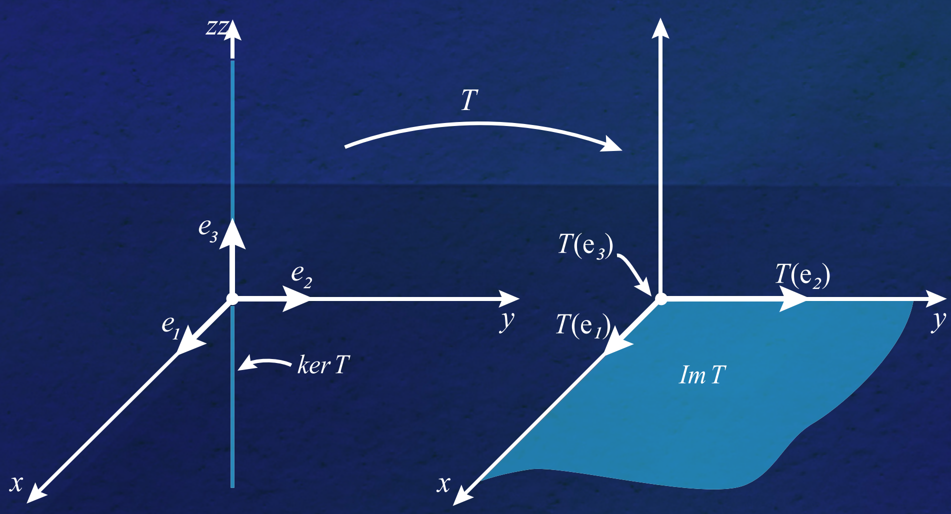
Antônio de Andrade e Silva

João Bosco Batista Lacerda

UM CURSO DE ÁLGEBRA LINEAR



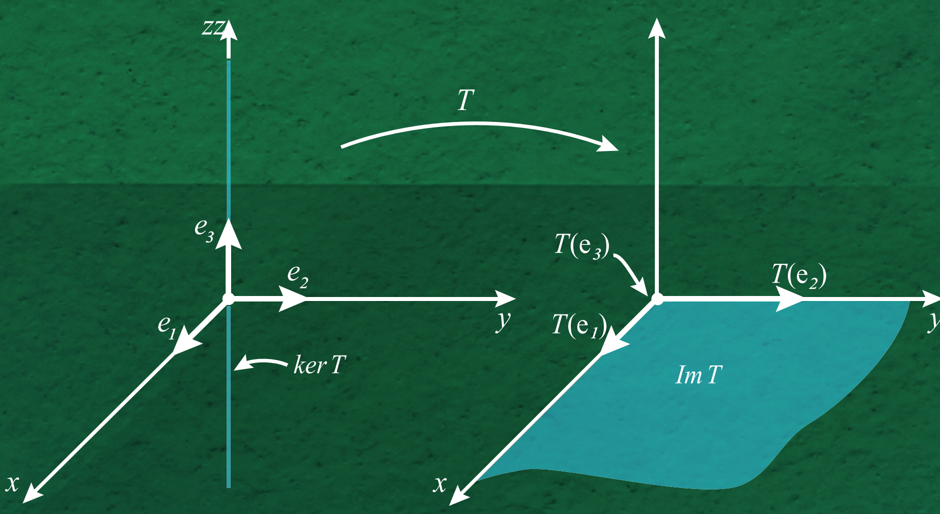
UNIFSM
CENTRO UNIVERSITÁRIO SANTA MARIA

pimenta cultural

Antônio de Andrade e Silva

João Bosco Batista Lacerda

UM CURSO DE ÁLGEBRA LINEAR



UNIFSM
CENTRO UNIVERSITÁRIO SANTA MARIA
I São Paulo I 2023 I
pimenta cultural

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)

AN553u

Andrade e Silva, Antônio de.

Um Curso de Álgebra Linear / Antônio de Andrade e Silva, João Bosco Batista Lacerda. – São Paulo: Pimenta Cultural, 2023.

Livro em PDF

ISBN 978-65-5939-779-2
DOI 10.31560/pimentacultural/2023.97792

1. Álgebra. 2. Matrizes. 3. Sistemas lineares. 4. Espaços vetoriais. 5. Transformações lineares. I. Andrade e Silva, Antônio de. II. Lacerda, João Bosco Batista. III. Título.

CDD: 512

Índice para catálogo sistemático:
I. Álgebra.
Jéssica Oliveira • Bibliotecária • CRB-034/2023

Direção editorial	Patricia Bieging Raul Inácio Busarello
Editora executiva	Patricia Bieging
Coordenadora editorial	Landressa Rita Schiefelbein
Assistente editorial	Bianca Bieging
Diretor de criação	Raul Inácio Busarello
Assistente de arte	Naiara Von Groll
Editoração eletrônica	Andressa Karina Voltolini Potira Manoela de Moraes
Bibliotecária	Jéssica Castro Alves de Oliveira
Imagens da capa	Acervo dos autores; Bedneyimages, Pikisuperstar, Zffoto - Freepik
Tipografias	Nimbus roman No9, Acumin, Balboa, Rockwell
Revisão	Os autores
Organizador	Antonio de Andrade e Silva João Bosco Batista Lacerda

PIMENTA CULTURAL
São Paulo • SP
+55 (11) 96766 2200
livro@pimentacultural.com
www.pimentacultural.com

pimenta cultural

2 0 2 3

PARECERISTAS E REVISORES(AS) POR PARES

Avaliadores e avaliadoras Ad-Hoc

PARECER E REVISÃO POR PARES

Os textos que compõem esta obra foram submetidos para avaliação do Conselho Editorial da Pimenta Cultural, bem como revisados por pares, sendo indicados para a publicação.

Prefácio

"Quando jovem, aprendemos.
Quando velho, compreendemos."

Albert Einstein

Este texto surgiu da experiência dos autores quando ministraram algumas vezes a disciplina Álgebra Linear e Geometria Analítica para vários cursos na Universidade Federal da Paraíba.

O principal objetivo deste texto é fazer uma apresentação rigorosa e clara das provas dos Teoremas e exemplos da Álgebra Linear no nível de graduação, desenvolvendo, também, a capacidade de modelagem de problemas e provas envolvendo combinações lineares, transformações lineares e matrizes, diagonalizações de operadores lineares e classificações de quádricas. Além disso, resolver problemas que envolvam matrizes utilizando a forma canônica de Jordan.

É nossa expectativa que este texto assuma o caráter de espinha dorsal de uma experiência permanentemente renovável, sendo, portanto, bem vindas as críticas e/ou sugestões apresentadas por todos - professores ou alunos quantos dele fizerem uso.

O leitor interessado em aprender a utilizar um programa de computação, por exemplo o Maple, como ferramenta na aprendizagem da Álgebra Linear e Geometria Analítica pode consultar uma das referências [1,3,5,8]

Para desenvolver a capacidade do estudante de pensar por si mesmo em termos das novas definições, incluímos no final de cada seção uma extensa lista de exercícios, onde a maioria dos exercícios dessas listas foram selecionados dos livros citados no final do texto. Devemos, porém, alertar aos leitores que os exercícios variam muito em grau de dificuldade, sendo assim, não é necessário resolver todos numa primeira leitura.

Nos capítulos 1 e 2 apresentaremos as principais definições e resultados sobre matrizes e sistemas de equações lineares que serão necessárias para o desenvolvimento deste texto.

No capítulo 3 apresentaremos definições abstratas de espaços vetoriais e subespaços, combinações lineares, conjuntos linearmente independentes e dependentes, bases e dimensão, coordenadas de um vetor e mudança de bases. Esse capítulo envolve o desenvolvimento axiomático de vetores, sendo assim, exigindo maior esforço no início do curso tanto do professor quanto do estudante.

No capítulo 4 apresentaremos transformações lineares, núcleo e imagem de uma transformação linear e representação matricial. A representação matricial proporciona um modo elegante de desenvolver a álgebra das matrizes e a geometria das transformações lineares.

No capítulo 5 apresentaremos as definições de autovalores e autovetores de um operador linear, o polinômio característico e minimal de um operador linear e operadores diagonalizáveis. Esse capítulo inicia o estudo das relações de equivalências e das formas canônicas, úteis nas aplicações que envolvem representações matriciais.

No capítulo 6 apresentaremos definições abstratas de espaços com produto interno, processo de ortogonalização de Gram-Schmide e o complemento ortogonal. Esse capítulo introduz a noção de conceitos métricos sobre um espaço vetorial qualquer.

No capítulo 7 apresentaremos o teorema da decomposição primária, operadores nilpotentes e a forma canônica de Jordan, a qual é uma ferramenta poderosa no estudo das relações de equivalência de matrizes.

Finalmente, no capítulo 8 apresentaremos operadores lineares especiais tais como: operador adjunto, ortogonais e simétricos e usá-los-emos para classificar as quádricas.

Agradecemos aos colegas e alunos do Departamento de Matemática que direta ou indiretamente contribuíram para a realização deste texto.

Antônio de Andrade e Silva
João Bosco Batista Lacerda

Aos nossos filhos:

José Augusto, Amanda e Fernanda.

Larissa e Natália.

Sumário

1
Matrizes

Neste capítulo relembramos alguns conceitos básicos sobre matrizes e determinantes, com o principal objetivo de fixarmos as notações e terminologias. Somente ideias e resultados necessários para a nossa discussão de matrizes especiais, muitos sem provas. O leitor interessado em mais detalhes pode consultar as referências no final do texto.

1.1 Notações Fundamentais

Nesta seção introduziremos algumas definições, notações e terminologias que serão usadas em todo texto.

Uma das maneiras de representar um *conjunto* é:

$$S = \{x \in U : P(x) \text{ é satisfeita}\},$$

em que U é o universo dos objetos (elementos), o símbolo ":" significa **tal que** e $P(x)$ é a propriedade ou condição que os elementos de U têm que satisfazer.

Um *corpo* é um conjunto F munido com duas operações binárias

$$\begin{array}{rccc} + : & F \times F & \to & F \\ & (x, y) & \longmapsto & x + y \end{array} \quad \text{e} \quad \begin{array}{rccc} \cdot : & F \times F & \to & F \\ & (x, y) & \longmapsto & x \cdot y \end{array}$$

chamadas de *adição* e *multiplicação*, tais que os seguintes axiomas (ou propriedades) são satisfeitos:

Axioma 1. (Leis comutativas)

$$x + y = y + x \text{ e } x \cdot y = y \cdot x, \; \forall \; x, y \in F.$$

Axioma 2. (Leis associativas)

$$x + (y + z) = (x + y) + z \text{ e } x \cdot (y \cdot z) = (x \cdot y) \cdot z, \; \forall \; x, y, z \in F.$$

Axioma 3. (Lei distributiva)

$$x \cdot (y + z) = x \cdot y + x \cdot z, \; \forall \; x, y, z \in F.$$

Axioma 4. (Elementos identidades) Existem dois elementos distintos 0 e 1 em F tais que

$$0 + x = x \text{ e } 1 \cdot x = x, \; \forall \; x \in F.$$

Axioma 5. (Elementos inversos) Se $x \in F$, então existe um único $-x \in F$ tal que $x + (-x) = 0$. Se $x \in F$ e $x \neq 0$, então existe um único $x^{-1} \in F$ tal que $x \cdot x^{-1} = 1$. Observe que as operações $+$ e $\cdot$ estão relacionadas pelo axioma (3).

Exemplo 1.1 *Seja $F = GF(2) = \{0, 1\}$. Definimos uma adição e uma multiplicação em F pelas tábuas*:

$$\begin{array}{c|cc} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \qquad e \qquad \begin{array}{c|cc} \cdot & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$

É fácil verificar que F com essas duas operações é um corpo, chamado de corpo de Galois. *Como uma ilustração, dados $x, y, z \in F$, obtemos*

$$x + (y + z) = \begin{cases} 0, & se \;\; x = 0 \;\; e \;\; y + z = 0 \;\; ou \;\; x = 1 \;\; e \;\; y + z = 1 \\ 1, & se \;\; x = 1 \;\; e \;\; y + z = 0 \;\; ou \;\; x = 0 \;\; e \;\; y + z = 1 \end{cases}$$

e

$$(x + y) + z = \begin{cases} 0, & se \;\; x + y = 0 \;\; e \;\; z = 0 \;\; ou \;\; x + y = 1 \;\; e \;\; z = 1 \\ 1, & se \;\; x + y = 1 \;\; e \;\; z = 0 \;\; ou \;\; x + y = 0 \;\; e \;\; z = 1. \end{cases}$$

Portanto, $x + (y + z) = (x + y) + z$, para todos $x, y, z \in F$.

Sejam F e K corpos. Diremos que K é uma *extensão* de F se $F \subseteq K$ e, neste caso, F é um *subcorpo* de K.

Agora apresentaremos as nossas principais notações e fatos:

- $\mathbb{N} = \{1, 2, 3, \ldots\}$ é o conjunto dos números naturais

- $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ é o conjunto dos números inteiros

- $\mathbb{Z}_+ = \{0, 1, 2, 3, \ldots\}$ é o conjunto dos números inteiros positivos

- $\mathbb{Z}^* = \mathbb{Z} - \{0\}$ ou $\mathbb{Z}^\times = \mathbb{Z} - \{0\}$

- $\mathbb{Q} = \left\{\frac{a}{b} : a, b \in \mathbb{Z}, \text{ com } b \neq 0\right\}$ é o conjunto dos números racionais. Neste caso, as operações são:

$$\begin{array}{rcl} \frac{a}{b} + \frac{c}{d} & = & \frac{ad+bc}{bd} \\ \frac{a}{b} \cdot \frac{c}{d} & = & \frac{ac}{bd} \\ \frac{a}{b} \div \frac{c}{d} & = & \frac{a}{b} \cdot \frac{d}{c}. \end{array}$$

- $\mathbb{R}$ é o conjunto dos números reais. No qual, suponhamos a existência de uma relação de ordem "$\leq$", com a seguinte propriedade: para cada $a \in \mathbb{R}$, com $a \geq 0$, existe um $x \in \mathbb{R}$ tal que $x^2 = a$. Neste caso, $x = \pm\sqrt{a}$.

- $\mathbb{C} = \left\{a + bi : a, b \in \mathbb{R}, \text{ com } i^2 = -1\right\}$ é o conjunto dos números complexos. Neste caso, as operações são:

$$\begin{array}{rcl} (a + bi) + (c + di) & = & (a + c) + (b + d)i \\ (a + bi) \cdot (c + di) & = & (ac - bd) + (bc + ad)i \end{array}$$

e $\bar{z} = a - bi$ é o *conjugado complexo* de $z = a + bi$. O número real

$$|z|^2 = z \cdot \bar{z} = a^2 + b^2 \text{ ou } |z| = \sqrt{a^2 + b^2},$$

chama-se *norma* ou *módulo* de z. Assim, se $z \neq 0$, então $z = rw$, com $r > 0$ e $|w| = 1$. Além disso,

$$\frac{1}{z} = \frac{1}{|z|^2}\bar{z}.$$

As notações $\mathrm{Re}(z) = \frac{1}{2}(z + \bar{z})$ e $\mathrm{Im}(z) = \frac{1}{2}(z - \bar{z})$ representam a *parte real* e a *parte imaginária* de z. Note que

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

Os elementos de um subcorpo F de $\mathbb{C}$ serão chamados de *escalares*. Em tudo que segue, salvo menção explícita em contrário, F representa um subcorpo do corpo dos números complexos $\mathbb{C}$.

- (**Princípio da Boa Ordenação**-PBO) Qualquer subconjunto não vazio de $\mathbb{Z}_+$ possui um menor elemento.

- (**Princípio de Indução Finita**-PIF) Seja S um subconjunto de $\mathbb{N}$ que goza das seguintes propriedades:

  (a) $1 \in S$ (base de indução).

  (b) Se $n \in S$, então $n + 1 \in S$ (PIF).

  Então $S = \mathbb{N}$.

- (**Princípio de Indução Completo**-PIC) Seja S um subconjunto de $\mathbb{N}$ que goze das seguintes propriedades:

  (a) $1 \in S$ (base de indução).

  (b) Para cada $n \in \mathbb{N}$, se $\{1, 2, \ldots, n\} \subseteq S$, então $n + 1 \in S$. (PIF)

  Então $S = \mathbb{N}$.

- Sejam S um conjunto, $n \in \mathbb{N}$ e $I_n = \{1, 2, \ldots, n\}$. Uma *lista* sobre S é qualquer função $\mathbf{x} : I_n \to S$ definida como $\mathbf{x}(i) = x_i$. Neste caso, identificamos o conjunto de todas as lista sobre S com o conjunto de todas as n-uplas ou "*vetores*" sobre S, $S^n = \{(x_1, \ldots, x_n) : x_i \in S\}$.

- Seja S um conjunto não vazio. Uma *relação* sobre S é qualquer função $R : S \times S \to S$. Diremos que R é uma *relação de equivalência* sobre S se R goza das seguintes propriedades:

  (a) xRx, para todo $x \in S$ (reflexividade).

  (b) Se xRy, então yRx, para todos $x, y \in S$. (simetria)

  (c) Se xRy e yRz, então xRz, para todos $x, y, z \in S$. (transitividade)

  Uma relação de equivalência pode ser pensada como uma generalização da relação de igualdade e é comum denotar R por $\sim$. Neste caso, se $x \sim y$, diremos que x *está relacionando* com y.

Seja $F[x]$ o conjunto de todas as *somas formais*

$$f(x) = a_m x^m + \cdots + a_1 x + a_0 \in F[x].$$

A expressão

$$f(x) = a_m x^m + \cdots + a_1 x + a_0 \text{ ou } f(x) = a_0 + a_1 x + \cdots + a_m x^m$$

chama-se *polinômio* sobre F e denotamos por $\text{co}_k(f) = a_k$, se $k = 0, \ldots, m$, e $\text{co}_k(f) = 0$, se $k > m$. Quando $a_m \neq 0$, diremos que $f(x)$ possui *grau* m e denotamos por $\partial(f) = m$ e convencionamos $\partial(0) = -\infty$. Seja

$$g(x) = b_0 + b_1 x + \cdots + b_m x^m \in F[x].$$

Diremos que $f(x)$ e $g(x)$ são *iguais* ($f = g$) se, e somente se, $\text{co}_k(f) = \text{co}_k(g)$, para cada $k = 0, 1, \ldots, m$. Dados

$$f(x) = a_0 + a_1 x + \cdots + a_m x^m, g(x) = b_0 + b_1 x + \cdots + b_m x^m \in F[x].$$

Definimos *adição de polinômio* como

$$(f + g)(x) = f(x) + g(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_m + b_m)x^m$$

e *multiplicação por escalar* como

$$(cf)(x) = (ca_0) + (ca_1)x + \cdots + (ca_m)x^m, \ \forall \ c \in F.$$

Como a adição e multiplicação por escalar de polinômios foram definidas em termos das operações de F, é natural que $F[x]$ herde algumas propriedades de F. Assim, é fácil verificar que $F[x]$ possui as seguintes propriedades:

1. $f(x)+(g(x)+h(x)) = (f(x)+g(x))+h(x)$, para todos $f(x), g(x), h(x) \in F[x]$.
2. $f(x) + g(x) = g(x) + f(x)$, para todos $f(x), g(x) \in F[x]$.
3. Existe $0 \in F[x]$ tal que $f(x) + 0 = f(x)$, para todos $f(x) \in F[x]$.
4. Para cada $f(x) \in F[x]$, existe $-f(x) \in F[x]$ tal que $f(x) + (-f(x)) = 0$, com $-f(x) = (-1)f(x)$.
5. $(a + b)f(x) = af(x) + bf(x)$, para todos $a, b \in F$ e $f(x) \in F[x]$.

6. $a(f(x) + g(x)) = af(x) + ag(x)$, para todos $f(x), g(x) \in F[x]$ e $a \in F$.
7. $a(bf(x)) = (ab)f(x)$, para todos $a, b \in F$ e $f(x) \in F[x]$.
8. $1 \cdot f(x) = f(x)$, para todo $f(x) \in F[x]$.

As propriedades de 1 à 8 constituem a definição de "*espaço vetorial*" sobre F.

Sejam

$$f(x) = a_0 + a_1x + \cdots + a_mx^m, g(x) = b_0 + b_1x + \cdots + b_nx^n \in F[x]$$

O *produto* de $f(x)$ por $g(x)$ é definido como

$$(f \cdot g)(x) = f(x)g(x) = c_0 + c_1x + \cdots + c_{m+n}x^{m+n},$$

em que

$$c_k = \sum_{i+j=k} a_ib_j = \sum_{i=0}^{k} a_ib_{k-i} \in F.$$

É fácil verificar que $F[x]$ munido com esse produto satisfaz as seguintes propriedades:

1. $f(x)(g(x)h(x)) = (f(x)g(x))h(x)$, para todos $f(x), g(x), h(x) \in F[x]$.
2. $f(x)(g(x) + h(x)) = f(x)g(x) + f(x)h(x)$, para todos $f(x), g(x), h(x) \in F[x]$.
3. $(f(x) + g(x))h(x) = f(x)h(x) + g(x)h(x)$, para todos $f(x), g(x), h(x) \in F[x]$.
4. $f(x)g(x) = g(x)f(x)$, para todos $f(x), g(x) \in F[x]$.
5. $c(f(x)g(x)) = f(x)(cg(x)) = (cf(x))g(x)$, para todos $f(x), g(x) \in F[x]$ e $c \in F$.

É pertinente observar que $F[x]$ possui todas as propriedades de F, exceto que $f(x) \in F[x]$, com $\partial(f) \geq 1$, não possui inverso multiplicativo. Por exemplo, se $|x| < 1$, então a fração, obtida por divisão usual,

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots = \sum_{n=0}^{\infty} x^n \notin F[x]. \tag{1.1}$$

Para cada

$$f(x) = c_0 + c_1x + \cdots + c_mx^m \in \mathbb{C}[x],$$

definimos o *polinômio conjugado* de $f(x)$ como

$$\bar{f}(x) = \bar{c}_0 + \bar{c}_1 x + \cdots + \bar{c}_m x^m \in \mathbb{C}[x].$$

Então

$$f(x)\bar{f}(x) = d_0 + d_1 x + \cdots + d_n x^n \in \mathbb{R}[x],$$

pois $d_j = \bar{d}_j$, com $j = 0, \ldots, n$. Portanto, não há perda de generalidade, em considerar polinômios apenas sobre $\mathbb{R}$.

Teorema 1.2 (Teorema Fundamental da Álgebra) *Cada polinômio sobre $\mathbb{R}$ possui uma raiz em $\mathbb{C}$.*

Prova. Confira a Bibliografia. ■

Sejam $f(x) = c_0 + \cdots + c_m x^m \in \mathbb{R}[x]$ um *polinômio mônico*, isto é, com $c_m = 1$, e $\alpha \in \mathbb{C}$ uma *raiz* de $f(x)$. Então é bem conhecido que $x - \alpha$ divide $f(x)$, ou seja, existe um $g(x) \in \mathbb{R}[x]$ tal que $f(x) = (x - \alpha)g(x)$. Assim, indutivamente,

$$f(x) = (x - \alpha_1) \cdots (x - \alpha_m), \tag{1.2}$$

onde $\alpha_1, \ldots, \alpha_m \in \mathbb{C}$ são as raízes, não necessariamente distintas, de $f(x)$. Note que se $\beta \in \mathbb{C} - \mathbb{R}$ é uma raiz complexa de $f(x)$, então

$$a_0 + a_1 \beta + \cdots + a_m \beta^m = 0.$$

Tomando a conjugação complexa desta equação, teremos

$$a_0 + a_1 \bar{\beta} + \cdots + a_m \bar{\beta}^m = 0.$$

Logo, $\bar{\beta}$ é também uma raiz de $f(x)$, ou seja, as raízes complexas de $f(x)$ ocorrem aos pares. Portanto,

$$f(x) = \prod_{j=1}^{r} (x - \alpha_j) \prod_{k=1}^{s} (x - \beta_k)(x - \bar{\beta}_k) \in \mathbb{C}[x],$$

onde $\alpha_j \in \mathbb{R}$, $\beta_k \in \mathbb{C} - \mathbb{R}$, $r, s \in \mathbb{N}$ e $r + 2s = m$. Neste caso,

$$f(x) = \prod_{j=1}^{r} (x - \alpha_j) \prod_{k=1}^{s} (x^2 - (\beta_k + \bar{\beta}_k)x + \beta_k \bar{\beta}_k) \in \mathbb{R}[x],$$

pois $\beta_k + \bar{\beta}_k, \beta_k\bar{\beta}_k \in \mathbb{R}$. Se $\beta = a + bi$, onde $a, b \in \mathbb{R}$, com $b \neq 0$, então

$$x^2 - (\beta + \bar{\beta})x + \beta\bar{\beta} = x^2 - 2ax + a^2 + b^2 \in \mathbb{R}[x]$$

é um *polinômio irredutível* sobre $\mathbb{R}$, pois $\Delta = -4b^2 < 0$. Portanto, podemos concluir que, qualquer polinômio mônico $f(x) = \sum_{i=0}^{m-1} c_i x^i + x^m$ sobre $\mathbb{R}[x]$ é fatorado em fatores lineares e quadráticos, a saber,

$$f(x) = \prod_{j=1}^{r}(x - \lambda_j)\prod_{k=1}^{s}(x^2 + b_k x + c_k) \in \mathbb{R}[x], \text{ onde } r, s \in \mathbb{N} \text{ e } r + 2s = m. \quad (1.3)$$

Exemplo 1.3 *Determine as raízes, se existirem sobre $\mathbb{R}$, de*

$$f(x) = x^3 - 7x^2 + 16x - 12 \in \mathbb{R}[x].$$

Solução. Um método bem conhecido para determinar as raízes do polinômio $f(x)$ é o *Dispositivo de Briot-Ruffine*[1]. Neste caso, as possíveis raízes racionais de $f(x)$ são os divisores de 12: $\pm 1, \pm 2, \pm 3, \pm 4, \pm 6, \pm 12$. Por exemplo, para $\lambda = 2$,

$$\begin{array}{c|cccc} 2 & 1 & -7 & 16 & -12 \\ \hline & 1 & 2\cdot 1 + (-7) = -5 & 2\cdot(-5) + 16 = 6 & 2\cdot 6 + (-12) = 0 \end{array}$$

Portanto, $\lambda = 2$ é uma raiz de $p(x)$. Agora, repete o mesmo Dispositivo com o poliômio $g(x) = x^2 - 5x + 6 \in \mathbb{R}[x]$, para obter as outras raízes 2 e 3 de $f(x)$. Neste caso, $f(x) = (x-2)^2(x-3)$. Consequentemente, $\lambda = 2$ possui *multiplicidade algébrica* igual a 2 e $\lambda = 3$ possui multiplicidade algébrica igual a 1. ■

Finalizaremos esta seção apresentando um fato conhecido como Relações de Girard[2]. Sejam $n, r \in \mathbb{N}$ e o conjunto de listas

$$Q_{r,n} = \{(k_1, \ldots, k_r) \in \mathbb{N}^r : 1 \leq k_1 < k_2 < \cdots < k_r \leq n\}.$$

Então existem $\binom{n}{r}$ tais conjuntos. A função $E_r : F^n \to F$ definida como

$$E_r(x_1, \ldots, x_n) = \sum_{(k_1,\ldots,k_r)\in Q_{r,n}} \prod_{i=1}^{r} x_{k_i}$$

[1] Charles Auguste Briot, 1817-1862, matemático francês; Paolo Ruffini, 1765-1822, médico e matemático italiano.

[2] Albert Girard, 1595-1632, matemático francês.

chama-se r-ésima *função simétrica elementar*. Por exemplo, se $n = 3$, então

$$\begin{array}{rcl} E_1(x_1, x_2, x_3) & = & x_1 + x_2 + x_3 \\ E_2(x_1, x_2, x_3) & = & x_1x_2 + x_1x_3 + x_2x_3 \\ E_3(x_1, x_2, x_3) & = & x_1x_2x_3. \end{array}$$

Observe que

$$E_r(x_1, \ldots, x_n) = \sum_{(k_1,\ldots,k_r)\in Q_{r,(n-1)}} \prod_{i=1}^{r} x_{k_i} + \sum_{(k_1,\ldots,k_{r-1})\in Q_{(r-1),(n-1)}} \prod_{i=1}^{r-1} x_{k_i} x_n.$$

Então obtemos a fórmula de recorrência

$$E_r(x_1, ..., x_n) = \begin{cases} E_r(x_1, ..., x_{n-1}) + x_n E_{r-1}(x_1, ..., x_{n-1}), & \text{se } r = 1, ..., n-1 \\ x_n E_{n-1}(x_1, \ldots, x_{n-1}), & \text{se } r = n. \end{cases}$$

Lema 1.4 *Seja $f(x) = (x - \alpha_1)\cdots(x - \alpha_n) \in F[x]$. Então*

$$f(x) = x^n - E_1(\alpha_1, \ldots, \alpha_n)x^{n-1} + \cdots + (-1)^n E_n(\alpha_1, \ldots, \alpha_n).$$

Prova. Vamos usar indução sobre o grau n. Se $n = 1$, nada há para ser provado. Suponhamos que o resultado seja válido para todo k, com $1 \leq k \leq n$ e $n > 1$. Então

$$\begin{array}{rcl} \prod_{i=1}^{n+1}(x - \alpha_i) & = & \prod_{i=1}^{n}(x - \alpha_i)(x - \alpha_{n+1}) \\ & = & \sum_{i=0}^{n}(-1)^i E_i(\alpha_1, \ldots, \alpha_n)x^{n-i}(x - \alpha_{n+1}) \\ & = & x^{n+1} - (E_1(\alpha_1, \ldots, \alpha_n) + \alpha_{n+1})x^n \\ & + & (-1)^r(E_r(\alpha_1, \ldots, \alpha_n) + \alpha_{n+1}E_{(r-1)}(\alpha_1, \ldots, \alpha_n))x^{n-r} \\ & + & \cdots + (-1)^{n+1}\alpha_{n+1}E_n(\alpha_1, \ldots, \alpha_{n+1}) \\ & = & x^{n+1} - E_1(\alpha_1, \ldots, \alpha_{n+1})x^n + \cdots \\ & + & (-1)^{n+1}\alpha_{n+1}E_n(\alpha_1, \ldots, \alpha_{n+1}), \end{array}$$

que é o resultado desejado. ■

É de grande importância prática e teórica observar que o Lema 1.4 nos fornece uma relação entre as raízes do polinômio e seus coeficientes. Mais precisamente, se $f(x) = c_0 + c_1x + \cdots + c_{n-1}x^{n-1} + x^n$, então

$$c_{n-k} = (-1)^k E_k(\alpha_1, \ldots, \alpha_n), \quad k = 1, \ldots, n.$$

Exemplo 1.5 *Resolver a equação* $x^3 - 6x^2 + 3x + 10 = 0$*, sabendo-se que a soma de duas raízes é* 1.

Solução. Sejam r_1, r_2 e r_3 as raízes da equação $x^3 - 6x^2 + 3x + 10 = 0$. Então obtemos as relações de Girard:

$$r_1 + r_2 + r_3 = 6, r_1r_2 + r_1r_3 + r_2r_3 = 3 \text{ e } r_1r_2r_3 = -10.$$

Podemos supor, sem perda de generalidade, que $r_1 + r_2 = 1$. Assim, $r_3 = 5$ implica que $5r_1r_2 = -10$ e $r_1 + r_2 = 1$. Portanto, $r_1 = -1.r_2 = 2$ e $r_3 = 5$. ■

Exercícios

1. (**Leis do Cancelamento**) Sejam $a, b, c \in F$, com $a \neq 0$. Mostre que

$$a + b = a + c \Rightarrow b = c \text{ e } ab = ac \Rightarrow b = c.$$

2. Sejam $z, w \in \mathbb{C}$. Mostre que:

   (a) $\overline{z + w} = \bar{z} + \bar{w}$, $\overline{zw} = \bar{z}\bar{w}$ e $\bar{\bar{z}} = z$.

   (b) $z + \bar{z} \in \mathbb{R}$ e $z - \bar{z} \in \mathbb{C} - \mathbb{R}$.

   (c) $|zw| = |z||w|$ e $|z + w| \leq |z| + |w|$.

3. Um subconjunto finito de $\mathbb{C}$ pode formar um corpo?

4. Mostre que o polinômio $f(x) = x^3 + ax^2 + bx + c \in \mathbb{R}[x]$ possui pelo menos uma raiz real. Conclua que qualquer polinômio de grau ímpar sobre $\mathbb{R}$ possui pelo menos uma raiz real.

5. Se $\{r_1, r_2, r_3\}$ é o conjunto solução da equação $2x^3 + 5x^2 + 8x + 11 = 0$ calcule $r_1^2 + r_2^2 + r_3^2$.

6. Calcular a soma e o produto das raízes da equação $2x^4 + 3x^3 + 4x^2 + 5x + 6 = 0$.

1.2 Matrizes

Ao trabalhar com um "sistema de equações lineares", somente os coeficientes e suas respectivas posições são importantes. Ademais, ao reduzir o sistema à forma escalonada, é essencial manter as equações cuidadosamente alinhadas. Assim, esses coeficientes podem ser e cientemente arrumados em uma disposição retangular chamada de “matriz”, termo devido a Sylvester[3] e Cayley[4].

Uma *matriz* $m \times n$ sobre F é um arranjo retangular com m linhas, n colunas e elementos ou entradas em F:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \text{ ou } \mathbf{A} = (a_{ij}).$$

Formalmente, uma matriz $m \times n$ é uma função

$$f : \{1, \ldots, m\} \times \{1, \ldots, n\} \to F$$

definida como $f(i, j) = a_{ij}$. A matriz $\mathbf{A} = (a_{ij})$ chama-se matriz de *ordem* ou *dimensão* mn. Em particular, se $m = n$, diremos que $\mathbf{A}$ é uma *matriz quadrada* de ordem n e os elementos

$$a_{11}, a_{22}, \ldots, a_{nn} \text{ e } a_{12}, a_{23}, \ldots, a_{(n-1)n} \ \left(a_{21}, a_{32}, \ldots, a_{n(n-1)}\right)$$

formam a *diagonal principal* e a *superdiagonal* (*subdiagonal*) de $\mathbf{A}$. Além disso, $a_{n1}, a_{(n-1)2}, \ldots, a_{1n}$ formam a *antidiagonal* ou a *diagonal secundária* de $\mathbf{A}$.

Sejam $\mathbf{A} = (a_{ij})$ uma matriz $m \times n$ e $\mathbf{B} = (b_{ij})$ uma matriz $r \times s$. Diremos que $\mathbf{A}$ e $\mathbf{B}$ são *iguais* ($\mathbf{A} = \mathbf{B}$) se, e somente se, $m = r$, $n = s$ e

$$a_{ij} = b_{ij}, \ \ i = 1, \ldots, m \text{ e } j = 1, \ldots, n.$$

Vamos denotar por $F^{m \times n}$ o conjunto de todas as matrizes de ordem mn sobre F. Em particular, $F^{1 \times n}$ o conjunto de todas as *matrizes linha* sobre F ou *vetores linhas* sobre F:

$$\mathbf{L}_i = \begin{pmatrix} a_{i1} & \cdots & a_{in} \end{pmatrix}$$

e $F^{m \times 1}$ o conjunto de todas as *matrizes colunas* sobre F ou *vetores colunas* sobre F:

$$\mathbf{C}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

[3] James Joseph Sylvester, 1814-1897, matemático inglês.
[4] Arthur Cayley, 1821-1895, matemático inglês.

Neste caso, obtemos as identificações: $F^n \longleftrightarrow F^{1\times n}$ e $F^m \longleftrightarrow F^{m\times 1}$, por exemplo, a função $T : F^n \to F^{1\times n}$ definida como

$$T(\mathbf{x}) = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} = \mathbf{L},$$

para todo $\mathbf{x} = (x_1, \ldots, x_n) \in F^n$, é bijetora e preserva as operações.

Dado $\mathbf{A} = (a_{ij}) \in F^{n\times n}$. Diremos que $\mathbf{A}$ é uma *matriz diagonal* se

$$a_{ij} = 0, \ \ i \neq j.$$

Usaremos a notação $\mathbf{D} = \operatorname{diag}(d_1, \ldots, d_n)$ para representar a matriz diagonal $\mathbf{A}$, com $a_{ii} = d_i$ e $i = 1, \ldots, n$. Quando $a_{ii} = d_i = d$, $i = 1, \ldots, n$, diremos que a matriz diagonal $\mathbf{D}$ é uma *matriz escalar*. Em particular, diremos que a matriz diagonal $\mathbf{D}$ é uma *matriz identidade* ou *matriz unidade* se

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

e denotamos por $\mathbf{I}_n = (\delta_{ij}) = \operatorname{diag}(1, \ldots, 1)$ ou simplesmente $\mathbf{I}$, em que o símbolo δ_{ij} chama-se *delta de Kronecker*[5]. Neste caso, $\mathbf{D} = (a_{ij}\delta_{ij})$.

A matriz $\mathbf{A} = (a_{ij}) \in F^{m\times n}$, com $a_{ij} = 0$, $i = 1, \ldots, m$ e $j = 1, \ldots, n$, chama-se *matriz nula* e denotamos por $\mathbf{O}_{m\times n}$ ou simplesmente $\mathbf{O}$.

Seja $\mathbf{A} \in F^{m\times n}$. Para cada $r \in \{1, \ldots, \min\{m, n\}\}$, uma *submatriz* de ordem r de $\mathbf{A}$ é qualquer matriz de ordem r obtida eliminando-se $(m - r)$ linhas e $(n - r)$ colunas de $\mathbf{A}$. Assim, um modo útil de escrever uma matriz é sob a forma de *blocos* ou *particionada*:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \end{pmatrix} \ \text{ ou } \ \mathbf{A} = \left(\begin{array}{c|c|c} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \end{array}\right),$$

cujas entradas são submatrizes de $\mathbf{A}$. Por exemplo,

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\operatorname{sen}\theta \\ 0 & \operatorname{sen}\theta & \cos\theta \end{pmatrix} = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & \cos\theta & -\operatorname{sen}\theta \\ 0 & \operatorname{sen}\theta & \cos\theta \end{array}\right), \ \ \forall \ \ \theta \in \mathbb{R}.$$

[5]Leopold Kronecker, 1823-1891, matemático alemão.

Uma matriz na forma de blocos

$$\mathbf{A} = \begin{pmatrix} \mathbf{B} & \mathbf{O} \\ \mathbf{O} & \mathbf{C} \end{pmatrix} \text{ ou } \mathbf{A} = \left(\begin{array}{c|c} \mathbf{B} & \mathbf{O} \\ \hline \mathbf{O} & \mathbf{C} \end{array}\right)$$

chama-se *soma direta* das submatrizes $\mathbf{B}$ e $\mathbf{C}$ e denotamos por $\mathbf{A} = \mathbf{B} \oplus \mathbf{C}$. Neste caso,

$$a_{ij} = \begin{cases} b_{ij}, & \text{se } i, j = 1, \ldots, k \\ c_{(i-k)(j-k)}, & \text{se } i, j = k+1, \ldots, n \\ 0, & \text{caso contrário.} \end{cases}$$

Dados $\mathbf{A} = (a_{ij}), \mathbf{B} = (b_{ij}) \in F^{m \times n}$. Definimos *adição de matrizes* como

$$\mathbf{A} + \mathbf{B} = (c_{ij}), \;\; c_{ij} = a_{ij} + b_{ij}, \;\; i = 1, \ldots, m \text{ e } j = 1, \ldots, n$$

e *multiplicação por escalar* como

$$c\mathbf{A} = (c_{ij}), \;\; c_{ij} = ca_{ij}, \;\; i = 1, \ldots, m \text{ e } j = 1, \ldots, n, \;\; \forall \; c \in F,$$

É pertinente lembrar que: como a adição e multiplicação por escalar de matrizes foram definidas em termos das operações de F, é natural que $F^{m \times n}$ herde algumas propriedades de F. Assim, é fácil verificar que $F^{m \times n}$ possui as seguintes propriedades:

1. $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$, para todos $\mathbf{A}, \mathbf{B}, \mathbf{C} \in F^{m \times n}$.

2. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$, para todos $\mathbf{A}, \mathbf{B} \in F^{m \times n}$.

3. Existe $\mathbf{O} \in F^{m \times n}$ tal que $\mathbf{A} + \mathbf{O} = \mathbf{A}$, para todos $\mathbf{A} \in F^{m \times n}$.

4. Para cada $\mathbf{A} \in F^{m \times n}$, existe $-\mathbf{A} \in F^{m \times n}$ tal que $\mathbf{A} + (-\mathbf{A}) = \mathbf{O}$, com $-\mathbf{A} = (-1)\mathbf{A} = (-a_{ij})$. Neste caso, definimos a *diferença de matrizes*

$$\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B}).$$

5. $(a + b)\mathbf{A} = a\mathbf{A} + b\mathbf{A}$, para todos $a, b \in F$ e $\mathbf{A} \in F^{m \times n}$.

6. $a(\mathbf{A} + \mathbf{B}) = a\mathbf{A} + a\mathbf{B}$, para todos $\mathbf{A}, \mathbf{B} \in F^{m \times n}$ e $a \in F$.

7. $a(b\mathbf{A}) = (ab)\mathbf{A}$, para todos $a, b \in F$ e $\mathbf{A} \in F^{m \times n}$.

8. $1 \cdot \mathbf{A} = \mathbf{A}$, para todo $\mathbf{A} \in F^{m \times n}$.

As propriedades de 1 à 8 constituem a definição de "*espaço vetorial*" sobre F.

Sejam $\mathbf{A} = (a_{ij}) \in F^{m\times n}$ e $\mathbf{B} = (b_{ij}) \in F^{n\times p}$. O *produto* de $\mathbf{A}$ por $\mathbf{B}$ é definido como

$$\mathbf{AB} = (c_{ij}), \quad c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}, \quad i = 1, \ldots, m \text{ e } j = 1, \ldots, p.$$

Note que $\mathbf{AB} \in F^{m\times p}$. O produto de matrizes possui as seguintes propriedades:

1. $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$, para todo $\mathbf{A} \in F^{m\times n}$, $\mathbf{B} \in F^{n\times p}$ e $\mathbf{C} \in F^{p\times q}$.

2. $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$, para todos $\mathbf{A}, \mathbf{B} \in F^{m\times n}$ e $\mathbf{C} \in F^{n\times p}$.

3. $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$, para todo $\mathbf{A} \in F^{m\times n}$ e $\mathbf{B}, \mathbf{C} \in F^{n\times p}$.

4. $c(\mathbf{AB}) = \mathbf{A}(c\mathbf{B}) = (c\mathbf{A})\mathbf{B}$, para todos $\mathbf{A}, \mathbf{B} \in F^{n\times n}$ e $c \in F$.

5. Existe $\mathbf{I} \in F^{n\times n}$ tal que $\mathbf{A} \cdot \mathbf{I} = \mathbf{I} \cdot \mathbf{A} = \mathbf{A}$, para todo $\mathbf{A} \in F^{n\times n}$.

6. $\mathbf{AO} = \mathbf{O}$ e $\mathbf{OB} = \mathbf{O}$, para todos $\mathbf{A}, \mathbf{O} \in F^{m\times n}$ e $\mathbf{B}, \mathbf{O} \in F^{n\times p}$.

7. Se $\mathbf{A} = (a_{ij}) \in F^{n\times n}$ e $\mathbf{D} = \operatorname{diag}(d_1, \ldots, d_n) \in F^{n\times n}$, então $\mathbf{DA} = (d_i a_{ij})$, ou seja, multiplica cada linha $\mathbf{L}_i$ de $\mathbf{A}$ por uma constante d_i.

8. Se $\mathbf{A} \in F^{m\times n}$ e $\mathbf{x} = (x_1, \ldots, x_m) \in F^m$, então

$$\mathbf{xA} = x_1\mathbf{L}_1 + \cdots + x_m\mathbf{L}_m$$

é uma *combinação linear* das linhas $\mathbf{L}_i$ de $\mathbf{A}$. Em particular, se $\mathbf{e}_j = (\delta_{ij}) \in F^m$ é a matriz com 1 na j-ésima coluna e as demais zeros, então $\mathbf{e}_j\mathbf{A} = \mathbf{L}_j$. O vetor $\mathbf{e}_j$ chama-se *vetor elementar* ou *canônico*.

9. Se $\mathbf{A} \in F^{m\times n}$ e $\mathbf{X} = (x_i) \in F^{n\times 1}$, então

$$\mathbf{AX} = x_1\mathbf{C}_1 + \cdots + x_n\mathbf{C}_n$$

é uma combinação linear das colunas $\mathbf{C}_j$ de $\mathbf{A}$. Em particular, se $\mathbf{E}_i = (\delta_{ij}) \in F^{n\times 1}$ é a matriz com 1 na i-ésima linha e as demais zeros, então $\mathbf{AE}_i = \mathbf{C}_i$.

Como ilustração provaremos os itens (1) e (3): sejam $\mathbf{A} = (a_{ij})$, $\mathbf{B} = (b_{ij})$ e $\mathbf{C} = (c_{ij})$. Então

$$(\mathbf{AB})\mathbf{C} = \left(\sum_{l=1}^{p}\left(\sum_{k=1}^{n} a_{ik}b_{kl}\right)c_{lj}\right) = \left(\sum_{k=1}^{n} a_{ik}\left(\sum_{l=1}^{p} b_{kl}c_{lj}\right)\right) = \mathbf{A}(\mathbf{BC})$$

e

$$\begin{aligned}\mathbf{A}(\mathbf{B}+\mathbf{C}) &= (\textstyle\sum_{k=1}^{n} a_{ik}(b_{kj}+c_{kj})) = (\sum_{k=1}^{n}(a_{ik}b_{kj}+a_{ik}c_{kj}))\\ &= (\textstyle\sum_{k=1}^{n} a_{ik}b_{kj}) + (\sum_{k=1}^{n} a_{ik}c_{kj}) = \mathbf{AB}+\mathbf{AC}.\end{aligned}$$

É pertinente notar que podemos escrever o produto $\mathbf{AB}$ sob a forma de blocos:

$$\begin{aligned}\mathbf{AB} &= \mathbf{A}\begin{pmatrix} \mathbf{C}_1 & \cdots & \mathbf{C}_p \end{pmatrix} = \begin{pmatrix} \mathbf{AC}_1 & \cdots & \mathbf{AC}_p \end{pmatrix}\\ &= \begin{pmatrix} \mathbf{L}_1 \\ \vdots \\ \mathbf{L}_n \end{pmatrix}\mathbf{B} = \begin{pmatrix} \mathbf{L}_1\mathbf{B} \\ \vdots \\ \mathbf{L}_n\mathbf{B} \end{pmatrix},\end{aligned}$$

ou ainda, $\mathbf{AB} = \sum_{k=1}^{n} \mathbf{C}_k\mathbf{L}_k$. Observe que as colunas de $\mathbf{AB}$ são combinações lineares das colunas de $\mathbf{A}$ e as linhas de $\mathbf{AB}$ são combinações lineares das linhas de $\mathbf{B}$. Além disso, se $\mathbf{A} \in F^{n\times n}$ então. indutivamente, obtemos

$$\mathbf{A}^{k+1} = \mathbf{A}^k\mathbf{A},\ \ \mathbf{A}^k\mathbf{A}^m = \mathbf{A}^{k+m} \ \text{ e } \ \left(\mathbf{A}^k\right)^m = \mathbf{A}^{km},\ \ \forall\ k, m \in \mathbb{N}.$$

Mais geralmente, sejam

$$f(x) = c_m x^m + \cdots + c_1 x + c_0 \in F[x]$$

um polinômio de grau $\partial(f) = m$ sobre F e $\mathbf{A} \in F^{n\times n}$. Então $f(\mathbf{A})$ é uma matriz de ordem n sobre F definida como

$$f(\mathbf{A}) = c_m\mathbf{A}^m + \cdots + c_1\mathbf{A} + c_0\mathbf{I} \in F^{n\times n},$$

a qual chama-se *função polinomial* em $\mathbf{A}$. Note que a matriz $f(\mathbf{A})$ foi obtida de $f(x)$ substituindo-se a variável x pela matriz $\mathbf{A}$ e o escalar c_0 pela matriz escalar $c_0\mathbf{I}$. Diremos que $f(x)$ é um *polinômio anulador* de $\mathbf{A}$ se $f(\mathbf{A}) = \mathbf{O}$ e $\mathbf{A}$ é um *zero* de $f(x)$. Por exemplo, se

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix} \in F^{2\times 2} \ \text{ e } \ f(x) = x^2 - 2x - 3 \in F[x],$$

então
$$\begin{aligned} f(\mathbf{A}) &= \mathbf{A}^2 - 2\mathbf{A} - 3\mathbf{I} \\ &= \begin{pmatrix} 5 & 2 \\ 8 & 5 \end{pmatrix} - 2\begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix} - 3\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{O}. \end{aligned}$$

É fácil verificar que $\mathbf{A}f(\mathbf{A}) = f(\mathbf{A})\mathbf{A}$, para todo $f(x) \in F[x]$. Mais geralmente,

$$f(\mathbf{A})g(\mathbf{A}) = g(\mathbf{A})f(\mathbf{A}), \ \ \forall \ f(x), g(x) \in F[x].$$

Observe que as operações de adição, multiplicação por escalar e produto de matrizes estão relacionadas pelas propriedades (2), (3) e (4). Neste caso, diremos que $F^{n\times n}$ munido com essas operações é uma *álgebra "linear"* sobre F.

Vale observar que $F^{n\times n}$ não herda todas as propriedade de F. Por exemplo, se

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \ \mathbf{B} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \ \text{e} \ \mathbf{C} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$$

então
$$\mathbf{AB} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \ \mathbf{AC} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \ \text{e} \ \mathbf{BA} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Assim, $\mathbf{AB} \neq \mathbf{BA}$, $\mathbf{BA} = \mathbf{O}$, com $\mathbf{A} \neq \mathbf{O}$ e $\mathbf{B} \neq \mathbf{O}$, e $\mathbf{AB} = \mathbf{AC}$ não implica que $\mathbf{B} = \mathbf{C}$, com $\mathbf{A} \neq \mathbf{O}$. Além disso, o polinômio

$$f(\mathbf{A}) = (\mathbf{A} + \mathbf{I})(\mathbf{A} - 3\mathbf{I}) = \mathbf{A}^2 - 2\mathbf{A} - 3\mathbf{I}, \ \ \mathbf{A} \in \mathbb{C}^{2\times 2},$$

ao contrário do polinômio $f(x) = (x+1)(x-3)$, possui para zeros, além das matrizes escalares $-\mathbf{I}$ e $3\mathbf{I}$, as matrizes não escalares

$$\mathbf{A} = \begin{pmatrix} 3 & a \\ 0 & -1 \end{pmatrix},$$

para todo $a \in \mathbb{R}$, ou seja, não vale o Teorema Fundamental da Álgebra sobre à álgebra linear $\mathbb{C}^{2\times 2}$.

Seja $\mathbf{A} = (a_{ij}) \in F^{m\times n}$. A *matriz transposta* de $\mathbf{A}$ é a matriz obtida de $\mathbf{A}$ refletido-se as entradas em relação à diagonal principal de $\mathbf{A}$, ou seja,

$$\mathbf{A}^t = (b_{ij}), \ \ b_{ij} = a_{ji}, \ \ i = 1, \ldots, n \ \text{e} \ j = 1, \ldots, m.$$

A transposta de matrizes possui as seguintes propriedades:

1. $(\mathbf{A}+\mathbf{B})^t = \mathbf{A}^t + \mathbf{B}^t$, para todos $\mathbf{A}, \mathbf{B} \in F^{m\times n}$.
2. $(a\mathbf{A})^t = a\mathbf{A}^t$, para todo $\mathbf{A} \in F^{m\times n}$ e $a \in F$.
3. $(\mathbf{A}^t)^t = \mathbf{A}$, para todo $\mathbf{A} \in F^{m\times n}$.
4. $(\mathbf{A}\mathbf{B})^t = \mathbf{B}^t\mathbf{A}^t$, para todos $\mathbf{A}, \mathbf{B} \in F^{n\times n}$.
5. A função $T : F^{n\times n} \to F^{n\times n}$ definida como $T(\mathbf{A}) = \mathbf{A}^t$ é bijetora e preserva as propriedades (1) e (2). Além disso, $T^2 = I$.

A *transposta conjugada* ou *adjunta* de $\mathbf{A}$ é a matriz $\mathbf{A}^* = \bar{\mathbf{A}}^t$, em que $\bar{\mathbf{A}} = (\bar{a}_{ij})$. Observe, para cada matriz $\mathbf{A} \in F^{m\times n}$ fixada, que

$$\mathbf{A} = \frac{1}{2}(\mathbf{A}+\mathbf{A}^*) + \frac{1}{2}(\mathbf{A}-\mathbf{A}^*),$$

desde que $2 \neq 0$ em F. Além disso, a *matriz antitransposta* de $\mathbf{A} = (a_{ij}) \in F^{n\times n}$ é a matriz obtida de $\mathbf{A}$ refletido-se as entradas em relação à antidiagonal de $\mathbf{A}$, ou seja,

$$\mathbf{A}^s = (b_{ij}), \quad b_{ij} = a_{(n-(j-1))(n-(i-1))}, \quad i,j = 1,\ldots,n.$$

Seja $\mathbf{E}_{ij} = (e_{pq}) \in F^{m\times n}$ a matriz definida como

$$e_{pq} = \delta_{ip}\delta_{qj} = \begin{cases} 1, & \text{se } (p,q) = (i,j) \\ 0, & \text{se } (p,q) \neq (i,j), \end{cases}$$

isto é, $\mathbf{E}_{ij}$ é a matriz cuja (i,j)-ésima entrada é igual a 1 e as demais zeros, ou ainda, $e_{ij} = 1$ e $e_{pq} = 0$, quando $i \neq p$ ou $j \neq q$. Se $m = n$, diremos que $\mathbf{E}_{ij}$ é uma *matriz unitária fundamental*. Por exemplo, para $m = n = 2$, obtemos

$$\mathbf{E}_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \ \mathbf{E}_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \ \mathbf{E}_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \text{ e } \mathbf{E}_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Proposição 1.6 *Sejam* $\mathbf{A} = (a_{ij}), \mathbf{E}_{ij} = (e_{pq}) \in F^{n\times n}$. *Então*:

1. $\mathbf{E}_{ij} = \mathbf{E}_{kl}$ *se, e somente se,* $(i,j) = (k,l)$.
2. $\mathbf{E}_{ij} = \mathbf{E}_i\mathbf{E}_j^t$ *e* $\sum_{i=1}^n \mathbf{E}_{ii} = \mathbf{I}_n$.
3. $\mathbf{E}_{ij}\mathbf{E}_{kl} = \delta_{jk}\mathbf{E}_{il}$. *Em particular,* $\mathbf{E}_{ii}^2 = \mathbf{E}_{ii}$ *e* $\mathbf{E}_{ij}^2 = \mathbf{0}$ *se* $i \neq j$. *Isto nos fornece uma tabela de multiplicação para à álgebra linear* $F^{n\times n}$.

4. *$\mathbf{E}_{ij}\mathbf{A} = \sum_{k=1}^{n} a_{jk}\mathbf{E}_{ik}$, isto é, $\mathbf{E}_{ij}\mathbf{A}$ é a matriz cuja i-ésima linha é igual a j-ésima linha da matriz $\mathbf{A}$ e as demais zeros.*

5. *$\mathbf{A}\mathbf{E}_{pq} = \sum_{i=1}^{n} a_{ip}\mathbf{E}_{iq}$, isto é, $\mathbf{A}\mathbf{E}_{pq}$ é a matriz cuja q-ésima coluna é igual a p-ésima coluna da matriz $\mathbf{A}$ e as demais zeros.*

6. *$\mathbf{E}_{pq}\mathbf{A}\mathbf{E}_{rs} = a_{qr}\mathbf{E}_{ps}$, isto é, $\mathbf{E}_{pq}\mathbf{A}\mathbf{E}_{rs}$ é a matriz cuja (p, s)-ésima entrada é igual a a_{qr} e as demais zeros.*

7. *Mostraremos no capítulo 3 que:*
$$\mathbf{A} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\mathbf{E}_{ij}.$$

Prova. Vamos provar apenas os itens (4) e (6): (4) Pondo $\mathbf{E}_{ij}\mathbf{A} = (c_{rs})$, obtemos
$$c_{rs} = \sum_{k=1}^{n} e_{rk}a_{ks} = \sum_{k=1}^{n} \delta_{ir}\delta_{kj}a_{ks} = \delta_{ir}a_{js}.$$
Assim, $c_{is} = a_{js}$ e $c_{rs} = 0$ se $i \neq r$.

(6) Note que
$$\begin{aligned}\mathbf{E}_{pq}\mathbf{A}\mathbf{E}_{rs} &= \mathbf{E}_{pq}\left(\textstyle\sum_{i=1}^{m} a_{ir}\mathbf{E}_{is}\right) = \sum_{i=1}^{m} a_{ir}\mathbf{E}_{pq}\mathbf{E}_{is} = \sum_{i=1}^{m} a_{ir}\delta_{qi}\mathbf{E}_{ps}\\ &= a_{qr}\mathbf{E}_{ps},\end{aligned}$$
que é o resultado desejado. ■

Já vimos, para cada $a \in F^{\times}$, que existia um único $a^{-1} \in F$ tal que $a \cdot a^{-1} = a^{-1} \cdot a = 1$. É fácil verificar que a função $T : F \to F^{n\times n}$ definida como $T(a) = a\mathbf{I}$ é injetora e preservas as operações. Assim, obtemos uma equação análoga para as matrizes escalares, a saber, $(a\mathbf{I}) \cdot (a^{-1}\mathbf{I}) = (a^{-1}\mathbf{I}) \cdot (a\mathbf{I}) = \mathbf{I}$. Isto motiva a seguinte definição.

Seja $\mathbf{A} \in F^{n\times n}$. Diremos que $\mathbf{A}$ é uma *matriz invertível* ou uma *matriz não singular* se existir um $\mathbf{X} \in F^{n\times n}$ tal que $\mathbf{A}\mathbf{X} = \mathbf{X}\mathbf{A} = \mathbf{I}$. Caso contrário, diremos que $\mathbf{A}$ é uma *matriz não invertível* ou uma *matriz singular*. A matriz $\mathbf{X}$ chama-se *matriz inversa* de $\mathbf{A}$. Observe que se a inversa existir, é única, pois se $\mathbf{X}$ e $\mathbf{Y}$ são duas inversas de $\mathbf{A}$, então
$$\mathbf{X} = \mathbf{X}\mathbf{I} = \mathbf{X}(\mathbf{A}\mathbf{Y}) = (\mathbf{X}\mathbf{A})\mathbf{Y} = \mathbf{I}\mathbf{Y} = \mathbf{Y}.$$

Denotamos a matriz inversa de $\mathbf{A}$ por $\mathbf{X} = \mathbf{A}^{-1}$. Por exemplo, se $d_i \neq 0$, para cada $i = 1, \ldots, n$, então as matrizes

$$\mathbf{D} = \operatorname{diag}(d_1, \ldots, d_n) \in F^{n\times n} \text{ e } \mathbf{A} = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} \in F^{2\times 2}$$

são não singulares, pois

$$\mathbf{D}^{-1} = \operatorname{diag}(d_1^{-1}, \ldots, d_n^{-1}) \text{ e } \mathbf{A}^{-1} = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}$$

Exemplo 1.7 *Seja* $\mathbf{A} \in F^{n\times n}$.

1. *Mostre que se* $\mathbf{A}$ *possui uma linha nula, então* $\mathbf{A}$ *é singular.*
2. *Mostre que se* $\mathbf{A}$ *possui duas linhas proporcionais, então* $\mathbf{A}$ *é singular.*

Solução. (1) Suponhamos, por absurdo, que $\mathbf{A}$ seja não singular. Então existe $\mathbf{X} \in F^{n\times n}$ tal que $\mathbf{AX} = \mathbf{XA} = \mathbf{I}$. Se $\mathbf{L}_i = \mathbf{O}$ é uma linha nula de $\mathbf{A}$, então

$$\mathbf{I} = \mathbf{AX} = \begin{pmatrix} \mathbf{L}_1\mathbf{X} & \cdots & \mathbf{L}_{i-1}\mathbf{X} & \mathbf{O} & \mathbf{L}_{i+1}\mathbf{X} & \cdots & \mathbf{L}_n\mathbf{X} \end{pmatrix}^t,$$

o que é uma contradição, pois $\mathbf{I}$ não possui linha nula.

(2) Suponhamos, por absurdo, que $\mathbf{A}$ seja não singular. Então existe um $\mathbf{X} \in F^{n\times n}$ tal que $\mathbf{AX} = \mathbf{XA} = \mathbf{I}$. Se $\mathbf{L}_j = c\mathbf{L}_i$, com $i \neq j$, são duas linhas proporcionais de $\mathbf{A}$, então

$$\mathbf{I} = \mathbf{AX} = \begin{pmatrix} \mathbf{L}_1\mathbf{X} & \cdots & \mathbf{L}_{j-1}\mathbf{X} & c(\mathbf{L}_i\mathbf{X}) & \mathbf{L}_{j+1}\mathbf{X} & \cdots & \mathbf{L}_n\mathbf{X} \end{pmatrix}^t,$$

o que é uma contradição, pois $\mathbf{I}$ não possui linhas proporcionais. ■

Teorema 1.8 *Sejam* $\mathbf{A}, \mathbf{B}, \mathbf{C} \in F^{n\times n}$.

1. *Se* $\mathbf{A}$ *for não singular, então* $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$ *e* $(\mathbf{A}^{-1})^t = (\mathbf{A}^t)^{-1}$.
2. *Se* $\mathbf{A}$ *for não singular, então a equação matricial* $\mathbf{AX} = \mathbf{B}$ *possui uma única solução em* $F^{n\times n}$.
3. *Se* $\mathbf{A}$ *e* $\mathbf{B}$ *são não singulares, então* $\mathbf{AB}$ *é não singular e* $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.
4. *Se* $\mathbf{AB} = \mathbf{O}$, *então ocorre exatamente uma das condições*: $\mathbf{A} = \mathbf{O}$ *ou* $\mathbf{B} = \mathbf{O}$ *ou* $\mathbf{A}$ *e* $\mathbf{B}$ *são singulares. Quando* $\mathbf{A} \neq \mathbf{O}$ *e* $\mathbf{B} \neq \mathbf{O}$, *diremos que* $\mathbf{A}$ *é um* divisor de zero.

5. *Se* $\mathbf{A}$ *for não singular, então* $\mathbf{AB} = \mathbf{AC}$ *implica que* $\mathbf{B} = \mathbf{C}$ *e* $\mathbf{BA} = \mathbf{CA}$ *implica que* $\mathbf{B} = \mathbf{C}$. *Em particular,* $\mathbf{AB} = \mathbf{O}$ *implica que* $\mathbf{B} = \mathbf{O}$.

6. *Se* $\mathbf{A}$ *for singular, então existe pelo menos uma* $\mathbf{B} \neq \mathbf{O}$ *tal que* $\mathbf{AB} = \mathbf{O}$.

Prova. Vamos provar apenas os itens (2), (3) e (4):

(2) É claro que $\mathbf{A}^{-1}\mathbf{B}$ é uma solução. Por outro lado, se $\mathbf{X} \in F^{n\times n}$ for uma solução, então

$$\mathbf{X} = \mathbf{IX} = \mathbf{A}^{-1}(\mathbf{AX}) = \mathbf{A}^{-1}\mathbf{B}.$$

Assim, qualquer solução é da forma $\mathbf{A}^{-1}\mathbf{B}$. Sejam $\mathbf{X}$ e $\mathbf{Y}$ duas soluções, então

$$\mathbf{X} = \mathbf{IX} = \mathbf{A}^{-1}(\mathbf{AX}) = \mathbf{A}^{-1}\mathbf{B} = \mathbf{Y},$$

ou seja, a solução é única.

(3) Note que

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

De modo análogo, $(\mathbf{B}^{-1}\mathbf{A}^{-1})(\mathbf{AB}) = \mathbf{I}$. Assim, pela unicidade da matriz inversa, $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ e $\mathbf{AB}$ é não singular.

