1, 2, \ldots, n$.[30]

EXERCÍCIOS

[29] Cf. (4.9), (4.10) e (4.11), p. 100.

[30] Confira o início da subseção 4.1.2.

1. Para o exemplo 2×2 anterior,

$$\begin{aligned} A &= \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix}, \\ \mathsf{P} &= \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 \end{bmatrix} \\ &= \begin{bmatrix} -4 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{e} \\ \mathsf{D} &= \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}. \end{aligned}$$

   Verifique que

$$\mathsf{P}^{-1} = \begin{bmatrix} -1/5 & 1/5 \\ 1/5 & 4/5 \end{bmatrix} \text{ e } \mathsf{P}^{-1}A\mathsf{P} = \mathsf{D}.$$

2. Para o exemplo 3×3 anterior,

$$\begin{aligned} A &= \begin{bmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{bmatrix}, \\ \mathsf{P} &= \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \end{bmatrix} \\ &= \begin{bmatrix} -2 & -1 & -2 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{e} \\ \mathsf{D} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{bmatrix}. \end{aligned}$$

   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

$$A = \begin{bmatrix} 0,75 & 0,5 \\ 0,25 & 0,5 \end{bmatrix}$$

   para obter A^{100}.[31]

RESOLUÇÃO DO EX. 3

A^{100}, sem diagonalização, pode ser obtida calculando-se A^n, $n = 2, 3, \ldots, 100$, ou seja,

$$\overbrace{\begin{bmatrix} 0,6875 & 0,625 \\ 0,3125 & 0,375 \end{bmatrix}}^{A^2=AA}, \overbrace{\begin{bmatrix} 0,671875 & 0,65625 \\ 0,328125 & 0,34375 \end{bmatrix}}^{A^3=AA^2}, \text{ etc.}$$

Existe um alto custo computacional nesse procedimento. Contudo, caso A seja diagonalizável,

$$A = \mathsf{PDP}^{-1},\ A^2 = AA = \mathsf{PD}^2\mathsf{P}^{-1},\ A^3 = AA^2 = \mathsf{PD}^3\mathsf{P}^{-1},\ \ldots,\ A^{100} = \mathsf{PD}^{100}\mathsf{P}^{-1}.$$

[31] Aqui, A^{100} representa a centésima potência de A.

Além disso, caso λ_1 e λ_2 sejam 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

$$\begin{aligned} \det(A - \lambda \mathsf{I}) &= \begin{vmatrix} \frac{3}{4} - \lambda & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} - \lambda \end{vmatrix} \\ &= \lambda^2 - \frac{5}{4}\lambda + \frac{1}{4} \\ &= (\lambda - 1)\left(\lambda - \frac{1}{4}\right) \\ &= 0. \end{aligned}$$

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,

$$\begin{aligned} \begin{bmatrix} -1/4 & 1/2 \\ 1/4 & -1/2 \end{bmatrix} &\longrightarrow \begin{bmatrix} 1/4 & -1/2 \\ 0 & 0 \end{bmatrix} \\ &\longrightarrow \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

Analogamente, para obtermos um autovetor $\mathbf{x}_2$ de A associado à $\lambda_2 = \frac{1}{4}$, podemos escalonar a matriz do sistema $\left(A - \frac{1}{4}\mathsf{I}\right)\mathbf{x}_2 = \mathbf{0}$, ou seja,

$$\begin{aligned} \begin{bmatrix} 1/2 & 1/2 \\ 1/4 & 1/4 \end{bmatrix} &\longrightarrow \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\ &\longrightarrow \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

Portanto, para $\mathbf{x}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ 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 $\mathsf{P} = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}$, $\mathsf{P}^{-1} = \begin{bmatrix} 1/3 & 1/3 \\ -1/3 & 2/3 \end{bmatrix}$ e

$$\begin{aligned} A^{100} &= \mathsf{P} \begin{bmatrix} 1^{100} & 0 \\ 0 & (1/4)^{100} \end{bmatrix} \mathsf{P}^{-1} \\ &\approx \mathsf{P} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \mathsf{P}^{-1} = \begin{bmatrix} 2/3 & 2/3 \\ 1/3 & 1/3 \end{bmatrix}, \end{aligned}$$

ou seja,

$$A^{100} \approx \begin{bmatrix} 0,\overline{6} & 0,\overline{6} \\ 0,\overline{3} & 0,\overline{3} \end{bmatrix}.$$

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 P invertível, com entradas reais e inversa igual à sua transposta é dita *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.[32]

AFIRMAÇÃO 6

A ortogonalidade de P é equivalente à ortonormalidade de suas colunas/linhas, ou seja, as três afirmações seguintes são equivalentes entre si:

1. *P é ortogonal;*
2. *$\{P(-,1),\ldots,P(-,n)\}$ é uma base ortonormal de $\mathbb{R}^n$;*
3. *$\{P(1,-),\ldots,P(n,-)\}$ é uma base ortonormal de $\mathbb{R}^n$.*

EXEMPLOS

Confira os três primeiros exemplos dessa subseção.

DEMONSTRAÇÃO DA AFIRMAÇÃO 6

Como

$$P^tP = \begin{bmatrix} P^t(1,-)\cdot P(-,1) & \cdots & P^t(1,-)\cdot P(-,n) \\ \vdots & \vdots & \vdots \\ P^t(n,-)\cdot P(-,1) & \cdots & P^t(n,-)\cdot P(-,n) \end{bmatrix},^{33}$$

$$P^tP = \mathsf{I} \iff P(-,i)\cdot P(-,j) = \begin{cases} 1 \text{ se } i = j \\ \quad\text{e} \\ 0 \text{ se } i \neq j \end{cases} \quad \text{(produto interno de vetores em } \mathbb{R}^n\text{)}$$

$$\iff \text{as colunas de } P \text{ formam uma base ortonormal de } \mathbb{R}^n.$$

Analogamente,

$$PP^t = \mathsf{I} \iff \text{as linhas de } P \text{ formam uma base ortonormal de } \mathbb{R}^n.$$

AFIRMAÇÃO 7

Demonstra-se que:

1. *A é uma matriz 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, $A^t = A$.*[34]
2. *A é simétrica $\Longrightarrow$ os autovalores de A são reais.*
3. *A é simétrica $\Longrightarrow$ os autovetores associados à eventuais autovalores distintos de A são ortogonais.*

[32] Para o último exemplo, confira o exercício 10 da seção 2.7.

[33] A entrada de índice ij dessa matriz é o "produto" da linha i de P^t pela coluna j de P, $i,j = 1,\ldots,n$, conforme a mnemônica do produto de matrizes. Além disso, note que

$$P^t(i,-) = P(-,i),\, i = 1,\ldots,n.$$

[34] Esse resultado é conhecido como *teorema espectral* para $A \in \mathbb{R}^{n\times n}$.

DEMONSTRAÇÃO DA AFIRMAÇÃO 7

1. A validade desse resultado (para operadores sobre espaços vetoriais reais) será demonstrada na subseção 7.4.7. Contudo, para a demonstração do item 1 "não passar em branco", note que a implicação "$\Longrightarrow$" segue de
$$\begin{aligned} A &= \mathsf{PDP}^t \\ &= \mathsf{PD}^t\mathsf{P}^t \\ &= \left(\mathsf{PDP}^t\right)^t \\ &= A^t. \end{aligned}$$
Para a implicação "$\Longleftarrow$", confira o teorema da subseção 5.3.2.
2. De fato, se $A\mathbf{x} = \lambda\mathbf{x}$, então
$$\begin{aligned} \lambda\overline{\mathbf{x}}^t\mathbf{x} &= \overline{\mathbf{x}}^t\lambda\mathbf{x} \\ &= \overline{\mathbf{x}}^t A\mathbf{x} \\ &= \overline{\mathbf{x}}^t A^t\mathbf{x} \\ &= (A\overline{\mathbf{x}})^t\,\mathbf{x} \\ &= \left(\overline{A}\overline{\mathbf{x}}\right)^t\mathbf{x} \\ &= \left(\overline{A\mathbf{x}}\right)^t\mathbf{x} \\ &= \left(\overline{\lambda\mathbf{x}}\right)^t\mathbf{x} \\ &= \left(\overline{\lambda}\overline{\mathbf{x}}\right)^t\mathbf{x} \\ &= \overline{\lambda}\overline{\mathbf{x}}^t\mathbf{x}, \end{aligned}$$
pelas propriedades de transposição de matrizes e conjugação complexa.[35] Portanto, $\left(\lambda - \overline{\lambda}\right)\overline{\mathbf{x}}^t\mathbf{x} = 0$ e, como $\mathbf{x} \neq \mathbf{0}$, $\lambda - \overline{\lambda} = 0$, 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 "$\cdot$", e produto de matrizes, quando não usamos "$\cdot$". Então, $(\lambda_1 - \lambda_2)(\mathbf{x}_1\cdot\mathbf{x}_2) = 0$. Assim, para $\lambda_1 \neq \lambda_2$, $\mathbf{x}_1\cdot\mathbf{x}_2 = 0$.

EXEMPLOS

· Considere a matriz $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. Por ser simétrica, A é diagonalizável.[36] Contudo, P não é ortogonal, ou seja, $\mathsf{P}^{-1} \neq \mathsf{P}^t$, caso seja calculada como na subseção 4.2.1.[37]

[35] Confira o capítulo 5.
[36] Confira o item 1 da afirmação 7.
[37] Verifique!

Assim, para podermos utilizar a afirmação 7, P^{-1} não será obtida por escalonamento. Para calcularmos os autovalores de A, observe que

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

Portanto, esses autovalores são $\lambda_1 = -1$ e $\lambda_2 = 3$, o que acarreta

$$\mathsf{D} = \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}.$$

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 $\lambda_2 = 3$.[38] 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.[39] 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

$$\det \begin{bmatrix} 2-\lambda & 2 & 1 \\ 2 & 5-\lambda & 2 \\ 1 & 2 & 2-\lambda \end{bmatrix} = (\lambda - 7)(\lambda - 1)^2.$$

[38] 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.

[39] Todavia, a opção do método a ser utilizado é sua!

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

$$(A - 7\mathsf{I})\mathbf{x} = \mathbf{0}, \text{ ou seja, } \begin{bmatrix} -5 & 2 & 1 \\ 2 & -2 & 2 \\ 1 & 2 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

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 $\lambda_2 = \lambda_3 = 1$.[40] 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

$$\mathsf{P} = \begin{bmatrix} \mathbf{x}_1' & \mathbf{x}_2'' & \mathbf{x}_3'' \end{bmatrix} \Longrightarrow \mathsf{P}^t A \mathsf{P} = \mathsf{D}.$$

[40]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,

$$\begin{aligned} L(\mathbf{e}_1) &= L\left(\frac{1}{5}\mathbf{u}\right) \\ &= \frac{1}{5}L(\mathbf{u}) \qquad \text{(LINEARIDADE)} \\ &= \begin{bmatrix} 1/5 \\ 2/5 \end{bmatrix}. \end{aligned}$$

Por outro, como

$$\begin{aligned} L(\mathbf{v}) &= L(6\mathbf{e}_1 + 7\mathbf{e}_2) \\ &= 6L(\mathbf{e}_1) + 7L(\mathbf{e}_2) \quad (\text{LINEARIDADE}) \\ &= \begin{bmatrix} 6/5 \\ 12/5 \end{bmatrix} + 7L(\mathbf{e}_2), \end{aligned}$$

$$\begin{aligned} L(\mathbf{e}_2) &= \frac{1}{7}\left\{ \begin{bmatrix} 3 \\ 4 \end{bmatrix} - \begin{bmatrix} 6/5 \\ 12/5 \end{bmatrix} \right\} \\ &= \begin{bmatrix} 9/35 \\ 8/35 \end{bmatrix}. \end{aligned}$$

Então,

$$\begin{aligned} A &= \begin{bmatrix} L(\mathbf{e}_1) & L(\mathbf{e}_2) \end{bmatrix} \\ &= \begin{bmatrix} 1/5 & 9/35 \\ 2/5 & 8/35 \end{bmatrix}. \end{aligned}$$

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.[41] 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.

[41] Cf. (4.6), p. 94.

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 $(0,-1,1,4)$.[42]

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 $[L(\mathbf{x})]_{\mathcal{B}'}$.[43]

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 $\mathsf{P} = [T]_{\mathcal{B}}$.[44] 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

$$\mathsf{P} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix},$$

[42] Faça como no exercício que precede a subseção 4.1.2. Observe que, dos quatro vetores dados, os dois primeiros determinam a base procurada.

[43] Basta utilizarmos a fórmula (4.8), da página 99, 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.

[44] Cf. (4.9), (4.10) e (4.11), p. 100.

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

$$[M|\mathsf{I}] = \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & -1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$
$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & -1 & 1 & -1 & 1 & 0 \end{array}\right]$$
$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right] = [\mathsf{I}|\mathsf{P}]$$

resolve (simultaneamente) os três sistemas supracitados,[45] ou seja, determina P, e (sem qualquer esforço adicional) calcula a inversa de P.[46] 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,

$$\begin{bmatrix} 4-\lambda_1 & \sqrt{3} \\ \sqrt{3} & 2-\lambda_1 \end{bmatrix} = \begin{bmatrix} 3 & \sqrt{3} \\ \sqrt{3} & 1 \end{bmatrix}$$
$$\longrightarrow \begin{bmatrix} 3 & \sqrt{3} \\ 3 & \sqrt{3} \end{bmatrix}$$
$$\longrightarrow \begin{bmatrix} 1 & \frac{\sqrt{3}}{3} \\ 0 & 0 \end{bmatrix}.$$

Os autovetores associados à $\lambda_2 = 5$ são múltiplos de

$$\mathbf{x}_2 = \begin{bmatrix} \sqrt{3} \\ 1 \end{bmatrix}.$$

[45] 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 $(t_{13}, t_{23}, t_{33}) = (-1, 1, 1)$.

[46] $\mathsf{P}^{-1} = M$.

De fato,

$$\begin{aligned}
\begin{bmatrix} 4-\lambda_2 & \sqrt{3} \\ \sqrt{3} & 2-\lambda_2 \end{bmatrix} &= \begin{bmatrix} -1 & \sqrt{3} \\ \sqrt{3} & -3 \end{bmatrix} \\
&\longrightarrow \begin{bmatrix} 1 & -\sqrt{3} \\ 1 & -\sqrt{3} \end{bmatrix} \\
&\longrightarrow \begin{bmatrix} 1 & -\sqrt{3} \\ 0 & 0 \end{bmatrix}.
\end{aligned}$$

10. Determine 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,

$$\begin{aligned}
\begin{bmatrix} 2-\lambda_1 & 5 \\ 3 & -4-\lambda_1 \end{bmatrix} &= \begin{bmatrix} 3-2\sqrt{6} & 5 \\ 3 & -\left(3+2\sqrt{6}\right) \end{bmatrix} \\
&\longrightarrow \begin{bmatrix} 1 & \frac{5}{3-2\sqrt{6}} \\ 3 & -\left(3+2\sqrt{6}\right) \end{bmatrix} \\
&\longrightarrow \begin{bmatrix} 1 & \frac{5}{3-2\sqrt{6}} \\ 0 & 0 \end{bmatrix}.
\end{aligned}$$

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,

$$\begin{aligned}
\begin{bmatrix} 2-\lambda_2 & 5 \\ 3 & -4-\lambda_2 \end{bmatrix} &= \begin{bmatrix} 3+2\sqrt{6} & 5 \\ 3 & -\left(3-2\sqrt{6}\right) \end{bmatrix} \\
&\longrightarrow \begin{bmatrix} 1 & \frac{5}{3+2\sqrt{6}} \\ 3 & -\left(3-2\sqrt{6}\right) \end{bmatrix} \\
&\longrightarrow \begin{bmatrix} 1 & \frac{5}{3+2\sqrt{6}} \\ 0 & 0 \end{bmatrix}.
\end{aligned}$$

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.[47]
- A é diagonalizável? Justifique (corretamente) a sua resposta.

RESOLUÇÃO

Pela equação (4.12),[48] $A\mathbf{v}_1 = \lambda_1\mathbf{v}_1$, $A\mathbf{v}_2 = \lambda_2\mathbf{v}_2$ e $A\mathbf{v}_3 = \lambda_3\mathbf{v}_3$. Assim, se $\mathbf{a}_1 = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right)$,[49] então

$$\begin{aligned} \mathbf{a}_1 \cdot \mathbf{v}_1 &= \text{1a. coordenada de } \lambda_1\mathbf{v}_1, \\ \mathbf{a}_1 \cdot \mathbf{v}_2 &= \text{1a. coordenada de } \lambda_2\mathbf{v}_2 \text{ e} \\ \mathbf{a}_1 \cdot \mathbf{v}_3 &= \text{1a. coordenada de } \lambda_3\mathbf{v}_3. \end{aligned}$$

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 $\mathbb{R}^3$.[50] De fato, os três vetores desse conjunto são LI, pela não nulidade do determinante da matriz com linhas (ou colunas) dadas por esses vetores.[51] Outra maneira de demonstrarmos a independência linear dos vetores supracitados é utilizarmos o resultado (R20) do capítulo 7.[52] 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,

$$\det(A - \lambda\mathsf{I}) = (2-\lambda)(1+\lambda)^2,$$

onde I é a matriz identidade de ordem três. 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

$$(A - 2\mathsf{I})\mathbf{x} = \mathbf{0} \;\text{ e }\; (A + \mathsf{I})\mathbf{x} = \mathbf{0}.$$

---

[47] SUGESTÃO

Não obtenha o polinômio característico de A. Utilize a definição de autovalores.

[48] Cf. p. 102.

[49] $\mathbf{a}_i$ representa a i-ésima linha de A, $i = 1, 2, 3$.

[50] Cf. a primeira sentença da subseção 4.2.1.

[51] Cf. o penúltimo exercício do capítulo 3.

[52] Cf. p. 182.

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) $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$;[53]

(b) $A = \begin{bmatrix} -1 & 6 & -12 \\ 0 & -13 & 30 \\ 0 & -9 & 20 \end{bmatrix}$;[54]

---

[53] 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.

[54] DICA
Em relação a A, verifique que -1, 2 e 5 são seus autovalores.

(c) $A = \begin{bmatrix} 5 & -2 & 0 \\ -2 & 6 & 2 \\ 0 & 2 & 7 \end{bmatrix}$.[55]

14. Em (R20),[56] 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.[57]

RESOLUÇÃO

Como estamos lidando com uma matriz triangular,

$$\det(A - \lambda \mathsf{I}) = (\lambda - (D+2))(\lambda - (E+3))^2(\lambda - (C+1)),$$

onde I é a matriz identidade de ordem quatro. 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. Por (R20), já temos (pelo menos) três autovetores LI em $\mathcal{B}$: 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).[58] 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 $D - E = 1$;

---

[55] DICA
Em relação a A, verifique que 3, 6 e 9 são seus autovalores.

[56] Cf. p. 182.

[57] Confira a nota de rodapé do enunciado do exercício 2 da subseção 7.2.7.

[58] 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.

$$\cdot \begin{bmatrix} 1 & (C+1)/(D-E-1) & (C+1)/(D-E-1) & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{caso } D-E \neq 1.$$

Assim, a matriz do sistema terá posto 2 apenas se $x = 0$.[59]

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,[60] x deve ser igual a zero.

15. Caso $A \in \mathbb{R}^{n \times n}$ seja *idempotente*, isto é, $A^2 = A$, verifique que 0 ou 1 são os (únicos) autovalores de A.[61]

RESOLUÇÃO

Seja $\mathbf{x}$ um autovetor de A associado ao autovalor λ. Como

$$A\mathbf{x} = \lambda\mathbf{x}, \tag{4.14}$$

$$A^2\mathbf{x} = \lambda^2\mathbf{x}. \tag{4.15}$$

De fato,

$$\begin{aligned} A^2\mathbf{x} &= AA\mathbf{x} \\ &= A(\lambda\mathbf{x}) \\ &= \lambda A\mathbf{x} \\ &= \lambda\lambda\mathbf{x} \\ &= \lambda^2\mathbf{x}. \end{aligned}$$

Como $A = A^2$,

$$A\mathbf{x} = A^2\mathbf{x}. \tag{4.16}$$

Por (4.14), (4.15) e (4.16),

$$\begin{aligned} \lambda\mathbf{x} = \lambda^2\mathbf{x} &\Longrightarrow \left(\lambda - \lambda^2\right)\mathbf{x} = \mathbf{0} \\ &\Longrightarrow \lambda - \lambda^2 = 0 \text{ (pois } \mathbf{x} \neq \mathbf{0}) \\ &\Longrightarrow \lambda(1-\lambda) = 0 \\ &\Longrightarrow \lambda = 0 \text{ ou } \lambda = 1. \end{aligned}$$

OBSERVAÇÃO

Caso o enunciado do exercício não tivesse a frase "os (únicos)", teríamos a seguinte resolução alternativa: Como

$$A = A^2 \Longleftrightarrow A - A^2 = \mathsf{O}$$

e o determinante do produto de matrizes é igual ao produto dos determinantes dessas matrizes,

$$\begin{aligned} \det\left(A - A^2\right) = \det \mathsf{O} &\Longrightarrow \det(A(A-\mathsf{I})) = 0 \\ &\Longrightarrow \det A \cdot \det(A-\mathsf{I}) = 0 \\ &\Longrightarrow \det(A - 0\mathsf{I}) \cdot \det(A - 1\mathsf{I}) = 0 \\ &\Longrightarrow \det(A - 0\mathsf{I}) = 0 \text{ ou } \det(A - 1\mathsf{I}) = 0 \\ &\Longrightarrow 0 \text{ ou } 1 \text{ são autovalores de } A. \end{aligned}$$

[59] 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.

[60] Lembre-se que, para qualquer matriz $m \times n$, posto + nulidade = n.

[61] SUGESTÃO

Relacione a definição de autovalores/autovetores, dada na página 102, com a idempotência de A.

16. Caso $A \in \mathbb{R}^{n\times n}$ seja arbitrária e A^2 tenha um autovalor não negativo λ^2, demonstre que $\pm\lambda$ são 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. Assim, ao menos um dos dois fatores do último produto, de cima para baixo, deve ser nulo. Portanto, $\mp\lambda$ são autovalores de A.

4.3.1 Resoluções de alguns exercícios que precedem a subseção 4.1.1

EX. 2

$$\begin{aligned}A=\begin{bmatrix}2&1&0&0&0&0\\0&0&4&-3&0&0\\0&0&0&0&-6&5\\-1&2&3&4&-5&6\\1&0&0&0&0&0\\0&-1&0&0&0&0\\0&0&1&0&0&0\\0&0&0&-1&0&0\end{bmatrix} &\Longrightarrow L=L_A\\ &\Longrightarrow L \text{ é linear pela afirmação 1.}\end{aligned}$$

EX. 5.(a)

$$\begin{aligned}A_x=\begin{bmatrix}1&0\\0&0\end{bmatrix} \text{ e } A_y=\begin{bmatrix}0&0\\0&1\end{bmatrix} &\Longrightarrow P_x=L_{A_x} \text{ e } P_y=L_{A_y}\\ &\Longrightarrow P_x \text{ e } P_y \text{ são lineares pela afirmação 1.}\end{aligned}$$

EX. 6.(a)

$$\begin{aligned}A_x=\begin{bmatrix}1&0\\0&-1\end{bmatrix} \text{ e } A_y=\begin{bmatrix}-1&0\\0&1\end{bmatrix} &\Longrightarrow R_x=L_{A_x} \text{ e } R_y=L_{A_y}\\ &\Longrightarrow R_x \text{ e } R_y \text{ são lineares pela afirmação 1.}\end{aligned}$$

EX. 5.(b), partes i. e ii., e EX. 6.(b)

Se

$$\mathbf{x}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix},\ \mathbf{a}=\begin{bmatrix}1/\sqrt{3}\\1/\sqrt{3}\\1/\sqrt{3}\end{bmatrix},\ \mathbf{a}_1=\begin{bmatrix}1/\sqrt{2}\\0\\-1/\sqrt{2}\end{bmatrix} \text{ e } \mathbf{a}_2=\begin{bmatrix}-1/\sqrt{6}\\2/\sqrt{6}\\-1/\sqrt{6}\end{bmatrix},$$

então

$$\begin{aligned}
P_{\mathcal{S}}(\mathbf{x}) &= (\mathbf{x}\cdot\mathbf{a})\mathbf{a} \\
&= \frac{x_1+x_2+x_3}{\sqrt{3}}\begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \\
&= \begin{bmatrix} \frac{x_1+x_2+x_3}{3} \\ \frac{x_1+x_2+x_3}{3} \\ \frac{x_1+x_2+x_3}{3} \end{bmatrix} \\
&= \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\
&= A_{\mathcal{S}}\mathbf{x} \\
&\text{e} \\
P_{\mathcal{S}^\perp}(\mathbf{x}) &= (\mathbf{x}\cdot\mathbf{a}_1)\mathbf{a}_1 + (\mathbf{x}\cdot\mathbf{a}_2)\mathbf{a}_2 \\
&= \frac{x_1-x_3}{\sqrt{2}}\begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix} + \left(\frac{-x_1+2x_2-x_3}{\sqrt{6}}\right)\begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix} \\
&= \begin{bmatrix} \frac{x_1-x_3}{2} \\ 0 \\ \frac{x_3-x_1}{2} \end{bmatrix} + \begin{bmatrix} \frac{x_1-2x_2+x_3}{6} \\ \frac{-x_1+2x_2-x_3}{3} \\ \frac{x_1-2x_2+x_3}{6} \end{bmatrix} \\
&= \begin{bmatrix} \frac{4x_1-2x_2-2x_3}{6} \\ \frac{-x_1+2x_2-x_3}{3} \\ \frac{-2x_1-2x_2+4x_3}{6} \end{bmatrix} \\
&= \begin{bmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\
&= A_{\mathcal{S}^\perp}\mathbf{x}.
\end{aligned}$$

Agora, basta observarmos que

$$\begin{aligned}
R_{\mathcal{S}}(\mathbf{x}) &= 2P_{\mathcal{S}}(\mathbf{x}) - \mathbf{x} \\
&= 2A_{\mathcal{S}}\mathbf{x} - \mathsf{I}\mathbf{x} \\
&= (2A_{\mathcal{S}} - \mathsf{I})\,\mathbf{x} \\
&\text{e} \\
R_{\mathcal{S}^\perp}(\mathbf{x}) &= 2P_{\mathcal{S}^\perp}(\mathbf{x}) - \mathbf{x} \\
&= 2A_{\mathcal{S}^\perp}\mathbf{x} - \mathsf{I}\mathbf{x} \\
&= (2A_{\mathcal{S}^\perp} - \mathsf{I})\,\mathbf{x},
\end{aligned}$$

onde I representa 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,

$$
\begin{aligned}
\mathbb{R}^2 \ni \mathbf{x} \mapsto L_2\left(L_1(\mathbf{x})\right) &= L_2\left(\begin{bmatrix} x_2 \\ x_1 \\ x_1 + x_2 \end{bmatrix}\right) \\
&= \begin{bmatrix} x_1 + x_2 \\ 2x_1 + x_2 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\
&= C\mathbf{x}.
\end{aligned}
$$

Agora, 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 antecipamos no capítulo 4, existem matrizes sem autovalores (reais). De fato, veremos que a matriz de (5.2) não tem autovalores reais, [1] pois não é possível resolvermos (em $\mathbb{R}$) a equação característica

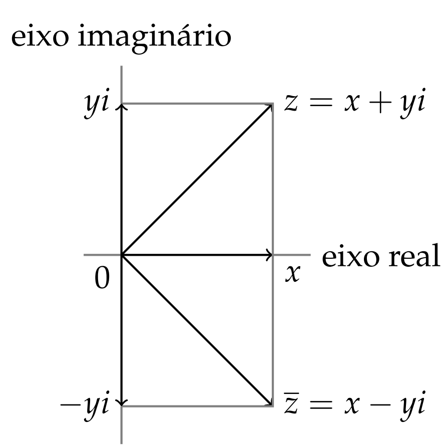
$$\lambda^2 + 1 = 0.$$

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. 134.