(4) Observe, pelo item (3), que $\mathbf{A}$ é não singular ou $\mathbf{B}$ é não singular ou $\mathbf{A}$ e $\mathbf{B}$ são singulares. Se $\mathbf{A}$ for não singular, então

$$\mathbf{B} = \mathbf{IB} = (\mathbf{A}^{-1}\mathbf{A})\mathbf{B} = \mathbf{A}^{-1}(\mathbf{AB}) = \mathbf{A}^{-1}\mathbf{O} = \mathbf{O}.$$

De modo análogo, se $\mathbf{B}$ for não singular, então $\mathbf{A} = \mathbf{O}$. ■

Corolário 1.9 *Seja* $\mathbf{A} \in F^{n\times n}$. *Então* $\mathbf{A}$ *possui uma inversa à direita se, e somente se, ela for não singular.*

Prova. Suponhamos que exista um $\mathbf{B} \in F^{n\times n}$ tal que $\mathbf{AB} = \mathbf{I}$ ou $\mathbf{B}^t\mathbf{A}^t = \mathbf{I}$. Note, pelo item (4) do Teorema 1.8, que $\mathbf{A}^t\mathbf{X} = \mathbf{O}$ implica que $\mathbf{X} = \mathbf{O}$, pois $\mathbf{X} = \mathbf{I}\cdot\mathbf{X} = (\mathbf{B}^t\mathbf{A}^t)\mathbf{X} = \mathbf{B}^t(\mathbf{A}^t\mathbf{X}) = \mathbf{B}^t\cdot\mathbf{O} = \mathbf{O}$. Assim, $\mathbf{B}$ é única. Como $\mathbf{ABA} = \mathbf{A}$ e

$$\mathbf{A}(\mathbf{B} + \mathbf{BA} - \mathbf{I}) = \mathbf{AB} + \mathbf{ABA} - \mathbf{A} = \mathbf{I}$$

temos que $\mathbf{B} + \mathbf{BA} - \mathbf{I} = \mathbf{B}$. Portanto, $\mathbf{BA} = \mathbf{I}$. A recíproca é clara. ■

Para apresentar um método de obter a inversa de uma matriz definimos um tipo bastante útil de matrizes. Uma *matriz elementar* sobre F é uma matriz quadrada de ordem n de um dos seguintes tipos:

$$\begin{array}{lcl}\mathbf{P}_{ij} & = & \mathbf{I}_n - \mathbf{E}_{ii} + \mathbf{E}_{ij} - \mathbf{E}_{jj} + \mathbf{E}_{ji}, \\ \mathbf{D}_i(c) & = & \mathbf{I}_n + (c-1)\mathbf{E}_{ii}, \quad c \neq 0, \\ \mathbf{T}_{ij}(c) & = & \mathbf{I}_n + c\mathbf{E}_{ij}, \quad i \neq j \text{ e } c \neq 0.\end{array}$$

$\mathbf{P}_{ij}$ chama-se *involução*, $\mathbf{D}_i(c)$ chama-se *alongamento* e $\mathbf{T}_{ij}(c)$ chama-se *cisalhamento*. Observe que as matrizes elementares são não singulares. De fato. Se $i \neq j$, então

$$\mathbf{T}_{ij}(c)\mathbf{T}_{ij}(-c) = (\mathbf{I}_n + c\mathbf{E}_{ij})(\mathbf{I}_n - c\mathbf{E}_{ij}) = \mathbf{I}_n \Rightarrow \mathbf{T}_{ij}^{-1}(c) = \mathbf{T}_{ij}(-c),$$

de modo análogo, $\mathbf{P}_{ij}^{-1} = \mathbf{P}_{ij}$ e $\mathbf{D}_i^{-1}(c) = \mathbf{D}_i(c^{-1})$. Além disso,

1. $\mathbf{P}_{ij}^t = \mathbf{P}_{ij}$, $\mathbf{D}_i^t(c) = \mathbf{D}_i(c)$ e $\mathbf{T}_{ij}^t(c) = \mathbf{T}_{ji}(c)$.
2. $\mathbf{D}_i(c)\mathbf{D}_i(d) = \mathbf{D}_i(cd)$, para todos $c, d \in F^*$.
3. $\mathbf{T}_{ij}(c)\mathbf{T}_{ij}(d) = \mathbf{T}_{ij}(c+d)$, para todos $c, d \in F^*$.

Por exemplo,

$$\mathbf{P}_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = (\mathbf{E}_2 \;\; \mathbf{E}_1 \;\; \mathbf{E}_3)^t$$

é a matriz obtida de $\mathbf{I}_3$ permutando-se ou trocando-se as linhas $\mathbf{L}_1$ e $\mathbf{L}_2$,

$$\mathbf{D}_2(c) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{pmatrix} = (\mathbf{E}_1 \;\; c\mathbf{E}_2 \;\; \mathbf{E}_3)^t$$

é a matriz obtida de $\mathbf{I}_3$ multiplicando-se a linha $\mathbf{L}_2$ por c e

$$\mathbf{T}_{21}(c) = \begin{pmatrix} 1 & 0 & 0 \\ c & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = (\mathbf{E}_1 \;\; c\mathbf{E}_1 + \mathbf{E}_2 \;\; \mathbf{E}_3)^t$$

é a matriz obtida de $\mathbf{I}_3$ multiplicando-se a linha $\mathbf{L}_1$ por c e somando-se com a linha $\mathbf{L}_2$, ou seja, é a substituição da linha $\mathbf{L}_2$ pela linha $\mathbf{L}_2$ mais c vezes a linha $\mathbf{L}_1$.

Proposição 1.10 *Seja* $\mathbf{A} \in F^{m \times n}$.

1. $\mathbf{P}_{ij}\mathbf{A}$ *é a matriz obtida de* $\mathbf{A}$ *trocando-se as linhas* $\mathbf{L}_i$ *e* $\mathbf{L}_j$.

2. $\mathbf{D}_i(c)\mathbf{A}$ *é a matriz obtida de* $\mathbf{A}$ *multiplicando-se a linha* $\mathbf{L}_i$ *por* c*.*

3. $\mathbf{T}_{ij}(c)\mathbf{A}$ *é a matriz obtida de* $\mathbf{A}$ *multiplicando-se a linha* $\mathbf{L}_j$ *por* c *e somando-se com a linha* $\mathbf{L}_i$*.*

Prova. Vamos provar apenas o item (3). Como $c\mathbf{E}_{ij}\mathbf{A} = (0 \cdots c\mathbf{L}_j \cdots 0)^t$, com $c\mathbf{L}_j$ na i-ésima linha, temos que

$$\begin{aligned}
\mathbf{T}_{ij}(c)\mathbf{A} &= (\mathbf{I}_n + c\mathbf{E}_{ij})\mathbf{A} \\
&= \mathbf{A} + c\mathbf{E}_{ij}\mathbf{A} \\
&= \begin{pmatrix} \mathbf{L}_1 & \cdots & \mathbf{L}_{i-1} & \mathbf{L}_i + c\mathbf{L}_j & \mathbf{L}_{i+1} & \cdots & \mathbf{L}_n \end{pmatrix}^t,
\end{aligned}$$

ou seja, o esquema $\mathbf{L}_i \to \mathbf{L}_i + c\mathbf{L}_j$. ■

Isto motiva a seguinte definição. Seja $\mathbf{A} \in F^{m\times n}$. As *operações elementares* sobre as linhas de $\mathbf{A}$ são:

1. Permutação das i-ésima e j-ésima linhas. $(L_i \leftrightarrow L_j)$

2. Multiplicação da i-ésima linha por um escalar não nulo c. $(L_i \to cL_i)$

3. Substituição da i-ésima linha pela i-ésima linha mais c vezes a j-ésima linha, $i \neq j$. $(L_i \to L_i + cL_j)$

É claro que essas operações elementares sobre as linhas de $\mathbf{A}$ são equivalentes a multiplicar $\mathbf{A}$ à esquerda por matrizes elementares.

Sejam $\mathbf{A}, \mathbf{R} \in F^{m\times n}$. Diremos que $\mathbf{R}$ é *equivalente por linha* a $\mathbf{A}$ se $\mathbf{R}$ for obtida de $\mathbf{A}$ através de uma sequência finita de operações elementares sobre as linhas de $\mathbf{A}$. É útil e didático representa esta relação por um *diagrama de flechas*:

$$\mathbf{A} = \mathbf{A}_0 \longrightarrow \mathbf{A}_1 \longrightarrow \cdots \longrightarrow \mathbf{A}_k \longrightarrow \mathbf{A}_{k+1} = \mathbf{R},$$

em que a flecha significa uma única operação elementar e $\mathbf{A}_i$ é equivalente por linha a $\mathbf{A}_j$, quando $i < j$. Por exemplo, as matrizes

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}, \mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in F^{2\times 2},$$

são equivalentes por linhas, pois

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} \xrightarrow{L_1 \to \frac{1}{2}L_1} \begin{pmatrix} 1 & \frac{1}{2} \\ 5 & 3 \end{pmatrix} \xrightarrow{L_2 \to L_2 - 5L_1} \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}$$
$$\xrightarrow{L_2 \to 2L_2} \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix} \xrightarrow{L_1 \to L_1 - \frac{1}{2}L_2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{I}.$$

Observe que

$$\mathbf{I} = \mathbf{T}_{12}\left(-\frac{1}{2}\right)\mathbf{D}_2(2)\mathbf{T}_{21}(-5)\mathbf{D}_1\left(\frac{1}{2}\right)\mathbf{A},$$

isto é, existe uma matriz invertível $\mathbf{P} \in F^{2\times 2}$ tal que $\mathbf{PA} = \mathbf{I}$. Mais geralmente, temos o seguinte resultado:

Proposição 1.11 *Sejam $\mathbf{A}, \mathbf{R} \in F^{m\times n}$. Então $\mathbf{R}$ é equivalente por linha a $\mathbf{A}$ se, e somente se, existir uma matriz invertível $\mathbf{P} \in F^{m\times m}$ tal que $\mathbf{PA} = \mathbf{R}$, em que $\mathbf{P}$ é um produto de matrizes elementares.*

Prova. Suponhamos que $\mathbf{R}$ seja equivalente por linha a $\mathbf{A}$. Então existem matrizes elementares $\mathbf{E}_1, \ldots, \mathbf{E}_k \in F^{m\times m}$ tais que

$$\mathbf{R} = \mathbf{E}_k \cdots \mathbf{E}_1\mathbf{A}.$$

Assim, existe uma matriz invertível $\mathbf{P} = \mathbf{E}_k \cdots \mathbf{E}_1 \in F^{m\times m}$ tal que $\mathbf{PA} = \mathbf{R}$.

Reciprocamente, suponhamos que $\mathbf{PA} = \mathbf{R}$, em que $\mathbf{P} = \mathbf{E}_k \cdots \mathbf{E}_1$ é um produto de matrizes elementares. Então $\mathbf{E}_1\mathbf{A}$ é equivalente por linha a $\mathbf{A}$ e $\mathbf{E}_2(\mathbf{E}_1\mathbf{A})$ é equivalente por linha a $\mathbf{E}_1\mathbf{A}$ implicam que $\mathbf{E}_2(\mathbf{E}_1\mathbf{A})$ é equivalente por linha a $\mathbf{A}$ e, indutivamente, $\mathbf{R} = \mathbf{PA}$ é equivalente por linha a $\mathbf{A}$. ■

Corolário 1.12 *Sejam $\mathbf{A}, \mathbf{R} \in F^{n\times n}$.*

1. *Se $\mathbf{R}$ for equivalente por linha a $\mathbf{A}$, então $\mathbf{A}$ e $\mathbf{R}$ são não singulares ou $\mathbf{A}$ e $\mathbf{R}$ são singulares.*

2. *Se $\mathbf{A}$ for equivalente por linha a $\mathbf{I}$, então $\mathbf{A}$ é não singular. Conclua que $\mathbf{A}$ é um produto de matrizes elementares.*

Prova. (1) Suponhamos que $\mathbf{A}$ seja equivalente por linha a $\mathbf{R}$. Então, pela Proposição 1.11, existe uma matriz invertível $\mathbf{P} \in F^{n\times n}$ tal que $\mathbf{PA} = \mathbf{R}$. Assim, se $\mathbf{A}$ é não singular, então $\mathbf{PA}$ é não singular. Logo, $\mathbf{R}$ é não singular, de modo análogo, trata o outro caso.

(2) Como $\mathbf{I}$ é não singular temos, pelo item (1), que $\mathbf{A}$ é não singular. Portanto, $\mathbf{A}^{-1} = \mathbf{P} = \mathbf{E}_k \cdots \mathbf{E}_1$ é um produto de matrizes elementares. ■

Seja $\mathbf{A} = (a_{ij}) \in F^{m\times n}$. Diremos que $\mathbf{A}$ é uma *matriz trapezoidal superior* se $a_{ij} = 0$, quando $i > j$. Em particular, se $m = n$, diremos $\mathbf{A}$ é uma *matriz triangular superior*. Observe que se $\mathbf{A} = (a_{ij})$ e $\mathbf{B} = (b_{ij})$ são matrizes triangulares superiores, então $\mathbf{A} + \mathbf{B}, c\mathbf{A}$ e $\mathbf{AB}$ são matrizes triangulares superiores. De fato. Pondo $\mathbf{AB} = (c_{ij})$, obtemos

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = \sum_{k=i}^{j} a_{ik}b_{kj},$$

pois se $i > j$, então $i > k$ ou $k > j$, para todo k, de modo que $a_{ik} = 0$ ou $b_{kj} = 0$. Portanto, $c_{ij} = 0$.

Teorema 1.13 *Qualquer $\mathbf{A} \in F^{m\times n}$ é equivalente por linha à uma matriz trapezoidal $($triangular quando $m = n)$ superior $\mathbf{U}$.*

Prova. Vamos usar indução sobre o número de colunas n de $\mathbf{A}$. Se $n = 1$, então

$$\mathbf{A} = \mathbf{C}_1 = \begin{pmatrix} a_{11} & \cdots & a_{m1} \end{pmatrix}^t .$$

Se $\mathbf{C}_1 = \mathbf{O}$, nada há para ser provado. Caso contrário, o seguinte conjunto $S = \{i \in \mathbb{N} : a_{i1} \neq 0\}$ é não vazio. Assim, pelo PBO, S contém um menor elemento, digamos $k \in S$. Efetuando sucessivas permutações de linhas, podemos supor que $k = 1$ e $a_{11} \neq 0$. Logo, a sequência de operações elementares

$$\mathbf{L}_i \longrightarrow \mathbf{L}_i - a_{11}^{-1}a_{i1}\mathbf{L}_1, \;\; i = 2, \ldots, m,$$

implicam que

$$\mathbf{A} \longrightarrow \cdots \longrightarrow \begin{pmatrix} a_{11} & 0 & \cdots & 0 \end{pmatrix}^t ,$$

ou seja, $\mathbf{A}$ é equivalente por linha à uma matriz trapezoidal superior. Suponhamos que o resultado seja válido para todo j, com $1 \leq j < n$ e $n > 1$. Então

$$\mathbf{C}_{j+1} = \begin{pmatrix} * & \cdots & * & a_{(j+1)(j+1)} & \cdots & a_{m(j+1)} \end{pmatrix}^t ,$$

em que $*$ significa que o elemento não interfere no resultado. Se $a_{k(j+1)} = 0$, quando $k = j + 2, \ldots, m$, nada há para ser provado. Caso contrário, pela base de indução, obtemos

$$\mathbf{C}_{j+1} \longrightarrow \cdots \longrightarrow \begin{pmatrix} * & \cdots & * & a_{(j+1)(j+1)} & 0 & \cdots & 0 \end{pmatrix}^t.$$

Portanto, o resultado é válido para $j+1$. Consequentemente, $\mathbf{A}$ é equivalente por linha à uma matriz trapezoidal superior $\mathbf{U}$. ■

Vamos aplicar o Teorema 1.13 a matriz $(\mathbf{I}|\mathbf{A})$, onde

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & -2 & 4 \\ 1 & 1 & -1 & 2 \\ 1 & 4 & -4 & 5 \end{pmatrix} \in F^{3\times 4}.$$

Note que

$$\left(\begin{array}{ccc|cccc} 1 & 0 & 0 & 1 & 1 & -2 & 4 \\ 0 & 1 & 0 & 1 & 1 & -1 & 2 \\ 0 & 0 & 1 & 1 & 4 & -4 & 5 \end{array}\right) \longrightarrow \cdots \longrightarrow \left(\begin{array}{ccc|cccc} 1 & 0 & 0 & 1 & 1 & -2 & 4 \\ -1 & 0 & 1 & 0 & 3 & -2 & 1 \\ -1 & 1 & 0 & 0 & 0 & 1 & -2 \end{array}\right)$$

ou $\mathbf{P}_{23}\mathbf{T}_{13}(-1)\mathbf{T}_{12}(-1)\mathbf{A} = \mathbf{U}$. Observe que

$$\mathbf{P}_{23}\mathbf{T}_{31}(-1)\mathbf{T}_{21}(-1) = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

é uma matriz triangular inferior permutada $(L_2 \leftrightarrow L_3)$ e $\mathbf{A} = \mathbf{PLU}$ ou simplesmente $\mathbf{A} = \mathbf{LU}$ chama-se a *decomposição LU* de $\mathbf{A}$.

Teorema 1.14 *Seja $\mathbf{U} \in F^{n\times n}$ uma matriz triangular superior. Então $\mathbf{U}$ é singular se, e somente se, a diagonal de $\mathbf{U}$ possui pelo menos um elemento nulo.*

Prova. Suponhamos que $\mathbf{U}$ possua todos os elementos não nulos em sua diagonal. Então $\mathbf{U}$ é equivalente por linhas a $\mathbf{I}_n$. Assim, pelo item (2) do Corolário 1.12, temos que $\mathbf{U}$ é não singular.

Reciprocamente, suponhamos que $\mathbf{U}$ possua pelo menos um elemento nulo em sua diagonal, digamos $a_{ii} = 0$. Se $a_{nn} = 0$, então $\mathbf{U}$ possui uma linha nula. Assim, pelo item (1) do Exemplo 1.7, $\mathbf{U}$ é singular. Se $a_{nn} \neq 0$, então a sequência de operações elementares

$$\mathbf{L}_i \longrightarrow \mathbf{L}_i - a_{nn}^{-1}a_{in}\mathbf{L}_n, \ \ i = 1, \ldots, n-1,$$

implicam que $\mathbf{U}$ é equivalente por linha a uma matriz com a última coluna igual a

$$\mathbf{C}_n = \begin{pmatrix} 0 & \cdots & 0 & a_{nn} \end{pmatrix}^t.$$

De modo análogo, trabalha com os elementos $a_{(n-j)(n-j)}$, para $j = 1, \ldots, n-i$. Logo, $\mathbf{U}$ é equivalente por linha a uma matriz com uma linha nula. Portanto, pelo item (1) do Corolário 1.12, $\mathbf{U}$ é singular. ∎

Proposição 1.15 (Critério da Inversa) *Seja $\mathbf{A} \in F^{n\times n}$.*

1. *$\mathbf{A}$ é singular se, e somente se, $\mathbf{A}$ é equivalente por linha a uma matriz triangular superior com pelo menos um elemento nulo em sua diagonal.*

2. *$\mathbf{A}$ é não singular se, e somente se, $\mathbf{A}$ é equivalente por linha a $\mathbf{I}$.*

3. *Se $\mathbf{A}$ é não singular, então*

$$\left(\mathbf{A} \mid \mathbf{I} \right) \longrightarrow \left(\mathbf{I} \mid \mathbf{P} \right).$$

   Neste caso, $\mathbf{P} = \mathbf{A}^{-1}$.

4. *Se existir uma matriz invertível $\mathbf{P} \in F^{n\times n}$ tal que $\mathbf{PA} = \mathbf{I}$ ou $\mathbf{AP} = \mathbf{I}$, então $\mathbf{A}$ é não singular.*

Prova. Vamos provar apenas os itens (2) e (3): (2) Pelo Teorema 1.13, $\mathbf{A}$ é equivalente por linha à uma matriz triangular superior $\mathbf{R}$ e, pelo item (1) do Corolário 1.12, $\mathbf{R}$ é não singular. Assim, $\mathbf{A} \longrightarrow \mathbf{R} \longrightarrow \mathbf{I}$.

(3) Pelo item (2), $\mathbf{A}$ é equivalente por linha a $\mathbf{I}$. Assim, pela Proposição 1.11, existe uma matriz invertível $\mathbf{P} = \mathbf{E}_k \cdots \mathbf{E}_1 \in F^{n\times n}$ tal que $\mathbf{PA} = \mathbf{I}$. Logo,

$$\mathbf{P}\left(\mathbf{A} \mid \mathbf{I} \right) = \left(\mathbf{PA} \mid \mathbf{P} \right) = \left(\mathbf{I} \mid \mathbf{P} \right).$$

Finalmente, como $\mathbf{PA} = \mathbf{I}$ temos que $\mathbf{P} = \mathbf{A}^{-1}$. ∎

A prova do item (3) da Proposição 1.15 nos fornece um método para obter a inversa de uma matriz invertível:

$$\left(\mathbf{A} \mid \mathbf{I} \right) \longrightarrow \left(\mathbf{E}_1\mathbf{A} \mid \mathbf{E}_1 \right) \longrightarrow \cdots \longrightarrow \left(\mathbf{I} \mid \mathbf{E}_k \cdots \mathbf{E}_1 = \mathbf{A}^{-1} \right).$$

É pertinente observar que no método acima devemos efetuar as operações elementares necessárias para reduzir $\mathbf{A}$ a $\mathbf{I}$. Além disso, ele é tedioso de escrever. porém todas as operações são fáceis.

Exemplo 1.16 *Seja*

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in F^{2\times 2}.$$

Mostre que $a \neq 0$ *e* $ad - bc = D \neq 0$ *se, e somente se, existem* $q, r, s, t \in F$*, com* $qs \neq 0$*, tais que* $\mathbf{A} = \mathbf{D}_1(q)\mathbf{T}_{21}(r)\mathbf{D}_2(s)\mathbf{T}_{12}(t)$*.*

Solução. É fácil verificar que

$$\mathbf{T}_{12}(-a^{-1}b)\mathbf{D}_2(aD^{-1})\mathbf{T}_{21}(-c)\mathbf{D}_1(a^{-1})\mathbf{A} = \mathbf{I}$$

Portanto, existem $q = a, r = c, s = a^{-1}D$ e $t = a^{-1}b$, com $qs = D \neq 0$. Reciprocamente, como

$$\mathbf{D}_1(q)\mathbf{T}_{21}(r)\mathbf{D}_2(s)\mathbf{T}_{12}(t) = \begin{pmatrix} q & qt \\ r & s + rt \end{pmatrix}$$

temos que $a = q, b = qt, c = r$ e $d = s + rt$. Assim, $a \neq 0$ e $D = q(s + rt) - qrt = qs \neq 0$. Observe que

$$\mathbf{A}^{-1} = \mathbf{T}_{12}(-a^{-1}b)\mathbf{D}_2(aD^{-1})\mathbf{T}_{21}(-c)\mathbf{D}_1(a^{-1}) = \frac{1}{qs}\begin{pmatrix} s + rt & -qt \\ -r & q \end{pmatrix}.$$

Note que se $a = 0$ e $c \neq 0$, o processo aplica-se a $\mathbf{B} = \mathbf{P}_{12}\mathbf{A}$. ■

Seja $\mathbf{A} \in F^{m\times n}$. As *operações elementares sobre as colunas* de $\mathbf{A}$ são:

1. Permutação das i-ésima e j-ésima colunas. ($C_i \leftrightarrow C_j$)

2. Multiplicação da j-ésima coluna por um escalar não nulo c. ($C_j \to cC_j$)

3. Substituição da j-ésima coluna pela j-ésima coluna mais c vezes a i-ésima coluna, $i \neq j$. ($C_j \to C_j + cC_i$)

De modo análogo, as operações elementares sobre as linhas de $\mathbf{A}$, é claro que essas operações elementares sobre as colunas de $\mathbf{A}$ são equivalentes a multiplicar $\mathbf{A}$ à direita por matrizes elementares. Por exemplo,

$$\begin{aligned} \mathbf{A}\mathbf{T}_{ij}(c) &= \mathbf{A}(\mathbf{I}_n + c\mathbf{E}_{ij}) = \mathbf{A} + c\mathbf{A}\mathbf{E}_{ij} \\ &= \begin{pmatrix} \mathbf{C}_1 & \cdots & \mathbf{C}_{j-1} & \mathbf{C}_j + c\mathbf{C}_i & \mathbf{C}_{j+1} & \cdots & \mathbf{C}_n \end{pmatrix}, \end{aligned}$$

ou seja, o esquema $\mathbf{C}_j \to \mathbf{C}_j + c\mathbf{C}_i$.

Sejam $\mathbf{A}, \mathbf{B} \in F^{m\times n}$. Diremos que $\mathbf{A}$ é *equivalente* a $\mathbf{B}$ se $\mathbf{B}$ for obtida de $\mathbf{A}$ através de um número finito de operações elementares sobre as linhas e as colunas de $\mathbf{A}$. Neste caso, $\mathbf{B}$ é equivalente a $\mathbf{A}$ se, e somente se, existirem matrizes invertíveis $\mathbf{P} \in F^{m\times m}$ e $\mathbf{Q} \in F^{n\times n}$ tais que $\mathbf{B} = \mathbf{PAQ}$, é fácil verificar que equivalência é uma relação de equivalência sobre $F^{m\times n}$. Por exemplo, as matrizes

$$\mathbf{A} = \begin{pmatrix} 1 & -1 \\ -1 & 5 \end{pmatrix}, \mathbf{D} = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} \in F^{2\times 2},$$

são equivalentes, pois

$$\begin{pmatrix} 1 & -1 \\ -1 & 5 \end{pmatrix} \xrightarrow{L_2 \to L_2 + L_1 (C_2 \to C_2 + C_1)} \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}.$$

Portanto,

$$\begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Sejam $\mathbf{A}, \mathbf{B} \in F^{n\times n}$ e $\mathbf{E} \in F^{n\times n}$ uma matriz elementar. Diremos que a matriz $\mathbf{B} = \mathbf{E}^{-1}\mathbf{AE}$ é uma *semelhança elementar* de $\mathbf{A}$. Por exemplo, se $c \neq 0$, então $\mathbf{D}_i(c^{-1})\mathbf{AD}_i(c)$ significa que $\mathbf{L}_i \to c^{-1}\mathbf{L}_i$ e $\mathbf{C}_i \to c\mathbf{C}_i$. Por outro lado, se $i \neq j$, então $\mathbf{T}_{ij}(-c)\mathbf{AT}_{ij}(c)$ significa que $\mathbf{L}_i \to \mathbf{L}_i - c\mathbf{L}_j$ e $\mathbf{C}_j \to \mathbf{C}_j + c\mathbf{C}_i$. Portanto, as matrizes $\mathbf{A}$ e $\mathbf{B}$ são *semelhantes* (ou *conjugadas*) se, e somente se, se $\mathbf{B}$ for obtida de $\mathbf{A}$ através de uma sequência finita de operações de semelhanças elementares. Neste caso, $\mathbf{B}$ é semelhante a $\mathbf{A}$ se, e somente se, existir uma matriz invertíveis $\mathbf{P} \in F^{n\times n}$ tal que $\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}$, é fácil verificar que semelhança é uma relação de equivalência sobre $F^{n\times n}$. Por exemplo, as matrizes

$$\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}, \mathbf{D} = \begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix} \in F^{2\times 2},$$

são semelhantes, pois

$$\begin{aligned}
&\begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \xrightarrow{L_2 \to L_2 - L_1 (C_1 \to C_1 + C_2)} \begin{pmatrix} 4 & 3 \\ 0 & -2 \end{pmatrix} \\
&\begin{pmatrix} 4 & 3 \\ 0 & -2 \end{pmatrix} \xrightarrow{L_2 \to -\frac{1}{2} L_2 (C_2 \to -2C_2)} \begin{pmatrix} 4 & -6 \\ 0 & -2 \end{pmatrix} \\
&\begin{pmatrix} 4 & -6 \\ 0 & -2 \end{pmatrix} \xrightarrow{L_1 \to L_1 - L_2 (C_2 \to C_2 + C_1)} \begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix}.
\end{aligned}$$

Portanto,

$$\begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

em que

$$\mathbf{P} = \mathbf{T}_{12}(1)\mathbf{D}_2(-2)\mathbf{T}_{21}(1) \text{ e } \mathbf{P}^{-1} = \mathbf{T}_{12}(-1)\mathbf{D}_2(-2^{-1})\mathbf{T}_{21}(-1).$$

É pertinente observar que matrizes semelhantes são sempre equivalentes. Por outro lado, já vimos que as matrizes

$$\mathbf{A} = \begin{pmatrix} 1 & -1 \\ -1 & 5 \end{pmatrix}, \mathbf{D} = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} \in F^{2\times 2},$$

são equivalentes, mas não são semelhantes, ou seja, a recíproca é falsa.

Finalizaremos esta seção com um esquema para obter as matrizes invertíveis $\mathbf{P} \in F^{m\times m}$ e $\mathbf{Q} \in F^{n\times n}$ tais que $\mathbf{PAQ} = \mathbf{N}_r$, em que $\mathbf{N}_r = (d_{ij})$, com $d_{ii} = 1$, para $i = 1, \ldots, r$ e $r \leq \min\{m, n\}$, e as demais zero, confira Proposição 2.6. Em particular, isto nos forne um método alternativo para obter a inversa de uma matriz invertível:

$$\begin{array}{c|c} \mathbf{I}_n & \mathbf{O} \\ \hline \mathbf{A} & \mathbf{I}_m \end{array} \to \cdots \to \begin{array}{c|c} \mathbf{I}_n & \mathbf{O} \\ \hline \mathbf{PA} & \mathbf{P} \end{array} \to \cdots \to \begin{array}{c|c} \mathbf{Q} & \mathbf{O} \\ \hline \mathbf{PAQ} & \mathbf{P} \end{array}$$

ou

$$\begin{array}{c|c} \mathbf{A} & \mathbf{I}_m \\ \hline \mathbf{I}_n & \mathbf{O} \end{array} \to \cdots \to \begin{array}{c|c} \mathbf{AQ} & \mathbf{I}_m \\ \hline \mathbf{Q} & \mathbf{O} \end{array} \to \cdots \to \begin{array}{c|c} \mathbf{PAQ} & \mathbf{P} \\ \hline \mathbf{Q} & \mathbf{O} \end{array}$$

A matriz $\mathbf{N}_r$ chama-se *forma normal de Hermite*[6] de $\mathbf{A}$. Observe que as operações de linhas sobre $\mathbf{A}$ são as mesmas sobre $(\mathbf{A}|\mathbf{I}_m)$ e as operações de colunas sobre $\mathbf{A}$ são as mesmas sobre $(\mathbf{I}_n|\mathbf{A})^t$ ou $(\mathbf{A}|\mathbf{I}_n)^t$. Neste caso, a ordem das operações não importa, pois

[6]Charles Hermite, 1822-1901, matemático francês.

$(\mathbf{EA})\mathbf{F} = \mathbf{E}(\mathbf{AF})$. Por exemplo,

$$
\begin{array}{rr|rr}
1 & 0 & & \\
0 & 1 & & \\
\hline
2 & 1 & 1 & 0 \\
5 & 3 & 0 & 1
\end{array}
\xrightarrow{C_1 \to C_2}
\begin{array}{rr|rr}
0 & 1 & & \\
1 & 0 & & \\
\hline
1 & 2 & 1 & 0 \\
3 & 5 & 0 & 1
\end{array}
\xrightarrow{L_2 \to L_2 - 3L_1}
$$

$$
\begin{array}{rr|rr}
0 & 1 & & \\
1 & 0 & & \\
\hline
1 & 2 & 1 & 0 \\
0 & -1 & -3 & 1
\end{array}
\xrightarrow{L_2 \to -L_2 (C_2 \to C_2 - 2C_1)}
\begin{array}{rr|rr}
0 & 1 & & \\
1 & -2 & & \\
\hline
1 & 0 & 1 & 0 \\
0 & 1 & 3 & -1
\end{array}
$$

Portanto, $\mathbf{PAQ} = \mathbf{I}$ ou $\mathbf{A} = \mathbf{P}^{-1}\mathbf{Q}^{-1} = (\mathbf{QP})^{-1}$, de modo que $\mathbf{A}^{-1} = \mathbf{QP}$.

Exercícios

1. Mostre todas as afirmações deixadas nesta seção.

2. Seja
$$M_2(\mathbb{R}) = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} : a, b \in \mathbb{R} \right\}.$$
Mostre que $M_2(\mathbb{R})$ com as operações usuais de matrizes é um corpo. Conclua que a função $f : \mathbb{C} \to M_2(\mathbb{R})$ definida como
$$f(a + bi) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$
é bijetora e preserva as operações.

3. Mostre que existem matrizes $\mathbf{A}, \mathbf{B} \in F^{2\times 2}$ tais que
$$(\mathbf{A} + \mathbf{B})^2 \neq \mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2 \text{ e } (\mathbf{A} - \mathbf{B})(\mathbf{A} + \mathbf{B}) \neq \mathbf{A}^2 - \mathbf{B}^2.$$

4. Seja $\mathbf{A} \in F^{n\times n}$. Mostre que $(\mathbf{P}^{-1}\mathbf{AP})^m = \mathbf{P}^{-1}\mathbf{A}^m\mathbf{P}$, para toda matriz invertível $\mathbf{P} \in F^{n\times n}$ e todo $m \in \mathbb{N}$.

5. Sejam $\mathbf{A}, \mathbf{B} \in F^{n\times n}$. Mostre que

$$\mathbf{B}^m - \mathbf{A}^m = \sum_{i=0}^{m-1} \mathbf{A}^i(\mathbf{B} - \mathbf{A})\mathbf{B}^{m-(i+1)}, \ \forall \ m \in \mathbb{N}.$$

   Conclua que $\mathbf{I} - \mathbf{A}^m = (\sum_{i=0}^{m-1} \mathbf{A}^i)(\mathbf{I} - \mathbf{A})$, para todo $m \in \mathbb{N}$.

6. Mostre que $\mathbf{A} = \mathbf{I}$ se, e somente se, $\mathbf{AX} = \mathbf{X}$, para todo $\mathbf{X} \in F^{n\times 1}$.

7. Sejam $\mathbf{A}, \mathbf{B} \in F^{n\times n}$. Diremos que $\mathbf{A}$ é *idempotente* se $\mathbf{A}^2 = \mathbf{A}$.
   (a) Determine todas as matrizes idempotentes em $\mathbb{R}^{2\times 2}$.
   (b) Mostre que se $\mathbf{AB} = \mathbf{A}$ e $\mathbf{BA} = \mathbf{B}$, então $\mathbf{A}$ e $\mathbf{B}$ são idempotentes.
   (c) Mostre que se $\mathbf{A} + \mathbf{B} = \mathbf{I}$ e $\mathbf{AB} = \mathbf{O}$, então $\mathbf{A}$ e $\mathbf{B}$ são idempotentes.

8. Seja $\mathbf{A} \in F^{n\times n}$. Diremos que $\mathbf{A}$ é *nilpotente* de índice $k \in \mathbb{N}$ se $\mathbf{A}^k = \mathbf{O}$, mas $\mathbf{A}^{k-1} \neq \mathbf{O}$.
   (a) Determine todas as matrizes nilpotentes de índice 2 em $\mathbb{R}^{2\times 2}$.
   (b) Mostre que se $\mathbf{A}$ é nilpotente, então $\mathbf{I} - \mathbf{A}$ é invertível.

9. Seja $\mathbf{A} \in F^{n\times n}$. Diremos que $\mathbf{A}$ é uma *involução* se $\mathbf{A}^2 = \mathbf{I}$.
   (a) Determine todas as matrizes involuções em $\mathbb{R}^{2\times 2}$.
   (b) Mostre que $\mathbf{A}$ é uma involução se, somente se, $(\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A}) = \mathbf{O}$.

10. Seja $\mathbf{A} \in F^{n\times n}$. Diremos que $\mathbf{A}$ é uma *matriz simétrica* se $\mathbf{A}^t = \mathbf{A}$ e que $\mathbf{A}$ é uma *matriz antissimétrica* se $\mathbf{A}^t = -\mathbf{A}$.
   (a) Determine todas as matrizes simétricas (antissimétricas) em $\mathbb{R}^{2\times 2}$.
   (b) Mostre que se $\mathbf{A}$ e $\mathbf{B}$ são simétricas (antissimétricas), então $\mathbf{A} + \mathbf{B}$ e $\mathbf{A} - \mathbf{B}$ são simétricas (antissimétricas).
   (c) Mostre que se $\mathbf{A}$ é simétricas (antissimétricas), então $\mathbf{P}^t\mathbf{AP}$ é simétrica (antissimétrica), para todo $\mathbf{P} \in F^{n\times n}$.
   (d) Mostre que se $\mathbf{A}$ e $\mathbf{B}$ são simétricas então $\mathbf{AB}$ é simétrica se, e somente se, $\mathbf{AB} = \mathbf{BA}$.
   (e) Mostre que $\mathbf{AA}^t$ e $\mathbf{A} + \mathbf{A}^t$ são simétrica e $\mathbf{A} - \mathbf{A}^t$ é antissemétrica.

11. Sejam $\mathbf{A}, \mathbf{B} \in F^{n\times n}$. Mostre que $\mathbf{A}(\mathbf{I}-\mathbf{BA}) = (\mathbf{I}-\mathbf{AB})\mathbf{A}$ e $(\mathbf{I}-\mathbf{BA})\mathbf{B} = \mathbf{B}(\mathbf{I}-\mathbf{AB})$. Conclua que se $\mathbf{I}-\mathbf{AB}$ é invertível, então $\mathbf{I}-\mathbf{BA}$ é invertível e
$$(\mathbf{I}-\mathbf{BA})^{-1} = \mathbf{I}+\mathbf{B}(\mathbf{I}-\mathbf{AB})^{-1}\mathbf{A}.$$

12. Sejam $\mathbf{A}, \mathbf{B}, \mathbf{P} \in F^{n\times n}$ tais que $\mathbf{B}$, $\mathbf{P}$ e $\mathbf{APA}^t + \mathbf{B}^{-1}$ sejam invertíveis. Mostre que $\mathbf{P}^{-1} + \mathbf{A}^t\mathbf{BA}$ é invertível e
$$(\mathbf{P}^{-1} + \mathbf{A}^t\mathbf{BA})^{-1} = \mathbf{P} - \mathbf{PA}^t(\mathbf{APA}^t + \mathbf{B}^{-1})^{-1}\mathbf{AP}.$$

13. Seja $\mathbf{A} \in F^{n\times n}$ invertível. Mostre que existem $\mathbf{X}, \mathbf{Y} \in F^{n\times n}$ tais que
$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{X} & \mathbf{I} \end{pmatrix}\begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{S} \end{pmatrix}\begin{pmatrix} \mathbf{I} & \mathbf{Y} \\ \mathbf{O} & \mathbf{I} \end{pmatrix},$$
com $\mathbf{S} = \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}$ o *complemento de Schur*[7].

14. Seja $\mathbf{A} \in F^{n\times n}$. Mostre que $\mathbf{AB} = \mathbf{BA}$, para todo $\mathbf{B} \in F^{n\times n}$ se, e somente se, $\mathbf{A} = a\mathbf{I}$, para algum $a \in F$.

15. Sejam $\mathbf{A}, \mathbf{D} \in F^{n\times n}$ tais que $\mathbf{D} = (\mathbf{I}+\mathbf{A})^{-1}\mathbf{A}$ é diagonal e invertível. Mostre que $\mathbf{A}$ é diagonal.

16. Seja $\mathbf{A} \in F^{2\times 2}$, com $a_{11} \neq a_{22}$. Mostre que qualquer $\mathbf{X} \in F^{2\times 2}$, com $x_{11} \neq x_{22}$, tal que $\mathbf{AX} = \mathbf{XA}$ é da forma $\mathbf{X} = u\mathbf{A} + v\mathbf{I}$, onde $u, v \in F$ e $u \neq 0$.

17. Seja $\mathbf{A} = (a_{ij}) \in F^{n\times n}$. O *traço* de $\mathbf{A}$ é definido como
$$\operatorname{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}.$$
Mostre que:
   (a) $\operatorname{tr}(\mathbf{A}+\mathbf{B}) = \operatorname{tr}(\mathbf{A}) + \operatorname{tr}(\mathbf{B})$, para todos $\mathbf{A}, \mathbf{B} \in F^{n\times n}$.
   (b) $\operatorname{tr}(c\mathbf{A}) = c\operatorname{tr}(\mathbf{A})$, para todo $\mathbf{A} \in F^{n\times n}$ e $c \in F$.
   (c) $\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA})$, para todos $\mathbf{A}, \mathbf{B} \in F^{n\times n}$.

[7]Issai Schur , 1875-1911, matemático israelita.

(d) Mostre que $\mathrm{tr}(\mathbf{A}^t\mathbf{A}) = \sum_{i=1}^n \sum_{j=1}^n a_{ii}^2$. Conclua que $\mathbf{A} = \mathbf{O}$ se, e somente se, $\mathrm{tr}(\mathbf{A}^t\mathbf{A}) = 0$.

(e) $\mathrm{tr}(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}) = \mathrm{tr}(\mathbf{A})$, para todos $\mathbf{A}, \mathbf{P} \in F^{n\times n}$, com $\mathbf{P}$ invertível.

(f) $\mathrm{tr}(\mathbf{AB} - \mathbf{BA}) = 0$, para todos $\mathbf{A}, \mathbf{B} \in F^{n\times n}$.

18. Seja $\mathbf{U}_n = (u_{ij}) \in F^{n\times n}$, com $u_{ij} = 1$, para todos i, j. Mostre que se $n > 1$, então $(\mathbf{I} - \mathbf{U}_n)^{-1} = \mathbf{I} - (n-1)^{-1}\mathbf{U}_n$.

1.3 Determinantes

Por volta de 250 a. C. já havia exemplos de resolução de sistemas de equações através de matrizes, no livro chinês Nove Capítulos sobre a Arte Matemática, cujo autor é desconhecido. Além disso, algumas noções ligadas a determinantes, o assunto que será objeto de estudo desta seção, já eram conhecidas na China antiga. É importante observar que os determinantes possuem pouco valor prático no cálculo em geral. Não obstante, possuem muito valor como instrumento teórico.

Seja $\mathbf{A} \in F^{n\times n}$. O *determinante* de $\mathbf{A}$ é um escalar associado com $\mathbf{A}$. Existem várias maneiras de definir o determinante. Mas, em termo de permutações é a mais concisa e às vezes conveniente, no entanto, isto pareça complicado em uma primeira leitura. Para isto, vamos apresentar alguns conceitos e resultados sobre permutações.

Seja $S = \{1, \ldots, n\}$. Uma *permutação* sobre S é qualquer função bijetora σ de S sobre S, ou seja, é um rearranjo dos elementos de S em alguma ordem sem repetições ou omissões. Vamos denotar por S_n o conjunto de todas as permutações de S. Então S_n é um *grupo de permutações*, isto é, S_n munido com a composição usual de funções satisfaz os seguintes axiomas:

1. $\sigma \circ (\tau \circ \phi) = (\sigma \circ \tau) \circ \phi$, para todos $\sigma, \tau, \phi \in S_n$.

2. Existe um $I = I_S \in S_n$ tal que $\sigma \circ I = I \circ \sigma = \sigma$, para todo $\sigma \in S_n$.

3. Para cada $\sigma \in S_n$, existe um $\sigma^{-1} \in S_n$ tal que $\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = I$, com $\sigma(i) = j$ se, e somente se, $\sigma^{-1}(j) = i$.

É fácil verificar que a ordem de S_n é igual a $n!$. Neste caso, S_n chama-se *grupo simétrico de grau* n.

Um elemento $\sigma \in S_n$ pode ser escrita sob a forma

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix},$$

em que a ordem das colunas não importa e $\sigma(1), \ldots, \sigma(n)$ são simplesmente os elementos de S em uma ordem diferente, ou seja, $S = \{\sigma(1), \ldots, \sigma(n)\} = \sigma(S)$. Por exemplo, para $n = 3$, temos que os seis elementos de S_3 são:

$$\begin{aligned} I &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \ \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \ \sigma^2 = \sigma \circ \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \\ \tau &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \ \sigma \circ \tau = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \ \sigma^2 \circ \tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}. \end{aligned}$$

Sejam $\sigma \in S_n$ e $i \in S$ fixado. Então, aplicando sucessivamente σ a i, obtemos

$$i, \ \sigma(i), \ \sigma^2(i), \ldots, \sigma^m(i), \ldots$$

Como $\sigma^m(i) \in S$ e S é um conjunto finito temos que esses elementos não são distintos. Assim, existem $r, s \in \mathbb{Z}_+$, com $s > r$, tais que

$$\sigma^r(i) = \sigma^s(i) \Rightarrow \sigma^{s-r}(i) = (\sigma^s \circ \sigma^{-r})(i) = (\sigma^r \circ \sigma^{-r})(i) = I(i) = i.$$

Logo, o conjunto $T = \{m \in \mathbb{N} : \sigma^m(i) = i\}$ é não vazio, Portanto, pelo PBO, T contém um menor elemento, digamos k. Neste caso,

$$\sigma^k(i) = i, \ \text{ mas } \ \sigma^r(i) \neq i, \ \text{ para } \ r = 1, \ldots, k-1.$$

O subconjunto $\mathcal{O}(i) = \{i, \sigma(i), \ldots, \sigma^{k-1}(i)\}$ de S chama-se uma *órbita* de i e pode ser visualizado, geometricamente, em volta de um círculo. Observe que se $i_1 = i, \sigma(i_1) = i_2, \ldots, \sigma^{k-1}(i_1) = i_k$, então

$$\sigma = \begin{pmatrix} 1 & \cdots & i_1 & \cdots & i_{k-1} & i_k & \cdots & n \\ 1 & \cdots & i_2 & \cdots & i_k & i_1 & \cdots & n \end{pmatrix} \ \text{ e } \ \sigma(x) = x, \ \forall \ x \notin \{i_1, \ldots, i_k\}.$$

Isto motiva a seguinte definição. Diremos que $\sigma \in S_n$ é um *k-ciclo* se existirem elementos distintos $i_1, \ldots, i_k \in S$ tais que

$$\sigma(i_1) = i_2, \ldots, \sigma(i_k) = i_1 \ \text{ e } \ \sigma(x) = x, \ \forall \ x \in S - \{i_1, \ldots, i_k\}$$

e denotamos por $\sigma = (i_1 \ldots i_k)$. O número k chama-se o *comprimento* do ciclo. Por exemplo,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 4 & 5 \end{pmatrix} = (123)(4)(5) = (123) \ \text{ e } \ \mathcal{O}(1) = \{1, 2, 3\}$$

é um 3-ciclo. Um ciclo de comprimento dois chama-se uma *transposição* ou uma *inversão*. É claro que nem todo elemento de S_n é um ciclo.

Note que se $\sigma = (i_1 \ldots i_k)$ é um k-ciclo em S_n, então seu inverso

$$\sigma^{-1} = (i_k i_{k-1} \ldots i_2 i_1) = (i_1 i_k i_{k-1} \ldots i_2).$$

é um k-ciclo. Além disso, σ pode ser escrito de k maneiras, a saber,

$$\sigma = (i_j i_{j+1} \ldots i_k i_1 \ldots i_{j-1}), \ \ j = 1, \ldots, k.$$

Assim, existem

$$\frac{n(n-1)\cdots(n-k+1)}{k} = (k-1)! \begin{pmatrix} n \\ k \end{pmatrix}$$

k-ciclos distintos em S_n. Observe que

$$\sigma(i_1) = i_2, \ \ \sigma^2(i_1) = \sigma(i_2) = i_3, \ldots, \sigma^{k-1}(i_1) = i_k \ \text{ e } \ \sigma^k(i_1) = i_1.$$

Portanto, qualquer k-ciclo σ dar origem a um *subgrupo cíclico*

$$H = \langle \sigma \rangle = \{I, \sigma, \ldots, \sigma^{k-1}\} = \{\sigma^m : m \in \mathbb{Z}\}$$

em S_n. Neste caso, H é um subgrupo cíclico de ordem k. Em particular, um elemento $\sigma \in S_n$ definido como

$$\sigma(i) = \begin{cases} i+1, & \text{se } \ 1 \leq i \leq n-1 \\ 1, & \text{se } \ i = n \end{cases}$$

chama-se n-*ciclo* e $\sigma = (123\cdots n)$. Consequentemente, $\sigma^n = I$ e $\sigma^k \neq I$, para todo k, com $k = 1 \ldots, n-1$, isto é, $H = \langle \sigma \rangle$ é um subgrupo cíclico de ordem n.

Seja $\sigma \in S_n$. O subconjunto $\{i \in S : \sigma(i) \neq i\}$ de S chama-se *suporte* de σ e denotamos por $\operatorname{supp}(\sigma)$. Assim, σ é um k-ciclo se

$$\operatorname{supp}(\sigma) = \{i_1, \ldots, i_k\}.$$

Sejam σ e τ dois ciclos em S_n. Diremos que σ e τ são *permutações disjuntas* se $\operatorname{supp}(\sigma) \cap \operatorname{supp}(\tau) = \emptyset$. Neste caso, $\sigma\tau = \tau\sigma$, isto é, ciclos disjuntos comutam.

Teorema 1.17 *Qualquer $\sigma \in S_n$, com $\sigma \neq I$, pode ser escrita como um produto de transposições.*

Prova. Vamos usar indução sobre n. Se $n = 1$, nada há para ser provado. Suponhamos que o resultado seja válido para todo k, com $1 \leq k \leq n-1$ e $n > 1$. Seja $\sigma \in S_n$, com $\sigma \neq I$, tal que $\sigma(n) = k$. Consideremos a transposição $\tau \in S_n$ tal que $\tau(n) = k$ e $\tau(k) = n$. Então $\tau\sigma \in S_n$ e $(\tau\sigma)(n) = \tau(k) = n$. Assim, podemos ver $\tau\sigma$ como um elemento de S_{n-1}. Logo, existem transposições $\tau_1, \ldots . \tau_m \in S_n$ tais que $\tau\sigma = \tau_1 \cdots \tau_m$, de modo que

$$\sigma = \tau^{-1}\tau_1 \cdots \tau_m = \tau\tau_1 \cdots \tau_m,$$

que é o resultado desejado. ■

Exemplo 1.18 *Escreva a permutação $\sigma = (1234)$ como um produto de transposições.*

Solução. Escolhendo $\tau_1 = (12)$, temos $\tau_1\sigma = (234)$. Tome $\tau_2 = (23)$. Então $\tau_2\tau_1\sigma = (34)$. Pondo $\tau_4 = (34)$, obtemos $\tau_3\tau_2\tau_1\sigma = I$. Portanto, $\sigma = \tau_1\tau_2\tau_3$. ■

Observe que se $\sigma \in S_n$ é um k-ciclo, então σ pode ser escrito como um produto de $k-1$ transposições disjuntas, pois, indutivamente, obtemos

$$\begin{aligned} \sigma &= (i_1 \ldots i_k) = (i_1 i_2)(i_1 i_3 \ldots i_k) = (i_1 i_2)(i_2 i_3) \cdots (i_{k-1} i_k) \\ &= \tau_1 \cdots \tau_{k-1}. \end{aligned}$$

Para cada $\sigma \in S_n$ associamos uma medida

$$\prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i} = \prod_{i=1}^{n-1} \prod_{j=i+1}^{n} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Por exemplo, se $n = 3$ e $\sigma = (123)$ é um 3-ciclo, então

$$\prod_{1 \leq i < j \leq 3} \frac{\sigma(j) - \sigma(i)}{j - i} = \frac{\sigma(2) - \sigma(1)}{2 - 1} \cdot \frac{\sigma(3) - \sigma(1)}{3 - 1} \cdot \frac{\sigma(3) - \sigma(2)}{3 - 2} = 1.$$

Mais geralmente,

$$\prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i} = \prod_{i=1}^{n-1} \prod_{j=i+1}^{n} \frac{\sigma(j) - \sigma(i)}{j - i} = \pm 1.$$

De fato. Como σ é bijetora temos que existem únicos k e l tais que $\sigma(k) = i$ e $\sigma(l) = j$ ou $\sigma(l) - \sigma(k) = j - i$. Se $k < l$, então o fator $\sigma(l) - \sigma(k) = j - i$ aparece no numerador do produto. Se $k > l$, então o fator

$$\sigma(k) - \sigma(l) = i - j = -(j - i)$$

aparece no numerador do produto. Portanto,

$$\frac{\sigma(l) - \sigma(k)}{j - i} = 1 \text{ ou } \frac{\sigma(k) - \sigma(l)}{j - i} = -1,$$

e fatores distintos no denominador dão origem a fatores distintos no numerador, ou seja, se $\sigma(k) = \sigma(l)$ e $\sigma(k') = \sigma(l')$, então $k = l$ e $k' = l'$.

Seja $\sigma \in S_n$. Definimos o *sinal* ou *índice de paridade* de σ como

$$\operatorname{sgn} \sigma = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i} = \prod_{i=1}^{n-1} \prod_{j=i+1}^{n} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Assim, $\operatorname{sgn} \sigma = -1$ ou $\operatorname{sgn} \sigma = 1$.

Proposição 1.19 *Sejam* $\sigma, \tau \in S_n$. *Então* $\operatorname{sgn}(\tau\sigma) = \operatorname{sgn} \tau \cdot \operatorname{sgn} \sigma$. *Conclua que* $\operatorname{sgn} \sigma = \operatorname{sgn} \sigma^{-1}$.

Prova. Basta provar que

$$\prod_{1 \leq i < j \leq n} \frac{(\tau\sigma)(j) - (\tau\sigma)(i)}{\sigma(j) - \sigma(i)} = \prod_{1 \leq k < l \leq n} \frac{\tau(l) - \tau(k)}{l - k}.$$

Como $\sigma \in S_n$ temos que existem únicos p e q tais que $\sigma(p) = k$ e $\sigma(q) = l$. Se $p > q$, então

$$\frac{(\tau\sigma)(p) - (\tau\sigma)(q)}{\sigma(p) - \sigma(q)} = \frac{\tau(\sigma(q)) - \tau(\sigma(p))}{\sigma(q) - \sigma(p)} = \frac{\tau(l) - \tau(k)}{l - k}.$$

Se $p < q$, então

$$\frac{(\tau\sigma)(q) - (\tau\sigma)(p)}{\sigma(q) - \sigma(p)} = \frac{\tau(l) - \tau(k)}{l - k},$$

ou seja, os termos dos produtos estão em correspondência biunívoca. ■

Seja $\sigma \in S_n$. Diremos que σ é uma *permutação par* se $\operatorname{sgn} \sigma = 1$ e uma *permutação ímpar* se $\operatorname{sgn} \sigma = -1$. Observe que transposições são sempre permutações ímpares, pois se

$$\tau = \begin{pmatrix} 1 & 2 & \cdots & i & \cdots & j & \cdots & n-1 & n \\ 1 & 2 & \cdots & j & \cdots & i & \cdots & n-1 & n \end{pmatrix} \in S_n, \ \text{ com } \ i < j,$$

então, pelo o esquema ou diagrama cruzado, confira Figura 1.1, o número de inversões (cruzamentos) é $1+2(j-i-1)$, que é um número ímpar, pois para cada um dos $(j-i-1)$ inteiros $i+1, \ldots, j-1$ existem dois cruzamentos, enquanto i e j adicionam mais um. Portanto, $\operatorname{sgn} \tau = -1$. Consequentemente, pelo Teorema 1.17,

$$\operatorname{sgn} \sigma = (-1)^N,$$

em que N é o número de transposições na decomposição de σ. Neste caso, σ é uma permutação par se, e somente se, N é um número par.

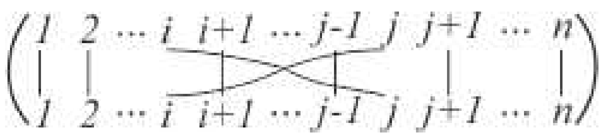


Figura 1.1: Diagrama cruzado.

Lema 1.20 *Sejam $\sigma, \tau \in S_n$ transposições. Então $\sigma\tau$ é um 3-ciclo ou um produto de dois 3-ciclos. Conclua que qualquer 3-ciclo é uma permutação par.*

Prova. Se $\sigma = \tau$, então $\sigma\tau = \sigma^2 = I$. Assim, claramente $\sigma\tau$ é um produto de dois 3-ciclos, digamos $\sigma\tau = (abc)(acb)$, para todos $a, b, c \in \{1, \ldots, n\}$. Se $\sigma \neq \tau$, então há dois caso a serem considerados. Se σ e τ não são disjuntos, digamos $\sigma = (ab)$ e $\tau = (ac)$, para todos $a, b, c \in \{1, \ldots, n\}$, então $\sigma\tau = (acb)$. Se σ e τ são disjuntos, digamos $\sigma = (ab)$ e $\tau = (cd)$, para todos $a, b, c, d \in \{1, \ldots, n\}$, então $\sigma\tau = (acb)(acd)$. ■

Seja $\sigma \in S_n$. A lista $(a_{1\sigma(1)}, \ldots, a_{n\sigma(n)})$ chama-se *diagonal* associada a σ. Enquanto, o produto $a_{1\sigma(1)} \cdots a_{n\sigma(n)}$ chama-se *produto diagonal* associada a σ. Observe que $a_{\sigma(1)\tau(\sigma(1))} \cdots a_{\sigma(n)\tau(\sigma(n))} = a_{1\tau(1)} \cdots a_{n\tau(n)}$, para todos $\sigma, \tau \in S_n$, pois cada elemento de 1 à n aparece exatamente uma vez entre $\sigma(1), \ldots, \sigma(n)$. Em particular, $a_{\sigma(1)1} \cdots a_{\sigma(n)n} = a_{1\sigma^{-1}(1)} \cdots a_{n\sigma^{-1}(n)}$. Consideremos a permutação $\sigma = (12) \in S_2$ e a matriz

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in F^{2\times 2}.$$

Então

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} = (a_{11}a_{22} - a_{12}a_{21}) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Assim, $\mathbf{A}$ é invertível se, e somente se, o escalar $a_{11}a_{22} - a_{12}a_{21} \neq 0$. Neste caso,

$$\mathbf{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

Observe que $a_{11}a_{22}$ é o produto diagonal associada ao elemento identidade I e $a_{12}a_{21} = a_{1\sigma(1)}a_{2\sigma(2)}$ é o produto diagonal associada a σ. Isto motiva a seguinte definição:

Seja $\mathbf{A} \in F^{n \times n}$. O *determinante* de $\mathbf{A}$ é definido como o escalar

$$\det \mathbf{A} = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) \prod_{i=1}^{n} a_{i\sigma(i)}.$$

Assim, $\det \mathbf{A}$ é a soma de $n!$ termos cada um dos quais envolvendo um produto diagonal, em que o sinal está bem definido e qualquer produto possui n elementos, um e somente um, de cada linha e de cada coluna de $\mathbf{A}$, pois o índice de cada linha (coluna) aparece exatamente uma vez. É muito importante e didático escrever a definição do determinante como

$$\det \mathbf{A} = a_{11} \cdots a_{nn} + \sum_{\sigma \in S_n, \sigma \neq I} (\operatorname{sgn} \sigma) \prod_{i=1}^{n} a_{i\sigma(i)}.$$

Logo, se $i \neq j$, então o produto diagonal $(\pm)a_{1*} \cdots a_{ij} \cdots a_{n*}$ não contém os fatores a_{ii} ou a_{jj}, caso contrário, ele possuiria dois fatores de uma mesma linha ou coluna. Portanto, qualquer produto diagonal em $\det \mathbf{A}$, exceto $a_{11} \cdots a_{nn}$, possui no máximo $n-2$ fatores da diagonal de $\mathbf{A}$. Comprove isto no seguinte exemplo, se $n = 3$, então

$$\begin{aligned} \det \mathbf{A} &= (-1)^0 a_{11}a_{22}a_{33} + (-1)^2 a_{12}a_{23}a_{31} + (-1)^2 a_{13}a_{21}a_{32} \\ &\quad + (-1)^1 a_{11}a_{23}a_{32} + (-1)^1 a_{12}a_{21}a_{33} + (-1)^1 a_{13}a_{22}a_{31} \\ &= (a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}) \\ &\quad - (a_{11}a_{23}a_{32} + a_{13}a_{22}a_{31} + a_{12}a_{21}a_{33}) \end{aligned}$$

que é exatamente a *Regra de Sarrus*[8]. É muito útil ver o determinante como uma função

[8]Pierre Frédéric Sarrus, 1798-1842, matemático framcês.

$\det : F^{n\times n} \to F$ definida como

$$\det \mathbf{A} = \det(\mathbf{L}_1, \ldots, \mathbf{L}_n) \ \text{ ou } \ \det \mathbf{A} = \det(\mathbf{C}_1, \ldots, \mathbf{C}_n),$$

em que $\mathbf{L}_i$ é a i-ésima linha de $\mathbf{A}$ e $\mathbf{C}_j$ é a j-ésima coluna de $\mathbf{A}$.

Lema 1.21 *Seja* $\mathbf{A} \in F^{n\times n}$.

1. *Se* $\mathbf{B}$ *for a matriz obtida de* $\mathbf{A}$ *permutando-se duas linhas* $($*colunas*$)$*, então* $\det \mathbf{B} = -\det \mathbf{A}$.
2. $\det \mathbf{A} = 0$*, quando* $\mathbf{A}$ *possui duas linhas* $($*colunas*$)$ *iguais.*
3. $\det \mathbf{A} = \det \mathbf{A}^t$.

Prova. (1) Seja $\mathbf{B} = (b_{ij})$ a matriz obtida de $\mathbf{A}$ permutando-se as linhas $\mathbf{L}_k$ e $\mathbf{L}_l$, com $k < l$. Então

$$b_{ij} = \begin{cases} a_{ij}, & \text{se } \ i \neq k \ \text{ ou } \ i \neq l \\ a_{lj}, & \text{se } \ i = k \\ a_{kj}, & \text{se } \ i = l. \end{cases}$$

Escolhendo $\tau \in S_n$, com $\tau(k) = l$, $\tau(l) = k$ e $\tau(x) = x$, para todo $x \notin \{k, l\}$, teremos

$$\begin{aligned} \det \mathbf{B} &= \textstyle\sum_{\sigma\in S_n} (\operatorname{sgn} \sigma) b_{1\sigma(1)} \cdots b_{k\sigma(k)} \cdots b_{l\sigma(l)} \cdots b_{n\sigma(n)} \\ &= \textstyle\sum_{\sigma\in S_n} (\operatorname{sgn} \sigma) a_{1\sigma(1)} \cdots a_{l\sigma(k)} \cdots a_{k\sigma(l)} \cdots a_{n\sigma(n)} \\ &= \textstyle\sum_{\sigma\in S_n} (\operatorname{sgn}(\sigma\tau)) a_{1\sigma(\tau(1))} \cdots a_{l\sigma(\tau(l))} \cdots a_{k\sigma(\tau(k))} \cdots a_{n\sigma(\tau(n))} \\ &= -\det \mathbf{A}, \end{aligned}$$

pois $\operatorname{sgn}(\sigma\tau) = -\operatorname{sgn} \sigma$, para todo $\sigma \in S_n$.

(2) Permutando as linhas iguais temos, pelo item (1), que $\det \mathbf{A} = -\det \mathbf{A}$, de modo que $2\det \mathbf{A} = 0$. Portanto, $\det \mathbf{A} = 0$.

(3) Seja $\mathbf{A}^t = (b_{ij})$, com $b_{ij} = a_{ji}$. Como $b_{i\sigma(i)} = a_{\sigma(i)i}$, para todo $\sigma \in S_n$, temos que

$$\begin{aligned} \det \mathbf{A}^t &= \textstyle\sum_{\sigma\in S_n} (\operatorname{sgn} \sigma) \prod_{i=1}^n b_{i\sigma(i)} = \sum_{\sigma\in S_n} (\operatorname{sgn} \sigma) \prod_{i=1}^n a_{\sigma(i)i} \\ &= \textstyle\sum_{\sigma\in S_n} (\operatorname{sgn} \sigma^{-1}) \prod_{i=1}^n a_{i\sigma^{-1}(i)} = \det \mathbf{A}, \end{aligned}$$

pois quando σ percorre todo S_n, σ^{-1} também o faz. ■

Observe que se $n = 3$, então podemos agrupar o $\det \mathbf{A}$ sob a forma:

$$\det \mathbf{A} = a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13} \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}.$$

Mais geralmente:

Teorema 1.22 (Fórmula de Laplace) [9] *Seja* $\mathbf{A} = (a_{ij}) \in F^{n\times n}$. *Então*:

1. *Expansão em relação à* i*-ésima linha de* $\mathbf{A}$*,*

$$\sum_{j=1}^{n} (-1)^{k+j} a_{kj} \det(\mathbf{A}_{kj}) = \delta_{ik} \det \mathbf{A}, \quad i, k = 1, \ldots, n.$$

2. *Expansão em relação à* j*-ésima coluna de* $\mathbf{A}$*,*

$$\sum_{i=1}^{n} (-1)^{i+k} a_{ik} \det(\mathbf{A}_{ik}) = \delta_{jk} \det \mathbf{A}, \quad j, k = 1, \ldots, n,$$

com $\mathbf{A}_{ij}$ *a submatriz de* $\mathbf{A}$ *obtida eliminando-se a* i*-ésima linha e* j*-ésima coluna.*

Prova. (1) Já vimos que cada produto diagonal de $\det \mathbf{A}$ contém exatamente um elemento de cada linha i. Assim, podemos escrever

$$\det \mathbf{A} = a_{i1}c_{i1} + \cdots + a_{in}c_{in}, \quad i = 1, \ldots, n,$$

em que c_{ij} é independente da linha i.

Afirmação. $c_{ij} = (-1)^{i+j} a_{ij} \det(\mathbf{A}_{ij})$.

De fato. Para $i = j = 1$, os termos de $\det \mathbf{A}$ que envolvem a_{11} são aqueles que aparecem na soma

$$\sum_{\sigma \in S_n} \operatorname{sgn} \sigma (a_{11}a_{2\sigma(2)} \cdots a_{n\sigma(n)}), \quad \text{com} \quad \sigma(1) = 1.$$

É fácil ver que esta soma é igual a

$$a_{11} \left(\sum_{\sigma \in S_n} \operatorname{sgn} \sigma (a_{2\sigma(2)} \cdots a_{n\sigma(n)}) \right), \quad \text{com} \quad \sigma(1) = 1,$$

[9] Pierre Simon Marquis de Laplace, 1749-1827, matemático francês.

ou seja, $c_{11} = a_{11}\det(\mathbf{A}_{11})$. Para i e j quaisquer, permutando a linha i sucessivamente com as linhas $i-1, i-2, \ldots, 2, 1$ até que a_{ij} esteja na posição $(1, j)$, por meio de $i-1$ trasposições. É claro que isso não altera a ordem relativa das $n-1$ linhas restantes. Então permutando a coluna j sucessivamente com as colunas $j-1, j-2, \ldots, 2, 1$ até que a_{ij} esteja na posição $(1,1)$, por meio de $j-1$ trasposições. Assim, obtemos uma matriz $\mathbf{B}$ que possui o elemento a_{ij} na posição $(1,1)$. Como a ordem relativa das outras linhas e colunas não foram alteradas temos, pelo item (1) do Lema 1.21, que

$$\det(\mathbf{B}_{11}) = (-1)^{(i-1)+(j-1)}\det(\mathbf{A}_{ij}) = (-1)^{i+j}\det(\mathbf{A}_{ij}).$$

Logo, pelo caso $i = j = 1$, $c_{ij} = a_{ij}\det(\mathbf{B}_{11}) = (-1)^{i+j}a_{ij}\det(\mathbf{A}_{ij})$. Portanto,

$$\det \mathbf{A} = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det(\mathbf{A}_{ij}), \quad i = 1, \ldots, n.$$

Finalmente, seja $\mathbf{B} = (b_{ij})$ a matriz obtida de $\mathbf{A}$ substituindo-se a linha $\mathbf{L}_i$ pela linha $\mathbf{L}_k$. Então

$$b_{ij} = \begin{cases} a_{kj}, & \text{se } i \neq k \\ a_{ij}, & \text{se } i = k. \end{cases}$$

e $\det \mathbf{B} = 0$, pois $\mathbf{B}$ possui duas linha iguais. Logo,

$$0 = \det \mathbf{B} = \sum_{j=1}^{n}(-1)^{i+j}b_{ij}\det(\mathbf{B}_{ij}) = \sum_{j=1}^{n}(-1)^{i+j}a_{kj}\det(\mathbf{A}_{ij}),$$

ou seja,

$$\sum_{j=1}^{n}(-1)^{i+j}a_{kj}\det(\mathbf{A}_{kj}) = \delta_{ik}\det \mathbf{A}, \quad i, k = 1, \ldots, n.$$

(2) Segue do fato de que $\det \mathbf{A} = \det \mathbf{A}^t$. ■

É importante observar que a fórmula de Laplace é um método bastante eficiente para o cálculo do determinante. Além disso, quando $F = \mathbb{R}$, obtemos

$$\frac{\partial \det}{\partial a_{ij}}(\mathbf{A}) = (-1)^{i+j}\det(\mathbf{A}_{ij}), \quad i = 1, \ldots, n.$$

Exemplo 1.23 *Seja* $\mathbf{U} = (a_{ij}) \in F^{n\times n}$ *uma matriz triangular superior. Então*

$$\det \mathbf{U} = a_{11} \cdots a_{nn}.$$

Solução. Vamos usar indução sobre n. Se $n = 1$, nada há para ser provado. Suponhamos que o resultado seja válido para todo k, com $1 \leq k \leq n-1$ e $n > 1$. Assim, pela fórmula de Laplace em relação primeira coluna de $\mathbf{A}$,

$$\det \mathbf{U} = a_{11} \det(\mathbf{U}_{11}) = a_{11}(a_{22} \cdots a_{nn}),$$

pois $\mathbf{U}_{11}$ é uma matriz triangular superior de ordem $n-1$. ■

Teorema 1.24 *Sejam* $\mathbf{A}, \mathbf{B} \in F^{n\times n}$ *e* $c \in F$, *com* $c \neq 0$.

1. *A função determinante é linear em cada linha* (*coluna*).
2. *A função determinante é homogênea em cada linha* (*coluna*).
3. $\det(\mathbf{P_{ij}A}) = -\det \mathbf{A}$.
4. $\det(\mathbf{D}_i(c)\mathbf{A}) = c \det \mathbf{A}$.
5. $\det(\mathbf{T}_{ij}(c)\mathbf{A}) = \det \mathbf{A}$, *se* $i \neq j$.
6. $\det(\mathbf{EA}) = \det \mathbf{E} \cdot \det \mathbf{A}$, *para qualquer matriz elementar* $\mathbf{E} \in F^{n\times n}$.
7. *Se* $\mathbf{A}$ *for equivalente por linha a uma matriz triangular* $\mathbf{U}$, *então* $\det \mathbf{A} = d \det \mathbf{U}$, *para algum* $d \in F$, *com* $d \neq 0$. *Conclua que* $\det \mathbf{A} = 0$ *se, e somente se,* $\det \mathbf{U} = 0$.
8. $\det(\mathbf{AB}) = \det \mathbf{A} \cdot \det \mathbf{B} = \det(\mathbf{BA})$ (***Fórmula de Binet-Cauchy***)[10].
9. *Se* $\mathbf{B}$ *for semelhante a* $\mathbf{A}$, *então* $\det \mathbf{B} = \det \mathbf{A}$.
10. *Fórmulas úteis*:

$$\det \begin{pmatrix} \mathbf{I}_n & \mathbf{O} \\ \mathbf{O} & \mathbf{A} \end{pmatrix} = \det(\mathbf{I}_n \oplus \mathbf{A}) = \det \mathbf{A} \;\; e \;\; \det \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{I}_n \end{pmatrix} = \det \mathbf{A}.$$

[10] Alfred Binet, 1857-1911, pedagogo e psicólogo francês; Augustin-Louis Cauchy, 1789-1857, matemático francês.

Prova. Vamos provar apenas os itens (1), (2) e (8): (1) Seja $\mathbf{C} = (c_{kj})$ a matriz obtida de $\mathbf{A}$ substituindo-se a linha $\mathbf{L}_i$ pela linha $\mathbf{L}_i + \mathbf{L}'_i$. Então

$$c_{kj} = \begin{cases} a_{ij}, & \text{se } i \neq k \\ a_{kj} + a'_{kj}, & \text{se } i = k. \end{cases}$$

Assim, pela fórmula de Laplace em relação à k-ésima linha de $\mathbf{C}$,

$$\begin{aligned} \det \mathbf{C} &= \textstyle\sum_{j=1}^{n} (-1)^{k+j} c_{kj} \det(\mathbf{C}_{kj}) \\ &= \textstyle\sum_{j=1}^{n} (-1)^{k+j} (a_{kj} + a'_{kj}) \det(\mathbf{A}_{kj}) \\ &= \textstyle\sum_{j=1}^{n} (-1)^{k+j} a_{kj} \det(\mathbf{A}_{kj}) + \sum_{j=1}^{n} (-1)^{k+j} a'_{kj} \det(\mathbf{A}_{kj}) \\ &= \det \mathbf{A} + \det \mathbf{A}'. \end{aligned}$$

(2) Seja $\mathbf{C} = (c_{kj})$ a matriz obtida de $\mathbf{A}$ substituindo-se a linha $\mathbf{L}_i$ pela linha $c\mathbf{L}_i$, para algum $c \in F$. Então

$$c_{kj} = \begin{cases} a_{ij}, & \text{se } i \neq k \\ ca_{kj}, & \text{se } i = k. \end{cases}$$

Assim, pela fórmula de Laplace em relação à k-ésima linha de $\mathbf{C}$,

$$\begin{aligned} \det \mathbf{C} &= \textstyle\sum_{j=1}^{n} (-1)^{k+j} c_{kj} \det(\mathbf{C}_{kj}) = \sum_{j=1}^{n} (-1)^{k+j} (ca_{kj}) \det(\mathbf{A}_{kj}) \\ &= c \left(\textstyle\sum_{j=1}^{n} (-1)^{k+j} a_{kj} \det(\mathbf{A}_{kj}) \right) = c \det \mathbf{A}. \end{aligned}$$

(8) É fácil verificar que $\mathbf{E}(\mathbf{AB}) = (\mathbf{EA})\mathbf{B}$. Vamos dividir a prova em dois casos. Se $\mathbf{A}$ for não singular, então $\mathbf{A} = \mathbf{E}_1 \cdots \mathbf{E}_k$. Assim, aplicando recursivamente o item (6), obtemos $\det(\mathbf{AB}) = \det((\mathbf{E}_1 \cdots \mathbf{E}_k)\mathbf{B}) = \det \mathbf{A} \cdot \det \mathbf{B}$. Se $\mathbf{A}$ for singular, então $\mathbf{A}$ é equivalente por linha a uma matriz triangular superior $\mathbf{U}$, com pelo menos um elemento diagonal nulo. Não há perda de generalidade, em supor que $u_{nn} = 0$. Como $\mathbf{AB} = \mathbf{E}_1 \cdots \mathbf{E}_k(\mathbf{UB})$ e $\mathbf{UB}$ possui a última linha nula temos que

$$\det(\mathbf{AB}) = \det(\mathbf{E}_1 \cdots \mathbf{E}_k(\mathbf{UB})) = 0 = \det \mathbf{A} \cdot \det \mathbf{B}.$$

Portanto, em qualquer caso, $\det(\mathbf{AB}) = \det \mathbf{A} \cdot \det \mathbf{B}$. ■

É pertinente, de um ponto de vista teórico, definir formalmente o determinante. Seja $n \in \mathbb{N}$. Um *determinante* é uma função $\det : F^{n \times n} \to F$ que satisfaz os seguintes axiomas:

1. $\det(\mathbf{L}_1, \ldots, \mathbf{L}_i + \mathbf{L}_j, \ldots, \mathbf{L}_n) = \sum_{k=1}^{2} \det(\mathbf{L}_1, \ldots, \mathbf{L}_k, \ldots, \mathbf{L}_n)$, se $i \neq j$.
2. $\det(\mathbf{L}_1, \ldots, c\mathbf{L}_j, \ldots, \mathbf{L}_n) = c \det(\mathbf{L}_1, \ldots, \mathbf{L}_j, \ldots, \mathbf{L}_n)$, para todo $c \in F$.
3. $\det \mathbf{I}_n = 1$.

Observe, por exemplo, que

$$\begin{aligned} \det(\mathbf{L}_1, \ldots, \mathbf{L}_i, \ldots, \mathbf{L}_j, \ldots, \mathbf{L}_n) &= \det(\mathbf{L}_1, \ldots, \mathbf{L}_i, \ldots, \mathbf{L}_i + \mathbf{L}_j, \ldots, \mathbf{L}_n) \\ &= \det(\mathbf{L}_1, \ldots, -\mathbf{L}_j, \ldots, \mathbf{L}_i + \mathbf{L}_j, \ldots, \mathbf{L}_n) \\ &= -\det(\mathbf{L}_1, \ldots, \mathbf{L}_j, \ldots, \mathbf{L}_i, \ldots, \mathbf{L}_n). \end{aligned}$$

Uma maneira de provar que a função $\det$ existe é usando indução sobre n e a expressão

$$\det \mathbf{A} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(\mathbf{A}_{ij}), \quad i = 1, \ldots, n.$$

Por exemplo, $\det \mathbf{I}_n = \det \mathbf{I}_{n-1} = 1$.

A matriz $\mathbf{V}_n = (a_{ij}) \in F^{n \times n}$, com $a_{ij} = x_i^{j-1}$:

$$\mathbf{V}_n = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{pmatrix}$$

chama-se *matriz de Vandermonde*[11] de ordem n.

Exemplo 1.25 *Mostre que*

$$\det \mathbf{V}_n = \prod_{1 \leq i < j \leq n} (x_j - x_i) = \prod_{i=1}^{n-1} \prod_{j=i+1}^{n} (x_j - x_i),$$

o qual chama-se determinante de Vandermonde.

Solução. Vamos usar indução sobre n. Se $n = 2$, nada há para ser provado. Suponhamos que o resultado seja válido para todo k, com $1 \leq k \leq n - 1$ e $n > 2$. Então a sequência de operações elementares

[11] Alexandre-Theóphile Vandermonde, 1735-1796, matemático francês.

$$\mathbf{C}_{j+1} \to \mathbf{C}_{j+1} - x_1\mathbf{C}_j, \;\; j = n-1, \ldots, 2, 1,$$

implicam que $\mathbf{V}_n$ é equivalente por colunas à matriz

$$\mathbf{V}_n' = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & x_2 - x_1 & x_2(x_2 - x_1) & \cdots & x_2^{n-2}(x_2 - x_1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n - x_1 & x_n(x_n - x_1) & \ldots & x_n^{n-2}(x_n - x_1) \end{pmatrix}.$$

Assim, pela fórmula de Laplace em relação à primeira linha, obtemos

$$\det \mathbf{V}_n' = (x_2 - x_1)\cdots(x_n - x_1) \det \begin{pmatrix} 1 & x_2 & x_2^2 & \cdots & x_2^{n-2} \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-2} \end{pmatrix}.$$

Portanto,

$$\det \mathbf{V}_n = \prod_{1<j\leq n} (x_j - x_1) \det \mathbf{V}_{n-1} = \prod_{1\leq i<j\leq n} (x_j - x_i).$$

Note, para $x = x_1$, que $f_1(x) = \det \mathbf{V}_n$ é um polinômio de grau $n-1$, cujas raízes são: $x_2, \ldots, x_n$. Consequentemente, $f_1(x) = a(-1)^{n-1}(x - x_2)\cdots(x - x_n)$, com $a = \det \mathbf{V}_{n-1}$ o coeficiente de $x^{n-1} = x_1^{n-1}$, obtido pela fórmula de Laplace em relação à primeira linha. ■

Seja $\mathbf{T}_n = (a_{ij}) \in F^{n\times n}$. Diremos que $\mathbf{T}_n$ é uma *matriz de Jacobi*[12] ou uma *matriz tridiagonal* se $a_{ij} = 0$, para $|j - i| > 1$. Pondo

$$\begin{array}{ll} a_i = a_{ii}, & \text{se } \; i = 1, \ldots, n \\ b_i = a_{i(i+1)}, & \text{se } \; i = 1, \ldots, n-1 \\ c_i = a_{(i+1)i}, & \text{se } \; i = 1, \ldots, n-1, \end{array}$$

[12]Carl Gustav Jakob Jacobi, 1804-1851, matemático alemão.

obtemos

$$\mathbf{T}_1 = (a_1), \mathbf{T}_2 = \begin{pmatrix} a_1 & b_1 \\ c_1 & a_2 \end{pmatrix} \text{ e } \mathbf{T}_n = \begin{pmatrix} a_1 & b_1 & 0 & 0 & \cdots & 0 & 0 \\ c_1 & a_2 & b_2 & 0 & \cdots & 0 & 0 \\ 0 & c_2 & a_3 & b_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & a_{n-1} & b_{n-1} \\ 0 & 0 & 0 & 0 & \ldots & c_{n-1} & a_n \end{pmatrix}$$

Então, pela fórmula de Laplace em relação à última linha, teremos

$$T_n = a_n T_{n-1} - b_{n-1} c_{n-1} T_{n-2}, \ \forall \ n \in \mathbb{N}, \ \text{com} \ n \geq 3 \ \text{e} \ T_n = \det \mathbf{T}_n.$$

Em particular, se $a_i = b_i = 1$ e $c_i = -1$, obtemos a *sequência de Fibonacci*[13]:

$$F_n = F_{n-1} + F_{n-2}, \ \forall \ n \in \mathbb{N}, \ \text{com} \ n \geq 3.$$

Além disso, se $b_i = 1$ e $c_i = -1$, obtemos o *contínuo* $C_n = a_n C_{n-1} + C_{n-2}$, para todo $n \in \mathbb{N}$, com $n \geq 3$.

Seja $\mathbf{A} = (a_{ij}) \in F^{n \times n}$. O escalar $c_{ij} = (-1)^{i+j} \det(\mathbf{A}_{ij})$ chama-se o *cofator* do termo a_{ij} em $\det \mathbf{A}$, o determinante $\det(\mathbf{A}_{ij})$ chama-se (i, j)-*menor* e a matriz $\mathbf{C} = (c_{ij})$ chama-se a *matriz dos cofatores* de $\mathbf{A}$. A transposta da matriz dos cofatores $\mathbf{C}^t$ chama-se de *adjunta clássica* de $\mathbf{A}$ e denotamos por $\operatorname{adj} \mathbf{A} = \mathbf{C}^t$. Agora, estamos em condições de apresentar um outro método de determinar a inversa de uma matriz. É importante observar que este método é ineficiente quando $n \geq 4$.

Teorema 1.26 *Seja* $\mathbf{A} \in F^{n \times n}$. *Então* $\mathbf{A}(\operatorname{adj} \mathbf{A}) = (\operatorname{adj} \mathbf{A})\mathbf{A} = (\det \mathbf{A})\mathbf{I}_n$.

Prova. Seja $d_{ij} = c_{ji} = (-1)^{i+j} \det(\mathbf{A}_{ji})$ e $\operatorname{adj} \mathbf{A} = (d_{ij})$. Então $\mathbf{A} \operatorname{adj} \mathbf{A} = (p_{ij})$, com

$$p_{ij} = \sum_{k=1}^{n} a_{ik} d_{kj} = \sum_{k=1}^{n} a_{ik} c_{jk} = \delta_{ij} \det \mathbf{A}.$$

Assim, $\mathbf{A}(\operatorname{adj} \mathbf{A}) = \operatorname{diag}(\det \mathbf{A}, \ldots, \det \mathbf{A})$. Portanto, $\mathbf{A}(\operatorname{adj} \mathbf{A}) = (\det \mathbf{A})\mathbf{I}_n$. ■

Corolário 1.27 *Seja* $\mathbf{A} = (a_{ij}) \in F^{n \times n}$. *Então* $\mathbf{A}$ *é não singular se, e somente se,*

[13] Leonardo Fibonacci, 1170-1250, matemático italiano.

$\det \mathbf{A} \neq 0$. *Neste caso,*

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \operatorname{adj} \mathbf{A} \;\; e \;\; a_{ij}^{-1} = \frac{1}{\det \mathbf{A}} c_{ji}.$$

Prova. Fica como um exercício. ■

Exemplo 1.28 *Seja*

$$\mathbf{A} = \begin{pmatrix} 4 & 2 & 0 \\ -1 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}.$$

Determine $\det \mathbf{A}$, *adj*$\mathbf{A}$ *e* $\mathbf{A}^{-1}$.

Solução. Vamos aplicar a fórmula de Laplace em relação à primeira linha de $\mathbf{A}$.

$$\begin{aligned} \det \mathbf{A} &= 4(-1)^{1+1} \det \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} + 2(-1)^{1+2} \det \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix} \\ &+ 0(-1)^{1+3} \det \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix} = 8 + 4 + 0 = 12. \end{aligned}$$

Note que

$$c_{11} = (-1)^{1+1} \det \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} = 2 \text{ e } .c_{12} = (-1)^{1+2} \det \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix} = 2$$

De modo análogo, obtemos $c_{13} = -1, c_{21} = -4, c_{22} = 8$, $c_{23} = -4$, $c_{31} = 0$, $c_{32} = 0$ e $c_{33} = 6$. Assim,

$$\mathbf{C} = \begin{pmatrix} 2 & 2 & -1 \\ -4 & 8 & -4 \\ 0 & 0 & 6 \end{pmatrix} \text{ e } \operatorname{adj}\mathbf{A} = \mathbf{C}^t = \begin{pmatrix} 2 & -4 & 0 \\ 2 & 8 & 0 \\ -1 & -4 & 6 \end{pmatrix}.$$

Finalmente,

$$\mathbf{A}^{-1} = \frac{1}{12} \begin{pmatrix} 2 & -4 & 0 \\ 2 & 8 & 0 \\ -1 & -4 & 6 \end{pmatrix},$$

que é o resultado desejado. ■

Finalizaremos esta seção com uma generalização da fórmula de Laplace. Seja $\mathbf{A} \in F^{m \times n}$. Para cada $r \in \{1, \ldots, \min\{m, n\}\}$, o determinante de uma submatriz de ordem r de $\mathbf{A}$ chama-se um *menor de ordem* r de $\mathbf{A}$ ou, equivalentemente, quando obtido eliminando-se $m-r$ linhas e $n-r$ colunas de $\mathbf{A}$. Mais precisamente, se as linhas e colunas mantidas são:

$$1 \leq i_1 < i_2 < \cdots < i_r \leq m \ \text{ e } \ 1 \leq j_1 < i_2 < \cdots < j_r \leq n,$$

então denotamos por

$$\mathbf{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix} = \det(\mathbf{A}_{rr}),$$

com $\mathbf{A}_{rr} = (a_{i_k j_l})$ a submatriz de ordem r obtida de $\mathbf{A}$ eliminando-se todas as linhas $\mathbf{L}_i$, onde $i \notin \{i_1, \ldots, i_r\}$, e todas as colunas $\mathbf{C}_j$, onde $j \notin \{j_1, \ldots, j_r\}$. Por exemplo, se $m = n$, então

$$\mathbf{A}\begin{pmatrix} 2 & 3 & \cdots & n \\ 1 & 2 & \cdots & n-1 \end{pmatrix} = \det(\mathbf{A}_{1n}).$$

Quando $i_k = j_k, k = 1, \ldots, r$, diremos que $\det(\mathbf{A}_{rr})$ são os *menores principais* de ordem r de $\mathbf{A}$. Em particular, quando $i_k = j_k = k, k = 1, \ldots, r$, diremos que $\det(\mathbf{A}_{rr})$ são os *menores principais líderes* de ordem r de $\mathbf{A}$. Os *menores complementares* de $\mathbf{A}$, com $m = n$, são definidos como

$$\mathbf{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}^c = \mathbf{A}\begin{pmatrix} i_{r+1} & i_{r+2} & \cdots & i_n \\ j_{r+1} & j_{r+2} & \cdots & j_n \end{pmatrix}$$

e os *cofatores complementares* de $\mathbf{A}$ são definidos como

$$\mathbf{A}^c\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix} = (-1)^s \mathbf{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}^c,$$

com $s = (i_1 + i_2 + \cdots + i_r) + (j_1 + j_2 + \cdots + j_r)$. Note que se $r = 1$ a definição reduz a do cofator de um elemento e se $r = n$ é conveniente definir como 1. Por exemplo, se $n = 5$, então

$$\mathbf{A}\begin{pmatrix} 2 & 3 & 5 \\ 1 & 2 & 4 \end{pmatrix}^c = \mathbf{A}\begin{pmatrix} 1 & 4 \\ 3 & 5 \end{pmatrix} = \det\begin{pmatrix} a_{13} & a_{15} \\ a_{43} & a_{45} \end{pmatrix}$$

e

$$\mathbf{A}^c\begin{pmatrix} 2 & 3 & 5 \\ 1 & 2 & 4 \end{pmatrix} = (-1)^{17} \det\begin{pmatrix} a_{13} & a_{15} \\ a_{43} & a_{45} \end{pmatrix} = -\det\begin{pmatrix} a_{13} & a_{15} \\ a_{43} & a_{45} \end{pmatrix}.$$

Teorema 1.29 (Teorema de Laplace) *Seja* $\mathbf{A} \in F^{n\times n}$ *e escolhemos qualquer* r *linhas de* $\mathbf{A}$*. Então*

$$\det \mathbf{A} = \sum_{(j_1,\ldots,j_r)\in Q_{r,n}} \mathbf{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix} \mathbf{A}^c \begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix},$$

onde a soma é estendida sobre todos os $\binom{n}{r}$ *subconjuntos* $Q_{r,n}$ *de* $\{1,\ldots,n\}$*. Em particular, se* $r=1$*, então*

$$\det \mathbf{A} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(\mathbf{A}_{ij}), i = 1,\ldots,n.$$

Prova. Imite a Prova do Teorema 1.22 e/ou confira a Bibliografia. ■

Corolário 1.30 (Fórmula de Binet-Cauchy) *Sejam* $\mathbf{A} \in F^{m\times n}$ *e* $\mathbf{B} \in F^{n\times m}$*. Se* $m \leq n$ *e* $\mathbf{C} = \mathbf{AB}$*, então*

$$\det \mathbf{C} = \sum_{(j_1,\ldots,j_m)\in Q_{m,n}} \mathbf{A}\begin{pmatrix} 1 & 2 & \cdots & m \\ j_1 & j_2 & \cdots & j_m \end{pmatrix} \mathbf{B}\begin{pmatrix} j_1 & j_2 & \cdots & j_m \\ 1 & 2 & \cdots & m \end{pmatrix}.$$

Em particular, se $m = n$*, então* $\det \mathbf{C} = \det \mathbf{A} \det \mathbf{B}$.

Exemplo 1.31 *Seja* $\mathbf{A} \in F^{n\times n}$*, com* $\mathbf{L}_{i_k} = \mathbf{e}_{i_k}, k = 1,\ldots,r < n$*. Mostre que* $\det \mathbf{A}$ *é igual a um menor principal de ordem* r *obtido de* $\mathbf{A}$ *eliminando-se as linhas e as colunas* $i_1,\ldots,i_r$*.*

Solução. Como $\mathbf{L}_{i_k} = \mathbf{e}_{i_k}$ possui 1 na i_k-ésima coluna e as demais zeros temos, pela fórmula de Laplace em relação à i_1-ésima linha de $\mathbf{A}$, que

$$\det \mathbf{A} = (-1)^{i_1+i_1} \mathbf{A}\begin{pmatrix} 1 & 2 & \cdots & i_1-1 & i_1+1 & \cdots & n \\ 1 & 2 & \cdots & i_1-1 & i_1+1 & \cdots & n \end{pmatrix} = \det \mathbf{A}_{i_1 i_1}.$$

Assim, aplicando, indutivamente, esse processo nas linhas i_k de $\mathbf{A}_{i_k i_k}$, para cada $k = 2,\ldots,r$, obtemos o resultado. ■

Exercícios

1. Mostre todas as afirmações deixadas nesta seção.

2. Sejam $\tau \in S_n$, com $\tau(i) = j$, $\tau(j) = i$ e $\tau(x) = x$, para todo $x \notin \{i, j\}$,

$$P = \{\sigma \in S_n : \sigma(i) < \sigma(j)\} \text{ e } I = \{\sigma \in S_n : \sigma(i) > \sigma(j)\}.$$

Mostre que a função $f : P \to I$ definida como $f(\sigma) = \sigma\tau$ é bijetora. Conclua que $S_n = P \cup I$ e $|P| = |I| = \frac{n!}{2}$.

3. Seja $\mathbf{A} \in F^{n\times n}$. Mostre que $\det(c\mathbf{A}) = c^n \det \mathbf{A}$, para todo $c \in F$.

4. Seja $\mathbf{A} \in F^{n\times n}$ idempotente. Mostre que $\det \mathbf{A} = 0$ ou $\det \mathbf{A} = 1$.

5. Seja $\mathbf{A} \in F^{n\times n}$ nilpotente. Mostre que $\det \mathbf{A} = 0$.

6. Sejam $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D} \in F^{n\times n}$, com $\mathbf{A}$ não singular. Use a identidade

$$\mathbf{E} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{C} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{O} & \mathbf{C} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{I} \end{pmatrix} \text{ e } \mathbf{F} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$$

para mostrar que $\det \mathbf{E} = \det \mathbf{A} \det \mathbf{C}$ e $\det \mathbf{F} = \det(\mathbf{AD} - \mathbf{ACA}^{-1}\mathbf{B})$.

7. Sejam $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{2\times 2}$ não singulares e $\mathbf{C} = \mathbf{ABA}^{-1}\mathbf{B}^{-1}$. Mostre que se $\mathbf{AC} = \mathbf{CA}$ e $\mathbf{BC} = \mathbf{CB}$, então $\mathbf{AB} = \mathbf{BA}$ ou $\mathbf{AB} = -\mathbf{BA}$.

8. Sejam $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{2\times 2}$ invertíveis e $\mathbf{C} = \mathbf{AB} - \mathbf{BA}$. Mostre que se $\mathbf{AC} = \mathbf{CA}$ e $\mathbf{BC} = \mathbf{CB}$, então $\mathbf{AB} = \mathbf{BA}$.

9. Seja $\mathbf{A} \in F^{2\times 2}$. Mostre que se existir um $\mathbf{X} \in F^{2\times 2}$, com $\mathbf{X} \neq \mathbf{O}$, tal que $\mathbf{AX} = -\mathbf{XA}$, então $\det \mathbf{A} = 0$ ou $\operatorname{tr}(\mathbf{A}) = 0$. Em particular, sejam

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in \mathbb{C}^{2\times 2}$$

as *matrizes de Pauli*.[14] Se $\mathbf{A} = a_0\mathbf{I} + a_j\sigma_j$, com $j = 1, 2, 3$ e $a_0^2 = a_1^2 + a_2^2 + a_3^2$, então $\mathbf{X} = u(a_0\mathbf{I} - a_j\sigma_j)$, para algum $u \in \mathbb{C}$, e $\mathbf{AX} = \mathbf{XA} = \mathbf{O}$. Se $\mathbf{A} = a_j\sigma_j$, com $j = 1, 2, 3$, então $\mathbf{X} = x_j\sigma_j$, em que $a_1x_1 + a_2x_2 + a_3x_3 = 0$. Note que $\sigma_j^\star = \sigma_j$, ou seja, elas são *matrizes Hermitianas*.

10. Seja $\mathbf{A} \in F^{n\times n}$. Mostre que $\det(x\mathbf{I} + \mathbf{A})$ é um polinômio de grau n sobre F. Conclua que existe um $c \in F$ tal que $c\mathbf{I} + \mathbf{A}$ seja não singular.

[14]Wolfgang Ernst Pauli, 1900-1958, físico austríaco.

11. Sejam $\mathbf{A}$, $\mathbf{B} \in F^{n\times n}$ não singulares. Mostre que $\mathbf{A} + x\mathbf{B}$ é não singular, para todo exceto uma quantidade finita de $x \in F$.

12. Seja $\mathbf{C}_n \in F^{n\times n}$ a *matriz de Frobenius*[15] ou a *matriz companheira*:

$$\mathbf{C}_n = \left(\begin{array}{c|ccc} \mathbf{O} & & \mathbf{I}_{n-1} & \\ \hline -a_0 & -a_1 & \cdots & -a_{n-1} \end{array}\right)$$

Mostre que

$$p(x) = \det(x\mathbf{I} - \mathbf{C}_n) = x^n + a_{n-1}x^{n-1} + \cdots +_2 x^2 + a_1 x + a_0.$$

Note que $\det(\lambda\mathbf{I} - \mathbf{C}_n) = 0$ se, e somente se, λ é uma raiz de $p(x)$. Conclua que para cada polinômio $p(x)$ de grau n existe uma matriz $\mathbf{A} \in F^{n\times n}$ tal que $p(x) = \det(x\mathbf{I} - \mathbf{A})$.

13. Sejam $\mathbf{A} = (a_{ij}) \in F^{n\times n}$ e $p_j(x) = a_{1j} + a_{2j}x + \cdots + a_{nj}x^{n-1}$. Mostre que se $\mathbf{B} = (b_{jk}) \in F^{n\times n}$, com $b_{jk} = p_j(x_k)$, então $\det \mathbf{B} = \det \mathbf{V}_n \det \mathbf{A}$.

14. Seja $f(x) = a_0 + a_1 x + \cdots + a_n x^n \in F[x]$. Mostre que se $f(x)$ possui $n+1$ raízes distintas em F, então $f(x)$ é identicamente nulo.

15. Seja $\mathbf{A}, \mathbf{B} \in F^{n\times n}$. Mostre que $\operatorname{adj}(\mathbf{AB}) = \operatorname{adj}\mathbf{A}\operatorname{adj}\mathbf{B}$, $\det(\operatorname{adj}\mathbf{A}) = (\det\mathbf{A})^{n-1}$ e $\operatorname{adj}(\operatorname{adj}\mathbf{A}) = (\det\mathbf{A})^{n-2}\mathbf{A}$.

16. Seja $\mathbf{A} \in F^{n\times n}$. Diremos que $\mathbf{A}$ é uma *matriz ortogonal* se $\mathbf{A}\mathbf{A}^t = \mathbf{A}^t\mathbf{A} = \mathbf{I}_n$. Mostre que se $\mathbf{A}$ é ortogonal, então $\det\mathbf{A} = \pm 1$.

17. Seja $\mathbf{A} \in \mathbb{C}^{n\times n}$. Diremos que $\mathbf{A}$ é uma *matriz unitária* se $\mathbf{A}\mathbf{A}^* = \mathbf{A}^*\mathbf{A} = \mathbf{I}_n$.
    (a) Mostre que $\det\mathbf{A}^* = \overline{\det\mathbf{A}}$.
    (b) Mostre que se $\mathbf{A}$ é unitária, então $\det\mathbf{A} = e^{i\theta}$, para algum $\theta \in \mathbb{R}$.

18. Mostre que as equações

$$\begin{cases} f(x) = a_0x^2 + a_1x + a_2 &= 0, \quad \text{com } a_0 \neq 0 \\ g(x) = b_0x^2 + b_1x + b_2 &= 0, \quad \text{com } b_0 \neq 0 \end{cases}$$

[15] Ferdinand Georg Frobenius, 1849-1917, matemático alemão.

possui uma raiz comum se, e somente se, o determinante da *matriz de Sylvester* associada a $f(x)$ e $g(x)$

$$\operatorname{res}(f,g)=\begin{pmatrix} a_0 & 0 & b_0 & 0\\ a_1 & a_0 & b_1 & b_0\\ a_2 & a_1 & b_2 & b_1\\ 0 & a_2 & 0 & b_2 \end{pmatrix}$$

é igual a zero. Conclua que as equações

$$\begin{cases} ax^2+x+(a-1) &= 0\\ (b-1)x^2+x+b &= 0 \end{cases}$$

possui uma raiz comum, para todos $a,b\in F$.

19. Seja

$$\mathbf{H}=\begin{pmatrix} 0 & a^2 & c^2 & 1\\ a^2 & 0 & b^2 & 1\\ c^2 & b^2 & 0 & 1\\ 1 & 1 & 1 & 0 \end{pmatrix}\in F^{4\times 4}.$$

Mostre que $\det\mathbf{H}=-16s(s-a)(s-b)(s-c)$, com $s=2^{-1}(a+b+c)$.

20. Sejam $a_i,b_i,c_i,d_i\in F$, com $i=1,\ldots,n$. Mostre que

$$\det\begin{pmatrix} \sum_{i=1}^n a_ic_i & \sum_{i=1}^n a_id_i\\ \sum_{i=1}^n b_ic_i & \sum_{i=1}^n b_id_i \end{pmatrix}=\sum_{1\leq i<j\leq n}\det\begin{pmatrix} a_i & a_j\\ b_i & b_j\end{pmatrix}\det\begin{pmatrix} c_i & c_j\\ d_i & d_j\end{pmatrix}$$

21. (**Identidade de Cauchy**) Sejam $a_i,b_i\in F$, com $i=1,\ldots,n$. Mostre que

$$\det\begin{pmatrix} \sum_{i=1}^n a_i^2 & \sum_{i=1}^n a_ib_i\\ \sum_{i=1}^n a_ib_i & \sum_{i=1}^n b_i^2 \end{pmatrix}=\sum_{1\leq i<j\leq n}\det\begin{pmatrix} a_i & a_j\\ b_i & b_j\end{pmatrix}^2.$$

22. (**Desigualdade de Cauchy-Schwarz**[16]) Sejam $a_i,b_i\in\mathbb{R}$, com $i=1,\ldots,n$. Mostre que

$$(a_1b_1+\cdots+a_nb_n)^2\leq(a_1^2+\cdots+a_n^2)(b_1^2+\cdots+b_n^2).$$

[16] Karl Hermann Amandus Schwarz, 1843-1921, matemático alemão.

Conclua que a igualdade ocorre se, e somente se, existir um $\lambda \in \mathbb{R}$ tal que $b_i = \lambda a_i$, para todo $i = 1, \ldots, n$, ou seja, os pontos (a_i, b_i) estão sobre uma reta vertical.

23. Sejam $a_i \in \mathbb{R}$, com $i = 1, \ldots, n$. Mostre que

$$\frac{1}{n}(a_1 + \cdots + a_n)^2 \leq a_1^2 + \cdots + a_n^2.$$

Mostre, também, que se $a_i > 0$, então $(a_1 + \cdots + a_n)(a_1^{-1} + \cdots + a_n^{-1}) \geq n^2$.

24. Seja $\mathbf{A} \in F^{n \times n}$ não singular. Mostre que $\mathbf{A}$ pode ser transformada em uma matriz diagonal $\mathbf{D}$ usando apenas uma sequância de operações elementares $\mathbf{T}_{ij}(c)$ e $\mathbf{P}_{ij}$ seguida pela multiplicação de uma das linhas por -1.

25. Seja $f : F^{n \times n} \to F$ uma função que satisfaz os seguintes axiomas:
    (a) $f(\mathbf{T}_{ij}(1)\mathbf{A}) = f(\mathbf{A})$, quando $i \neq j$.
    (b) $f(\mathbf{L}_i(c)\mathbf{A}) = cf(\mathbf{A})$, em que $\mathbf{L}_i(c)$ é o esquema $\mathbf{L}_i \to c\mathbf{L}_i$, para todo $c \in F$.
    (c) $f(\mathbf{I}_n) = 1$.

    Mostre que $f(\mathbf{A}) = \det \mathbf{A}$.

26. Seja $f : F^{n \times n} \to F$ uma função tal que

$$f(\mathbf{AB}) = f(\mathbf{A})f(\mathbf{B}), \ \forall \ \mathbf{A}, \mathbf{B} \in F^{n \times n},$$

e existem $\mathbf{X}, \mathbf{Y} \in F^{n \times n}$, com $f(\mathbf{X}) \neq 0$ e $f(\mathbf{Y}) \neq 1$ (f não é constante). Mostre que $f(\mathbf{A}) \neq 0$ se, e somente se, $\mathbf{A}$ for não singular.

27. Mostre que

$$\Delta = \prod_{1 \leq i < j \leq n} (x_j - x_i)^2 = \det \begin{pmatrix} s_0 & s_1 & \cdots & s_{n-1} \\ s_1 & s_2 & \cdots & s_n \\ \vdots & \vdots & \ddots & \vdots \\ s_{n-1} & s_n & \cdots & s_{2n-1} \end{pmatrix},$$

com $s_k = x_1^k + \cdots + x_n^k$, para todo $k \in \mathbb{Z}_+$. Quando os x_i são raízes de um polinômio $p(x)$, diremos que Δ é o *discriminante* de $p(x)$.

28. Seja $\mathbf{A} = \mathbf{I} - \mathbf{U}_n = (a_{ij}) \in F^{n\times n}$, com $a_{ij} = \delta_{ij} - 1$. Mostre que $\det \mathbf{A} = (-1)^n(n-1)$.

2
Equações Lineares

O coração da álgebra linear está em duas operações, ambas com vetores. Adicionamos vetores $\mathbf{u}$ e $\mathbf{v}$ para obter

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

e multiplicamos eles por escalares c e d para obter

$$c\mathbf{u} \text{ e } d\mathbf{v}$$

Combinamos essas operações para obter a combinação linear

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

a qual é uma das mais importantes propriedade neste contexto. Portanto, a *álgebra linear* é o ramos da matemática que trata das propriedades comuns a sistemas algébricos constituídos por um conjunto mais uma razoável noção de uma combinação de vetores do conjunto.

Veremos que a álgebra linear em um nível elementar (espaço de dimensão finita) é basicamente resolver os problemas:

$$\begin{array}{lrcl} (\mathrm{P}_1) & \mathbf{AX} & = & \mathbf{B} \\ (\mathrm{P}_2) & \mathbf{AX} & = & \lambda\mathbf{X} \Leftrightarrow (\lambda\mathbf{I}_n-\mathbf{A})\mathbf{X}=\mathbf{O}. \end{array}$$

A transição entre os problemas P_1 e P_2 é via determinantes:

$$\begin{array}{lcl} \text{(Regra de Cramer)} & \longleftrightarrow & \mathbf{X} = \mathbf{A}^{-1}\mathbf{B} \\ \text{(Polinomial)} & \longleftrightarrow & p(\lambda) = \det(\lambda \mathbf{I}_n - \mathbf{A}). \end{array}$$

2.1 Sistemas de Equações Lineares

O principal objetivo desta seção é um dos aspecto básico da álgebra linear, o qual é o estudo de sistemas de equações lineares.

Uma *equação linear* é uma igualdade cujos membros são polinômios de grau 1 em uma ou mais variáveis ou números. Se o número de variáveis é grande ou indefinido, então o uso de índices se faz necessário no sentido de tornar mais precisa a escrita. Dessa forma, uma equação nas variáveis

$$x_1, x_2, \ldots, x_n$$

é uma igualdade do tipo

$$a + a_1x_1 + a_2x_2 + \cdots + a_nx_n = b + b_1x_1 + b_2x_2 + \cdots + b_nx_n, \tag{2.1}$$

onde $a, a_1, a_2, \ldots, a_n, b, b_1, b_2, \ldots, b_n \in F$ chamam-se *coeficientes das variáveis*. Observe, via operações elementares, que a equação (2.1) pode ser reescrita sob a forma

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b.$$

A expressão à esquerda dessa equação chama-se *forma linear*. Quando se atribuem valores numéricos às variáveis, a equação se transforma em uma igualdade entre números. Se essa igualdade é verdadeira, então a n-upla que satisfaz esta equação

$$(y_1, y_2, \ldots, y_n) \in F^n$$

chama-se *solução* da equação (2.1) ou então que é um ponto do conjunto definido pela equação ou ainda que está no lugar geométrico que ela descreve.

Sejam $m, n \in \mathbb{N}$. Um *sistema de equações lineares* com m equações e n incógnitas é um conjunto de equações lineares sob a forma:

$$\left\{\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1}x_1 & + & \cdots & + & a_{mn}x_n & = & b_m, \end{array}\right. \tag{2.2}$$

ou, abreviadamente, para $\mathbf{x} = (x_1, \ldots, x_n) \in F^n$,

$$f_i(\mathbf{x}) := \sum_{j=1}^{n} a_{ij}x_j \ \text{ e } \ f_i(\mathbf{x}) = b_i, \tag{2.3}$$

onde $a_{ij}, b_i \in F$, $i = 1, \ldots, m, j = 1, \ldots, n$. Seja S_i o conjunto solução da equação (2.3), para cada $i = 1, \ldots, m$. Então uma *solução* do sistema (2.2) é uma n-upla

$$(y_1, \ldots, y_n) \in F^n$$

que satisfaz cada uma das m equações, isto é,

$$(y_1, \ldots, y_n) \in S_1 \cap \cdots \cap S_m.$$

Neste caso, temos n graus de liberdade e no máximo m restrições. Observe que quando $b_1 = b_2 = \cdots = b_m = 0$, diremos que o sistema (2.2) é um *sistema homogêneo*. Neste caso, a n-upla $(0, \ldots, 0)$ é sempre uma solução do sistema homogêneo e chama-se *solução trivial*.

O sistema (2.2) pode ser escrito sob a forma matricial

$$\mathbf{AX} = \mathbf{B} \Leftrightarrow x_1\mathbf{C}_1 + \cdots + x_n\mathbf{C}_n = \mathbf{B}, \tag{2.4}$$

em que

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \ \ \mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \ \text{ e } \ \mathbf{B} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

são a matriz dos coeficientes, a matriz das incógnitas e a matriz dos termos independentes, respectivamente. Neste caso, a equação (2.4) afirma que $\mathbf{B}$ é uma combinação linear das colunas de $\mathbf{A}$. A matriz associada ao sistema (2.2) ou (2.4)

$$\mathbf{A}' = (\mathbf{A} \mid \mathbf{B}) = \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ a_{21} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right)$$

chama-se *matriz ampliada do sistema* ou *matriz aumentada do sistema*. Note que a equação (2.3) pode ser reescrita sob a forma

$$f_i(\mathbf{x}) := \sum_{j=1}^{n} \mathbf{L}_i x_j = b_i, i = 1, \ldots, m. \tag{2.5}$$

O sistema de equações lineares (2.5) chama-se *sistema consistente* se para qualquer escolha de $r_i \in F$ tal que $\sum_{i=1}^m r_i f_i = \mathbf{0}$ se, e somente se, $\sum_{i=1}^m r_i \mathbf{L}_i = \mathbf{0}$, então necessariamente $\sum_{i=1}^m r_i b_i = 0$, ou seja, qualquer dependência linear entre as formas lineares f_i permanece entre os correspondentes membros da direita. Caso contrário, ele chama-se *inconsistente*.

Se o sistema de equações lineares (2.5) possui pelo menos uma solução, então ele é consistente, pois se $\mathbf{Y}$ for uma solução do sistema e $\sum_{i=1}^m r_i \mathbf{L}_i = \mathbf{0}$, então

$$\sum_{i=1}^{m} r_i b_i = \sum_{i=1}^{m} r_i(\mathbf{L}_i \mathbf{Y}) = \sum_{i=1}^{m} (r_i \mathbf{L}_i)\mathbf{Y} = \left(\sum_{i=1}^{m} r_i \mathbf{L}_i\right)\mathbf{Y} = \mathbf{0}\mathbf{Y} = 0.$$

A técnica básica para determinar a solução de um sistema de equações lineares é o método por eliminação ou por determinantes. É importante ressaltar que o método por determinantes emprega fórmulas específicas que são fáceis de entender, de um ponto de vista teórico, mas a computação é bastante tediosa, exceto em casos muito simples. Veremos neste capítulo que uma compreensão elementar das equações lineares e de suas soluções nos ajudará a construir uma teoria dos espaços vetoriais.

Exemplo 2.1 *Resolva o sistema*

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

Solução. Para resolver este sistema vamos usar o sistema e a matriz ampliada ao lado, nos quais vamos aplicar as mesmas operações:

$$\left\{\begin{array}{rcl} x_1+x_2-2x_3 & = & 4 \\ x_1+x_2-x_3 & = & 2 \\ x_1+4x_2-4x_3 & = & 5 \end{array}\right. \qquad \left(\begin{array}{rrr|r} 1 & 1 & -2 & 4 \\ 1 & 1 & -1 & 2 \\ 1 & 4 & -4 & 5 \end{array}\right)$$

Multiplicando a primeira equação por -1 e somando com a segunda, obtemos

$$\left\{\begin{array}{rcr} x_1+x_2-2x_3 & = & 4 \\ x_3 & = & -2 \\ x_1+4x_2-4x_3 & = & 5 \end{array}\right. \qquad \left(\begin{array}{rrr|r} 1 & 1 & -2 & 4 \\ 0 & 0 & 1 & -2 \\ 1 & 4 & -4 & 5 \end{array}\right)$$

Multiplicando a primeira equação por -1 e somando com a terceira, teremos

$$\left\{\begin{array}{rcr} x_1+x_2-2x_3 & = & 4 \\ x_3 & = & -2 \\ 3x_2-2x_3 & = & 1 \end{array}\right. \qquad \left(\begin{array}{rrr|r} 1 & 1 & -2 & 4 \\ 0 & 0 & 1 & -2 \\ 0 & 3 & -2 & 1 \end{array}\right)$$

Trocando-se (ou permutando-se) as equações dois e três, temos

$$\left\{\begin{array}{rcr} x_1+x_2-2x_3 & = & 4 \\ 3x_2-2x_3 & = & 1 \\ x_3 & = & -2 \end{array}\right. \qquad \left(\begin{array}{rrr|r} 1 & 1 & -2 & 4 \\ 0 & 3 & -2 & 1 \\ 0 & 0 & 1 & -2 \end{array}\right)$$

Multiplicando a segunda equação por $\frac{1}{3}$, obtemos

$$\left\{\begin{array}{rcr} x_1+x_2-2x_3 & = & 4 \\ x_2-\frac{2}{3}x_3 & = & \frac{1}{3} \\ x_3 & = & -2 \end{array}\right. \qquad \left(\begin{array}{rrr|r} 1 & 1 & -2 & 4 \\ 0 & 1 & -\frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 1 & -2 \end{array}\right)$$

Multiplicando a terceira equação por 2 e somando com a primeira, teremos

$$\left\{\begin{array}{rcr} x_1+x_2 & = & 0 \\ x_2-\frac{2}{3}x_3 & = & \frac{1}{3} \\ x_3 & = & -2 \end{array}\right. \qquad \left(\begin{array}{rrr|r} 1 & 1 & 0 & 0 \\ 0 & 1 & -\frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 1 & -2 \end{array}\right)$$

Multiplicando a terceira equação por $\frac{2}{3}$ e somando com a segunda, temos

$$\left\{\begin{array}{rcr} x_1 + x_2 & = & 0 \\ x_2 & = & -1 \\ x_3 & = & -2 \end{array}\right. \qquad \left(\begin{array}{ccc|r} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -2 \end{array}\right)$$

Finalmente, multiplicando a segunda equação por -1 e somando com a primeira, obtemos

$$\left\{\begin{array}{ccccr} x_1 & & & = & 1 \\ & x_2 & & = & -1 \\ & & x_3 & = & -2 \end{array}\right. \qquad \left(\begin{array}{ccc|r} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -2 \end{array}\right)$$

Portanto, $(1, -1, -2)$ é a única solução deste sistema. Reciprocamente, é fácil verificar que $(1, -1, -2)$ é a única solução do nosso sistema. ■

O exemplo 2.1 ilustra as manipulações básicas usadas na resolução de um dado sistema de equações lineares por eliminação, as quais reduzem-se ao uso de operações elementares sobre as linhas da matriz ampliada do sistema.

Diremos que dois sistemas de equações lineares são *equivalentes* se eles admitem as mesmas soluções.

Teorema 2.2 *Se um sistema de equações lineares for obtido de outro através de uma sequência finita de operações elementares, então eles são equivalentes.*

Prova. É claro que basta provar que uma operação elementar sempre produz um sistema equivalente. As operações (1) e (2) são facilmente provadas. Suponhamos que $\mathbf{T}_{ij}(c)$, com $i < j$ e $c \neq 0$, seja efetuada. Então o sistema (2.5) pode ser escrito sob a forma

$$(\mathbf{L}_i + c\mathbf{L}_j)\mathbf{X} = b_i + cb_j \text{ e } \mathbf{L}_k\mathbf{X} = b_k, k = 1, \dots i-1, i+1, \dots m. \tag{2.6}$$

Se $\mathbf{Y}$ é uma solução do sistema (2.5), então é claro que $\mathbf{Y}$ também é uma solução do sistema (2.6). Reciprocamente, seja $\mathbf{Y}$ uma solução do sistema (2.6), de modo que

$$(\mathbf{L}_i + c\mathbf{L}_j)\mathbf{Y} = b_i + cb_j \text{ e } \mathbf{L}_k\mathbf{Y} = b_k, k = 1, \dots i-1, i+1, \dots m.$$

Como $(\mathbf{L}_i + c\mathbf{L}_j)\mathbf{Y} = \mathbf{L}_i\mathbf{Y} + c\mathbf{L}_j\mathbf{Y}$ temos que $\mathbf{L}_i\mathbf{Y} = b_i$, pois $c\mathbf{L}_j\mathbf{Y} = cb_j$. Portanto, $\mathbf{Y}$ é uma solução do sistema (2.5). ■

Seja $\mathbf{R} \in F^{m \times n}$. Diremos que $\mathbf{R}$ é *linha reduzida à forma em escada* ou simplesmente em *forma escalonada* se:

1. O primeiro elemento não nulo em cada linha não nula de $\mathbf{R}$ for igual a 1, o qual chama-se *termo líder*.
2. Cada coluna de $\mathbf{R}$ que contém o primeiro elemento não nulo de alguma linha possui todos os outros elementos nulos, a qual chama-se *coluna líder*.
3. Qualquer linha de $\mathbf{R}$ cujos elementos são todos nulos, se existir, ocorre abaixo de todas as linhas que possuem pelo menos um elemento não nulo.
4. Se as linhas $i = 1, \ldots, r$, com $r \leq m$, são as linhas não nulas de $\mathbf{R}$ e se o primeiro elemento não nulo da linha i ocorre na coluna k_i, então
$$k_1 < k_2 < \cdots < k_r.$$

Esta condição impõe a forma escada (ou escada descendente) à matriz. Observe que se $\mathbf{R} = (c_{ij})$, então $\mathbf{R} = \mathbf{O}$ ou existe um inteiro r, com $1 \leq r \leq m$, e existem inteiros $k_1, k_2, \ldots, k_r$, com $1 \leq k_i \leq n$, tais que

- $c_{ij} = 0$ se $i > r$ e $c_{ij} = 0$ se $j < k_i$.
- $c_{ik_j} = \delta_{ij}$, $1 \leq i \leq r$ e $1 \leq j \leq r$.
- $k_1 < k_2 < \cdots < k_r$.

É de grande importância o seguinte: pela condição (4), $\mathbf{R}$ possui exatamente r linhas não nulas. Assim, pela condição (3), $\mathbf{C}_{k_i} = \mathbf{E}_i$ são as colunas líderes de $\mathbf{R}$. Além disso, uma matriz $\mathbf{R}$ está em *forma de degrau* se ela satisfaz às condições (1), (3), (4) e a condição $(2')$ Cada coluna de $\mathbf{R}$ que contém o primeiro elemento não nulo de alguma linha possui todos os elementos abaixo nulos. Portanto, qualquer matriz em forma escalonada está na forma de degrau. Por exemplo, a matriz

$$\mathbf{R}_1 = \begin{pmatrix} 1 & 4 & 1 & 3 \\ 0 & 1 & 2 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

está na forma de degrau, mas não está em forma escalonada. Por outro lado, dadas as matriz

$$\mathbf{R}_2 = \begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 2 \end{pmatrix} \text{ e } \mathbf{R}_3 = \begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & -2 \\ 0 & 1 & 0 & 4 \end{pmatrix},$$

Então $\mathbf{R}_2$ está em forma escalonada e degrau, mas $\mathbf{R}_3$ não está em forma escalonada e nem de degrau, pois $k_1 = 1$, $k_2 = 3$ e $k_3 = 2$ não implica que $k_1 < k_2 < k_3$.

Lema 2.3 (Lema da Substituição) *Sejam $\mathbf{A} = (a_{ij}) \in F^{m\times n}$ e $\mathbf{B} = \mathbf{EA}$, com $\mathbf{E}$ uma operação elementar. Então a j-ésima coluna de $\mathbf{A}$ é uma combinação linear das colunas $k_1, \dots, k_r$ de $\mathbf{A}$ se, e somente se, a j-ésima coluna de $\mathbf{B}$ é uma combinação linear das colunas $k_1, \dots, k_r$ de $\mathbf{B}$. Conclua que os escalares dessas combinações são os mesmos.*

Prova. É claro que as operações (1) e (2) são facilmente provadas. Suponhamos que $\mathbf{T}_{pq}(c)$, com $p \neq q$ e $c \neq 0$, seja efetuada. Então $b_{pj} = a_{pj} + ca_{qj}$ e $b_{ij} = a_{ij}$, se $i \neq p$. Note que a condição: a j-ésima coluna de $\mathbf{A}$ é uma combinação linear das colunas $k_1, \dots, k_r$ de $\mathbf{A}$ é equivalente a: existem $x_1, \dots, x_r \in F$ tais que

$$a_{ij} = \sum_{l=1}^{r} x_l a_{ik_l},\ i = 1, \dots, m, \ \text{ em particular, } \ a_{qj} = \sum_{l=1}^{r} x_l a_{qk_l}.$$

Se $i \neq p$, nada há para ser provado. Se $i = p$, então

$$a_{pj} + ca_{qj} = \sum_{l=1}^{r} x_l (a_{pk_l} + ca_{qk_l}) \Rightarrow b_{pj} = \sum_{l=1}^{r} x_l b_{pk_l},$$

pois $ca_{qj} = c\sum_{l=1}^{r} x_l a_{qk_l}$. A última afirmação será provada no capítulo 3. ■

Teorema 2.4 *Qualquer matriz é equivalente por linha a uma matriz na forma em escada.*

Prova. Seja $\mathbf{A} = (a_{ij}) \in F^{m\times n}$. Se $\mathbf{A} = \mathbf{O}$, nada há para ser provado. Se $\mathbf{A} \neq \mathbf{O}$, então existe um a_{ij} em $\mathbf{A}$ tal que $a_{ij} \neq 0$. Entre todas as linhas de $\mathbf{A}$, escolhemos aquela em que k_1 seja o primeiro j para o qual o elemento $a_{ij} \neq 0$. Assim, permutando a i-ésima linha com a primeira linha ($L_i \leftrightarrow L_1$) movemos o elemento a_{ik_1} para a posição $(1, k_1)$. Multiplicando a primeira linha de $\mathbf{A}$ por $a_{ik_1}^{-1}$, obtemos uma matriz cuja primeira linha é

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & b_{1(k_1+1)} & \cdots & b_{1n} \end{pmatrix},$$

com $b_{1j} = a_{ik_1}^{-1} a_{ij}$, $j = k_1 + 1, \dots, n$. Logo, nesta nova matriz, efetuando a operação

$\mathbf{T}_{i1}(-a_{ik_1})$, obtemos uma matriz sob a forma

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & b_{1(k_1+1)} & \cdots & b_{1n} \\ 0 & \cdots & 0 & 0 & b_{2(k_1+1)} & \cdots & b_{2n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & b_{m(k_1+1)} & \cdots & b_{mn} \end{pmatrix},$$

com $b_{ij} = a_{ij} - a_{ik_1}b_{1j}, i = 2, \ldots, m$ e $j = k_1, \ldots, n$. Se todos os b_{ij} são iguais a 0, acabou. Se algum $b_{ij} \neq 0$, então o processo acima pode ser repetido, obtendo uma matriz sob a forma

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & b_{1(k_1+1)} & \cdots & b_{1(k_2-1)} & 0 & c_{1(k_2+1)} & \cdots & c_{1n} \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & c_{2(k_2+1)} & \cdots & c_{2n} \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & c_{3(k_2+1)} & \cdots & c_{3n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & c_{m(k_2+1)} & \cdots & c_{mn} \end{pmatrix}.$$

E assim sucessivamente. ■

Corolário 2.5 *Se $\mathbf{R}$ e $\mathbf{S}$ são formas escalonadas de $\mathbf{A}$, então, a menos da sequência de operações elementares sobre as linhas, $\mathbf{R} = \mathbf{S}$.*

Prova. Como o procedimento para obter $\mathbf{R}$ e $\mathbf{S}$ não altera as colunas nulas (zeros) temos que elas possuem as mesmas colunas nulas. Assim, $\mathbf{R}$ e $\mathbf{S}$ possuem as colunas líderes $\mathbf{E}_1, \mathbf{E}_2, \ldots$, ocorrendo pela primeira vez nas mesmas posições, digamos, $k_1, k_2, \ldots, k_r$. Logo, pelo Lema 2.3, cada coluna $\mathbf{C}_j$ de $\mathbf{R}$ e de $\mathbf{S}$ ocorrendo entre k_i e k_{i+1} ou fora é uma combinação linear das colunas $\mathbf{E}_{k_1}, \ldots, \mathbf{E}_{k_r}$. Portanto, pelo Lema 2.3, $\mathbf{R} = \mathbf{S}$. ■

Proposição 2.6 *Qualquer $\mathbf{A} \in F^{m\times n}$ é equivalente a uma matriz sob a forma normal*

$$\mathbf{N}_r = \begin{pmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}, \quad \textit{com} \ \ r \leq \min\{m, n\}.$$

Prova. Com objetivos didáticos apresentaremos a prova. Seja $\mathbf{A} = (a_{ij}) \in F^{m\times n}$. Se $\mathbf{A} = \mathbf{O}$, nada há para ser provado. Se $\mathbf{A} \neq \mathbf{O}$, então existe um a_{ij} em $\mathbf{A}$ tal que $a_{ij} \neq 0$. Então permutando a i-ésima linha com a primeira linha ($L_i \leftrightarrow L_1$) e a j-ésima coluna com a primeira coluna ($C_j \leftrightarrow C_1$) movemos o elemento a_{ij} para a posição $(1, 1)$.

Assim, efetuando a operação $\mathbf{D}_1(a_{ij}^{-1})$, obtemos uma matriz cuja primeira linha é

$$\begin{pmatrix} 1 & b_{12} & \cdots & b_{1n} \end{pmatrix}.$$

Efetuando a sequência de operações $\mathbf{T}_{i1}(-a_{ij})$ à esquerda e $\mathbf{T}_{j1}(-a_{ij})$ à direita, obtemos uma matriz sob a forma

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & b_{m2} & \cdots & b_{mn} \end{pmatrix}.$$

Se todos os b_{ij} são iguais a 0, acabou. Se algum $b_{ij} \neq 0$, então o processo acima pode ser repetido com a submatriz (b_{ij}). E assim sucessivamente. ■

Teorema 2.7 *Sejam* $\mathbf{AX} = \mathbf{B}$ *e* $\mathbf{CX} = \mathbf{D}$ *sistemas. Então eles são equivalentes se suas matrizes ampliadas são equivalentes por linhas.*

Prova. Suponhamos que as matrizes ampliadas sejam equivalentes por linhas. Então, pela Proposição 1.11, existe uma matriz invertível $\mathbf{P} \in F^{m \times m}$ tal que

$$(\mathbf{PA} \mid \mathbf{PB}) = \mathbf{P}(\mathbf{A} \mid \mathbf{B}) = (\mathbf{C} \mid \mathbf{D}).$$

Assim, $\mathbf{PA} = \mathbf{C}$ e $\mathbf{PB} = \mathbf{D}$. Logo, multiplicando ambos os membros do sistema $\mathbf{AX} = \mathbf{B}$ à esquerda por $\mathbf{P}$, obtemos $\mathbf{PAX} = \mathbf{PB}$ se, e somente se, $\mathbf{CX} = \mathbf{D}$. Se $\mathbf{X}$ for qualquer solução do sistema $\mathbf{AX} = \mathbf{B}$, então ela é também uma solução do sistema $\mathbf{CX} = \mathbf{D}$. Reciprocamente, se $\mathbf{X}$ for qualquer solução do sistema $\mathbf{CX} = \mathbf{D}$, então

$$\mathbf{AX} = \mathbf{P}^{-1}(\mathbf{PA})\mathbf{X} = \mathbf{P}^{-1}(\mathbf{CX}) = \mathbf{P}^{-1}\mathbf{D} = \mathbf{B},$$

ou seja, ela é também uma solução do sistema $\mathbf{AX} = \mathbf{B}$. ■

É muito útil, de um ponto de vista teórico e didático, a seguinte simplificação de notação. Se o sistema $\mathbf{AX} = \mathbf{B}$ for equivalente ao sistema $\mathbf{RX} = \mathbf{C}$, em que $\mathbf{R}$ é a matriz em forma escalonada de $\mathbf{A}$, então renumerando, se necessário, as variáveis podemos supor que

$$\mathbf{R} = \left(\begin{array}{c|c} \mathbf{I}_r & \mathbf{D} \\ \hline \mathbf{O} & \mathbf{O} \end{array}\right), \quad \text{com } r \leq \min\{m, n\}.$$

Mais precisamente, se efetuamos a operação $\mathbf{P} = \mathbf{P_{ij}}$ sobre $\mathbf{X}$, então

$$\mathbf{Y} = \mathbf{PX} = \begin{pmatrix} x_1 & \cdots & x_j & \cdots & x_i & \cdots & x_n \end{pmatrix}^t$$

e

$$\mathbf{D} = \mathbf{AP} = \begin{pmatrix} \mathbf{C}_1 & \cdots & \mathbf{C}_j & \cdots & \mathbf{C}_i & \cdots & \mathbf{C}_n \end{pmatrix}.$$

Como $\mathbf{P}^2 = \mathbf{I}$ temos que

$$\mathbf{AX} = \mathbf{AP}^2\mathbf{X} = (\mathbf{AP})(\mathbf{PX}) = \mathbf{DY}.$$

Portanto, indutivamente, obtemos a forma desejada.