[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 $w = u + vi$, definimos 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/radiciação em $\mathbb{C}$, confira qualquer bom livro sobre o tema.[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 $k^{-1} \in \mathbb{K}$.

EXEMPLOS DE CORPOS

$\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$. (Embora existam outros exemplos,[10] consideraremos (sem perda de generalidade) $\mathbb{K} = \mathbb{R}, \mathbb{C}$ nos capítulos 6 e 7.)

---

[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!

[10]Verifique que

$$\mathbb{K} = \left\{ x + y\sqrt{2} \mid x, y \in \mathbb{Q} \right\}$$

é um corpo, denotado por $\mathbb{K} = \mathbb{Q}\left[\sqrt{2}\right]$.

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. $(\lambda + \beta)\mathbf{x} = \lambda\mathbf{x} + \beta\mathbf{x}$;
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 $\mathbb{K}^{m\times n}$.

[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.

5.2 Exercícios

1. Sejam:

   · $\mathbb{K}$ um corpo;

   · $\mathbf{x} = (x_1, \ldots, x_n), \mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{K}^n$;

   · $\mathbf{x} \cdot \mathbf{y} := x_1y_1 + \cdots + x_ny_n$.

   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 exemplos onde a propriedade (d) supracitada não seja satisfeita.[16]

   1.2. Para que $\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$ esteja *bem definido* em $\mathbb{K}^n$, considere

   $$\mathbf{x} \cdot \mathbf{y} := x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n} \tag{5.1}$$

   e verifique que (d) é satisfeita para qualquer corpo $\mathbb{K}$.[17]

   1.3. Ao considerarmos (5.1), como ficam as propriedades (a) e (c) supracitadas?[18]

   OBSERVAÇÃO

   Como a conjugação complexa não altera números reais, se $\mathbb{K} = \mathbb{R}$ e $\mathbf{x}, \mathbf{y} \in \mathbb{K}^n$, então

   $$x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n} = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

   Portanto, o produto interno estudado até o capítulo 4 é, exatamente, aquele definido em (5.1), isto é, independentemente do espaço $\mathbb{K}^n$ considerado, $\mathbf{x} \cdot \mathbf{y}$ conjuga as coordenadas de $\mathbf{y}$.

---

[16] RESOLUÇÃO

· $\mathbf{x} = (2i, 1) \Longrightarrow \mathbf{x} \cdot \mathbf{x} = -3 < 0$;

· $\mathbf{x} = (i, 1) \Longrightarrow \mathbf{x} \cdot \mathbf{x} = 0$.

[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 $\mathbf{a}_2 = (1,2,1-i)$. Como nenhum desses vetores é um múltiplo escalar do outro, $\mathcal{B} = \{\mathbf{a}_1, \mathbf{a}_2\}$ 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 que uma base ortogonal $\mathcal{B}' = \{\mathbf{a}_1', \mathbf{a}_2'\}$ de $\mathcal{S}$ seja obtida, 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 que uma base ortonormal $\mathcal{B}'' = \{\mathbf{a}_1'', \mathbf{a}_2''\}$ de $\mathcal{S}$ seja obtida, 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

$$\mathbf{x} = \left((C+1)+(D+1)i, E+1, F+1\right).$$[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 nota de rodapé do enunciado do exercício 2 da subseção 7.2.7.

[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 (3.14),[23] então

$$[A|\mathrm{I}] = \left[\begin{array}{ccc|ccc} 1 & 1+i & 0 & 1 & 0 & 0 \\ 0 & 1 & i & 0 & 1 & 0 \\ -i & 1-2i & 2 & 0 & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|ccc} 1 & 1+i & 0 & 1 & 0 & 0 \\ 0 & 1 & i & 0 & 1 & 0 \\ 0 & -i & 2 & i & 0 & 1 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 1-i & 1 & -(1+i) & 0 \\ 0 & 1 & i & 0 & 1 & 0 \\ 0 & 0 & 1 & i & i & 1 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -i & -2(1+i) & -(1-i) \\ 0 & 1 & 0 & 1 & 2 & -i \\ 0 & 0 & 1 & i & i & 1 \end{array}\right] = [\mathrm{I}|A'].$$

[23] Cf. p. 76.

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, $T = R_{\pi/2}$, a rotação de $\pi/2$ radianos 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 nota de rodapé do enunciado do exercício 2 da subseção 7.2.7.

[26] Cf. (4.4), p. 88.

- 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:

$$b_{ij} = \overline{a_{ij}},\, i,j = 1,\dots,n.$$

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] Cf. a figura 4.1, p. 88 ... 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 dita *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 à existência de uma matriz unitária* P *com*

$$\mathsf{P}^*A\mathsf{P} = \mathsf{D}$$

diagonal.[32]

Demonstraremos (na seção 5.3) uma forma mais fraca do resultado (5.6) (para matrizes hermiteanas) e (na seção 7.4) a forma geral desse resultado (para matrizes/operadores normais).

12.3. Para 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 com diagonal principal formada pelos 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

$$\begin{aligned} \det(A - \lambda \mathsf{I}) &= \begin{vmatrix} \frac{1}{\sqrt{2}} - \lambda & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} - \lambda \end{vmatrix} \\ &= \lambda^2 - \sqrt{2}\lambda + 1, \end{aligned}$$

$\lambda_{1,2} = (1 \pm i)\sqrt{2}$ 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

$$\begin{aligned} \mathsf{P}_1^{-1} A \mathsf{P}_1 &= \mathsf{P}_1^* A \mathsf{P}_1 \\ &= \mathsf{P}_1^* \left[A\mathbf{x}\; A\mathbf{u}_2 \cdots A\mathbf{u}_r\right] \\ &= \begin{bmatrix} \overline{\mathbf{x}}^T \\ \overline{\mathbf{u}}_2^T \\ \vdots \\ \overline{\mathbf{u}}_r^T \end{bmatrix} \left[A\mathbf{x}\; A\mathbf{u}_2 \cdots A\mathbf{u}_r\right] \\ &= \begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & B \end{bmatrix}, \end{aligned}$$

[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 $A \in \mathbb{C}^{n\times n}$

$$\boxed{\textit{A é hermiteana} \Longrightarrow \textit{A é unitariamente diagonalizável.}}$$

DEMONSTRAÇÃO

Note que, pelo lema de Schur, existe uma matriz P unitária tal que $\mathsf{P}^*A\mathsf{P}$ é triangular superior. Além disso, como A é hermiteana, $\mathsf{P}^*A\mathsf{P}$ também o é, pois

$$\begin{aligned} (\mathsf{P}^*A\mathsf{P})^* &= \mathsf{P}^*A^*\mathsf{P} \\ &= \mathsf{P}^*A\mathsf{P}. \end{aligned}$$

Portanto, 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 especificarmos o espaço vetorial $\mathcal{V}$, denotaremos 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, os vetores de $\mathcal{V} = \mathbb{K}^{\mathbb{N}}$ são as sequências infinitas em $\mathbb{K}$, isto é, $x \in \mathcal{V}$ se, e somente se, x é uma função definida por
$$\begin{array}{rccc} x: & \mathbb{N} & \to & \mathbb{K} \\ & n & \mapsto & x_n \end{array}$$
e denotada por
$$x = (x_1, x_2, \ldots)\,.$$
Em $\mathcal{V}$, a adição de vetores e a multiplicação de escalares por vetores são definidas como em $\mathbb{K}^n$, ou seja, componente a componente. Portanto,
$$x + y : n \mapsto x_n + y_n \text{ e } \lambda x : n \mapsto \lambda x_n,$$
para quaisquer $x, y \in \mathcal{V}$ e qualquer escalar $\lambda \in \mathbb{K}$. Além disso, $0 = (0, 0, \ldots)$ é o vetor nulo de $\mathcal{V}$ e $-x = (-x_1, -x_2, \ldots)$.

- Os vetores de $\mathcal{V} = \mathcal{P}_n(\mathbb{K})$ são os polinômios com coeficientes em $\mathbb{K}$ e de graus menores do que n. Caso p e q sejam polinômios pertencentes à $\mathcal{V}$, isto é, para cada $t \in \mathbb{K}$,
$$p(t) = c_0 + c_1 t + c_2 t^2 + \cdots + c_{n-1} t^{n-1} \;\text{ e }\; q(t) = d_0 + d_1 t + d_2 t^2 + \cdots + d_{n-1} t^{n-1},$$
com coeficientes c_i e d_i em $\mathbb{K}$, $i = 0, 1, 2, \ldots, n-1$,[4] e α seja algum escalar pertencente à $\mathbb{K}$,
$$(c_i + d_i)\, t^i \;\text{ e }\; \alpha c_i t^i \tag{6.1}$$
são as i-ésimas parcelas da soma $(p+q)(t)$ e do produto $(\alpha p)(t)$, respectivamente, para cada $t \in \mathbb{K}$. Em particular, 0 é o polinômio nulo, ou seja, de grau $n-1$ e com n coeficientes nulos, e a i-ésima parcela de $(-p)(t)$ (simétrico de $p(t)$), com $t \in \mathbb{K}$ arbitrário, é dada por $-c_i t^i$, $i = 0, 1, 2, \ldots, n-1$.

EXERCÍCIO

Para cada exemplo dessa subseção, inclua os argumentos faltantes para que o $\mathcal{V}$ considerado seja 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.

EXEMPLOS

[3] Verifique!

[4] Alguns desses coeficientes podem ser nulos. (Por exemplo, $-2 + t^2 + \frac{1}{2}t^3 + t^5 \in \mathcal{P}_6(\mathbb{K})$.) Além disso, observamos que, em outros livros, $\mathcal{P}_n(\mathbb{K})$ representa o espaço dos polinômios com coeficientes em $\mathbb{K}$ e de graus menores do que *ou iguais a n*. Nesse caso, esse espaço será $(n+1)$-dimensional. Como veremos, na nossa representação, esse espaço tem dimensão n.

- Em relação ao primeiro exemplo da subseção 6.1.1, caso $\mathbb{C}^n$ seja um espaço vetorial sobre $\mathbb{C}$, $\dim(\mathbb{C}^n) = n$ 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 os vetores da lista $\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}, \mathbf{e}_n$ são LI. Contudo, caso $\mathbb{C}^n$ seja um 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 os vetores da lista $\mathbf{e}_1, i\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}, i\mathbf{e}_{n-1}, \mathbf{e}_n, i\mathbf{e}_n$ são LI.

- Uma base para o espaço do segundo exemplo da subseção 6.1.1, dita *base canônica*, é obtida ao considerarmos os vetores da lista

$$e_1 = (1,0,0,0,\ldots), e_2 = (0,1,0,0,\ldots), e_3 = (0,0,1,0,\ldots), \ldots, \tag{6.2}$$

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] Logo, para $x \in \mathbb{K}^{\mathbb{N}}$, $x = x_1e_1 + x_2e_2 + x_3e_3 + \cdots$. Diremos que esse espaço tem dimensão *infinita*, mas não nos aprofundaremos nesse tipo de infinitude.

- O espaço do terceiro exemplo da subseção 6.1.1 tem dimensão n e uma de suas bases é constituída pelos polinômios da lista $1, t, t^2, \ldots, t^{n-1}$. De fato, esses n polinômios geram $\mathcal{P}_n(\mathbb{K})$, pois o polinômio $p(t)$ daquele exemplo é uma CL deles.[6] Além disso, os polinômios da lista supracitada são LI, pois, igualando-se o $p(t)$ supracitado ao polinômio nulo, $c_i = 0$, $i = 0, 1, 2, \ldots, n-1$.[7]

- Os vetores de $\mathcal{V} = \mathcal{P}(\mathbb{K})$ são os polinômios com coeficientes em $\mathbb{K}$.[8] Em $\mathcal{V}$, a adição de dois polinômios e a multiplicação de escalares por polinômios são definidas como em (6.1), acrescentando-se parcelas com coeficientes nulos ao polinômio de menor grau, no caso da adição.[9] Podemos visualizar essas operações pelos gráficos das combinações lineares de polinômios reais. Por exemplo, na figura 6.1, ilustramos os (ou partes dos) gráficos do polinômio cúbico $p(t) = \frac{t^3}{3} - 2$, com apenas uma raiz real, da parábola $q(t) = 3t^2 - \frac{11t}{2}$, com duas raízes reais, e da CL $3p(t) - 2q(t) = t^3 - 6t^2 + 11t - 6$, com três raízes reais.[10]

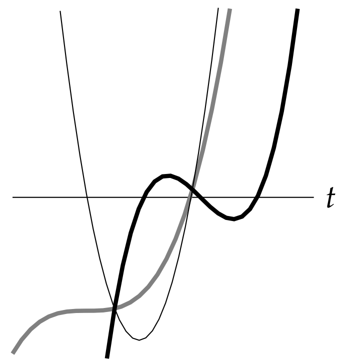
  Note que $\mathcal{S} = \mathcal{P}_n(\mathbb{K})$ é um subespaço de dimensão n de $\mathcal{V}$ e, analogamente à lista (6.2) de vetores de $\mathbb{K}^{\mathbb{N}}$, os vetores da lista

$$1, t, t^2, \ldots, t^{n-1}, t^n, t^{n+1}, \ldots$$

formam a base *canônica* de $\mathcal{V}$ e, por isso, diremos que $\mathcal{V}$ tem dimensão *infinita*.

[5] Note a semelhança com a base canônica de $\mathbb{K}^m$!
[6] O mesmo vale para o polinômio $q(t)$!
[7] Iguale os coeficientes de $p(t)$ aos do polinômio $0 + (0)(t) + (0)(t^2) + \cdots + (0)(t^{n-1})$.
[8] Agora, os graus dos polinômios pertencentes à $\mathcal{V}$ são dados por inteiros não negativos quaisquer.
[9] Verifique que $\mathcal{V}$ é um espaço vetorial sobre $\mathbb{K}$.
[10] $3p(t) - 2q(t) = (t-1)(t-2)(t-3)$.

Figura 6.1: Gráficos de $q(t)$, $p(t)$ (em cinza) e da CL supracitada (em negrito)



- 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 $\mathcal{S}$ e, caso λ seja um escalar pertencente à $\mathbb{K}$ e x e y sejam vetores pertencentes à $\mathcal{S}$,[11] 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.

OBSERVAÇÃO

Se $\mathcal{S}$ é um subespaço de $\mathcal{V}$, então $\mathcal{S}$ também é um espaço vetorial (sobre $\mathbb{K}$).[12]

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

[11] x tem m componentes não nulas e y tem n componentes não nulas.

[12] De fato, $0 \in \mathcal{S}$ e, como

$$\lambda \in \mathbb{K} \text{ e } u, v \in \mathcal{S} \Longrightarrow \lambda u, u + v \in \mathcal{S},$$

podemos considerar, em $\mathcal{S}$, operações de adição de vetores e 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, pois, 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 $\mathcal{S}$.

- $a + bt + ct^2 + dt^3$.

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 $\mathcal{V} = \mathbb{R}^4, \mathbb{R}^{2\times 2}, \mathcal{P}_4(\mathbb{R})$.

EXEMPLO

Nos três casos supracitados, considere a mesma CL, representada nas três formas seguintes:

- $2(1,1,0,0) - (0,1,1,0) = (2,1,-1,0)$;
- $2\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}$;
- $2(1+t) - (t+t^2) = 2 + t - t^2$.

A indistinguibilidade supracitada é conhecida como *isomorfismo*. Em geral, dizemos que dois espaços vetoriais sobre $\mathbb{K}$, digamos $\mathcal{V}_1$ e $\mathcal{V}_2$, são *isomorfos* quando existe uma função linear $L : \mathcal{V}_1 \to \mathcal{V}_2$ invertível, chamada *isomorfismo (entre $\mathcal{V}_1$ e $\mathcal{V}_2$)*. A linearidade de L, como no capítulo 4, é equivalente às seguintes propriedades:

- A imagem da soma de vetores (pela L) é igual a soma das imagens desses vetores;
- A imagem do produto de um escalar qualquer por um vetor arbitrário (pela L) é igual ao produto desse escalar pela imagem desse vetor.

A invertibilidade de L é equivalente à existência de sua inversa $L^{-1} : \mathcal{V}_2 \to \mathcal{V}_1$, cuja composição com L resulta numa função identidade.[13]

EXEMPLOS

- É fácil ver que,
$$\begin{array}{cccc} L & : & \mathbb{R}^{2\times 2} & \to & \mathbb{R}^4 \\ & & \begin{bmatrix} a & b \\ c & d \end{bmatrix} & \mapsto & (a,b,c,d) \end{array}$$
é um isomorfismo entre $\mathbb{R}^{2\times 2}$ e $\mathbb{R}^4$.[14]
- É fácil ver que,
$$\begin{array}{ccccc} L & : & \mathcal{P}_4(\mathbb{R}) & \to & \mathbb{R}^4 \\ & & a + bt + ct^2 + dt^3 & \mapsto & (a,b,c,d) \end{array} \tag{6.3}$$
é um isomorfismo entre $\mathcal{P}_4(\mathbb{R})$ e $\mathbb{R}^4$.[15]

[13] Por exemplo, se $\mathcal{V}_1 = \mathbb{R}^n$, $L^{-1} \circ L = S_1$ é a semelhança de razão $\lambda = 1$, vista no início do capítulo 4.
[14] Verifique!
[15] Idem!

· A correspondência biunívoca entre $\mathbb{K}^{m\times n}$ e $\mathbb{K}^{mn}$ da subseção 3.1.2 (para $\mathbb{K} = \mathbb{R}$) é, claramente, um isomorfismo entre esses espaços.[16]

Esses três isomorfismos transformam bases canônicas em bases canônicas e preservam dimensões. 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,[17] 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 $\mathrm{Im}(L) = \mathcal{V}_2$.

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

$$\begin{array}{cccc} L : & \mathbb{R}^{2\times 3} & \to & \mathbb{R}^6 \\ & \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} & \mapsto & (a, b, c, d, e, f) \end{array}$$

reduz o problema a determinação de duas bases do subespaço de $\mathbb{R}^6$, gerado por

$$(1, -1, 1, -1, 0, 1),\ (-1, 1, 0, -1, 1, -1),\ (1, -1, 2, -3, 1, 1) \text{ e } (3, -3, 1, 1, -2, 3),$$

[16] Idem!

[17] Cf. p. 156.

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),[18] defina $\|A\| := \|L(A)\|$ (onde a segunda norma é a norma em $\mathbb{K}^{mn}$ definida no exercício 1 da seção 5.2).[19] Portanto,

$$\begin{aligned} \|A\| &= \sqrt{|a_{11}|^2 + \cdots + |a_{1n}|^2 + |a_{21}|^2 + \cdots + |a_{2n}|^2 + \cdots + |a_{m1}|^2 + \cdots + |a_{mn}|^2} \\ &= \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2} \end{aligned}$$

calcula o "comprimento" de A.

Final da parte 1 (P_1) do livro

- A parte 2 (P_2) desse livro terá início na próxima seção. Essencialmente, P_2 é um curso mais avançado de AL, sobre o que estudamos desde o capítulo 2, acrescido da *forma de Jordan*. Contudo, no lugar da abordagem matricial empregada até agora, adotaremos o enfoque de operadores.

- Para o material já estudado nesse capítulo, recomendo que sejam trabalhados apenas os exercícios 1–8 e 13–14 da seção 6.4. Os seus outros exercícios, que são mais conceituais, são recomendados apenas para aqueles que estudarem o restante desse capítulo.

- Alguns dos exercícios que podem ser resolvidos com o que estudamos até essa seção, utilizam alguns conceitos e resultados que serão aprofundados a partir da seção 6.3. Especificamente, analogamente 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 $L : \mathcal{V} \to \mathcal{W}$ pelos 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. Além disso, L é dita:

  · *sobrejetiva* quando $\mathrm{Im}(L) = \mathcal{W}$;

  · *injetiva* quando
$$u, v \in \mathcal{V} \text{ com } L(u) = L(v) \Longrightarrow u = v.$$

[18]Cf. p. 51.

[19]Confira o exercício 4 da seção 6.4, para um melhor entendimento de *norma*.

Sobrejetividade e injetividade foram estudadas para funções reais do ensino médio e (na subseção 6.3.2) demonstraremos que:

- $\text{Nu}(L)$ é um subespaço de $\mathcal{V}$;
- $\text{Im}(L)$ é um subespaço de $\mathcal{W}$;
- L é injetiva $\Longleftrightarrow \text{Nu}(L) = \{0_{\mathcal{V}}\}$.

Agora, "mãos à obra", ou seja, tentem resolver os exercícios que acabamos de recomendar.

6.3 Alguns resultados e algumas demonstrações

P_2 começa aqui.[20] P_1 será utilizada para exemplificarmos alguns conceitos e resultados que estudaremos em P_2.

6.3.1 Espaços finitamente gerados, vetores LI e LD, bases

O subespaço de $\mathcal{V}$ gerado pelos m vetores

$$v_j \in \mathcal{V}, j = 1, \dots, m,$$

é denotado por

$$\begin{aligned}[v_1, \dots, v_m] &= \{v \in \mathcal{V} \,:\, v \text{ é uma CL dos vetores da lista } v_1, \dots, v_m\} \\ &= \left\{ \sum_{j=1}^{m} a_j v_j \,:\, a_1, \dots, a_m \in \mathbb{K} \right\}.\end{aligned}$$

EXERCÍCIO

Verifique que, de fato,

$$0_{\mathcal{V}}, \lambda v, u + v \in [v_1, \dots, v_m],$$

para quaisquer $\lambda \in \mathbb{K}$ e $u, v \in [v_1, \dots, v_m]$.

EXEMPLOS

Confira as seções 2.5.1 e 2.5.2 do capítulo 2.

DEFINIÇÃO

Caso $\mathcal{V} = \{0\}$ ou

$$[v_1, \dots, v_m] = \mathcal{V}, \tag{6.4}$$

dizemos que $\mathcal{V}$ é *finitamente gerado*.

EXEMPLOS

$\mathbb{K}^m = [\mathbf{e}_1, \dots, \mathbf{e}_m]$ e $\mathcal{P}_m(\mathbb{K}) = \left[1, t, t^2, \dots, t^{m-1}\right]$ são finitamente gerados.

OBSERVAÇÃO SOBRE (IN)DEPENDÊNCIA LINEAR

É possível que os m geradores de $\mathcal{V}$ em (6.4) sejam LI, ou seja, a única CL nula desses geradores seja a trivial, isto é, a condição

$$\sum_{j=1}^{m} a_j v_j = 0_{\mathcal{V}}, \text{ com } a_j \in \mathbb{K}, \; j = 1, \dots, m,$$

seja válida apenas para

$$a_1 = \cdots = a_m = 0.$$

Nesse caso, $\{v_1, \dots, v_m\}$ é uma *base* de $\mathcal{V}$. Portanto, caso sejam LD, os geradores supracitados não formam uma base de $\mathcal{V}$.

[20] Confira a página anterior.

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 os vetores da lista $v_1, \ldots, v_m$ são LD *e $v_1 \neq 0$, então, para algum índice $j \in \{2, \ldots, m\}$, as condições seguintes são válidas:*

1. $v_j \in [v_1, \ldots, v_{j-1}]$;
2. $[v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m] = [v_1, \ldots, v_m]$.

EXERCÍCIO

Analise o exemplo (com $\mathbb{K}^3$) supracitado, utilizando (R1).

DEMONSTRAÇÃO DO ITEM 1 DE (R1)

Por hipótese, existe uma lista de escalares $a_1, \ldots, a_m$, não todos nulos, tal que

$$\sum_{i=1}^{m} a_i v_i = 0.$$ [21]

Na verdade, como $v_1 \neq 0$, deve existir algum escalar não nulo na lista $a_2, \ldots, a_m$. Assim, caso j seja o maior índice em $\{2, \ldots, m\}$ tal que $a_j \neq 0$, como

$$v_j = -\frac{a_1}{a_j} v_1 - \cdots - \frac{a_{j-1}}{a_j} v_{j-1}, \tag{6.5}$$

a condição 1 é válida.

DEMONSTRAÇÃO DO ITEM 2 DE (R1)

Para $v \in [v_1, \ldots, v_m]$, existe uma lista de escalares $b_1, \ldots, b_m$ tal que

$$v = b_1 v_1 + \cdots + b_j v_j + \cdots + b_m v_m. \tag{6.6}$$

Portanto, para provarmos a validade da condição 2, basta substituirmos o lado direito de (6.5) no lugar do v_j da equação (6.6).

(R2) Em (6.4),[22] os m geradores majoram n vetores LI

Se $\mathcal{V}$ é finitamente gerado por m vetores e tem n vetores LI, *então $n \leq m$.*

[21] Está implícito, no enunciado e na demonstração de (R1), que os vetores pertencem ao espaço $\mathcal{V}$ e os escalares pertencem ao corpo $\mathbb{K}$.

[22] Cf. p. 149.

EXERCÍCIO

Via (R2), analise o exemplo (com $\mathbb{K}^3$) supracitado.

DEMONSTRAÇÃO DE (R2)

Considere os m geradores listados em (6.4) e que os vetores da lista

$$u_1, \ldots, u_n \tag{6.7}$$

são LI em $\mathcal{V}$.

Passo 1. Os $m+1$ vetores da lista ordenada $u_1, v_1, \ldots, v_m$ são LD.[23] Assim, podemos excluir algum vetor, diferente de u_1, da lista supracitada.[24] 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,[25] começando por u_1, u_2. Assim, podemos excluir algum vetor, diferente de u_1 e u_2, da lista supracitada.[26] 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,[27] 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.[28] Portanto, podemos obter uma lista ordenada ℓ_j, começando por $u_1, \ldots, u_j$, formada por m geradores de $\mathcal{V}$.