O Teorema 2.7 nos fornece um método simples para analisar e resolver o sistema $\mathbf{AX} = \mathbf{B}$, quando $\mathbf{C} = \mathbf{R}$ é a matriz em forma escalonada de $\mathbf{A}$, e conhecido como o método da *redução de Gauss-Jordan*.[1] Por outro lado, uma maneira mais eficiente de resolver o sistema $\mathbf{AX} = \mathbf{B}$ é o método da *substituição reversa* ou o método da *eliminação Gaussiana*, em que $\mathbf{C} = \mathbf{R}$ é a matriz em forma de degrau. Neste caso, se $\mathbf{A} = \mathbf{LU}$, então podemos encontrar $\mathbf{X}$ resolvendo $\mathbf{LY} = \mathbf{B}$ e depois $\mathbf{UX} = \mathbf{Y}$

Exemplo 2.8 *Resolva o sistema*

$$\left\{\begin{array}{rcr} x_1 + 2x_2 + 3x_3 & = & 5 \\ 2x_1 + 5x_2 + 3x_3 & = & 3 \\ x_1 + 0x_2 + 8x_3 & = & 17 \end{array}\right. \Leftrightarrow x_1\mathbf{C}_1 + x_2\mathbf{C}_2 + x_3\mathbf{C}_3 = \mathbf{B}.$$

Solução. Para resolver este sistema vamos usar o método da substituição reversa. Para isto, basta reduzir a matriz ampliada em forma de degrau:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 5 \\ 2 & 5 & 3 & 3 \\ 1 & 0 & 8 & 17 \end{array}\right) \longrightarrow \cdots \longrightarrow \mathbf{R} = \left(\begin{array}{ccr|r} 1 & 2 & 3 & 5 \\ 0 & 1 & -3 & -7 \\ 0 & 0 & 1 & 2 \end{array}\right)$$

Assim, $x_3 = 2$, então $x_2 = -7 + 3x_3 = -1$. Finalmente, $x_1 = 5 - 2x_2 - 3x_3 = 1$. Portanto, $(1, -1, 2)$ é a única solução do nosso sistema. Neste caso, a matriz $\mathbf{B}$ pode ser escrita de modo único como uma combinação linear das colunas de $\mathbf{A}$. ■

[1] Johann Carl Friedrich Gauss, 1777-1855, matemático alemão; Camille Marie Ennemond Jordan. 1838-1922, matemático francês.

Exemplo 2.9 *Resolva o sistema*

$$\left\{\begin{array}{rcl} x_1+x_2+2x_3 & = & 3 \\ x_1+0x_2+x_3 & = & 5 \\ 2x_1+x_2+3x_3 & = & 8 \end{array}\right. \Leftrightarrow x_1\mathbf{C}_1+x_2\mathbf{C}_2+x_3\mathbf{C}_3=\mathbf{B}.$$

Solução. Basta reduzir a matriz ampliada em forma de degrau:

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 1 & 0 & 1 & 5 \\ 2 & 1 & 3 & 8 \end{array}\right) \longrightarrow \cdots \longrightarrow \mathbf{R}=\left(\begin{array}{rrr|r} 1 & 1 & 2 & 3 \\ 0 & -1 & -1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Assim, $x_2=-2-x_3$, então $x_1=3-x_2-2x_3=5-x_3$. Logo, pondo $x_3=t$,

$$S=\{(5-t,-2-t,t):t\in F\}$$

é a *descrição paramétrica* do conjunto de todas as soluções do nosso sistema. Portanto, o nosso sistema possui infinitas soluções. Neste caso, a matriz $\mathbf{B}$ pode ser escrita de várias maneiras como uma combinação linear das colunas de $\mathbf{A}$. ■

Exemplo 2.10 *Resolva o sistema*

$$\left\{\begin{array}{rcr} x_1-x_2+2x_3 & = & 3 \\ x_1+2x_2-x_3 & = & -3 \\ 0x_1+2x_2-2x_3 & = & 1 \end{array}\right. \Leftrightarrow x_1\mathbf{C}_1+x_2\mathbf{C}_2+x_3\mathbf{C}_3=\mathbf{B}.$$

Solução. Basta reduzir a matriz ampliada em forma de degrau:

$$\left(\begin{array}{rrr|r} 1 & -1 & 2 & 3 \\ 1 & 2 & -1 & -3 \\ 0 & 2 & -2 & 1 \end{array}\right) \longrightarrow \cdots \longrightarrow \mathbf{R}=\left(\begin{array}{rrr|r} 1 & -1 & 2 & 3 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 5 \end{array}\right)$$

o que é impossível, pois a equação $0x_1+0x_2+0x_3=5$ é falsa. Portanto, o nosso sistema não possui solução. Neste caso, a matriz $\mathbf{B}$ não pode ser escrita como uma combinação linear das colunas de $\mathbf{A}$. ■

Exemplo 2.11 *Sejam*

$$\left\{\begin{array}{l} a_{11}x_1+a_{12}x_2=0 \\ a_{21}x_1+a_{22}x_2=0 \end{array}\right. \quad e \quad \left\{\begin{array}{l} b_{11}x_1+b_{12}x_2=0 \\ b_{21}x_1+b_{22}x_2=0 \end{array}\right.$$

sistemas homogêneos. Mostre que se eles são equivalentes, então cada equação do segundo é uma combinação linear das equações do primeiro e vice-versa.

Solução. Para resolver basta provar que os sistemas

$$(1)\ \left\{\begin{array}{l} a_{11}y_1 + a_{12}y_2 = b_{11} \\ a_{21}y_1 + a_{22}y_2 = b_{12} \end{array}\right. \quad \text{e} \quad (2)\ \left\{\begin{array}{l} a_{11}z_1 + a_{12}z_2 = b_{21} \\ a_{21}z_1 + a_{22}z_2 = b_{22} \end{array}\right.$$

possuem soluções. Se $a_{ij} = 0$, para todos $i, j \in \{1, 2\}$, nada há para ser provado. Suponhamos que existam $i, j \in \{1, 2\}$ tais que $a_{ij} \neq 0$, digamos $a_{11} \neq 0$. Então

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \xrightarrow{L_1 \to a_{11}^{-1}L_1} \begin{pmatrix} 1 & a_{11}^{-1}a_{12} \\ a_{21} & a_{22} \end{pmatrix} \xrightarrow{L_2 \to L_2 - a_{21}L_1} \begin{pmatrix} 1 & a_{11}^{-1}a_{12} \\ 0 & a_{11}^{-1}(a_{11}a_{22} - a_{21}a_{12}) \end{pmatrix}.$$

Assim, nosso sistema é equivalente ao sistema

$$\left\{\begin{array}{r} x_1 + a_{11}^{-1}a_{12}x_2 = 0 \\ a_{11}^{-1}(a_{11}a_{22} - a_{21}a_{12})x_2 = 0 \end{array}\right.$$

Se $a_{11}a_{22} - a_{21}a_{12} \neq 0$, então, pela Regra de Cramer, os sistemas (1) e (2) possuem soluções (únicas). Se $a_{11}a_{22} - a_{21}a_{12} = 0$, então as soluções do nosso sistema são da forma

$$S = \{(-a_{12}, a_{11})t : t \in F\}.$$

Considerando as matrizes ampliadas dos sistemas (1) e (2) e usando as mesmas operações, obtemos

$$\begin{array}{l} (1)\ \left\{\begin{array}{r} y_1 + a_{11}^{-1}a_{12}y_2 = a_{11}^{-1}b_{11} \\ a_{11}^{-1}(a_{11}a_{22} - a_{21}a_{12})y_2 = b_{12} - a_{11}^{-1}a_{12}b_{11} \end{array}\right. \\ (2)\ \left\{\begin{array}{r} z_1 + a_{11}^{-1}a_{12}z_2 = a_{11}^{-1}b_{21} \\ a_{11}^{-1}(a_{11}a_{22} - a_{21}a_{12})z_2 = b_{22} - a_{11}^{-1}a_{12}b_{21} \end{array}\right. \end{array}$$

Como $(-a_{12}, a_{11})$ é uma solução do nosso sistema temos que os sistemas (1) e (2) possuem soluções

$$S_1 = \{(b_{11} - a_{12}, a_{11})s : s \in F\} \text{ e } S_2 = \{(b_{21} - a_{12}, a_{11})t : t \in F\},$$

respectivamente. ■

O restante desta seção será dedicada a discussão dos sistemas de equações lineares. Sejam $\mathbf{A} \in F^{m \times n}$ e $\mathbf{R}$ a matriz em forma escalonada de $\mathbf{A}$. O *posto linha* de $\mathbf{A}$ e denotamos por $\rho_l(\mathbf{A})$, é o número de linhas não nulas de $\mathbf{R}$ o qual, pelo Corolário 2.5, está bem definido. A *nulidade* de $\mathbf{A}$ e denotamos por $\nu(\mathbf{A})$, é igual

$$\nu(\mathbf{A}) = n - \rho_l(\mathbf{A})$$

e como veremos é o número de variáveis livres ou o grau de liberdade. Observe que somente a matriz nula possui posto linha zero. Mas, qualquer outra matriz possui posto linha pelo menos 1 e no máximo m. Por outro lado, cada linha não nula de $\mathbf{R}$ define um termo líder e, portanto, uma coluna líder, de modo que o posto linha é também no máximo n. Consequentemente, $0 \leq \rho_l(\mathbf{A}) \leq \min\{m, n\}$. Em particular, $\rho_l(\mathbf{N}_r) = r$. Por exemplo, o posto linha e a nulidade da matriz

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 & 0 \\ -1 & 0 & 3 & 5 \\ 1 & -2 & 1 & 1 \end{pmatrix}$$

são $\rho_1(\mathbf{A}) = 3$ e $\nu(\mathbf{A}) = 4 - 3 = 1$, pois

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 & 0 \\ -1 & 0 & 3 & 5 \\ 1 & -2 & 1 & 1 \end{pmatrix} \longrightarrow \cdots \longrightarrow \mathbf{R} = \begin{pmatrix} 1 & 0 & 0 & -\frac{7}{8} \\ 0 & 1 & 0 & -\frac{1}{4} \\ 0 & 0 & 1 & \frac{11}{8} \end{pmatrix}.$$

Teorema 2.12 *O sistema homogêneo $\mathbf{AX} = \mathbf{O}$ possui uma solução não trivial se, e somente se, $\rho_1(\mathbf{A}) < n$.*

Prova. Suponhamos que $\mathbf{A} \in F^{m \times n}$, $\mathbf{R}$ seja a matriz em forma escalonada de $\mathbf{A}$ e $r = \rho_1(\mathbf{A}) < n$. Então já vimos que

$$\mathbf{R} = \left(\begin{array}{c|c} \mathbf{I}_r & \mathbf{C} \\ \hline \mathbf{O} & \mathbf{O} \end{array}\right).$$

Como $r < n$ e existe pelo menos $r + 1$ colunas temos que a matriz $\mathbf{C}$ possui pelo menos uma coluna. Assim,

$$\mathbf{RX} = \left(\begin{array}{c|c} \mathbf{I}_r & \mathbf{C} \\ \hline \mathbf{O} & \mathbf{O} \end{array}\right) \left(\begin{array}{c} \mathbf{Z} \\ \hline \mathbf{Y} \end{array}\right)$$

possui uma solução não trivial

$$\mathbf{X} = \left(\frac{-\mathbf{CY}}{\mathbf{Y}} \right), \quad \text{para cada escolha de } \mathbf{Y} \neq \mathbf{O}.$$

Reciprocamente, se $r = \rho_1(\mathbf{A}) = n$, então

$$\mathbf{RX} = \left(\frac{\mathbf{I}_n}{\mathbf{O}} \right) \mathbf{X} = \mathbf{O}.$$

Logo, a única solução possível do sistema é a trivial $\mathbf{X} = \mathbf{O}$. ■

Seja $\mathbf{A} \in F^{m \times n}$. Diremos que as colunas de $\mathbf{A}$ são *linearmente dependentes* (LD) se o sistema homogêneo $\mathbf{AX} = \mathbf{O}$ possui pelo menos uma solução não trivial. Caso contrário, diremos que as colunas de $\mathbf{A}$ são *linearmente independentes* (LI). O conjunto

$$C = \{x_1\mathbf{C}_1 + \cdots + x_n\mathbf{C}_n : x_i \in F\}$$

gerado pelas colunas de $\mathbf{A}$ chama-se *espaço coluna*. Neste caso, o número de colunas LI de $\mathbf{A}$ chama-se *posto coluna* de $\mathbf{A}$ ou *dimensão* de C e denotamos por $\rho_c(\mathbf{A})$ ou $\dim(C)$. De modo análogo, o conjunto

$$L = \{x_1\mathbf{L}_1 + \cdots + x_m\mathbf{L}_m : x_i \in F\}$$

gerado pelas linhas de $\mathbf{A}$ chama-se *espaço linha*. Neste caso, $\dim(L) = \rho_l(\mathbf{A})$.

Lema 2.13 *Sejam* $\mathbf{A} \in F^{m \times n}$ *e* $\mathbf{X} \in F^{n \times 1}$.

1. $\mathbf{X}^t\mathbf{X} = 0$ *se, e somente se,* $\mathbf{X} = \mathbf{O}$.
2. $\mathbf{AX} = \mathbf{O}$ *se, e somente se,* $\mathbf{A}^t\mathbf{AX} = \mathbf{O}$.
3. $\rho_c(\mathbf{A}) \leq n$ *e* $\rho_c(\mathbf{A}) = \rho_c(\mathbf{A}^t\mathbf{A})$.
4. $\rho_c(\mathbf{A}) = \rho_c(\mathbf{A}^t) = \rho_l(\mathbf{A})$.

Prova. Vamos provar apenas o item (4). Já sabemos que

$$\mathbf{A}^\mathbf{t}\mathbf{A} = \left(\ \mathbf{A}^\mathbf{t}\mathbf{C}_1 \ \middle| \ \cdots \ \middle| \ \mathbf{A}^\mathbf{t}\mathbf{C}_n \ \right),$$

isto é, a j-ésima coluna de $\mathbf{A^t A}$ é igual a $\mathbf{A^t C}_j$, a qual é uma combinação linear das colunas de $\mathbf{A}^t$. Assim, pelo item (3),

$$\begin{aligned}
\rho_c(\mathbf{A}) &= \rho_c(\mathbf{A}^t\mathbf{A}) = \dim\{x_1(A^t\mathbf{C}_1) + \cdots + x_n(A^t\mathbf{C}_n) : x_i \in F\} \\
&\leq \dim\{A^t\mathbf{Y}_1 + \cdots + A^t\mathbf{Y}_n : \mathbf{Y}_i = x_i\mathbf{C}_i \in F^{m\times 1}\} \\
&= \rho_c(\mathbf{A}^t).
\end{aligned}$$

Por simetria, $\rho_c(\mathbf{A}^t) \leq \rho_c((\mathbf{A}^t)^t) = \rho_c(\mathbf{A})$. Portanto, $\rho_c(\mathbf{A}) = \rho_c(\mathbf{A}^t) = \rho_l(\mathbf{A})$. ■

Seja $\mathbf{A} \in F^{m\times n}$. O *posto* de $\mathbf{A}$ é definido como $\rho_c(\mathbf{A}) = \rho_l(\mathbf{A})$ e denotamos simplesmente por $\rho(\mathbf{A})$. Neste caso, $\nu(\mathbf{A})$ é igual ao número de colunas não líderes de $\mathbf{R}$. Em particular, se $\rho(\mathbf{A}) = n$, então a forma escalonada de $\mathbf{A}$ é

$$\mathbf{R} = \begin{pmatrix} \mathbf{E}_1 & \cdots & \mathbf{E}_n \end{pmatrix}$$

e $\mathbf{R} = \mathbf{I}_n$ quando $m = n$. Observe que $\rho(\mathbf{A}) = r$ se, e somente se, r é o maior número natural tal que o menor

$$\mathbf{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix} = \det(\mathbf{A}_{rr}) \neq 0,$$

ou seja, pelo menos um dos seus $\nu(\mathbf{A}) = n - r$ menores é não nulo, enquanto todos os seus $\nu(\mathbf{A}) - 1 = n - (r+1)$ menores são nulos.

Teorema 2.14 *Sejam $\mathbf{A} \in F^{m\times n}$, $\mathbf{AX} = \mathbf{B}$ um sistema e $\mathbf{A}'$ sua matriz ampliada. Então o sistema é consistente se, e somente se, $\rho(\mathbf{A}) = \rho(\mathbf{A}')$.*

Prova. Suponhamos que $\rho(\mathbf{A}) = r$ e que $\mathbf{R}$ seja a matriz em forma escalonada de $\mathbf{A}$. Então

$$\rho(\mathbf{A} \mid \mathbf{B}) = \rho(\mathbf{R} \mid \mathbf{C}) = r \text{ ou } r+1.$$

Como

$$(\mathbf{R} \mid \mathbf{C}) = \left(\begin{array}{ccc|ccc|c}
1 & \cdots & 0 & * & \cdots & * & c_1 \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & \cdots & 1 & * & \cdots & * & c_r \\
\hline
0 & \cdots & 0 & 0 & \cdots & 0 & c \\
0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & \cdots & 0 & 0 & \cdots & 0 & 0
\end{array}\right)$$

temos que o sistema é inconsistente se, e somente se, $c \neq 0$. Mas, $c \neq 0$ se, e somente se, $\rho(\mathbf{A} \mid \mathbf{B}) = r + 1$. Assim, $\mathbf{C} = \mathbf{E}_{r+1}$, pois $(\mathbf{R} \mid \mathbf{C})$ está em forma escalonada. Portanto, o sistema é consistente se, e somente se, $\rho(\mathbf{A}) = \rho(\mathbf{A}')$. ■

O Teorema 2.14 nos fornece um método simples para **discutir** o sistema não homogêneo $\mathbf{AX} = \mathbf{B}$:

- Se $\rho(\mathbf{A}) = \rho(\mathbf{A}') = n$, então o sistema possui uma única solução. O sistema é consistente e *determinado*.
- Se $\rho(\mathbf{A}) = \rho(\mathbf{A}') < n$, então o sistema possui infinitas soluções. Neste caso, existem $\nu(\mathbf{A}) = n - \rho(\mathbf{A})$ variáveis livres. O sistema é consistente e *indeterminado*.
- Se $\rho(\mathbf{A}) < \rho(\mathbf{A}')$, então o sistema é inconsistente, ou seja, não possui solução.

Exemplo 2.15 *Discuta e resolva o sistema*

$$\begin{cases} x_1 - 2x_2 + 3x_3 = 0 \\ 2x_1 - 5x_2 + 6x_3 = 0 \end{cases}$$

Solução. A matriz ampliada do sistema é

$$\left(\begin{array}{ccc|c} 1 & -2 & 3 & 0 \\ 2 & -5 & 6 & 0 \end{array}\right) \longrightarrow \cdots \longrightarrow \mathbf{R} = \left(\begin{array}{ccc|c} 1 & 0 & 3 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right).$$

Assim, nosso sistema é equivalente ao sistema

$$\begin{cases} x_1 + 3x_3 = 0 \\ x_2 = 0 \end{cases}$$

Como $\rho(\mathbf{A}) = \rho(\mathbf{A}') = 2$ temos que o sistema é consistente. Sendo $\rho(\mathbf{A}) < 3$, teremos $\nu(\mathbf{A}) = 3 - 2 = 1$ variáveis livres. Portanto, $S = \{(-3, 0, 1)t : t \in F\}$ é a *descrição paramétrica* do conjunto solução do sistema. ■

JÃ¡ vimos que uma matriz $\mathbf{A} \in F^{n\times n}$ era não singular se existisse uma matriz $\mathbf{X} \in F^{n\times n}$ tal que $\mathbf{AX} = \mathbf{I}$. Mas isto significa que devemos resolver simultaneamente n sistemas

$$\mathbf{AX}_j = \mathbf{E}_j, \quad j = 1, \ldots, n,$$

ou seja, escalonando cada matriz

$$\left(\begin{array}{c|c} \mathbf{A} & \mathbf{E}_1 \end{array}\right), \left(\begin{array}{c|c} \mathbf{A} & \mathbf{E}_2 \end{array}\right), \ldots, \left(\begin{array}{c|c} \mathbf{A} & \mathbf{E}_n \end{array}\right).$$

Em cada caso, efetuamos as mesmas operações elementares sobre $\mathbf{A}$. Isto sugere um método muito eficiente de resolver e analisar uma sequência finita de sistemas de equações lineares, o qual é muito útil no *Princípio da Superposição*:

$$\mathbf{AX} = \mathbf{B}_1, \ldots, \mathbf{AX} = \mathbf{B}_k,$$

a saber, escalonando a matriz aumentada

$$\left(\begin{array}{c|c|c|c} \mathbf{A} & \mathbf{B}_1 & \cdots & \mathbf{B}_k \end{array}\right).$$

Exemplo 2.16 *Resolva os sistemas*

$$\left\{\begin{array}{rcr} x_1 + 3x_2 & = & -9 \\ -4x_1 - 5x_2 & = & 1 \end{array}\right. \quad e \quad \left\{\begin{array}{rcr} y_1 + 3y_2 & = & -5 \\ -4y_1 - 5y_2 & = & -1 \end{array}\right.$$

Solução. Pelo exposto acima, basta escalonar a matriz

$$\left(\begin{array}{rr|r|r} 1 & 3 & -9 & -5 \\ -4 & -5 & 1 & -1 \end{array}\right) \longrightarrow \cdots \longrightarrow \left(\begin{array}{rr|r|r} 1 & 0 & 6 & 4 \\ 0 & 1 & -5 & -3 \end{array}\right).$$

Assim, $(6, -5)$ e $(4, -3)$ são as soluções dos sistemas. ■

Teorema 2.17 (Teorema da Matriz Inversa) *Seja $\mathbf{A} \in \mathbb{R}^{n\times n}$. Então as seguintes condições são equivalentes*:

1. $\mathbf{A}$ *é invertível*;
2. $\mathbf{AX} = \mathbf{O}$ *possui somente a solução trivial*;
3. $\rho(\mathbf{A}) = n$ *e* $\nu(\mathbf{A}) = 0$;
4. $\mathbf{A}$ *pode ser escrita como um produto de matrizes elementares*;
5. *Para cada* $\mathbf{B} \in \mathbb{R}^{n\times 1}$*, o sistema* $\mathbf{AX} = \mathbf{B}$ *possui uma solução, ou seja,* $\mathbf{B}$ *pode ser escrita como uma combinação linear das colunas de* $\mathbf{A}$;
6. *Para cada* $\mathbf{B} \in \mathbb{R}^{n\times 1}$*, o sistema* $\mathbf{AX} = \mathbf{B}$ *possui exatamente uma solução, ou seja,* $\mathbf{B}$ *pode ser escrita de modo único como uma combinação linear das colunas de* $\mathbf{A}$.

Prova. Vamos provar apenas que $(3 \Leftrightarrow 4)$. Suponhamos que $\rho(\mathbf{A}) = n$. Enrão $\mathbf{A}$ é equivalente por linha a $\mathbf{I}$. Assim, pela Proposição 1.11, existe uma matriz invertível $\mathbf{P} \in \mathbb{R}^{n\times n}$ tal que $\mathbf{PA} = \mathbf{I}$, em que $\mathbf{P} = \mathbf{E}_k \cdots \mathbf{E}_1$. Logo,

$$\mathbf{P}\left(\begin{array}{c|c} \mathbf{A} & \mathbf{I} \end{array}\right) = \left(\begin{array}{c|c} \mathbf{PA} & \mathbf{P} \end{array}\right) = \left(\begin{array}{c|c} \mathbf{I} & \mathbf{P} \end{array}\right).$$

Finalmente, como $\mathbf{PA} = \mathbf{I}$ temos que $\mathbf{P} = \mathbf{A}^{-1}$. A recíproca é clara. ■

A prova da equivalência dos itens (3) e (4) do Teorema 2.17 nos fornece um método para obter a inversa de uma matriz invertível:

$$\left(\begin{array}{c|c} \mathbf{A} & \mathbf{I} \end{array}\right) \longrightarrow \left(\begin{array}{c|c} \mathbf{E}_1\mathbf{A} & \mathbf{E}_1 \end{array}\right) \longrightarrow \cdots \longrightarrow \left(\begin{array}{c|c} \mathbf{I} & \mathbf{E}_k \cdots \mathbf{E}_1 = \mathbf{A}^{-1} \end{array}\right).$$

É pertinente observar que no método acima devemos efetuar as operações elementares necessárias para reduzir $\mathbf{A}$ a $\mathbf{I}$. Além disso, a solução do sistema $\mathbf{AX} = \mathbf{B}$ é dada pelo diagrama de flechas

$$\left(\begin{array}{c|c|c} \mathbf{A} & \mathbf{I} & \mathbf{B} \end{array}\right) \longrightarrow \cdots \longrightarrow \left(\begin{array}{c|c|c} \mathbf{I} & \mathbf{A}^{-1} & \mathbf{X} \end{array}\right).$$

Exemplo 2.18 *Determine a inversa da matriz*

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

Solução. Pelo exposto acima, basta escalonar a matriz

$$\left(\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 5 & 3 & 0 & 1 \end{array}\right) \longrightarrow \cdots \longrightarrow \left(\begin{array}{cc|cc} 1 & 0 & 3 & -1 \\ 0 & 1 & -5 & 2 \end{array}\right).$$

Portanto,

$$\mathbf{A}^{-1} = \mathbf{T}_{12}\left(-\frac{1}{2}\right)\mathbf{D}_2(2)\mathbf{T}_{21}(-5)\mathbf{D}_1\left(\frac{1}{2}\right) = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}.$$

Este método de determinar a inversa de uma matriz é tedioso de escrever. porém todas as operações são fáceis. ■

Exercícios

1. Sejam $\mathbf{A} \in F^{n\times n}$ e $\mathbf{R}$ a matriz em forma escalonada de $\mathbf{A}$. Mostre que $\mathbf{R}$ possui uma linha nula ou $\mathbf{R} = \mathbf{I}$.

2. Determine $a, b, c \in F$, de modo que as matrizes sejam equivalentes por linhas.

$$\begin{pmatrix} 1 & 0 & a \\ 0 & -1 & 0 \\ b & c & 0 \end{pmatrix} \text{ e } \begin{pmatrix} 1 & 1 & 2 \\ -2 & 0 & -4 \\ 1 & 3 & 2 \end{pmatrix}$$

3. Discuta e resolva os sistemas:

$$\left\{ \begin{array}{rcr} x_1 + x_2 + x_3 & = & 3 \\ 2x_1 - x_2 + 2x_3 & = & 5 \\ x_1 + 4x_2 - 5x_3 & = & 4 \end{array} \right. \text{ e } \left\{ \begin{array}{rcr} 2x_1 + x_2 - 4x_3 & = & 3 \\ 2x_1 + 3x_2 + 2x_3 & = & -1 \\ 0x_1 + x_2 + 3x_3 & = & -2 \end{array} \right.$$

4. Determine $a, b \in F$, de modo que o sistema tenha infinitas soluções.

$$\left\{ \begin{array}{rcl} x_1 + 2x_2 - 2x_3 & = & 7 \\ 3x_1 + x_2 - 5x_3 & = & b \\ -x_1 + ax_2 + x_3 & = & 3 \end{array} \right.$$

5. Determine condições sobre b_1, b_2 e b_3, de modo que o sistema tenha solução.

$$\left\{ \begin{array}{rcl} x_1 - 2x_2 + x_3 & = & b_1 \\ 2x_1 + x_2 + x_3 & = & b_2 \\ 0x_1 + 5x_2 - x_3 & = & b_3 \end{array} \right.$$

6. Determine $\lambda \in F$, de modo que exista uma matriz $\mathbf{B} \in F^{3\times 2}$ tal que

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & \lambda \end{pmatrix} \mathbf{B} = \begin{pmatrix} 1 & 2 \\ 3 & 1 \\ 5 & 5 \end{pmatrix}.$$

7. Seja $\mathbf{H} = (h_{ij}) \in F^{n\times n}$. Diremos que $\mathbf{H}$ é uma *matriz de Hilbert*[2] se $h_{ij} = \frac{1}{i+j-1}$. Determine a inversa de $\mathbf{H}$ quando $n = 3$. Seja $\mathbf{T} = (a_{ij}) \in F^{n\times n}$ a

[2]David Hilbert, 1862-1943, matemático alemão.

matriz tridiagonal, com $a_{ii} = 2(i-1)i + \frac{3}{4}$ e $a_{i(i+1)} = a_{(i+1)i} = -i^2$. Mostre que $\mathbf{HT} = \mathbf{TH}$.

8. Sejam as matrizes

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \mathbf{C} = \begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix} \in F^{2\times 2}.$$

Determine uma matriz $\mathbf{X} \in F^{2\times 2}$, de modo que

$$\mathbf{XA} - 2\mathbf{X} + \mathbf{XB}^2 = \mathbf{C}^2 - \mathbf{XA} - \mathbf{XB}^2.$$

9. Seja a matriz

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 & 0 \\ -1 & 0 & 3 & 5 \\ 1 & -2 & 1 & 1 \end{pmatrix} \in F^{3\times 4}.$$

Determine uma matriz $\mathbf{R}$ em forma escalonada de $\mathbf{A}$ e uma matriz invertível $\mathbf{P} \in F^{3\times 3}$ tal que $\mathbf{R} = \mathbf{PA}$.

10. Seja a matriz

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -4 \\ -3 & -4 & 8 \end{pmatrix} \in F^{3\times 3}.$$

Determine uma matriz invertível $\mathbf{P} \in F^{3\times 3}$ tal que

$$\mathbf{P}^t\mathbf{AP} = \mathbf{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{pmatrix}.$$

Note que $\mathbf{A}^t = \mathbf{A}$ e $\mathbf{D}$ é diagonal.

11. Seja $t \in F$ fixado e considere os conjuntos

$$U = \{(x,y,z) \in \mathbb{R}^3 : x - y + tz = 2\}, V = \{(x,y,z) \in \mathbb{R}^3 : y + z = 1\}$$
$$W = \{(x,y,z) \in \mathbb{R}^3 : x - y - ty = t\}.$$

Determine $U \cap V \cap W$. Dê uma interpretação geométrica desse problema.

12. Sejam $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{X} \in F^{n\times n}$. Mostre que a equação $\mathbf{AX} = \mathbf{B}$ possui solução

se, e somente se, $\rho(\mathbf{A}) = \rho(\mathbf{A}|\mathbf{B})$.

(a) Se $\mathbf{AD} = \mathbf{BC}$, então as equações $\mathbf{AX} = \mathbf{B}$ e $\mathbf{XC} = \mathbf{D}$ possuem uma solução comum.

(b) Se $\mathbf{A}$ ou $\mathbf{B}$ é não singular, então vale a recíproca do item (a).

2.2 Um Método Alternativo

Nesta seção apresentamos um método alternativo para resolver e analisar o sistema $\mathbf{AX} = \mathbf{B}$ ou o *sistema adjunto* $\mathbf{A}^t\mathbf{Y} = \mathbf{C}$. Para isto, iniciaremos descrevendo o conjunto geral de soluções do sistema homogêneo.

Sejam $\mathbf{A} \in F^{m\times n}$, $\mathbf{X}, \mathbf{Y} \in F^{n\times 1}$ soluções do sistema $\mathbf{AX} = \mathbf{O}$ e $c, d \in F$. Então

$$\mathbf{A}(c\mathbf{X} + d\mathbf{Y}) = c\mathbf{AX} + d\mathbf{AY} = c\mathbf{O} + d\mathbf{O} = \mathbf{O}.$$

Isto significa que $c\mathbf{X} + d\mathbf{Y}$ é também uma solução do sistema. Assim, o conjunto

$$S = \{\mathbf{X} \in F^{n\times 1} : \mathbf{AX} = \mathbf{O}\}$$

herda as propriedades de $F^{n\times 1}$ e chama-se *espaço solução*, de modo que a nulidade $\nu(\mathbf{A}) = \dim(S)$. Seja $\mathbf{R} \in F^{n\times m}$ a matriz em forma escalonada de $\mathbf{A}^t$. Então, pelo Proposição 1.11, existe uma matriz invertível $\mathbf{P} \in F^{n\times n}$ tal que

$$\left(\begin{array}{c|c} \mathbf{PA}^t & \mathbf{P} \end{array}\right) = \mathbf{P}\left(\begin{array}{c|c} \mathbf{A}^t & \mathbf{I} \end{array}\right) = \left(\begin{array}{c|c} \mathbf{R} & \mathbf{N} \end{array}\right).$$

Assim, $\mathbf{PA}^t = \mathbf{R}$ e $\mathbf{N} = \mathbf{P}$. Logo, $\mathbf{NA}^t = \mathbf{R}$. Como $\mathbf{N}$ é invertível temos que $\rho(\mathbf{R}) = \rho(\mathbf{NA}^t) = \rho(\mathbf{A}^t) = r$. Isto significa que $\mathbf{R}$ possui $n - r$ linhas nulas. Pondo

$$\mathbf{R} = \left(\begin{array}{c} \mathbf{R}_1 \\ \hline \mathbf{O} \end{array}\right) \text{ e } \mathbf{N} = \left(\begin{array}{c} \mathbf{N}_1 \\ \hline \mathbf{N}_2 \end{array}\right),$$

onde $\mathbf{R}_1 \in F^{r\times m}$, possui todas as linhas não nulas, e $\mathbf{N}_2 \in F^{(n-r)\times n}$. Neste caso,

$$\left(\begin{array}{c|c} \mathbf{NA}^t & \mathbf{N} \end{array}\right) = \left(\begin{array}{c|c} \mathbf{R} & \mathbf{N} \end{array}\right) = \left(\begin{array}{c|c} \mathbf{R}_1 & \mathbf{N}_1 \\ \hline \mathbf{O} & \mathbf{N}_2 \end{array}\right) \text{ e } \mathbf{R} = \mathbf{NA}^t = \left(\begin{array}{c} \mathbf{N}_1\mathbf{A}^t \\ \hline \mathbf{N}_2\mathbf{A}^t \end{array}\right)$$

implicam que

$$\left(\begin{array}{c|c} \mathbf{N}_1\mathbf{A}^t & \mathbf{N}_1 \\ \hline \mathbf{N}_2\mathbf{A}^t & \mathbf{N}_2 \end{array}\right) = \left(\begin{array}{c|c} \mathbf{R}_1 & \mathbf{N}_1 \\ \hline \mathbf{O} & \mathbf{N}_2 \end{array}\right).$$

Logo, $\mathbf{N}_2\mathbf{A}^t = \mathbf{O}$ ou $\mathbf{A}\mathbf{N}_2^t = \mathbf{O}$. Portanto, as $\nu(\mathbf{A}) = n - r$ colunas de $\mathbf{N}_2^t$ são as soluções LI do sistema $\mathbf{AX} = \mathbf{O}$, pois $\mathbf{R}_1$ possui r linhas e $\mathbf{N}$ é invertível. O processo de determinar a solução geral do sistema homogêneo pode ser resumido no diagrama de fechas:

$$\left(\begin{array}{c|c} \mathbf{A}^t & \mathbf{I} \end{array}\right) \longrightarrow \cdots \longrightarrow \left(\begin{array}{c|c} \mathbf{R} & \mathbf{N} \end{array}\right).$$

Exemplo 2.19 *No processo de fotossíntese, o dióxido de carbono (CO_2) reage com a água (H_2O) para produzir glicose $(C_6H_{12}O_6)$ e oxigênio (O_2). Quanto de carbono, hidrogênio e oxigênio precisamos*?

Solução. O nosso problema consiste em resolver a equação química:

$$x_1CO_2 + x_2H_2O \to x_3C_6H_{12}O_6 + x_4O_2.$$

Como o número de átomos de cada elemento no início da reação deve ser igual ao número de átomos desse mesmo elemento no fim da reação temos que escolher x_1, x_2, x_3 e x_4, de modo que $x_1 = 6x_3, 2x_2 = 12x_3$ e $2x_1 + x_2 = 6x_3 + 2x_4$. Mas isto é equivalente a resolver o sistema homogêneo:

$$\left\{\begin{array}{rcl} x_1 + 0x_2 - 6x_3 + 0x_4 & = & 0 \\ 0x_1 + 2x_2 - 12x_3 + 0x_4 & = & 0 \\ 2x_1 + x_2 - 6x_3 - 2x_4 & = & 0 \end{array}\right.$$

Pelo exposto acima, basta escalonar a matriz

$$\left(\begin{array}{rrr|rrrr} 1 & 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 1 & 0 & 0 \\ -6 & -12 & -6 & 0 & 0 & 1 & 0 \\ 0 & 0 & -2 & 0 & 0 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|cccc} 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{4} \\ 0 & 0 & 1 & 0 & 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 & 6 & 6 & 1 & 6 \end{array}\right)$$

Portanto, $S = \{(6c, 6c, c, 6c) : c \in \mathbb{R}\}$ é a descrição paramétrica do conjunto solução do sistema. Em particular, tomando $c = 1$, obtemos $(6, 6, 1, 6)$ uma solução particular do sistema. Neste caso,

$$6CO_2 + 6H_2O \to C_6H_{12}O_6 + 6O_2.$$

Note que $\rho(\mathbf{A}) = 3$ e $\nu(\mathbf{A}) = 4 - 3 = 1$ é o número de variáveis livres, o qual é igual ao número de linhas LI de $\mathbf{N}_2$. ■

Teorema 2.20 *Sejam* $\mathbf{A} \in F^{m\times n}$ *e* $\mathbf{AX} = \mathbf{B}$ *um sistema. Então o sistema é consistente e determinado se, e somente se, o sistema homogêneo* $\mathbf{AX} = \mathbf{O}$ *possui apenas a solução trivial.*

Prova. Suponhamos, por absurdo, que o sistema $\mathbf{AX} = \mathbf{O}$ possua uma solução $\mathbf{X}_h \neq \mathbf{O}$ e que $\mathbf{X}_p$ seja uma solução de $\mathbf{AX} = \mathbf{B}$. Então, para cada $c \in F$, teremos

$$\mathbf{A}(\mathbf{X}_p + c\mathbf{X}_h) = \mathbf{AX}_p + c\mathbf{AX}_0 = \mathbf{B},$$

isto é, $\mathbf{X}_p + c\mathbf{X}_h$ é uma solução de $\mathbf{AX} = \mathbf{B}$, o que contradiz a unicidade da solução. Reciprocamente, suponhamos que $\mathbf{X}_1$ e $\mathbf{X}_2$ sejam soluções de $\mathbf{AX} = \mathbf{B}$. Então

$$\mathbf{A}(\mathbf{X}_2 - \mathbf{X}_1) = \mathbf{AX}_2 - \mathbf{AX}_1 = \mathbf{B} - \mathbf{B} = \mathbf{O}.$$

Assim, $\mathbf{X}_2 - \mathbf{X}_1 = \mathbf{O}$, ou seja, $\mathbf{X}_2 = \mathbf{X}_1$. Portanto, o sistema é consistente e determinado. ■

Pela prova do Teorema 2.20, a solução geral do sistema $\mathbf{AX} = \mathbf{B}$ é

$$\mathbf{X} = \mathbf{X}_p + \mathbf{X}_h,$$

em que

$$\mathbf{X}_h = \sum_{i=r+1}^{n} c_i\mathbf{L}_i, \quad c_i \in F,$$

$r = \rho(\mathbf{A})$ e $\mathbf{L}_i, i = r+1, \ldots, n$, são as linhas da matriz $\mathbf{N}_2$, as quais são todas não nulas, pois são LI. O próximo resultado apresenta uma prova alternativa de que $\rho_c(\mathbf{A}) = \rho_l(\mathbf{A})$.

Teorema 2.21 *Sejam* $\mathbf{A} \in F^{m\times n}$ *e* $\mathbf{AX} = \mathbf{B}$ *um sistema consistente. Então o diagrama de flechas*

$$\left(\begin{array}{c|c} \mathbf{A}^t & \mathbf{I} \\ \hline -\mathbf{B}^t & \mathbf{O}^t \end{array}\right) \longrightarrow \cdots \longrightarrow \left(\begin{array}{c|c} \mathbf{A}^t & \mathbf{I} \\ \hline \mathbf{O}^t & \mathbf{C}^t \end{array}\right).$$

implica que $\mathbf{C}$ *é uma solução de* $\mathbf{AX} = \mathbf{B}$.

Prova. Observe que $\mathbf{AX} = \mathbf{B}$ é consistente se, e somente se, $\mathbf{X^t A^t} = \mathbf{B}^t$ for consistente. Assim, o diagrama de flechas é possível. Logo, pela Proposição 1.11,

$$\left(\begin{array}{c|c} \mathbf{A}^t & \mathbf{I} \\ \hline \mathbf{S^t A^t} - \mathbf{B}^t & \mathbf{S}^t \end{array}\right) = \left(\begin{array}{c|c} \mathbf{I} & \mathbf{O} \\ \hline \mathbf{S}^t & \mathbf{I} \end{array}\right) \left(\begin{array}{c|c} \mathbf{A}^t & \mathbf{I} \\ \hline -\mathbf{B}^t & \mathbf{O}^t \end{array}\right) = \left(\begin{array}{c|c} \mathbf{A}^t & \mathbf{I} \\ \hline \mathbf{O}^t & \mathbf{C}^t \end{array}\right).$$

Isto implica que $\mathbf{S^t A^t} - \mathbf{B}^t = \mathbf{O}^t$ e $\mathbf{S}^t = \mathbf{C}^t$. Portanto, $\mathbf{AC} = \mathbf{B}$. ■

Já vimos que o processo de determinar a solução geral do sistema homogêneo era dado pelo diagrama de fechas:

$$\left(\begin{array}{c|c} \mathbf{A}^t & \mathbf{I} \end{array}\right) \longrightarrow \cdots \longrightarrow \left(\begin{array}{c|c} \mathbf{R} & \mathbf{N} \end{array}\right).$$

Portanto, pelo Teorema 2.21, a solução geral do sistema $\mathbf{AX} = \mathbf{B}$ é dada pelo diagrama de fechas:

$$\left(\begin{array}{c|c} \mathbf{A}^t & \mathbf{I} \\ \hline -\mathbf{B}^t & \mathbf{O}^t \end{array}\right) \longrightarrow \cdots \longrightarrow \left(\begin{array}{c|c} \mathbf{A}^t & \mathbf{I} \\ \hline \mathbf{O}^t & \mathbf{C}^t \end{array}\right) \longrightarrow \cdots \longrightarrow \left(\begin{array}{c|c} \mathbf{R} & \mathbf{N} \\ \hline \mathbf{O} & \mathbf{C}^t \end{array}\right).$$

É pertinente ressaltar que no primeiro estágio do diagrama de flechas usamos algumas combinações lineares das linhas de $\mathbf{A}^t$ para zerar a linha $-\mathbf{B}^t$, sem que a última linha fosse permutada com qualquer outra linha e nem multiplicada por constante. Além disso, $\mathbf{C}$ é uma solução de $\mathbf{AX} = \mathbf{B}$ e não do sistema homogêneo $\mathbf{AX} = \mathbf{O}$, as quais são dadas pelas linhas de $\mathbf{N}$.

Exemplo 2.22 *Resolva o sistema*

$$\begin{cases} x_1 + x_2 + 2x_3 &= 3 \\ x_1 + 0x_2 + x_3 &= 5 \\ 2x_1 + x_2 + 3x_3 &= 8 \end{cases}$$

Solução. Basta reduzir a matriz em forma escalonada:

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 2 & 1 & 3 & 0 & 0 & 1 \\ \hline -3 & -5 & -8 & 0 & 0 & 0 \end{array}\right) \to \cdots \to \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 & -1 & 1 \\ \hline 0 & 0 & 0 & 5 & -2 & 0 \end{array}\right)$$

Note que para zerarmos a última linha usamos a combinação linear $5\mathbf{L}_1 - 2\mathbf{L}_2$. O terno $(5, -2, 0)$ é uma solução do sistema. Enquanto, o terno $(-1, -1, 1)$ é uma solução do sistema homogêneo. Portanto,

$$\mathbf{X}^t = (5, -2, 0) + t(-1, -1, 1), \ \ \forall \ t \in F, \tag{2.7}$$

é a solução geral do sistema. Geometricamente, a equação (2.7) representa à “reta” que passa por $(5, -2, 0)$ na direção do vetor $(-1, -1, 1)$. Portanto, o conjunto solução de $\mathbf{AX} = \mathbf{B}$ é uma reta através de $(5, -2, 0)$ paralela ao conjunto solução do sistema homogêneo $\mathbf{AX} = \mathbf{O}$. ■

Exercícios

1. Resolva os sistemas

$$\left\{\begin{array}{lcr} x_1 + 2x_2 - 2x_3 & = & 1 \\ 2x_1 + x_2 - 2x_3 & = & 6 \\ x_1 + 8x_2 - 6x_3 & = & -7 \end{array}\right. \quad \text{e} \quad \left\{\begin{array}{lcl} x_1 + 2x_2 - x_3 - 2x_4 & = & 1 \\ x_1 + 2x_2 + x_3 - x_4 & = & 0 \\ 0x_1 + x_2 + x_3 + x_4 & = & 0 \\ x_1 + 2x_2 + 2x_3 + 2x_4 & = & 0. \end{array}\right.$$

2. Seja $\mathbf{A} \in F^{m\times n}$. Mostre que se $\mathbf{A}$ possui pelo menos $m - r$ linhas nulas, então $\rho(\mathbf{A}) \leq r$.

3. Sejam $\mathbf{A}, \mathbf{B} \in F^{n\times n}$. Mostre que $\rho(\mathbf{AB}) \leq \min\{\rho(\mathbf{A}), \rho(\mathbf{B})\}$. Conclua que se $\mathbf{B}$ é não singular, então $\rho(\mathbf{BA}) = \rho(\mathbf{A}) = \rho(\mathbf{AB})$.

4. Sejam $\mathbf{A} \in F^{m\times m}$, $\mathbf{B} \in F^{m\times n}$, $\mathbf{C} \in F^{n\times m}$, $\mathbf{D} \in F^{n\times n}$,

$$\mathbf{E} = \begin{pmatrix} \mathbf{I}_m & \mathbf{B} \\ \mathbf{C} & \mathbf{O} \end{pmatrix} \ \text{e} \ \mathbf{F} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$$

   Mostre que $\rho(\mathbf{B}) + \rho(\mathbf{C}) \leq \rho(\mathbf{E})$ e se $\mathbf{A}$ é não singular, então $\rho(\mathbf{F}) = m$ se, e somente se, $\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B} = \mathbf{O}$.

5. (**Desigualdade de Sylvester**) Sejam $\mathbf{A} \in F^{m\times n}$ e $\mathbf{B} \in F^{n\times p}$. Mostre que $\rho(\mathbf{A}) + \rho(\mathbf{B}) - n \leq \rho(\mathbf{AB})$.

6. Determine todas as funções $f : \mathbb{R} \to \mathbb{R}$ da forma $f(x) = a + bx + cx^2 + dx^3 + ex^4$, de modo que $f + f' + f'' + f''' = 1$.

7. Uma matriz

$$\mathbf{A} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \in F^{3\times 3}$$

é um *quadrado mágico de ordem* 3 se a soma dos elementos das três linhas, das três colunas e das duas diagonais são todas iguais ao mesmo *peso* s.

   (a) Reescreva as condições para um quadrado mágico como um sistema de 8 equações lineares nas variáveis s, a_i, b_i e c_i, $i = 1, 2, 3$ e resolva esse sistema.

   (b) Mostre que $3b_2 = s$.

   (c) Substitua as estrelas por números, de modo que a matriz

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

   seja um quadrado mágico.

8. Seja $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$ e $\mathbf{y} = (y_1, y_2)$ sua reflexão em volta do eixo dos x. Determine $\mathbf{A} \in \mathbb{R}^{2\times 2}$ tal que $\mathbf{y}^t = \mathbf{A}\mathbf{x}^t$.

9. Que condições os inteiros a, b e c devem satisfazer para que o sistema de equações

$$\left\{ \begin{array}{rcl} 2x_1 - x_2 + x_3 - 3x_4 & = & a \\ x_1 + x_2 - x_3 & = & b \\ 4x_1 + x_2 - x_3 - 3x_4 & = & c \end{array} \right.$$

tenha soluções inteiras?

10. Sejam $\mathbf{A} \in \mathbb{R}^{m\times n}$ e $\mathbf{B} \in \mathbb{R}^{m\times 1}$. Mostre que se $\mathbf{AX} = \mathbf{B}$ possui uma solução em $\mathbb{C}^{n\times 1}$, então ele também possui uma solução em $\mathbb{R}^{n\times 1}$.

11. Sejam $x_1, \ldots, x_n, y_1, \ldots, y_n \in F$. Mostre que se $x_1, \ldots, x_n$ são distintos aos pares, então existe um único polinômio $f(x) \in F[x]$ de grau no máximo $n - 1$ tal que $f(x_i) = y_i$, $i = 1, \ldots, n$.

12. Sejam $f(x) = (x_1 - x) \cdots (x_n - x) \in F[x]$ e

$$\mathbf{D}_n = \begin{pmatrix} x_1 & a & a & a & \cdots & a & a \\ b & x_2 & a & a & \cdots & a & a \\ b & b & x_3 & a & \cdots & a & a \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ b & b & b & b & \ldots & x_{n-1} & a \\ b & b & b & b & \ldots & b & x_n \end{pmatrix}.$$

(a) Mostre que se $a \neq b$, então $(b - a)\det(\mathbf{D}_n) = bf(a) - af(b)$.

(b) Mostre que se $a = b$, então $\det(\mathbf{D}_n) = a\sum_{i=1}^{n} f_i(a) + x_n f_n(a)$, com $f_i(x) = (x_1 - x) \cdots (x_{i-1} - x)(x_{i+1} - x) \cdots (x_n - x)$.

3
Vetores

O principal objetivo deste capítulo é levar o aluno a compreender o conceito de espaço vetorial de um ponto de vista axiomático, isto é, o conceito abstrato de espaço vetorial como objeto com uma estrutura algébrica específica, o qual é devido a Peano[1] em 1888. Além disso, serão vistos os conceitos de subespaços vetoriais, dependência e independência linear, bases e dimensão de um espaço vetorial e relações entre bases de um mesmo espaço vetorial.

3.1 Espaços Vetoriais

Um *espaço vetorial* sobre o corpo F é um conjunto não vazio V munido com duas operações: *adição*

$$\begin{array}{rccc} +: & V \times V & \rightarrow & V \\ & (\mathbf{u}, \mathbf{v}) & \mapsto & \mathbf{u}+\mathbf{v} \end{array}$$

e *multiplicação por escalar*

$$\begin{array}{rccc} \cdot: & F \times V & \rightarrow & V \\ & (a, \mathbf{u}) & \mapsto & a\mathbf{u} \end{array}$$

[1]Giuseppe Peano, 1858-1932, matemático italiano.

tais que os seguintes axiomas são satisfeitos:

1. $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$, para todos $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.

2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$, para todos $\mathbf{u}, \mathbf{v} \in V$.

3. Existe um $\mathbf{0} \in V$ tal que $\mathbf{u} + \mathbf{0} = \mathbf{u}$, para todo $\mathbf{u} \in V$.

4. Para cada $\mathbf{u} \in V$, existe um $-\mathbf{u} \in V$ tal que $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.

5. $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$, para todos $a, b \in F$ e $\mathbf{u} \in V$.

6. $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$, para todos $\mathbf{u}, \mathbf{v} \in V$ e $a \in F$.

7. $a(b\mathbf{u}) = (ab)\mathbf{u}$, para todos $a, b \in F$ e $\mathbf{u} \in V$.

8. $1 \cdot \mathbf{u} = \mathbf{u}$, para todo $\mathbf{u} \in V$.

Note que o próprio corpo F com as operações usuais é um espaço vetorial. Além disso, $V = \{\mathbf{0}\}$ é um espaço vetorial, o qual chama-se *espaço vetorial trivial*. Observe que se $\mathbf{w} = -\mathbf{u} + \mathbf{u}$, então

$$\begin{aligned} \mathbf{w} + \mathbf{w} &= (-\mathbf{u} + \mathbf{u}) + (-\mathbf{u} + \mathbf{u}) = -\mathbf{u} + ([\mathbf{u} + (-\mathbf{u})] + \mathbf{u}) \\ &= -\mathbf{u} + (\mathbf{0} + \mathbf{u}) = -\mathbf{u} + \mathbf{u} = \mathbf{w}. \end{aligned}$$

Assim,

$$\begin{aligned} \mathbf{0} &= \mathbf{w} + (-\mathbf{w}) = [\mathbf{w} + \mathbf{w}] + (-\mathbf{w}) = \mathbf{w} + [\mathbf{w} + (-\mathbf{w})] \\ &= \mathbf{w} + \mathbf{0} = \mathbf{w}. \end{aligned}$$

Portanto, $-\mathbf{u} + \mathbf{u} = \mathbf{0}$, para todo $\mathbf{u} \in V$. Além disso,

$$\mathbf{0} + \mathbf{u} = [\mathbf{u} + (-\mathbf{u})] + \mathbf{u} = \mathbf{u} + [-\mathbf{u} + \mathbf{u}] = \mathbf{u} + \mathbf{0} = \mathbf{u},$$

isto é, $\mathbf{0} + \mathbf{u} = \mathbf{u}$, para todo $\mathbf{u} \in V$. No item (3) da Proposição 3.5, provaremos que podemos escrever

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

para todos $\mathbf{u}, \mathbf{v} \in V$, para representar a *diferença* entre elementos de V. Os elementos de V chamam-se, por conveniência, de vetores. As propriedades associativa e comutativa da adição de vetores implicam que a soma de um certo número de vetores é independente da maneira pela qual esses vetores são combinados ou associados. Por exemplo, se $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ e $\mathbf{t}$ são vetores quaisquer em V, então

$$(\mathbf{u}+\mathbf{v})+(\mathbf{w}+\mathbf{t})=[\mathbf{v}+(\mathbf{u}+\mathbf{w})]+\mathbf{t}$$

e essa pode ser escrita sem confusão como

$$\mathbf{u}+\mathbf{v}+\mathbf{w}+\mathbf{t}.$$

O próximo resultado é nosso **exemplo modelo**. Por isto, faremos todos os detalhes da prova:

Exemplo 3.1 *Seja $V=F^n$ o conjunto das n-uplas ou das listas sobre F. Se $\mathbf{u}=(x_1,\ldots,x_n)\in V$ e $\mathbf{v}=(y_1,\ldots,y_n)\in V$, então V, munido com as **operações usuais** de adição*

$$\mathbf{u}+\mathbf{v}=(x_1+y_1,\ldots,x_n+y_n)$$

e multiplicação por escalar

$$a\mathbf{u}=(ax_1,\ldots,ax_n),$$

é um espaço vetorial sobre F.

Solução. O leitor que tenha dificuldade em trabalhar com o caso geral pode iniciar esse exemplo com $n=2$ ou $n=3$. (1) Dado $\mathbf{w}=(z_1,\ldots,z_n)\in V$, obtemos

$$\begin{aligned}
\mathbf{u}+(\mathbf{v}+\mathbf{w}) &= (x_1,\ldots,x_n)+(y_1+z_1,\ldots,y_n+z_n)\\
&= (x_1+(y_1+z_1),\ldots,x_n+(y_n+z_n))\\
&\overset{\text{em } F}{=} ((x_1+y_1)+z_1,\ldots,(x_n+y_n)+z_n)\\
&= (x_1+y_1,\ldots,x_n+y_n)+(z_1,\ldots,z_n)\\
&= (\mathbf{u}+\mathbf{v})+\mathbf{w}.
\end{aligned}$$

(2) Dados $\mathbf{u}=(x_1,\ldots,x_n),\mathbf{v}=(y_1,\ldots,y_n)\in V$,

$$\mathbf{u}+\mathbf{v}=(x_1+y_1,\ldots,x_n+y_n)\overset{\text{em } F}{=}(y_1+x_1,\ldots,y_n+x_n)=\mathbf{v}+\mathbf{u}$$

(3) Dado $\mathbf{u}=(x_1,\ldots,x_n)\in V$, devemos encontrar $\mathbf{v}=(y_1,\ldots,y_n)\in V$ tal que $\mathbf{u}+\mathbf{v}=\mathbf{u}$. Assim,

$$(x_1+y_1,\ldots,x_n+y_n)=(x_1,\ldots,x_n)\Rightarrow x_i+y_i=x_i,\ \ i=1,\ldots,n.$$

Logo, $y_1 = y_2 = \cdots = y_n = 0$. Portanto, existe um $\mathbf{0} = (0, \ldots, 0) \in V$ tal que $\mathbf{u} + \mathbf{0} = \mathbf{u}$, para todo $\mathbf{u} \in V$.

(4) Dado $\mathbf{u} = (x_1, \ldots, x_n) \in V$, devemos encontrar $\mathbf{v} = (y_1, \ldots, y_n) \in V$ tal que $\mathbf{u} + \mathbf{v} = \mathbf{0}$. Assim,

$$(x_1 + y_1, \ldots, x_n + y_n) = (0, \ldots, 0) \Rightarrow x_i + y_i = 0, \;\; i = 1, \ldots, n.$$

Logo, $y_i = -x_i$, $i = 1, \ldots, n$. Portanto, existe um $-\mathbf{u} = (-x_1, \ldots, -x_n) \in V$ tal que $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$, para todo $\mathbf{u} \in V$.

(5) Dado $\mathbf{u} = (x_1, \ldots, x_n) \in V$ e $a, b \in F$, teremos

$$\begin{aligned}
(a+b)\mathbf{u} &= ((a+b)x_1, \ldots, (a+b)x_n) \\
&\overset{\text{em } F}{=} (ax_1 + bx_1, \ldots, ax_n + bx_n) \\
&= (ax_1, \ldots, ax_n) + (bx_1, \ldots, bx_n) \\
&= a\mathbf{u} + b\mathbf{u}.
\end{aligned}$$

(6) Dados $\mathbf{u} = (x_1, \ldots, x_n), \mathbf{v} = (y_1, \ldots, y_n) \in V$ e $a \in F$, obtemos

$$\begin{aligned}
a(\mathbf{u} + \mathbf{v}) &= a(x_1 + y_1, \ldots, x_n + y_n) \\
&= (a(x_1 + y_1), \ldots, a(x_n + y_n)) \\
&\overset{\text{em } F}{=} (ax_1 + ay_1, \ldots, ax_n + ay_n) \\
&= (ax_1, \ldots, ax_n) + (ay_1, \ldots, ay_n) \\
&= a\mathbf{u} + a\mathbf{v}.
\end{aligned}$$

(7) Dados $\mathbf{u} = (x_1, \ldots, x_n) \in V$ e $a, b \in F$, temos que

$$\begin{aligned}
a(b\mathbf{u}) &= a(bx_1, \ldots, bx_n) = (a(bx_1), \ldots, a(bx_n)) \\
&\overset{\text{em } F}{=} ((ab)x_1, \ldots, (ab)x_n) = (ab)\mathbf{u}.
\end{aligned}$$

(8) Dado $\mathbf{u} = (x_1, \ldots, x_n) \in V$, teremos

$$1 \cdot \mathbf{u} = (1 \cdot x_1, \ldots, 1 \cdot x_n) \overset{\text{em } F}{=} (x_1, \ldots, x_n) = \mathbf{u},$$

que é o resultado desejado. ■

Já vimos que os conjuntos de matrizes $F^{m \times n}$ e dos polinômios $F[x]$ eram espaços vetoriais sobre F. Em particular, o conjunto

$$V = \mathbb{P}_n(F) = \{p(x) \in F[x] : \partial(p) \leq n\},$$

com as operações herdadas de $F[x]$, é um espaço vetorial sobre F.

Devido a sua grande importância teórica faremos todos os detalhes da prova do resultado a seguir:

Exemplo 3.2 (Espaço das Funções) *Sejam S um conjunto não vazio e*

$$V = \mathcal{F}(S, F) = F^S = \{f : S \to F : f \text{ é uma função}\}.$$

o conjunto de todas as funções de valores sobre F. Se $f \in V$ e $g \in V$, então V, munido com as operações de adição $f + g$ definida como

$$(f + g)(x) = f(x) + g(x), \ \ \forall \ x \in S,$$

e multiplicação por escalar af definida como

$$(af)(x) = af(x), \ \ \forall \ x \in S,$$

é um espaço vetorial sobre F.

Solução. (1) Dados $f, g, h \in V$. Como a adição em F é associativa temos que

$$\begin{aligned} [f + (g + h)](x) &= f(x) + (g + h)(x) = f(x) + [g(x) + h(x)] \\ &\overset{\text{em } F}{=} [f(x) + g(x)] + h(x) = (f + g)(x) + h(x) \\ &= [(f + g) + h](x), \ \ \forall \ x \in S. \end{aligned}$$

Assim, $f + (g + h) = (f + g) + h$.

(2) Dados $f, g \in V$. Como a adição em F é comutativa temos que

$$\begin{aligned} (f + g)(x) &= f(x) + g(x) \overset{\text{em } F}{=} g(x) + f(x) \\ &= (g + f)(x), \ \ \forall \ x \in S. \end{aligned}$$

Logo, $f + g = g + f$.

(3) Seja $\mathbf{0}$ a função nula, isto é, $\mathbf{0}(x) = 0$, para todo $x \in S$. Então

$$(f + \mathbf{0})(x) = f(x) + \mathbf{0}(x) = f(x) + 0 \overset{\text{em } F}{=} f(x), \ \ \forall \ x \in S.$$

Portanto, $f + \mathbf{0} = f$, para todo $f \in V$.

(4) Seja $-f$ a função definida como $(-f)(x) = -f(x)$, para todo $x \in S$. Então

$$(f + (-f))(x) = f(x) + (-f)(x) = f(x) - f(x) \overset{\text{em } F}{=} 0, \ \ \forall \ x \in S.$$

Consequentemente, $f + (-f) = \mathbf{0}$, para todo $f \in V$,

(5) Dado $f \in V$ e $a, b \in F$. Como a adição e a multiplicação em F são distributivas temos que

$$\begin{aligned}[(a+b)f](x) &= (a+b)f(x) \overset{\text{em } F}{=} af(x) + bf(x) \\ &= (af)(x) + (bf)(x) = [af + bf](x), \ \ \forall \ x \in S.\end{aligned}$$

Assim, $(a+b)f = af + bf$.

(6) Dados $f, g \in V$ e $a \in F$. Como a adição e a multiplicação em F são distributivas temos que

$$\begin{aligned}[a(f+g)](x) &= a(f+g)(x) = a[f(x) + g(x)] \\ &\overset{\text{em } F}{=} af(x) + ag(x) = (af)(x) + (ag)(x) \\ &= [af + ag](x), \ \ \forall \ x \in S.\end{aligned}$$

Logo, $a(f+g) = af + ag$.

(7) Dados $f \in V$ e $a, b \in F$. Como a multiplicação em F é associativa temos que

$$\begin{aligned}[a(bf)](x) &= a[(bf)(x)] = a[bf(x)] \overset{\text{em } F}{=} (ab)f(x) \\ &= [(ab)f](x), \ \ \forall \ x \in S.\end{aligned}$$

Portanto, $a(bf) = (ab)f$.

(8) Dado $f \in V$, temos que

$$(1 \cdot f)(x) = 1 \cdot f(x) \overset{\text{em } F}{=} f(x), \ \ \forall \ x \in S.$$

Consequentemente, $1 \cdot f = f$. ■

É importante ressaltar, de um ponto de vista teórico e didático, que os espaços vetoriais $F^n, F[x]$ e $F^{m\times n}$ são casos particulares de $\mathcal{F}(S, F)$. com $S = \{1, \ldots, n\}$, $S = F$ e $S = \{1, \ldots, m\} \times \{1, \ldots, n\}$, respectivamente.

Exemplo 3.3 *Sejam $V = F^2$, $\mathbf{u} = (x_1, x_2) \in V$ e $\mathbf{v} = (y_1, y_2) \in V$. Verifique se V munido com as operações de adição usual e multiplicação por escalar $a \odot \mathbf{u} = (ax_1, x_2)$ é um espaço vetorial sobre F.*

Solução. É claro que a operação de adição satisfaz as propriedades de (1) à (4). Assim, devemos verificar as propriedades relativas à multiplicação por escalar. Note que

$$(a+b)\odot\mathbf{u} = ((a+b)x_1, x_2) = (ax_1+bx_1, x_2);\ a\odot\mathbf{u}+b\odot\mathbf{u} = (ax_1+bx_1, 2x_2)$$

Logo, se $x_2 \neq 0$, então $(a+b)\odot\mathbf{u} \neq a\odot\mathbf{u}+b\odot\mathbf{u}$, pois $x_2 \neq 2x_2$. Portanto, V não é um espaço vetorial sobre F. ■

Exemplo 3.4 *Sejam $V = F^2$, $\mathbf{u} = (x_1, x_2) \in V$ e $\mathbf{v} = (y_1, y_2) \in V$. Verifique se V munido com as operações de adição usual e multiplicação por escalar $a\odot\mathbf{u} = (ax_1 - a + 1, ax_2)$ é um espaço vetorial sobre F.*

Solução. Fica como um exercício. ■

Proposição 3.5 *Seja V um espaço vetorial sobre F. Então:*

1. *Existe um único vetor nulo em V (elemento neutro).*
2. *Cada vetor $\mathbf{u} \in V$ admite um único vetor simétrico $-\mathbf{u}$.*
3. *Existe um único $\mathbf{x} \in V$ tal que $\mathbf{u}+\mathbf{x} = \mathbf{v}$, para todos $\mathbf{u}, \mathbf{v} \in V$, a saber, $\mathbf{x} = \mathbf{v} + (-\mathbf{u})$.*
4. *$a\mathbf{0} = \mathbf{0}$, para todo $a \in F$ e $\mathbf{0} \in V$.*
5. *$0\mathbf{u} = \mathbf{0}$, para todo $\mathbf{u} \in V$ e $0 \in F$.*
6. *Se $a\mathbf{u} = \mathbf{0}$, então $a = 0$ ou $\mathbf{u} = \mathbf{0}$, onde $a \in F$ e $\mathbf{u} \in V$.*
7. *$-\mathbf{u} = (-1)\mathbf{u}$, para todo $\mathbf{u} \in V$.*
8. *$(-a)\mathbf{u} = a(-\mathbf{u}) = -(a\mathbf{u})$, para todo $a \in F$ e $\mathbf{u} \in V$.*
9. *Se $\mathbf{u} = x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n$ e $\mathbf{v} = y_1\mathbf{u}_1 + \cdots + y_n\mathbf{u}_n$, onde $\mathbf{u}_i \in V$ e $x_i, y_i \in F$, $i = 1, \ldots, n$, então*

   $$\mathbf{u}+\mathbf{v} = (x_1+y_1)\mathbf{u}_1 + \cdots + (x_n+y_n)\mathbf{u}_n \text{ e } a\mathbf{u} = (ax_1)\mathbf{u}_1 + \cdots + (ax_n)\mathbf{u}_n$$

10. *$a(n\mathbf{u}) = n(a\mathbf{u}) = (na)\mathbf{u}$, para todo $n \in \mathbb{Z}$, $a \in F$ e $\mathbf{u} \in V$.*

Prova. Vamos provar apenas os itens $(1), (4)$ e (5): (1) Suponhamos que exista outro vetor $\mathbf{0}' \in V$ tal que $\mathbf{u} + \mathbf{0}' = \mathbf{u}$, para todo $\mathbf{u} \in V$. Então $\mathbf{0} = \mathbf{0} + \mathbf{0}' = \mathbf{0}'$. (4) e (5) Note, pela leis distributivas, que

$$\begin{aligned} a\mathbf{u} + a\mathbf{0} &= a(\mathbf{u} + \mathbf{0}) = a\mathbf{u} \\ a\mathbf{u} + 0\mathbf{u} &= (a + 0)\mathbf{u} = a\mathbf{u} \end{aligned}$$

Portanto, pelo item (1), $a\mathbf{0} = \mathbf{0}$ e $0\mathbf{u} = \mathbf{0}$. ■

Sejam $\mathbf{u}, \mathbf{v} \in V$. A *diferença* entre $\mathbf{u}$ e $\mathbf{v}$ é definida como

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

a qual, pelo item (3) da Proposição 3.5, está bem definida.

Exercícios

1. Mostre todas as afirmações deixadas nesta seção.

2. Seja $V = F^2$. Se $\mathbf{u} = (x_1, x_2) \in V$ e $\mathbf{v} = (y_1, y_2) \in V$, então V, munido com as operações de adição $\mathbf{u} \oplus \mathbf{v} = (3x_2 + 3y_2, -x_1 - y_1)$ e multiplicação por escalar $a \odot \mathbf{u} = (3ax_2, -ax_1)$, é um espaço vetorial sobre F?

3. Seja $V = F^2$. Se $\mathbf{u} = (x_1, x_2) \in V$, então V, munido com as operações de adição usual e multiplicação por escalar $a \odot \mathbf{u} = (a^2x_1, a^2x_2)$, é um espaço vetorial sobre F?

4. Seja $V = F^2$. Se $\mathbf{u} = (x_1, x_2) \in V$, então V, munido com as operações de adição usual e multiplicação por escalar $a \odot \mathbf{u} = (5ax_1, 5ax_2)$, é um espaço vetorial sobre F?

5. Seja $V = F^2$. Se $\mathbf{u} = (x_1, x_2) \in V$, então V, munido com as operações de adição usual e multiplicação por escalar $a \odot \mathbf{u} = (ax_1, 0)$, é um espaço vetorial sobre F?

6. Sejam $V = F^n$ e $\mathbf{u} = (x_1, \dots, x_n) \in V$. Verifique que se V munido com as operações de adição usual e multiplicação por escalar $a \odot \mathbf{u} = (0, \dots, 0)$ é um espaço vetorial sobre F.

7. Seja $V = \{(t, 2t, \ldots, nt) \in F^n : t \in F\}$. Verifique se V munido com as operações usuais de adição e multiplicação por escalar é um espaço vetorial sobre F.

8. Seja $V = F^n$. Se $\mathbf{u} = (x_1, \ldots, x_n) \in V$ e $\mathbf{v} = (y_1, \ldots, y_n) \in V$. Verifique se V munido com as operações de adição $\mathbf{u} \oplus \mathbf{v} = (x_1, \ldots, x_n)$ e multiplicação por escalar usual é um espaço vetorial sobre F.

9. Seja $V = F^2$. Se $\mathbf{u} = (x_1, x_2) \in V$ e $\mathbf{v} = (y_1, y_2) \in V$, então V munido com as operações de adição $\mathbf{u} \oplus \mathbf{v} = (x_1 - y_1, x_2 - y_2)$ e multiplicação por escalar $a \odot \mathbf{u} = (-ax_1, -ax_2)$ é um espaço vetorial sobre F?

10. Mostre que a propriedade de comutatividade para a adição de vetores é redundante, isto é, pode ser provada a partir das outras propriedades.

11. Por que um espaço vetorial V sobre F possui um elemento ou infinitos elementos?

12. Sejam V qualquer espaço vetorial sobre $\mathbb{R}$ e $\tilde{V} = V^{\mathbb{C}} = V \times V$. Mostre que se $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2) \in \tilde{V}$ e $\mathbf{v} = (\mathbf{v}_1, \mathbf{v}_2) \in \tilde{V}$, então $\tilde{V}$, munido com as operações de adição $\mathbf{u} \oplus \mathbf{v} = (\mathbf{u}_1 + \mathbf{v}_1, \mathbf{u}_2 + \mathbf{v}_2)$ e multiplicação por escalar $(a + ib) \odot \mathbf{u} = (a\mathbf{u}_1 - b\mathbf{u}_2, a\mathbf{u}_2 + b\mathbf{u}_1)$, é um espaço vetorial sobre $\mathbb{C}$ e chama-se *complexificação* de V. Neste caso, $\tilde{V} = \{\mathbf{u} + i\mathbf{v} : \mathbf{u}, \mathbf{v} \in V\}$.

13. Sejam $V_1, \ldots, V_n$ espaços vetoriais sobre F e $V = \prod_{i=1}^{n} V_i$. Mostre que se $\mathbf{u} = (\mathbf{x}_1, \ldots, \mathbf{x}_n) \in V$ e $\mathbf{v} = (\mathbf{y}_1, \ldots, \mathbf{y}_n) \in V$, então V, munido com as operações de adição
$$\mathbf{u} \oplus \mathbf{v} = (\mathbf{x}_1 + \mathbf{y}_1, \ldots, \mathbf{x}_n + \mathbf{y}_n)$$
e multiplicação por escalar
$$a \odot \mathbf{u} = (a\mathbf{x}_1, \ldots, a\mathbf{x}_n),$$
é um espaço vetorial sobre F e chama-se *soma direta externa* dos espaços V_i e denotamos por $V_1 \otimes \cdots \otimes V_n$.

3.2 Subespaços Vetoriais

Sejam V um espaço vetorial sobre F e W um subconjunto não vazio de V. Então as operações de adição e multiplicação por escalar são definidas para todos os elementos de V. Em particular, para aqueles pertencente a W. O resultado dessas operações sobre elementos de W estão em V. Portanto, muitas vezes, é possível que elas sejam fechadas em W, ou seja, a adição de dois elementos de W devem sempre ser um elemento de W, e a multiplicação de um escalar e um elemento de W devem sempre ser um elemento de W. Neste caso, diremos que W é um *subespaço* de V. Mais precisamente. um subconjunto de V é um *subespaço* de V se as seguintes condições são satisfeitas:

1. $W \neq \emptyset$.

2. $\mathbf{u} + \mathbf{v} \in W$, para todos $\mathbf{u}, \mathbf{v} \in W$.

3. $a\mathbf{u} \in W$, para todo $a \in \mathbb{R}$ e $\mathbf{u} \in W$.

Seja V um espaço vetorial sobre F. Qualquer subespaço W de V contém o vetor nulo $\mathbf{0}$, pois quando $a = 0$, temos que

$$\mathbf{0} = 0 \cdot \mathbf{u} \in W, \ \ \forall \ \ \mathbf{u} \in W.$$

Pode ser provado que, se admitirmos essas duas propriedades em W, os oito axiomas de espaço vetorial são válidos em W. Dessa forma, W é também um espaço vetorial com as propriedades herdadas de V. Note que o espaço vetorial V admite pelo menos dois subespaços, a saber, $\{\mathbf{0}\}$ e V, os quais chamam-se *subespaços triviais* ou *impróprios*. Os demais subespaços de V chamam-se *subespaços não triviais* ou *próprios*.

Proposição 3.6 (Critério de Subespaço) *Sejam V um espaço vetorial sobre F e W um subconjunto não vazio de V. Então W é um subespaço de V se, e somente se, $\mathbf{u} + a\mathbf{v} \in W$, para todos $\mathbf{u}, \mathbf{v} \in W$ e $a \in F$.*

Prova. Fica como um exercício. ■

Exemplo 3.7 *Sejam $V = F^n$ e $W = \{(x_1, \dots, x_n) \in V : x_1 = 0\}$. Mostre que W é um subespaço de V.*

Solução. É claro que $W \neq \emptyset$, pois $\mathbf{0} = (0, \dots, 0) \in W$. Dados $\mathbf{u}, \mathbf{v} \in W$ e $a \in F$. Como $\mathbf{u}, \mathbf{v} \in W$ temos que $\mathbf{u} = (0, x_2, \dots, x_n)$ e $\mathbf{v} = (0, y_2, \dots, y_n)$. Assim,

$$\mathbf{u} + \mathbf{v} = (0 + 0, x_2 + y_2, \dots, x_n + y_n) = (0, x_2 + y_2, \dots, x_n + y_n) \in W$$

e

$$a\mathbf{u} = (a0, ax_2, \dots, ax_n) = (0, ax_2, \dots, ax_n) \in W.$$

Portanto, W é um subespaço de V. ■

Exemplo 3.8 *Sejam* $V = F^{n\times n}$ *e* $W = \{\mathbf{A} \in V : \mathbf{A}^t = \mathbf{A}\}$ *o conjunto das matrizes simétricas. Mostre que* W *é um subespaço de* V.

Solução. É claro que $W \neq \emptyset$, pois $\mathbf{O}^t = \mathbf{O}$ implica que $\mathbf{O} \in W$. Dados $\mathbf{A}, \mathbf{B} \in W$ e $a \in F$. Como $\mathbf{A}, \mathbf{B} \in W$ temos que $\mathbf{A}^t = \mathbf{A}$ e $\mathbf{B}^t = \mathbf{B}$. Assim,

$$(\mathbf{A}+\mathbf{B})^t = \mathbf{A}^t + \mathbf{B}^t = \mathbf{A}+\mathbf{B} \Rightarrow \mathbf{A}+\mathbf{B} \in W$$

e $(a\mathbf{A})^t = a\mathbf{A}^t = a\mathbf{A}$ implica que $a\mathbf{A} \in W$. Portanto, W é um subespaço de V. ■

Exemplo 3.9 *Sejam* $\mathbf{A} \in F^{m\times n}$ *uma matriz fixada,* $V = F^{n\times 1}$ *e* $S = \{\mathbf{X} \in V : \mathbf{AX} = \mathbf{O}\}$ *o espaço solução do sistema homogêneo* $\mathbf{AX} = \mathbf{O}$. *Mostre que* W *é um subespaço de* V. *Em particular, se* $\mathbf{a} = (a, b, c) \in \mathbb{R}^3$, *com* $a^2 + b^2 + c^2 > 0$, *então o plano em* $\mathbb{R}^3$ *passando através da origem* $\pi = \{\mathbf{x} = (x, y, z) \in \mathbb{R}^3 : \mathbf{a}\mathbf{x}^t = 0\}$ *é um subespaço de* $\mathbb{R}^3$.

Solução. Fica como um exercício. ■

Exemplo 3.10 *Sejam* $V = \mathcal{F}(S, F)$ *e* $V_p = \{f \in V : f(-x) = f(x),\ \forall\ x \in F\}$ *o conjunto das* funções pares. *Mostre que* V_p *é um subespaço de* V.

Solução. É claro que $V_p \neq \emptyset$, pois $\mathbf{0}(-x) = 0 = \mathbf{0}(x)$, para todo $x \in F$, implica que $\mathbf{0} \in V_p$. Dados $f, g \in V_p$ e $a \in F$. Como f, $g \in W$ temos que $f(x) = f(-x)$ e $g(x) = g(-x)$, para todo $x \in F$. Assim,

$$(f+g)(-x) = f(-x) + g(-x) = f(x) + g(x) = (f+g)(x),\ \ \forall\ x \in F,$$

de modo que $f + g \in V_p$, e

$$(af)(-x) = af(-x) = af(x) = (af)(x),\ \ \forall\ x \in F,$$

implica que $af \in V_p$. Portanto, V_p é um subespaço de V. ■

Exemplo 3.11 (Espaço das Sequências) *Sejam* $V = \mathrm{Seq}(F) = \mathcal{F}(\mathbb{N}, F)$ *o espaço das sequências e*

$$\mathrm{Seq}_f(F) = \{\mathbf{x} = (x_1, \ldots, x_n, \ldots) \in V : x_1, x_2 \in F \ \text{ e } \ x_{n+2} = x_n + x_{n+1}\}$$

o conjunto das sequências do tipo Fibonacci. *Mostre que* $\mathrm{Seq}_f(F)$ *é um subespaço de*

V. *Note, em particular, que*

$$\mathbf{e} = (1, 0, 1, 1, 2, \ldots) \in \mathrm{Seq}_f(F) \ \ e \ \ \mathbf{f} = (0, 1, 1, 2, 3, \ldots) \in \mathrm{Seq}_f(F).$$

Solução. Fica como um exercício. ■

Exemplo 3.12 *Sejam $V = \mathbb{P}_n(F)$, com $n \geq 2$, e $W = \{p(x) \in V : p(1) = p(7) = 0\}$. Mostre que W é um subespaço de V.*

Solução. É claro que $W \neq \emptyset$, pois $\mathbf{0}(1) = \mathbf{0}(7) = 0$ implica que $\mathbf{0} \in W$. Dados $p(x), q(x) \in W$ e $a \in F$. Como $p(x), q(x) \in W$ temos que $p(1) = p(7) = 0$ e $q(1) = q(7) = 0$. Assim,

$$(p+q)(1) = p(1) + q(1) = 0 + 0 = 0 \ \text{ e } \ (p+q)(7) = p(7) + q(7) = 0 + 0 = 0,$$

de modo que $p(x) + q(x) \in W$, e $(ap)(1) = ap(1) = 0$ e $(ap)(7) = ap(7) = 0$ implicam que $ap(x) \in W$. Portanto, W é um subespaço de V. ■

Exemplo 3.13 *Sejam $V = F^n$ e $W = \{(x_1, \ldots, x_n) \in V : x_2 = x_1 + 1\}$. Então W não é um subespaço de V, pois $\mathbf{0} = (0, \ldots, 0) \notin W$.*

Exemplo 3.14 *Sejam $V = \mathbb{R}^2$ e $W = \{(x_1, x_2) \in V : x_2 = |x_1|\}$. Então W não é um subespaço de V, pois $\mathbf{u} = (-1, 1) \in W$ e $\mathbf{v} = (2, 2) \in W$, mas $\mathbf{u} + \mathbf{v} = (1, 3) \notin W$. Note que $\mathbf{0} = (0, 0) \in W$. Portanto, $\mathbf{0} \in W$ é condição **necessária**, mas **não suficiente** para que W seja um subespaço de V.*

Teorema 3.15 *Seja V um espaço vetorial sobre F. Se W_1 e W_2 são subespaços de V, então $W_1 \cap W_2$ é um subespaço de V. Conclua que $W_1 \cap W_2$ é o maior subespaço de V contido em W_1 e W_2.*

Prova. É claro que $W_1 \cap W_2 \neq \emptyset$, pois $\mathbf{0} \in W_1$ e $\mathbf{0} \in W_2$ implicam que $\mathbf{0} \in W_1 \cap W_2$. Dados $\mathbf{u}, \mathbf{v} \in W_1 \cap W_2$ e $a \in F$. Como $\mathbf{u}, \mathbf{v} \in W_1 \cap W_2$ temos que $\mathbf{u}, \mathbf{v} \in W_1$ e $\mathbf{u}, \mathbf{v} \in W_2$. Assim, por hipótese, $\mathbf{u} + \mathbf{v} \in W_1$, $\mathbf{u} + \mathbf{v} \in W_2$ e $a\mathbf{u} \in W_1$, $a\mathbf{u} \in W_2$, de modo que $\mathbf{u} + \mathbf{v} \in W_1 \cap W_2$ e $a\mathbf{u} \in W_1 \cap W_2$. Portanto, $W_1 \cap W_2$ é um subespaço de V. ■

Exemplo 3.16 *Sejam $V = F^3, W_1 = \{(x, y, z) \in V : x = 0\}$ e $W_2 = \{(x, y, z) \in V : y = 0\}$ subespaços de V (prove isto!). Determine $W_1 \cap W_2$.*

Solução. Dado $\mathbf{u} = (x, y, z) \in W_1 \cap W_2$, obtemos $\mathbf{u} = (x, y, z) \in W_1$ e $\mathbf{u} = (x, y, z) \in W_2$. Assim, $x = 0$ e $y = 0$. Logo, $\mathbf{u} = (x, y, z) \in W_1 \cap W_2$ se, e somente se, $x = y = 0$ e z qualquer. Portanto,

$$W_1 \cap W_2 = \{(x, y, z) \in V : x = y = 0\},$$

que é o resultado desejado. ■

Exemplo 3.17 *Sejam* $V = F^{2\times 2}$,

$$W_1 = \left\{ \begin{pmatrix} a & b \\ c & 0 \end{pmatrix} \in V : a, b, c \in F \right\} \text{ e } W_2 = \left\{ \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} \in V : a, d \in F \right\}$$

subespaços de V (*prove isto*!). *Determine* $W_1 \cap W_2$.

Solução. Observe que qualquer

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in W_1 \cap W_2 \Leftrightarrow \mathbf{A} \in W_1 \text{ e } \mathbf{A} \in W_2.$$

Assim, $d = 0$, $b = 0$ e $c = 0$. Logo,

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in W_1 \cap W_2 \Leftrightarrow b = c = d = 0$$

e a qualquer elemento de F. Portanto,

$$W_1 \cap W_2 = \{a\mathbf{E}_{11} \in V : a \in F\},$$

que é o resultado desejado. ■

Pergunta. $W_1 \cup W_2$ é um subespaço de V? A resposta dessa pergunta é, em geral, não. Por exemplo, sejam $V = \mathbb{R}^2$,

$$W_1 = \{(x, y) \in V : y = 0\} \text{ e } W_2 = \{(x, y) \in V : x = 0\}$$

subespaços de V (prove isto!). Então $W_1 \cup W_2$ não é um subespaço de V, pois $\mathbf{u} = (1, 0) \in W_1 \cup W_2$ e $\mathbf{v} = (0, 1) \in W_1 \cup W_2$, mas $\mathbf{u} + \mathbf{v} = (1, 1) \notin W_1 \cup W_2$.

Teorema 3.18 *Seja V um espaço vetorial sobre F. Se W_1 e W_2 são subespaços de V, então o conjunto*

$$W_1 + W_2 = \{\mathbf{u}_1 + \mathbf{u}_2 : \mathbf{u}_1 \in W_1 \ \ e \ \ \mathbf{u}_2 \in W_2\}$$

é um subespaço de V. Note que $W_1 \cap W_2 \subseteq W_1 \cup W_2 \subseteq W_1 + W_2$.

Prova. Como $\mathbf{0} \in W_1$ e $\mathbf{0} \in W_2$ temos que $\mathbf{0} = \mathbf{0}+\mathbf{0} \in W_1+W_2$. Assim, $W_1+W_2 \neq \emptyset$. Dados $\mathbf{u}, \mathbf{v} \in W_1 + W_2$ e $a \in F$, de modo que existem $\mathbf{u}_1, \mathbf{v}_1 \in W_1$ e $\mathbf{u}_2, \mathbf{v}_2 \in W_2$ tais que $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$ e $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$. Logo, por hipótese, $\mathbf{u}_1 + \mathbf{v}_1 \in W_1$, $\mathbf{u}_2 + \mathbf{v}_2 \in W_2$ e $a\mathbf{u}_1 \in W_1$, $a\mathbf{u}_2 \in W_2$. Portanto,

$$\mathbf{u} + \mathbf{v} = (\mathbf{u}_1 + \mathbf{u}_2) + (\mathbf{v}_1 + \mathbf{v}_2) = (\mathbf{u}_1 + \mathbf{v}_1) + (\mathbf{u}_2 + \mathbf{v}_2) \in W_1 + W_2$$

e

$$a\mathbf{u} = a(\mathbf{u}_1 + \mathbf{u}_2) = a\mathbf{u}_1 + a\mathbf{u}_2 \in W_1 + W_2.$$

Consequentemente, $W_1 + W_2$ é um subespaço de V. ■

Exemplo 3.19 *Sejam $V = F$, $W_1 = \{(x, y, z) \in V : x = 0\}$ e $W_2 = \{(x, y, z) \in V : y = z = 0\}$ subespaços de V (prove isto!). Determine $W_1 \cap W_2$ e $W_1 + W_2$.*

Solução. Note que $\mathbf{u} = (x, y, z) \in W_1 \cap W_2$ se, e somente, se $\mathbf{u} \in W_1$ e $\mathbf{u} \in W_2$. Assim, $x = 0$ e $y = z = 0$. Portanto, $W_1 \cap W_2 = \{(0, 0, 0)\}$. Dado $\mathbf{u} \in W_1 + W_2$, existem $\mathbf{u}_1 = (0, y, z) \in W_1$ e $\mathbf{u}_2 = (x, 0, 0) \in W_2$, onde $x, y, z \in F$, tais que $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2 = (x, y, z) \in V$. Consequentemente, $W_1 + W_2 = V$. ■

Sejam V um espaço vetorial sobre F e W_1, W_2 subespaços de V. Diremos que V é decomposto em *soma direta interna* de W_1 e W_2, em símbolos $V = W_1 \oplus W_2$, se as seguintes condições são satisfeitas:

1. $V = W_1 + W_2$.
2. $W_1 \cap W_2 = \{\mathbf{0}\}$.

Neste caso, W_1 chama-se *complemento direto* de W_2 e vice-versa. Por exemplo, confira o Exemplo 3.19 para ver que $V = W_1 \oplus W_2$.

Exemplo 3.20 *Sejam $V = F^{n\times n}$, $W_1 = \{\mathbf{A} \in V : \mathbf{A}^t = \mathbf{A}\}$ e $W_2 = \{\mathbf{A} \in V : \mathbf{A}^t = -\mathbf{A}\}$ subespaços de V. Mostre que $V = W_1 \oplus W_2$.*

Solução. Dado $\mathbf{A} \in W_1 \cap W_2$, temos que $\mathbf{A} \in W_1$ e $\mathbf{A} \in W_2$. Assim, $\mathbf{A}^t = \mathbf{A}$ e $\mathbf{A}^t = -\mathbf{A}$, de modo que $\mathbf{A} = -\mathbf{A}$ implica que $2\mathbf{A} = \mathbf{O}$, ou seja, $\mathbf{A} = \mathbf{O}$. Logo, $W_1 \cap W_2 = \{\mathbf{O}\}$. Primeiro note que a inclusão $W_1 + W_2 \subseteq V$ é sempre verdadeira. Assim, basta provar a inclusão $V \subseteq W_1 + W_2$. Para isto, dado $\mathbf{A} \in V$, temos que

$$\begin{aligned} \mathbf{A} &= 1 \cdot \mathbf{A} = (\tfrac{1}{2} + \tfrac{1}{2})\mathbf{A} = \tfrac{1}{2}\mathbf{A} + \tfrac{1}{2}\mathbf{A} = \tfrac{1}{2}\mathbf{A} + \tfrac{1}{2}\mathbf{A}^t - \tfrac{1}{2}\mathbf{A}^t + \tfrac{1}{2}\mathbf{A} \\ &= \tfrac{1}{2}(\mathbf{A} + \mathbf{A}^t) + \tfrac{1}{2}(\mathbf{A} - \mathbf{A}^t). \end{aligned}$$

É fácil verificar que $2^{-1}(\mathbf{A} + \mathbf{A}^t) \in W_1$ e $2^{-1}(\mathbf{A} - \mathbf{A}^t) \in W_2$, de modo que $\mathbf{A} \in W_1 + W_2$. Portanto, $V = W_1 + W_2$. ■

É muito importante ressaltar que: quando as propriedades que caracterizam um subconjunto W de um espaço vetorial V não são "equações lineares homogêneas", em geral, W não é um subespaço de V. Como um exemplo, confira o Exemplo 3.14. Além disso, note a semelhança entre os vetores de F^4, $F^{2\times 2}$ e $\mathbb{P}_3(F)$:

$$(a, b, c, d), \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{e} \quad a + bx + cx^2 + dx^3.$$

Exercícios

1. Seja $V = \mathbf{R}^3$. Verifique quais dos subconjuntos abaixo são subespaços de V.

   (a) $W = \{(x, y, z) \in V : x + y + z = 0\}$.

   (b) $W = \{(x, y, z) \in V : x \leq y \leq z\}$.

   (c) $W = \{(x, y, z) \in V : x - 3z = 0\}$.

   (d) $W = \{(x, y, z) \in V : x \in \mathbb{Z}\}$.

   (e) $W = \{(x, y, z) \in V : x^2 + y^2 + z^2 \leq 1\}$.

   (f) $W = \{(x, y, z) \in V : x \geq 0\}$.

   (g) $W = \{(x, y, z) \in V : xy = 0\}$.

   (h) $W = \{(x, y, z) \in V : x = z^2\}$.

2. Seja $V = \mathbb{R}^{n\times n}$, $n \geq 2$. Verifique quais dos subconjuntos abaixo são subespaços de V.

   (a) $W = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in V : a = c \text{ e } b + d = 0 \right\}$.

(b) $W = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in V : a + d \leq b + c \right\}$.

(c) $W = \{\mathbf{A} \in V : \mathbf{AB} = \mathbf{BA},\ \ \mathbf{B}$ uma matriz fixa em $V\}$.

(d) $W = \{\mathbf{A} \in V : \mathbf{A}^2 = \mathbf{A}\}$.

(e) $W = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in V : ad - bc \neq 0 \right\}$.

(f) $W = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in V : ad - bc = 0 \right\}$.

3. Seja $V = \mathbb{P}_n(\mathbb{R})$, $n \geq 2$. Verifique quais dos subconjuntos abaixo são subespaços de V.

   (a) $W = \{p(x) \in V : p(0) = 0\}$.

   (b) $W = \{p(x) \in V : p(0) = 2p(1)\}$.

   (c) $W = \{p(x) \in V : p(x) + p'(x) = 0\}$

   (d) $W = \{p(x) \in V : p(2) = 0$ e $p(5) \neq 0\}$.

   (e) $W = \{p(x) \in V : p(x) = a_0 + a_2x^2 + \cdots + a_{2k}x^{2k}$ e $2k \leq n\}$, ou seja, W é o conjunto de todos os polinômios de potências pares.

4. Seja $V = \mathcal{F}(\mathbb{R}, \mathbb{R})$ o espaço vetorial de todas as funções reais. Verifique quais dos subconjuntos abaixo são subespaços de V.

   (a) $W = \{f \in V : f(0) = 1\}$.

   (b) $W = \{f \in V : f(5) = 0\}$.

   (c) $W = \{f \in V : f(3) = f(5)\}$.

   (d) $W = \{f \in V : f$ é contínua$\}$.

   (e) $W = \{f \in V : f$ é derivável$\}$.

   (f) $W = \{f \in V : f$ é integrável$\}$.

5. Sejam W_1, W_2 e W_3 os seguintes subespaços de $V = F^3$:

$$\begin{aligned} W_1 &= \{(x, y, z) \in V : x = z\},\ \ W_2 = \{(x, y, z) \in V : x = y = 0\}, \\ W_3 &= (x, y, z) \in V : x + y + z = 0\}. \end{aligned}$$

   É verdade que $W_1 + W_2 = W_1 + W_3 = W_2 + W_3 = F^3$? Em algum dos casos a soma é direta?

6. Sejam V um espaço vetorial sobre F e W_1, W_2 subespaços de V. Mostre que $V = W_1 \oplus W_2$ se, somente se, todo vetor $\mathbf{v}$ em V pode ser escrito de modo único sob a forma $\mathbf{v} = \mathbf{u}_1 + \mathbf{u}_2$, onde $\mathbf{u}_1 \in W_1$ e $\mathbf{u}_2 \in W_2$.

7. Considere
$$W_1 = \{(x, y) \in F^2 : y = x\}.$$
Encontre um subespaço W_2 de F tal que $F^2 = W_1 \oplus W_2$.

8. Sejam $V = \mathcal{F}(S, F)$ e
$$\begin{array}{rcl} V_p & = & \{f \in V : f(-x) = f(x),\ \ \forall\ x \in S\}, \\ V_i & = & \{f \in V : f(-x) = -f(x),\ \ \forall\ x \in S\} \end{array}$$
subespaços de V, com V_i o conjunto das *funções ímpares*. Mostre que $V = V_p \oplus V_i$.

9. Sejam $V = \mathcal{F}(\mathbb{R}, \mathbb{R})$ e $r \in \mathbb{R}_+^*$ fixado. Mostre que o conjunto
$$W_r = \{f \in V : f(x) = 0,\ \ \forall\ x \in [-r, r]\}$$
é um subespaço de V.

10. Sejam V um espaço vetorial sobre F e W_1, W_2 subespaços de V. Mostre que $W_1 \cup W_2$ é um subespaço de V se, e somente se, $W_1 \subseteq W_2$ ou $W_2 \subseteq W_1$.

11. Sejam V um espaço vetorial sobre F e W_1, W_2, W_3 subespaços de V.
    (a) Mostre que $(W_1 \cap W_2) + (W_1 \cap W_3) \subseteq W_1 \cap (W_2 + W_3)$.
    (b) Mostre que $W_1 + (W_2 \cap W_3) \subseteq (W_1 + W_2) \cap (W_1 + W_3)$.
    (c) Mostre, com um exemplo, que as inclusões acima podem ser estritas.
    (d) Mostre que se $W_3 \subseteq W_1$, então vale a igualdade.

12. Sejam V um espaço vetorial sobre F e W_1, W_2 subespaços de V tais que $V = W_1 \oplus W_2$. Diremos que um subespaço U de V é *adaptado* (mais tarde invariante sob um “operador linear” $T : V \to V$) a essa decomposição se
$$U = (U \cap W_1) \oplus (U \cap W_2).$$
    (a) Determine um exemplo de uma decomposição e um subespaço que não seja adaptado à decomposição.
    (b) Mostre que se $W_1 \subseteq U$ ou $W_2 \subseteq U$, então U é adaptado a decomposição.

3.3 Combinações Lineares

Sejam F um corpo e S um conjunto não vazio qualquer. Uma *soma formal* sobre S é uma expressão sob a forma $\sum_{s\in S} a_s s$, com $a_s = 0$, para todos exceto uma quantidade finita de $s \in S$. Pode ser provado que o conjunto das somas formais $F[S]$, munido com as operações

$$\sum_{s\in S} a_s s \oplus \sum_{s\in S} b_s s = \sum_{s\in S}(a_s + b_s)s \text{ e } c \odot \left(\sum_{s\in S} a_s s\right) = \sum_{s\in S}(ca_s)s,$$

é um espaço vetorial sobre F. O conjunto S chama-se *conjunto de geradores* e $F[S]$ chama-se *espaço vetorial livre* sobre S. Neste texto, nos restringimos, em geral, aos casos em que S é finito e $F[S]$ chama-se *espaço vetorial do tipo finito*. Caso contrário, *espaço vetorial do tipo infinito*.

Sejam V um espaço vetorial sobre F e $\alpha = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ uma lista sobre V ou em V. Diremos que um vetor $\mathbf{u} \in V$ é uma *combinação linear* dos vetores da lista α se existir uma lista $\mathbf{x} = (x_1, \ldots, x_n)$ sobre F tal que

$$\mathbf{u} = x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n = \sum_{i=1}^{n} x_i\mathbf{u}_i.$$

É muito importante o seguinte: cada lista $\alpha = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ sobre V fixada, induz uma função $L_\alpha : F^n \to V$ definida como $L_\alpha(\mathbf{x}) = x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n$, a qual preserva as operações e $L_\alpha(\mathbf{0}) = \mathbf{0}$. Por exemplo,

$$L_\alpha(\mathbf{x} + \mathbf{y}) = \sum_{i=1}^{n}(x_i + y_i)\mathbf{u}_i = \sum_{i=1}^{n} x_i\mathbf{u}_i + \sum_{i=1}^{n} y_i\mathbf{u}_i = L_\alpha(\mathbf{x}) + L_\alpha(\mathbf{y}).$$

Portanto, $F[\alpha] = \operatorname{Im} L_\alpha = \{x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n : x_i \in F\}$ é um subespaço de V, (prove isto!) o qual chama-se *subespaço gerado* por α e denotamos, também, por $F[\alpha] = F[\mathbf{u}_1, \ldots, \mathbf{u}_n]$. Além disso, $F[\alpha]$ é o menor subespaço de V contento os $\mathbf{u}_i$. Neste contexto, $F[S]$ é o subespaço gerado por S. Observe que $\mathbf{u} \in V$ é uma combinação linear dos vetores da lista α se, e somente se, $\mathbf{u} \in F[\alpha]$. Em particular, $V = F[\alpha]$ se, e somente se, L_α for sobrejetora. Portanto, quando V é um espaço vetorial do tipo finito como F^n, $F^{m\times n}$ ou $\mathbb{P}_n(F)$, o problema de decidir se $\mathbf{u} \in F[\alpha]$ é equivalente a discutir o sistema não homogêneo $\mathbf{AX} = \mathbf{B}$, com

$$\mathbf{A} = \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{pmatrix}, \ \ \mathbf{X} = (\mathbf{x}) \ \text{ e } \ \mathbf{B} = (\mathbf{u}).$$

Exemplo 3.21 *Sejam $V = F^4$ e $\alpha = ((1,1,-2,1),(3,0,4,-1),(-1,2,5,2))$ uma lista sobre V. Quais dos vetores $\mathbf{u} = (4,-5,9,-7)$, $\mathbf{v} = (3,1,-4,4)$ e $\mathbf{w} = (-1,1,0,1)$ são combinações lineares dos vetores da lista α.*

Solução. Sejam $\mathbf{u}_1 = (1,1,-2,1)$, $\mathbf{u}_2 = (3,0,4,-1)$e $\mathbf{u}_3 = (-1,2,5,2)$. Então, pelo exposto, para resolver esse problema devemos escalonar a matriz

$$\left(\begin{array}{rrr|l} 1 & 3 & -1 & b_1 \\ 1 & 0 & 2 & b_2 \\ -2 & 4 & 5 & b_3 \\ 1 & -1 & 2 & b_4 \end{array}\right) \to \cdots \to \left(\begin{array}{ccc|c} 1 & 0 & 0 & \frac{8b_1+19b_2-6b_3}{39} \\ 0 & 1 & 0 & \frac{3b_1-b_2+b_3}{13} \\ 0 & 0 & 1 & \frac{-4b_1+10b_2+3b_3}{39} \\ 0 & 0 & 0 & \frac{3b_1-14b_2+b_3+13b_4}{13} \end{array}\right),$$

com $\mathbf{u} = (b_1,b_2,b_3,b_4) \in V$. Assim, pelo Teorema 2.14, o vetor $\mathbf{u}$ é uma combinação linear dos vetores da lista α se, e somente se,

$$\frac{3b_1 - 14b_2 + b_3 + 13b_4}{13} = 0 \Leftrightarrow b_3 = -3b_1 + 14b_2 - 13b_4.$$

Logo, $\mathbf{u}$ é uma combinação linear dos vetores da lista α, pois

$$9 = -12 - 70 + 91 \ \text{ e } \ \mathbf{u} = -3\mathbf{u}_1 + 2\mathbf{u}_2 - \mathbf{u}_3.$$

É fácil verificar que $\mathbf{v}$ e $\mathbf{w}$ não são combinações lineares dos vetores da lista α. Uma outra solução seria escalonando a matriz

$$\left(\begin{array}{rrr|rrr} 1 & 3 & -1 & 4 & 3 & -1 \\ 1 & 0 & 2 & -5 & 1 & 1 \\ -2 & 4 & 5 & 9 & -4 & 0 \\ 1 & -1 & 2 & -7 & 4 & 1 \end{array}\right) \to \cdots \to \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & -3 & 0 & \frac{19}{43} \\ 0 & 1 & 0 & 2 & 0 & -\frac{12}{43} \\ 0 & 0 & 1 & -1 & 0 & \frac{14}{43} \\ 0 & 0 & 0 & 0 & 1 & -\frac{4}{43} \end{array}\right).$$

Portanto, $\mathbf{u} = -3\mathbf{u}_1 + 2\mathbf{u}_2 - \mathbf{u}_3$, mas $\mathbf{v}$ e $\mathbf{w}$ não são. ■

Exemplo 3.22 *Sejam $V = F^3$ e $\alpha = (\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)$ a lista dos vetores canônicos $\mathbf{e}_i = (\delta_{i1},\delta_{i2},\delta_{i3}) \in V$, com $i = 1,2,3$. Determine $W = F[\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3]$.*

Solução. Note que

$$\begin{aligned} W &= \{x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3 : x, y, z \in F\} \\ &= \{x(1,0,0) + y(0,1,0) + z(0,0,1) : x, y, z \in F\} \\ &= \{(x,y,z) : x, y, z \in F\}. \end{aligned}$$

Portanto, $W = V$, isto é, $L_\alpha : F^3 \to V$ é sobrejetora. ■

Exemplo 3.23 *Sejam $V = F^{2\times 2}$ e $\alpha = (\mathbf{E}_{11}, \mathbf{E}_{12}, \mathbf{E}_{21}, \mathbf{E}_{22})$ a lista dos vetores canônicos $\mathbf{E}_{pq} = (\delta_{ip}\delta_{qj}) \in V$, com $p, q = 1, 2$. Determine o subespaço gerado por α, ou seja, $W = F[\mathbf{E}_{11}, \mathbf{E}_{12}, \mathbf{E}_{21}, \mathbf{E}_{22}]$.*

Solução. Note que

$$\begin{aligned} W &= \{a\mathbf{E}_{11} + b\mathbf{E}_{12} + c\mathbf{E}_{21} + d\mathbf{E}_{22} : a, b, c, d \in \mathbb{R}\} \\ &= \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in F \right\}. \end{aligned}$$

Portanto, $W = V$, isto é, $L_\alpha : F^4 \to V$ é sobrejetora. ■

Exemplo 3.24 *Sejam $V = \mathbb{P}_3(F)$ e $\alpha = (p_0, p_1, p_2, p_3)$ a lista dos vetores canônicos $p_i = x^i \in V$, com $i = 0, 1, 2, 3$. Determine $W = F[p_0, p_1, p_2, p_3]$.*

Solução. Note que

$$\begin{aligned} W &= \{a_0p_0 + a_1p_1 + a_2p_2 + a_3p_3 : a_0, a_1, a_2, a_3 \in F\} \\ &= \{a_0 + a_1x + a_2x^2 + a_3x^3 : a_0, a_1, a_2, a_3 \in F\}. \end{aligned}$$

Portanto, $W = V$, isto é, $L_\alpha : F^4 \to V$ é sobrejetora. ■

Exemplo 3.25 *Sejam V um espaço vetorial sobre F e W_1, W_2 subespaços de V. Mostre que $W_1 + W_2$ é o menor subespaço de V contendo W_1 e W_2, isto é,*

$$W_1 + W_2 = [W_1, W_2] = [W_1 \cup W_2].$$

Solução. Já vimos, pelo Teorema 3.18, que $W_1 + W_2$ é um subespaço de V. Como $\mathbf{w}_1 = \mathbf{w}_1 + \mathbf{0} \in W_1 + W_2$ e $\mathbf{w}_2 = \mathbf{0} + \mathbf{w}_2 \in W_1 + W_2$ temos que $W_1 \subseteq W_1 + W_2$ e $W_2 \subseteq W_1 + W_2$. Logo,

$$W_1 \cup W_2 \subseteq W_1 + W_2 \text{ e } [W_1 \cup W_2] \subseteq W_1 + W_2.$$

Por outro lado, se $\mathbf{w} \in W_1 + W_2$, então existe um $\mathbf{w}_1 \in W_1$ e $\mathbf{w}_2 \in W_2$ tal que

$$\mathbf{w} = \mathbf{w}_1 + \mathbf{w}_2 = 1 \cdot \mathbf{w}_1 + 1 \cdot \mathbf{w}_2.$$

Assim, qualquer vetor $\mathbf{w} \in W_1 + W_2$ é uma combinação linear de vetores em $W_1 \cup W_2$. Consequentemente, $W_1 + W_2 \subseteq [W_1 \cup W_2]$. Portanto, $W_1 + W_2 = [W_1 \cup W_2]$. Finalmente, seja W qualquer subespaço de V tal que $W_1 \subseteq W$ e $W_2 \subseteq W$. Então $W_1 \cup W_2 \subseteq W$ e $[W_1 \cup W_2] \subseteq W$, pois qualquer vetor de $[W_1 \cup W_2]$ é uma combinação linear de vetores em $W_1 \cup W_2$ e W é um subespaço de V. Portanto, $W_1 + W_2 \subseteq W$. ■

Exemplo 3.26 *Determine todos os subespaços de F^2.*

Solução. Seja W um subespaço qualquer de F^2. Se $W = \{(0,0)\}$, nada há para ser provado. Se $W \neq \{(0,0)\}$, então W contém um vetor $\mathbf{u} = (a,b)$, com $\mathbf{u} \neq (0,0)$. Em particular, $F[\mathbf{u}] \subseteq W$. Se W contém um vetor $\mathbf{v} = (c,d)$ tal que $\mathbf{v} \neq x\mathbf{u}$, para todo $x \in F$, então $ad - bc \neq 0$. Caso contrário, existiria $t \in F^\star$ tal que $ad = bc = t$, de modo que $\mathbf{v} = t\mathbf{u}$, o que é impossível. Neste caso, $W = F[\mathbf{u},\mathbf{v}] = F^2$. De fato, dado $\mathbf{w} = (x,y) \in F^2$, devemos provar que existem $r, s \in F$ tais que $\mathbf{w} = r\mathbf{u} + s\mathbf{v}$. Mas isto é equivalente a determinar condições sobre r e s, de modo que o sistema não homogêneo

$$\begin{cases} ar + cs = x \\ br + ds = y. \end{cases}$$

tenha solução. Como $ad - bc \neq 0$ temos, pela Regra de Cramer a seguir, que o sistema possui uma única solução

$$r = \frac{dx - by}{ad - bc} \text{ e } r = \frac{ax - cy}{ad - bc}.$$

Uma outra solução, seja W um subespaço qualquer de F^2. Então

$$\begin{aligned} W_1 &= \{x \in F : x\mathbf{e}_1 + y\mathbf{e}_2 = (x,y) \in W, \text{ para algum } y \in F\} \text{ e} \\ W_2 &= \{y \in F : y\mathbf{e}_2 = (0,y) \in W\} \end{aligned}$$

são subespaços de F (prove isto!). Assim, existem $x_0, y_1 \in F$ tais que $W_1 = F[x_0]$ e $W_2 = F[y_1]$. Logo, pela definição, podemos encontrar $y_0 \in F$ tal que $\mathbf{u}_0 = (x_0, y_0) \in W$ e $\mathbf{u}_1 = (0, y_1) \in W$.

Afirmação. $W = F[\mathbf{u}_0, \mathbf{u}_1]$.

De fato, dado $\mathbf{u} = (x,y) \in W$, com $x \in W_1$, de modo que $x = ax_0$, para algum $a \in F$. Logo, $\mathbf{u} - a\mathbf{u}_0 = (0, y - ay_0) \in W$ implica que $y - ay_0 \in W_2$. Portanto, $y - ay_0 = by_1$, para algum $b \in F$. Consequentemente,

$$\mathbf{u} = (x, y) = (ax_0, ay_0 + by_1) = a\mathbf{u}_0 + b\mathbf{u}_1,$$

isto é, $W = F[\mathbf{u}_0, \mathbf{u}_1]$. ■

Exercícios

1. Mostre que todo vetor em F^2 pode ser escrito como combinação linear dos vetores $(1, 2)$ e $(5, 0)$. Que relação existe entre F^2 e $F[(1, 2), (5, 0)]$?

2. Sejam $V = \mathbb{P}_2(F)$ e $\alpha = (f(x) = 2-3x+5x^2, g(x) = -8+5x-2x^2)$ uma lista de vetores em V. Quais dos vetores $p(x) = -26+11x+7x^2$ e $q(x) = 1+x+x^2$ são combinações lineares dos vetores da lista α?

3. Sejam $V = F^3$ e $\alpha = (\mathbf{u}_1 = (1, 1, -2), \mathbf{u}_2 = (3, 0, 4))$ uma lista de vetores em V. Quais dos vetores $\mathbf{u} = (4, -5, 9)$, $\mathbf{v} = (3, 1, -4)$ e $\mathbf{w} = (-1, 1, 0)$ são combinações lineares dos vetores da lista α?

4. Sejam $V = F^3$ e $\alpha = (\mathbf{u}_1 = (1, 1, -2), \mathbf{u}_2 = (3, 0, 4), \mathbf{u}_3 = (-1, 1, 0))$ uma lista de vetores em V. Determine o valor de k de modo que $(4, -5, k) \in F[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3]$.

5. Sejam $V = \mathbb{P}_3(F)$ e $\alpha = (p_0(x), p_1(x), p_2(x), p_3(x))$ uma lista de vetores $p_i(x) = (1-x)^i \in V$, com $i = 0, 1, 2, 3$. Quais dos vetores em V são combinações lineares dos vetores $p_0(x), p_1(x), p_2(x)$ e $p_3(x)$?

6. Sejam $V = F^{2\times 2}$ e

$$\alpha = \left(\mathbf{A}_1 = \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix}, \mathbf{A}_2 = \begin{pmatrix} 3 & 0 \\ 4 & -1 \end{pmatrix}, \mathbf{A}_3 = \begin{pmatrix} -1 & 2 \\ 5 & 2 \end{pmatrix}\right)$$

uma lista de vetores em V.Quais dos vetores

$$\mathbf{A} = \begin{pmatrix} 4 & -5 \\ 9 & -7 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} 3 & 1 \\ -4 & 4 \end{pmatrix} \text{ e } \mathbf{C} = \begin{pmatrix} -1 & 1 \\ -2 & 1 \end{pmatrix}$$

são combinações lineares dos vetores da lista α?

7. Encontre os geradores para os seguintes subespaços de F^3:

   (a) $W_1 = \{(x, y, z) \in F^3 : x - y = 0\}$.

(b) $W_2 = \{(x, y, z) \in F^3 : x + z = x - 2y = 0\}$.

(c) $W_3 = \{(x, y, z) \in F^3 : x + 2y - 3z = 0\}$.

(d) $W_1 \cap W_2$.

(e) $W_2 + W_3$.

8. Sejam $V = F^4$ e $W = \{(x, y, z, t) \in V : x + 2y - 2z = 0 \text{ e } t = 0\}$ um subespaço de V. Quais dos vetores $\mathbf{u} = (-2, 4, 3, 0), \mathbf{v} = (6, 2, 4, 1)$ e $\mathbf{w} = (-2, 1, 0, 0)$ estão em W?

9. Sejam V um espaço vetorial sobre $\mathbb{R}$, $\tilde{V}$ sua complexificação e W qualquer subespaço de $\tilde{V}$. Mostre que existe um subespaço U de V tal que $W = \tilde{U} = \{\mathbf{u} + i\mathbf{v} : \mathbf{u}, \mathbf{v} \in U\}$ se, e somente se, $J(W) \subseteq W$, em que $J : \tilde{V} \to \tilde{V}$ definida como $J(\mathbf{u} + i\mathbf{v}) = \mathbf{u} - i\mathbf{v}$ é a *função conjugação*.

10. Sejam V um espaço vetorial sobre F e W um subespaço de V. Diremos que W é um *subespaço minimal* de V se W é um subespaço próprio de V e não existir um subespaço U de V tal que $\{\mathbf{0}\} \subset U \subset W$. Mostre que qualquer subespaço minimal de V é da forma $[\mathbf{u}]$, para algum $\mathbf{u} \in V - \{\mathbf{0}\}$.

3.4 Dependência e Independência Linear

Sejam V um espaço vetorial sobre F e $\alpha = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ uma lista sobre V. Diremos que os vetores da lista α são *linearmente dependentes* (LD) se existir uma lista $\mathbf{x} = (x_1, \ldots, x_n)$ não nula sobre F tal que

$$x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n = \mathbf{0} \tag{3.1}$$

ou, equivalentemente, a equação vetorial (3.1) admite uma solução não nula. Caso contrário, diremos que os vetores da lista α são *linearmente independentes* (LI) ou, equivalentemente, a equação vetorial (3.1) admite apenas a solução nula.

É muito importante o seguinte: a equação vetorial (3.1) é equivalente ao conjunto

$$\ker L_\alpha = \{\mathbf{x} \in F^n : L_\alpha(\mathbf{x}) = \mathbf{0}\}.$$

Lembrando que $L_\alpha : F^n \to V$ é a função definida como

$$L_\alpha(\mathbf{x}) = x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n,$$

Neste caso, $\ker L_\alpha$ é um subespaço de F^n (prove isto!). Portanto, a lista α é LI se, e somente se, L_α é injetora, ou seja, $\ker L_\alpha = \{\mathbf{0}\}$. Quando V é um espaço vetorial do tipo finito como F^n, $F^{m\times n}$ ou $\mathbb{P}_n(F)$, o problema de decidir se $\mathbf{x} \in \ker L_\alpha$ é equivalente a discutir o sistema homogêneo $\mathbf{AX} = \mathbf{O}$, com

$$\mathbf{A} = \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{pmatrix} \text{ e } \mathbf{X} = (\mathbf{x}).$$

Portanto, $\ker L_\alpha$ é o espaço solução de $\mathbf{AX} = \mathbf{O}$ e a nulidade $\nu(\mathbf{A}) = \dim(\ker L_\alpha)$. É interessante, de um ponto de vista teórico e didático, termos uma noção contável de um conjunto LI ou LD. Sejam $\beta = (\mathbf{u}_n)_{n\in\mathbb{N}}$ qualquer sequência em $\operatorname{Seq}(V)$ e $\mathbf{x} = (x_n)_{n\in\mathbb{N}}$ qualquer sequência de suporte finito em $\operatorname{Seq}(F)$. Diremos que β é LI se

$$L_\beta(\mathbf{x}) = \sum_{n\in\mathbb{N}} x_n\mathbf{u}_n = \mathbf{0} \Rightarrow x_n = 0, \;\; \forall \;\; n \in \mathbb{N}.$$

Intuitivamente, qualquer subconjunto finito de β é LI. Caso contrário, β é LD.

Exemplo 3.27 *Sejam $V = F^3$ e $\alpha = (\mathbf{u}_1 = (3, 0, -3), \mathbf{u}_2 = (-1, 1, 2), \mathbf{u}_3 = (4, 2, -2), \mathbf{u}_4 = (2, 1, 1))$ uma lista de vetores sobre V. Verifique se os vetores da lista α são LI ou LD.*

Solução. Pelo exposto, para resolver esse problema devemos escalonar a matriz

$$\begin{pmatrix} 3 & -1 & 4 & 2 \\ 0 & 1 & 2 & 1 \\ -3 & 2 & -2 & 1 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Assim, a nulidade $\nu(\mathbf{A}) = 1$ e a lista α é LD. Outro modo, nosso sistema é equivalente ao sistema

$$\begin{cases} x_1 + 2x_3 & = & 0 \\ x_2 + 2x_3 & = & 0 \\ x_4 & = & 0. \end{cases}$$

Escolhendo, $x_3 = t \in F$, temos que

$$\ker L_\alpha = \{(-2t, -2t, t, 0) \in F_4 : t \in F\}$$

é o conjunto solução do sistema. Em particular, se $t = 1$, então $(-2, -2, 1, 0)$ é uma solução não nula do sistema. Portanto, a lista α é LD, isto é, $L_\alpha : F^4 \to V$ não é injetora e $-2\mathbf{u}_1 - 2\mathbf{u}_2 + \mathbf{u}_3 + 0\mathbf{u}_4 = \mathbf{0}$. ■

Exemplo 3.28 *Sejam* $V = \mathcal{F}(\mathbb{R}, \mathbb{R})$ *e* $\alpha = (\mathbf{u}_1 = e^x, \mathbf{u}_2 = e^{2x})$ *uma lista de vetores sobre* V. *Verifique se os vetores da lista* α *são* LI *ou* LD. *Note que* $\mathbf{u}_1$ *e* $\mathbf{u}_2$ *são soluções da equação diferencial* $y'' - 3y' + 2y = 0$.

Solução. Para resolver esse problema devemos discutir a equação vetorial

$$ae^x + be^{2x} = \mathbf{0}, \ \ \forall \ x \in \mathbb{R},$$

onde $\mathbf{0}$ é a função identicamente nula. Diferenciando ambos os membros dessa equação, temos que

$$ae^x + 2be^{2x} = \mathbf{0}, \ \ \forall \ x \in \mathbb{R}.$$

Assim, subtraindo a primeira equação da segunda, resulta que $be^{2x} = \mathbf{0}$, para todo $x \in \mathbb{R}$. Em particular, se $x = 0$, então $b = 0$ e, da primeira equação, $ae^x = \mathbf{0}$. Logo, $a = 0$. Portanto, os vetores da lista α são LI. ■

Exemplo 3.29 *Seja* $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{n\times n}$ *tal que* $a_{ij} < 0$*, se* $i \neq j$ *e* $\sum_{k=1}^{n} a_{ik} > 0$*, para* $i = 1, \ldots, n$. *Mostre que* $\mathbf{A}$ *é não singular.*

Solução. Suponhamos, por absurdo, que $\mathbf{A}$ seja singular. Então, pelo Teorema 2.12, $\rho(\mathbf{A}) < n$, ou seja, as colunas de $\mathbf{A}$ são LD. Assim, existem escalares $x_1, \ldots, x_n \in \mathbb{R}$, não todos nulos, tais que

$$x_1\mathbf{C}_1 + \cdots + x_n\mathbf{C}_n = \mathbf{O} \Leftrightarrow \sum_{k=1}^{n} a_{ik}x_k = 0, \ \ i = 1, \ldots, n. \tag{3.2}$$

Pondo $|x_j| = \max\{|x_1|, , \ldots, |x_n|\}$ e multiplicando a solução do sistema (3.2) por -1, se necessário, podemos supor que $x_j > 0$. Considerando a j-ésima equação do sistema (3.2) temos, para $j \neq k$, que $a_{jk} < 0$, de modo que

$$0 = \sum_{k=1}^{n} a_{jk}x_k = a_{jj}x_j + \sum_{k\neq j} a_{jk}x_k \geq a_{jj}x_j + \sum_{k\neq j} a_{jk}x_j = (\sum_{k=1}^{n} a_{jk})x_j > 0,$$

o que é uma contradição. ■

Exemplo 3.30 (Regra de Cramer) [2] *Seja* $\mathbf{A} \in F^{n\times n}$. *Mostre que se existir um* $\mathbf{X} = (x_1, \ldots, x_n)^t \in F^{n\times 1}$ *tal que* $\mathbf{B} = \mathbf{AX} = x_1\mathbf{C}_1 + \cdots + x_n\mathbf{C}_n \in F^{n\times 1}$*, então*

[2] Gabriel Cramer, 1704-1752, matemático suíço.

$x_j \det \mathbf{A} = \det \mathbf{I}_j(\mathbf{X}) \det \mathbf{A} = \det \mathbf{A}_j(\mathbf{B})$, *com* $\mathbf{A}\mathbf{E}_i = \mathbf{C}_i$ *e*

$$\mathbf{A}_j(\mathbf{B}) = \begin{pmatrix} \mathbf{C}_1 & \cdots & \mathbf{C}_{j-1} & \mathbf{B} & \mathbf{C}_{j+1} & \cdots & \mathbf{C}_n \end{pmatrix} = \mathbf{A} \cdot \mathbf{I}_j(\mathbf{X}).$$

Em particular, se $\det \mathbf{A} \neq 0$, *então*

$$x_j = a_{j1}^{-1} b_1 + \cdots + a_{jn}^{-1} b_n = \frac{1}{\det \mathbf{A}}(c_{1j} b_1 + \cdots + c_{nj} b_n), c_{ij} = (-1)^{i+j} \det \mathbf{A}_{ij}.$$

Solução. Note que $\mathbf{B} = x_1 \mathbf{C}_1 + \cdots + x_n \mathbf{C}_n$ implica que

$$x_1 \mathbf{C}_1 + \cdots + x_{j-1} \mathbf{C}_{j-1} + 1 \cdot (x_j \mathbf{C}_j - \mathbf{B}) + x_{j+1} \mathbf{C}_{j+1} + \cdots + x_n \mathbf{C}_n = \mathbf{O}.$$

Assim, as colunas da matriz

$$\begin{pmatrix} \mathbf{C}_1 & \cdots & \mathbf{C}_{j-1} & x_j \mathbf{C}_j - \mathbf{B} & \mathbf{C}_{j+1} & \cdots & \mathbf{C}_n \end{pmatrix}$$

são LD. Logo, pelos itens (1) e (2) do Teorema 1.24, temos que

$$\begin{aligned} 0 &= \det \begin{pmatrix} \mathbf{C}_1 & \cdots & \mathbf{C}_{j-1} & x_j \mathbf{C}_j - \mathbf{B} & \mathbf{C}_{j+1} & \cdots & \mathbf{C}_n \end{pmatrix} \\ &= x_j \det \mathbf{A} - \det \begin{pmatrix} \mathbf{C}_1 & \cdots & \mathbf{C}_{j-1} & \mathbf{B} & \mathbf{C}_{j+1} & \cdots & \mathbf{C}_n \end{pmatrix}. \end{aligned}$$

Portanto, $x_j \det \mathbf{A} = \det \mathbf{A}_j(\mathbf{B})$. Uma outra prova é aplicando a Fórmula de Binet-Cauchy. ■

Observe que se $\mathbf{A} \in F^{n \times n}$ for não singular, então $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$ e a j-ésima coluna $\mathbf{X}_j$ da matriz inversa $\mathbf{A}^{-1} = (b_{ij})$ é tal que $\mathbf{A}\mathbf{X}_j = \mathbf{E}_j$. Portanto, pela Regra de Cramer,

$$b_{ij} = x_i = \frac{\det \mathbf{A}_i(\mathbf{E}_j)}{\det \mathbf{A}} = \frac{(-1)^{i+j} \det \mathbf{A}_{ji}}{\det \mathbf{A}}.$$

Já vimos, no Exemplo 3.28, como provar que um conjunto simples de funções era LI ou LD. É muito útil, de um ponto de vista teórico e didático, a sua generalização baseada no *determinante Wronskiano*.[3] Sejam $V = \mathcal{C}^\infty([a, b], \mathbb{R})$ o espaço das funções de classe infinita, isto é, $f^{(n)}$ existe, para todo $n \in \mathbb{Z}_+$. O determinante Wronskiano de $f_1, \ldots, f_n \in V$ é definido como $W(x) = \det \mathbf{A}(x)$, com $\mathbf{A}(x) = (f_j^{(i-1)}(x)) \in \mathbb{R}^{n \times n}$ e $f_j^{(0)}(x) = f_j(x)$, para todo $x \in [a, b]$.

[3]Josef Maria Hoëné-Wronski, 1776-1853, filósofo e matemático polonês.

Exemplo 3.31 *Sejam $\alpha = (f_1, \ldots, f_n)$ uma lista em $\mathcal{C}^\infty([a,b], \mathbb{R})$ e $W(x_0) \neq 0$, para algum $x_0 \in [a,b]$. Então α é LI.*

Solução. Suponhamos que $f = c_1 f_1 + \cdots + c_n f_n = 0$. Então $f(x) = 0$, para todo $x \in [a,b]$. Assim, por derivação sucessiva, obtemos o sistema homogêneo

$$\left\{\begin{array}{rcl} c_1 f_1^{(0)}(x) + \cdots + c_n f_n^{(0)}(x) & = & 0 \\ c_1 f_1^{(1)}(x) + \cdots + c_n f_n^{(1)}(x) & = & 0 \\ & \vdots & \\ c_1 f_1^{(n-1)}(x) + \cdots + c_n f_n^{(n-1)}(x) & = & 0 \end{array}\right. \Leftrightarrow \mathbf{A}(x)\mathbf{X} = \mathbf{O},$$

para todo $x \in [a,b]$. Em particular, avaliando em $x_0 \in [a,b]$, obtemos $\det \mathbf{A}(x_0) = W(x_0) \neq 0$ e, pela Regra de Cramer, $c_1 = \cdots = c_n = 0$. Portanto, α é LI. ■

Sejam V um espaço vetorial sobre F e $\alpha = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ uma lista sobre V. Diremos que uma combinação linear dos vetores da lista α, $\mathbf{u} = \sum_{i=1}^n x_i \mathbf{u}_i$, é *dependente* de $\mathbf{u}_k$ se $x_k \neq 0$. Caso contrário, é *independente* de $\mathbf{u}_k$. Note que se $x_k \neq 0$, então

$$\mathbf{u}_k = \sum_{i \neq k} (-x_i x_k^{-1}) \mathbf{u}_i + x_k^{-1} \mathbf{u}.$$

Portanto, podemos substituir $\mathbf{u}_k$ por $\mathbf{u}$ para obter a "nova" lista

$$\beta = (\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}, \mathbf{u}, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n).$$

Neste caso, $F[\alpha] = F[\beta]$. De fato, $\mathbf{v} \in F[\alpha]$ se, e somente se,

$$\mathbf{v} = \sum_{i \neq k} y_i \mathbf{u}_i + y_k \mathbf{u}_k = \sum_{i \neq k} (y_i - x_i x_k^{-1} y_k) \mathbf{u}_i + (x_k^{-1} y_k) \mathbf{u}$$

se, e somente se, $\mathbf{v} \in F[\beta]$. Por um argumento indutivo, podemos substituir $\mathbf{u}_{k_1}, \ldots, \mathbf{u}_{k_m}$, com $1 \leq k_1 < \cdots < k_m \leq n$, por $\mathbf{v}_1, \ldots, \mathbf{v}_m$ para obter a lista

$$\beta = (\mathbf{u}_1, \ldots, \mathbf{u}_{k_1 - 1}, \mathbf{v}_1, \mathbf{u}_{k_1 + 1}, \ldots, \ldots, \mathbf{v}_m, \ldots, \mathbf{u}_n).$$

Teorema 3.32 *Sejam V um espaço vetorial sobre F e $\alpha = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ uma lista sobre V. Então α é LD se, e somente se, pelo menos um desses vetores for zero ou uma combinação linear dos outros.*

Prova. Suponhamos que α seja LD. Então, por definição, a combinação linear

$$x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n = \mathbf{0}$$

é dependente de algum vetor, digamos $\mathbf{u}_k \neq \mathbf{0}$. Assim, $\mathbf{u}_k = \sum_{i\neq k}(-x_i x_k^{-1})\mathbf{u}_i$. Portanto, $\mathbf{u}_k$ é uma combinação linear dos outros. Reciprocamente, suponhamos que um desses vetores seja uma combinação linear dos outros, digamos

$$\mathbf{u}_k = x_1\mathbf{u}_1 + \cdots + x_{k-1}\mathbf{u}_{k-1} + x_{k+1}\mathbf{u}_{k+1} + \cdots + x_n\mathbf{u}_n.$$

Logo, a combinação linear

$$x_1\mathbf{u}_1 + \cdots + x_{k-1}\mathbf{u}_{k-1} + (-1)\mathbf{u}_k + x_{k+1}\mathbf{u}_{k+1} + \cdots + x_n\mathbf{u}_n = \mathbf{0}$$

é dependente de $\mathbf{u}_k$. Portanto, α é LD. ■

Corolário 3.33 *Sejam V um espaço vetorial sobre F e $\alpha = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ uma lista sobre V. Então α é LD se, e somente se, um desses vetores for zero ou uma combinação linear dos precedentes.*

Prova. Suponhamos que α seja LD. Então, por definição, a combinação linear

$$x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n = \mathbf{0}$$

é dependente de algum vetor. Assim, o conjunto $S = \{m \in \mathbb{N} : x_m \neq 0\}$ é não vazio e possui um maior elemento, digamos $k \in S$ e $x_k \neq 0$. Se $k = 1$, então $x_1\mathbf{u}_1 = \mathbf{0}$, de modo que $\mathbf{u}_1 = \mathbf{0}$, acabou. Caso contrário, $k > 1$ e

$$\mathbf{u}_k = (-x_1 x_k^{-1})\mathbf{u}_1 + \cdots + (-x_{k-1} x_k^{-1})\mathbf{u}_{k-1}.$$

A recíproca é clara. ■

Por exemplo, se $V = F^2$ e $\alpha = (\mathbf{u}_1 = (1, -1), \mathbf{u}_2 = (1, 1), \mathbf{u}_3 = (1, 0))$ uma lista sobre V, então α é LD, pois $\mathbf{u}_3 = 2^{-1}(\mathbf{u}_1 + \mathbf{u}_2)$.