Na lista ℓ_j, à esquerda dos $m-j$ geradores da lista (6.4) que não foram excluídos, foi incluída a lista $u_1, \ldots, u_j$. Assim, o número máximo de passos é m,[29] que majora n.[30] De fato, caso o passo $m+1$ pudesse ocorrer, todos os geradores listados em (6.4) já teriam sido excluídos no passo m e os vetores da lista

$$u_1, \ldots, u_m, u_{m+1}$$

seriam LD, contradizendo a hipótese da independência linear dos vetores da lista (6.7).

(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.

[23] De fato, $u_1 \in [v_1, \ldots, v_m]$.

[24] Cf. (R1), p. 150.

[25] De fato, como $u_2 \in \mathcal{V}$ e $\mathcal{V}$ é gerado pelos vetores de ℓ_1, u_2 é uma CL dos vetores de ℓ_1.

[26] Como u_1 e u_2 são LI, nenhum deles pode ser excluído, por (R1).

[27] De fato, como $u_j \in \mathcal{V}$ e $\mathcal{V}$ é gerado pelos vetores de ℓ_{j-1}, u_j é uma CL dos vetores de ℓ_{j-1}.

[28] Como os vetores da lista $u_1, \ldots, u_j$ são LI, nenhum deles pode ser excluído, por (R1).

[29] m = número de geradores listados em (6.4).

[30] n = número de vetores da lista (6.7).

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).

Caso $\mathcal{W}$ não seja finitamente gerado no passo $j-1$, devemos prosseguir para o passo j. Contudo, caso $\mathcal{V}$ seja gerado por m vetores, o maior valor que o índice j pode assumir é $m+1$ e, caso assuma esse valor, $\mathcal{W} = [w_1, \ldots, w_m]$ e, pelo passo m, os vetores da lista $w_1, \ldots, w_m$ são LI. De fato, caso o passo $m+2$ pudesse ocorrer, 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 (finita) é um conjunto finito de geradores LI, o seguinte resultado é válido:

> *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 $v_j \in \left[v_1, \ldots, v_{j-1}\right]$, 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 uma 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 nesse livro tiveram bases. Por (R5), qualquer espaço da forma (6.4) tem alguma base (finita).[31]

(R6) LI estendidos a bases

Caso seja necessário, para obtermos uma base de $\mathcal{V} = [v_1, \dots, v_m]$, podemos estender uma lista, digamos, $w_1, \dots, w_n$, constituída de vetores LI *em $\mathcal{V}$.*

EXEMPLO

Para $w_1 = \mathbf{e}_1$, $w_2 = \mathbf{e}_2 + \mathbf{e}_3$ e $\mathcal{V} = \mathbb{R}^3$, considere $w_3 = \mathbf{e}_1 \times (\mathbf{e}_2 + \mathbf{e}_3)$.

DEMONSTRAÇÃO DE (R6)

Passo 1. Caso v_1 seja uma CL dos vetores da lista supracitada, considere

$$\mathcal{B}_1 = \{w_1, \dots, w_n\}\,.^{32}$$

Caso contrário, considere

$$\mathcal{B}_1 = \{w_1, \dots, w_n, v_1\}\,.^{33}$$

Em qualquer caso, se $\mathcal{B}_1$ não for uma base de $\mathcal{V}$, siga para o passo 2.

Passo 2. Caso v_2 seja uma CL dos vetores de $\mathcal{B}_1$, considere $\mathcal{B}_2 = \mathcal{B}_1$. Caso contrário, considere

$$\mathcal{B}_2 = \mathcal{B}_1 \cup \{v_2\}\,.^{34}$$

Em qualquer caso, se $\mathcal{B}_2$ não for uma base de $\mathcal{V}$, siga para o passo 3.

Para $j \geq 2$, considere o seguinte:

Passo *j*. Caso v_j seja uma CL dos vetores de $\mathcal{B}_{j-1}$, considere $\mathcal{B}_j = \mathcal{B}_{j-1}$. Caso contrário, considere

$$\mathcal{B}_j = \mathcal{B}_{j-1} \cup \{v_j\}\,.^{35}$$

Em qualquer caso, se $\mathcal{B}_j$ não for uma base de $\mathcal{V}$, siga para o próximo passo.

No passo $j \geq 1$, $\mathcal{B}_j$ é uma base de um subespaço $\mathcal{S}_j$ que contem os vetores da lista $v_1, \dots, v_j$. Obviamente, esse algoritmo deve terminar em algum passo $j \in \{1, \dots, m\}$. De fato, caso o passo m ocorra, $\mathcal{S}_m \supset [v_1, \dots, v_m]$.

(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

[31] Cf. p. 149.
[32] $n \leq m$, por (R2).
[33] $n + 1 \leq m$, por (R2).
[34] O número de vetores de $\mathcal{B}_2$ não pode ser maior do que m, por (R2).
[35] O número de vetores de $\mathcal{B}_j$ não pode ser maior do que m, por (R2).

DEMONSTRAÇÃO DE (R7)

Sejam $\mathcal{B}_1$ e $\mathcal{B}_2$ duas bases de $\mathcal{V}$ com m e n vetores, respectivamente. Assim, como m vetores geram $\mathcal{V}$ e existem n vetores LI em $\mathcal{V}$, $n \leq m$.[36] Analogamente, como n vetores geram $\mathcal{V}$ e existem m vetores LI em $\mathcal{V}$, $m \leq n$.[37] 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),[38] 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).[39]

OBSERVAÇÃO

Caso $\mathcal{W}$ e $\mathcal{V}$ sejam os espaços dados em (R8),

$$\dim \mathcal{W} = \dim \mathcal{V} \iff \mathcal{W} = \mathcal{V}.$$[40]

(R9) Verificação para que $\{v_1, \ldots, v_{\dim \mathcal{V}}\}$ seja uma base de $\mathcal{V}$

Para essa verificação, é suficiente que os vetores da lista $v_1, \ldots, v_{\dim \mathcal{V}}$ sejam LI ou gerem $\mathcal{V}$.

EXEMPLO

$\{(1,2),(3,4)\}$ é base de $\mathbb{R}^2$ pois seus dois vetores são LI.

DEMONSTRAÇÃO DE (R9)

Caso os vetores da lista supracitada sejam LI ou $m = \dim \mathcal{V}$ em (6.4),[41] essa lista pode ser reduzida ou estendida a uma base de $\mathcal{V}$, por (R4) ou (R6).[42] 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).[43]

[36] Cf. (R2), p. 150.
[37] Idem!
[38] Cf. pp. 151 e 153.
[39] Cf. p. 150.
[40] Essa equivalência foi antecipada na subseção 2.5.6 (para $\mathcal{V} = \mathbb{R}^n$).
[41] Cf. p. 149.
[42] Cf. pp. 152 e 153.
[43] Cf. p. 153.

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 onde $L \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$, confira os capítulos 4 e 5. Para exemplos onde $L \in \mathcal{L}(\mathcal{P}(\mathbb{K}), \mathcal{P}(\mathbb{K}))$, $L \in \mathcal{L}(\mathcal{P}(\mathbb{K}), \mathbb{K})$ ou $L \in \mathcal{L}(\mathbb{K}^{\mathbb{N}}, \mathbb{K}^{\mathbb{N}})$, confira os exercícios 5–8 da seção 6.4. Além disso, note que, nos exemplos de isomorfismos da seção 6.2, $L \in \mathcal{L}(\mathbb{R}^{2\times 2}, \mathbb{R}^4)$, $L \in \mathcal{L}(\mathcal{P}_4(\mathbb{R}), \mathbb{R}^4)$ e $L \in \mathcal{L}(\mathbb{R}^{m\times n}, \mathbb{R}^{mn})$.

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, λT e $T + L$ são transformações lineares, pois, 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

$$\begin{aligned}(T + L)(u + v) &= T(u + v) + L(u + v)\\ &= T(u) + T(v) + L(u) + L(v)\\ &= T(u) + L(u) + T(v) + L(v)\\ &= (T + L)(u) + (T + L)(v).\end{aligned}$$

Além disso, $\mathcal{L}(\mathcal{V}, \mathcal{W})$ é um espaço vetorial, pois 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 seis propriedades (elencadas no comecinho da seção 6.1) 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 uma 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

Para $\mathcal{V} = \mathbb{R}^2$, $\mathcal{W} = \mathbb{R}^3$, $v_j = \mathbf{e}_j$, $j = 1, 2$, $w_1 = (1, 1, 0)$ e $w_2 = (1, 0, 1)$, considere a função linear "multiplicação por matriz", dada por

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \overset{L}{\mapsto} \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$ [44]

DEMONSTRAÇÃO DE (R10)

L está bem definida, pois, qualquer $v \in \mathcal{V}$ pode ser escrito, de modo único, como uma CL dos vetores da base supracitada, ou seja, existe uma única lista $a_1, \ldots, a_n$ de escalares (pertencentes ao corpo $\mathbb{K}$ e dependentes de v) tal que

$$v = \sum_{j=1}^{n} a_j v_j. \tag{6.8}$$

Por outro lado, caso $a_j = 1$ e os outros escalares da lista supracitada sejam nulos,

$$v_j \overset{L}{\mapsto} w_j.$$

L é linear, pois, caso $\lambda \in \mathbb{K}$, v seja dado pelo somatório (6.8) e

$$u = \sum_{j=1}^{n} c_j v_j$$

também seja uma CL dos vetores da base supracitada,

$$\begin{aligned} L(\lambda v) &= L\left(\sum_{j=1}^{n} (\lambda a_j)\, v_j\right) \\ &= \sum_{j=1}^{n} (\lambda a_j)\, w_j \\ &= \lambda \sum_{j=1}^{n} a_j w_j \\ &= \lambda L(v) \end{aligned}$$

[44]Confira o início do capítulo 4.

e

$$\begin{aligned} L(u+v) &= L\left(\sum_{j=1}^{n}(c_j+a_j)\,v_j\right) \\ &= \sum_{j=1}^{n}(c_j+a_j)\,w_j \\ &= \sum_{j=1}^{n} c_j w_j + \sum_{j=1}^{n} a_j w_j \\ &= L(u)+L(v). \end{aligned}$$

Para provarmos a unicidade de L, seja $S \in \mathcal{L}(\mathcal{V},\mathcal{U})$ tal que $v_j \overset{S}{\mapsto} w_j$, $j=1,\dots,n$. Assim,

$$S(v_j) = L(v_j)\,,$$

$j=1,\dots,n$. Portanto, caso v seja dado pelo somatório (6.8),

$$\begin{aligned} S(v) &= S(a_1v_1+\cdots+a_nv_n) \\ &= a_1S(v_1)+\cdots+a_nS(v_n) \\ &= a_1L(v_1)+\cdots+a_nL(v_n) \\ &= L(a_1v_1+\cdots+a_nv_n) \\ &= L(v), \end{aligned}$$

pela linearidade de S e L.

(R11) $\mathrm{Nu}(L)$ **é um subespaço de** $\mathcal{V}$ **e** $\mathrm{Im}(L)$ **é um subespaço de** $\mathcal{W}$

Analogamente 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 um subespaço de* $\mathcal{V}$*;*
>
> (II) *A imagem de L, isto é, o conjunto*
>
> $$\mathrm{Im}(L) = \{L(v) \,:\, v \in \mathcal{V}\}\,,$$
>
> *é um subespaço de W.*

OBSERVAÇÃO

Para exemplos, confira o capítulo supracitado e a seção 6.4.

DEMONSTRAÇÃO DO ITEM (I) DE (R11)[45]

Note que $0_{\mathcal{V}} \in \mathrm{Nu}(L)$, pois, como

$$\begin{aligned} L(0_{\mathcal{V}}) &= L(0_{\mathcal{V}}+0_{\mathcal{V}}) \\ &= L(0_{\mathcal{V}})+L(0_{\mathcal{V}})\,, \end{aligned}$$

[45] A demonstração do item (II) fica como exercício.

$L(0_{\mathcal{V}}) = 0_{\mathcal{W}}$. Além disso, se $\lambda \in \mathbb{K}$ e $u, v \in \mathrm{Nu}(L)$, então $\lambda u, u+v \in \mathrm{Nu}(L)$, pois

$$\begin{aligned} L(\lambda u) &= \lambda L(u) \\ &= \lambda \cdot 0_{\mathcal{W}} \\ &= 0_{\mathcal{W}} \end{aligned}$$

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

$$\boxed{L \in \mathcal{L}(\mathcal{V}, \mathcal{W}) \textit{ é injetiva} \iff \mathrm{Nu}(L) = \{0_{\mathcal{V}}\}.}$$

OBSERVAÇÃO

Como antecipado na subeção 6.2.2, A "injetividade" de L significa o seguinte:

$$u, v \in \mathcal{V} \textit{ com } L(u) = L(v) \implies u = v.$$

EXEMPLOS

Confira os exercícios 5–8 da seção 6.4.

DEMONSTRAÇÃO DE (R12)

Para a implicação "$\Longrightarrow$", considere L 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 injetividade de L,

$$v = 0_{\mathcal{V}},$$

ou seja, $v \in \{0_{\mathcal{V}}\}$.
Para a implicação "$\Longleftarrow$", considere

$$\mathrm{Nu}(L) = \{0_{\mathcal{V}}\}\,.$$

Portanto, como, para $u, v \in \mathcal{V}$,

$$\begin{aligned} L(u) = L(v) &\implies L(u) - L(v) = 0_{\mathcal{W}} \\ &\implies L(u-v) = 0_{\mathcal{W}} \\ &\implies u - v \in \mathrm{Nu}(L) \\ &\implies u - v = 0_{\mathcal{V}} \\ &\implies u = v, \end{aligned}$$

L é injetiva.[46]

[46]Note que, na segunda implicação, de cima para baixo, usamos a linearidade de L. Na penúltima, de cima para baixo, usamos a hipótese da nulidade do núcleo.

Além de injetividade, precisaremos do conceito de “sobrejetividade”, ou seja:

DEFINIÇÃO

Como antecipamos na subseção 6.2.2,

$$L \in \mathcal{L}(\mathcal{V}, \mathcal{W}) \text{ é } \textit{sobrejetiva} \iff \mathrm{Im}(L) = \mathcal{W}.$$

EXEMPLOS

Confira os exercícios 5–8 da seção 6.4.

(R13) Isomorfismo e subespaços triviais

$$\boxed{L \in \mathcal{L}(\mathcal{V}, \mathcal{W}) \textit{ é isomorfismo} \iff \mathrm{Nu}(L) = \{0_{\mathcal{V}}\} \textit{ e } \mathrm{Im}(L) = \mathcal{W}.}$$

DEMONSTRAÇÃO

Considere o resultado (R12), a definição supracitada e que, conforme o AXLER,[47] para $L \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, vale o seguinte resultado:

$$\boxed{L \textit{ é um isomorfismo} \iff L \textit{ é injetiva e sobrejetiva.}}$$

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

$$\boxed{\mathcal{V} \textit{ é finitamente gerado e } L \in \mathcal{L}(\mathcal{V}, \mathcal{W}) \Longrightarrow \dim \mathrm{Nu}(L) + \dim \mathrm{Im}(L) = \dim \mathcal{V}.}$$

EXEMPLOS

Confira a subseção 4.1.1.

DEMONSTRAÇÃO DE (R14)

Por (I) de (R11), (R3) e (R5),[48] podemos considerar uma base

$$\{v_1, \ldots, v_m\} \tag{6.9}$$

de $\mathrm{Nu}(L)$. Assim, por (R6),[49] podemos estender a lista $v_1, \ldots, v_m$, caso necessário, para obtermos uma base

$$\{v_1, \ldots, v_m, u_1, \ldots, u_n\} \tag{6.10}$$

de $\mathcal{V}$.[50] Para concluirmos a demonstração, basta provarmos que o conjunto

$$\{L(u_1), \ldots, L(u_n)\}$$

---

[47] Confira o capítulo 1.

[48] Cf. pp. 157, 151 e 153, respectivamente.

[49] Cf. p. 153.

[50] 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.

é uma base de $\mathrm{Im}(L)$. Assim, por um lado, as n imagens por L listadas no conjunto supracitado são LI. De fato,

$$\begin{aligned}\sum_{i=1}^{n} a_i L(u_i) = 0_{\mathcal{W}} &\Longrightarrow L\left(\sum_{i=1}^{n} a_i u_i\right) = 0_{\mathcal{W}}\\ &\Longrightarrow \sum_{i=1}^{n} a_i u_i \in \mathrm{Nu}(L)\\ &\Longrightarrow \sum_{i=1}^{n} a_i u_i = \sum_{j=1}^{m} c_j v_j,\end{aligned}$$

onde:

- os escalares da lista $c_1, \ldots, c_m$ dependem dos escalares da lista $a_1, \ldots, a_n$, que são arbitrários;[51]
- na primeira implicação, de cima para baixo, utilizamos a linearidade de L;
- na última implicação, de cima para baixo, utilizamos a consideração inicial de que (6.9) é uma base de $\mathrm{Nu}(L)$.

Portanto, como

$$\sum_{j=1}^{m} c_j v_j + \sum_{i=1}^{n} (-a_i)\, u_i = 0_V \tag{6.11}$$

e os $m+n$ vetores de (6.10) são LI, todos os escalares de (6.11) são nulos e, em particular,

$$a_i = 0, i = 1, \ldots, n.$$

Por outro lado,

$$\mathrm{Im}(L) = \left[L(u_1), \ldots, L(u_n)\right].$$

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;
- na última igualdade, de cima para baixo, utilizamos a consideração inicial de que (6.9) é uma base de $\mathrm{Nu}(L)$.

[51] Todos esses escalares pertencem ao corpo $\mathbb{K}$.

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

Caso $\mathcal{B} = \{v_1, \ldots, v_n\}$ e $\mathcal{B}' = \{w_1, \ldots, w_m\}$ sejam bases ordenadas de $\mathcal{V}$ e $\mathcal{W}$, respectivamente,[52] podemos considerar a transformação linear

$$\begin{array}{rcccc} \mathcal{M}(-,\mathcal{B},\mathcal{B}') & : & \mathcal{L}(\mathcal{V},\mathcal{W}) & \longrightarrow & \mathbb{K}^{m\times n} \\ & & L & \longmapsto & \mathcal{M}(L,\mathcal{B},\mathcal{B}') \end{array}, \tag{6.12}$$

onde, para cada $j \in \{1, \ldots, n\}$, se

$$L(v_j) = a_{1j}w_1 + \cdots + a_{mj}w_m,$$

com $a_{ij} \in \mathbb{K}$, $i = 1\ldots, m$, 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 não haja dúvidas sobre as bases consideradas, podemos denotar

  $$\mathcal{M}(-) := \mathcal{M}(-,\mathcal{B},\mathcal{B}').$$ [53]

- $\mathcal{M}(-)$ 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 $T, L \in \mathcal{L}(\mathcal{V},\mathcal{W})$.[54]

- Para que $\mathcal{M}(-)$ seja um isomorfismo, basta que seja invertível.[55]

- Como no capítulo 4, podemos demonstrar que

  $$\mathcal{M}(TL,\mathcal{B},\mathcal{B}'') = \mathcal{M}(T,\mathcal{B}',\mathcal{B}'')\,\mathcal{M}(L,\mathcal{B},\mathcal{B}'), \tag{6.13}$$

  para quaisquer transformações $L \in \mathcal{L}(\mathcal{U},\mathcal{V})$ e $T \in \mathcal{L}(\mathcal{V},\mathcal{W})$ e quaisquer bases ordendas $\mathcal{B}$ 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.

[52] 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.

[53] Mas não podemos esquecer que $\mathcal{M}(-)$ depende das bases $\mathcal{B}$ e $\mathcal{B}'$.

[54] Verifique!

[55] A invertibilidade de $\mathcal{M}(-)$ será demonstrada em (R16), p. 162. A linearidade de $\mathcal{M}(-)^{-1}$ é uma consequência da propriedade simétrica do exercício 12 da seção 6.4.

(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 $\alpha_i \in \mathbb{K}$, $i = 1, \ldots, n$, tais que

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n.$$

A *matriz de v na base $\mathcal{B}$* é definida por

$$\mathcal{M}(v, \mathcal{B}) := \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}. \tag{6.14}$$

[56]

Assim,

$$\mathcal{M}\big(L(v), \mathcal{B}'\big) = \mathcal{M}\big(L, \mathcal{B}, \mathcal{B}'\big)\, \mathcal{M}(v, \mathcal{B}). \tag{6.15}$$

[57]

De fato, seja

$$\mathcal{M}(L) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}. \tag{6.16}$$

Logo, para cada índice $j \in \{1, \ldots, n\}$,

$$L\big(v_j\big) = \sum_{i=1}^{m} a_{ij} w_i.$$

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,

$$\mathcal{M}(L(v)) = \begin{bmatrix} a_{11}\alpha_1 + \cdots + a_{1n}\alpha_n \\ \vdots \\ a_{m1}\alpha_1 + \cdots + a_{mn}\alpha_n \end{bmatrix}. \tag{6.17}$$

Assim, para que (6.15) seja válida, basta verificarmos que a matriz coluna (6.17) é igual ao produto da matriz (6.16) pela matriz coluna (6.14).

(R16) $\mathcal{M}(-)$ é invertível

DEMONSTRAÇÃO DE (R16)

Por (R12),[58] 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

$$\mathcal{M}(L) = 0_{\mathbb{K}^{m\times n}}.$$

[56] Caso não haja dúvidas sobre a base considerada, podemos denotar $\mathcal{M}(v) := \mathcal{M}(v, \mathcal{B})$.

[57] Cf. (4.8), p. 99.

[58] Cf. p. 158.

Então, para cada índice $j \in \{1, \ldots, n\}$,

$$L(v_j) = 0_{\mathcal{W}}.$$

Portanto, como $\mathcal{B}$ é uma base de $\mathcal{V}$,

$$L(v) = 0_{\mathcal{W}}, \text{ para cada } v \in \mathcal{V},$$

ou seja,

$$L = 0_{\mathcal{L}(\mathcal{V},\mathcal{W})}.$$

Agora, para demonstrarmos que $\mathcal{M}(-)$ é sobrejetiva, basta verificarmos que

$$\mathrm{Im}\,(\mathcal{M}(-)) = \mathbb{K}^{m\times n}.$$

De fato, se

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{K}^{m\times n}$$

e $L \in \mathcal{L}(\mathcal{V},\mathcal{W})$ é definida por

$$L(v_j) = \sum_{i=1}^{m} a_{ij} w_i,\ j = 1, \ldots, n,$$

então, claramente,

$$\mathcal{M}(L) = A.$$

(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 $\mathcal{L}(\mathcal{V},\mathcal{W}) \xrightarrow{\mathcal{M}(-)} \mathbb{K}^{m\times n}$ é um isomorfismo,

$$\begin{aligned} \dim \mathcal{L}(\mathcal{V},\mathcal{W}) &= \dim \mathbb{K}^{m\times n} \\ &= mn \\ &= (\dim \mathcal{V})(\dim \mathcal{W}), \end{aligned}$$

pelo teorema 6.2.1.[59]

[59] Cf. p. 146.

6.4 Exercícios

1. Caso 0 seja o polinômio nulo e m seja 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 $m = 2$, $p(t) = t^2 + t$ e $q(t) = -t^2 + 1$, $p(t) + q(t) = t + 1 \notin \mathcal{S}$.

2. Seja $\mathbb{R}^{[a,b]}$ o conjunto de todas as funções reais definidas no intervalo real $[a, b]$. Para $f, g \in \mathbb{R}^{[a,b]}$ e $\lambda \in \mathbb{R}$, defina

$$(f + g)(x) := f(x) + g(x) \quad \text{e} \quad (\lambda f)(x) := \lambda f(x), \tag{6.18}$$

para qualquer $x \in [a, b]$. Demonstre que, munido das operações (6.18) 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 $[a,b]$. Para $f,g \in C[a,b]$ e $\lambda \in \mathbb{R}$, defina $f+g$ e λg como em (6.18). Verifique que, munido dessa soma e desse produto, $C[a,b]$ é um espaço vetorial real.

   RESOLUÇÃO

   Basta verificarmos que $C[a,b]$ é um 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 $\mathbb{K}$. 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)[60]
   · $\langle u,u\rangle \geq 0$; $\langle u,u\rangle = 0 \iff u = 0$. (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 $\langle\ ,\ \rangle$ real ou complexo, dizemos que a função

$$\mathcal{V}\ni v \stackrel{\|\ \|}{\longmapsto} \|v\| = \sqrt{\langle v,v\rangle} \in \mathbb{R}$$

   é uma *norma em* $\mathcal{V}$.

   EXEMPLOS

[60] 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,

$$\mathcal{V} \ni \mathbf{x} \longmapsto \|\mathbf{x}\| = \sqrt{|x_1|^2 + \cdots + |x_n|^2} \in \mathbb{R} \tag{6.19}$$

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 (6.19), ou seja,

$$\|A\| = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2}.$$

Portanto, $\|A\| = \sqrt{\langle A, A\rangle}$, onde o produto interno $\langle\ ,\ \rangle$ é definido por

$$\mathcal{V} \times \mathcal{V} \ni (A, B) \longmapsto \langle A, B\rangle := \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\overline{b_{ij}} \in \mathbb{K}.$$

$\boxed{\text{OBSERVAÇÕES}}$

- Pela definição de produto interno, dada no início desse exercício, podemos verificar que uma norma em $\mathcal{V}$ satisfaz as condições seguintes:[61]
  - · (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,[62] 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.[63]

---

[61] Verifique-as!

[62] $\mathbb{K}^n \times \mathbb{K}^n \ni (\mathbf{x}, \mathbf{y}) \longmapsto \mathbf{x} \cdot \mathbf{y} \in \mathbb{K}$.

[63] Confira as seções 2.7 e 5.2. Além disso, para uma demonstração do caso geral, indico o AXLER, referenciado no capítulo 1.

Resolva as seguintes questões:

4.1. Seja $\mathcal{V}$ um espaço vetorial com produto interno $\langle\ ,\ \rangle$ real ou complexo. Demonstre a validade da *desigualdade de Cauchy-Schwarz* para vetores de $\mathcal{V}$, ou seja,

$$u, v \in \mathcal{V} \Longrightarrow |\langle u, v\rangle| \leq \|u\|\|v\|.^{64}$$

4.2. Considere o espaço da questão 3 dessa seção.

(a) Demonstre que $C[a,b]$ é munido do produto interno $\langle\ ,\ \rangle$ real, 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) Caso $\mathcal{S}$ seja um subespaço de $C[a,b]$ e $\{f_1, f_2, \ldots, f_r\}$ seja uma base ortonormal de $\mathcal{S}$, a projeção ortogonal, ou seja, a melhor aproximação,[65] de um vetor $f \in C[a,b]$ 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 $f : [0,1] \to \mathbb{R}$ dada por $f(t) = e^t$, $t \in [0,1]$.