Exercícios

1. Sejam $V = F^n$ e $\mathbf{u} = (x_1, \ldots, x_n), \mathbf{v} = (y_1, \ldots, y_n) \in V$. Mostre que $\mathbf{u}$ e $\mathbf{v}$ são LD se, e somente se, existir um $a \in F$ tal que $y_i = ax_i$, $i = 1, \ldots, n$.

2. Sejam V um espaço vetorial sobre F e $\alpha = (\mathbf{u}, \mathbf{v}, \mathbf{w})$ uma lista sobre V. Mostre que se α é LI, então:

   (a) $\beta = (\mathbf{u} + \mathbf{v} - 2\mathbf{w}, \mathbf{u} - \mathbf{v} - \mathbf{w}, \mathbf{u} + \mathbf{w})$ é uma lista LI.

   (b) $\gamma = (\mathbf{u} + \mathbf{v} - 3\mathbf{w}, \mathbf{u} + 3\mathbf{v} - \mathbf{w}, \mathbf{v} + \mathbf{w})$ é uma lista LD.

3. Seja $\alpha = ((a, b), (c, d))$ uma lista sobre F^2. Mostre que α é LD se, e somente se, $ad = bc$.

4. A lista $\alpha = (1, x, x^2, 2 + x + 2x^2)$ é LI ou LD sobre $\mathbb{P}_2(F)$? O que se pode afirmar a respeito de qualquer uma de suas listas com três elementos?

5. Sejam $W_1 = F[(1, 0, 0), (0, 1, 0)]$ e $W_2 = F[(1, 2, 3), (1, -1, 1)]$ subespaços de F^3. Encontre um vetor $\mathbf{u} \in F^3$ tal que $F[\mathbf{u}] = W_1 \cap W_2$.

6. Em quais condições sobre k, a lista $\alpha = ((1, 0, k), (1, 1, k), (1, 1, k^2))$ é LI sobre F^3?

7. Seja $V = \mathcal{C}(\mathbb{R}, \mathbb{R})$ o espaço das funções reais contínuas. Quais das listas abaixo são LI sobre V.

   (a) $(x, x + 1, x^2 - 1)$.

   (b) $(x + 5, x^2 - x, x^2 + x - 10)$.

   (c) $((x + 1)^2, 2x, x + \frac{1}{2})$.

   (d) $((x + 1)^2, x^2 - 1, x + 1)$.

   (e) $(1 - x, x(1 - x), 1 - x^2)$.

   (f) $(1, e^x, e^{-x})$.

   (g) $(\operatorname{sen} x, \cos x, \operatorname{tg} x)$.

8. Seja V um espaço vetorial sobre F. Mostre que a lista $\alpha = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ sobre V é LI se, e somente se, a lista $\beta = (\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}, \mathbf{u}, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n)$ sobre V é LI, para toda substituição $\mathbf{u}$.

9. Sejam V um espaço vetorial sobre F e $\alpha = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ uma lista sobre V. Mostre que se $\mathbf{v}_1, \ldots, \mathbf{v}_m \in F[\alpha]$ são LI, então alguns dos m $\mathbf{u}_k$ podem ser substituídos pelos $\mathbf{v}_i$, com $i = 1, \ldots, m$.

3.5 Bases e Dimensão

Sejam V um espaço vetorial sobre F e $\alpha = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ uma lista sobre V. Diremos que α é uma *base* de V se as seguintes condições são satisfeitas:

1. α é LI.
2. $V = F[\alpha]$.

ou, equivalentemente, a função $L_\alpha : F^n \to V$ é bijetora. Note que

$$V = F[\mathbf{u}_1] \oplus \cdots \oplus F[\mathbf{u}_n].$$

Sejam $\alpha = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ uma lista sobre V e $\sigma \in S_n$, então α é uma base de V se, e somente se, $\beta = \alpha \circ \sigma = (\mathbf{u}_{\sigma(1)}, \ldots, \mathbf{u}_{\sigma(n)})$ for uma base de V. De fato, como a adição sobre V é associativa e comutativa temos que

$$\sum_{i=1}^{n} x_i \mathbf{u}_i = \sum_{i=1}^{n} x_{\sigma(i)} \mathbf{u}_{\sigma(i)},$$

para todo $\sigma \in S_n$. Portanto, podemos considerar uma base ordenada α em V como um subconjunto, sem nenhuma ordem particular, $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ uma base de V. Vale lembrar que os elementos do conjunto são distintos.

É muito importante, de um ponto de vista teórico e didático, termos uma noção geral de uma base qualquer. Sejam V um espaço vetorial sobre F e S um conjunto não vazio qualquer. Uma *família* (indexada) sobre V é qualquer função $f : S \to V$ definida como $f(s) = \mathbf{u}_s$ e $f = \{\mathbf{u}_s : s \in S\} = (\mathbf{u}_s)_{s \in S}$. Neste contexto, o conjunto

$$F^{(S)} = \{\mathbf{x} \in \mathcal{F}(S, F) : \operatorname{supp}(\mathbf{x}) \text{ é finito}\}$$

é um subespaço de $\mathcal{F}(S, F)$, pois

$$\operatorname{supp}(\mathbf{x} + \mathbf{y}) \subseteq \operatorname{supp}(\mathbf{x}) \cup \operatorname{supp}(\mathbf{y}) \text{ e } \operatorname{supp}(a\mathbf{x}) \subseteq \operatorname{supp}(\mathbf{x}).$$

Observe que cada $s \in S$ fixado, induz uma função $\mathbf{e}_s : S \to F$ definida como

$$\mathbf{e}_s(t) = \delta_{st} = \begin{cases} 1, & \text{se } s = t \\ 0, & \text{se } s \neq t. \end{cases}$$

Como $\operatorname{supp}(\mathbf{e}_s) = \{s\}$ temos que $\mathbf{e}_s \in F^{(S)}$, para todo $s \in S$. Por outro lado, dado

$\mathbf{x} \in F^{(S)}$, com $\operatorname{supp}(\mathbf{x}) = \{s_1, \ldots, s_n\}$, é fácil verificar que

$$\mathbf{x} = x_{s_1}\mathbf{e}_{s_1} + \cdots + x_{s_n}\mathbf{e}_{s_n},$$

em que $x_{s_i} = \mathbf{x}(s_i), i = 1, \ldots, n$. Portanto, $F^{(S)} = F[\alpha]$, com $\alpha = \{\mathbf{e}_s : s \in S\}$. Em particular, se $S = I_n = \{1, \ldots, n\}$, então $F^{(S)} = F^n$ e $\alpha = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ é a *base canônica* de F^n, pois

$$\mathbf{x} = (x_1, \ldots, x_n) = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n.$$

Finalmente, cada família $\alpha = \{\mathbf{u}_s : s \in S\} = (\mathbf{u}_s)_{s\in S}$ sobre V induz uma função $L_\alpha : F^{(S)} \to V$ definida como

$$L_\alpha(\mathbf{x}) = \sum_{s\in S} x_s\mathbf{u}_s, \quad \forall \ \mathbf{x} = (x_s)_{s\in S} \in F^{(S)}.$$

Consequentemente, α é um base de V se, e somente se, L_α for bijetora.

Teorema 3.34 (Existência de Base) *Qualquer espaço vetorial possui uma base.*

Prova. Usa o Lema de Zorn.[4] ■

Exemplo 3.35 *Sejam $V = F[x]$ e $\alpha = \{x^{n-1} : n \in \mathbb{N}\}$ uma sequência sobre V. Mostre que α é uma base infinita de V, a qual chama-se* base canônica *de V.*

Solução. Sejam $p_i(x) = x^i, \ldots, p_{i+n}(x) = x^{i+n}$ vetores distintos de V, com $i \geq 0$. Se

$$c_1 p_i(x) + \cdots + c_n p_{i+n}(x) = \mathbf{0},$$

então $c_1 = \cdots = c_n = 0$. Assim, α é LI. É claro que $F[\alpha] = V$, pois qualquer vetor $p(x)$ em V é da forma $p(x) = a_0 + a_1 x + \cdots + a_n x^n$. Portanto, α é uma base infinita de V. ■

Seja V um espaço vetorial sobre F. Diremos que V é de *dimensão finita* se ele possui uma base finita, por exemplo, $V = F^n$ é de dimensão finita. Caso contrário, V é de *dimensão infinita*.

Seja $\operatorname{Seq}_r(F)$ o espaço das sequências $(x_n)_{n\in\mathbb{N}}$ do *tipo recorrência*:

$$x_{n+k} = a_1 x_n + \cdots + a_k x_{n+k-1}, \quad k, n \in \mathbb{N} \ \text{ e } \ a_1, \ldots, a_k \in F.$$

[4] Max August Zorn, 1906-1993, matemático alemão.

Observe que se $x_1, \ldots, x_k \in F$ são dados, então a fórmula de recorrência determina unicamente todos os termos da sequência. Seja $\mathbf{e}_m = (e_n^{(m)})_{n\in\mathbb{N}}$ a sequência, com $m = 1, \ldots, k$, em que os primeiros k termos são:

$$e_n^{(m)} = \begin{cases} 1, & \text{se } m = n \\ 0, & \text{se } n \leq k \text{ e } m \neq k. \end{cases}$$

Então $\alpha = \{\mathbf{e}_1, \ldots, \mathbf{e}_k\}$ é uma base de $\operatorname{Seq}_r(F)$. Em particular, quando $k = 2$ e $a_1 = a_2 = 1$, temos o espaço das sequências do tipo Fibonacci $\operatorname{Seq}_f(F)$. Neste caso, uma base $\alpha = \{\mathbf{e}_1, \mathbf{e}_2\}$, em que

$$\mathbf{e}_1 = \mathbf{e} = (1, 0, 1, 1, 2, \ldots) \text{ e } \mathbf{e}_2 = \mathbf{f} = (0, 1, 1, 2, 3, \ldots).$$

De fato, consideremos a sequência

$$b_1\mathbf{e}_1 + \cdots + b_k\mathbf{e}_k = (b_1 e_n^{(1)} + \cdots + b_k e_n^{(k)})_{n\in\mathbb{N}} = (0).$$

Se $m \leq k$, então o m-ésimo termo dela é igual a b_m. Assim, $b_m = 0$ e α é LI. Para cada $(x_n)_{n\in\mathbb{N}}$ em $\operatorname{Seq}_r(F)$, definimos a sequência

$$(y_n)_{n\in\mathbb{N}} = (x_1 e_n^{(1)} + \cdots + x_k e_n^{(k)})_{n\in\mathbb{N}}.$$

Se $m \leq k$, então o m-ésimo termo dela é igual a x_m. Logo, as duas sequências são iguais. Portanto, $\operatorname{Seq}_r(F) = F[\alpha]$, de modo que $\operatorname{Seq}_r(F)$ é de dimensão finita.

Teorema 3.36 (Teorema da Existência de Base) *Sejam V um espaço vetorial sobre F e $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ um conjunto de vetores não nulos em V tal que $V = F[\alpha]$. Então, dentre esses vetores, podemos extrair uma base de V.*

Prova. Se α for LI, nada há para ser provado. Caso contrário, pelo Teorema 3.32, temos que um desses vetores é combinação linear dos outros, digamos

$$\mathbf{u}_n = x_1\mathbf{u}_1 + \cdots + x_{n-1}\mathbf{u}_{n-1}.$$

Assim, $V = [\alpha] = [\alpha - \{\mathbf{u}_n\}]$. Logo, aplicando sucessivamente o Teorema 3.32, (em no máximo $n-1$ etapas), obtemos uma base de V. ■

Exemplo 3.37 *Sejam $V = F^3$ e $\mathbf{u}_1 = (1, 0, 0), \mathbf{u}_2 = (1, 1, 0), \mathbf{u}_3 = (0, 0, 1)$ e $\mathbf{u}_4 = (1, 1, 1)$ vetores em V tais que $V = [\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4]$. Determine dentre esses vetores uma base de V.*

Solução. Pelo Teorema 3.36, basta verificar se os vetores $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ e $\mathbf{u}_4$ são LI ou LD ou, equivalentemente, escalonar a matriz

$$\begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Assim, a nulidade da matriz é igual a 1, de modo que os vetores são LD. Neste caso, é fácil verificar que $\mathbf{u}_4 = 0\mathbf{u}_1 + \mathbf{u}_2 + \mathbf{u}_3$. Portanto, o conjunto $\alpha = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ é uma base de V. ■

Teorema 3.38 (Lema da Mudança de Steinitz) *Seja V um espaço vetorial sobre F tal que $V = F[\mathbf{u}_1, \ldots, \mathbf{u}_m]$. Então qualquer conjunto com mais de m vetores em V é LD. Portanto, um conjunto de vetores LI em V possui no máximo m vetores.*

Prova. Como $V = [\mathbf{u}_1, \ldots, \mathbf{u}_m]$ temos, pelo Teorema 3.36, que existe uma base de V dentre esses vetores. Assim, renumerando, se necessário, podemos supor que $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$, com $k \leq m$, seja uma base de V. Seja $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ um conjunto qualquer de vetores em V, com $n > m$. Como $\mathbf{v}_j \in V$ e α é uma base de V temos que existem únicos $a_{ij} \in F$ tais que $\mathbf{v}_j = a_{1j}\mathbf{u}_1 + \cdots + a_{kj}\mathbf{u}_k$, para $j = 1, \ldots, n$. Vamos estudar a combinação linear

$$\begin{aligned} x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n &= \textstyle\sum_{j=1}^n x_j\mathbf{v}_j = \sum_{j=1}^n x_j \left(\sum_{i=1}^k a_{ij}\mathbf{u}_i\right) \\ &= \textstyle\sum_{i=1}^k \left(\sum_{j=1}^n x_j a_{ij}\right) \mathbf{u}_i. \end{aligned}$$

Sendo α um conjunto LI, obtemos

$$x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n = \mathbf{0} \Leftrightarrow x_1 a_{i1} + \cdots + x_n a_{in} = 0, i = 1, \ldots, k,$$

ou seja, basta discutir o sistema homogêneo com k equações e n incógnitas. Como $n > m \geq k = \rho(\mathbf{A})$, com $\mathbf{A} = (a_{ij}) \in F^{k\times n}$, temos, pelo Teorema 2.12, que esse sistema possui pelo menos uma solução não trivial $(y_1, \ldots, y_n)$. Logo,

$$\sum_{j=1}^n y_j\mathbf{v}_j = \sum_{i=1}^k \left(\sum_{j=1}^n y_j a_{ij}\right) \mathbf{u}_i = \sum_{i=1}^k 0\mathbf{u}_i = \mathbf{0}.$$

Portanto, β é LD. ■

Corolário 3.39 *Seja V um espaço vetorial de dimensão finita sobre F. Se $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ e $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ são duas bases quaisquer de V, então $m = n$.*

Prova. Como $V = F[\alpha]$ e β é um conjunto LI temos, pelo Teorema 3.38, que $n \leq m$. Por outro lado, como $V = F[\beta]$ e α é um conjunto LI temos, pelo Teorema 3.38, que $m \leq n$. Portanto, $m = n$. ■

Seja V um espaço vetorial de dimensão finita sobre F. A *dimensão* de V é o número de elementos em alguma base de V e denotamos por $\dim V$ ou $\dim_F V$. Note, pelo Corolário 3.39, que essa definição não depende da base de V, isto é, está bem definida. Quando $V = \{\mathbf{0}\}$, convencionamos que $\dim V = 0$.

Sejam V um espaço vetorial sobre F e $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ um subconjunto qualquer de vetores de V. O *posto* de α é definido como $\rho(\alpha) = \dim F[\alpha]$.

Lema 3.40 *Seja V um espaço vetorial sobre F. Seja $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ um subconjunto qualquer LI em V. Então $\mathbf{u} \in V - F[\alpha]$ se, e somente se, $\beta = \alpha \cup \{\mathbf{u}\}$ for um conjunto LI.*

Prova. Suponhamos que $\mathbf{u} \in V - F[\alpha]$ e existam $x_1, \ldots, x_m, y \in F$ tais que

$$x_1\mathbf{u}_1 + \cdots + x_m\mathbf{u}_m + y\mathbf{u} = \mathbf{0}.$$

Então $y = 0$, pois se $y \neq 0$, então a combinação linear é dependente de $\mathbf{u}$, de modo que $\mathbf{u} \in F[\alpha]$, o que é impossível. Assim, $y = 0$ e $x_1 = \cdots = x_m = 0$, pois α LI. Portanto, β é LI. A recíproca é clara. ■

Sejam V um espaço vetorial sobre F e $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ um conjunto qualquer de V. Uma *parte* de α significa que: existe uma função injetora $h : I_m \to I_n$, com $m \leq n$, tal que $\beta = \alpha \circ h = \{\mathbf{u}_{h(1)}, \ldots, \mathbf{u}_{h(m)}\} \subseteq \alpha$.

Teorema 3.41 (Teorema da Extensão) *Sejam V um espaço vetorial de dimensão finita sobre F e W um subespaço de V. Então qualquer conjunto de vetores LI em W é parte de uma base de W. Conclua que existe um subespaço U de V tal que $V = W \oplus U$.*

Prova. Seja $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ um conjunto qualquer de vetores LI em W. Se $W = F[\alpha]$, acabou. Caso contrário, existe, pelo Lema 3.40, um $\mathbf{u}_{m+1} \in W - F[\alpha]$ tal que $\alpha \cup \{\mathbf{u}_{m+1}\}$ é LI em W. Assim, aplicando sucessivamente o Lema 3.40 (em no máximo $\dim V - 1$ etapas), obtemos uma base $\{\mathbf{u}_1, \ldots, \mathbf{u}_m, \mathbf{u}_{m+1}, \ldots, \mathbf{u}_n\}$ de W. Finalmente, qualquer base α de W é parte de uma base β de V. Então existe um $U = F[\beta - \alpha]$

tal que $V = W \oplus U$. Por exemplo, como $\beta = \alpha \cup (\beta - \alpha)$ temos que $V = F[\beta] = F[\alpha] + F[\beta - \alpha] = W + U$. ■

Corolário 3.42 *Seja V um espaço vetorial de dimensão finita sobre F. Se W for um subespaço próprio de V, então $\dim W < \dim V$. Além disso, se $\dim V = n$, então qualquer conjunto com n vetores LI em V é uma base de V.*

Prova. Como $W \neq \{\mathbf{0}\}$ temos que existe um $\mathbf{u}$ em W, com $\mathbf{u} \neq \mathbf{0}$. É claro que $\{\mathbf{u}\}$ é LI em W. Assim, pelo Teorema 3.41, existe uma base de W contendo $\mathbf{u}$ e no máximo $\dim V$ elementos. Logo, $\dim W \leq \dim V$. Como $W \neq V$ temos que existe um $\mathbf{v} \in V$ tal que $\mathbf{v} \notin W$, de modo que acrescentando $\mathbf{v}$ a uma base de W, obtemos um conjunto LI de V. Portanto, $\dim W < \dim V$. ■

Exemplo 3.43 *Seja $V = F^3$. Verifique se os vetores $(1, 1, 0)$ e $(0, 1, 1)$ é parte de uma base de V.*

Solução. Pelo Teorema 3.41, basta escalonar a matriz

$$\left(\begin{array}{cc|c} 1 & 0 & a \\ 1 & 1 & b \\ 0 & 1 & c \end{array}\right) \to \cdots \to \left(\begin{array}{cc|c} 1 & 0 & a \\ 0 & 1 & b-a \\ 0 & 0 & a-b+c \end{array}\right).$$

Assim, os vetores $(1, 1, 0)$ e $(0, 1, 1)$ são LI, de modo que $\alpha = \{(1, 1, 0), (0, 1, 1)\}$ é parte de uma base de V. Logo, $\mathbf{u} = (a, b, c) \in V - F[\alpha]$ se, e somente se, $a - b + c \neq 0$. Em particular, $\mathbf{u} = (1, 1, 1) \in V$. Portanto, $\{(1, 1, 0), (0, 1, 1), (1, 1, 1)\}$ é uma base de V. ■

Teorema 3.44 *Seja V um espaço vetorial de dimensão finita sobre F. Se W_1 e W_2 são subespaços de V, então*

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

Prova. Como $W_1 \cap W_2$ é um subespaço de W_1 e W_2 temos, pelo Teorema 3.41, que $W_1 \cap W_2$ contém uma base $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ que é parte de uma base $\alpha \cup \beta$, com $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ de W_1 e parte de uma base $\alpha \cup \gamma$, com $\gamma = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ de W_2. Note que os conjuntos α, β e γ são dois a dois disjuntos, confira Figura 3.1.

Afirmação. O conjunto $\delta = \alpha \cup \beta \cup \gamma$ é uma base de $W_1 + W_2$.

De fato, é claro que o conjunto δ gera $W_1 + W_2$. Suponhamos que

$$(x_1\mathbf{u}_1 + \cdots + x_k\mathbf{u}_k) + (y_1\mathbf{v}_1 + \cdots + y_m\mathbf{v}_m) + (z_1\mathbf{w}_1 + \cdots + z_n\mathbf{w}_n) = \mathbf{0}$$

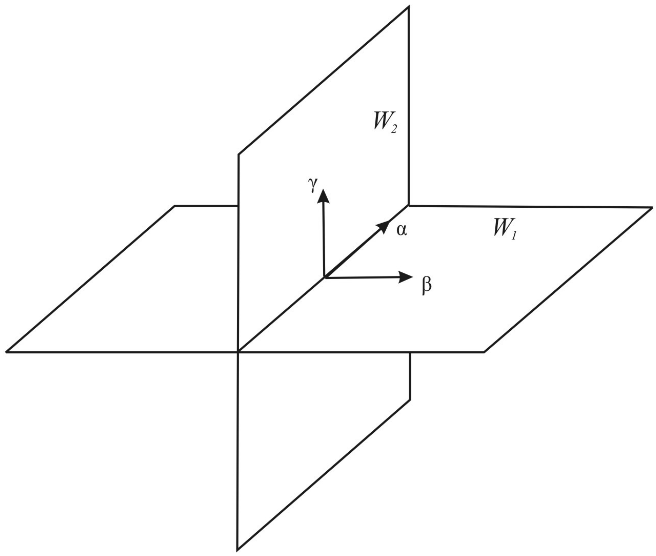


Figura 3.1: Representação gráfica de $W_1 \cap W_2$.

Então $-(z_1\mathbf{w}_1 + \cdots + z_n\mathbf{w}_n) \in W_1 \cap W_2$, pois

$$-(z_1\mathbf{w}_1 + \cdots + z_n\mathbf{w}_n) = (x_1\mathbf{u}_1 + \cdots + x_k\mathbf{u}_k) + (y_1\mathbf{v}_1 + \cdots + y_m\mathbf{v}_m) \in W_1.$$

Assim, existem $t_1, \ldots, t_k \in F$ tais que

$$(t_1\mathbf{u}_1 + \cdots + t_k\mathbf{u}_k) + (z_1\mathbf{w}_1 + \cdots + z_n\mathbf{w}_n) = \mathbf{0}.$$

Como γ é LI temos que $z_1 = \cdots = z_n = 0$. Logo,

$$(x_1\mathbf{u}_1 + \cdots + x_k\mathbf{u}_k) + (y_1\mathbf{v}_1 + \cdots + y_m\mathbf{v}_m) = \mathbf{0},$$

de modo que $x_1 = \cdots = x_k = y_1 = \cdots = y_m = 0$, pois β é LI. Portanto, δ é um conjunto LI. Consequentemente,

$$\begin{aligned} \dim W_1 + \dim W_2 &= (m+k) + (n+k) = (m+n+k) + k \\ &= \dim(W_1 + W_2) + \dim(W_1 \cap W_2), \end{aligned}$$

que é o resultado desejado. ■

Exemplo 3.45 *Sejam* $V = F^4$, $W_1 = \{(x, y, z, t) \in V : y + z + t = 0\}$ *e* $W_2 = \{(x, y, z, t) \in V : x + y = 0$ *e* $z - 2t = 0\}$ *subespaços de* V.

1. *Determine uma base de* $W_1 + W_2$ *e* $\dim(W_1 + W_2)$.
2. V *é soma direta de* W_1 *e* W_2?

Solução. (2) Já vimos que para determinar uma base de $W_1 \cap W_2$ basta resolver o sistema

$$\begin{cases} y+z+t & = & 0 \\ x+y & = & 0 \\ z-2t & = & 0 \end{cases}.$$

Assim, $W_1 \cap W_2 = F[(3, -3, 2, 1)]$. Portanto, V não é soma direta de W_1 e W_2. (1) Note que

$$\begin{aligned} W_1 &= \{(x, y, z, t) \in V : y + z + t = 0\} \\ &= \{(x, y, z, -y - z) \in V : x, y, z \in F\} \\ &= \{(x, 0, 0, 0) + (0, y, 0, -y) + (0, 0, z, -z) : x, y, z \in F\} \\ &= F[(1, 0, 0, 0), (0, 1, 0, -1), (0, 0, 1, -1)]. \end{aligned}$$

e $\dim W_1 = 3$. De modo análogo, $W_2 = F[(1, -1, 0, 0), (0, 0, 2, 1)]$ e $\dim W_2 = 2$. Neste caso,

$$\dim(W_1 + W_2) = \dim W_1 + W_2 - \dim(W_1 \cap W_2) = 3 + 2 - 1 = 4.$$

Portanto, $V = W_1 + W_2$, pois $W_1 + W_2 \subseteq V$. Para determinar uma base de $W_1 + W_2$, basta escalonar a matriz

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 2 & 1 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Assim, $\alpha = \{(1, 0, 0, 0), (0, 1, 0, -1), (0, 0, 1, -1), (1, -1, 0, 0)\}$ é uma base de $W_1 + W_2$. ■

Exemplo 3.46 *Sejam* $V = F^3$, $W_1 = F[(1, 0, -1), (0, 1, 2)]$ *e*
$W_2 = F[(1, 2, 3), (1, -1, 1)]$. *subespaços de* V.

1. *Determine uma base de* $W_1 \cap W_2$ *e a* $\dim(W_1 \cap W_2)$.

2. V *é soma direta de* W_1 *e* W_2?

Solução. (2) É fácil verificar que $\dim W_1 = 2$ e $\dim W_2 = 2$. Assim,

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2) = 4 - \dim(W_1 \cap W_2)$$

implica que $\dim(W_1 \cap W_2) \geq 1$, pois $\dim(W_1 + W_2) \leq 3 = \dim V$. Portanto, V não é soma direta de W_1 e W_2. (1) Para determinar uma base para $W_1 \cap W_2$, devemos primeiro determinar os vetores $\mathbf{u} = (x, y, z) \in F^3$ que estão nos subespaços W_1 e W_2, isto é, escalonar as matrizes

$$\left(\begin{array}{rr|c} 1 & 0 & x \\ 0 & 1 & y \\ -1 & 2 & z \end{array}\right) \to \cdots \to \left(\begin{array}{cc|c} 1 & 0 & x \\ 0 & 1 & y \\ 0 & 0 & x - 2y + z \end{array}\right)$$

e

$$\left(\begin{array}{rr|c} 1 & 1 & x \\ 2 & -1 & y \\ 3 & 1 & z \end{array}\right) \to \cdots \to \left(\begin{array}{cc|c} 1 & 0 & \frac{x+y}{3} \\ 0 & 1 & \frac{2x-y}{3} \\ 0 & 0 & \frac{-5x-2y+3z}{3} \end{array}\right).$$

Assim, pelo Teorema 2.14, $W_1 = \{(x, y, z) \in V : x - 2y + z = 0\}$ e $W_2 = \{(x, y, z) \in V : -5x - 2y + 3z = 0\}$. Logo, basta resolver o sistema

$$\left\{\begin{array}{rcl} x - 2y + z & = & 0 \\ -5x - 2y + 3z & = & 0 \end{array}\right..$$

Portanto, $W_1 \cap W_2 = F[(1, 2, 3)]$ e $\dim(W_1 \cap W_2) = 1$. ■

Exercícios

1. Responda verdadeiro (V) ou falso (F). Justifique.

   ( ) Todo conjunto que contém um subconjunto LD é LD?

   ( ) Todo subconjunto de um conjunto LI é LI?

   ( ) Todo conjunto que contém dois vetores iguais é LI?

   ( ) Todo conjunto que contém o vetor nulo é LI?

2. Sejam $V = F^3$ e W_1, W_2 subespaços de V tais que $\dim W_1 = \dim W_2 = 2$. É possível obtermos $W_1 \cap W_2 = \{(0,0,0)\}$?

3. Sejam $V = F^3$ e W_1, W_2 subespaços V tais que $\dim W_1 = 1$, $\dim W_2 = 2$ e $W_1 \not\subseteq W_2$. Mostre que $V = W_1 \oplus W_2$.

4. Sejam V um espaço vetorial sobre F e W_1, W_2 subespaços V, com $\dim W_1 = 4$, $\dim W_2 = 5$ e $\dim V = 7$. Determine os possíveis valores para a dimensão de $W_1 \cap W_2$.

5. Seja $V = F^4$. Determine uma base e a dimensão dos subespaços

$$\begin{aligned} W_1 &= F[(1,4,-1,3),(2,1,-3,-1),(0,2,1,-5)] \text{ e} \\ W_2 &= F[(1,-4,-2,1),(1,-3,-1,2),(3,-8,-2,7)]. \end{aligned}$$

6. Sejam $V = F^3$, $W_1 = \{(x,y,z) \in V : x = 0\}$ e $W_2 = F[(1,2,0),(3,1,2)]$ subespaços de V. Determine uma base e a dimensão para $W_1, W_2, W_1 + W_2$ e $W_1 \cap W_2$.

7. Sejam $V = F^{2\times 2}$,

$$W_1 = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in V : b = -a \right\} \text{ e } W_2 = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in V : c = -a \right\}$$

subespaços de V. Determine uma base e a dimensão para $W_1, W_2, W_1 + W_2$ e $W_1 \cap W_2$. É verdade que $F = W_1 \oplus W_2$?

8. Sejam $V = \mathbb{P}_3(F)$ e $W = \{p(x) \in V : p'(x) = 0\}$ um subespaço de V. Determine uma base e a dimensão de W.

9. Sejam $V = F^2$ e o conjunto de vetores $\alpha = \{\mathbf{u}, \mathbf{v}\}$ em V, em que $\mathbf{u} = (1-a, 1+a)$ e $\mathbf{v} = (1+a, 1-a)$. Determine o valor de $a \in F$ para que α não seja uma base de V.

10. Seja $V = \mathbb{P}_2(F)$. O conjunto $\alpha = \{p(x), p'(x), p''(x)\}$, onde $p(x) = 2x^2 - 3x + 1 \in V$, é uma base de V?

11. Mostre que o conjunto $\alpha = \{(1-x)^3, (1-x)^2, 1-x, 1\}$ é uma base de $V = \mathbb{P}_3(F)$.

12. Seja $V = F^4$. Quais dos subconjuntos abaixo são bases de V?

(a) $\{(1,1,0,0),(0,1,1,0),(0,0,1,1),(1,0,0,1)\}$.

(b) $\{(1,3,-2,4),(1,1,5,9),(2,0,-13,23),(1,5,1,-2)\}$.

(c) $\{(1,1,1,1),(3,2,0,3),(0,-1,0,3),(4,2,1,7)\}$.

(d) $\{(1,-2,0,1),(0,0,2,5),(-2,4,2,3),(-1,2,4,9)\}$.

13. Em cada um dos subconjuntos abaixo determine uma base de W e estenda-a a uma base de V.

    (a) $W = \{(x,y,z) \in F^3 : x - 3y + 3z = x + 5y - z = x + y + z = 0\}$.

    (b) Se $V =\in F^4$ e $W = F[(1,-2,0,1),(0,0,2,5),(-2,4,2,3)]$.

    (c) Se $V = F^4$ e $W = F[(1,1,1,1),(3,2,0,3),(0,-1,0,3)]$.

14. Seja W o conjunto de todos os quadrados mágicos de ordem 3 (confira Exercício (7) da Seção 2.2).

    (a) Mostre que W é um subespaço de $F^{3\times 3}$ e que o conjunto

    $$\alpha = \left\{ \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} \right\}$$

    é uma base de W.

    (b) Mostre que qualquer matriz

    $$\mathbf{A} = \begin{pmatrix} a_1 & a_2 & a_3 \\ * & * & * \\ * & * & * \end{pmatrix}$$

    pode ser transformada em um quadrado mágico. Existe outra maneira de fazê-la?

15. Sejam V um espaço vetorial sobre F e W_1, W_2, W_3 subespaços de V. Mostre que

$$\begin{aligned} \dim(W_1 + W_2 + W_3) \quad \leq \quad & \dim(W_1) + \dim(W_2) + \dim(W_3) \\ - \quad & \textstyle\sum_{1 \leq i < j \leq 3} \dim(W_i \cap W_j) \\ + \quad & \dim(W_1 \cap W_2 \cap W_3) \end{aligned}$$

16. Sejam V um espaço vetorial sobre $\mathbb{C}$ e $\mathbf{u} \in V$, com $\mathbf{u} \neq \mathbf{0}$. Mostre que os vetores $\mathbf{u}$ e $i\mathbf{u}$ são LI sobre $\mathbb{R}$. Conclua que se $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ for uma base de V,

então $\beta = \{\mathbf{u}_1, \ldots, \mathbf{u}_n, i\mathbf{u}_1, \ldots, i\mathbf{u}_n\}$ é uma base de V visto como um espaço vetorial sobre $\mathbb{R}$.

17. Sejam V um espaço vetorial de dimensão finita sobre $\mathbb{R}$ e $\tilde{V}$ sua complexificação. Mostre que $\dim \tilde{V} = \dim V$, mas $\dim \tilde{V} = 2 \dim V$ quando $\tilde{V}$ é visto como um espaço vetorial sobre $\mathbb{R}$.

18. Sejam V um espaço vetorial sobre F, com $V \neq \{\mathbf{0}\}$, e α um subconjunto não vazio de V. Mostre que as seguintes condições são equivalentes:

    (a) α é um conjunto *linearmente independente maximal* de V, no seguinte sentido: não existe um subconjunto LI β de V tal que $\alpha \subset \beta$;

    (b) α é um conjunto de *geradores minimal* de V, no seguinte sentido: não existe um subconjunto de geradores β de V tal que $\beta \subset \alpha$;

    (c) α é uma base de V.

19. Sejam V um espaço vetorial sobre F, W um subespaço de V e o *subespaço afim* ou a *classe lateral* $\mathbf{u} + W = \{\mathbf{u} + \mathbf{w} : \mathbf{w} \in W\}$, para todo $\mathbf{u} \in V$.

    (a) Mostre que $\mathbf{u} + W = \mathbf{v} + W$ se, e somente se, $\mathbf{v} - \mathbf{u} \in W$ e chama-se *congruência módulo* W, a qual é uma *relação de equivalência* sobre V.

    (b) Mostre que o conjunto $\overline{V} = V/W = \{\mathbf{u} + W : \mathbf{u} \in V\}$, munido com as operações de adição $(\mathbf{u} + W) \oplus (\mathbf{v} + W) = (\mathbf{u} + \mathbf{v}) + W$ e multiplicação por escalar $a \odot (\mathbf{u} + W) = a\mathbf{u} + W$ é um espaço vetorial sobre F, o qual chama-se *espaço quociente* de V módulo W. Quando W for o subespaço trivial identificamos $\overline{V}$ com V ou $\{\mathbf{0}\}$.

    (c) Se α for uma base de W e se β for um subconjunto de V tal que $\gamma = \{\mathbf{u} + W : \mathbf{u} \in \beta\}$ seja uma base de $\overline{V}$, então $\alpha \cap \beta = \emptyset$ e $\alpha \cup \beta$ é uma base de V.

    (d) Se β for uma base de V tal que $\alpha \subseteq \beta$ seja uma base de W, então $\gamma = \{\mathbf{u} + W : \mathbf{u} \in \beta - \alpha\}$ é uma base de $\overline{V}$.

    (e) Mostre que $\dim V = \dim \overline{V} + \dim W$ ou $\dim \overline{V} = \dim V - \dim W$.

    (f) Mostre que se W_1, W_2 são subespaços de V e $U = W_1 \cap W_2$, então $(W_1 + W_2)/U = W_1/U \oplus W_2/U$. Use isto para dar outra prova ao Teorema 3.44.

3.6 Mudança de Bases

Em toda esta seção, salvo menção explícita em contráro, identificamos uma lista de vetores com um conjunto de vetores.

Sejam V um espaço vetorial de dimensão finita sobre F e $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ uma base de V. Então a função $L_\alpha : F^n \to V$ definida como

$$L_\alpha(\mathbf{x}) = x_1\mathbf{u}_1 + \ldots + x_n\mathbf{u}_n$$

é bijetora. Assim, dado $\mathbf{u} \in V$, existe uma única lista $\mathbf{x} = (x_1, \ldots, x_n)$ em F^n tal que

$$\mathbf{u} = L_\alpha(\mathbf{x}) = x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n.$$

Esta decomposição chama-se *representação* de u em relação à base α e os escalares $x_1, \ldots, x_n$ chamam-se as *coordenadas* (ou *pesos*) de $\mathbf{u}$ em relação à base α e o vetor coluna

$$\mathbf{X} = [\mathbf{u}]_\alpha = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}^t \in F^{n\times 1}$$

chama-se *vetor coordenada* em relação à base α. Neste caso, obtemos um *sistema de coordenadas* $T_\alpha : V \to F^n$ definido como $T_\alpha(\mathbf{u}) = [\mathbf{u}]_\alpha$, quando $F^{n\times 1}$ é identificado com F^n, preserva as operações e $T_\alpha = L_\alpha^{-1}$, pois

$$[\mathbf{u}+\mathbf{v}]_\alpha = [\mathbf{u}]_\alpha + [\mathbf{v}]_\alpha \;\text{ e }\; [a\mathbf{u}]_\alpha = a[\mathbf{u}]_\alpha,\;\forall\; \mathbf{u},\mathbf{v} \in V \;\text{ e }\; a \in F.$$

Concluímos que $\beta = \{[\mathbf{u}_1]_\alpha, \ldots, [\mathbf{u}_n]_\alpha\}$ é uma base de F^n. Note que se $p \in \mathbb{N}$ e $\beta = (\mathbf{v}_1, \ldots, \mathbf{v}_p)$ for qualquer lista de V, então existem únicos $a_{ij} \in F$ tais que

$$\mathbf{v}_j = a_{1j}\mathbf{u}_1 + \cdot + a_{nj}\mathbf{u}_n = \sum_{i=1}^{n} a_{ij}\mathbf{u}_i,\;\; j = 1, \ldots, p. \tag{3.3}$$

Logo, $\mathbf{C}_j = [\mathbf{v}_j]_\beta \in F^{n\times 1}$, com $j = 1, \ldots, p$. Portanto, obtemos uma matriz $[I]_\alpha^\beta = (a_{ij}) \in F^{n\times p}$, a qual chama-se *representação* de β em relação à base α e $\dim F[\beta] = \rho([I]_\alpha^\beta)$. Em particular, se $n = p$ e β for uma base de V, então veremos, via a equação (3.3), que associado a matriz $[I]_\alpha^\beta$ existe uma única $I_{\alpha\beta} : V \to V$ tal que $I_{F^n} \circ L_\beta = I_{\alpha\beta} \circ L_\alpha$ ou $I_{\alpha\beta} = L_\beta \circ L_\alpha^{-1}$: isto significa que o diagrama (a) abaixo comuta. Observe que $\mathbf{v}_j = L_\beta(\mathbf{e}_j) = I_{\alpha\beta}(\mathbf{u}_j)$, de modo que $I_{\alpha\beta}$ é bijetora e preserva as operações, ou seja, um "isomorfismo linear".

$$\begin{array}{cc}
\begin{tikzcd} F^n \arrow[r, "L_\alpha"] \arrow[d, dashed, "I_{F^n}"'] & V \arrow[d, dashed, "I_{\alpha\beta}"] \\ F^n \arrow[r, "L_\beta"'] & V \end{tikzcd}
&
\begin{tikzcd} V \arrow[r, "T_\beta"] \arrow[d, dashed, "I_V"'] & F^{n\times 1} \arrow[d, dashed, "T_{\beta\alpha}"] \\ V \arrow[r, "T_\alpha"'] & F^{n\times 1} \end{tikzcd}
\\
(a) & (b)
\end{array}$$

Exemplo 3.47 *Sejam* $V = F^3$ *e* $\alpha = \{(1,0,-1),(1,1,1),(1,0,0)\}$ *uma base de* V. *Determine* $[(a,b,c)]_\alpha$.

Solução. Já vimos, para resolver esse problema, que basta escalonar a matriz

$$\left(\begin{array}{rrr|c} 1 & 1 & 1 & a \\ 0 & 1 & 0 & b \\ -1 & 1 & 0 & c \end{array}\right) \to \cdots \to \left(\begin{array}{ccc|c} 1 & 0 & 0 & b-c \\ 0 & 1 & 0 & b \\ 0 & 0 & 1 & a-2b+c \end{array}\right).$$

Portanto, $[(a,b,c)]_\alpha = \begin{pmatrix} b-c & b & a-2b+c \end{pmatrix}^t$. ■

Exemplo 3.48 *Seja* $V = \mathbb{P}_2(F)$. *Mostre que* $\alpha = \{1, 1+x, (1+x)^2\}$ *é uma base de* V *e determine* $[a+bx+cx^2]_\alpha$.

Solução. Como $(1+x)^2 = 1+2x+x^2$, basta escalonar a matriz

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & a \\ 0 & 1 & 2 & b \\ 0 & 0 & 1 & c \end{array}\right) \to \cdots \to \left(\begin{array}{ccc|c} 1 & 0 & 0 & a-b+c \\ 0 & 1 & 0 & b-2c \\ 0 & 0 & 1 & c \end{array}\right).$$

Portanto, α é LI e $\dim V = 3$ implica que α é uma base de V. Além disso, $[a+bx+cx^2]_\alpha = \begin{pmatrix} a-b+c & b-2c & c \end{pmatrix}^t$. ■

Sejam V um espaço vetorial sobre F e $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$, $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ bases de V. Então qualquer vetor $\mathbf{u} \in V$ pode ser escrito de modo único sob a forma

$$\left\{\begin{array}{rcl} \mathbf{u} = I_\alpha(\mathbf{x}) & = & x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n \\ \mathbf{u} = I_\beta(\mathbf{y}) & = & y_1\mathbf{v}_1 + \cdots + y_n\mathbf{v}_n. \end{array}\right. \tag{3.4}$$

Pergunta. Qual a relação entre $[\mathbf{u}]_\alpha$ e $[\mathbf{u}]_\beta$? Para responder essa pergunta consideremos a representação dos elementos de β em relação à base α. Assim, para cada $\mathbf{v}_i \in \beta \subset V$, temos que existem únicos escalares $a_{ji} \in F$ tais que

$$\mathbf{v}_i = \sum_{j=1}^{n} a_{ji}\mathbf{u}_j, \quad i = 1, \ldots, n. \tag{3.5}$$

Logo, pela segunda condição da equação (3.4), temos que

$$\mathbf{u} = \sum_{i=1}^{n} y_i \mathbf{v}_i = \sum_{i=1}^{n} y_i (\sum_{j=1}^{n} a_{ji}\mathbf{u}_j) = \sum_{j=1}^{n} (\sum_{i=1}^{n} a_{ji} y_i)\mathbf{u}_j.$$

Portanto, pela primeira condição da equação (3.4) e pela unicidade das coordenadas, temos que

$$x_j = \sum_{i=1}^{n} a_{ji} y_i, \quad j = 1, \ldots, n, \tag{3.6}$$

ou seja, um sistema de equações lineares com n equações e n incógnitas. Sob forma matricial

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

Portanto,

$$[\mathbf{u}]_\alpha = [I]_\alpha^\beta [\mathbf{u}]_\beta.$$

É conveniente representar a relação entre as coordenadas de $\mathbf{u}$ por uma tabela:

$V \backslash N$	y_1	$\cdots$	y_n
x_1	a_{11}	$\cdots$	a_{1n}
$\vdots$	$\vdots$	$\ldots$	$\vdots$
x_n	a_{n1}	$\cdots$	a_{nn}

Por razões técnicas, a matriz $[I]_\alpha^\beta$ chama-se a *matriz de mudança de base* ou *matriz de transição* da base "velha" (partida) α para a base "nova" β (chegada). Note, pela segunda condição da equação (3.4), que

$$[\mathbf{u}]_\alpha = [y_1\mathbf{v}_1 + \cdots + y_n\mathbf{v}_n]_\alpha = y_1[\mathbf{v}_1]_\alpha + \cdots + y_2[\mathbf{v}_n]_\alpha,$$

de modo que

$$[\mathbf{u}]_\alpha = \begin{pmatrix} [\mathbf{v}_1]_\alpha & \cdots & [\mathbf{v}_n]_\alpha \end{pmatrix} [\mathbf{u}]_\beta \Leftrightarrow [\mathbf{u}]_\alpha = [I]_\alpha^\beta [\mathbf{u}]_\beta.$$

Assim, $[I]_\alpha^\beta$ é a matriz obtida colocando as coordenadas em relação à base α do vetor $\mathbf{v}_i$ na i-ésima coluna e a matriz velha $[\mathbf{u}]_\alpha$ é obtida quando a matriz nova $[\mathbf{u}]_\beta$ é multiplicada à esquerda por $[I]_\alpha^\beta$. Portanto, se V for um espaço vetorial de dimensão finita como $F^n, F^{n\times n}$ ou $\mathbb{P}_n(F)$, então o problema de determinar $[I]_\alpha^\beta \in F^{n\times n}$ é equivalente a discutir a sequência finita de sistemas não homogêneos $\mathbf{AX} = \mathbf{B}_i$, com

$$\mathbf{A} = \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{pmatrix}, \;\; \mathbf{X} = (\mathbf{x}) \;\text{ e }\; \mathbf{B}_i = (\mathbf{v}_i), \;\; i = 1, \ldots, n.$$

Já vimos que isso é dado pelo esquema:

$$\left(\begin{array}{c|c} \alpha & \beta \end{array} \right) = \left(\begin{array}{ccc|ccc} \mathbf{u}_1 & \cdots & \mathbf{u}_n & \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{array} \right) \to \cdots \to \left(\begin{array}{c|c} \mathbf{I} & [I]_\alpha^\beta \end{array} \right).$$

Por exemplo, se $\mathbf{u}_1 = (-9, 1), \mathbf{u}_2 = (-5, -1)$ e $\mathbf{v}_1 = (1, -4), \mathbf{v}_2 = (3, -5)$, então

$$\left(\begin{array}{rr|rr} -9 & -5 & 1 & 3 \\ 1 & -1 & -4 & -5 \end{array} \right) \to \cdots \to \left(\begin{array}{rr|rr} 1 & 0 & -\frac{3}{2} & -2 \\ 0 & 1 & \frac{5}{2} & 3 \end{array} \right).$$

Logo, $[\mathbf{v}_1]_\alpha = \frac{1}{2}(-3, 5)^t$ e $[\mathbf{v}_2]_\alpha = (-2, 3)^t$. Observe que a matriz $[I]_\alpha^\beta$ é sempre invertível, pois

$$\mathbf{v}_i = \sum_{j=1}^n a_{ij}\mathbf{u}_j \;\text{ e }\; \mathbf{u}_j = \sum_{k=1}^n b_{jk}\mathbf{v}_k$$

implicam que

$$\mathbf{v}_i = \sum_{j=1}^n a_{ij}(\sum_{k=1}^n b_{jk}\mathbf{v}_k) = \sum_{k=1}^n(\sum_{j=1}^n a_{ij}b_{jk})\mathbf{v}_k,$$

de modo que

$$\sum_{j=1}^n a_{ij}b_{jk} = \delta_{ik}, i, k = 1, \ldots, n.$$

Portanto, $[I]_\alpha^\beta[I]_\beta^\alpha = (\delta_{ik}) = \mathbf{I}$. Veremos que $[I_V(\mathbf{u}_j]_\beta = [\mathbf{u}_j]_\beta$ implica que $[I_V]_\alpha^\beta = [I]_\beta^\alpha$ e os dois sistemas de coordenadas estão relacionados pelo diagrama comutativo (b). Note que $T_\alpha \circ I_V = T_{\beta\alpha} \circ T_\beta$ e $T_{\beta\alpha}$ é a função associada a matriz $[I]_\alpha^\beta$, usando a equação

(3.5), obtemos

$$(T_\alpha^{-1} \circ T_\beta)(\mathbf{e}_j) = T_\alpha^{-1}(\mathbf{v}_j) = \sum_{i=1}^{n} a_{ji} T_\alpha^{-1}(\mathbf{u}_i) = \sum_{i=1}^{n} a_{ji}\mathbf{e}_i. \tag{3.7}$$

É importante ressaltar que $[I]_\alpha^\beta = \mathbf{I}$ quando $\alpha = \beta$.

Exemplo 3.49 *Sejam $V = F^2$ e $\alpha = \{\mathbf{e}_1, \mathbf{e}_2\}, \beta = \{\mathbf{u}_1, \mathbf{u}_2\}$ bases de V, com $\mathbf{u}_1 = (2, -1)$ e $\mathbf{u}_2 = (3, 4)$. Determine $[(a, b)]_\beta$. Em particular, $[(5, -8)]_\beta$.*

Solução. Como $\mathbf{u}_1 = 2\mathbf{e}_1 - \mathbf{e}_2$ e $\mathbf{u}_2 = 3\mathbf{e}_1 + 4\mathbf{e}_2$ temos que $\mathbf{e}_1 = \frac{1}{11}(4\mathbf{u}_1 + \mathbf{u}_2)$ e $\mathbf{e}_2 = \frac{1}{11}(-3\mathbf{u}_1 + 2\mathbf{u}_2)$. Observe que esse processo é equivalente a escalonar a matriz

$$\left(\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ -1 & 4 & 0 & 1 \end{array}\right) \to \cdots \to \left(\begin{array}{cc|cc} 1 & 0 & \frac{4}{11} & -\frac{3}{11} \\ 0 & 1 & \frac{1}{11} & \frac{2}{11} \end{array}\right).$$

Pondo $\mathbf{u} = a\mathbf{e}_1 + b\mathbf{e}_2$, obtemos $[\mathbf{u}]_\beta = [a\mathbf{e}_1 + b\mathbf{e}_2]_\beta = a[\mathbf{e}_1]_\beta + b[\mathbf{e}_2]_\beta$ ou, equivalentemente, $[\mathbf{u}]_\beta = [I]_\beta^\alpha[\mathbf{u}]_\alpha$. Logo, para determinar $[\mathbf{u}]_\beta$ basta conhecer as colunas da matriz $[I]_\beta^\alpha$. Portanto, $[\mathbf{u}]_\beta = \frac{1}{11}(4 - 3b, a + 2b)^t$. Em particular, $[(5, -8)]_\beta = (4, -1)^t$. ■

Exemplo 3.50 *Sejam $V = \mathbb{R}^2, \alpha = \{\mathbf{e}_1, \mathbf{e}_2\}$ e $\beta = \{\mathbf{f}_1, \mathbf{f}_2\}$ uma base de V obtida de α pela rotação anti-horária de um ângulo θ através da origem. Determine as novas coordenadas $[\mathbf{u}]_\beta$ cujas coordenadas velhas são $[\mathbf{u}]_\alpha = (x, y)^t$.*

Solução. Pelo exposto a resposta é dada pela equação matricial $[\mathbf{u}]_\beta = [I]_\beta^\alpha[\mathbf{u}]_\alpha$, pois conhecemos $[\mathbf{u}]_\alpha$. Como $\mathbf{f}_1 = (\cos\theta, \operatorname{sen}\theta)$ e $\mathbf{f}_2 = (-\operatorname{sen}\theta, \cos\theta))$ temos que

$$\mathbf{e}_1 = \cos\theta\mathbf{f}_1 - \operatorname{sen}\theta\mathbf{f}_2 \ \text{ e } \ \mathbf{e}_2 = \operatorname{sen}\theta\mathbf{f}_1 + \cos\theta\mathbf{f}_2,$$

confira Figura 3.2. Assim,

$$[I]_\beta^\alpha = \left(\ [\mathbf{e}_1]_\beta \quad [\mathbf{e}_2]_\beta\ \right) = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix} \text{ e } [\mathbf{u}]_\beta = \begin{pmatrix} x\cos\theta + y\operatorname{sen}\theta \\ -x\operatorname{sen}\theta + y\cos\theta \end{pmatrix}.$$

Em particular, se $\mathbf{u} = (x, y)$ e $\theta = \frac{\pi}{4}$, então

$$[I]_\beta^\alpha = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \ \text{ e } \ [\mathbf{u}]_\beta = \frac{\sqrt{2}}{2}\begin{pmatrix} x + y \\ -x + y \end{pmatrix}.$$

Note que se $[\mathbf{u}]_\beta = (u, v)^t$, então a mudança de coordenadas ou uma rotação de eixos $[\mathbf{u}]_\alpha = ([I]_\beta^\alpha)^{-1}[\mathbf{u}]_\beta$ implica que $x = \frac{\sqrt{2}}{2}(u - v)$ e $y = \frac{\sqrt{2}}{2}(u + v)$. Por exemplo, se $x^2 - y^2 = 1$ for a equação de uma hipérbole no plano cartesiano "velho" xy, então $2uv + 1 = 0$ for a equação da mesma hipérbole no plano cartesiano "novo" uv. Portanto, isto mudou apenas a localização do gráfico da hipérbole, mas não sua natureza (forma).
■

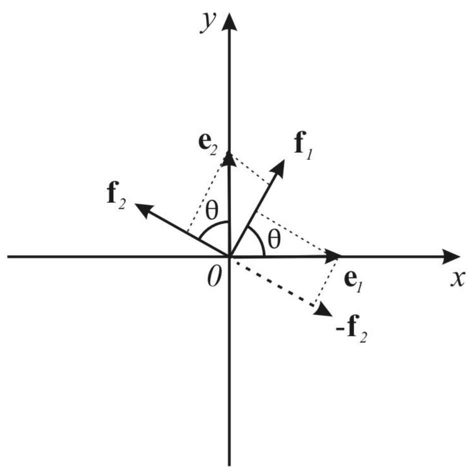


Figura 3.2: Representação gráfica da rotação.

Exemplo 3.51 *Sejam $V = F^2$ e $\alpha = \{(1, 2), (3, -6)\}, \beta$ bases de V. A matriz de mudança de base da base α para a base β é $[I]_\alpha^\beta = \mathbf{E}_{11} + \mathbf{E}_{12} + \mathbf{E}_{21} - \mathbf{E}_{22}$. Determine a base β.*

Solução. Seja $\beta = \{\mathbf{v}_1, \mathbf{v}_2\}$ a base desejada. Então, pelo exposto, $[\mathbf{u}]_\alpha = [I]_\alpha^\beta[\mathbf{u}]_\beta$ se, e somente se, $[\mathbf{u}]_\beta = ([I]_\alpha^\beta)^{-1}[\mathbf{u}]_\alpha$. Como $[(1, 2)]_\alpha = (1, 2)^t$ e $[(3, -6)]_\alpha = (3, -6)^t$ temos que determinar a inversa de $[I]_\alpha^\beta$, ou seja, escalonar a matriz

$$\left(\begin{array}{rr|rr} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{array}\right) \to \cdots \to \left(\begin{array}{rr|rr} 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} & -\frac{1}{2} \end{array}\right).$$

Assim, $[\mathbf{v}_1]_\beta = (2, -2)^t$ e $[\mathbf{v}_2]_\beta = (-1, 4)^t$. Portanto, $\beta = \{(2, -2), (-1, 4)\}$. ■

Teorema 3.52 *Sejam V um espaço vetorial sobre F e $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$, $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, $\gamma = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ bases de V. Então $[I]_\alpha^\gamma = [I]_\alpha^\beta[I]_\beta^\gamma$.*

Prova. De modo análogo a equação (3.5), obtemos

$$\mathbf{v}_i = \sum_{j=1}^{n} a_{ji}\mathbf{u}_j \ \text{ e } \ \mathbf{w}_k = \sum_{l=1}^{n} b_{lk}\mathbf{v}_l, \ \ i.k = 1, \ldots, n.$$

Assim,

$$\mathbf{w}_k = \sum_{l=1}^{n} b_{lk}(\sum_{j=1}^{n} a_{jl}\mathbf{u}_j) = (\sum_{l=1}^{n} a_{1l}b_{lk})\mathbf{u}_1 + \cdots + (\sum_{l=1}^{n} a_{nl}b_{lk})\mathbf{u}_n, \ \ k = 1, \ldots, n.$$

Logo,

$$[I]_\alpha^\gamma = \begin{pmatrix} [\mathbf{w}_1]_\alpha & \cdots & [\mathbf{w}_n]_\alpha \end{pmatrix} = \begin{pmatrix} \sum_{l=1}^{n} a_{1l}b_{l1} & \cdots & \sum_{l=1}^{n} a_{nl}b_{ln} \\ \vdots & \ddots & \vdots \\ \sum_{l=1}^{n} a_{nl}b_{l1} & \cdots & \sum_{l=1}^{n} a_{nl}b_{ln} \end{pmatrix}.$$

Portanto, $[I]_\alpha^\gamma = [I]_\alpha^\beta [I]_\beta^\gamma$. ■

Observe que se $\mathcal{B}$ for o conjunto das bases de V e $\mathrm{GL}_n(F)$ for o conjunto das matrizes não singulares de $F^{n\times n}$, então, cada base $\alpha \in \mathcal{B}$ fixada, induz uma função bijetora $f_\alpha : \mathcal{B} \to \mathrm{GL}(F)$ definida como $f_\alpha(\beta) = [I]_\alpha^\beta$, pois dado $\mathbf{P} = (a_{ij}) \in \mathrm{GL}(F)$, use a equação (3.5) para obter β. O conjunto $\mathrm{GL}_n(F)$ chama-se *grupo linear geral*.

Exercícios

1. Sejam $V = F^2$ e $\alpha = \{(2, 1), (1, -1)\}$ em V. Mostre que α é uma base de V e calcule $[(4, -1)]_\alpha$ e $[(x, y)]_\alpha$.

2. Seja $V = F^2$. Então calcule $[(6, 2)]_\alpha$ e $[(x, y)]_\alpha$, onde

   (a) $\alpha = \{(2, 1), (1, -1)\}$.
   (b) $\alpha = \{(2, 0), (0, -1)\}$.
   (c) $\alpha = \{(1, 0), (0, 1)\}$.
   (d) $\alpha = \{(2, 1), (1, 2)\}$.

3. Sejam $V = \mathbb{R}^2$ e $\mathbf{u} = (a, b), \mathbf{v} = (c, d) \in V$ tais que

$$ac + bd = 0 \ \text{ e } \ a^2 + b^2 = c^2 + d^2 = 1.$$

Mostre que $\beta = \{\mathbf{u}, \mathbf{v}\}$ é uma base de V. Além disso, calcule $[(x, y)]_\alpha$.

4. Sejam $V = \mathbb{P}_3(F)$ e $\beta = \{(1-x)^3, (1-x)^2, 1-x, 1\}$ uma base de V. Determine $[-x^2 - 2x + 3]_\beta$.

5. Determine a matriz de transição da base $\beta = \{(1, 1, 0), (0, 1, 0), (0, 0, 3)\}$ para a base canônica de F^3.

6. Sejam $V = F^3$ e $\beta = \{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}$ uma base de V. Determine $[(x, y, z)].\beta$.

7. Sejam $V = F^2$ e $\alpha = \{(1, 3), (2, -4)\}$ uma base de V. Se

$$[I]_\beta^\alpha = \begin{pmatrix} -7 & 6 \\ -11 & 8 \end{pmatrix},$$

então determine a base β.

8. Sejam $V = F^2$ e $\beta = \{(3, 5), (1, 2)\}$ uma base de V. Se

$$[I]_\beta^\alpha = \begin{pmatrix} -1 & 4 \\ 4 & -11 \end{pmatrix},$$

então determine a base α.

9. Seja $V = F^3$ e $\alpha = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}, \beta = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ bases de V, em que $\mathbf{v}_1 = \mathbf{u}_1 + \mathbf{u}_3$, $\mathbf{v}_2 = 2\mathbf{u}_1 + \mathbf{u}_2 + \mathbf{u}_3$ e $\mathbf{v}_3 = \mathbf{u}_1 + 2\mathbf{u}_2 + \mathbf{u}_3$. Determine as matrizes de transições de α paraβ e de β para α.

10. Considere os dados do exercício anterior. Se $[\mathbf{v}]_\alpha^t = (1, 2, 3)$, então determine $[\mathbf{v}]_\beta$.

11. A matriz de transição da base α de F^2 para a base $\beta = \{(1, 1), (0, 2)\}$ é

$$[\mathbf{I}]_\alpha^\beta = \begin{pmatrix} 1 & 0 \\ -\frac{2}{3} & \frac{1}{3} \end{pmatrix}.$$

Determine a base α.

12. Sejam $\alpha = \{\mathbf{e}_1, \mathbf{e}_2\}, \beta = \{-\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2\}$ e $\gamma = \{2\mathbf{e}_1, 2\mathbf{e}_2\}$ bases de F^2. Se $[\mathbf{u}]_\beta^t = (-1, 3)$, então determine $[\mathbf{u}]_\alpha$ e $[\mathbf{u}]_\gamma$.

13. Sejam $\alpha = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ e $\beta = \{\mathbf{e}_1 + 2\mathbf{e}_2, \mathbf{e}_1 + 3\mathbf{e}_2 + 2\mathbf{e}_3, \mathbf{e}_2 + 3\mathbf{e}_3\}$ bases de F^3. Determine $[I]_\beta^\alpha$, $[I]_\alpha^\beta$ e $[I]_\beta^\alpha \cdot [I]_\alpha^\beta$.

14. Sejam $\alpha = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}, \beta = \{-\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3\}$ e $\gamma = 2\alpha$ bases de F^3. Se $[\mathbf{u}]_\beta^t = (1, 3, 1)$, então determine $[\mathbf{u}]_\alpha$ e $[\mathbf{u}]_\gamma$.

15. Sejam V um espaço vetorial de dimensão n sobre F e α uma base de V. Determine $[I]_\alpha^\alpha$.

16. Seja $V = F[\mathbf{E}_{11}, \mathbf{E}_{12}, \mathbf{E}_{22}]$ um espaço vetorial sobre F.

    (a) $\alpha = \{\mathbf{E}_{11}, \mathbf{E}_{12}, \mathbf{E}_{22}\}$ e $\beta = \{\mathbf{E}_{11}, \mathbf{E}_{11} + \mathbf{E}_{12}, \mathbf{E}_{11} + \mathbf{E}_{12} + \mathbf{E}_{22}\}$ são bases de V?

    (b) Se sua resposta ao item (a) foi positiva, determine $[I]_\beta^\alpha$ e $[I]_\alpha^\beta$.

17. Sejam $\alpha = (\mathbf{u}_1, \ldots \mathbf{u}_n)$ uma base de $F^{1\times n}$ e $\mathbf{A} \in F^{n\times n}$. Mostre que as linhas da matriz $\alpha^t\mathbf{A}$ formam uma base de $F^{1\times n}$ se, e somente se, $\mathbf{A}$ for não singular.

18. Sejam V um espaço vetorial sobre F e $\{W_n : n \in \mathbb{N}\}$ uma sequência de subespaços de V. Mostre que $W = \cap_{n\in\mathbb{N}} W_n$ é um subespaço de V.

19. Sejam V um espaço vetorial sobre F e $\{W_n : n \in \mathbb{N}\}$ uma sequência de subespaços de V tais que
$$W_1 \subseteq W_2 \subseteq \cdots \subseteq W_n \subseteq \cdots .$$
    Mostre que $W = \cup_{n\in\mathbb{N}} W_n$ é um subespaço de V.

20. Seja V um espaço vetorial de dimensão finita sobre F. Suponhamos que $\alpha = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ uma lista de V, com $n \geq 3$, seja LD, mas quaisquer $n - 1$ desses vetores sejam LI.

    (a) Dê um exemplo de uma tal lista em F^3!

    (b) Mostre que existe uma lista $\mathbf{x} = (x_1, \ldots, x_n)$ em F^n, todos diferentes de zeros, tais que $x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n = \mathbf{0}$.

    (c) Suponhamos que $y_1\mathbf{u}_1 + \cdots + y_n\mathbf{u}_n = \mathbf{0}$. Mostre que existe um $a \in F^*$ tal que $y_i = ax_i$, com $i = 1, \ldots n$.

    (d) Mostre que
$$W = \{(y_1, \ldots, y_n) \in F^n : y_1\mathbf{u}_1 + \cdots + y_n\mathbf{u}_n = \mathbf{0}\}$$
    é um subespaço de F^n e determine a dimensão de W.

21. Sejam $V = F^{n\times n}$ e $\alpha = \{\mathbf{F}_{ij} : i, j = 1, \ldots, n\}$, em que $\mathbf{F}_{ij} = \mathbf{E}_{ij} + \mathbf{I}$. Mostre que α é uma base de V, com elementos não singulares.

22. Sejam $a, b \in F$ fixados, com $a \neq b$, e $m, n \in \mathbb{Z}_+$. Mostre que

$$F_f(x) = \left\{ \frac{f(x)}{(x-a)^m(x-b)^n} : f(x) = a_0 + \cdots + a_{m+n-1}x^{m+n-1} \in F[x] \right\}$$

é um subespaço de $\mathcal{F}(F, F)$ e $\dim F_f(x) = m + n$. Conclua que se $p_i(x) = (x-a)^i$ e $q_j(x) = (x-b)^j$, então a lista

$$\beta = (p_1(x), \ldots, p_m(x), q_1(x), \ldots, q_n(x))$$

é também uma base de $F_f(x)$.

4 Transformações

Neste capítulo vamos estudar um tipo especial de funções, as quais chamam-se de "transformações lineares" e que é um dos objetos fundamentais da álgebra linear. Em cálculo, por exemplo, costuma-se aproximar uma função diferenciável por uma transformação linear. Veremos, também, que resolver um sistema $\mathbf{AX} = \mathbf{B}$ de equações lineares é equivalente a encontrar todos os elementos $\mathbf{X} \in F^{n\times 1}$ tais que

$$T_{\mathbf{A}}(\mathbf{X}) = \mathbf{B},$$

em que $T_{\mathbf{A}} : F^{n\times 1} \to F^{m\times 1}$ definida como $T_{\mathbf{A}}(\mathbf{X}) = \mathbf{AX}$ é uma transformação linear.

4.1 Transformações Lineares

Sejam V um espaço vetorial sobre F e $\alpha = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ qualquer lista sobre V. Então, já vimos que, a função $L_\alpha : F^n \to V$ definida como

$$L_\alpha(\mathbf{x}) = x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n$$

satisfazia as seguintes condições:

$$L_\alpha(\mathbf{x}+\mathbf{y}) = L_\alpha(\mathbf{x}) + L_\alpha(\mathbf{y}) \ \text{ e } \ L_\alpha(a\mathbf{x}) = aL_\alpha(\mathbf{x}).$$

Isto motiva a seguinte definição.

Sejam V e W espaços vetoriais sobre F. Uma função $T : V \to W$ é uma *transformação linear* ou simplesmente *linear* se as seguintes condições são satisfeitas:

1. $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$, para todos $\mathbf{u}, \mathbf{v} \in V$ (aditividade).
2. $T(a\mathbf{u}) = aT(\mathbf{u})$, para todo $a \in F$ e $\mathbf{u} \in V$ (homogeneidade).

Observe, intuitivamente, que uma transformação linear é uma função que preserva as operações dos espaços vetoriais e $T(\mathbf{0}) = \mathbf{0}$, pois

$$T(\mathbf{0}) = T(0 \cdot \mathbf{u}) = 0 \cdot T(\mathbf{u}) = \mathbf{0}.$$

Note que $T : V \to W$ é linear se, e somente se,

$$T(a\mathbf{u} + b\mathbf{v}) = aT(\mathbf{u}) + bT(\mathbf{v}), \ \ \forall \ a, b \in F \ \text{ e } \ \mathbf{u}, \mathbf{v} \in V,$$

pois

$$T(a\mathbf{u} + b\mathbf{v}) = T(a\mathbf{u}) + T(b\mathbf{v}) = aT(\mathbf{u}) + bT(\mathbf{v}).$$

Conclua que T é linear se, e somente se, o gráfico $G_T = \{(\mathbf{u}, T(\mathbf{u})) : \mathbf{u} \in V\}$ de T é um subespaço de $V \times W$. Mais geralmente,

$$T(a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n) = a_1T(\mathbf{u}_1) + \cdots + a_nT(\mathbf{u}_n), \ \ \forall \ a_i \in F \ \text{ e } \ \mathbf{u}_i \in V.$$

Finalmente, quando $V = W$, diremos que T é um *operador linear* sobre V.

Exemplo 4.1 (Transformação Nula) *Sejam V e W espaços vetoriais sobre F. A função $0 : V \to W$ definida como $0(\mathbf{u}) = \mathbf{0}$, para todo $\mathbf{u} \in V$, é linear.*

Solução. Dados $\mathbf{u}, \mathbf{v} \in V$ e $a \in F$, obtemos

$$0(\mathbf{u}+\mathbf{v}) = \mathbf{0} = \mathbf{0} + \mathbf{0} = 0(\mathbf{u}) + 0(\mathbf{v}) \ \text{ e } \ 0(a\mathbf{u}) = \mathbf{0} = a0(\mathbf{u}).$$

Portanto, 0 é linear. ■

Exemplo 4.2 (Operador Identidade) *Seja V um espaço vetorial sobre F. A função $I = I_V : V \to V$ definida como $I_V(\mathbf{u}) = \mathbf{u}$, para todo $\mathbf{u} \in V$, é um operador linear.*

Solução. Dados $\mathbf{u}, \mathbf{v} \in V$ e $a \in F$, obtemos

$$I_V(\mathbf{u}+\mathbf{v}) = \mathbf{u} + \mathbf{v} = I_V(\mathbf{u}) + I_V(\mathbf{v}) \ \text{ e } \ I_V(a\mathbf{u}) = a\mathbf{u} = aI_V(\mathbf{u}).$$

Portanto, I_V é um operador linear. ■

Exemplo 4.3 *Qualquer transformação linear $T : F \to F$ é da forma ax, para algum $a \in F$ fixado.*

Solução. É claro que a função $T : F \to F$ definida como $T(x) = ax$, para todo $x \in F$, é linear. Reciprocamente, seja $T : F \to F$ qualquer transformação linear. Então

$$T(x) = T(1 \cdot x) = T(1)x, \ \forall \ x \in F.$$

Pondo $a = T(1) \in F$, obtemos $T(x) = ax$, para todo $x \in F$. É importante ressaltar que o gráfico da função $T : \mathbb{R} \to \mathbb{R}$ definida como $f(x) = ax + b$ é uma reta, mas f não é linear, a menos que $b = 0$. ■

Exemplo 4.4 *Sejam $V = F^{n\times 1}, W = F^{m\times 1}$ espaços vetoriais sobre F e $\mathbf{A} \in F^{m\times n}$ fixada. A função $T_{\mathbf{A}} : V \to W$ definida como $T_{\mathbf{A}}(\mathbf{X}) = \mathbf{AX}$, para todo $\mathbf{X} \in V$, é linear.*

Solução. Dados $\mathbf{X}, \mathbf{Y} \in V$ e $a \in F$, obtemos

$$T_{\mathbf{A}}(\mathbf{X} + \mathbf{Y}) = \mathbf{A}(\mathbf{X} + \mathbf{Y}) = \mathbf{AX} + \mathbf{AY} = T_{\mathbf{A}}(\mathbf{X}) + T_{\mathbf{A}}(\mathbf{Y})$$

e

$$T_{\mathbf{A}}(a\mathbf{X}) = \mathbf{A}(a\mathbf{X}) = a(\mathbf{AX}) = aT_{\mathbf{A}}(\mathbf{X}).$$

Portanto, $T_{\mathbf{A}}$ é linear. Note que $S_{\mathbf{A}} : F^n \to F^m$ definida como $S_{\mathbf{A}}(\mathbf{x}) = (\mathbf{Ax}^t)^t = \mathbf{xA}^t$, para todo $\mathbf{x} \in F^n$, é linear. ■

Exemplo 4.5 (Operador Diferencial) *Seja $V = \mathbb{P}_n(\mathbb{R})$ um espaço vetorial sobre $\mathbb{R}$. A função $D : V \to V$ definida como $D(p(x)) = p'(x)$, para todo $p(x) \in V$, é um operador linear.*

Solução. Dados $p(x), q(x) \in V$ e $a \in F$, obtemos

$$D(p(x) + aq(x)) = (p(x) + aq(x))' = p'(x) + aq'(x) = D(p(x)) + aD(q(x)).$$

Portanto, D é um operador linear. ■

Exemplo 4.6 (Operador Semelhança) *A função $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida como $T(x, y) = c(x, y)$, para todo $c \in \mathbb{R}$, é um operador linear. Quando $c > 0$, T chama-se* homotetia.

Solução. Fica como um exercício. ■

Exemplo 4.7 (Rotação por um ângulo θ) *Seja $V = \mathbb{R}^2$. Determine o operador linear $R_\theta : V \to V$, em que $R_\theta(\mathbf{u})$ é uma rotação anti-horário de um ângulo θ, $0 \leq \theta < 2\pi$, do vetor $\mathbf{u} \in V$.*

Solução. Sejam $\mathbf{u} = (x, y)$ e $R_\theta(x, y) = (u, v)$. Então, pela Figura 4.1, temos que $u = r\cos(\alpha + \theta), x = r\cos\alpha$ e $y = r\operatorname{sen}\alpha$. Assim, $u = x\cos\theta - y\operatorname{sen}\theta$. De modo análogo, $v = x\operatorname{sen}\theta + y\cos\theta$. Portanto,

$$R_\theta(x, y) = (x\cos\theta - y\operatorname{sen}\theta, x\operatorname{sen}\theta + y\cos\theta),$$

que é o resultado desejado. ■

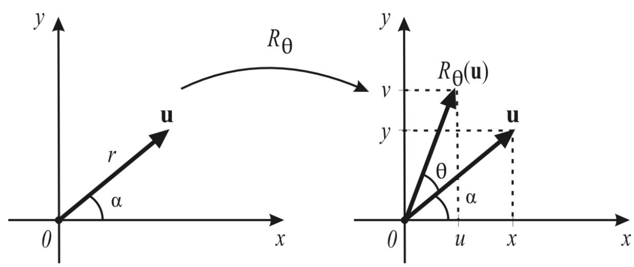


Figura 4.1: Representação gráfica da rotação.

Exemplo 4.8 (Operador Translação) *Sejam V um espaço vetorial sobre F e $\mathbf{v} \in V$ um vetor fixado. Então a função $T_\mathbf{v} : V \to V$ definida como $T(\mathbf{u}) = \mathbf{u} + \mathbf{v}$ não é um operador linear, a menos que $\mathbf{v} = \mathbf{0}$.*

Solução. Basta notar, em geral, que $T(\mathbf{0}) = \mathbf{v} \neq \mathbf{0}$. Confira Figura 4.2, quando $V = \mathbb{R}^2$ e $\mathbf{v} = (a, b)$. ■

Exemplo 4.9 *Seja $V = F^2$. A função $T : V \to V$ definida como $T(x, y) = (x, |y|)$ não é linear.*

Solução. Pondo $\mathbf{u} = (2, 1)$ e $\mathbf{v} = (3, -1)$, obtemos

$$T(\mathbf{u} + \mathbf{v}) = (5, 0) \neq (5, 2) = T(\mathbf{u}) + T(\mathbf{v}).$$

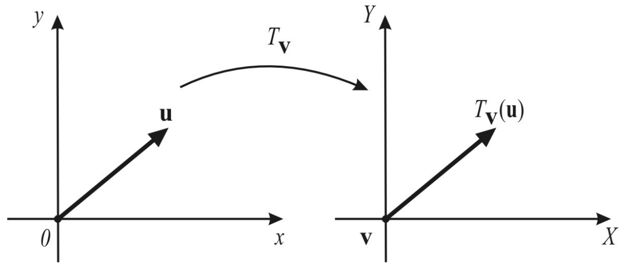


Figura 4.2: Representação gráfica da translação.

Note que $T(0,0) = (0,0)$. Portanto, $T(\mathbf{0}) = \mathbf{0}$ é condição necessária, mas não suficiente para que T seja uma transformação linear. ■

Seja $T : \mathbb{R} \to \mathbb{R}$ uma função. Diremos que T é uma *função aditiva* sobre $\mathbb{R}$ se ela satisfaz a seguinte condição (equação de Cauchy):

$$T(x+y) = T(x) + T(y), \ \ \forall \ \ x, y \in \mathbb{R},$$

Pelo Exemplo 4.3, para cada $a \in \mathbb{R}$ fixado, a função $T_a : \mathbb{R} \to \mathbb{R}$ definida como $T_a(x) = ax$ é aditiva. Note que $T : \mathbb{R}^2 \to \mathbb{R}$ definida como $T(x,y) = x + y\cos(\frac{y}{x})$ se $x \neq 0$ e $T(x,y) = 0$ se $x = 0$, satisfaz $T(a\mathbf{u}) = aT(\mathbf{u})$, mas não é aditiva.

Exemplo 4.10 *Seja $T : \mathbb{R} \to \mathbb{R}$ uma função. Mostre que T é linear sobre $\mathbb{Q}$ se, e somente se, T for aditiva.*

Solução. Se T for linear sobre $\mathbb{Q}$, então claramente T é aditiva. Reciprocamente, suponhamos que T seja aditiva. Então, para cada $x \in \mathbb{R}$, $T(2x) = T(x) + T(x) = 2T(x)$ e, indutivamente, $T(nx) = nT(x)$, para todo $n \in \mathbb{N}$. Como $T(0) = T(0+0) = T(0) + T(0)$ temos que $T(0) = 0$ e $T(-y) = -T(y)$, para todo $y \in \mathbb{R}$. Assim, $T(nx) = nT(x)$, para todo $n \in \mathbb{Z}$. Para cada $n \in \mathbb{N}$,

$$T(x) = T(n(n^{-1}x)) = T(n^{-1}x + \cdots + n^{-1}x) = nT(n^{-1}x)$$

implica que $T(n^{-1}x) = n^{-1}T(x)$. Portanto, $T(rx) = rT(x)$, para todo $r \in \mathbb{Q}$. Mais geralmente, se V, W são espaços vetoriais sobre $\mathbb{Q}$ e a função $T : V \to W$ for aditiva, então T é linear sobre $\mathbb{Q}$. Neste caso, podemos concluir que toda função definida em espaço vetorial sobre $\mathbb{Q}$, satisfazendo à condição aditiva, é sempre linear sobre $\mathbb{Q}$. Provaremos, no exemplo 4.15, que esse resultado não é, em geral, verdade. ■

Teorema 4.11 *Sejam V, W espaços vetoriais sobre F. Sejam $\alpha = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ uma base de V e $\beta = (\mathbf{w}_1, \dots, \mathbf{w}_n)$ uma lista qualquer em W. Então existe uma única transformação linear $T : V \to W$ tal que*

$$T(\mathbf{u}_i) = \mathbf{w}_i, \quad i = 1, \dots, n.$$

Prova. (Existência) Como α é uma base de V temos que cada vetor $\mathbf{u} \in V$ pode ser escrito de modo único sob a forma

$$\mathbf{u} = x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n.$$

Vamos definir $T : V \to W$ como

$$T(\mathbf{u}) = x_1\mathbf{w}_1 + \cdots + x_n\mathbf{w}_n = \sum_{i=1}^{n} x_i\mathbf{w}_i.$$

É claro que T está bem definida e $T(\mathbf{u}_i) = \mathbf{w}_i$, com $i = 1, \dots, n$, pois

$$\mathbf{u}_i = 0\mathbf{u}_1 + \cdots + 0\mathbf{u}_{i-1} + 1\mathbf{u}_i + 0\mathbf{u}_{i+1} + \cdots + 0\mathbf{u}_n, \quad i = 1, \dots, n.$$

Dados $\mathbf{v} \in V$, digamos $\mathbf{v} = y_1\mathbf{u}_1 + \cdots + y_n\mathbf{u}_n$, e $c \in \mathbb{R}$, temos que

$$T(\mathbf{u} + \mathbf{v}) = \sum_{i=1}^{n}(x_i + y_i)\mathbf{w}_i = \sum_{i=1}^{n} x_i\mathbf{w}_i + \sum_{i=1}^{n} y_i\mathbf{w}_i = T(\mathbf{u}) + T(\mathbf{v})$$

e

$$T(c\mathbf{u}) = \sum_{i=1}^{n}(cx_i)\mathbf{w}_i = c\sum_{i=1}^{n} x_i\mathbf{w}_i = cT(\mathbf{u}).$$

Portanto, T é linear.

(Unicidade) Seja $S : V \to W$ outra transformação linear tal que $S(\mathbf{u}_i) = \mathbf{w}_i$, com $i = 1, \dots, n$. Então

$$S(\mathbf{u}) = S(\sum_{i=1}^{n} x_i\mathbf{u}_i) = \sum_{i=1}^{n} x_iS(\mathbf{u}_i) = \sum_{i=1}^{n} x_i\mathbf{w}_i = T(\mathbf{u}),$$

para todo $\mathbf{u} \in V$. Portanto, $S = T$. ■

$$\begin{tikzcd} \alpha \arrow[r, "\lambda"] \arrow[dr, "f"'] & V \arrow[d, dashed, "T"] \\ & W \end{tikzcd}$$

O teorema 4.11 nos diz que uma transformação linear $T : V \to W$ goza da seguinte propriedade "universal": ela é unicamente determinada pela sua ação sobre os vetores de uma base de V. Explicitamente, pelos vetores $T(\mathbf{u}_i)$, com $i = 1 \ldots, n$, de alguma base $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ de V. Portanto, para determinar $T(\mathbf{u})$ basta encontrar $[\mathbf{u}]_\alpha$. Além disso, duas transformações lineares sobre V são **iguais** se elas coincidem em qualquer base de V. É muito importante, de um ponto de vista teórico e didático, termos uma noção de transformaçôes lineares sobre espaços vetoriais quaisquer. Para isto, sejam V, W espaços vetoriais sobre F e $\alpha = \{\mathbf{u}_i\}_{i \in I}$ uma base de V. Então para qualquer função $f : \alpha \to W$, existe uma única transformação linear $T : V \to W$ tal que $f = T \circ \lambda$, com $\lambda : \alpha \to V$, $\lambda(\mathbf{u}) = \mathbf{u}$, a *função inclusão*, ou seja, o diagrama acima comuta. Note que $T(\mathbf{u}_i) = f(\mathbf{u}_i) = \mathbf{w}_i$, para todo $i \in I$, e $\beta = \{\mathbf{w}_i\}_{i \in I}$ é uma família qualquer de vetores em W.

Exemplo 4.12 *Determine a transformação linear* $T : F^2 \to F^3$ *tal que* $T(1, 2) = (3, 2, 1)$ *e* $T(3, 4) = (6, 5, 4)$.

Solução. É fácil verificar que $\alpha = \{(1, 2), (3, 4)\}$ é uma base de F^2. Assim, pelo Teorema 4.11, existe uma única transformação linear $T : T^2 \to F^3$ tal que $T(1, 2) = (3, 2, 1)$ e $T(3, 4) = (6, 5, 4)$. Logo, para determinar $T(\mathbf{u})$ basta encontrar $[\mathbf{u}]_\alpha$ ou, equivalentemente, escalonar a matriz

$$\left(\begin{array}{cc|c} 1 & 3 & x \\ 2 & 4 & y \end{array}\right) \to \cdots \to \left(\begin{array}{cc|c} 1 & 0 & \frac{-4x+3y}{2} \\ 0 & 1 & \frac{2x-y}{2} \end{array}\right).$$

Portanto,

$$T(x, y) = \frac{-4x + 3y}{2}(3, 2, 1) + \frac{2x - y}{2}(6, 5, 4) = \left(\frac{3}{2}y, x + \frac{1}{2}y, 2x - \frac{1}{2}y\right),$$

que é o resultado desejado. ■

Exemplo 4.13 (Operador Projeção) *Determine a projeção de um vetor* $\mathbf{u} \in \mathbb{R}^2$ *sobre a reta* $y = ax$, *com* $a \in \mathbb{R}$.

Solução. É fácil verificar que $\alpha = \{(1, a), (-a, 1)\}$ e/ou $\alpha = \{(1, a), (1, b)\}$ é uma base de $\mathbb{R}^2$, para todos $a, b \in \mathbb{R}$, com $a \neq b$. Então, pelo Teorema 4.11, existe uma única transformação linear $P : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $P(1, a) = (1, a)$ e $P(-a, 1) = (0, 0)$. Logo, para determinar $T(\mathbf{u})$ basta encontrar $[\mathbf{u}]_\alpha$ ou, equivalentemente, escalonar a matriz

$$\left(\begin{array}{cc|c} 1 & -a & x \\ a & 1 & y \end{array}\right) \to \cdots \to \left(\begin{array}{cc|c} 1 & 0 & \frac{x+ay}{1+a^2} \\ 0 & 1 & \frac{-ax+y}{1+a^2} \end{array}\right).$$

Portanto, $P(\mathbf{u}) = \mathbf{u} - \mathbf{v}$ ou

$$P(x, y) = \left(\frac{x + ay}{1 + a^2}, \frac{ax + a^2 y}{1 + a^2}\right) = \frac{(x, y)(1, a)^t}{(1, a)(1, a)^t}(1, a).$$

Como $\mathbb{R}^2 = \mathbb{R}[(1, a)] \oplus \mathbb{R}[(-a, 1)]$, diremos que P é a *projeção* sobre $\mathbb{R}[(1, a)]$ na direção de (paralela a) $\mathbb{R}[(-a, 1)]$, onde $a \in \mathbb{R}$, confira Figura 4.3, ou seja, $P(x, y)$ é o ponto de interseção da reta $r : y = ax$ com reta que passa por (x, y) e paralela à reta (perpendicular a r) $y = -\frac{1}{a}x$, quando $a \neq 0$, e/ou $y = bx$. Além disso, é fácil verificar que $P^2 = P$. ■

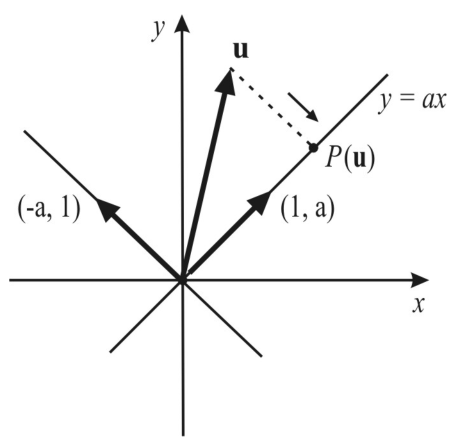


Figura 4.3: Representação gráfica da projeção.

Exemplo 4.14 (Operador Reflexão) *Determine a reflexão de um vetor* $\mathbf{u} \in \mathbb{R}^2$ *em torno de uma reta* $y = ax$*, com* $a \in \mathbb{R}$*.*

Solução. É fácil verificar que $\alpha = \{(1, a), (-a, 1)\}$ é uma base de $\mathbb{R}^2$, para todo $a \in \mathbb{R}$.

Então, pelo Teorema 4.11, existe uma única transformação linear $R : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $R(1, a) = (1, a)$ e $R(-a, 1) = (a, -1)$. Logo, para determinar $T(\mathbf{u})$ basta encontrar $[\mathbf{u}]_\alpha$ ou, equivalentemente, escalonar a matriz

$$\left(\begin{array}{cc|c} 1 & -a & x \\ a & 1 & y \end{array}\right) \to \cdots \to \left(\begin{array}{cc|c} 1 & 0 & \frac{x+ay}{1+a^2} \\ 0 & 1 & \frac{-ax+y}{1+a^2} \end{array}\right).$$

Portanto, $\mathbf{u} - R(\mathbf{u}) = 2(\mathbf{u} - P(\mathbf{u}))$ ou

$$R(x, y) = \left(\frac{(1-a^2)x + 2ay}{1+a^2}, \frac{2ax - (1-a^2)y}{1+a^2}\right).$$

Como $\mathbb{R}^2 = \mathbb{R}[(1, a)] \oplus \mathbb{R}[(-a, 1)]$,, diremos que R é a *reflexão* em $\mathbb{R}[(1, a)]$ na direção de (paralela a) $\mathbb{R}[(-a, 1)]$, onde $a \in \mathbb{R}$, confira Figura 4.4. Além disso, é fácil verificar que $R^2 = I$. ■

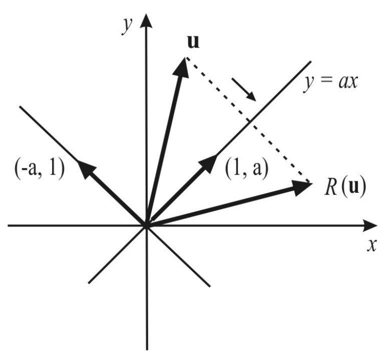


Figura 4.4: Representação gráfica da refleção.

Observe que se $\mathbf{d} = (1, a)$ for o vetor diretor da reta $y = ax$ e $T_\mathbf{d} : \mathbb{R}^2 \to \mathbb{R}^2$ for o operador translação, então $R_\mathbf{d} = R \circ T_\mathbf{d}$ chama-se *reflexão de deslizamento* (glide), ou seja, $R_\mathbf{d}$ é a composição de uma translação na direção de uma reta por uma distância d seguida por uma reflexão em torno dessa reta. Além disso, se θ for o ângulo que a reta $y = ax$ faz com o eixo dos x, então $a = \tan\theta$ e é fácil verificar que a reflexão é definida como

$$R(x, y) = (x \cos 2\theta + y \operatorname{sen} 2\theta, x \operatorname{sen} 2\theta - y \cos 2\theta).$$

Em particular, quando $\theta = \frac{\pi}{4}$, temos que $R(x, y) = (y, x)$.

Exemplo 4.15 *Mostre que existe uma função aditiva $T : \mathbb{R} \to \mathbb{R}$ tal que $T(x) \neq ax$, para todo $a \in \mathbb{R}$.*

Solução. É fácil verificar que $\mathbb{R}$ com as operações usuais é um espaço vetorial sobre $\mathbb{Q}$. Seja α uma base de “Hamel[1]” de $\mathbb{R}$ sobre $\mathbb{Q}$ e escolhendo $x_0 \in \alpha$. Então vamos definir $T : \mathbb{R} \to \mathbb{R}$ como

$$T(x) = \begin{cases} r_i, & \text{se } \ x = r_1x_1 + \cdots + r_ix_i + \cdots + r_nx_n \ \text{ e } \ x_i = x_0 \\ 0, & \text{caso contrário.} \end{cases}$$

Então T é aditiva. Por exemplo, se x_0 ocorre na representação básica de x e y, então $T(x+y) = r_i + s_j = T(x) + T(y)$. Em particular, $T(x_0) = 1$ e $T(x_i) = 0$, para todo $x_i \in \alpha$, com $i \neq 0$. Se $T(x) = ax$, para algum $a \in \mathbb{R}$, então $1 = T(x_0) = ax_0$, de modo que $a \neq 0$. Por outro lado, $0 = T(x_i) = ax_i$ implica que $a = 0$, pois $0 \notin \alpha$, o que é impossível. Portanto, $T(x) \neq ax$, para todo $a \in \mathbb{R}$. ■

Exercícios

1. Verifique quais das funções abaixo são lineares.

   (a) $T : F^2 \to F^2$, $T(x, y) = (2x - y, 0)$.

   (b) $T : F^3 \to F^2$, $T(x, y, z) = (x - 1, y + z)$.

   (c) $T : F \to F^3$, $T(x) = (x, 2x, -x)$.

   (d) $T : F^2 \to F^2$, $T(x, y) = (y, x^3)$.

   (e) $T : F^2 \to F^2$, $T(x, y) = (ax + by, cx + dy)$, onde $a, b, c, d \in F$.

2. Sejam $V = F^{n \times n}$ e $\mathbf{B}$ é uma matriz não nula fixada em V, quais das seguintes funções são lineares?

   (a) $T(\mathbf{A}) = \mathbf{BA}$.

   (b) $T(\mathbf{A}) = \mathbf{BA} - \mathbf{AB}$.

   (c) $T(\mathbf{A}) = \mathbf{B} + \mathbf{A}$.

   (d) $T(\mathbf{A}) = \mathbf{A}^t$.

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

[1] Georg Karl Wilhelm Hamel, 1877-1954, matemático alemão.

3. Sejam $V = \mathcal{F}(\mathbb{R}, \mathbb{R})$ e $h \in \mathbb{R}$ fixado. Mostre que cada uma das funções $T : V \to V$ abaixo é linear:

   (a) $(Tf)(x) = f(x + h)$. (Deslocamento)

   (b) $(Tf)(x) = f(x + h) - f(x)$. (Diferença para frente)

   (c) $(Tf)(x) = f(x) - f(x - h)$. (Diferença para trás)

   (d) $(Tf)(x) = f(x + \frac{h}{2}) - f(x - \frac{h}{2})$. (Diferença central)

   (e) $(Tf)(x) = \frac{1}{2}\left(f(x + \frac{h}{2}) - f(x - \frac{h}{2})\right)$. (Valor médio)

4. (**Operador Integração**) Seja $V = \mathcal{C}(\mathbb{R}, \mathbb{R})$ o espaço das funções reais contínuas. Mostre que a função $J : V \to V$ definida como

$$(Jf)(x) = \int_0^x f(t)dt$$

   é linear.

5. (**Operador Cisalhamento (shear) na direção de** x) Determine a função linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ que satisfaça $T(1, 0) = (1, 0)$ e $T(0, 1) = (a, 1)$, onde $a \in \mathbb{R}^*$. Defina **Operador Cisalhamento na direção de** y.

6. Determine o operador linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ que satisfaça $T(1, 2) = (1, 1)$ e $T(0, 1) = (1, 0)$.

7. Determine o operador linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ que satisfaça $T(1, 0) = (a, b)$ e $T(0, 1) = (c, d)$.

8. Seja $V = F[x]$ Mostre que cada uma das funções $T : V \to V$ abaixo é linear:

   (a) $T(p(x)) = xp(x)$ (Multiplicação por x).

   (b) $T(p(x)) = \frac{p(x)-a_0}{x}$ (Eliminação do termo constante e divisão por x).

9. Sejam $S : V \to W$ e $T : V \to W$ transformações lineares. Mostre que $S + T$ e aT, para todo $a \in \mathbb{R}$, são lineares. Conclua que o conjunto de todas as transformações lineares $\mathcal{L}(V, W)$ é um espaço vetorial sobre F.

10. Se $\dim V = 2$ e $\dim W = 3$, determine uma base de $L(V, W)$. Generalize.

11. Sejam $R : U \to V$, $S : U \to V$ e $T : V \to W$ transformações lineares. Mostre que $T \circ S$ é linear e $T \circ (R + S) = T \circ R + T \circ S$.

12. Sejam $V = \mathbb{R}^2$ e $R, S, T : V \to V$ operadores lineares definidos como $R(x, y) = (x, 0)$, $S(x, y) = (y, x)$ e $T(x, y) = (0, y)$. Determine:

    (a) $S + T$ e $3S - 5T$.

    (b) $R \circ S$, $S \circ R$, $R \circ T$, $T \circ R$, $S \circ T$ e $T \circ S$.

    (c) R^2, S^2 e T^2.

    (d) Mostre que S e T são LI.

13. Sejam $V = \mathbb{R}[x]$ e $D, M : V \to V$ operadores lineares definidos como $D(p(x)) = p'(x)$ e $M(p(x)) = xp(x)$. Mostre que $DM - MD = I$ e $(DM)^2 = D^2M^2 - DM$.

14. Sejam V, W espaços vetoriais sobre F e $f : V \to W$ uma função. Mostre que as seguintes condições são equivalentes:

    (a) Se $\mathbf{w} - \mathbf{u} = c(\mathbf{v} - \mathbf{w})$, então $f(\mathbf{w}) - f(\mathbf{u}) = c(f(\mathbf{v}) - f(\mathbf{w}))$, para todos $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ e $c \in F$;

    (b) $f(\mathbf{z}) = T(\mathbf{z}) + \mathbf{x}$, para todo $\mathbf{z} \in V$, onde $\mathbf{x} \in W$ e $T : V \to W$ é uma transformação linear;

    (c) $f(\sum_{i=1}^{n} c_i\mathbf{u}_i) = \sum_{i=1}^{n} c_i f(\mathbf{u}_i)$, para todo $\mathbf{u}_i \in V$ e $c_i \in F$, com $i = 1, \ldots, n$ e $c_1 + \cdots + c_n = 1$.

    Uma função que satisfaz uma das (todas) condições chama-se *função afim*.

15. Seja $T : V \to V$ um operador linear tal que $T^k = T^{k-1} \circ T = 0$, para algum $k \in \mathbb{N}$.

    (a) Mostre que se existir um $\mathbf{u} \in V$ tal que $T^{k-1}(\mathbf{u}) \neq \mathbf{0}$, então o conjunto $\alpha = \{\mathbf{u}, T(\mathbf{u}), \ldots, T^{k-1}(\mathbf{u})\}$ é LI.

    (b) Mostre que se $W = F[\mathbf{u}, T(\mathbf{u}), \ldots, T^{k-1}(\mathbf{u})]$, então $T(\mathbf{v}) \in W$, para todo $\mathbf{v} \in W$.

4.2 Núcleo e Imagem

Sejam V e W espaços vetoriais sobre F, com o objetivo de simplificar as notações, usaremos $\mathcal{L}(V, W)$ para representar o conjunto de todas as transformações lineares de V em W. Em particular, $\mathcal{L}(V)$ quando $V = W$. Já vimos que $\mathcal{L}(V, W)$, munido com a soma e multiplicação por escalar usuais de funções é um espaço vetorial sobre F. Em particular, $\mathcal{L}(V)$ mais a composição usual de função é um álgebra linear sobre F.

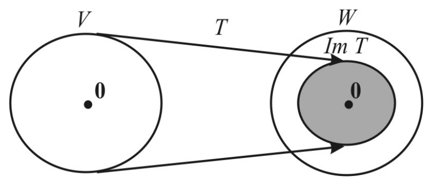


Figura 4.5: Representação gráfica da imagem.

Seja $T \in \mathcal{L}(V, W)$. A *imagem* de T é o conjunto

$$\begin{aligned} \operatorname{Im} T &= \{\mathbf{w} \in W : \mathbf{w} = T(\mathbf{u}), \text{ para algum } \mathbf{u} \in V\} \\ &= \{T(\mathbf{u}) : \mathbf{u} \in V\} = T(V), \end{aligned}$$

confira Figura 4.5. Observe que $\mathbf{w} \in \operatorname{Im} T$ significa que existe um $\mathbf{u} \in V$ tal que $\mathbf{w} = T(\mathbf{u})$. Portanto, se $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ for uma base de V, então, pelo Teorema 4.11, $\operatorname{Im} T = F[T(\alpha)] = F[T(\mathbf{u}_1), \ldots, T(\mathbf{u}_n)]$ e *posto* de T é definido como $\rho(T) = \dim \operatorname{Im} T$.

O *núcleo* (kernel) de T é o conjunto

$$\ker T = \{\mathbf{u} \in V : T(\mathbf{u}) = \mathbf{0}\} = T^{-1}(\mathbf{0}),$$

confira Figura 4.6. Observe que $\mathbf{u} \in \ker T$ significa que $T(\mathbf{u}) = \mathbf{0}$ e a *nulidade* ou *defeito* de T é definida como $\nu(T) = \dim \ker T$.

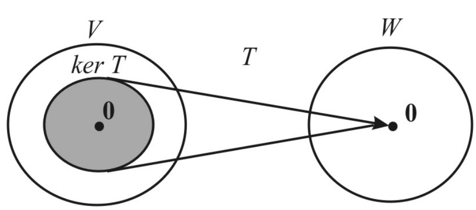


Figura 4.6: Representação gráfica do núcleo.

Teorema 4.16 *Seja $T \in \mathcal{L}(V, W)$. Então $\operatorname{Im} T$ é um subespaço de W e $\ker T$ é um subespaço de V.*

Prova. Vamos provar apenas que $\operatorname{Im} T$ é um subespaço de W. Note que $\operatorname{Im} T \neq \emptyset$, pois $\mathbf{0} = T(\mathbf{0}) \in \operatorname{Im} T$. Dados $\mathbf{w}_1, \mathbf{w}_2 \in \operatorname{Im} T$ e $a \in F$, existem $\mathbf{u}_1, \mathbf{u}_2 \in V$ tais que $\mathbf{w}_1 = T(\mathbf{u}_1)$ e $\mathbf{w}_2 = T(\mathbf{u}_2)$. Assim,

$$\mathbf{w}_1 + a\mathbf{w}_2 = T(\mathbf{u}_1) + aT(\mathbf{u}_2) = T(\mathbf{u}_1 + a\mathbf{u}_2) \in \operatorname{Im} T,$$

pois $\mathbf{u}_1 + a\mathbf{u}_2 \in V$. Portanto, $\operatorname{Im} T$ é um subespaço de W. ■

Observe que se $\mathbf{A} \in F^{m\times n}$ for fixada, então a função $T_{\mathbf{A}} : F^{n\times 1} \to F^{m\times 1}$ definida como $T_{\mathbf{A}}(\mathbf{X}) = \mathbf{AX}$ é, pelo Exemplo 4.4, uma transformação linear. Neste caso, $\mathbf{B} \in \operatorname{Im} T_{\mathbf{A}} = \{\mathbf{AX} : \mathbf{X} \in F^{n\times 1}\}$ significa que o sistema não homogêneo $\mathbf{AX} = \mathbf{B}$ possui pelo menos uma solução. Por outro lado, $\mathbf{X} \in \ker T_{\mathbf{A}} = \{\mathbf{X} : \mathbf{AX} = \mathbf{0}\}$ significa que o sistema homogêneo $\mathbf{AX} = \mathbf{O}$ possui pelo menos uma solução. Portanto,

$$\rho(\mathbf{A}) = \rho(T_{\mathbf{A}}) = \dim \operatorname{Im} T_{\mathbf{A}} \text{ e } \nu(\mathbf{A}) = \nu(T_{\mathbf{A}}) = \dim \ker T_{\mathbf{A}}.$$

Exemplo 4.17 *Seja $T \in: \mathbb{R}^3 \to \mathbb{R}^3$ definida como $T(x,y,z) = (x, 2y, 0)$. Mostre que T é linear e determine o núcleo e a imagem de T.*

Solução. É fácil verificar que T é linear. Assim,

$$\begin{aligned} \ker T &= \{(x,y,z) \in \mathbb{R}^3 : T(x,y,z) = (0,0,0)\} \\ &= \{(x,y,z) \in \mathbb{R}^3 : (x,2y,0) = (0,0,0)\} \\ &= \{(0,0,z) : z \in \mathbb{R}\} = \mathbb{R}[(0,0,1)] \end{aligned}$$

e

$$\begin{aligned} \operatorname{Im} T &= \{T(x,y,z) : (x,y,z) \in \mathbb{R}^3\} \\ &= \{(x,2y,0) : x,y \in \mathbb{R}\} = \mathbb{R}[(1,0,0),(0,2,0)]. \end{aligned}$$

Como $T(1,0,0) = (1,0,0), T(0,1,0) = (0,2,0)$ e $T(0,0,1) = (0,0,0)$ temos que $\operatorname{Im} T = \mathbb{R}[T(1,0,0), T(0,1,0)]$, confira Figura 4.7. ■

Exemplo 4.18 *Determine $T \in \mathcal{L}(F^3, F^4)$ tal que*

$$\operatorname{Im} T = F[(1,0,0,-1),(0,1,1,0)].$$

Solução. É fácil verificar que $\alpha = \{(1,0,0,-1),(0,1,1,0)\}$ é uma base de $\operatorname{Im} T$. Como $(1,0,0,-1),(0,1,1,0) \in \operatorname{Im} T$ temos que existem $\mathbf{u}_1, \mathbf{u}_2 \in F^3$ tais que

$$T(\mathbf{u}_1) = (1,0,0,-1) \text{ e } T(\mathbf{u}_2) = (0,1,1,0).$$

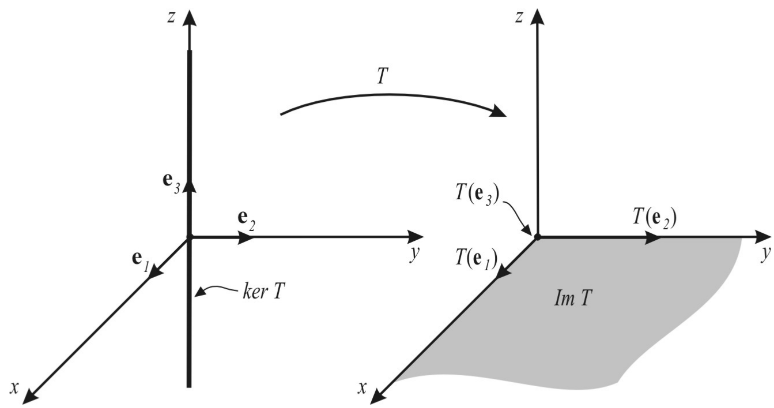


Figura 4.7: Representação gráfica da imagem e do núcleo.

Seja $W = F[\mathbf{u}_1, \mathbf{u}_2]$. Então $\{\mathbf{u}_1, \mathbf{u}_2\}$ é uma base de W, pois α é uma base de $\operatorname{Im} T$.

Afirmação. $F^3 = W \oplus \ker T$ e W chama-se *complemento direto* de $\ker T$.
De fato. Dado $\mathbf{u} \in F^3$, temos que $T(\mathbf{u}) \in \operatorname{Im} T$. Assim, existem $y_1, y_2 \in F$ tais que

$$T(\mathbf{u}) = y_1(1,0,0,-1) + y_2(0,1,1,0) = y_1 T(\mathbf{u}_1) + y_2 T(\mathbf{u}_2) = T(y_1\mathbf{u}_1 + y_2\mathbf{u}_2).$$

Logo, $\mathbf{u} - (y_1\mathbf{u}_1 + y_2\mathbf{u}_2) \in \ker T$, pois

$$T\left(\mathbf{u} - (y_1\mathbf{u}_1 + y_2\mathbf{u}_2)\right) = T(\mathbf{u}) - T(y_1\mathbf{u}_1 + y_2\mathbf{u}_2) = T(\mathbf{u}) - T(\mathbf{u}) = \mathbf{0},$$

Portanto, existe um $\mathbf{v} \in \ker T$ tal que

$$\mathbf{u} - (y_1\mathbf{u}_1 + y_2\mathbf{u}_2) = \mathbf{v} \Rightarrow \mathbf{u} = (y_1\mathbf{u}_1 + y_2\mathbf{u}_2) + \mathbf{v} \in W + \ker T,$$

ou seja, $F^3 = W + \ker T$. É fácil verificar que $W \cap \ker T = \{\mathbf{0}\}$. Escolhendo uma base $\{\mathbf{u}_3\}$ para $\ker T$, obtemos uma base $\beta = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ de F^3. Em particular, pondo $\mathbf{u}_1 = (1,0,0), \mathbf{u}_2 = (0,1,0)$ e $\mathbf{u}_3 = (0,0,1)$ temos, pelo Teorema 4.11, que existe uma única transformação linear $T : F^3 \to F^4$ tal que

$$T(\mathbf{u}_1) = (1,0,0,-1), T(\mathbf{u}_2) = (0,1,1,0) \text{ e } T(\mathbf{u}_3) = (0,0,0,0).$$

Finalmente, para determinar T, dado $\mathbf{u} = (x,y,z) \in F^3$, obtemos

$$T(x,y,z) = xT(\mathbf{u}_1) + yT(\mathbf{u}_2) + zT(\mathbf{u}_3) = (x,y,y,-x),$$

que é o resultado desejado. ■

Seja $T \in \mathcal{L}(V, W)$. Diremos que T é *injetora* se

$$T(\mathbf{u}) = T(\mathbf{v}) \Rightarrow \mathbf{u} = \mathbf{v}, \ \forall \ \mathbf{u}, \mathbf{v} \in V$$

ou, equivalentemente, $\mathbf{u} \neq \mathbf{v}$ implica que $T(\mathbf{u}) \neq T(\mathbf{v})$, para todos $\mathbf{u}, \mathbf{v} \in V$. Diremos que T é *sobrejetora* se dado $\mathbf{w} \in W$, existir um $\mathbf{u} \in V$ tal que $T(\mathbf{u}) = \mathbf{w}$, isto é, $\operatorname{Im} T = W$. Finalmente, dizemos que T é *bijetora* se T é injetora e sobrejetora. Neste caso, obtemos a **relação fundamental**:

$$\mathbf{w} = T(\mathbf{u}) \Leftrightarrow \mathbf{u} = T^{-1}(\mathbf{w}) \Leftrightarrow T \circ T^{-1} = I_W \text{ e } T^{-1} \circ T = I_V.$$

Seja $T : F^2 \to F$ definida como $T(x, y) = x$. Então T é linear e sobrejetora, pois

$$\operatorname{Im} T = \{T(x, y) : (x, y) \in F^2\} = \{x \cdot 1 : x \in F\} = F[1] = F.$$

Mas não é injetora, pois $T(0, 1) = 0 = T(0, -1)$ e $(0, 1) \neq (0, -1)$. Por outro lado, seja $T : F \to F^2$ definida como $T(x) = (x, 0)$. Então T é linear e injetora, pois

$$T(x) = T(y) \Rightarrow (x, 0) = (y, 0) \Rightarrow x = y.$$

Mas não é sobrejetora, pois $T(x) \neq (0, 1)$, para todo $x \in F$, isto é, $\operatorname{Im} T \neq F^2$. Por fim, seja $T : F^3 \to F^3$ definida como $T(x, y, z) = (x, 2y, 0)$. Então T é linear, não é injetora e nem sobrejetora, pois $T(0, 0, 1) = (0, 0, 0) = T(0, 0, -1)$, com $(0, 0, 1) \neq (0, 0, -1)$ e $T(x, y, z) \neq (0, 0, 1)$, para todo $(x, y, z) \in F^3$, isto é, $\operatorname{Im} T \neq F^3$.

Sejam $T \in \mathcal{L}(V, W)$. Diremos que T é *não singular* se $\ker T = \{\mathbf{0}\}$. Caso contrário, diremos que T é *singular*.

Teorema 4.19 *Seja $T \in \mathcal{L}(V, W)$. Então T é não singular se, e somente se, T for injetora.*

Prova. Suponhamos que T seja não singular, isto é, $\ker T = \{\mathbf{0}\}$. Dados $\mathbf{u}, \mathbf{v} \in V$, se $T(\mathbf{u}) = T(\mathbf{v})$, então

$$T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v}) = T(\mathbf{u}) - T(\mathbf{u}) = \mathbf{0}.$$

Assim, $\mathbf{u} - \mathbf{v} \in \ker T = \{\mathbf{0}\}$. Portanto, $\mathbf{u} = \mathbf{v}$, ou seja, T é injetora. Reciprocamente, suponhamos que T seja injetora. Dado $\mathbf{u} \in \ker T$, obtemos $T(\mathbf{u}) = \mathbf{0}$. Como $T(\mathbf{0}) = \mathbf{0}$

temos que $T(\mathbf{u}) = \mathbf{0} = T(\mathbf{0})$ implica que $\mathbf{u} = \mathbf{0}$. Logo, $\ker T = \{\mathbf{0}\}$. Portanto, T é não singular. ■

Corolário 4.20 *Sejam $T \in \mathcal{L}(V, W)$. Então T é não singular se, e somente se, T leva qualquer conjunto LI de V em algum conjunto LI de W.*

Prova. Suponhamos que T seja não singular. Seja $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ conjunto qualquer LI de V. Então devemos provar que $T(\alpha) = \{T(\mathbf{u}_1), \ldots, T(\mathbf{u}_n)\}$ é um conjunto LI de W. Sejam $x_1, \ldots, x_n \in F$ tais que

$$x_1T(\mathbf{u}_1) + \cdots + x_nT(\mathbf{u}_n) = \mathbf{0}.$$

Assim,

$$T(x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n) = x_1T(\mathbf{u}_1) + \cdots + x_nT(\mathbf{u}_n) = \mathbf{0}.$$

Logo, $x_1\mathbf{u}_1+\cdots+x_n\mathbf{u}_n \in \ker T = \{\mathbf{0}\}$, de modo que $x_1\mathbf{u}_1+\cdots+x_n\mathbf{u}_n = \mathbf{0}$. Portanto, $x_1 = 0, \ldots, x_n = 0$, pois α é LI, e $T(\alpha)$ é um conjunto LI de W. Reciprocamente, seja $\mathbf{u} \in \ker T$, com $\mathbf{u} \neq \mathbf{0}$. Então $\{\mathbf{u}\}$ é um conjunto LI de V. Assim, $\{T(\mathbf{u})\} = \{\mathbf{0}\}$ é um conjunto LI de W, o que é impossível. Portanto, $\mathbf{u} = \mathbf{0}$ e T é não singular. é importante ressaltar que esta prova vale para um conjunto LI qualquer. ■

Teorema 4.21 (Teorema do Posto) *Seja $T \in \mathcal{L}(V, W)$, com $\dim V = n$. Então*

$$\dim V = \dim \ker T + \dim \operatorname{Im} T = \nu(T) + \rho(T).$$

Prova. Como $\ker T$ é um subespaço de V temos que $\ker T$ contém uma base $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ que é parte de uma base $\beta = \alpha \cup \{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ de V.

Afirmação. $\gamma = \{T(\mathbf{u}_{k+1}), \ldots, T(\mathbf{u}_n)\}$ é uma base de $\operatorname{Im} T$.

De fato. Dado $\mathbf{w} \in \operatorname{Im} T$, existe um $\mathbf{u} \in V$ tal que $\mathbf{w} = T(\mathbf{u})$. Como $\mathbf{u} \in V$ e β é uma base de V temos que existem $x_1, \ldots, x_n \in F$ tais que

$$\mathbf{u} = x_1\mathbf{u}_1 + \cdots + x_k\mathbf{u}_k + x_{k+1}\mathbf{u}_{k+1} + \cdots + x_n\mathbf{u}_n.$$

Assim,

$$\mathbf{w} = T(\mathbf{u}) = x_{k+1}T(\mathbf{u}_{k+1}) + \cdots + x_nT(\mathbf{u}_n),$$

pois $T(\mathbf{u}_i) = \mathbf{0}$, $i = 1, \ldots, k$. Logo, γ gera $\operatorname{Im} T$. Sejam $y_{k+1}, \ldots, y_n \in F$ tais que

$$y_{k+1}T(\mathbf{u}_{k+1}) + \cdots + y_nT(\mathbf{u}_n) = \mathbf{0}.$$

Então $T(y_{k+1}\mathbf{u}_{k+1} + \cdots + y_n\mathbf{u}_n) = \mathbf{0}$, de modo que $y_{k+1}\mathbf{u}_{k+1} + \cdots + y_n\mathbf{u}_n \in \ker T$. Assim, existem $x_1, \ldots, x_k \in \mathbb{R}$ tais que

$$y_{k+1}\mathbf{u}_{k+1} + \cdots + y_n\mathbf{u}_n = x_1\mathbf{u}_1 + \cdots + x_k\mathbf{u}_k,$$

ou seja,

$$x_1\mathbf{u}_1 + \cdots + x_k\mathbf{u}_k + (-y_{k+1})\mathbf{u}_{k+1} + \cdots + (-y_n)\mathbf{u}_n = \mathbf{0}.$$

Como β é uma base de V temos que $y_{k+1} = \cdots = y_n = 0$ e γ é um conjunto LI. Portanto,

$$\dim V = n = k + (n - k) = \dim \ker T + \dim \operatorname{Im} T$$

e $V = \ker T \oplus [\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n]$. ■

É muito importante, de um ponto de vista teórico e didático, observar que o Teorema do Núcleo e da Imagem continua válido quando $\dim V$ é infinito. Neste caso, se β for uma base de $\operatorname{Im} T$, então, para cada $\mathbf{w} \in \beta$, existe um $\mathbf{u} = \mathbf{u}_\mathbf{w} \in V$ tal que $\mathbf{u} \in T^{-1}(\mathbf{w})$. Assim, a família α dos tais $\mathbf{u}$, isto é, $\alpha = \{\mathbf{u}_\mathbf{w} : \mathbf{w} \in \beta\}$, é LI em V. Portanto, $V = F[\alpha] \oplus \ker T$, confira Exemplo 4.18.

Corolário 4.22 *Seja $T \in \mathcal{L}(V, W)$, com $\dim V = n$ e $\dim W = m$. Então o posto $k = \rho(T) \leq \min\{m, n\}$. Conclua que existe uma base $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ de V e existe uma base $\beta = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ de W tal que $T(\mathbf{u}_i) = \mathbf{w}_i$, com $i = 1, \ldots, k$, e $T(\mathbf{u}_j) = \mathbf{0}$, com $j = i + 1, \ldots, n$.*

Prova. Como $\operatorname{Im} T \subseteq W$ e, pelo Teorema 4.21, $\dim \operatorname{Im} T \leq n$ temos que o posto de T $k = \rho(T) \leq \min\{m, n\}$. Assim, pela prova do Teorema 4.21, existe uma base $\alpha = \{\mathbf{u}_1, \ldots \mathbf{u}_p, \mathbf{v}_1, \ldots, \mathbf{v}_{n-p}\}$ de V, com $\{\mathbf{u}_1, \ldots \mathbf{u}_p\}$ uma base de $\ker T$, tal que $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_{n-p})\}$ é uma base de $\operatorname{Im} T$. Pondo $\mathbf{w}_i = T(\mathbf{v}_i)$, com $i = 1, \ldots, k = n - p$, e estendendo para obter a base $\beta = \{\mathbf{w}_1, \ldots, \mathbf{w}_k, \mathbf{w}_{k+1}, \ldots, \mathbf{w}_m\}$ de W com as propriedades desejadas. ■

Corolário 4.23 *Seja $T \in \mathcal{L}(V, W)$, com $\dim V = \dim W = n$. Então T é injetora se, e somente se, T for sobrejetora.*

Prova. Suponhamos que T seja injetora. Então, pelo Teorema 4.19, $\ker T = \{\mathbf{0}\}$. Assim,

$$\dim W = \dim V = \dim \ker T + \dim \operatorname{Im} T = \dim \operatorname{Im} T.$$

Como $\operatorname{Im} T \subseteq W$ e $\dim W = \dim \operatorname{Im} T$ temos que $\operatorname{Im} T = W$. Portanto, T é sobrejetora. Reciprocamente, suponhamos que T seja sobrejetora. Então $\operatorname{Im} T = W$ e

$\dim W = \dim \operatorname{Im} T$. Logo,

$$\dim \operatorname{Im} T = \dim V = \dim \ker T + \dim \operatorname{Im} T \Rightarrow \dim \ker T = 0.$$

Portanto, $\ker T = \{\mathbf{0}\}$ e, pelo Teorema 4.19, T é injetora. ■

Corolário 4.24 *Seja $T \in \mathcal{L}(V, W)$, com $\dim V = \dim W = n$. Então as seguintes condições são equivalentes:*

1. *T é bijetora.*
2. *T é sobrejetora.*
3. *T é não singular.*
4. *T leva qualquer base de V em alguma base de W.*

Prova. Fica como um exercício. ■

Exemplo 4.25 *Determine $T \in \mathcal{L}(F^3, F^4)$ tal que* $\ker T = \{(x, y, z) \in \mathbb{R}^3 : x+y+z = 0\}$.

Solução. É fácil verificar que $\ker T = F[(1, 0, -1), (0, 1, -1)]$. Então

$$\left(\begin{array}{rr|c} 1 & 0 & a \\ 0 & 1 & b \\ -1 & -1 & c \end{array}\right) \to \cdots \to \left(\begin{array}{cc|c} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & a+b+c \end{array}\right).$$

Assim, $\alpha = \{(1, 0, -1), (0, 1, -1)\}$ é uma base de $\ker T$. Como $\ker T \subseteq F^3$ temos que α é parte de uma base de F^3. Logo, para estender α a uma base de F^3, basta escolher $\mathbf{u} = (a, b, c) \in F^3$, com $a + b + c \neq 0$, digamos $\beta = \alpha \cup \{(0, 0, 1)\}$. Definindo arbitrariamente $T(0, 0, 1) = \mathbf{w}$, temos, pelo Teorema 4.11, que existe uma única transformação linear $T : F^3 \to F^4$ tal que

$$T(1, 0, -1) = (0, 0, 0, 0), T(0, 1, -1) = (0, 0, 0, 0) \text{ e } T(0, 0, 1) = \mathbf{w}.$$

Logo, para determinar $T(\mathbf{u})$ basta encontrar $[\mathbf{u}]_\alpha$ ou, equivalentemente, escalonar a matriz

$$\left(\begin{array}{rrr|c} 1 & 0 & 0 & x \\ 0 & 1 & 0 & y \\ -1 & -1 & 1 & z \end{array}\right) \to \cdots \to \left(\begin{array}{ccc|c} 1 & 0 & 0 & x \\ 0 & 1 & 0 & y \\ 0 & 0 & 1 & x+y+z \end{array}\right).$$

Portanto, $T(x, y, z) = (x + y + z)\mathbf{w}$. ■

Teorema 4.26 *Sejam $S \in \mathcal{L}(U, V)$ e $T \in \mathcal{L}(V, W)$. Então $T \circ S \in \mathcal{L}(U, W)$. Conclua se T for bijetora, então $T^{-1} \in \mathcal{L}(W, V)$.*

Prova. É claro que $T^{-1}(\mathbf{0}) = \mathbf{0}$, pois $T(\mathbf{0}) = \mathbf{0}$. Dados $\mathbf{w}_1, \mathbf{w}_2 \in W$ e $a \in F$. Como T é bijetora temos que existem únicos $\mathbf{u}_1, \mathbf{u}_2 \in V$ tais que

$$\mathbf{w}_1 = T(\mathbf{u}_1) \Leftrightarrow \mathbf{u}_1 = T^{-1}(\mathbf{w}_1) \ \text{ e } \ \mathbf{w}_2 = T(\mathbf{u}_2) \Leftrightarrow \mathbf{u}_2 = T^{-1}(\mathbf{w}_2).$$

Note que $T(\mathbf{u}_1 + \mathbf{u}_2) = \mathbf{w}_1 + \mathbf{w}_2$ e $T(a\mathbf{u}_1) = a\mathbf{w}_1$ implicam que

$$T^{-1}(\mathbf{w}_1 + \mathbf{w}_2) = T^{-1}(\mathbf{w}_1) + T^{-1}(\mathbf{w}_2) \ \text{ e } \ T^{-1}(a\mathbf{w}_1) = aT^{-1}(\mathbf{w}_1).$$

Portanto, T^{-1} é linear. ■

Seja $T \in \mathcal{L}(V, W)$. Diremos que T é um *isomorfismo* se T é bijetora. Se existir um isomorfismo de V sobre W, diremos que V é *isomorfo* a W e denotamos por $V \simeq W$. Intuitivamente, um isomorfismo T de V sobre W é uma regra que consiste em renomear os elementos de V, isto é, o nome do elemento sendo $T(\mathbf{u})$ ao invés de $\mathbf{u} \in V$. É muito importante ressaltar o seguinte: se T for um isomorfismo e $S \subseteq V$. Então S é LI se, e somente se, $T(S)$ for LI. Portanto, ao decidirmos que S é LI (base) não importa se consideramos S ou $T(S)$, confira Corolário 4.24. Por exemplo, a função $T : F^{n\times 1} \to F^{1\times n}$ definida como $T(\mathbf{X}) = \mathbf{X}^t$ é claramente um isomorfismo. Portanto, isto justifica que sempre podemos identificar os espaços vetoriais F^n, $F^{1\times n}$ e $F^{n\times 1}$.

Exemplo 4.27 *Mostre que $T \in \mathcal{L}(F^3)$ definida como $T(x, y, z) = (x - 2y, z, x + y)$ é um isomorfismo e determine uma regra para T^{-1} como a que define T.*

Solução. Note que $\mathbf{v} \in \operatorname{Im} T$ é equivalente a resolver a equação vetorial $\mathbf{v} = T(\mathbf{u})$, para algum $\mathbf{u} \in F^3$. Mas, pelo Teorema 4.11, isto é equivalente a escalonar a matriz $(T(\mathbf{e}_1) = (1, 0, 1) \ldots)$

$$\left(\begin{array}{ccc|c} 1 & -2 & 0 & a \\ 0 & 0 & 1 & b \\ 1 & 1 & 0 & c \end{array}\right) \to \cdots \to \left(\begin{array}{ccc|c} 1 & 0 & 0 & \frac{a+2c}{3} \\ 0 & 1 & 0 & \frac{c-a}{3} \\ 0 & 0 & 1 & b \end{array}\right).$$

Assim, a equação possui uma solução e T é sobrejetora, de modo que T é um isomorfismo. Logo, para determinar uma regra para T^{-1} devemos usar a relação $\mathbf{v} = T(\mathbf{u}) \Leftrightarrow \mathbf{u} = T^{-1}(\mathbf{v})$, ou seja, dado $\mathbf{v} = (a, b, c) \in F^3$, existe um único $\mathbf{u} = (x, y, z) \in F^3$ tal que

$$T(x, y, z) = (a, b, c) \Leftrightarrow T^{-1}(a, b, c) = (x, y, z).$$

Portanto,

$$T^{-1}(a,b,c) = \left(\frac{a+2c}{3}, \frac{-a+c}{3}, b \right),$$

que é o resultado desejado. ■

Teorema 4.28 *Qualquer espaço vetorial de dimensão n sobre F é isomorfo a F^n.*

Prova. Seja $\alpha = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ uma base de V. Então já vimos que a função $L_\alpha : F^n \to V$ definida como $L_\alpha(\mathbf{x}) = x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n$ é linear e bijetora. Portanto, V é isomorfo a F^n. ■

A transformação linear $L_\alpha : F^n \to V$ chama-se *parametrização* de V induzida pela base α. Note que $L_\alpha(\mathbf{e}_i) = \mathbf{u}_i, i = 1, \dots, n$.