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

$$\begin{aligned} 0 \leq \langle u + \lambda v, u + \lambda v\rangle &= \langle u, u\rangle + \overline{\lambda}\langle u, v\rangle + \lambda\langle v, u\rangle + \lambda\overline{\lambda}\langle v, v\rangle \\ &= \|u\|^2 + \overline{\lambda\langle v, u\rangle} + \lambda\langle v, u\rangle + |\lambda|^2\|v\|^2 \\ &= \|u\|^2 + 2\,\mathrm{Re}(\lambda\langle v, u\rangle) + |\lambda|^2\|v\|^2, \end{aligned}$$

onde $\mathrm{Re}(z)$ representa 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 $z \in \mathbb{K}$,

$$\begin{aligned} 0 \leq \|u\|^2 - 2\,\mathrm{Re}\left(\frac{|\langle v, u\rangle|^2}{\|v\|^2}\right) + \frac{|\langle u, v\rangle|^2}{\|v\|^2} &= \|u\|^2 - \frac{2|\langle u, v\rangle|^2}{\|v\|^2} + \frac{|\langle u, v\rangle|^2}{\|v\|^2} \\ &= \|u\|^2 - \frac{|\langle u, v\rangle|^2}{\|v\|^2}. \end{aligned}$$

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.

[64] 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.

[65] 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

$$\|q_i\| = \sqrt{\langle q_i, q_i\rangle}, i = 1, 2.$$ [66]

[66] 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.[67]

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.[68] 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:

$$\mathrm{Im}(L) = \mathcal{P}(\mathbb{R})?$$ [69]

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.

[67] Deixamos o cálculo dessa projeção para o leitor!

[68] dp/dt é a derivada do polinômio $p(t)$ em relação à variável t.

[69] 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: $\mathrm{Im}(L) = \mathbb{R}$?[70]
Para a injetividade, responda: $\mathrm{Nu}(L) = \{0\}$?[71]

7. Demonstre que o operador

$$\begin{array}{rccc} L: & \mathcal{P}(\mathbb{R}) & \rightarrow & \mathcal{P}(\mathbb{R}) \\ & p(t) & \mapsto & t^2 p(t) \end{array}$$

é linear. L é injetivo? É sobrejetivo?

RESOLUÇÃO PARCIAL

A linearidade de L é trivial.[72] Para a injetividade, devemos responder a pergunta: $\mathrm{Nu}(L) = \{0\}$? Como

$$p(t) \neq 0 \Longrightarrow L(p(t)) = t^2 p(t) \text{ tem grau no mínimo 2,}$$

$L(p(t)) = 0$ apenas se $p(t) = 0$. Então, L é injetiva.
Quanto a sobrejetividade, devemos responder a pergunta: $\mathrm{Im}(L) = \mathcal{P}(\mathbb{R})$?
A resposta é "não!", pois polinômios de grau 1 ou polinômios não nulos de grau 0 não podem ser imagens (por L) de polinômios em $\mathcal{P}(\mathbb{R})$!

8. Demonstre que o operador

$$\begin{array}{rccc} L: & \mathbb{K}^{\mathbb{N}} & \longrightarrow & \mathbb{K}^{\mathbb{N}} \\ & (x_1, x_2, \ldots) & \longmapsto & (x_2, x_3, \ldots) \end{array}$$

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

$$\begin{aligned} L(\lambda x) &= L(\lambda x_1, \lambda x_2, \ldots) \\ &= (\lambda x_2, \lambda x_3, \ldots) \\ &= \lambda (x_2, x_3, \ldots) \\ &= \lambda L(x_1, x_2, \ldots) \\ &= \lambda L(x) \end{aligned}$$

e

$$\begin{aligned} L(x+y) &= L(x_1 + y_1, x_2 + y_2, \ldots) \\ &= (x_2 + y_2, x_3 + y_3, \ldots) \\ &= (x_2, x_3, \ldots) + (y_2, y_3, \ldots) \\ &= L(x_1, x_2, \ldots) + L(y_1, y_2, \ldots) \\ &= L(x) + L(y). \end{aligned}$$

L é sobrejetivo, pois $\mathrm{Im}(L) = \mathbb{K}^{\mathbb{N}}$, ou seja, para cada $y \in \mathbb{K}^{\mathbb{N}}$, existe $x \in \mathbb{K}^{\mathbb{N}}$ tal que $L(x) = y$. De fato, qualquer sequência $x = (x_1, x_2, \ldots)$ onde $x_{n+1} = y_n$ para cada inteiro positivo n, isto é, $x_2 = y_1$, $x_3 = y_2$,

[70] Todo número real r pode ser escrito da forma $\int_0^1 r\,dt$.
[71] $\int_0^1 (2t-1)dt = 0$.
[72] Verifique!

etc., é uma pré-imagem de y (por L), pois

$$\begin{aligned} L(x) &= L(x_1, x_2, \ldots) \\ &= (x_2, x_3, \ldots) \\ &= (y_1, y_2, \ldots) \\ &= y. \end{aligned}$$

Além disso, L não é injetivo, ou seja, $\mathrm{Nu}(L) \neq \{0\}$, pois

$$\mathrm{Nu}(L) = \{a(1, 0, 0, \ldots) \, : \, a \in \mathbb{K}\}.$$

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

$$\begin{array}{rcccccc} L_2 L_1 & := & L_2 \circ L_1 & : & \mathcal{V}_1 & \longrightarrow & \mathcal{V}_3 \\ & & & & v & \longmapsto & L_2(L_1(v)) \end{array}$$

é linear.[73]

RESOLUÇÃO

Sejam $u, v \in \mathcal{V}_1$ e $\lambda \in \mathbb{K}$ arbitrários. Portanto,

$$\begin{aligned} (L_2 L_1)(u + v) &= L_2(L_1(u + v)) \\ &= L_2(L_1(u) + L_1(v)) \qquad (L_1 \text{ É LINEAR}) \\ &= L_2(L_1(u)) + L_2(L_1(v)) \; (L_2 \text{ É LINEAR}) \\ &= (L_2 L_1)(u) + (L_2 L_1)(v) \end{aligned}$$

[73] 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 $\mathcal{V}_1$;[74]

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 $\mathcal{V}_3$.[75]

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 demonstrarmos 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

$$u' = L^{-1}(u) \ \text{ e } \ v' = L^{-1}(v),$$

$$L^{-1}(u+v) = L^{-1}(u) + L^{-1}(v).$$

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,[76] a identidade

$$(L_2L_1)^{-1} = L_1^{-1}L_2^{-1}$$

e que, como acabamos de demonstrar, isomorfismo é uma relação simétrica.

[74] 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.

[75] 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.

[76] 6.4.

13. Podemos considerar $\mathbb{K}^n$ como subespaço de $\mathbb{K}^{\mathbb{N}}$, isto é, existe um 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 (6.3).[77]

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;[78]

   (b) $\dim \mathcal{V} < \dim \mathcal{W} \Longrightarrow \nexists\, L \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ sobrejetiva.[79]

RESOLUÇÃO

Seja $L \in \mathcal{L}(\mathcal{V}, \mathcal{W})$. Por (R14),[80]

$$\dim \mathrm{Nu}(L) + \dim \mathrm{Im}(L) = \dim \mathcal{V}.$$

DEMONSTRAÇÃO DE (a)

Como, por (II) de (R11),[81] $\mathrm{Im}(L)$ é um subespaço de $\mathcal{W}$, temos

$$\dim \mathrm{Im}(L) \leq \dim \mathcal{W},$$

---

[77] Cf. p. 145.

[78]

$$\mathbb{K}^3 \ni (x_1, x_2, x_3) \stackrel{L}{\longmapsto} (x_1, x_2) \in \mathbb{K}^2$$

não pode ser injetiva pois, por exemplo, $(0, 0, 1) \in \mathrm{Nu}(L)$.

[79]

$$\mathbb{K}^2 \ni (x_1, x_2) \stackrel{L}{\longmapsto} (x_1, x_2, x_1 + x_2) \in \mathbb{K}^3$$

não pode ser sobrejetiva pois, por exemplo, $(0, 0, 1) \notin \mathrm{Im}(L)$.

[80] Cf. p. 159.

[81] Cf. p. 157.

por (R8).[82] Portanto, como

$$\begin{aligned}\dim \mathrm{Nu}(L) &= \dim \mathcal{V} - \dim \mathrm{Im}(L)\\ &\geq \dim \mathcal{V} - \dim \mathcal{W}\\ &> 0,\end{aligned}$$

$\mathrm{Nu}(L) \neq \{0\}$. Logo, por (R12),[83] L não é injetiva.

DEMONSTRAÇÃO DE (b)

Como, por (I) de (R11), $\mathrm{Nu}(L)$ é um 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, como $\mathrm{Im}(L) \subsetneqq \mathcal{W}$, L não é sobrejetiva.

16. Considere uma lista formada por m subespaços de $\mathcal{V}$, digamos

$$\mathcal{U}_1, \ldots, \mathcal{U}_m, \tag{6.20}$$

e o conjunto $\mathcal{U}_1 + \cdots + \mathcal{U}_m$ formado pelas somas de m vetores tais que a j-ésima parcela de cada uma dessas somas pertence à $\mathcal{U}_j$, $j = 1, \ldots, m$, ou seja,

$$\mathcal{U}_1 + \cdots + \mathcal{U}_m := \left\{ u_1 + \cdots + u_m \,|\, u_j \in \mathcal{U}_j,\, j = 1, \ldots, m \right\}.^{84} \tag{6.21}$$

Diremos que o conjunto (6.21) é a *soma* dos subespaços da lista (6.20).

EXEMPLO

Se $\mathcal{V} = \mathbb{K}^4$, $\mathcal{U}_1 = \{x\mathbf{e}_1 \,|\, x \in \mathbb{K}\}$, $\mathcal{U}_2 = \{y\mathbf{e}_2 \,|\, y \in \mathbb{K}\}$ e $\mathcal{U}_3 = \{z\mathbf{e}_3 \,|\, z \in \mathbb{K}\}$, então $\mathcal{U}_1 + \mathcal{U}_2 + \mathcal{U}_3 = \{(x, y, z, 0) \,|\, x, y, z \in \mathbb{K}\}$.

(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) $\sum_{j=1}^{m} \mathcal{U}_j$ é uma *soma direta* dos subespaços da lista (6.20) quando

$$u_1 + \cdots + u_m = u_1' + \cdots + u_m' \Longrightarrow u_1 = u_1', \ldots, u_m = u_m'.$$

Nesse caso, denotaremos (6.21) por

$$\mathcal{U}_1 \oplus \cdots \oplus \mathcal{U}_m.$$

EXEMPLO

Se $\mathcal{V} = \mathbb{K}^3$, $\mathcal{U}_1 = \{x\mathbf{e}_1 + y\mathbf{e}_2 \,|\, x, y \in \mathbb{K}\}$ e $\mathcal{U}_2 = \{z\mathbf{e}_3 \,|\, z \in \mathbb{K}\}$, então

$$\mathcal{V} = \mathcal{U}_1 \oplus \mathcal{U}_2.$$

Para esse exemplo, se $\mathcal{U}_3 = \{\lambda\left(\mathbf{e}_2 + \mathbf{e}_3\right) \,|\, \lambda \in \mathbb{K}\}$, verifique que $\mathbb{K}^3$ não é soma direta de $\mathcal{U}_1, \mathcal{U}_2$ e $\mathcal{U}_3$.

[82] Cf. p. 154.

[83] Cf. p. 158.

[84] Nesse exercício, ao considerarmos qualquer vetor $u_1 + \cdots + u_m$, estará implícito que $u_i \in \mathcal{U}_i$, $i = 1, \ldots, m$.

(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, $\mathcal{U} \cap \mathcal{W} = \{0_{\mathcal{V}}\}$.

(e) Sejam $\mathcal{U}$ e $\mathcal{W}$ subespaços de $\mathcal{V}$ tais que $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$. Sejam $\{u_1, \ldots, u_m\}$ e $\{w_1, \ldots, w_n\}$ bases de $\mathcal{U}$ e $\mathcal{W}$, respectivamente. Demonstre que

$$\{u_1, \ldots, u_m, w_1, \ldots, w_n\}$$

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 um subespaço de $\mathcal{V}$. De fato:

- $0_{\mathcal{V}} \in \sum_{j=1}^{m} \mathcal{U}_j$;[85]
- $\lambda \in \mathbb{K}$ e $u, v \in \sum_{j=1}^{m} \mathcal{U}_j \Longrightarrow \lambda u, u + v \in \sum_{j=1}^{m} \mathcal{U}_j$.[86]

Agora, claramente, $\mathcal{U}_i \subset \sum_{j=1}^{m} \mathcal{U}_j$, para cada $i \in \{1, \ldots, m\}$.[87] Por outro lado, seja $\mathcal{U}$ um subespaço de $\mathcal{V}$ que contenha $\mathcal{U}_j$, $j = 1, \ldots, m$. Portanto,

$$\begin{aligned} u_j \in \mathcal{U}_j, j = 1, \ldots, m &\Longrightarrow u_j \in \mathcal{U}, j = 1, \ldots, m \\ &\Longrightarrow u_1 + \cdots + u_m \in \mathcal{U} \text{ (pois } \mathcal{U} \text{ é subespaço)} \\ &\Longrightarrow \sum_{j=1}^{m} \mathcal{U}_j \subset \mathcal{U}. \end{aligned}$$

(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

$$\begin{aligned} u_1 + \cdots + u_m &= 0_{\mathcal{V}} \\ &= 0_{\mathcal{V}} + \cdots + 0_{\mathcal{V}}, \end{aligned}$$

[85] Verifique!
[86] Idem!
[87] De fato, basta considerarmos somas $u_1 + \cdots + u_m$ onde todos os u's sejam nulos, exceto um deles.

onde estamos considerando m parcelas nulas na última igualdade (de cima para baixo),

$$u_j = 0_{\mathcal{V}},\ j = 1, \ldots, m.$$

Para a recíproca, considere a seguinte hipótese:

$$0_{\mathcal{V}} = u_1 + \cdots + u_m \Longrightarrow u_j = 0_{\mathcal{V}},\ j = 1, \ldots, m.$$

Verificaremos que a soma é direta pela definição dada no item (b) supracitado. Assim, caso

$$u_1 + \cdots + u_m = u_1' + \cdots + u_m',$$

$$0_{\mathcal{V}} = u_1 - u_1' + \cdots + u_m - u_m'.$$

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,[88] $v = 0_{\mathcal{V}}$. 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

$$0_{\mathcal{V}} = u + w \text{ com } u \in \mathcal{U} \text{ e } w \in \mathcal{W} \tag{6.22}$$

é aquele em que $u = w = 0_{\mathcal{V}}$. De fato, caso (6.22) 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

$$v = u + w. \tag{6.23}$$

Por outro lado, considerando as bases de $\mathcal{U}$ e $\mathcal{W}$ dadas, existe uma lista de escalares $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n$ pertencentes ao corpo $\mathbb{K}$ tal que

$$u = \alpha_1 u_1 + \cdots + \alpha_m u_m \quad \text{e} \quad w = \beta_1 w_1 + \cdots + \beta_n w_n. \tag{6.24}$$

Ao substituirmos (6.24) em (6.23), vemos que v é uma CL dos $m + n$ vetores dados. Vamos demonstrar, agora, que estes $m + n$ vetores são LI. Assim, seja

$$\alpha_1 u_1 + \cdots \alpha_m u_m + \beta_1 w_1 + \cdots + \beta_n w_n = 0_{\mathcal{V}}, \tag{6.25}$$

onde os $m + n$ escalares, como antes, estão em $\mathbb{K}$. Denote u e w como em (6.24). Portanto, $0_{\mathcal{V}} = u + w$ e, como a soma é direta,

$$u = w = 0_{\mathcal{V}}, \tag{6.26}$$

pelo item (c) supracitado. Então, ao substituirmos (6.26) em (6.24) e utilizarmos a hipótese da independência linear dos m vetores da base de $\mathcal{U}$ dada e dos n vetores da base de $\mathcal{W}$ dada, todos os $m + n$ escalares da CL nula (6.25) devem ser nulos.

(f)
i. Note que toda matriz pertencente ao subespaço $\mathcal{U}$ é uma CL da forma

$$x \begin{pmatrix} E & D \\ 0 & 0 \end{pmatrix} + y \begin{pmatrix} 0 & 0 \\ D & F \end{pmatrix},$$

[88] Cf. o item (c) supracitado!

com $x, y \in \mathbb{K}$. Assim, como as duas matrizes dessa CL são LI,[89] uma base $\mathcal{B}$ 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.[90] Considere, então, o conjunto $\mathcal{B}'$ formado por elas e pelas duas matrizes da base $\mathcal{B}$ supracitada. Então, por uma lado, como a identidade

$$\begin{pmatrix} Ex & Dx \\ Dy & Fy \end{pmatrix} = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}$$

acarreta $x = y = a = b = 0$,

$$\mathcal{U} \cap W = \{0_{\mathbb{K}^{2\times 2}}\}$$

e, portanto, pelo item (d) desse exercício, a soma $\mathcal{U} + \mathcal{W}$ é direta. Por outro, $\mathcal{B}'$ é uma base de $\mathcal{U} \bigoplus \mathcal{W}$, pois seus quatro vetores são LI. De fato, considere o determinante da matriz (de ordem 4) cujas colunas sejam dadas pelos vetores colunas formados pelas 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 cada um dos subespaços da lista (6.20) seja finitamente gerado e a soma (6.21) 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 da seguinte afirmação:

[89] De fato, nenhuma delas é um múltiplo escalar da outra.
[90] Idem.

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 $u_i \in \mathcal{U}_i$, $i = 1, \ldots, m$. Logo, 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 $\mathcal{U}$. Assim, resta verificarmos que os vetores de $\mathcal{B}$ são LI. Para isso, considere a seguinte CL nula:

$$c_{1_1} u_{1_1} + \cdots + c_{1_{n_1}} u_{1_{n_1}} + \cdots + c_{m_1} u_{m_1} + \cdots + c_{m_{n_m}} u_{m_{n_m}} = 0, \tag{6.27}$$

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 $j = 1, \ldots, n_i$. Portanto, $c_{i_1} u_{i_1} + \cdots + c_{i_{n_i}} u_{i_{n_i}} = 0$, $i = 1, \ldots, m$, pelo item (c) do exercício supracitado. Assim, como $\mathcal{B}_i$ é uma base de $\mathcal{U}_i$, $i = 1, \ldots, m$, cada escalar c_{i_j} de (6.27) 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 e $L \in \mathcal{L}(\mathcal{V})$, 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

Caso $L \in \mathcal{L}(\mathcal{V})$ e $\mathcal{U}$ seja um subespaço de $\mathcal{V}$, definimos 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. 159.
[2] Cf. p. 158.
[3] Cf. p. 159.
[4] Cf. p. 154.

ou seja, 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:

$$\begin{aligned} v \in \mathrm{Nu}(S) &\Longrightarrow S(v) = 0 \\ &\Longrightarrow T(S(v)) = T(0) \\ &\Longrightarrow TS(v) = 0 \\ &\Longrightarrow ST(v) = 0 \\ &\Longrightarrow S(T(v)) = 0 \\ &\Longrightarrow T(v) \in \mathrm{Nu}(S) \end{aligned}$$

  e

$$\begin{aligned} v \in \mathrm{Im}(S) &\Longrightarrow \exists w \in V \text{ tal que } v = S(w) \\ &\Longrightarrow T(v) = T(S(w)) \\ &\Longrightarrow T(v) = TS(w) \\ &\Longrightarrow T(v) = ST(w) \\ &\Longrightarrow T(v) = S(T(w)) \\ &\Longrightarrow T(v) = S(u) \text{ com } u = T(w) \\ &\Longrightarrow \exists u \in V \text{ tal que } T(v) = S(u) \\ &\Longrightarrow T(v) \in \mathrm{Im}(S). \end{aligned}$$

- Se $T \in \mathcal{L}(\mathcal{V})$ e $\mathcal{U}_i$ é um subespaço de $\mathcal{V}$ invariante por T, $i = 1, \ldots, m$, então $\sum_{i=1}^{m} \mathcal{U}_i$ é invariante por T.[5]

[5] Confira o último exercício da seção 6.4.

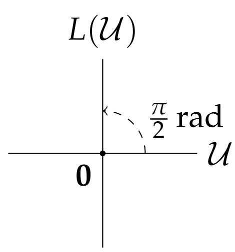
De fato,

$$\begin{aligned} v \in \sum_{i=1}^{m} \mathcal{U}_i &\Longrightarrow v = \sum_{i=1}^{m} u_i, \text{ onde } u_i \in \mathcal{U}_i,\ i = 1, \ldots, m \\ &\Longrightarrow T(v) = T\left(\sum_{i=1}^{m} u_i\right) \\ &\Longrightarrow T(v) = \sum_{i=1}^{m} T(u_i)\text{ , onde } T(u_i) \in \mathcal{U}_i,\ i = 1, \ldots, m, \text{ pois cada } \mathcal{U}_i \text{ é invariante por } T \\ &\Longrightarrow T(v) \in \sum_{i=1}^{m} \mathcal{U}_i. \end{aligned}$$

CONTRA EXEMPLO

Se $\mathcal{V} = \mathbb{R}^2$, $L = R_{\pi/2}$ é a rotação de $\pi/2$ radianos em torno da origem e $\mathcal{U}$ é uma reta que passa pela origem, então $\mathcal{U}$ não é invariante por L.[6] Por exemplo, caso $\mathcal{U}$ seja o eixo das abscissas, $L(\mathcal{U})$ é o eixo das ordenadas, conforme a figura 7.1.

Figura 7.1: Exemplo de $\mathcal{U}$ não invariante por L



7.2 Autovalores e autovetores

Como no capítulo 4, dizemos que $\lambda \in \mathbb{K}$ é um *autovalor de* $L \in \mathcal{L}(\mathcal{V})$ quando existe 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

  $$L(u) = \lambda u.$$

[6] Confira o exercício 3 da seção 4.1.

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),

$$\begin{aligned} v \text{ é autovetor de } L \text{ associado à } \lambda &\iff Lv = \lambda v \\ &\iff Lv = \lambda I v \\ &\iff Lv - \lambda I v = 0_{\mathcal{V}} \\ &\iff (L - \lambda I)v = 0_{\mathcal{V}}. \end{aligned}$$

EXEMPLOS

Confira os exemplos da seção 4.2.[7]

(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).[8] As equivalências $3 \iff 4 \iff 5$ seguem de (R18).[9]

(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),[10] podemos considerar o menor índice j tal que

$$v_j \in [v_1, \ldots, v_{j-1}]\,. \tag{7.2}$$

Assim, existe uma lista de escalares $\alpha_1, \ldots, \alpha_{j-1}$ tal que

$$v_j = \alpha_1 v_1 + \cdots + \alpha_{j-1} v_{j-1}. \tag{7.3}$$

[7] 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.*

[8] Cf. p. 158.

[9] Cf. p. 179.

[10] Cf. p. 150.

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,

$$\alpha_1 = \cdots = \alpha_{j-1} = 0.$$

Portanto, por (7.3), $v_j = 0_{\mathcal{V}}$ e chegamos a uma contradição, pois v_j é um autovetor, por hipótese.

(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).[11]

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.*[12] 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 $j \neq 0$,

$$T^j := \underbrace{T \cdots T}_{j \text{ fatores}};$$[13]

Assim, $T^0 = I$, $T^1 = T$, $T^2 = TT$, $T^3 = TTT$, etc.

- Caso

$$p(t) = \sum_{j=0}^{m} a_j t^j \in \mathcal{P}(\mathbb{K}),$$

$$p(T) := \sum_{j=0}^{m} a_j T^j.$$

[11]Cf. p. 150.

[12]Confira a seção 4.2.

[13]Confira o exercício 11 da seção 6.4.

EXEMPLO

Se $p(t) = 3 - 2t + t^2$, então $p(T) = 3I - 2T + T^2$.

EXERCÍCIOS

1. Demonstre que
$$T^i T^j = T^{i+j} \text{ e } \left(T^i\right)^j = T^{ij}.$$[14]

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.[15]

EXEMPLOS

· Se $T^2 = I$ e -1 não é autovalor de T, então $T = I$;[16]

[14]Caso T seja invertível, essas duas igualdades também são válidas para potências negativas. De fato, para cada inteiro k, podemos definir
$$T^{-k} := \left(T^{-1}\right)^k.$$

[15]De fato,
$$\begin{aligned} p(T)q(T) &= (pq)(T) \\ &= (qp)(T) \\ &= q(T)p(T). \end{aligned}$$

[16]De fato, suponha que $T - I \neq 0$. Assim, existe um vetor não nulo $v \in \mathcal{V}$ 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 desse exemplo.

· 9 é autovalor de $T^2 \Longleftrightarrow$ 3 é autovalor de T ou -3 é autovalor de T.[17]

7.2.3 Caso $\mathcal{V} \neq \{0_{\mathcal{V}}\}$ seja um espaço vetorial complexo finitamente gerado, $T \in \mathcal{L}(\mathcal{V})$ tem autovalor

Embora esse resultado não seja válido para espaços vetoriais reais,[18] podemos demonstrá-lo (para espaços vetoriais complexos) 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 AXLER, mencionado no capítulo 1.

DEMONSTRAÇÃO DE 7.2.3

Sejam $v \in \mathcal{V} - \{0_{\mathcal{V}}\}$ e $n = \dim \mathcal{V}$. Logo, os $n+1$ vetores da lista

$$v, T(v), \ldots, T^n(v)$$

são LD, por (R2).[19] Assim, existe uma lista $a_0, a_1, \ldots, a_n$ de escalares complexos, não todos nulos, tal que

$$\sum_{j=0}^{n} a_j T^j(v) = 0_{\mathcal{V}}. \tag{7.7}$$

Caso m seja o máximo do conjunto

$$\{j \in \{0,1,\ldots,n\} \,:\, a_j \neq 0\}\,,$$

$m \neq 0$. 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 (*),

$$p(t) = \sum_{j=0}^{n} a_j t^j$$

---

[17] De fato, como

$$\begin{aligned} T^2 - 9I &= (T+3I)(T-3I) \\ &= (T-3I)(T+3I), \end{aligned}$$

$$\begin{aligned} 9 \text{ é autovalor de } T^2 &\Longleftrightarrow T^2 - 9I \text{ não é injetiva} \\ &\Longleftrightarrow (T-3I) \text{ não é injetiva ou } (T+3I) \text{ não é injetiva} \\ &\Longleftrightarrow 3 \text{ é autovalor de } T \text{ ou } -3 \text{ é autovalor de } T, \end{aligned}$$

por (R19).

[18] Cf. (5.2), p. 134.

[19] Cf. p. 150.

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).[20] Assim, também por (R19), algum λ_i é um autovalor de T.

7.2.4 Matrizes triangulares e operadores

(R22) Matrizes triangulares e invariância de bases

Se $T \in \mathcal{L}(\mathcal{V})$ e $\mathcal{B} = \{v_1, \dots, v_n\}$ é uma base de $\mathcal{V}$, então as condições seguintes são equivalentes:

1. *$\mathcal{M}(T, \mathcal{B})$ é triangular superior;*
2. *$T(v_j) \in [v_1, \dots, v_j]$, $j = 1, \dots, n$;*
3. *$[v_1, \dots, v_j]$ é invariante por T, $j = 1, \dots, n$.*

EXEMPLO

Considere $\mathcal{V} = \mathbb{K}^2$ e

$$\mathcal{M}(T, \mathcal{B}) = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}.$$

Como

$$\begin{array}{rcrcl} T(v_1) & = & v_1 & & \text{e} \\ T(v_2) & = & 2v_1 & + & 3v_2, \end{array}$$

a condição 2 supracitada é satisfeita. Logo, como

$$\begin{aligned} T(\alpha_1 v_1 + \alpha_2 v_2) &= \alpha_1 T(v_1) + \alpha_2 T(v_2) \\ &= (\alpha_1 + 2\alpha_2) v_1 + 3\alpha_2 v_2, \end{aligned}$$

onde $\alpha_1, \alpha_2 \in \mathbb{K}$, a condição 3 supracitada também é satisfeita.

DEMONSTRAÇÃO DE (R22)

$1 \Longleftrightarrow 2$ Segue de

$$\begin{aligned} &T(v_1) = a_{11}v_1; \\ &T(v_2) = a_{12}v_1 + a_{22}v_2; \\ &T(v_3) = a_{13}v_1 + a_{23}v_2 + a_{33}v_3; \\ &\cdots \\ &T(v_n) = a_{1n}v_1 + a_{2n}v_2 + a_{3n}v_3 + \cdots + a_{nn}v_n. \end{aligned}$$

[20] Cf. p 182.

$\boxed{3 \Longrightarrow 2}$ Óbvia.

$\boxed{2 \Longrightarrow 3}$ Fixe $j \in \{1, \ldots, n\}$. Assim, pela condição 2,

$$\begin{aligned} &T(v_1) \in [v_1] \subset [v_1, \ldots, v_j]\,; \\ &T(v_2) \in [v_1, v_2] \subset [v_1, \ldots, v_j]\,; \\ &\ldots \\ &T(v_j) \in [v_1, \ldots, v_j]\,. \end{aligned}$$

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,[21] considere

$$\mathcal{U} := \operatorname{Im}(T - \lambda I).$$

Note que:

- $T - \lambda I$ não é sobrejetiva, por (R19).[22] 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, $\mathcal{M}(T|_{\mathcal{U}}, \mathcal{B}')$ 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 $\mathcal{V}$,[23] temos

$$T(v_j) = (T - \lambda I)(v_j) + \lambda v_j,$$

com $(T - \lambda I)(v_j) \in \mathcal{U} = [u_1, \ldots, u_n]$,[24] $j = 1, \ldots, m$. 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

[21] Cf. 7.2.3, p. 185.
[22] Cf. p. 182.
[23] Cf. (R6), p. 153.
[24] Pela definição de $\mathcal{U}$.

· 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.[25] 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:

$$\boxed{T \text{ é invertível} \iff \lambda_i \neq 0,\, i = 1, \ldots, n.}$$

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.[26]

$\boxed{\Longleftarrow}$ 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 $\lambda_3 \neq 0$ e, para escalares β e γ adequados,

$$T\left(\frac{1}{\lambda_3} v_3\right) = \frac{\beta}{\lambda_3} v_1 + \frac{\gamma}{\lambda_3} v_2 + v_3,^{27}$$

$v_3 \in \mathrm{Im}(T)$.

[25] Confira a subseção 5.3.1.
[26] Obviamente, essas entradas são nulas.
[27] Por (7.11).

Ao repetirmos esse raciocínio mais $n-3$ vezes, concluímos que $v_j \in \operatorname{Im}(T)$, $j = 1, \ldots, n$, e, como $\mathcal{B}$ é uma base de $\mathcal{V}$, $\operatorname{Im}(T) = \mathcal{V}$, isto é, T é sobrejetivo, ou seja, T é invertível, por (R18).[28]

$\boxed{\Longrightarrow}$ 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, $\operatorname{Nu}(T) \neq \{0\}$, isto é, T não é injetivo,[29] ou seja, T não é invertível,[30] contradizendo a suposição supracitada. Agora, considere que, para algum índice $j > 1$, $\lambda_j = 0$. Logo,

$$\operatorname{Im}\Big(T|_{[v_1,\ldots,v_j]}\Big) \subset [v_1, \ldots, v_{j-1}]\,,$$

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,[31] 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,[32] contradizendo a suposição supracitada. Portanto, $\lambda_i \neq 0$, $i = 1, \ldots, n$.

(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:

$$\boxed{\lambda \text{ é autovalor de } T \Longleftrightarrow \lambda \in \{\mathsf{T}_{11}, \ldots, \mathsf{T}_{nn}\}.}$$

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),[33]

$$\mathcal{M}(T - \lambda I) = \begin{bmatrix} \lambda_1 - \lambda & & * \\ & \ddots & \\ 0 & & \lambda_n - \lambda \end{bmatrix},$$

$$\begin{aligned} \lambda \text{ é autovalor de } T &\Longleftrightarrow T - \lambda I \text{ não é invertível} \\ &\Longleftrightarrow \lambda_i = \lambda, \text{ para algum } i \in \{1, \ldots, n\}\,, \end{aligned}$$

por (R19) e (R24).[34]

OBSERVAÇÃO

Noutros livros, (R25) é (geralmente) demonstrado via determinantes.

[28] Cf. p. 179.
[29] Cf. (R12), p. 158.
[30] Cf. (R18), p. 179.
[31] Confira o penúltimo exercício da seção 6.4.
[32] Cf. (R18).
[33] Cf. p. 188.
[34] Cf. pp. 182 e 188.

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 $\mathcal{M}(T, \mathcal{B}) = \mathsf{D}$,[35] e $d_{ii} = \lambda_i$, $i = 1, \ldots, n$,

$$\boxed{\lambda \text{ é autovalor de } T \Longleftrightarrow \lambda \in \{\lambda_1, \ldots, \lambda_n\}}$$

e $\mathcal{B}$ é constituída de autovetores de T.[36]

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

$$\boxed{\lambda_1, \ldots, \lambda_n \text{ é uma lista de autovalores distintos de } T \in \mathcal{L}(\mathcal{V}) \text{ e } \dim \mathcal{V} = n \Longrightarrow T \text{ é diagonalizável.}}$$

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).[37] Portanto, $\mathcal{B} = \{v_1, \ldots, v_n\}$ é uma base de $\mathcal{V}$, por (R9).[38] 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

Doravante, o último exercício da seção 6.4 é imprescindível.

---

[35] Cf. (3.2), p. 58.

[36] De fato, $T(v_i) = \lambda_i v_i$, $i = 1, \ldots, n$, pois

$$\mathsf{D} = \begin{bmatrix} \lambda_1 & & & & \\ & \lambda_2 & & \text{zeros} & \\ & & \ddots & & \\ & \text{zeros} & & \lambda_{n-1} & \\ & & & & \lambda_n \end{bmatrix}.$$

[37] Cf. p. 182.

[38] Cf. p. 154.

(R28) Somas diretas e autovalores

Caso $\mathcal{V}$ seja finitamente gerado, $T \in \mathcal{L}(\mathcal{V})$ e $\lambda_1, \dots, \lambda_m$ 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

$$\sum_{i=1}^{m} \dim \mathrm{Nu}(T - \lambda_i I) \leq \dim \mathcal{V}.$$

EXEMPLO

Confira o exercício 2 da seção 7.5.

DEMONSTRAÇÃO DE (R28)

Para demonstrarmos que a soma supracitada é direta, consideremos

$$\sum_{i=1}^{m} u_i = 0_{\mathcal{V}}, \tag{7.13}$$

onde

$$u_i \in \mathrm{Nu}(T - \lambda_i I)\,, i = 1, \dots, m. \tag{7.14}$$

Note que, caso alguns dos vetores listados em (7.14) fossem não nulos, eles seriam LI, por (R20).[39] Como, nesse caso, teríamos uma contradição de (7.13),

$$u_i = 0_{\mathcal{V}}, i = 1, \dots, m.$$

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 um 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 $\lambda_1, \dots, \lambda_m$ 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. $\dim \mathcal{V} = \sum_{i=1}^{m} \dim \mathrm{Nu}(T - \lambda_i I)$.

[39] Cf. p. 182.

EXEMPLO

Conforme os três primeiros exemplos da seção 4.2, onde

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

temos

$$\mathbb{R}^2 \ni \mathbf{x} \longmapsto L(\mathbf{x}) = A\mathbf{x} \in \mathbb{R}^2$$

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 $\mathrm{Nu}(T - \lambda_i I)$, $i = 1, \ldots, m$.[40]

$\boxed{1 \Longrightarrow 3}$ Caso T seja diagonalizável, podemos considerar que a base $\mathcal{B}$ de (R26) é da forma (7.15). Portanto, a condição 3 supracitada é válida.

$\boxed{3 \Longrightarrow 2}$ Como a soma direta dos núcleos (mencionada na última nota de rodapé) é um subespaço de $\mathcal{V}$,[41] a condição 2 supracitada é válida, pela condição 3 supracitada e 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 $\mathcal{V}$.[42]

(R30) Coordenadas na base ortonormal

$$\boxed{\textit{O escalar } c_j \textit{ da } \mathrm{CL}\ u = \textstyle\sum_{i=1}^{r} c_i u_i \textit{ é dado por } c_j = \langle u, u_j \rangle,\ j = 1 \ldots, r.}$$

EXEMPLOS

Confira (2.22) (da página 41) e os exercícios 7 e 8 da seção 2.7.

DEMONSTRAÇÃO DE (R30)

[40] A soma desses núcleos é direta, por (R28).

[41] Confira o item (a) do último exercício da seção 6.4.

[42] A existência dessa base é uma consequência do processo de Gram-Schmidt!

Para cada índice $j \in \{1, \dots, r\}$, 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 um subespaço de* $\mathcal{V}$*;*
2. $\emptyset \neq \mathcal{W} \subset \mathcal{U} \Longrightarrow \mathcal{U}^{\perp} \subset \mathcal{W}^{\perp}$;
3. $\mathcal{U} = \mathcal{B} \Longrightarrow \mathcal{U}^{\perp} = [u_1, \dots, u_r]^{\perp}$;
4. $v \in \mathcal{U} \cap \mathcal{U}^{\perp} \Longrightarrow v = 0_{\mathcal{V}}$.

EXEMPLOS

Confira o exercício 14 da seção 2.7.

DEMONSTRAÇÃO DO ITEM 1 DE (R31)

$0_{\mathcal{V}} \in \mathcal{U}^{\perp}$, isto é, para $u \in \mathcal{U}$ arbitrário, $\langle 0_{\mathcal{V}}, u \rangle = 0$, pois

$$\begin{aligned}\langle 0_{\mathcal{V}}, u \rangle &= \langle 0_{\mathcal{V}} + 0_{\mathcal{V}}, u \rangle \\ &= \langle 0_{\mathcal{V}}, u \rangle + \langle 0_{\mathcal{V}}, u \rangle.\end{aligned}$$

Além disso, 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)

$$\begin{aligned}v \in \mathcal{U}^{\perp} &\Longrightarrow \langle v, u \rangle = 0 \text{ para todo } u \in \mathcal{U} \\ &\Longrightarrow \langle v, u \rangle = 0 \text{ para todo } u \in \mathcal{W}, \text{ pois } \mathcal{W} \subset \mathcal{U} \\ &\Longrightarrow v \in \mathcal{W}^{\perp}.\end{aligned}$$

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, se $v \in \mathcal{U}^\perp$ e u é uma CL dos r vetores de $\mathcal{U}$, ou seja, existe uma lista de escalares, digamos, $c_1, \ldots, c_r$, tal que

$$u = c_1 u_1 + \ldots + c_r u_r,$$

então, como

$$\begin{aligned} \langle v, u \rangle &= \overline{c_1} \langle v, u_1 \rangle + \cdots + \overline{c_r} \langle v, u_r \rangle \\ &= \overline{c_1} \cdot 0 + \cdots + \overline{c_r} \cdot 0 \\ &= 0, \end{aligned}$$

$v \in [u_1, \ldots, u_r]^\perp$. Portanto,

$$\mathcal{U}^\perp \subset [u_1, \ldots, u_r]^\perp.$$ [43]

DEMONSTRAÇÃO DO ITEM 4 DE (R31)

$$\begin{aligned} v \in \mathcal{U} \cap \mathcal{U}^\perp &\Longrightarrow \langle v, v \rangle = 0 \\ &\Longrightarrow v = 0_{\mathcal{V}}. \end{aligned}$$

(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 da seção 2.7.

DEMONSTRAÇÃO DE (R32)

Na primeira parte 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

$$w = v - u \in \mathcal{U}^\perp.$$ [44]

Portanto, basta demonstrarmos que

$$w \in \mathcal{B}^\perp,$$

[43] Essa demonstração também é válida para qualquer subconjunto finito $\mathcal{U} \neq \mathcal{B}$ de $\mathcal{V}$.

[44] De fato, nesse caso,

$$v = u + w, \tag{7.19}$$

com $u \in \mathcal{U}$ e $w \in \mathcal{U}^\perp$.

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 (7.18).
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 $\mathcal{V}$, $0_{\mathcal{V}} \in \mathcal{U} \cap \mathcal{U}^{\perp}$, isto é,

$$\{0_{\mathcal{V}}\} \subset \mathcal{U} \cap \mathcal{U}^{\perp}.$$

(R33) Corolário de (R32)

$$\boxed{\mathcal{V} \textit{ é finitamente gerado e } \mathcal{U} \textit{ é subespaço de } \mathcal{V} \Longrightarrow \dim \mathcal{V} = \dim \mathcal{U} + \dim \mathcal{U}^{\perp}.}$$

De fato, (R33) é uma consequência de (R3),[45] 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

$$\mathbf{y}_i \cdot \mathbf{x}_j = 0, i, j \in \{1, 2\},$$

[45] Cf. p. 151.

$\{\mathbf{y}_1, \mathbf{y}_2\}$ 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

$$\begin{aligned} \mathbf{y}_1 &= (15, -9, 1, 0), \\ \mathbf{y}_2 &= (-14, 9, 0, 1), \\ \mathbf{y}_1' &= \mathbf{y}_1, \\ \mathbf{y}_2' &= \mathbf{y}_2 - \frac{\mathbf{y}_2 \cdot \mathbf{y}_1'}{\|\mathbf{y}_1'\|^2}\mathbf{y}_1', \\ \mathbf{y}_1'' &= \mathbf{y}_1'/\|\mathbf{y}_1'\| \text{ e} \\ \mathbf{y}_2'' &= \mathbf{y}_2'/\|\mathbf{y}_2'\|, \end{aligned}$$

$\{\mathbf{y}_1'', \mathbf{y}_2''\}$ é uma base ortonormal de $\mathcal{U}^\perp$.

$\boxed{\text{DEFINIÇÃO}}$

Seja $\mathcal{U}$ o subespaço dado em (R32). Para $v \in \mathcal{V}$, seja u o vetor definido em (7.18).[46] 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 7.2.

$\boxed{\text{EXEMPLO}}$

Confira o item (b) do exercício 5 que antecede a subseção 4.1.1.

(R34) Propriedades das projeções ortogonais

Considere a definição supracitada. Assim:

---

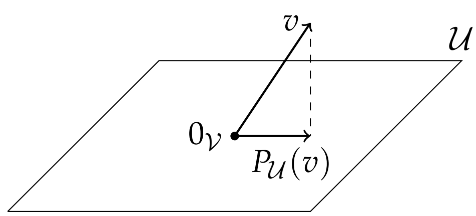
[46]Cf. p. 194.

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 7.2.

Figura 7.2: Projeção ortogonal de $v \in \mathcal{V}$ sobre $\mathcal{U}$



EXEMPLO

Confira o item (c) do exercício 4.2 da seção 6.4.

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).[47]

DEMONSTRAÇÃO DO ITEM 3 DE (R34)

$$\begin{aligned} v_1, v_2 \in \mathcal{V} &\Longrightarrow v_1 - P_{\mathcal{U}}(v_1)\,,\, v_2 - P_{\mathcal{U}}(v_2) \in \mathcal{U}^{\perp}, \text{ pelo item 2 de (R34)} \\ &\Longrightarrow (v_1 + v_2) - [P_{\mathcal{U}}(v_1) + P_{\mathcal{U}}(v_2)] \in \mathcal{U}^{\perp} \\ &\Longrightarrow v_1 + v_2 = [P_{\mathcal{U}}(v_1) + P_{\mathcal{U}}(v_2)] + w_{v_1+v_2}, \\ &\qquad \text{onde a parcela entre colchetes pertence à } \mathcal{U} \\ &\qquad \text{(pela definição supracitada) e } w_{v_1+v_2} \in \mathcal{U}^{\perp} \\ &\Longrightarrow P_{\mathcal{U}}(v_1 + v_2) = P_{\mathcal{U}}(v_1) + P_{\mathcal{U}}(v_2)\,; \end{aligned}$$

[47] Cf. p. 194.

$$\begin{aligned}
v \in \mathcal{V} \text{ e } \lambda \in \mathbb{K} &\Longrightarrow \lambda\,(v - P_{\mathcal{U}}(v)) \in \mathcal{U}^{\perp}, \text{ pelo item 2 de (R34)}\\
&\Longrightarrow \lambda v = [\lambda P_{\mathcal{U}}(v)] + w_{\lambda v},\\
&\quad\ \text{onde a parcela entre colchetes pertence à } \mathcal{U}\\
&\quad\ \text{(pela definição supracitada) e } w_{\lambda v} \in \mathcal{U}^{\perp}\\
&\Longrightarrow P_{\mathcal{U}}(\lambda v) = \lambda P_{\mathcal{U}}(v).
\end{aligned}$$

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

$$\begin{aligned}
\|v_1 + v_2\|^2 &= \langle v_1 + v_2, v_1 + v_2 \rangle\\
&= \langle v_1, v_1 + v_2 \rangle + \langle v_2, v_1 + v_2 \rangle\\
&= \overline{\langle v_1 + v_2, v_1 \rangle} + \overline{\langle v_1 + v_2, v_2 \rangle}\\
&= \overline{\langle v_1, v_1 \rangle} + \overline{\langle v_2, v_1 \rangle} + \overline{\langle v_1, v_2 \rangle} + \overline{\langle v_2, v_2 \rangle}\\
&= \overline{\|v_1\|^2} + \langle v_1, v_2 \rangle + \overline{\langle v_1, v_2 \rangle} + \overline{\|v_2\|^2}.
\end{aligned}$$

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,

$$\begin{aligned}
\|v - x\|^2 &= \|(v - P_{\mathcal{U}}(v)) + (P_{\mathcal{U}}(v) - x)\|^2\\
&= \|v - P_{\mathcal{U}}(v)\|^2 + \|P_{\mathcal{U}}(v) - x\|^2\\
&\geq \|v - P_{\mathcal{U}}(v)\|^2
\end{aligned}$$

e, assim,

$$\begin{aligned}
\|v - x\| = \|v - P_{\mathcal{U}}(v)\| &\Longleftrightarrow \|P_{\mathcal{U}}(v) - x\| = 0\\
&\Longleftrightarrow x = P_{\mathcal{U}}(v).
\end{aligned}$$

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,

$$\begin{aligned}
\mathbf{u} &= P_{\mathcal{U}}(1,2,3,4)\\
&= \left((1,2,3,4) \cdot \mathbf{x}_1''\right)\mathbf{x}_1'' + \left((1,2,3,4) \cdot \mathbf{x}_2''\right)\mathbf{x}_2''.
\end{aligned}$$

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 $F = 6$.[48]
  - 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

$$Y = \begin{bmatrix}a & b\\ c & d\end{bmatrix} \in \mathcal{U}^\perp,$$

[48] 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 *on-line*. 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.

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 $d = 1 \Longrightarrow Y = Y_2$. [49]

Agora, ao considerarmos $C = 3$, $D = 4$, $E = 5$ e $F = 6$, temos

$$\begin{aligned}
Y_1' &= Y_1 \\
&= \begin{bmatrix} 3 & -4 \\ 1 & 0 \end{bmatrix}, \\
Y_2' &= Y_2 - \frac{\langle Y_2, Y_1' \rangle}{\|Y_1'\|^2} Y_1' \\
&= \begin{bmatrix} 2 & -2 \\ 0 & 1 \end{bmatrix} - \frac{7}{13} \begin{bmatrix} 3 & -4 \\ 1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 5/13 & 2/13 \\ -7/13 & 1 \end{bmatrix}, \\
Y_1'' &= \frac{Y_1'}{\|Y_1'\|} \\
&= \frac{1}{\sqrt{26}} \begin{bmatrix} 3 & -4 \\ 1 & 0 \end{bmatrix} \ \text{ e} \\
Y_2'' &= \frac{Y_2'}{\|Y_2'\|} \\
&= \frac{13}{\sqrt{247}} \begin{bmatrix} 5/13 & 2/13 \\ -7/13 & 1 \end{bmatrix},
\end{aligned}$$

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 $\frac{\partial f}{\partial d} = 0$.[50] Em seguida, para garantirmos que o ponto (c_0, d_0) calculado seja um minimizador local de f,[51] podemos utilizar, nesse

[49] Como Y_1 e Y_2 não são múltiplas uma da outra, elas são LI.

[50] Para o exercício 2 supracitado, existe (apenas) um ponto crítico.

[51] Esse ponto é, de fato, o mínimo global!

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).[52]

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.[53]

RESOLUÇÃO

Como $p(0) = 0$, $c = 0$. Assim, claramente, $p \in \mathcal{U} = [q, r]$, onde $q(x) = x$ e $r(x) = x^2$.[54] 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

$$\begin{aligned}
q_1(x) &= q(x) \\
&= x, \\
r_1(x) &= r(x) - \frac{\int_0^1 r(x)q_1(x)\,dx}{\int_0^1 q_1(x)^2\,dx}\, q_1(x) \\
&= x^2 - \frac{\int_0^1 x^3\,dx}{\int_0^1 x^2\,dx}\, x \\
&= x^2 - \frac{1/4}{1/3} x \\
&= x^2 - \frac{3}{4} x.
\end{aligned}$$

[52] A propósito, para um estudo de *otimização não linear*, no âmbito de graduação, confira o meu livro de cálculo de várias variáveis, ISBN 978-65-980575-5-8, disponibilizado gratuitamente em formato digital, na página de acesso aberto do *site* da editora da UFPR.

[53] Confira a primeira nota de rodapé do exercício 2 supracitado.

[54] Verifique que q e r são LI.

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),

$$\begin{aligned}
p(x) &= \left(\int_0^1 (8+12x)(\sqrt{3}x)\,dx\right)\sqrt{3}x + \left(\int_0^1 (8+12x)(\sqrt{5}(4x^2-3x))\,dx\right)(\sqrt{5}(4x^2-3x)) \\
&= 3\left(8\int_0^1 x\,dx + 12\int_0^1 x^2\,dx\right)x \\
&\qquad + 5\left(32\int_0^1 x^2\,dx - 24\int_0^1 x\,dx + 48\int_0^1 x^3\,dx - 36\int_0^1 x^2\,dx\right)(4x^2-3x) \\
&= 3(4+4)x + 5((32/3) - 12 + 12 - 12)(4x^2 - 3x) \\
&= 24x - (20/3)(4x^2 - 3x) \\
&= 44x - (80/3)x^2.
\end{aligned}$$

4. Para $C, D, E \in \mathbb{R}$, considere $\mathcal{V} = \mathcal{C}[1, C+2]$,[55] 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

$$[1, C+1] \ni x \mapsto f(x) = (D+1)x^{\frac{1}{E+1}} \in \mathbb{R}.^{56}$$

Determine o polinômio $p \in \mathcal{U} = \mathcal{P}_2(\mathbb{R})$ mais próximo de f.[57,58]

RESOLUÇÃO

Como p é a projeção ortogonal de f sobre $\mathcal{U}$,

$$p = \langle f, p_1\rangle p_1 + \langle f, p_2\rangle p_2, \tag{7.23}$$

onde $\{p_1, p_2\}$ é uma base ortonormal de $\mathcal{U}$, que pode ser obtida pela aplicação de Gram-Schmidt na

[55] $f \in \mathcal{V} \iff f : [1, C+2] \to \mathbb{R}$ é contínua.

[56] Note que $f \in \mathcal{V}$.

[57] $\mathcal{U}$ é o espaço dos polinômios constantes ou lineares.

[58] Confira a primeira nota de rodapé do exercício 2 supracitado.

base canônica de $\mathcal{U}$, ou seja, $\{1, x\}$. Assim, como $p_i = q_i/\|q_i\|$, $i = 1, 2$, com

$$\begin{aligned} q_1(x) &= 1 \qquad \text{e} \\ q_2(x) &= x - \frac{\langle x, 1\rangle}{\langle 1, 1\rangle} 1 \\ &= x - \frac{\int_1^{C+2} x\,dx}{\int_1^{C+2} dx} \\ &= x - \frac{\frac{(C+2)^2-1}{2}}{C+1} \\ &= x - \frac{C+3}{2}, \end{aligned}$$

$$\begin{aligned} p_1(x) &= \frac{1}{\sqrt{\langle 1, 1\rangle}} \\ &= \frac{1}{\sqrt{C+1}} \qquad \text{e} \\ p_2(x) &= \frac{1}{\sqrt{\langle x - \frac{C+3}{2}, x - \frac{C+3}{2}\rangle}} \left(x - \frac{C+3}{2}\right) \\ &= \frac{1}{\sqrt{\int_1^{C+2} \left(x - \frac{C+3}{2}\right)^2 dx}} \left(x - \frac{C+3}{2}\right) \\ &= \frac{1}{\sqrt{\frac{u^3}{3}\Big|_{-\frac{C+1}{2}}^{\frac{C+1}{2}}}} \left(x - \frac{C+3}{2}\right) \qquad \left(\text{onde } u = x - \frac{C+3}{2}\right) \\ &= \frac{1}{\sqrt{\frac{(C+1)^3}{12}}} \left(x - \frac{C+3}{2}\right). \end{aligned}$$

Portanto, ao substituirmos f, p_1 e p_2 na expressão (7.23), obtemos

$$\begin{aligned} p(x) = (D+1)\Bigg(& \langle x^{1/(E+1)}, 1/\sqrt{C+1}\rangle \frac{1}{\sqrt{C+1}} \\ & + \quad \langle x^{1/(E+1)}, (1/\sqrt{(C+1)^3/12})(x - (C+3)/2)\rangle \frac{1}{\sqrt{\frac{(C+1)^3}{12}}} \left(x - \frac{C+3}{2}\right)\Bigg), \end{aligned}$$

isto é,

$$\begin{aligned} p(x) = (D+1)\Bigg(& \frac{1}{C+1}\int_1^{C+2} x^{1/(E+1)} dx \\ & + \quad \frac{12}{(C+1)^3}\int_1^{C+2} x^{1/(E+1)}(x - (C+3)/2)\,dx \left(x - \frac{C+3}{2}\right)\Bigg). \end{aligned}$$

Assim, para $C = 3$, $D = 4$ e $E = 5$, por exemplo, temos

$$\begin{aligned} p(x) &= 5\left(\frac{1}{4}\int_1^5 x^{1/6} dx + \frac{12}{4^3}\left(\int_1^5 x^{1/6}(x-3)\,dx\right)(x-3)\right) \\ &= 5\left(\frac{1}{4}\left[\frac{x^{7/6}}{7/6}\right]_1^5 + \frac{3}{16}\left[\frac{x^{13/6}}{13/6} - \frac{3x^{7/6}}{7/6}\right]_1^5 (x-3)\right) \\ &= 5\left(\frac{3}{14}\left(5^{7/6} - 1\right) + \frac{3}{16}\left(\frac{6}{13}\left(5^{13/6} - 1\right) - \frac{18}{7}\left(5^{7/6} - 1\right)\right)(x-3)\right). \end{aligned}$$

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 $F = 6$,[59] 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 $\mathcal{U}$.[60]

(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 $c = 1$.[61,62]

---

[59] Confira a primeira nota de rodapé do exercício 2 supracitado.

[60] Para $D = 4$, $E = 5$ e $F = 6$,

$$\mathbf{x}' = \frac{1}{\sqrt{5^2 + 6^2 + 7^2}}(5, 6, 7). \tag{7.24}$$

[61] Como $\mathbf{y}_1$ e $\mathbf{y}_2$ não são múltiplos um do outro, eles são LI.

[62] Para $D = 4$, $E = 5$ e $F = 6$,

$$\mathbf{y}_1 = \left(-\frac{6}{5}, 1, 0\right) \text{ e } \mathbf{y}_2 = \left(-\frac{7}{5}, 0, 1\right).$$

Agora, se

$$\begin{aligned}
\mathbf{y}_1' &= \mathbf{y}_1 \\
&= \left(-\frac{E+1}{D+1}, 1, 0\right), \\
\mathbf{y}_2' &= \mathbf{y}_2 - \frac{\langle \mathbf{y}_2, \mathbf{y}_1' \rangle}{\|\mathbf{y}_1'\|^2}\mathbf{y}_1' \\
&= \left(-\frac{F+1}{D+1}, 0, 1\right) - \frac{\frac{(E+1)(F+1)}{(D+1)^2}}{\left(\frac{E+1}{D+1}\right)^2 + 1}\left(-\frac{E+1}{D+1}, 1, 0\right) \\
&= \left(-\frac{F+1}{D+1}, 0, 1\right) + \left(\frac{(E+1)^2(F+1)}{\left((E+1)^2 + (D+1)^2\right)(D+1)}, -\frac{(E+1)(F+1)}{(E+1)^2 + (D+1)^2}, 0\right) \\
&= \left(-\frac{(D+1)(F+1)}{(E+1)^2 + (D+1)^2}, -\frac{(E+1)(F+1)}{(E+1)^2 + (D+1)^2}, 1\right), \\
\mathbf{y}_1'' &= \frac{\mathbf{y}_1'}{\|\mathbf{y}_1'\|} \\
&= \frac{1}{\sqrt{\left(\frac{E+1}{D+1}\right)^2 + 1}}\left(-\frac{E+1}{D+1}, 1, 0\right) \text{ e} \\
\mathbf{y}_2'' &= \frac{\mathbf{y}_2'}{\|\mathbf{y}_2'\|} \\
&= \frac{1}{\sqrt{\left(\frac{(D+1)(F+1)}{(E+1)^2+(D+1)^2}\right)^2 + \left(\frac{(E+1)(F+1)}{(E+1)^2+(D+1)^2}\right)^2 + 1}}\left(-\frac{(D+1)(F+1)}{(E+1)^2 + (D+1)^2}, -\frac{(E+1)(F+1)}{(E+1)^2 + (D+1)^2}, 1\right),
\end{aligned}$$

$\{\mathbf{y}_1'', \mathbf{y}_2''\}$ é ortonormal, por Gram-Schmidt.[63]

(c) Sejam $D = 4$, $E = 5$, $F = 6$ e $\mathbf{x} = (x, y, z)$.

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.[64] Assim,

$$\begin{aligned}
P_{\mathcal{U}^\perp}(\mathbf{x}) &= \mathbf{x} - P_{\mathcal{U}}(\mathbf{x}) \\
&= (x, y, z) - \frac{1}{110}(25x + 30y + 35z, 30x + 36y + 42z, 35x + 42y + 49z) \\
&= \left(x - \frac{1}{110}(25x + 30y + 35z), y - \frac{1}{110}(30x + 36y + 42z), z - \frac{1}{110}(35x + 42y + 49z)\right) \\
&= \left(\frac{85}{110}x - \frac{30}{110}y - \frac{35}{110}z, -\frac{30}{110}x + \frac{74}{110}y - \frac{42}{110}z, -\frac{35}{110}x - \frac{42}{110}y + \frac{61}{110}z\right).
\end{aligned}$$

[63] Para $D = 4$, $E = 5$ e $F = 6$,

$$\mathbf{y}_1'' = \frac{1}{\sqrt{(6/5)^2 + 1}}\left(-\frac{6}{5}, 1, 0\right) \text{ e } \mathbf{y}_2'' = \frac{1}{\sqrt{(35/61)^2 + (42/61)^2 + 1}}\left(-\frac{35}{61}, -\frac{42}{61}, 1\right). \tag{7.25}$$

[64] Confira o item (e) do exercício 4 da seção 7.5.

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. Em $\mathbb{R}^4$, munido do produto interno canônico, se $\mathbf{x}_0 = (1/2, 1/2, 1/2, 1/2)$, $\mathcal{U} = [\mathbf{x}_0]$ e $\mathbf{u} = (C+1, D+1, E+1, F+1)$, onde C, D, E e F são números reais diferentes de -1, determine $P_{\mathcal{U}^\perp}(\mathbf{u})$.

RESOLUÇÃO

Como, para cada $\mathbf{y} \in \mathcal{U}^\perp$, $\langle \mathbf{y}, \mathbf{x}_0 \rangle = 0$, podemos obter uma base ortonormal $\{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3\}$ de $\mathcal{U}^\perp$, por Gram-Schmidt, para calcularmos

$$P_{\mathcal{U}^\perp}(\mathbf{u}) = \langle \mathbf{u}, \mathbf{y}_1 \rangle \mathbf{y}_1 + \langle \mathbf{u}, \mathbf{y}_2 \rangle \mathbf{y}_2 + \langle \mathbf{u}, \mathbf{y}_3 \rangle \mathbf{y}_3.$$

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.[65] 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 $\mathbf{x}_0$ é unitário,

$$\begin{aligned} P_{\mathcal{U}}(\mathbf{u}) &= \langle \mathbf{u}, \mathbf{x}_0 \rangle \mathbf{x}_0 \\ &= \left(\frac{C+1}{2} + \frac{D+1}{2} + \frac{E+1}{2} + \frac{F+1}{2}\right)(1/2, 1/2, 1/2, 1/2) \\ &= \frac{C+D+E+F+4}{2}(1/2, 1/2, 1/2, 1/2) \end{aligned}$$

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 $\mathbb{K}$.

[65] Confira o item (e) do exercício 4 da seção 7.5.

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,[66] 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 de $\langle\ ,\ \rangle$ no primeiro fator.[67]

- 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).[68]

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

$$\varphi = \langle -, u\rangle.$$

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).[69] 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

$$u := \sum_{i=1}^{r} \overline{\varphi(v_i)} v_i. \tag{7.27}$$

[66] Confira o exercício 4 de 6.4.

[67] Idem.

[68] Aqui, "$\cdot$" representa o produto interno canônico em $\mathbb{R}^n$.

[69] Cf. p. 192.

A unicidade de u é decorrente do lema (R35), que veremos a seguir.

OBSERVAÇÕES

- Pela unicidade supracitada, φ não depende da base ortonormal $\mathcal{B}$ considerada;
- Pela demonstração supracitada, para representarmos um funcional linear φ arbitrário por (7.26), basta determinarmos (7.27).

(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:

- $\mathcal{V}$ finitamente gerado e dotado do produto interno $\langle\ ,\ \rangle$;[70]
- $T \in \mathcal{L}(\mathcal{V})$.

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,

$$\varphi = \langle -, w\rangle \circ T \in \mathcal{L}(\mathcal{V}, \mathbb{K}). \tag{7.28}$$

Por outro lado, por 7.3.1, existe um único $u \in \mathcal{V}$ tal que

$$\varphi = \langle -, u\rangle. \tag{7.29}$$

Portanto, se

$$u := T^*(w),$$

[70]Como no exercício 4 da seção 6.4.

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 (7.28) e (7.29).

DEFINIÇÃO T^* é dito *adjunto* de T.

EXEMPLO
Determinaremos T^* para

$$\mathbb{R}^2 \ni (x_1, x_2) \xmapsto{T} (x_1 - x_2, x_1 + x_2) \in \mathbb{R}^2,$$

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,

$$\mathbb{R}^2 \ni (y_1, y_2) \xmapsto{T^*} (y_1 + y_2, y_2 - y_1) \in \mathbb{R}^2.$$ [71]

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

$$\begin{aligned}
\langle v, T^*(w_1 + w_2) \rangle &= \langle T(v), w_1 + w_2 \rangle \\
&= \langle T(v), w_1 \rangle + \langle T(v), w_2 \rangle \\
&= \langle v, T^*(w_1) \rangle + \langle v, T^*(w_2) \rangle \\
&= \langle v, T^*(w_1) + T^*(w_2) \rangle,
\end{aligned}$$

$$T^*(w_1 + w_2) = T^*(w_1) + T^*(w_2).$$ [72]

Para $\lambda \in \mathbb{K}$ e $w \in \mathcal{V}$ fixos e $v \in \mathcal{V}$ arbitrário, como

$$\begin{aligned}
\langle v, T^*(\lambda w) \rangle &= \langle T(v), \lambda w \rangle \\
&= \overline{\lambda} \langle T(v), w \rangle \\
&= \overline{\lambda} \langle v, T^*(w) \rangle \\
&= \langle v, \lambda T^*(w) \rangle,
\end{aligned}$$

$$T^*(\lambda w) = \lambda T^*(w).$$ [73]

[71] Cf. (R35), p. 208.
[72] Idem.
[73] Idem.

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 $\mathcal{B} = \{v_1, \dots, v_n\}$ 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^*.*

Demonstraremos as propriedades 3 e 6.[74]

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).[75]

DEMONSTRAÇÃO DE 7.4.2.6

Como vimos na subseção 6.3.3, para $j \in \{1, \dots, n\}$ 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

$$\begin{bmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{bmatrix},$$

[74] As demonstrações das outras propriedades estão propostas como exercícios da seção 7.5.

[75] Cf. p. 208.

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).[76] 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),

  $$\mathbb{R}^n \ni \mathbf{x} \overset{T}{\longmapsto} A\mathbf{x} \in \mathbb{R}^n$$

  é autoadjunto;

- Para $A \in \mathbb{C}^{n \times n}$ hermiteana (como nos exercícios 12.3, 14 e 15 da seção 5.2),

  $$\mathbb{C}^n \ni \mathbf{x} \overset{T}{\longmapsto} A\mathbf{x} \in \mathbb{C}^n$$

  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 $\mathbf{x}, \mathbf{y} \in \mathbb{K}^n$,

$$\begin{aligned} \langle T(\mathbf{x}), \mathbf{y} \rangle &= \langle \mathbf{x}, T^*(\mathbf{y}) \rangle \text{ (pela definição de } T^*) \\ &= \langle \mathbf{x}, A^*\mathbf{y} \rangle \quad \text{(via 7.4.3.6)} \\ &= \langle \mathbf{x}, A\mathbf{y} \rangle \quad \text{(pois } A^* = A) \\ &= \langle \mathbf{x}, T(\mathbf{y}) \rangle. \end{aligned}$$

OBSERVAÇÃO

Os autovalores dos operadores dos dois exemplos supracitados são reais. Na verdade, vale o seguinte resultado mais geral:

[76] Cf. p. 192.

(R36) Autovalor real

$$\boxed{\lambda \text{ é um autovalor do operador hermiteano } T \Longrightarrow \lambda \in \mathbb{R}.}$$

DEMONSTRAÇÃO

Seja v um autovetor de T associado ao autovalor λ.[77] Assim, como

$$\begin{aligned}\lambda\|v\|^2 &= \lambda\langle v, v\rangle \\ &= \langle \lambda v, v\rangle \\ &= \langle T(v), v\rangle \\ &= \langle v, T(v)\rangle \\ &= \langle v, \lambda v\rangle \\ &= \overline{\lambda}\langle v, v\rangle \\ &= \overline{\lambda}\|v\|^2,\end{aligned}$$

$\lambda = \overline{\lambda}$ 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

$$\boxed{T \text{ é autoadjunto e, para todo } v \in \mathcal{V},\ \langle T(v), v\rangle = 0 \Longrightarrow T = 0.}$$

DEMONSTRAÇÃO

Sejam $u, w \in \mathcal{V}$ arbitrários. Portanto, como $\langle T(v), v\rangle = 0$ (para cada $v \in \mathcal{V}$),

$$\langle T(u+w), u+w\rangle - \langle T(u), u\rangle - \langle T(w), w\rangle = 0.$$

Assim, como

$$\langle T(u), w\rangle + \langle T(w), u\rangle = 0, \tag{7.30}$$

$$\begin{aligned}0 &= \langle T(u), w\rangle + \langle w, T(u)\rangle \ (T \text{ é hermiteano}) \\ &= \langle T(u), w\rangle + \overline{\langle T(u), w\rangle} \\ &= \text{ duas vezes a parte real de } \langle T(u), w\rangle.\end{aligned}$$

Portanto,

$$\text{a parte real de } \langle T(u), w\rangle \text{ é nula.} \tag{7.31}$$

Agora, ao substituirmos w por iw em (7.30), temos

$$\begin{aligned}0 &= \langle T(u), iw\rangle + \langle iT(w), u\rangle \\ &= \overline{i}\langle T(u), w\rangle + i\langle T(w), u\rangle,\end{aligned}$$

que, ao multiplicarmos cada um de seus membros por i, pode ser reescrita como

$$\begin{aligned}0 &= \langle T(u), w\rangle - \langle T(w), u\rangle \\ &= \langle T(u), w\rangle - \langle w, T(u)\rangle \ (T \text{ é hermiteano}) \\ &= \langle T(u), w\rangle - \overline{\langle T(u), w\rangle} \\ &= \text{ duas vezes a parte imaginária de } \langle T(u), w\rangle.\end{aligned}$$

[77] Lembre-se que autovetores são não nulos!

Assim,

$$\text{a parte imaginária de } \langle T(u), w\rangle \text{ é nula.} \tag{7.32}$$

Portanto, como

$$\langle T(u), w\rangle = 0,$$

por (7.31) e (7.32), e w é arbitrário,[78]

$$\langle T(u), T(u)\rangle = 0,$$

isto é,

$$T(u) = 0.$$

Consequentemente, como u é arbitrário,

$$T = 0.$$

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.[79]

(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)

$$\begin{aligned} \|T(v)\|^2 = \|T^*(v)\|^2 \ \forall v \in \mathcal{V} &\Longleftrightarrow \langle T(v), T(v)\rangle = \big\langle T^*(v), T^*(v)\big\rangle \,\forall v \in \mathcal{V} \\ &\Longleftrightarrow \big\langle\, (T^*T)(v), v\big\rangle = \big\langle\, (TT^*)(v), v\big\rangle \,\forall v \in \mathcal{V} \\ &\Longleftrightarrow \big\langle\, (T^*T - TT^*)(v), v\big\rangle = 0 \,\forall v \in \mathcal{V} \\ &\Longleftrightarrow T^*T - TT^* = 0, \end{aligned}$$

[78] Assim, podemos considerar $w = T(u)$.
[79] Considere o exercício 12 da seção 5.2.

pois, como $T^*T - TT^*$ é autoadjunto,[80] podemos usar (R37) na implicação "$\Longrightarrow$" da última equivalência, de cima para baixo.

DEMONSTRAÇÃO DE (R38.2)

É uma consequência direta de (R38.1).[81]

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^*,[82]

$$\mathrm{Nu}(T-\lambda I) = \mathrm{Nu}\left(T^*-\overline{\lambda}I\right), \tag{7.33}$$

por (R38.4), (R38.3) e 7.4.3.[83]

DEMONSTRAÇÃO DE (R38.6)

Como

$$\begin{aligned}(\lambda_1-\lambda_2)\left\langle v_1,v_2\right\rangle &= \lambda_1\langle v_1,v_2\rangle - \lambda_2\langle v_1,v_2\rangle\\ &= \left\langle \lambda_1 v_1, v_2\right\rangle - \left\langle v_1, \overline{\lambda_2} v_2\right\rangle\\ &= \left\langle T(v_1), v_2\right\rangle - \left\langle v_1, T^*(v_2)\right\rangle\\ &= \left\langle T(v_1), v_2\right\rangle - \left\langle T(v_1), v_2\right\rangle\\ &= 0,\end{aligned}$$

onde usamos (7.33) na terceira igualdade, e $\lambda_1 - \lambda_2 \neq 0$, temos

$$\left\langle v_1, v_2\right\rangle = 0.$$

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:

[80] Verifique!

[81] Confira a propriedade 3 de 7.4.3.

[82] Confira a propriedade 8 de 7.4.3.

[83] Cf. 7.2.1, p. 182.

Para $\mathbb{K} = \mathbb{C}$,

$$\text{existe alguma base ortonormal } \mathcal{B} \text{ de } \mathcal{V} \text{ com } \mathcal{M}(T,\mathcal{B}) \text{ diagonal}$$

$$\Longleftrightarrow$$

$$T \text{ é normal.}$$

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,[84] T é normal.

DEMONSTRAÇÃO DE 7.4.6

Para a implicação "$\Longrightarrow$", seja $\mathsf{D} = \mathcal{M}(T,\mathcal{B})$, isto é, $\mathsf{D}^* = \mathcal{M}(T^*,\mathcal{B})$,[85] que, obviamente, também é diagonal. Então, como D e D^* comutam, T e T^* também comutam.[86]

Para a implicação "$\Longleftarrow$", note que, pelo lema de Schur,[87] existe uma base ortonormal $\mathcal{B} = \{v_1, \ldots, v_n\}$ de $\mathcal{V}$ tal que

$$\mathcal{M}(T,\mathcal{B}) = \begin{bmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{bmatrix} \tag{7.34}$$

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,

$$\mathcal{M}(T^*,\mathcal{B}) = \begin{bmatrix} \overline{a_{11}} & & 0 \\ & \ddots & \\ \overline{*} & & \overline{a_{nn}} \end{bmatrix} \tag{7.35}$$

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,

$$T(v_1) = a_{11} v_1$$

---

[84] Confira o segundo item dos exemplos da subseção 7.4.4.

[85] Considere a propriedade 6 de 7.4.3.

[86] De fato, considere $\mathcal{V} = \mathcal{W}$ e $\mathcal{B} = \mathcal{B}' = \mathcal{B}''$ na subseção 6.3.3. Assim, o isomorfismo (6.12) satisfaz a condição (6.13) 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, $TT^* = T^*T$.

[87] Considere a observação que precede (R24), p. 188.

e
$$T^*(v_1) = \overline{a_{11}}v_1 + \overline{a_{12}}v_2 + \cdots + \overline{a_{1n}}v_n,$$
por (7.34) e (7.35), 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 (7.34) e (7.35), 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, (7.34) é diagonal.

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

$$\Longleftrightarrow$$

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:*

$$p(t) = c\,(t-\lambda_1)\cdots(t-\lambda_m)\left(t^2+b_1t+c_1\right)\cdots\left(t^2+b_Mt+c_M\right),$$

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 AXLER, referenciado no capítulo 1, que também foi a "fonte inspiradora" das demonstrações de alguns dos próximos resultados.

(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.[88]

Portanto, como $v \neq 0_{\mathcal{V}}$,

$$\|T(v)\|^2 - |b|\|T(v)\|\|v\| + c\|v\|^2 = \left(\|T(v)\| - \frac{|b|\|v\|}{2}\right)^2 + \left(c - \frac{b^2}{4}\right)\|v\|^2 > 0,$$ [89]

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, $\mathrm{Nu}\left(T^2 + bT + cI\right) = \{0_{\mathcal{V}}\}$, ou seja, $T^2 + bT + cI$ é invertível.[90]

(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

$$v, T(v), T^2(v), \ldots, T^{\dim \mathcal{V}}(v)$$

[88] Confira o exercício 4.1, seção 6.4.

[89] Completando quadrados, obtemos o segundo membro da igualdade, que é positivo, pois $b^2 < 4c$.

[90] Cf. (R12) e (R18), pp. 158 e 179, respectivamente.

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

$$\sum_{j=0}^{\dim \mathcal{V}} a_j T^j(v) = 0_{\mathcal{V}}. \tag{7.36}$$

Por outro lado, se

$$p(t) = \sum_{j=0}^{\dim \mathcal{V}} a_j t^j \in \mathcal{P}(\mathbb{R}), \tag{7.37}$$

então, pelo fato algébrico (**) supracitado,

$$p(t) = c\left(t^2 + b_1 t + c_1\right) \cdots \left(t^2 + b_M t + c_M\right)(t-\lambda_1)\cdots(t-\lambda_m), \tag{7.38}$$

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

$$\left(c\left(T^2 + b_1 T + c_1 I\right) \cdots \left(T^2 + b_M T + c_M I\right)(T-\lambda_1 I)\cdots(T-\lambda_m I)\right)(v) = 0_{\mathcal{V}}, \tag{7.39}$$

por 7.2.2, (7.36), (7.37) e (7.38). Portanto:

- Caso $M = 0$, temos $m > 0$ e (7.39) reduzida a equação

$$((T-\lambda_1 I)\cdots(T-\lambda_m I))(v) = 0_{\mathcal{V}}; \tag{7.40}$$

- Caso $M > 0$, como $T^2 + b_i T + c_i I$ é invertível, $i = 1, \ldots, M$,[91] temos

$$m > 0,$$

  pois, como $c \neq 0$, se $m = 0$, o operador invertível

$$c\left(T^2 + b_1 T + c_1 I\right) \cdots \left(T^2 + b_M T + c_M I\right) \tag{7.41}$$

  leva $v \neq 0_{\mathcal{V}}$ em $0_{\mathcal{V}}$, por (7.39), que é um absurdo.[92] Portanto, para que (7.39) seja válida, (7.40) deve ser verdadeira.[93]

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, (7.40) é válida apenas para $v = 0_{\mathcal{V}}$,[94] 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.[95]

DEMONSTRAÇÃO DE 7.4.7

Para a implicação "$\Longrightarrow$", seja $\mathsf{D} = \mathcal{M}(T, \mathcal{B})$, isto é, $\mathsf{D}^* = \mathcal{M}(T^*, \mathcal{B})$,[96] que, obviamente, também é diagonal. Logo, como

$$\begin{aligned} \mathsf{D}^* &= \overline{\mathsf{D}}^t \text{ } (\mathsf{D}^* \text{ é a transposta conjugada de } \mathsf{D}) \\ &= \mathsf{D}^t \text{ } (\mathsf{D} \text{ tem apenas entradas reais}) \\ &= \mathsf{D} \text{ } (\text{a transposta de uma matriz diagonal é a própria matriz}), \end{aligned}$$

[91] Cf. (R39).
[92] Cf. (R18) e (R12), pp. 179 e 158, respectivamente.
[93] De fato, a única pré-imagem de $0_{\mathcal{V}}$ pelo operador invertível (7.41) é o próprio $0_{\mathcal{V}}$.
[94] Cf. (R18) e (R12), pp. 179 e 158, respectivamente.
[95] Cf. (R19), p. 182.
[96] Confira a propriedade 6 de 7.4.3.

$T^* = T$, 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.[97]
Demonstraremos a implicação "$\Longleftarrow$" por indução sobre $n = \dim \mathcal{V}$. Na verdade, basta demonstrarmos que

$$T \text{ é autoadjunto} \Longrightarrow \mathcal{V} \text{ possui alguma base ortonormal } \mathcal{B} \text{ consistindo de autovetores de } T.^{98} \tag{7.42}$$

Obviamente, (7.42) é verdadeira para $n = 1$. Suponha, então, $n > 1$ e (7.42) verdadeira para qualquer espaço com produto interno real cuja dimensão pertença ao conjunto $\{1, 2 \ldots, n-1\}$.[99] Agora, por (R40), T admite algum autovetor u que, sem perda de generalidade, pode ser considerado unitário.[100] Portanto, $\mathcal{U} = [u]$ é invariante por T e

$$T|_{\mathcal{U}^\perp} \in \mathcal{L}\Big(\mathcal{U}^\perp\Big)$$

é autoadjunto, pelo exercício 8 da seção 7.5. Logo, podemos aplicar (7.42) ao operador $T|_{\mathcal{U}^\perp}$, pela hipótese de indução supracitada. Portanto, para concluirmos a demonstração, basta considerarmos o resultado (R32),[101] 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 $T|_{\mathcal{U}^\perp}$.

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

[97] Confira a subseção 6.3.3.
[98] Cf. (R26), p. 190.
[99] Essa é a nossa *hipótese de indução*.
[100] Caso u não seja unitário, considere o versor de u.
[101] Cf. p. 194.

· 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.[102]

· 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.[103]
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,

$$\mathcal{U}_i(\lambda) = \mathcal{U}_i.^{104}$$

Portanto,

$$\mathcal{U}_1 \subset \mathcal{U}_2 \subset \ldots \subset \mathcal{U}_i \subset \ldots^{105} \tag{7.43}$$

Logo, como $\mathcal{V}$ é finitamente gerado, a sequência (7.43) 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

$$\mathcal{U}_\ell = \mathcal{U}_{\ell+1}.^{106} \tag{7.44}$$

Portanto,

$$\mathcal{U}_{\ell+1} = \mathcal{U}_{\ell+2} = \mathcal{U}_{\ell+3} = \ldots^{107} \tag{7.45}$$

[102] Confira a seção 7.6.
[103] $r = 1$ é uma possibilidade!
[104] Note que,

$$v \in \mathcal{U}_i \iff (T - \lambda I)^i\,(v) = 0.$$

[105] Confira o antepenúltimo exercício da seção 7.5.
[106] Observe que ℓ depende do autovalor λ considerado.
[107] Confira o antepenúltimo exercício da seção 7.5.

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 (7.43) e (7.44), 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 (7.43), (7.44), (7.45) 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,

$$v = (T-\lambda I)^{\dim \mathcal{V}}(u) = 0_{\mathcal{V}}.$$

PARTE 3
Pela parte 2 supracitada e por (R14),[108]

$$\mathcal{V} = \mathrm{Nu}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big) \bigoplus \mathrm{Im}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big).$$

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.[109] 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,

$$v \in \mathrm{Nu}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big)$$

$$\Downarrow$$

$$\begin{aligned} (T-\lambda I)^{\dim \mathcal{V}} T(v) &= T(T-\lambda I)^{\dim \mathcal{V}}(v) \\ &= T(0) \\ &= 0. \end{aligned}$$

$$\Downarrow$$

[108] Cf. p. 159.

[109] Logo, a restrição de T a cada um desses subespaços é um operador.

$$T(v) \in \operatorname{Nu}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big).$$

Por outro,

$$\begin{gathered} v \in \operatorname{Im}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big) \\ \Downarrow \\ \text{existe } u \in \mathcal{V} \text{ tal que } (T-\lambda I)^{\dim \mathcal{V}}(u) = v \\ \Downarrow \\ \begin{aligned} (T-\lambda I)^{\dim \mathcal{V}} T(u) &= T(T-\lambda I)^{\dim \mathcal{V}}(u) \\ &= T(v) \end{aligned} \\ \Downarrow \\ T(v) \in \operatorname{Im}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big). \end{gathered}$$

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 $\operatorname{Im}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big)$, que é um operador sobre esse subespaço (pela parte 4 supracitada), são $\lambda_2, \ldots, \lambda_r$,[110]

$$\operatorname{Im}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big) = \bigoplus_{j=2}^{r} \operatorname{Nu}\Big((T-\lambda_j I)^{\dim \mathcal{V}}\Big), \tag{7.46}$$

pela hipótese de indução. Logo, para concluirmos a demonstração, basta substituirmos (7.46) 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

$$(z_1, z_2, z_3) \overset{T}{\longmapsto} (z_2, 0, z_3).$$

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.

[110] De fato, pela parte 2 supracitada, nenhum dos autovetores generalizados de $T|_{\operatorname{Im}\left((T-\lambda I)^{\dim \mathcal{V}}\right)}$ pode estar associado a λ, pois os autovetores generalizados de T associados a λ pertencem a $\operatorname{Nu}\Big((T-\lambda I)^{\dim \mathcal{V}}\Big)$.

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. Pela resolução do 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

$$\left\{L^{p_1-1}(v_1), \ldots, L(v_1), v_1, \ldots, L^{p_s-1}(v_s), \ldots, L(v_s), v_s\right\}$$

é uma base de $\mathcal{V}$.

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

$$\begin{aligned} \mathcal{U}_i &= \mathcal{U}_\ell \\ &= \mathcal{U}_{\dim \mathcal{V}} \end{aligned} \tag{7.47}$$

é invariante por T,[111] $\mathcal{U}_i$ é invariante por $T - \lambda I$,[112] e sua nilpotência é trivial, pois, para cada $v \in \mathcal{U}_i$,

$$\begin{aligned} \left((T - \lambda I)|_{\mathcal{U}_i}\right)^\ell (v) &= (T - \lambda I)^\ell (v) \\ &= 0_{\mathcal{V}}. \end{aligned}$$

DEMONSTRAÇÃO DO LEMA 2

Note que

$$\text{o lema 2 é válido para } L = 0. \tag{7.48}$$

De fato, basta considerarmos uma base $\{v_1, \ldots, v_s\}$ arbitrária de $\mathcal{V}$ e $p_1 = \cdots = p_s = 1$.[113] Por outro lado, para $L \neq 0$, podemos aplicar indução sobre $\dim \mathcal{V}$. Assim, se $\dim \mathcal{V} = 1$, então

$$L^p = 0 \iff L = 0,^{114} \tag{7.49}$$

ou seja, o lema 2 é válido para $\mathcal{V}$ unidimensional, por (7.48). Agora, se $\dim \mathcal{V} > 1$, suponha que

$$\text{o lema 2 é válido para qualquer espaço que tenha dimensão inferior a } \dim \mathcal{V} \tag{7.50}$$

e considere

$$\mathcal{U} = \mathrm{Im}(L). \tag{7.51}$$

· Caso $\mathcal{U}$ seja nulo, $L = 0$ e, para tal operador, o lema 2 é válido.[115]

[111] Cf. (7.44), (7.45), a informação adicional e a parte 4 da demonstração do lema 1.
[112] Qualquer subespaço $\mathcal{U}$ de $\mathcal{V}$ é invariante por λI.
[113] Nesse caso, $s = \dim \mathcal{V}$.
[114] Confira a resolução do penúltimo exercício da seção 7.5.
[115] Cf. (7.48).

- Caso $\mathcal{U} = \mathcal{V}$, L é injetivo.[116] 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 (7.50) 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

$$L^{q_j}\left(u_j\right) = 0_{\mathcal{V}},\ j = 1, \ldots, r, \tag{7.52}$$

e

$$\left\{L^{q_1-1}(u_1), \ldots, L(u_1), u_1, \ldots, L^{q_r-1}(u_r), \ldots, L(u_r), u_r\right\} \tag{7.53}$$

é uma base de $\mathcal{U}$. Portanto, como existe $v_j \in \mathcal{V}$, $j = 1, \ldots, r$, tal que $L\left(v_j\right) = u_j$, por (7.51), e

$$L^i\left(v_j\right) = L^{i-1}\left(u_j\right), \tag{7.54}$$

para todo inteiro positivo i, os vetores da lista $L^{q_1}(v_1), \ldots, L^{q_r}(v_r)$ são LI em $\mathrm{Nu}(L)$, por (7.52) e (7.53). Logo, podemos estender a lista supracitada, caso seja necessário, a uma base

$$\left\{L^{q_1}(v_1), \ldots, L^{q_r}(v_r), w_1, \ldots, w_q\right\} \tag{7.55}$$

de $\mathrm{Nu}(L)$.[117] Assim,

$$\left\{L^{q_1}(v_1), \ldots, L(v_1), v_1, \ldots, L^{q_r}(v_r), \ldots, L(v_r), v_r, w_1, \ldots, w_q\right\} \tag{7.56}$$

é uma base de $\mathcal{V}$.[118]

DEMONSTRAÇÃO DO TEOREMA 7.4.8

Considere λ_i como no lema 1,

$$\mathcal{U}_{\lambda_i} = \mathrm{Nu}\left((T - \lambda_i I)^{\dim \mathcal{V}}\right) \tag{7.57}$$

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

$$T|_{\mathcal{U}_{\lambda_i}} \in \mathcal{L}(\mathcal{U}_{\lambda_i}),\ i = 1, \ldots, r,^{119}$$

temos

$$\mathcal{M}(T, \mathcal{B}) = \begin{pmatrix} \mathcal{M}\left(T|_{\mathcal{U}_{\lambda_1}}, \mathcal{B}_1\right) & & \text{zeros} \\ & \ddots & \\ \text{zeros} & & \mathcal{M}\left(T|_{\mathcal{U}_{\lambda_r}}, \mathcal{B}_r\right) \end{pmatrix}.$$

Portanto, para $i = 1, \ldots, r$, como

$$L_i := (T - \lambda_i I)\,|_{\mathcal{U}_{\lambda_i}}$$

é nilpotente,[120] caso $\mathcal{B}_i$ seja uma base como a do lema 2,

$$\begin{aligned} \mathcal{M}\left(T|_{\mathcal{U}_{\lambda_i}}, \mathcal{B}_i\right) &= \mathcal{M}\left(\lambda_i I|_{\mathcal{U}_{\lambda_i}}, \mathcal{B}_i\right) + \mathcal{M}(L_i, \mathcal{B}_i) \\ &= \begin{pmatrix} J_{i,1} & & \text{zeros} \\ & \ddots & \\ \text{zeros} & & J_{i,s_i} \end{pmatrix}, \end{aligned}$$

[116] Cf. (R12) e (R14), pp. 158 e 159.

[117] Cf. (R6), p. 153.

[118] Confira a resolução do último exercício da seção 7.5.

[119] Confira a parte 4 da demonstração do lema 1.

[120] Confira o exemplo de nilpotência que precede a demonstração do lema 2.

onde, para $j = 1, \ldots, s_i$, as colunas do bloco

$$J_{i,j} = \begin{pmatrix} \lambda_i & 1 & & & \text{zeros} \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ \text{zeros} & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix}$$

são obtidas ao aplicarmos

$$T|_{\mathcal{U}_{\lambda_i}} = L_i + \lambda_i I|_{\mathcal{U}_{\lambda_i}}$$

à cada vetor da base

$$\mathcal{B}_{i,j} := \left\{ L_i^{p_j - 1}(v_j), \ldots, L_i(v_j), v_j \right\}, \tag{7.58}$$

com

$$\mathcal{B}_i = \bigcup_{j=1}^{s_i} \mathcal{B}_{i,j}. \tag{7.59}$$

OBSERVAÇÃO

O primeiro vetor da base (7.58) é um autovetor de T associado ao autovalor λ_i.[121]

DEFINIÇÃO

A lista de vetores de (7.58) é um *ciclo* da base $\mathcal{B}_i$, $i = 1, \ldots, r$, de *comprimento* p_j.

EXEMPLO.1[122]

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

$m_i := \dim(\mathcal{U}_{\lambda_i})$ é a *multiplicidade algébrica* de λ_i, $i = 1, \ldots, r$.[123]

OBSERVAÇÕES

[121] Confira a primeira coluna de $J_{i,j}$.

[122] Para mais exemplos, confira as subseções 7.6.2 e 7.6.3.

[123] Que fique claro:

a multiplicidade algébrica de λ_i é a dimensão do autoespaço generalizado de T associado à λ_i, $i = 1, \ldots, r$.

· 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 (7.59).

· Pelo lema 1,

$$\dim \mathcal{V} = \sum_{i=1}^{r} m_i. \tag{7.60}$$

DEFINIÇÃO

$\mu_i := \dim(\mathrm{Nu}\,(T - \lambda_i I))$ é a *multiplicidade geométrica* de λ_i, $i = 1, \ldots, r$.[124]

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 $\mu = 1$.[125]

DEFINIÇÃO

Para um espaço vetorial $\mathcal{V}$ sobre $\mathbb{C}$, caso m_i seja a multiplicidade algébrica do autovalor λ_i de $T \in \mathcal{L}(\mathcal{V})$, $i = 1, \ldots, r$, definimos o *polinômio característico* de T por

$$p(t) := \prod_{i=1}^{r} (t - \lambda_i)^{m_i}\,. \tag{7.61}$$

OBSERVAÇÃO

Por (7.61) e (7.60),

$$\begin{aligned} \text{grau de } p(t) &= \sum_{i=1}^{r} m_i \\ &= \dim \mathcal{V}. \end{aligned}$$

EXEMPLO.3

Em relação ao exemplo.1, $p(t) = (t - \lambda)^3$.[126]

[124]Que fique claro:

$\boxed{\textit{a multiplicidade geométrica de } \lambda_i \textit{ é a dimensão do autoespaço de } T \textit{ associado à } \lambda_i,\ i = 1, \ldots, r.}$

[125]Verifique!

[126]Para mais exemplos de polinômios característicos, confira os capítulos 4 e 5 e a subseção 7.6.2.

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 $\mathcal{U}_{\lambda_i}$,[127]

$$(T - \lambda_i I) \mid_{\mathcal{U}_{\lambda_i}}$$

é nilpotente, pelo exemplo de nilpotência que antecede à demonstração ao lema 2, e, portanto,

$$(T - \lambda_i I)^{m_i} \mid_{\mathcal{U}_{\lambda_i}} = 0, \tag{7.62}$$

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 (7.61). Assim, como

$$p(T)|_{\mathcal{U}_{\lambda_i}} = 0,\text{[128]}$$

$$p(T) = 0,$$

pois $p(T)$ é linear e $\mathcal{V} = \bigoplus_{j=1}^{r} \mathcal{U}_{\lambda_j}$.[129]

OBSERVAÇÃO

Seja $\mathcal{V}_{\mathbb{K}}$ um espaço vetorial sobre o corpo $\mathbb{K}$. No 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 aqui no capítulo 7. 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, dada 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 em (7.61).

[127] Cf. (7.57), p. 225.
[128] Cf. (7.62).
[129] Confira a observação que segue a demonstração do lema 1.

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).[130] 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).[131] 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).[132] 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,[133] para cada $v \in \mathcal{U}^{\perp}$,

$$\begin{aligned} u \in \mathcal{U} &\Longrightarrow \langle v, u \rangle = 0 \\ &\Longrightarrow \langle u, v \rangle = 0 \\ &\Longrightarrow u \in \left(\mathcal{U}^{\perp}\right)^{\perp}, \end{aligned}$$

[130]Cf. p. 190.
[131]Cf. p. 182.
[132]Cf. p. 191.
[133]Cf. (7.16), p. 193.

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,[134] existe um (único) par de vetores $(u, v) \in \mathcal{U} \times \mathcal{U}^{\perp}$ tal que

$$w = u + v,$$

por (R32).[135] 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 $\Longrightarrow P_{\mathcal{U}^{\perp}} = I - P_{\mathcal{U}}$;[136]

   (f) $P_{\mathcal{U}}$ é idempotente, ou seja, $(P_{\mathcal{U}})^2 = P_{\mathcal{U}}$;[137]

   (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)

$$\begin{aligned} w \in \mathcal{U}^{\perp} &\Longrightarrow \mathcal{V} \ni w = 0_{\mathcal{V}} + w, \text{ com } 0_{\mathcal{V}} \in \mathcal{U} \text{ e } w \in \mathcal{U}^{\perp} \\ &\Longrightarrow P_{\mathcal{U}}(w) = 0_{\mathcal{V}}, \text{ pela definição de } P_{\mathcal{U}}. \end{aligned}$$

[134] De fato, pelo exercício 4 da seção 6.4, podemos aplicar Gram-Schmidt a uma base qualquer de $\mathcal{U}$.

[135] Cf. p. 194.

[136] Cf. as matrizes das projeções $P_{\mathcal{S}}$ e $P_{\mathcal{S}^{\perp}}$ da subseção 4.3.1.

[137] 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),[138] e $P_{\mathcal{U}}(v) \in \mathcal{U} = \left(\mathcal{U}^{\perp}\right)^{\perp}$, pelo exercício 3 dessa seção.[139]

RESOLUÇÃO DE 4.(f)

Caso u e v sejam como na demonstração de (R32),[140]

$$\begin{aligned} P_{\mathcal{U}}(P_{\mathcal{U}}(v)) &= P_{\mathcal{U}}(u) \text{ (pois } P_{\mathcal{U}}(v) = u\text{, pela definição de } P_{\mathcal{U}}) \\ &= u \qquad \text{(pelo item (a) desse exercício)} \\ &= P_{\mathcal{U}}(v) \text{ (pois } P_{\mathcal{U}}(v) = u\text{, pela definição de } P_{\mathcal{U}}). \end{aligned}$$

RESOLUÇÃO DE 4.(g)

Sejam u, v e w como na demonstração de (R32). Logo, pelo teorema de Pitágoras,[141]

$$\begin{aligned} \|v\|^2 &= \|u\|^2 + \|w\|^2 \\ &\geq \|u\|^2 = \|P_{\mathcal{U}}(v)\|^2, \end{aligned}$$

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).[142]

---

[138] Cf. p. 196.
[139] Seção 7.5.
[140] Cf. p. 194.
[141] Cf. a demonstração do item 4 de (R34).
[142] Cf. p. 208.

2. Sejam $\lambda \in \mathbb{K}$ e $w \in \mathcal{V}$ fixos. Logo, para $v \in \mathcal{V}$ arbitrário,

$$\begin{aligned}\langle v, (\lambda T)^*(w)\rangle &= \langle (\lambda T)(v), w\rangle \\ &= \lambda\langle T(v), w\rangle \\ &= \lambda\langle v, T^*(w)\rangle \\ &= \langle v, \left(\overline{\lambda}T^*\right)(w)\rangle.\end{aligned}$$

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,[143] 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,[144]

$$\begin{aligned}\lambda \text{ é autovalor de } T &\iff T - \lambda I \text{ não é invertível} \\ &\iff (T-\lambda I)^* = T^* - \overline{\lambda} I \text{ não é invertível} \\ &\iff \overline{\lambda} \text{ é autovalor de } T^*.\end{aligned}$$

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 à L*.[145]
Demonstre que:

(a) $\mathrm{Nu}(L^*) = (\mathrm{Im}(L))^\perp$;

[143] Cf. p. 210.
[144] Cf. pp. 182 e 210.
[145] Conforme o 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. 210, 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 $T \in \mathcal{L}(\mathbb{C}^2)$ 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 (6.12) 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, $T|_{\mathcal{U}^{\perp}} \in \mathcal{L}(\mathcal{U}^{\perp})$;

(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).[146]

RESOLUÇÃO

Para a implicação "$\Longrightarrow$", suponha

$$T^* = T$$

e que λ seja um autovalor de T. Logo,

$$\begin{aligned}\mathrm{Nu}(T - \lambda I) &= \mathrm{Nu}\big(T^* - \overline{\lambda} I\big) \\ &= \mathrm{Nu}\big(T - \overline{\lambda} I\big)\,,\end{aligned} \tag{7.63}$$

pelos itens 3 e 4 de (R38),[147] e, assim, $\overline{\lambda}$ também é autovalor de T. Além disso, como T é normal,[148] T é diagonalizável, por 7.4.6.[149] Portanto, os autovalores λ e $\overline{\lambda}$ não podem ser distintos, por (7.63) e pelo item 2 de (R29).[150] Assim, como $\lambda = \overline{\lambda}$,

$$\lambda \in \mathbb{R}.$$

Para a implicação "$\Longleftarrow$", suponha que

$$\lambda \text{ é autovalor de } T \Longrightarrow \lambda \in \mathbb{R}. \tag{7.64}$$

Logo, por 7.4.6,[151] 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 (7.64), 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.[152] Assim,

$$T^* = T,$$

pelo isomorfismo (6.12).[153]

[146] Cf. p. 212.
[147] Cf. p. 213.
[148] Por hipótese ou por ser autoadjunto!
[149] Cf. p. 214.
[150] Cf. p. 191.
[151] T é normal, por hipótese.
[152] Cf. p. 210.
[153] Cf. p. 161.

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
$$v, T(v), T^2(v), \ldots, T^{m-1}(v).$$

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
$$T^{m-1}(v) \neq 0; \tag{7.65}$$
- $T^m(v) = 0$ acarreta
$$T^{m+j}(v) = 0,\ j = 0, 1, 2, \ldots \tag{7.66}$$

Assim, para
$$\sum_{j=0}^{m-1} a_{j+1} T^j(v) = 0, \tag{7.67}$$
onde $a_{j+1} \in \mathbb{K}$, $j = 0, 1, \ldots, m-1$, 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 (7.66) e (7.65). Portanto, (7.67) 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 (7.66) e (7.65). 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,[154] pela informação adicional da parte 1 da demonstração do lema 1.[155]

    (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.

[154] A existência desse autovalor segue de (R23) e (R25), pp. 187 e 189.

[155] O lema 1 encontra-se na página 220.

RESOLUÇÃO

Sejam α e β autovalores distintos de T.
(a) Para demonstrarmos que

$$\mathrm{Nu}(T-\alpha I)\cap \mathrm{Nu}\left((T-\beta I)^{\dim \mathcal{V}}\right)=\{0\}\,, \tag{7.68}$$

consideremos v pertencente à interseção do primeiro membro de (7.68). 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, (7.68) é válida.
(b) Caso

$$\mathcal{W}=\mathrm{Nu}\left((T-\alpha I)^{\dim \mathcal{V}}\right)\cap \mathrm{Nu}\left((T-\beta I)^{\dim \mathcal{V}}\right)$$

seja não nulo, como $T|_{\mathcal{W}}$ possui algum autovalor,[156] $\mathrm{Nu}\left((T-\alpha I)^{\dim \mathcal{V}}\right)$ e $\mathrm{Nu}\left((T-\beta I)^{\dim \mathcal{V}}\right)$ 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.[157] 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. [158]

RESOLUÇÕES

(a) Existe algum inteiro positivo p tal que

$$\begin{aligned}(L-0I)^p&=L^p\\&=0.\end{aligned}$$

Então, como, para $\lambda=0$,

$$\mathrm{Nu}\left((L-\lambda I)^p\right)=\mathcal{V}$$

é um dos subespaços da sequência (7.43) da demonstração do lema 1,[159]

$$\mathrm{Nu}\left(L^{\dim \mathcal{V}}\right)=\mathcal{V},$$

[156]Cf. (R23) e (R25).
[157]Cf. lema 2, p. 224.
[158]Esse resultado também é válido para $\mathbb{K}=\mathbb{R}$, conforme demonstrado no AXLER, citado no capítulo 1.
[159]Cf. p. 220.

por (7.44), (7.45) 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 (7.43) e (7.45) da demonstração do lema 1.[160]

RESOLUÇÃO

Para cada inteiro positivo i,

$$\mathcal{U}_i \subset \mathcal{U}_{i+1},$$

pois

$$\begin{aligned} v \in \mathcal{U}_i &\Longrightarrow v \in \operatorname{Nu}\Big((T - \lambda I)^i\Big) \\ &\Longrightarrow (T - \lambda I)^i(v) = 0_{\mathcal{V}} \\ &\Longrightarrow (T - \lambda I)\Big((T - \lambda I)^i(v)\Big) = (T - \lambda I)(0_{\mathcal{V}}) \\ &\Longrightarrow (T - \lambda I)^{i+1}(v) = 0_{\mathcal{V}} \\ &\Longrightarrow v \in \operatorname{Nu}\Big((T - \lambda I)^{i+1}\Big) \\ &\Longrightarrow v \in \mathcal{U}_{i+1}. \end{aligned}$$

Portanto, como a sequência de inclusões (7.43) é, de fato, crescente, em relação a (7.45), 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,

$$\begin{aligned} v \in \mathcal{U}_{\ell+2} &\Longrightarrow v \in \operatorname{Nu}\Big((T - \lambda I)^{\ell+2}\Big) \\ &\Longrightarrow (T - \lambda I)^{\ell+2}(v) = 0_{\mathcal{V}} \\ &\Longrightarrow (T - \lambda I)^{\ell+1}((T - \lambda I)(v)) = 0_{\mathcal{V}} \\ &\Longrightarrow (T - \lambda I)(v) \in \operatorname{Nu}\Big((T - \lambda I)^{\ell+1}\Big) \\ &\Longrightarrow (T - \lambda I)(v) \in \mathcal{U}_{\ell+1} \\ &\Longrightarrow (T - \lambda I)(v) \in \mathcal{U}_{\ell} \text{ (via (7.44))} \\ &\Longrightarrow (T - \lambda I)(v) \in \operatorname{Nu}\Big((T - \lambda I)^{\ell}\Big) \\ &\Longrightarrow (T - \lambda I)^{\ell}((T - \lambda I)(v)) = 0_{\mathcal{V}} \\ &\Longrightarrow (T - \lambda I)^{\ell+1}(v) = 0_{\mathcal{V}} \\ &\Longrightarrow v \in \operatorname{Nu}\Big((T - \lambda I)^{\ell+1}\Big) \\ &\Longrightarrow v \in \mathcal{U}_{\ell+1}. \end{aligned}$$

14. Prove a equivalência (7.49) da demonstração do lema 2.

RESOLUÇÃO

[160] Cf. p. 220.

Como $a \in \mathbb{K}$ é uma "matriz" 1×1 e toda transformação linear é do tipo "multiplicação por matriz",[161] 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 (7.56), da demonstração do lema 2, é base de $\mathcal{V}$.

RESOLUÇÃO

Por (R9),[162] basta demonstrarmos os seguintes fatos:

I. Os vetores de (7.56) são LI;
II. $\dim \mathcal{V}$ é igual ao número de vetores de (7.56).

DEMONSTRAÇÃO DO FATO I

Ao aplicarmos L a ambos os membros da CL nula

$$\sum_{i=1}^{r} \left(\sum_{j=0}^{q_i} a_{i,j} L^j(v_i) \right) + \sum_{k=1}^{q} b_k w_k = 0_{\mathcal{V}},^{163} \tag{7.69}$$

podemos reduzi-la a

$$\sum_{i=1}^{r} \left(\sum_{j=0}^{q_i - 1} a_{i,j} L^{j+1}(v_i) \right) = 0_{\mathcal{V}},^{164}$$

isto é,[165]

$$\sum_{i=1}^{r} \left(\sum_{j=0}^{q_i - 1} a_{i,j} L^{j}(u_i) \right) = 0_{\mathcal{V}},$$

que é uma CL nula dos vetores da base (7.53) de $\mathrm{Im}(T)$,[166] ou seja,

$$a_{1,0} = a_{1,1} = \ldots = a_{1,q_1 - 1} = \ldots = a_{r,0} = a_{r,1} = \cdots = a_{r,q_r - 1} = 0. \tag{7.70}$$

Portanto, ao substituirmos (7.70) em (7.69), obtemos a CL nula

$$\sum_{i=1}^{r} a_{i,q_i} L^{q_i}(v_i) + \sum_{k=1}^{q} b_k w_k = 0_{\mathcal{V}}. \tag{7.71}$$

Logo, como (7.55) é base de $\mathrm{Nu}(L)$,[167] os escalares de (7.71) são nulos e, portanto, todos os escalares de (7.69) também são nulos.

DEMONSTRAÇÃO DO FATO II

$$\begin{aligned} \dim \mathcal{V} &= \dim(\mathrm{Im}(L)) + \dim(\mathrm{Nu}(L)) && \text{(por (R14) do final da subseção 6.3.2)} \\ &= (q_1 + \cdots + q_r) + (r + q) && \text{(por (7.53) e (7.55) da demonstração do lema 2)} \\ &= (q_1 + 1) + \cdots + (q_r + 1) + q, \end{aligned}$$

que é exatamente o número de vetores de (7.56).

[161] Cf. (R15), p. 162.
[162] Cf. p. 154.
[163] Para todos os índices i, j, k admissíveis, $a_{i,j}, b_k \in \mathbb{C}$.
[164] De fato, os vetores de (7.55) pertencem a $\mathrm{Nu}(L)$, conforme a p. 225.
[165] Cf. (7.54), p. 225.
[166] Cf. p. 225.
[167] 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 $\mathcal{V}$,[168]

$$\begin{aligned}\mathcal{M}(I,\mathcal{B}',\mathcal{B})\,\mathcal{M}(I,\mathcal{B},\mathcal{B}') &= \mathcal{M}(I,\mathcal{B}) \\ &= \begin{bmatrix} 1 & & \text{zeros} \\ & \ddots & \\ \text{zeros} & & 1 \end{bmatrix},\end{aligned}$$

por (6.13).[169] 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 $L \in \mathcal{L}(\mathcal{V})$,

$$\mathcal{M}(L,\mathcal{B}') = \mathsf{P}^{-1}\mathcal{M}(L,\mathcal{B})\mathsf{P}, \tag{7.72}$$

por (6.13).[170]

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 (7.72).[171]

OBSERVAÇÃO

O operador (4.9) (da página 100) não é o operador identidade. Caso fosse, $[T]_{\mathcal{B}}$ e $[T^{-1}]_{\mathcal{B}'}$ seriam a matriz identidade em $\mathbb{R}^{n\times n}$ e, portanto, o uso de (4.11) (da página 100) estaria comprometido.[172] 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 (7.72), pois:

> *Caso $\mathcal{B}$ e $\mathcal{B}'$ sejam as bases ordenadas supracitadas e o operador $T \in \mathcal{L}(\mathcal{V})$ satisfaça a condição*
>
> $$T(v_k) = w_k,\ k = 1, \ldots, n,$$
>
> $$\mathcal{M}(T,\mathcal{B}) = \mathcal{M}(I,\mathcal{B}',\mathcal{B}).$$

[168] Para cada $v \in \mathcal{V}$, $I(v) = v$.

[169] Cf. p. 161.

[170] Matrizes que representam L (em bases distintas) são ditas *semelhantes (entre si)*.

[171] Essa abordagem também é válida para a forma de Jordan de uma matriz quadrada A, a ser estudada na subseção 7.6.2.

[172] Além disso, exemplos (nas subseções 4.1.2 e 4.2.1) comprovam que essas matrizes não podem ser a identidade.

De fato,

$$\begin{aligned}\mathcal{M}(T,\mathcal{B}) &= \mathcal{M}(IT,\mathcal{B},\mathcal{B})\\ &= \mathcal{M}\big(I,\mathcal{B}',\mathcal{B}\big)\,\mathcal{M}\big(T,\mathcal{B},\mathcal{B}'\big) \text{ (por (6.13))}\\ &= \mathcal{M}\big(I,\mathcal{B}',\mathcal{B}\big)\,\mathsf{I},\end{aligned}$$

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, o item 1 não é uma possibilidade viável, pelo exemplo que precede o teorema de Cayley-Hamilton.[173] 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 no item 2 supracitado.

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,

$$\{(A - \mathsf{I})v_2, v_2\} = \left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

[173] Esse teorema encontra-se na página 228.

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),[174] $\lambda = -1$ 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 o item 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 no item 1 supracitado.

[174] Cf. p. 189.

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

$$\mathsf{P}^{-1} A \mathsf{P} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

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

$$\mathrm{Nu}\left((A - 3\mathsf{I})^2\right) = \mathrm{Nu}\begin{pmatrix} 0 & -2 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 1 \end{pmatrix}$$
$$= \left[\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \end{pmatrix}\right],$$

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.

Comentário sobre a ciência e a arte da álgebra linear

> *So as is frequently true in life, the nature of the answer depends on how the question is posed.*
>
> *Shlomo Sternberg*

Na matemática, é possível que a importância do método utilizado na resolução de um dado problema tenha um peso equivalente ao do modo como esse método tenha sido empregado. Por exemplo, no escalonamento de uma dada matriz, embora sejam utilizadas operações elementares sobre as suas linhas, na obtenção de sua escalonada reduzida, o número de passos desse escalonamento, em geral, depende da pessoa que esteja escalonando. Do mesmo modo, no método para a obtenção da forma de Jordan de uma dada matriz, que acabamos de estudar, as escolhas dos autovetores, inclusive dos generalizados, para a composição dos ciclos, são totalmente individuais.
Na próxima subseção,[175] analisaremos um outro método para obtermos a forma de Jordan, baseado num diagrama de pontos, que é mais eficiente, ou seja, mais direto e menos trabalhoso, mas que ainda depende da "arte" do método, isto é, do modo como ele é utilizado.

[175] A subseção 7.6.3 é de autoria do Professor Marcelo Alves do DMAT da UFPR, conforme antecipado no capítulo 1.

7.6.3 Determinação da forma de Jordan via diagrama de pontos de um operador

Caso $\mathcal{V}$ seja um espaço vetorial complexo finitamente gerado e $\mathcal{B}$ seja uma base de Jordan para $T \in \mathcal{L}(\mathcal{V})$, isto é, dada como na demonstração do teorema 7.4.8, podemos associá-la à um diagrama de pontos dotado de toda a informação necessária para obtermos a forma de Jordan J de T. O mais interessante, nesse caso, é que podemos obter esse diagrama de outro modo e, assim, determinarmos J sem a necessidade de calcularmos $\mathcal{B}$. Além disso, o diagrama de pontos também pode ser útil no cálculo dessa base, caso seja necessário.

Nessa última subseção do livro, $\mathcal{V}$ representa o espaço supracitado e todos os operadores pertencem ao espaço $\mathcal{L}(\mathcal{U})$, onde $\mathcal{U}$ é um subespaço de $\mathcal{V}$. Além disso, o diagrama supracitado será construído, primeiramente, para operadores nilpotentes e, em seguida, imediatamente estendido para operadores em geral.

DEFINIÇÕES/OBSERVAÇÃO

- Seja $L \in \mathcal{L}(\mathcal{V})$ nilpotente.
  - O *grau de nilpotência* de L é o menor expoente $n(L)$ tal que $L^{n(L)} = 0$;
  - $n(L) \leq \dim(V)$;[176]
  - Considere os vetores v_j e os números inteiros p_j como no lema 2 e as bases $\mathcal{B}_{i,j}$ como na equação (7.58). Suponha que os inteiros supracitados estejam em uma sequência ordenada, digamos

    $$p_1 \geq p_2 \geq \ldots \geq p_s.$$

    Nesse caso, o *diagrama de pontos* de L tem s colunas e a coluna $j \in \{1, \ldots, s\}$ tem p_j pontos, ou seja, cada vetor do j-ésimo ciclo

    $$L^{p_j - 1}(v_j), \ldots, L(v_j), v_j$$

    pode ser representado por um ponto na coluna j.
- Caso $T \in \mathcal{L}(\mathcal{V})$ e a lista $\lambda_1, \ldots, \lambda_r$ represente seus autovalores distintos, o *diagrama de pontos* de T associado ao autovalor λ_i é o diagrama de pontos do operador nilpotente $(T - \lambda_i)\mid_{U_i} : U_i \to U_i$.

EXEMPLOS

- Caso um operador nilpotente $L : \mathbb{C}^5 \to \mathbb{C}^5$ possua uma base de Jordan composta por dois ciclos, digamos

  $$\mathcal{B}_{1,1} = \{L^2(v_1), L(v_1), v_1\} \text{ e } \mathcal{B}_{1,2} = \{L(v_2), v_2\},$$

  podemos representar esses ciclos em colunas por:

  $$\begin{array}{cc} L^2(v_1) & L(v_2) \\ L(v_1) & v_2 \\ v_1 & \end{array}$$

[176] Confira o exercício 12 da seção 7.5.

Essa representação pode ser simplificada no seguinte diagrama de pontos:

$$\begin{array}{cc} \bullet & \bullet \\ \bullet & \bullet \\ \bullet & \end{array}$$

- Caso $T : \mathbb{C}^7 \to \mathbb{C}^7$ tenha dois autovalores, digamos $\lambda_1 = 2$ e $\lambda_2 = i$, e o diagrama de pontos

$$\begin{array}{ccccccc} \multicolumn{2}{c}{\lambda_1 = 2} & \quad & \multicolumn{3}{c}{\lambda_2 = i} \\ \bullet & \bullet & & \bullet & \bullet & \bullet \\ \bullet & & & \bullet & & \end{array}$$

sabemos que, para λ_1, temos os ciclos

$$\mathcal{B}_{1,1} = \{(T - 2\mathsf{I})(v_1), v_1\} \text{ e } \mathcal{B}_{1,2} = \{v_2\}$$

e, para λ_2, os ciclos

$$\mathcal{B}_{2,1} = \{(T - i\mathsf{I})(w_1), w_1\}, \ \mathcal{B}_{2,2} = \{w_2\} \text{ e } \mathcal{B}_{2,3} = \{w_3\}.$$

Portanto, T possui a seguinte matriz de Jordan

$$J = \left(\begin{array}{cc|c|cc|c|c} 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & i & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & i & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & i & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & i \end{array} \right).$$

Veremos, agora, que o diagrama de pontos de um operador pode ser obtido e, portanto, a forma de Jordan desse operador, sem que seja necessário calcularmos uma base de Jordan.

Teorema do diagrama de pontos

Seja L um operador nilpotente, de grau de nilpotência $n(L)$, e considere v_j, p_j e $\mathcal{B}_{i,j}$ como definidos na página 247. Se r_i é o número de pontos que estão na i-ésima linha do diagrama de pontos de L, então:

1. *O diagrama possui $n(L)$ linhas.*

2. *Para cada r, os vetores associados aos pontos nas primeiras r linhas formam uma base de $\mathrm{Nu}(L^r)$.*

3. $r_1 = \dim(\mathrm{Nu}(L))$.

4. $r_i = \dim(\mathrm{Nu}(L)^i) - \dim(\mathrm{Nu}(L)^{i-1})$ *para todo $i > 1$ admissível.*

DEMONSTRAÇÃO

1. Esse item segue do fato que a coluna mais longa possui p_1 pontos e $n(L) = p_1$.

2. Os pontos de cada coluna j correspondem ao ciclo

$$L^{p_j-1}(v_j), L^{p_j-2}(v_j), \ldots, L(v_j), v_j,$$

nessa ordem, do início ao final da coluna. Desse modo, um vetor que corresponde a um ponto na linha i do diagrama de pontos de L é da forma $L^{p_j-i}(v_j)$ para algum j. Como $L^k(v_j) = 0$ se, e somente se, $k \geq p_j$, temos

$$L^r(L^{p_j-i}(v_j)) = L^{p_j+(r-i)}(v_j) = 0$$

se, e somente se, $r - i \geq 0$, ou seja, $r \geq i$. Assim, $L^{p_j-i}(v_j) \in \mathrm{Nu}(L^r)$ se, e somente se, $r \geq i$. Logo, existe um conjunto formado por vetores LI dados por

$$L^{p_j-i}(v_j),\ 1 \leq i \leq r,\ 1 \leq j \leq s,$$

contido no núcleo de L^r. Esse conjunto é uma base de $\mathrm{Nu}(L^r)$. De fato, como

$$\mathcal{B} = \left\{ L^{p_j-i}(v_j) \ : \ 1 \leq i \leq p_j, 1 \leq j \leq s \right\}$$

é uma base de $\mathcal{V}$,[177] ao considerarmos $v \in \mathrm{Nu}(L^r)$, temos a CL

$$v = \sum_{1 \leq j \leq s} \sum_{1 \leq i \leq p_j} a_{i,j} L^{p_j-i}(v_j)$$

e, portanto,

$$\begin{aligned} 0 = L^r(v) &= \sum_{1 \leq j \leq s} \sum_{1 \leq i \leq p_j} a_{i,j} L^{p_j+(r-i)}(v_j) \\ &= \sum_{1 \leq j \leq s} \sum_{1 \leq i \leq r} a_{i,j} L^{p_j+(r-i)}(v_j) + \sum_{1 \leq j \leq s} \sum_{r < i \leq p_j} a_{i,j} L^{p_j+(r-i)}(v_j) \\ &= \sum_{1 \leq j \leq s} \sum_{r < i \leq p_j} a_{i,j} L^{p_j+(r-i)}(v_j), \end{aligned}$$

que acarreta $a_{i,j} = 0$ se $i > r$, pois os vetores de $\mathcal{B}$ são LI. Consequentemente, os vetores

$$L^{p_j-i}(v_j),\ 1 \leq i \leq r,\ 1 \leq j \leq s,$$

formam uma base de $\mathrm{Nu}(L^r)$, por serem geradores LI desse núcleo.

3. Esse item é imediato, pelo item (2).

4. Para $i > 1$, pelo item (1), o número de pontos nas i primeiras linhas do diagrama é a dimensão de $\mathrm{Nu}(L^i)$, o mesmo valendo para $\dim(\mathrm{Nu}(L^{i-1}))$ e o número de pontos nas $i-1$ primeiras linhas. Consequentemente, o número de pontos da i-ésima linha é $\dim(\mathrm{Nu}(L^{i-1})) - \dim(\mathrm{Nu}(L^i))$.

Utilizando-se o teorema do núcleo e da imagem,[178] temos um segundo modo de calcular os números r_i supracitados.

Corolário

Caso L seja um operador nilpotente e r_i seja o número de pontos que estão na i-ésima linha do diagrama de pontos de L,

1. $r_1 = \dim \mathcal{V} - \dim(\mathrm{Im}(L))$;
2. $r_i = \dim(\mathrm{Im}(L^{i-1})) - \dim(\mathrm{Im}(L^i))$ *para todo* $i > 1$ *admissível.*

[177] Confira o lema 2, página 224.
[178] Cf. (R14), p. 159.

OBSERVAÇÃO

Note que, pelo teorema supracitado,[179] é possível construirmos o diagrama de pontos de um operador nilpotente L a partir do cálculo dos números r_i. Esse cálculo termina quando obtemos, pela primeira vez, a igualdade $L^i = 0$, ou seja, $i = n(L)$, pois, para $j \geq i$, temos $r_j = 0$. Por outro lado, a construção do diagrama de pontos passa a ser realmente interessante quando, por algum motivo, *não* temos conhecimento de qualquer base de Jordan para L. De fato, a determinação do número de pontos em cada *linha* fornece o diagrama, que fornece as *colunas*, e, a partir delas, sabemos quantos ciclos existem e qual o comprimento de cada um deles. Assim, podemos obter, imediatamente, a forma de Jordan do operador. Além disso, note que, ao compararmos o modo como a forma supracitada foi obtida (nessa subseção) com os passos I–V estudados na subseção 7.6.2, o cálculo dos núcleos das potências de $(T - \lambda_i \mathsf{I})$ foi acrescentado ao passo II, enquanto que o passo III passa a ser realizado pelo diagrama de pontos. Por outro lado, o passo IV, onde uma base composta por ciclos é obtida, é executado exatamente como em 7.6.2. Essa análise e o fato de podermos utilizar as informações sobre o comprimento dos ciclos para construir uma base de Jordan são ilustrados nos próximos exemplos.

EXEMPLOS

· Caso a matriz de um operador nilpotente $L : \mathbb{C}^5 \to \mathbb{C}^5$, na base canônica, seja dada por

$$M = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 \\ 4 & -2 & 0 & 0 & 0 \\ -3 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 & 4 \\ 0 & 0 & 0 & -1 & 2 \end{pmatrix},$$

temos

$$M^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

e $M^3 = \mathsf{O}$. Isso confirma a nilpotência de L e calcula o seu índice: $n(L) = n(M) = 3$. Logo, calculando-se os núcleos dessas matrizes (não nulas), temos

$$\begin{aligned} \mathrm{Nu}(M) &= [(0,0,1,0,0),(0,0,0,2,1)] \text{ e} \\ \mathrm{Nu}(M^2) &= [(1,2,0,0,0),(0,0,1,0,0),(0,0,0,1,0),(0,0,0,0,1)]. \end{aligned}$$

Portanto, como as suas dimensões são

$$\begin{aligned} \dim \mathrm{Nu}(M) &= 2 \text{ e} \\ \dim \mathrm{Nu}(M^2) &= 4, \end{aligned}$$

o número de pontos em cada linha será

$$\begin{aligned} r_1 &= \dim \mathrm{Nu}(M) = 2, \\ r_2 &= \dim \mathrm{Nu}(M^2) - \dim \mathrm{Nu}(M) = 2 \text{ e} \\ r_3 &= \dim \mathrm{Nu}(M^3) - \dim \mathrm{Nu}(M^2) = 1. \end{aligned}$$

[179] Cf. p. 248.

Assim, o diagrama de pontos correspondente é dado por

$$\begin{array}{cc} \bullet & \bullet \\ \bullet & \bullet \\ \bullet & \end{array}$$

Consequentemente, a forma de Jordan de L possui um bloco de ordem 3 e um bloco de ordem 2, dados por

$$J = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Agora, construiremos uma base de Jordan via ciclos. Note que é imprescindível que os vetores obtidos a cada ciclo sejam LI com os vetores obtidos nos ciclos anteriores. Para que isso ocorra, basta verificarmos que os vetores pertencentes ao núcleo de M, em cada ciclo, são LI.[180]

PRIMEIRO CICLO

Esse ciclo terá comprimento 3 e, para gerá-lo, precisaremos de um vetor

$$v_1 \in \mathbb{C}^5 - \mathrm{Nu}(M^2).$$

Pela descrição dos vetores do núcleo de M^2, podemos considerar $v_1 = (1,0,0,0,0)$. Portanto, temos o seguinte ciclo (gerado a partir de v_1):

$$\begin{aligned} v_1 &= (1,0,0,0,0), \\ Mv_1 &= (2,4,-3,0,0), \\ M^2v_1 &= (0,0,2,0,0). \end{aligned}$$

SEGUNDO CICLO

Esse ciclo terá comprimento 2 e, para gerá-lo, precisaremos de um vetor $v_2 \in \mathbb{C}^5$ tal que

$$v_2 \in \mathrm{Nu}(M^2) - \mathrm{Nu}(M), \text{ com } M^2v_1 \text{ e } Mv_2 \text{ LI}.$$

[180]Verifique essa independência linear.

SUGESTÃO

Suponha que obtivemos uma lista de vetores, digamos $v_1, v_2, \ldots, v_k$, tal que $L^{p_j}(v_j) = 0$, $j = 1,2,\ldots,k$, e os vetores da lista $L^{p_1-1}(v_1), \ldots, L^{p_k-1}(v_k)$ são LI. Use essa suposição para provar que os vetores dados por

$$L^{\ell_j}(v_j), \text{ com } 1 \leq j \leq k \text{ e } 0 \leq \ell_j \leq p_j - 1,$$

são LI.

Como podemos representar $v \in \mathrm{Nu}(M^2)$ por $(\alpha, 2\alpha, \gamma, \delta, \varepsilon)$, onde α, γ, δ e ε são escalares que dependem do vetor v, temos

$$Mv = (0, 0, \alpha, -2\delta + 4\varepsilon, -\delta + 2\varepsilon).$$

Assim, para obtermos um vetor v_2 tal que Mv_2 e M^2v_1 sejam LI, podemos escolher $\alpha = \gamma = \delta = 0$ e $\varepsilon = 1$. Logo, $v_2 = (0, 0, 0, 0, 1)$ e $Mv_2 = (0, 0, 0, 4, 2)$, que são LI com $M^2v_1 = (0, 0, 2, 0, 0)$.

Portanto, combinando os dois ciclos supracitados, uma base de Jordan completa é dada por

$$\mathcal{B} = \{M^2v_1, Mv_1, v_1, Mv_2, v_2\}.$$

· Caso a matriz de um operador não nilpotente $T : \mathbb{C}^4 \to \mathbb{C}^4$, na base canônica, seja dada por

$$M = \begin{pmatrix} 3 & 1 & 1 & -6 \\ 0 & 3 & 0 & -8 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

$p(t) = (t-3)^3(t+1)$ é o seu polinômio característico e, assim, 3 e -1 são os seus autovalores. Além disso, o expoente de cada fator $(t - \lambda_i)$ é a dimensão do autoespaço generalizado de T associado ao autovalor λ_i.[181]

DIAGRAMA DE PONTOS DE T ASSOCIADO AO AUTOVALOR $\lambda_1 = 3$

Note que

$$M - 3\mathsf{I} = \begin{pmatrix} 0 & 1 & 1 & -6 \\ 0 & 0 & 0 & -8 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -4 \end{pmatrix}$$

não é nilpotente.[182] Mesmo assim, podemos calcular os números r_i a partir das potências de $M - 3\mathsf{I}$. De fato, o posto de $M - 3\mathsf{I}$ é igual a $\dim(\mathrm{Im}(M - 3\mathsf{I})) = 2$, pois temos apenas duas colunas LI, e

$$(M - 3\mathsf{I})^2 = \begin{pmatrix} 0 & 0 & 0 & 16 \\ 0 & 0 & 0 & 32 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 16 \end{pmatrix},$$

acarretando $\dim(\mathrm{Im}(M - 3\mathsf{I})^2) = 1$. Consequentemente,

$$r_1 = 4 - 2 = 2, r_2 = 2 - 1 = 1 \text{ e } r_1 + r_2 = 3,$$

[181] Confira a definição que segue o exemplo.1 da página 226.

[182] O operador $T - 3\mathsf{I}$ é nilpotente apenas quando restrito ao autoespaço generalizado associado à $\lambda_1 = 3$.

onde, portanto, atingimos a dimensão do autoespaço generalizado de T associado à $\lambda_1 = 3$ e, assim, concluímos o cálculo do diagrama de pontos de T associado à esse autovalor:

$$\begin{array}{cc} \bullet & \bullet \\ \bullet & \end{array}$$

Note que, se continuarmos o cálculo das potências de $M - 3\mathsf{I}$, a próxima será (novamente) uma matriz com uma única coluna não nula e, então, teremos $r_3 = 1 - 1 = 0$.

$\boxed{\text{DIAGRAMA DE PONTOS DE } T \text{ ASSOCIADO AO AUTOVALOR } \lambda_2 = -1}$

Como a dimensão do autoespaço generalizado de T associado à λ_2 é 1, o diagrama de pontos supracitado tem apenas um ponto.

$\boxed{\text{DIAGRAMA DE PONTOS DE } T}$

Como

$$\begin{array}{ccc} \lambda_1 = 3 & \qquad & \lambda_2 = -1 \\ \begin{array}{cc} \bullet & \bullet \\ \bullet & \end{array} & & \bullet \end{array}$$

a matriz de Jordan de T pode ser representada por

$$J = \left(\begin{array}{cc|c|c} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ \hline 0 & 0 & 3 & 0 \\ \hline 0 & 0 & 0 & -1 \end{array}\right).$$

Nesse caso, também podemos obter uma base de Jordan pelo diagrama de pontos: há um ciclo de comprimento 2 e um ciclo de comprimento 1, associados ao autovalor λ_1, e um ciclo de comprimento 1, associado ao autovalor λ_2. De fato, note que, ao escalonarmos $M - 3\mathsf{I}$ e $(M - 3\mathsf{I})^2$, obtemos

$$\mathrm{Nu}(M - 3\mathsf{I}) = [(1,0,0,0), (0,1,-1,0)] \quad \text{e}$$
$$\mathrm{Nu}((M - 3\mathsf{I})^2) = [(1,0,0,0), (0,1,0,0)(0,0,1,0)].$$

Logo, como o vetor $v_1 = (0,0,1,0)$ está em $\mathrm{Nu}((M-3\mathsf{I})^2) - \mathrm{Nu}(M-3\mathsf{I})$, temos o ciclo

$$\mathcal{B}_{1,1} = \{(M - 3\mathsf{I})v_1, v_1\} = \{(1,0,0,0), (0,0,1,0)\}.$$

Por outro lado, como o segundo ciclo de T associado ao autovalor $\lambda_1 = 3$ tem apenas um elemento, que é um autovetor, podemos considerar $v_2 = (0,1,-1,0)$,[183] que é LI com o autovetor $(M - 3\mathsf{I})v_1 = (1,0,0,0)$ (obtido no primeiro ciclo). Assim, temos o ciclo

$$\mathcal{B}_{1,2} = \{(0,1,-1,0)\}.$$

[183] $v_2 \in \mathrm{Nu}(M - 3\mathsf{I})$.

Finalmente, precisamos obter um autovetor de T associado ao autovalor $\lambda_2 = -1$. Como a forma escalonada de

$$M + \mathsf{I} = \begin{pmatrix} 4 & 1 & 1 & -6 \\ 0 & 4 & 0 & -8 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

é a matriz

$$N = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

temos

$$\mathrm{Nu}(M + \mathsf{I}) = \mathrm{Nu}(N) = [(1,2,0,1)].$$

Portanto, como $v_3 = (1,2,0,1)$ é um autovetor de T associado à λ_2, temos o ciclo

$$\mathcal{B}_{2,1} = \{(1,2,0,1)\}.$$

Assim, ao juntarmos todos os ciclos dos autovalores supracitados, temos a seguinte base de Jordan para T:

$$\mathcal{B} = \{(1,0,0,0), (0,0,1,0), (0,1,-1,0), (1,2,0,1)\}.$$

Consequentemente, caso P seja a matriz de T com colunas dadas pelos vetores de $\mathcal{B}$, temos, de fato, $\mathsf{P}^{-1}M\mathsf{P} = J$.

· À guisa de comparação de métodos, consideraremos o operador cuja matriz, na base canônica, é dada no exemplo 3 da subseção 7.6.2, ou seja,

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

com polinômio característico dado por $p(t) = (t-2)^3(t-3)$. Como o autovalor $\lambda = 3$ de A tem multiplicidade algébrica igual a 1, sua multiplicidade geométrica também é 1. Assim, o seu autoespaço generalizado coincide com o seu autoespaço e o diagrama de pontos do operador considerado possui apenas um ponto. Por outro lado, como vimos no exemplo supracitado, para $\lambda = 2$, temos

$$A - 2\mathsf{I} = \begin{pmatrix} 0 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}, \ (A - 2\mathsf{I})^2 = \begin{pmatrix} 0 & -2 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 1 \end{pmatrix},^{184}$$

$$\mathrm{Nu}(A - 2\mathsf{I}) = \left[\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right], \ \mathrm{Nu}((A - 2\mathsf{I})^2) = \left[\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \end{pmatrix} \right]$$

[184] Note que $(A - 2\mathsf{I})^2 = (A - 2\mathsf{I})^3$.

e $\mathrm{Nu}((A-2\mathrm{I})^3) = \mathrm{Nu}((A-2\mathrm{I})^2)$.[185] Portanto, o diagrama de pontos de $A - 2\mathrm{I}$ possui apenas duas linhas e é da forma

$$\begin{array}{cc} \bullet & \bullet \\ \bullet & \end{array}$$

e o diagama de pontos de A é

$$\begin{array}{ccc} \lambda_1 = 2 & \qquad & \lambda_2 = 3 \\ \begin{array}{cc} \bullet & \bullet \\ \bullet & \end{array} & & \bullet \end{array}$$

Logo, para A, temos a forma de Jordan

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

e a base de Jordan como obtida no passo IV do exemplo 3 supracitado.

[185] Confira a nota de rodapé anterior.