Vamos finalizar esta seção com algumas considerações sobre soma direta externa e espaços quocientes. Sejam V e W espaços vetoriais sobre F. Então já vimos, no Exercício (13) da Seção 3.1, que $V \otimes W = V \times W$, munido com as operações de adição

$$\mathbf{u} \oplus \mathbf{v} = (\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}_1 + \mathbf{v}_2)$$

e multiplicação por escalar

$$a \odot \mathbf{u} = (a\mathbf{u}_1, a\mathbf{v}_1),$$

era um espaço vetorial sobre F. Então, naturalmente, obtemos quatro funções lineares, a saber, as inclusões $\lambda_1 : V \to V \otimes W$ e $\lambda_2 : W \to V \otimes W$ definidas como $\lambda_1(\mathbf{u}) = (\mathbf{u}, \mathbf{0})$ e $\lambda_2(\mathbf{v}) = (\mathbf{0}, \mathbf{v})$; as projeções $\pi_1 : V \otimes W \to V$ e $\pi_2 : V \otimes W \to W$ definidas como $\pi_1(\mathbf{u}, \mathbf{v}) = \mathbf{u}$ e $\pi_2(\mathbf{u}, \mathbf{v}) = \mathbf{v}$. É fácil verificar que elas satisfazem as relações:

$$\pi_1 \circ \lambda_1 = I_V, \quad \pi_2 \circ \lambda_2 = I_W, \quad \pi_i \circ \lambda_j = 0, i \neq j, \quad \text{e} \quad \lambda_1 \circ \pi_1 + \lambda_2 \circ \pi_2 = I_Z,$$

com $Z = V \otimes W$, confira o diagrama:

$$\begin{tikzcd}
V \arrow[dr, "\lambda_1"] & & V \\
& V \otimes W \arrow[ur, "\pi_1"] \arrow[dr, "\pi_2"] & \\
W \arrow[ur, "\lambda_2"] & & W
\end{tikzcd}$$

Teorema 4.29 *Sejam V e W espaços vetoriais de dimensões finitas sobre F. Então* $\dim(V \otimes W) = \dim V + \dim W$.

Prova. Sejam $\alpha = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ uma base de V e $\beta = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ base de W. Então é fácil verificar que

$$\gamma = \{(\mathbf{v}_1, \mathbf{0}), \ldots, (\mathbf{v}_m, \mathbf{0}), (\mathbf{0}, \mathbf{w}_1), \ldots, (\mathbf{0}, \mathbf{w}_n)\}$$

é uma base de $V \times W$. Portanto, $\dim(V \otimes W) = \dim V + \dim W$. ■

Sejam V um espaço vetorial sobre F e W um subespaço de V. Então já vimos, no Exercício (19) da Seção 3.5, que $\overline{V} = V/W = \{\mathbf{u} + W : \mathbf{u} \in V\}$, munido com as operações de adição

$$(\mathbf{u} + W) \oplus (\mathbf{v} + W) = (\mathbf{u} + \mathbf{v}) + W$$

e multiplicação por escalar

$$a \odot (\mathbf{u} + W) = (a\mathbf{u}) + W,$$

era um espaço vetorial sobre F. Então, naturalmente, obtemos uma função linear e sobrejetora $\pi : V \to \overline{V}$ definida como $\pi(\mathbf{u}) = \mathbf{u} + W$, a qual chama-se *projeção canônica*.

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\pi"'] & W \\ V/kerT \arrow[r, dashed, "\sigma"'] & \operatorname{Im} T \arrow[u, "\lambda"'] \end{tikzcd}$$

Teorema 4.30 *Seja $T \in \mathcal{L}(V, W)$. Então $V/\ker T \simeq \operatorname{Im} T$. Conclua que* $\dim V = \dim \ker T + \dim \operatorname{Im} T$.

Prova. Note, pelo Teorema 4.16, que $\ker T$ é um subespaço de V e $\operatorname{Im} T$ é um subespaço de W. Assim, $V/\ker T$ é um espaço vetorial sobre F. Logo, pelo diagrama, $\lambda \circ \sigma \circ \pi = T$ significa que: $T(\mathbf{u}) = \sigma(\mathbf{u} + \ker T)$, para todo $\mathbf{u} \in V$. Portanto, a função $\sigma : V/\ker T \to W$ definida como $\sigma(\mathbf{u} + \ker T) = T(\mathbf{u})$ possui as propriedades desejadas. Por exemplo, $\operatorname{Im} \sigma = \operatorname{Im} T$ implica que σ é sobrejetora e $\ker \sigma = \ker T$ implica que σ é injetora, pois $T(\mathbf{u}) = \mathbf{0}$ se, e somente se, $\mathbf{u} + \ker T = \ker T$ é o zero em $V/\ker T$. ■

Vimos acima que $\pi_2 \circ \lambda_2 = I_W$ significa que π_2 possui uma inversa à direita e

$\pi_2 \circ \lambda_1 = 0$ significa que $\operatorname{Im} \lambda_1 = \ker \pi_2$, mas isto é equivalente a sequência:

$$\{\mathbf{0}\} \longrightarrow V \xrightarrow{\lambda_1} V \otimes W \xrightarrow{\pi_2} W \longrightarrow \{\mathbf{0}\}.$$

Isto motiva a definição. Sejam U, V e W espaços vetoriais sobre F. Diremos que

$$\{\mathbf{0}\} \longrightarrow U \xrightarrow{S} V \xrightarrow{T} W \longrightarrow \{\mathbf{0}\}.$$

é uma *sequência exata curta* "em V" se S for injetora e T for sobrejetora, ou seja, $\operatorname{Im} S = \ker T$. Neste caso, obtemos o diagrama:

$$\begin{tikzcd}
\{\mathbf{0}\} \arrow[r] & \operatorname{Im} S \arrow[r, "\lambda"] & V \arrow[r, "\pi"] \arrow[dr, "T"'] & W/\operatorname{Im} T \arrow[r] \arrow[d, dashed, "\phi"] & \{\mathbf{0}\} \\
\{\mathbf{0}\} \arrow[r] & U \arrow[u, dashed, "\sigma"] \arrow[ur, "S"'] & & W \arrow[r] & \{\mathbf{0}\}
\end{tikzcd}$$

Note que σ e ϕ são isomorfismos, de modo que $V = U \oplus W$ é uma soma direta.

Exercícios

1. Seja $T \in \mathcal{L}(V, W)$. Mostre se U for um subespaço de V, então a *imagem direta* de U: $T(U) = \{T(\mathbf{u}) : \mathbf{u} \in U\}$, é um subespaço de W.

2. Seja $T \in \mathcal{L}(V, W)$. Mostre que se Z for um subespaço de W, então a *imagem inversa* de Z: $T^{-1}(Z) = \{\mathbf{u} \in V : T(\mathbf{u}) \in Z\}$, é um subespaço de V.

3. Sejam $T \in \mathcal{L}(\mathbb{R}^2)$ definido como $T(x, y) = (x + y, y)$ e

$$\begin{aligned} A &= \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}, B = \{(x,y) \in \mathbb{R}^2 : |x| + |y| = 1\}, \\ C &= \{(x,y) \in \mathbb{R}^2 : \max\{|x|, |y|\} = 1\}. \end{aligned}$$

   Determine $T(A)$, $T(B)$ e $T(C)$.

4. Para cada transformação linear abaixo determine o núcleo e a imagem:

   (a) $T \in \mathcal{L}(F^2, F^3)$ definida como $T(x, y) = (y - x, 0, 5x)$.

   (b) $T \in \mathcal{L}(F^3, F^2)$ definida como $T(x, y, z) = (x + y + z, z)$.

5. Seja $T \in \mathcal{L}(V, W)$. Mostre que se $V = F[\mathbf{u}_1, \ldots, \mathbf{u}_n] = F[\alpha]$, então $\operatorname{Im} T = F[T(\alpha)]$.

6. Seja $T \in \mathcal{L}(F^3, F^3)$ definida como $T(x, y, z) = (x-y+2z, 2x+y, -x-2y+2z)$.

   (a) Se (a, b, c) for um vetor em F^3, quais as condições sobre a, b e c, para que o vetor esteja na imagem de T? Qual é o posto de T?

   (b) Quais as condições sobre a, b e c, para que o vetor esteja no núcleo de T? Qual é a nulidade de T?

7. Determine $S, T \in \mathcal{L}(F^3)$ tais que $\operatorname{Im} S = F[(1, 0, -1), (1, 2, 2)]$ e $\operatorname{Im} T = [(1, 2, 3), (4, 0, 5)]$.

8. Determine $T \in \mathcal{L}(F^3)$ tal que $\ker T = F[(1, 1, 0)]$.

9. Determine $T \in \mathcal{L}(F^3, F^2)$ sobrejetora tal que $T(1, 1, 0) = T(0, 0, 1)$.

10. Existe um $T \in \mathcal{L}(F^3, F^2)$ tal que $T(1, -1, 1) = (1, 0)$ e $T(1, 1, 1) = (0, 1)$?

11. Existe um $T \in \mathcal{L}(F^2)$ tal que $T(1, -1) = (1, 0), T(2, -1) = (0, 1)$ e $T(-3, 2) = (1, 1)$?

12. Seja $T \in \mathcal{L}(V, W)$, com $\dim W = m$. Mostre que se $\beta = \{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ for uma base de $\operatorname{Im} T$ e $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$, com $T(\mathbf{u}_i) = \mathbf{w}_i$, então α é LI e $V = F[\alpha] \oplus \ker T$. Conclua que $\dim T(U) = \dim U - \dim(U \cap \ker T)$, para todo subespaço U de V.

13. Existe $T \in \mathcal{L}(F^5, F^3)$, com $T(\mathbf{e}_1) = (1, 0, 0), T(\mathbf{e}_2) = (0, 1, 0)$ e cujo núcleo consiste dos vetores $(x_1, x_2, x_3, x_4, x_5) \in F^5$ tais que

$$\left\{ \begin{array}{rcl} x_1 - 2x_2 + x_3 + x_4 - x_5 & = & 0 \\ x_1 + x_2 - 2x_3 + x_4 - x_5 & = & 0 \\ -2x_1 + x_2 + x_3 - 2x_4 + 2x_5 & = & 0? \end{array} \right.$$

14. Sejam $T \in \mathcal{L}(V, F^3)$ um isomorfismo e $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4 \in V$ tais que $T(\mathbf{u}_1) = (1, 0, 1), T(\mathbf{u}_2) = (-2, 1, 0), T(\mathbf{u}_3) = (-1, 1, 1)$ e $T(\mathbf{u}_4) = (2, 1, 3)$.

    (a) $\mathbf{u}_1$ está no subespaço gerado por $\mathbf{u}_2$ e $\mathbf{u}_3$?

    (b) Sejam $W_1 = F[\mathbf{u}_1, \mathbf{u}_2]$ e $W_2 = F[\mathbf{u}_3, \mathbf{u}_4]$. Qual é $W_1 \cap W_2$?

    (c) Determine uma base de $F[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4]$.

15. Sejam $S, T \in \mathcal{L}(V)$, com $\dim V = n$, tais que $\nu(S) = \nu(T) = 0$. Mostre que $\nu(T \circ S) = 0$.

16. Sejam $S \in \mathcal{L}(U, V)$ e $T \in \mathcal{L}(V, W)$, com $\dim V = n$ e $\dim W = m$.

    (a) Mostre que $\operatorname{Im}(T \circ S) \subseteq \operatorname{Im} T$ e $\dim \operatorname{Im}(T \circ S) \leq \dim \operatorname{Im} T$.

    (b) Mostre que $\ker S \subseteq \ker(T \circ S)$ e $\dim kerS \leq \dim \ker(S \circ T)$.

17. Sejam $T \in \mathcal{L}(V, W)$ e $S \in \mathcal{L}(W, V)$, com $m = \dim W > \dim V = n$. Mostre que $T \circ S$ não pode ser injetora.

18. Seja $T \in \mathcal{L}(V, W)$, com $\dim V = n$ e $\dim W = m$. Mostre que V e W são isomorfos se, e somente se, $m = n$.

19. Seja $T \in \mathcal{L}(V, W)$, com $\dim V = n$ e $\dim W = m$.

    (a) Mostre que se $\dim V < \dim W$, então T não pode ser sobrejetora.

    (b) Mostre que se $\dim V > \dim W$, então T não pode ser injetora.

20. Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Mostre que se existir um $S \in \mathcal{L}(V)$ tal que $S \circ T = I_V$, então T^{-1} existe e $T^{-1} = S$.

21. Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Mostre que se $T^2 - T + I = 0$, então T é não singular. Determine T^{-1} em função de T.

22. Sejam U, V e W espaços vetoriais sobre F, com $\dim U = k, \dim V = n$ e $\dim W = m$.

    (a) Sejam $S \in \mathcal{L}(U, W)$ e $T \in \mathcal{L}(V, W)$. Mostre que $\operatorname{Im} S \subseteq \operatorname{Im} T$ se, e somente se, existir um $R \in \mathcal{L}(U, V)$ tal que $S = T \circ R$.

    (b) Sejam $S \in \mathcal{L}(U, W)$ e $T \in \mathcal{L}(U, V)$. Mostre que $\ker T \subseteq \ker S$ se, e somente se, existir um $R \in \mathcal{L}(V, W)$ tal que $S = R \circ T$.

23. Sejam $S, T \in \mathcal{L}(V)$, com $\dim V = n$. Mostre que:

    (a) $\rho(T + S) \leq \rho(S) + \rho(T)$.

    (b) $\nu(S) + \nu(T) - n \leq \nu(S + T)$.

    (c) $\max\{\nu(S), \nu(T)\} \leq \nu(S \circ T) \leq \nu(S) + \nu(T)$.

    (d) $\rho(S) + \rho(T) - n \leq \rho(S \circ T) \leq \min\{\rho(S), \rho(T)\}$.

24. Seja $T \in \mathcal{L}(V, W)$, com $\dim W = m$. Mostre que T é injetora se, e somente se, existir um $S \in \mathcal{L}(W, V)$ tal que $S \circ T = I_V$. Conclua que se $V = W$ e T for não injetora, então existe um $R \in \mathcal{L}(V)$ tal que $T \circ R = 0$.

25. Seja $T \in \mathcal{L}(V, W)$, com $\dim V = n$. Mostre que T é sobrejetora se, e somente se, existir um $S \in \mathcal{L}(W, V)$ tal que $T \circ S = I_W$. Conclua que se $V = W$ e T for não sobrejetora, então existe um $R \in \mathcal{L}(V)$ tal que $R \circ T = 0$.

26. Sejam $S \in \mathcal{L}(U, V)$ e $T \in \mathcal{L}(V, W)$, com $\dim U = n$ e $\dim V = m$. Mostre que S for sobrejetora e T for injetora se, e somente se, existir um $R \in \mathcal{L}(W, U)$ tal que $S \circ R \circ T = I_V$.

27. Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Mostre que $\operatorname{Im} T = \ker T$ se, e somente se, $n = 2 \dim \operatorname{Im} T$ e $T^2 = 0$, mas $T \neq 0$. Determine vários $T \in \mathcal{L}(\mathbb{R}^2)$ que satisfaça essa condição.

28. Sejam $T \in \mathcal{L}(V, W)$ e $\mathbf{w}_0 \in W$. Mostre que se a equação $T(\mathbf{u}) = \mathbf{w}_0$ possuir uma solução $\mathbf{u}_0 \in V$, então qualquer solução desta equação é da forma $\mathbf{u}_0 + \mathbf{v}$, para algum $\mathbf{v} \in \ker T$. Conclua que $V = T^{-1}(\mathbf{w}_0) \oplus \ker T$.

29. Sejam $V = \mathbb{R}[x]$ e $D, E, T, U \in \mathcal{L}(V)$ definidas como

$$\begin{array}{rcl} D(\sum_{i=0}^{n} a_i x^i) & = & \sum_{i=1}^{n} i a_i x^{i-1}, \;\; E(\sum_{i=0}^{n} a_i x^i) = \sum_{i=0}^{n} \frac{a_i}{i+1} x^{i+1}, \\ T(\sum_{i=0}^{n} a_i x^i) & = & \sum_{i=0}^{n} a_i x^{i+1} \text{ e } U(\sum_{i=0}^{n} a_i x^i) = \sum_{i=1}^{n} a_i x^{i-1}. \end{array}$$

Mostre que E e T sâo injetoras, mas não são sobrejetoras. Além disso, $DE = I$ e $ED \neq I$; $UT = I$ e $TU \neq I$.

30. Sejam $S, T \in \mathcal{L}(V)$, com $\dim V = n$. Mostre que S e T são não singulares se, e somente se, $S \circ T$ e $T \circ S$ são não singulares.

31. Sejam $S_i, T_i \in \mathcal{L}(V)$, com $i = 1, 2$ e $\dim V = n$. Mostre que se $S_1 + S_2$ e $S_1 - S_2$ forem não singulares, então existem $X_i \in \mathcal{L}(V)$, com $i = 1, 2$, tais que

$$S_1 \circ X_1 + S_2 \circ X_2 = T_1 \text{ e } S_2 \circ X_1 + S_1 \circ X_2 = T_2.$$

32. Seja $S \in \mathcal{L}(W, V)$ um isomorfismo. Mostre que a função $f : \mathcal{L}(V) \to \mathcal{L}(W)$ definida como $f(T) = S^{-1} \circ T \circ S$ é um isomorfismo.

33. Sejam W_1 e W_2 subespaços de V, com $\dim V = n$.

(a) Mostre que a função $T : W_1 \times W_2 \to V$ definida como $T(\mathbf{w}_1, \mathbf{w}_2) = \mathbf{w}_1 + \mathbf{w}_2$ é linear.

(b) Mostre que T é injetora se, e somente se, $W_1 \cap W_2 = \{\mathbf{0}\}$.

(c) Mostre que T é sobrejetora se, e somente se, $V = W_1 + W_2$.

(d) Mostre que

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

34. Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Mostre que existe um $k \in \mathbb{N}$ tal que $\operatorname{Im} T^k \cap \ker T^k = \{\mathbf{0}\}$. Conclua que $V = \operatorname{Im} T^k \oplus \ker T^k$.

35. Sejam $T \in \mathcal{L}(V, W)$ e $S \in \mathcal{L}(W, V)$, com $\dim V = n$ e $\dim W = m$. Mostre que $\operatorname{Im}(T \circ S) = T(\operatorname{Im} S)$ e $\dim \operatorname{Im}(T \circ S) = \dim \operatorname{Im} S - \dim(\operatorname{Im} S \cap \ker T)$.

36. Seja V um espaço vetorial sobre F tal que $V = U \oplus W$.

(a) Mostre que a função $P : V \to V$ definida como $P(\mathbf{u} + \mathbf{w}) = \mathbf{u}$ é linear, a qual chama-se *projeção* de V sobre U paralela a W. Conclua que $Q = I_V - P$ é a projeção de V sobre W paralela a U e $\ker Q = \operatorname{Im} P$.

(b) Mostre que a função $R : V \to V$ definida como $R(\mathbf{u}+\mathbf{w}) = \mathbf{u}-\mathbf{w}$ é linear, a qual chama-se *reflexão* de V em U paralela a W.

37. Seja $T \in \mathcal{L}(V, W)$. Mostre que existe um $S \in \mathcal{L}(W, V)$ tal que $S \circ T = I_V - P$ e $T \circ S = I_W - Q$, em que $P \in \mathcal{L}(V)$ é a projeção sobre $\ker T$ e $Q \in \mathcal{L}(W)$ é a projeção paralela a $\operatorname{Im} T$.

38. Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Mostre que existe um $S \in \mathcal{L}(V)$ não singular tal que $S \circ T = P$, com P uma projeção de V.

39. Sejam $T \in \mathcal{L}(V, W)$ e $S \in \mathcal{L}(W, V)$. Mostre que se $T \circ S = I_W$, então

(a) S é injetora.

(b) T é sobrejetora.

(c) $\operatorname{Im} S \cap \ker T = \{\mathbf{0}\}$.

(d) $V = \operatorname{Im} S \oplus \ker T$.

40. Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Mostre que as seguintes condições são equivalentes:

(a) $V = \ker T + \operatorname{Im} T$;

(b) $V = \ker T \oplus \operatorname{Im} T$;

(c) $\ker T \cap \operatorname{Im} T = \{\mathbf{0}\}$;

(d) $\ker T^2 = \ker T$;

(e) $\operatorname{Im} T^2 = \operatorname{Im} T$;

(f) $T \circ S \circ T = T$ e $\operatorname{Im}(S \circ T) = \operatorname{Im} T$, para algum $S \in \mathcal{L}(V)$ não singular.

Conclua que $T^2 = cT$, para algum $c \in F^*$, satisfaz essas condições e determine vários $T \in \mathcal{L}(\mathbb{R}^2)$ que satisfaça essas condições.

41. Seja $\mathbb{I} = [0, 1]$ um intervalo em $\mathbb{R}$. Uma *rede* sobre $\mathbb{I}$ é qualquer lista $(a_i)_{i=0}^n$ tal que $0 = a_0 < a_1 < \cdots < a_n = 1$. Se existirem uma rede $(a_i)_{i=0}^n$ sobre $\mathbb{I}$ e uma lista $(b_i)_{i=0}^n$ sobre $\mathbb{R}$, então a função $f : \mathbb{I}_1 \to \mathbb{R}$, com $\mathbb{I}_1 = \mathbb{I} - \{1\}$, definida como $f(x) = b_i$, para todo $x \in [a_i, a_{i+1})$, chama-se *função degrau*. Mostre que o conjunto das funções degraus $\mathcal{F}_d(\mathbb{I}_1, \mathbb{R})$ é um espaço vetorial sobre $\mathbb{R}$ e determine uma base para ele.

42. Se existirem uma rede $(a_i)_{i=0}^n$ sobre $\mathbb{I}$ e listas $(b_i)_{i=0}^n, (c_i)_{i=0}^n$ sobre $\mathbb{R}$, então a função $f : \mathbb{I}_1 \to \mathbb{R}$ definida como $f(x) = c_i + b_i x$, para todo $x \in [a_i, a_{i+1})$, chama-se *função linear por partes*. Mostre que o conjunto das funções lineares por partes $\mathcal{F}_l(\mathbb{I}_1, \mathbb{R})$ é um espaço vetorial. Conclua que o subconjunto $\mathcal{C}(\mathbb{I}_1, \mathbb{R})$ das $g \in \mathcal{F}_l(\mathbb{I}_1, \mathbb{R})$ tal que g é contínua e $g(0) = 0$ é um subespaço, com $\mathcal{F}_l(\mathbb{I}_1, \mathbb{R}) = \mathcal{F}_d(\mathbb{I}_1, \mathbb{R}) \oplus \mathcal{C}(\mathbb{I}_1, \mathbb{R})$ e a função $T : \mathcal{F}_d(\mathbb{I}_1, \mathbb{R}) \to \mathcal{C}(\mathbb{I}_1, \mathbb{R})$ definida como
$$T(f) = \int_0^x f(t)dt$$
é um isomorfismo.

4.3 Transformações Lineares e Matrizes

Nesta seção mostraremos, de um ponto de vista matemático, que o estudo de transformações lineares em espaços vetoriais de dimensão finita pode ser reduzido ao estudo de matrizes. Para isto vamos lembrar alguns fatos:

Sejam $\mathbf{X} \in F^{n\times 1}$ e $\mathbf{A} \in F^{m\times n}$. Então
$$\mathbf{E}_i^t\mathbf{X} = x_i \text{ e } \mathbf{E}_i^t\mathbf{A}\mathbf{E}_j = \begin{pmatrix} \mathbf{E}_i^t\mathbf{C}_1 & \cdots & \mathbf{E}_i^t\mathbf{C}_n \end{pmatrix}\mathbf{E}_j = a_{ij},\ \ \mathbf{E}_i^t\mathbf{C}_k = a_{ik}.$$

Observe que a função $\sigma : F^{n\times 1} \to F^n$ definida como $\sigma(\mathbf{X}) = \mathbf{X}^t$ é um isomorfismo. Portanto, não há perda de generalidade, em fazermos $F^n = F^{n\times 1}$.

$$\begin{tikzcd}
F^{n\times 1} \arrow[r, "T_{\mathbf{A}}"] \arrow[d, "\sigma"'] & F^{m\times 1} \arrow[d, "\phi"] \\
F^n \arrow[r, dashed, "S_{\mathbf{A}}"'] & F^m
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
V \arrow[r, "T"] \arrow[d, shift left, "T_\alpha"] & W \arrow[d, "T_\beta"] \\
F^n \arrow[u, dashed, shift left, "T_\alpha^{-1}"] \arrow[r, dashed, "S"'] & F^m
\end{tikzcd}$$

(a) (b)

Já vimos no Exemplo 4.4 que, para cada $\mathbf{A} \in F^{m\times n}$ fixada, existia uma única transformação linear $T_{\mathbf{A}} : F^{n\times 1} \to F^{m\times 1}$ definida como

$$T_{\mathbf{A}}(\mathbf{X}) = \mathbf{A}\mathbf{X},\ \ \forall\ \mathbf{X} \in F^{n\times 1},$$

e, também, $S_{\mathbf{A}} : F^n \to F^m$ definida como $S_{\mathbf{A}}(\mathbf{x}) = \mathbf{x}\mathbf{A}^t$. Logo, obtemos o diagrama comutativo (a): isto significa que $\phi \circ T_{\mathbf{A}} = S_{\mathbf{A}} \circ \sigma$ ou $S_{\mathbf{A}} = \phi \circ T_{\mathbf{A}} \circ \sigma^{-1}$. Portanto, quando $m = n$ temos que $T_{\mathbf{A}}$ é um isomorfismo se, e somente se, $\mathbf{A}$ é não singular.

Teorema 4.31 *Seja $T \in \mathcal{L}(F^n, F^m)$. Então existe uma única $\mathbf{A} \in F^{m\times n}$ tal que $T(\mathbf{X}) = \mathbf{A}\mathbf{X}$, para todo $\mathbf{X} \in F^n$. Neste caso, $a_{ij} = \mathbf{E}_i^t T(\mathbf{E}_j)$.*

Prova. (Existência) Note que $T(\mathbf{X})_i = \mathbf{E}_i^t T(\mathbf{X}) = \mathbf{E}_i^t x_j T(\mathbf{E}_j) = (\mathbf{E}_i^t T(\mathbf{E}_j))x_j$. Pondo $a_{ij} = \mathbf{E}_i^t T(\mathbf{E}_j)$, obtemos $T(\mathbf{X}) = \mathbf{A}\mathbf{X}$, com $\mathbf{A} = (a_{ij})$. Mais explicitamente, $T(\mathbf{E}_j) = a_{1j}\mathbf{E}_1 + \cdots + a_{mj}\mathbf{E}_m$ e $T(\mathbf{X}) = x_1 T(\mathbf{E}_1) + \cdots + x_n T(\mathbf{E}_n)$ implicam que $\mathbf{E}_i^t T(\mathbf{E}_j) = a_{ij}$ e

$$T(\mathbf{X}) = \begin{pmatrix} T(\mathbf{E}_1) & \cdots & T(\mathbf{E}_n) \end{pmatrix} \mathbf{X}.$$

Portanto, $T(\mathbf{X}) = \mathbf{A}\mathbf{X}$.

(Unicidade) Seja outra $\mathbf{B} \in F^{m\times n}$ tal que $T(\mathbf{X}) = \mathbf{B}\mathbf{X}$, para todo $\mathbf{X} \in F^n$. Em particular, $b_{ij} = \mathbf{E}_i^t \mathbf{B}\mathbf{E}_j = \mathbf{E}_i^t T(\mathbf{E}_j) = \mathbf{E}_i^t \mathbf{A}\mathbf{E}_j = a_{ij}$. Portanto, $\mathbf{A} = \mathbf{B}$. ■

É muito importante observar a ordem de m e n. Enquanto, $\mathbf{A}$ é de ordem $m \times n$, T é de F^n em F^m. Além disso, os vetores $\mathbf{C}_j = T(\mathbf{E}_j)$, $j = 1, \ldots, n$, são as colunas de $\mathbf{A}$, a qual chama-se *matriz canônica* para T. Em geral, sejam $T \in \mathcal{L}(V, W)$, $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ uma base de V e $\beta = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ uma base de W. Então $T(\mathbf{u}_j) \in W$, com $j = 1, \ldots, n$. Como β é uma base de W temos que existem únicos $a_{ij} \in F$ tais que

$$T(\mathbf{u}_j) = \sum_{i=1}^{m} a_{ij}\mathbf{w}_i, j = 1, \ldots, n. \tag{4.1}$$

Portanto, a ação de T sobre α é completamente determinada pelos mn escalares a_{ij} ou, equivalentemente, a ação de T é completamente determinada conhecendo uma matriz $\mathbf{A} = (a_{ij}) \in F^{n\times m}$. Por razões técnicas, a matriz transposta $\mathbf{A}^t$ chama-se *representação matricial* de T em relação às bases α e β, e denotamos por

$$[T]_\alpha^\beta = \begin{pmatrix} [T(\mathbf{u}_1)]_\beta & \cdots & [T(\mathbf{u}_n)]_\beta \end{pmatrix} = (a_{ij}). \tag{4.2}$$

Observe a ordem de α (domínio) e β (contradomínio) em $[T]_\alpha^\beta$. Reciprocamente, dado $\mathbf{A} = (a_{ij}) \in F^{m\times n}$ e $\mathbf{u} = x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n \in V$. Então, pelo Teorema 4.11, existe uma única $T \in \mathcal{L}(V, W)$ tal que

$$T(\mathbf{u}) = \sum_{j=1}^{n} x_j(\sum_{i=1}^{m} a_{ij}\mathbf{w}_i) = \sum_{i=1}^{m}(\sum_{j=1}^{n} x_j a_{ij})\mathbf{w}_i.$$

Neste caso, $\mathbf{A} = \mathbf{A}_T$. Portanto, a função $\varphi : L(V, W) \to F^{m\times n}$ definida como $\varphi(T) = \mathbf{A}_T$ é claramente um isomorfismo. Por outro lado, é fácil verificar que a função $\phi_\alpha : L(V, W) \to W^n$ definida como $\phi_\alpha(T) = (T(\mathbf{u}_1), \ldots, T(\mathbf{u}_n))$ é um isomorfismo induzido por α.

Teorema 4.32 *Seja $T \in \mathcal{L}(V, W)$, com $\dim V = n$ e $\dim W = m$. Então $W^n \simeq \mathcal{L}(V, W) \simeq F^{m\times n}$.*

Prova. Confira o exposto acima. ∎

É muito importante, de um ponto de vista teórico e didático, observar o seguinte: sejam $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ e $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ bases de V. Então $S = T_\beta \circ T_\alpha^{-1}$ está em $\mathcal{L}(F^n)$ e $\mathbf{A} = [S]$ é a matriz de transição $[I]_\beta^\alpha$, pois $\mathbf{u}_j = \sum_{i=1}^n a_{ij}\mathbf{v}_i$ implica que

$$S(\mathbf{e}_j) = T_\beta(\mathbf{u}_j) = \sum_{i=1}^{n} a_{ij}T_\beta(\mathbf{v}_i) = \sum_{i=1}^{n} a_{ij}\mathbf{e}_i.$$

Portanto, $\mathbf{A} = [S] = (a_{ij})$, com $\mathbf{C}_j = [S(\mathbf{e}_j)]$ sua j-ésima coluna. Por outro lado, como $[\mathbf{u}_j]_\beta = T_\beta(\mathbf{u}_j) = S(\mathbf{v}_j)$ temos, pela equação (3.6), que

$$[I]_\beta^\alpha = \begin{pmatrix} [\mathbf{u}_1]_\beta & \cdots & [\mathbf{u}_n]_\beta \end{pmatrix} = (a_{ij}).$$

Neste caso, $S = T_\beta \circ T_\alpha^{-1}$ chama-se *transformação de coordenadas*, confira o diagrama comutativo (b) acima, com $V = W$ e $T = I_V$. É bom lembrar que $S = L_\beta^{-1} \circ$

L_α. No caso geral, sejam $T \in \mathcal{L}(V, W)$, $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ uma base de V e $\beta = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ uma base W. Então, pelo diagrama comutativo (b) acima, $S = T_\beta \circ T \circ T_\alpha^{-1} \in \mathcal{L}(F^n, F^m)$ ou $S = L_\beta^{-1} \circ T \circ L_\alpha \in \mathcal{L}(F^n, F^m)$ e sua representação matricial em relação às bases canônicas é $\mathbf{A} = [S] = [T]_\alpha^\beta$, pois, usando a equação (4.1),

$$S(\mathbf{e}_j) = T_\beta(T(\mathbf{u}_j)) = \sum_{i=1}^{m} a_{ij} T_\beta(\mathbf{w}_i) = \sum_{i=1}^{m} a_{ij}\mathbf{e}_i.$$

Portanto, se identificamos V com F^n via L_α e W com F^m via L_β, então T pode ser identificada com a matriz $[T]_\beta^\alpha$ e concluirmos que: o problema de determinar $[T]_\alpha^\beta$ é equivalente, pela equação (4.1), a discutir a sequência finita de sistemas não homogêneos $\mathbf{AX} = \mathbf{B}_i$, com

$$\mathbf{A} = \begin{pmatrix} \mathbf{w}_1 & \cdots & \mathbf{w}_m \end{pmatrix}, \ \ \mathbf{X} = (\mathbf{x}) \ \text{ e } \ \mathbf{B}_i = (T(\mathbf{u}_i)), \ \ i = 1, \ldots, n.$$

Já vimos que isso é dado pelo esquema:

$$\left(\begin{array}{c|c} \beta & T(\alpha) \end{array} \right) = \left(\begin{array}{ccc|ccc} \mathbf{w}_1 & \cdots & \mathbf{w}_m & T(\mathbf{u}_1) & \cdots & T(\mathbf{u}_n) \end{array} \right) \to \cdots \to \left(\begin{array}{c|c} \mathbf{I} & [T]_\alpha^\beta \end{array} \right).$$

Exemplo 4.33 *Sejam* $\alpha = \{(1,1,1), (1,1,0), (1,0,0)\}, \beta = \{(1,3), (1,4)\}$ *bases de* F^3, F^2 *e* $T \in \mathcal{L}(F^3, F^2)$ *definida como* $T(x, y, z) = (2x + y - z, 3x - 2y + 4z)$. *Determine* $[T]_\alpha^\beta$.

Solução. Note que $T(1,1,1) = (2,5), T(1,1,0) = (3,1)$ e $T(1,0,0) = (2,3)$. Assim, pelo exposto, basta escalonar a matriz

$$\left(\begin{array}{cc|ccc} 1 & 1 & 2 & 3 & 2 \\ 3 & 4 & 5 & 1 & 3 \end{array} \right) \to \cdots \to \left(\begin{array}{cc|ccc} 1 & 0 & 3 & 11 & 5 \\ 0 & 1 & -1 & -8 & -3 \end{array} \right).$$

Portanto,

$$[T]_\alpha^\beta = \begin{pmatrix} 3 & 11 & 5 \\ -1 & -8 & -3 \end{pmatrix}$$

é a matriz de T em relação às bases dadas. ■

Exemplo 4.34 *Sejam* $\alpha = \{(1,1), (0,1)\}$ *e* $\beta = \{(0,3,0), (-1,0,0), (0,1,1)\}$ *bases* *de F^2 e F^3. Determine $T \in \mathcal{L}(F^2, F^3)$ tal que*

$$[T]_\alpha^\beta = \begin{pmatrix} 0 & 2 \\ -1 & 0 \\ -1 & 3 \end{pmatrix}.$$

Solução. Por definição, $T(1,1) = (1,-1,-1)$ e $T(0,1) = (0,9,3)$. Assim, para determinar $T(\mathbf{u})$ basta encontrar $[\mathbf{u}]_\alpha$ ou, equivalentemente, escalonar a matriz

$$\left(\begin{array}{cc|c} 1 & 0 & x \\ 1 & 1 & y \end{array}\right) \to \cdots \to \left(\begin{array}{cc|c} 1 & 0 & x \\ 0 & 1 & y-x \end{array}\right).$$

Portanto,

$$T(x,y) = x(1,-1,-1) + (y-x)(0,9,3) = (x, -10x+9y, -4x+3y),$$

que é o resultado desejado. ■

Exemplo 4.35 *Seja $T \in \mathcal{L}(F^2)$ tal que $T(\mathbf{e}_1) = (1,-3)$ e $T(\mathbf{e}_2) = (-2,1)$. Determine a base β de F^2 tal que*

$$[T]_\alpha^\beta = \begin{pmatrix} 1 & -1 \\ 4 & 3 \end{pmatrix},$$

com α a base canônica de F^2.

Solução. Seja $\beta = \{(a,b),(c,d)\}$ a base desejada. Então $T(\mathbf{e}_1) = 1(a,b) + 4(c,d)$ e $T(\mathbf{e}_2) = -1(a,b) + 3(c,d)$. Assim, obtemos os sistemas

$$\left\{\begin{array}{rcr} a+4c & = & 1 \\ -a+3c & = & -2 \end{array}\right. \quad \text{e} \quad \left\{\begin{array}{rcr} b+4d & = & -3 \\ -b+3d & = & 1 \end{array}\right.$$

Logo, basta escalonar a matriz

$$\left(\begin{array}{rr|rr} 1 & 4 & 1 & -3 \\ -1 & 3 & -2 & 1 \end{array}\right) \to \cdots \to \left(\begin{array}{rr|rr} 1 & 0 & \frac{11}{7} & -\frac{13}{7} \\ 0 & 1 & -\frac{1}{7} & -\frac{2}{7} \end{array}\right).$$

Portanto,

$$\beta = \left\{\frac{1}{7}(11,-13), \frac{1}{7}(-1,-2)\right\}.$$

é a base procurada. ■

Teorema 4.36 *Sejam $T \in \mathcal{L}(V, W)$, $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ e $\beta = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ bases de V e W. Então*

$$[T(\mathbf{u})]_\beta = [T]_\alpha^\beta [\mathbf{u}]_\alpha, \quad \forall \ \mathbf{u} \in V.$$

Prova. Visualize no diagrama comutativo (b) acima. Dado $\mathbf{u} \in V$, existem únicos $x_j \in F$ tais que $\mathbf{u} = \sum_{j=1}^n x_j \mathbf{u}_j$. Assim, pela equação (4.1),

$$T(\mathbf{u}) = \sum_{j=1}^n x_j T(\mathbf{u}_j) = \sum_{i=1}^m (\sum_{j=1}^n x_j a_{ij}) \mathbf{w}_i.$$

Portanto, $[T(\mathbf{u})]_\beta = [T]_\alpha^\beta [\mathbf{u}]_\alpha$, para todo $\mathbf{u} \in V$. ■

Exemplo 4.37 *Sejam α a base canônica de F^2, $\beta = \{(1,0,1), (-2,0,1), (0,1,0)\}$ uma base de F^3 e $T \in \mathcal{L}(F^2, F^3)$ tal que*

$$[T]_\beta^\alpha = \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ -2 & 3 \end{pmatrix}.$$

Determine $T(a, b)$.

Solução. Aplicando o Teorema 4.36, obtemos $T(a, b) = (a - 3b, -2a + 3b, a)$. ■

Exemplo 4.38 *Seja $\mathbf{A} \in F^{m \times n}$. Mostre que $\rho(\mathbf{A}) = \rho(\mathbf{A}^t \mathbf{A}) = \rho(\mathbf{A}\mathbf{A}^t)$.*

Solução. Sejam $\mathbf{B} = \mathbf{A}^t \mathbf{A}$, $T_\mathbf{A} : F^{n \times 1} \to F^{m \times 1}$ e $T_\mathbf{B} : F^{n \times 1} \to F^{n \times 1}$. Então, pelo Teorema 4.21, $\rho(\mathbf{A}) = n - \dim \ker T_A$ e $\rho(\mathbf{B}) = n - \dim \ker T_B$. Assim, é suficiente provar que $\ker T_A = \ker T_B$. Se $\mathbf{X} \in \ker T_A$, então $\mathbf{AX} = \mathbf{0}$ implica que $\mathbf{BX} = \mathbf{A}^t(\mathbf{AX}) = \mathbf{0}$. Logo, $\ker T_A \subseteq \ker T_B$. Por outro lado, se $\mathbf{X} \in \ker T_B$, então $\mathbf{BX} = \mathbf{0}$ ou $\mathbf{A}^t \mathbf{AX} = \mathbf{0}$ e, pelo item (1) do Lema 2.13, $(\mathbf{AX})^t(\mathbf{AX}) = 0$ implica que $\mathbf{AX} = \mathbf{0}$ e $\mathbf{X} \in \ker T_A$. Portanto, $\ker T_B \subseteq \ker T_A$. ■

Já vimos no Teorema 2.21 um método alternativo de resolver um sistema não homogêneo usando uma matriz adequada. Algo semelhante a esta mesma técnica pode ser utilizada para obter simultaneamente bases para o núcleo e a imagem de $T \in \mathcal{L}(F^n, F^m)$ ou $T \in \mathcal{L}(V, W)$. Para isto vamos introduz alguns conceitos, sejam $\mathbf{A} = (a_{ij}) \in F^{n \times m}$ e $\mathbf{B} = (b_{ij}) \in F^{n \times n}$. Diremos que a matriz

$$\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \end{array} \right) \in F^{n \times (m+n)}$$

é T-*associada* se

$$T(b_{i1}, \ldots, b_{in}) = (a_{i1}, \ldots, a_{im}), \ \ i = 1, \ldots, n.$$

Diremos que a matriz $\left(\begin{array}{c|c} \mathbf{R} & \mathbf{S} \end{array} \right)$ é *reduzida por linha à T-forma em escada* da matriz T-associada $\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \end{array} \right)$ se $\mathbf{R}$ for reduzida por linha à forma em escada de $\mathbf{A}$. Observe que este conceito surge naturalmente: se $\mathbf{u}_i = (b_{i1}, \ldots, b_{in}) \in F^n$ e $\mathbf{w}_j = \mathbf{e}_j \in F^m$, então, pela equação (4.1), obtemos a matriz

$$\left(\begin{array}{c|c} T(\mathbf{u}_i) & \mathbf{u}_i \end{array} \right) = \left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \end{array} \right)$$

T-associada, com $\mathbf{A} = [T]^t$. Note que a matriz reduzida por linha à T-forma em escada $\left(\begin{array}{c|c} \mathbf{R} & \mathbf{S} \end{array} \right)$ também é T-associada, pois as linhas de $\mathbf{S}$ são combinações lineares das linhas de $\mathbf{B}$, as quais formam uma base de F^n. Por exemplo, sejam $\mathbf{u}_1 = (1, 1, 1, 1)$, $\mathbf{u}_2 = (1, 1, 1, 0)$, $\mathbf{u}_3 = (1, 1, 0, 0)$ e $\mathbf{u}_4 = (1, 0, 0, 0)$ uma base de F^4 e $T \in \mathcal{L}(F^4, F^3)$ definida como

$$T(x, y, z, t) = (x - y + z + t, x + 2z - t, x + y + 3z - 3t).$$

Então a matriz

$$\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \end{array} \right) = \left(\begin{array}{ccc|cccc} 2 & 2 & 2 & 1 & 1 & 1 & 1 \\ 1 & 3 & 5 & 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \end{array} \right)$$

é T-associada, pois $T(\mathbf{u}_1) = (2, 2, 2)$ etc. Neste caso,

$$\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \end{array} \right) \to \cdots \to \left(\begin{array}{c|c} \mathbf{R} & \mathbf{S} \end{array} \right) = \left(\begin{array}{ccc|cccc} 1 & 0 & -1 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{3}{4} \\ 0 & 1 & 2 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ 0 & 0 & 0 & \frac{3}{4} & \frac{3}{4} & -\frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{array} \right).$$

Note que

$$\operatorname{Im} T = F[(2, 2, 2), (1, 3, 5)] \ \text{ e } \ \ker T = F[(3, 3, -1, 1), (1, -1, -1, -1)].$$

Voltando ao caso geral. É claro que $\operatorname{Im} T = F[T(\mathbf{u}_1), \ldots, T(\mathbf{u}_n)]$. Então, pelo Teorema 3.36, dentre esses vetores, podemos extrair uma base de $\operatorname{Im} T$, digamos as linhas não nula de $\mathbf{R}$ ou $\{T(\mathbf{u}_1), \ldots, T(\mathbf{u}_k)\}$, com $k \leq \min\{m, n\}$, a qual é parte de uma base de F^m, a saber, $\beta = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$. Portanto, pelo Corolário 4.22, podemos escolher uma base $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}_1, \ldots, \mathbf{v}_{n-k}\}$ de F^n tal que $T(\mathbf{u}_i) = \mathbf{w}_i$, com $i = 1, \ldots, k$, e $T(\mathbf{v}_j) = \mathbf{0}$, com $j = 1, \ldots, n-k$. Este processo é dado pelo esquema:

$$\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \end{array} \right) \to \cdots \to \left(\begin{array}{c|c} \mathbf{R} & \mathbf{S} \end{array} \right)$$

Vamos resumir isto no próximo teorema:

Teorema 4.39 *Sejam $T \in \mathcal{L}(F^n, F^m)$ e*

$$\left(\begin{array}{c|c} [T]^t & \mathbf{I}_n \end{array} \right) \to \cdots \to \left(\begin{array}{c|c} \mathbf{R} & \mathbf{S} \end{array} \right).$$

Então $\{\mathbf{r}_1, \ldots, \mathbf{r}_k\}$, com $k \leq \min\{m, n\}$, é uma base de $\operatorname{Im} T$ e $\{\mathbf{s}_{k+1}, \ldots, \mathbf{s}_n\}$ é uma base de $\ker T$.

Prova. Fica como um exercício. ■

Já observamos que $\dim \operatorname{Im} T$ é o posto coluna da matriz $[T]$. Mas, o Teorema 4.39 afirma que $\dim \operatorname{Im} T$ é o posto linha da matriz $[T]$. Portanto, isto justifica mais uma vez que o posto linha é igual ao posto coluna.

Exemplo 4.40 *$T \in \mathcal{L}(F^{2\times 2})$ definido como $T(\mathbf{A}) = \mathbf{BA}$, com*

$$\mathbf{B} = \begin{pmatrix} 2 & -2 \\ -1 & 1 \end{pmatrix} = 2\mathbf{E}_{11} - 2\mathbf{E}_{12} - \mathbf{E}_{21} + \mathbf{E}_{22}.$$

Determine bases para o núcleo e a imagem de T.

Solução. Primeiro note que a função $\sigma : F^{2\times 2} \to F^4$ definida como $\sigma(a_{ij}) = (a_{11}, a_{12}, a_{21}, a_{22})$ é um isomorfismo. Portanto, podemos identificar $F^{2\times 2}$ com F^4. Assim, pelo Teorema 4.39 e $T(\mathbf{E}_{ij}) = \mathbf{BE_{ij}} = b_{1i}\mathbf{E}_{1j} + b_{2i}\mathbf{E}_{2j}$,

$$\left(\begin{array}{c|c} [T]^t & \mathbf{I}_4 \end{array} \right) \to \cdots \to \left(\begin{array}{cccc|cccc} 1 & 0 & -\frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 1 & 0 & -\frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{array} \right).$$

Portanto,

$$\left\{ \begin{pmatrix} 1 & 0 \\ -\frac{1}{2} & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & -\frac{1}{2} \end{pmatrix} \right\} \text{ e } \left\{ \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \right\}$$

são bases de $\operatorname{Im} T$ e $\ker T$. ■

Teorema 4.41 *Sejam $R, S \in \mathcal{L}(U, V)$ e $T \in \mathcal{L}(V, W)$, α, β e γ bases de U, V e W. Então*

$$[R+S]_\alpha^\beta = [R]_\alpha^\beta + [S]_\alpha^\beta, \quad [aR]_\alpha^\beta = a[R]_\alpha^\beta, \ \forall \ a \in F \ \text{ e } \ [(T \circ S)]_\alpha^\gamma = [T]_\beta^\gamma [S]_\alpha^\beta.$$

Prova. Dado $\mathbf{u} \in U$ temos, pelo Teorema 4.36, que

$$[(T \circ S)(\mathbf{u})]_\gamma = [T \circ S]_\alpha^\gamma [\mathbf{u}]_\alpha \ \text{ e } \ [T(S(\mathbf{u}))]_\gamma = [T]_\beta^\gamma [S(\mathbf{u})]_\beta$$

Como $(T \circ S)(\mathbf{u}) = T(S(\mathbf{u}))$ e $[S(\mathbf{u})]_\beta = [S]_\alpha^\beta [\mathbf{u}]_\alpha$ temos que

$$[(T \circ S)]_\alpha^\gamma [\mathbf{u}]_\alpha = [T]_\beta^\gamma [S]_\alpha^\beta [\mathbf{u}]_\alpha.$$

Portanto, pela unicidade das coordenadas, $[(T \circ S)]_\alpha^\gamma = [T]_\beta^\gamma [S]_\alpha^\beta$. ■

Corolário 4.42 *Sejam $T \in \mathcal{L}(V, W)$ e α, β bases de V e W.*

1. *Se T for um isomorfismo, então $[T^{-1}]_\beta^\alpha = ([T]_\alpha^\beta)^{-1}$.*
2. *T é um isomorfismo se, e somente se, $\det([T]_\alpha^\beta) \neq 0$.*
3. *T é singular se, e somente se, $\det(\mathbf{A}) = 0$, da matriz $\mathbf{A} = (^\beta_\alpha)$.*

Prova. Vamos provar apenas o item (1). Pelo Teorema 4.21, T^{-1} existe e $T^{-1} \in \mathcal{L}(W, V)$. Assim, $T^{-1} \circ T = I_V$ e $T \circ T^{-1} = I_W$. Logo, pelo Teorema 4.41,

$$\mathbf{I} = [I_V]_\alpha^\alpha = [T^{-1}]_\beta^\alpha [T]_\alpha^\beta \ \text{ e } \ \mathbf{I} = [I_W]_\beta^\beta = [T]_\alpha^\beta [T^{-1}]_\beta^\alpha.$$

Portanto, pela unicidade da inversa, $[T^{-1}]_\beta^\alpha = ([T]_\alpha^\beta)^{-1}$. ■

Exemplo 4.43 *Seja $p_k(x) \in \mathbb{P}_n(F)$, com $\partial(p_k) = k$, para $k = 0, 1, , \ldots, n$. Mostre que $T \in \mathcal{L}(\mathbb{P}_n(F))$ definido como $T(x^k) = p_k(x)$ é um isomorfismo. Conclua que $\{p_0(x), \ldots, p_n(x)\}$ é um base de $\mathbb{P}_n(F)$.*

Solução. Como $\partial(p_k) = k$ temos que o coeficiente líder de $p_k(x)$, $a_k \neq 0$. Sendo

$$[T] = \begin{pmatrix} a_0 & \star & \cdots & \star \\ 0 & a_1 & \cdots & \star \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}$$

uma matriz triangular superior, obtemos $\det([T]) = a_0 a_1 \cdots a_n \neq 0$. Portanto, T é um isomorfismo. ■

Exemplo 4.44 *Determine todos os isomorfismos sobre F^2.*

Solução. Seja $T \in \mathcal{L}(F^2)$. Então $T(\mathbf{e}_1) = (a, b)$ e $T(\mathbf{e}_2) = (c, d)$ completamente determina T, de modo que

$$[T] = \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

Assim, pelo item (2) do Corolário 4.42, T é um isomorfismo sobre F^2 se, e somente se, $D = ad - bc \neq 0$. Note, pelo Exemplo 1.16, que

$$[T] = \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & a^{-1}D \end{pmatrix} \begin{pmatrix} 1 & a^{-1}b \\ 0 & 1 \end{pmatrix}$$

quando $a \neq 0$. ■

É muito importante observar o seguinte: se $\mathbf{A} \in F^{m\times m}$ é não singular e $\mathbf{B}$ é qualquer elemento de $F^{m\times n}$, então $\mathbf{A}^{-1}\mathbf{B}$ é obtida via o esquema:

$$\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \end{array} \right) \to \cdots \to \left(\begin{array}{c|c} \mathbf{I} & \mathbf{A}^{-1}\mathbf{B} \end{array} \right).$$

Além disso, sejam $I_V \in \mathcal{L}(V)$, $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ e $\beta = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ bases de V. Então $[I_V]_\alpha^\beta = [I]_\beta^\alpha$, pois $\mathbf{u}_i = I_V(\mathbf{u}_i)$ implica que

$$[I_V]_\alpha^\beta = \begin{pmatrix} [I_V(\mathbf{u}_1)]_\beta & \cdots & [I_V(\mathbf{u}_n)]_\beta \end{pmatrix} = \begin{pmatrix} [\mathbf{u}_1]_\beta & \cdots & [\mathbf{u}_n]_\beta \end{pmatrix} = [I]_\beta^\alpha.$$

Sejam $T \in \mathcal{L}(V)$ e α, β bases de V. Qual a relação entre $[T]_\alpha^\alpha$ e $[T]_\beta^\beta$? Para responder esta questão, vamos considerar o diagrama comutativo (a):

$$\begin{tikzcd}
(\alpha)\ V \arrow[r, "T"] & V\ (\alpha) \arrow[d, "I_V"] \\
(\beta)\ V \arrow[u, "I_V"] \arrow[r, dashed, "S"'] & V\ (\beta)
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
F^n \arrow[r, "L_\alpha"] \arrow[d, "\mathbf{A}"'] & V \arrow[d, "T"] & F^n \arrow[l, "L_\beta"'] \arrow[d, "\mathbf{B}"] \\
F^m \arrow[r, "L_\gamma"'] & W & F^m \arrow[l, "L_\delta"]
\end{tikzcd}$$

$$(a) \qquad\qquad\qquad\qquad (b)$$

Assim, $S = I_V \circ T \circ I_V$ e, pelo Teorema 4.41, $[S]_\beta^\beta = [I_V]_\alpha^\beta [T]_\alpha^\alpha [I_V]_\beta^\alpha$. Logo, pondo $\mathbf{P} = [I_V]_\beta^\alpha = [I]_\alpha^\beta$, obtemos a *fórmula de mudança de bases* $[T]_\beta^\beta = [S]_\beta^\beta = \mathbf{P}^{-1}[T]_\alpha^\alpha \mathbf{P}$, isto é, as matrizes $[T]_\alpha^\alpha$ e $[T]_\beta^\beta$ são semelhantes ou conjugadas. Neste caso,

$$\det([T]_\beta^\beta) = \det([T]_\alpha^\alpha) \ \text{ e } \ \operatorname{tr}([T]_\beta^\beta) = \operatorname{tr}([T]_\alpha^\alpha).$$

Portanto, o *determinante* (*o traço*) de T é definido como o determinante (o traço) de qualquer representação matricial de T em relação a alguma base de V e denotamos por $\det(T)$ e $\operatorname{tr}(T)$. Já vimos que um modo de determinar o operador não singular $P \in \mathcal{L}(V)$ era dado pelo esquema:

$$\left(\begin{array}{c|c} \alpha & \beta \end{array} \right) \to \cdots \to \left(\begin{array}{c|c} \mathbf{I} & \mathbf{P} \end{array} \right).$$

Exemplo 4.45 *Sejam α a base canônica, $\beta = \{(1,2),(1,-1)\}$ outra base de F^2 e $T \in \mathcal{L}(F^2)$ definido como $T(x,y) = (x+2y, y)$. Determine $\mathbf{P} \in F^{2\times 2}$ não singular tal que $[T]_\beta^\beta = \mathbf{P}^{-1}[T]_\alpha^\alpha \mathbf{P}$.*

Solução. Pelo exposto,

$$\mathbf{P} = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}.$$

Assim, para determinar $\mathbf{P}^{-1}$, basta escalonar a matriz

$$\left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 2 & -1 & 0 & 1 \end{array} \right) \to \cdots \to \left(\begin{array}{cc|cc} 1 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & \frac{2}{3} & -\frac{1}{3} \end{array} \right).$$

Portanto,

$$\mathbf{P}^{-1}[T]_\alpha^\alpha \mathbf{P} = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 7 & -2 \\ 8 & -1 \end{pmatrix} = [T]_\beta^\beta$$

e $P(x, y) = (x + y, 2x - y)$. ■

Seja $T \in \mathcal{L}(V, W)$. Então, pelo Corolário 4.22, podemos escolher uma base $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ de V tal que $\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ seja uma base de $\ker T$, de modo que $\{T(\mathbf{u}_1), \ldots, T(\mathbf{u}_k)\}$ é uma base de $\operatorname{Im} T$, a qual pode ser estendida a uma base $\beta = \{T(\mathbf{u}_1), \ldots, T(\mathbf{u}_k), \mathbf{w}_{k+1} \ldots, \mathbf{w}_m\}$ de W. Portanto,

$$[T]_\alpha^\beta = \mathbf{I}_k \oplus \mathbf{O} = \begin{pmatrix} \mathbf{I}_k & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix} = \mathbf{N}_k, \quad k \leq \min\{m, n\}.$$

Em particular, se $V = W$, então existe um $S \in \mathcal{L}(V)$ definido como $S(T(\mathbf{u}_i)) = \mathbf{u}_i$, com $i = 1, \ldots, k$, e $S(\mathbf{w}_j) = \mathbf{u}_j$, com $j = k + 1, \ldots, n$, tal que $T \circ S \circ T = T$, de modo que $[T]_\alpha^\beta = [T]_\alpha^\beta [S]_\beta^\alpha [T]_\alpha^\beta$. Em geral, temos o seguinte resultado:

Teorema 4.46 *Sejam α e γ bases fixadas de V e W. Seja $T \in \mathcal{L}(V, W)$ tal que $\mathbf{A} = [T]_\alpha^\gamma$. Então existem β e δ bases de V e W tais que $\mathbf{B} = [T]_\beta^\delta$ se, e somente se, existem $\mathbf{P} \in F^{n \times n}$ e $\mathbf{Q} \in F^{m \times m}$ não singulares tais que $\mathbf{B} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{P}$.*

Prova. Pelo diagrama comutativo (b) acima e o Teorema 4.41, obtemos

$$\mathbf{A} = [T]_\alpha^\gamma = [L_\gamma^{-1} \circ T \circ L_\alpha] \text{ e } \mathbf{B} = [T]_\beta^\delta = [L_\delta^{-1} \circ T \circ L_\beta].$$

Assim, existem $P = L_\alpha^{-1} \circ L_\beta \in \mathcal{L}(F^n)$ e $Q = L_\gamma^{-1} \circ L_\delta \in \mathcal{L}(F^m)$ tais que $\mathbf{P} = [P] = [I]_\alpha^\beta$ e $\mathbf{Q} = [Q] = [I]_\gamma^\delta$ ou $\mathbf{Q}^{-1} = [I]_\delta^\gamma$, de modo que

$$\mathbf{Q}^{-1}\mathbf{A}\mathbf{P} = [L_\delta^{-1} \circ L_\gamma][L_\gamma^{-1} \circ T \circ L_\alpha][L_\alpha^{-1} \circ L_\beta] = [L_\delta^{-1} \circ T \circ L_\beta] = \mathbf{B}.$$

Reciprocamente, sejam $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}, \gamma = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$, $\mathbf{P} = (p_{ij}) \in F^{n \times n}$ e $\mathbf{Q} = (q_{ij}) \in F^{m \times m}$. Então, pela a equação (3.5), os vetores

$$\mathbf{v}_i = \sum_{j=1}^{n} p_{ji}\mathbf{u}_j \text{ e } \mathbf{z}_i = \sum_{j=1}^{m} q_{ji}\mathbf{w}_j$$

possuem as propriedades desejadas, por exemplo, $T_\mathbf{P} : V \to V$ definido como $\mathbf{v}_i = T_\mathbf{P}(\mathbf{u}_i)$ é um isomorfismo. Seja $T \in \mathcal{L}(V, W)$ tal que $\mathbf{A} = [T]_\alpha^\gamma$. Então existem β e δ bases de V e W tais que $\mathbf{B} = [T]_\beta^\delta$. ■

É muito importante lembrar que um modo alternativo de determinar as matrizes não singulares $\mathbf{P} \in F^{n\times n}$ e $\mathbf{Q} \in F^{m\times m}$ era dado pelo esquema:

$$\begin{array}{c|c} \mathbf{A} & \mathbf{I}_m \\ \hline \mathbf{I}_n & \mathbf{O} \end{array} \to \cdots \to \begin{array}{c|c} \mathbf{QA} & \mathbf{Q} \\ \hline \mathbf{I}_n & \mathbf{O} \end{array} \to \cdots \to \begin{array}{c|c} \mathbf{QAP} & \mathbf{Q} \\ \hline \mathbf{P} & \mathbf{O} \end{array}$$

Exemplo 4.47 *Sejam $\alpha = \{(1,0,-1),(0,2,0),(1,2,3)\}$ uma base de F^3, $\gamma = \{(-1,1),(2,0)\}$ uma base de F^2 e $T \in \mathcal{L}(F^3, F^2)$ é tal que*

$$[T]_\alpha^\gamma = \begin{pmatrix} 2 & -1 & 3 \\ 3 & 1 & 0 \end{pmatrix}.$$

Determine $[T]$.

Solução. Sejam β e δ as bases canônicas de F^3 e F^2. Então, pelo exposto, devemos determinar $\mathbf{P} = [I]_\alpha^\beta$ e $\mathbf{Q} = [I]_\gamma^\delta$ ou $\mathbf{Q}^{-1} = [I]_\delta^\gamma$. Observe que $\mathbf{Q}^{-1}$ é a matriz cujas colunas são os vetores da base γ. Assim, é suficiente escalonar a matriz

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 1 & 0 \\ -1 & 0 & 3 & 0 & 0 & 1 \end{array}\right) \to \cdots \to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{3}{4} & 0 & -\frac{1}{4} \\ 0 & 1 & 0 & -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \\ 0 & 0 & 1 & \frac{1}{4} & 0 & \frac{1}{4} \end{array}\right).$$

Portanto,

$$[T] = \begin{pmatrix} -1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & -1 & 3 \\ 3 & 1 & 0 \end{pmatrix} \frac{1}{4} \begin{pmatrix} 3 & 0 & -1 \\ -1 & 2 & -1 \\ 1 & 0 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 3 & 3 & -5 \\ 5 & -1 & 1 \end{pmatrix}$$

e $T(x,y,z) = \frac{1}{2}(3x+3y-5z, 5x-y+z)$. ■

Vamos finalizar esta seção com uma aplicação geométrica do determinante bastante usada no Cálculo Diferencial e Integral. Cada vetor fixado $\mathbf{w} = (a,b,c) \in \mathbb{R}^3$ induz um $T_\mathbf{w} \in \mathcal{L}(\mathbb{R}^3)$ definido como $T_\mathbf{w}(\mathbf{x}) = \mathbf{w} \times \mathbf{x}$ (*produto vetorial*). Como $T_\mathbf{w}(\mathbf{e}_1) = (0,c,-b)$, $T_\mathbf{w}(\mathbf{e}_2) = (-c,0,a)$ e $T_\mathbf{w}(\mathbf{e}_3) = (b,-a,0)$ temos que

$$[T_\mathbf{w}] = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix} \text{ e } (\mathbf{w} \times \mathbf{x})^t = [T_\mathbf{w}]\mathbf{x}^t.$$

Então $\operatorname{Im} T_\mathbf{w} = F[(0,c,-b),(-c,0,a)]$ e $\ker T_\mathbf{w} = F[\mathbf{w}]$, com $a^2+b^2+c^2 \neq 0$. É bem conhecido que se $\mathbf{u}_1 = (a_1,b_1,c_1)$, $\mathbf{u}_2 = (a_2,b_2,c_2)$ e $\mathbf{u}_3 = (a_3,b_3,c_3)$ são *LI* em $\mathbb{R}^3$, então o volume do paralelepípedo determinado por eles é o valor absoluto do *produto misto*

$$V = |[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3]| = \left|\det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}\right|.$$

Em particular, sendo $T : \mathbb{R}^2 \to \mathbb{R}^3$ definida como $T(x, y) = (x, y, 0)$ linear e injetora, temos que a área do retângulo determinado pelos vetores $\mathbf{u}_1 = (a_1, b_1)$ e $\mathbf{u}_2 = (a_2, b_2)$ em $\mathbb{R}^2$ é o valor absoluto do produto vetorial

$$A = |\mathbf{u}_1 \times \mathbf{u}_2| = |[\mathbf{u}_1, \mathbf{u}_2, \mathbf{e}_3]| = \left|\det \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}\right|.$$

Sejam $\mathcal{R}$ um retângulo (ou uma região que possa ser dividida em retângulos) em $\mathbb{R}^2$, determinado pelos vetores $\mathbf{u}$ e $\mathbf{v}$, ou seja, $\mathcal{R} = \{r\mathbf{u} + s\mathbf{v} : r, s \in [0, 1]\}$ e $T \in \mathcal{L}(\mathbb{R}^2)$ não singular. Qual a relação entre as áreas de $\mathcal{R}$ e $T(\mathcal{R})$? Para responder esta questão, note, pelo Exemplo 4.44, que $T(\mathbf{e}_1) = (a, b)$ e $T(\mathbf{e}_2) = (c, d)$, com $D = ad - bc \neq 0$. Como, pelo Teorema 4.41, $[T(\mathbf{x})] = [T][\mathbf{x}]$ temos que

$$|T(\mathbf{u}) \times T(\mathbf{v})| = |\det[T]||\mathbf{u} \times \mathbf{v}| = |D| \cdot |\mathbf{u} \times \mathbf{v}|.$$

Exercícios

1. Seja $T \in \mathcal{L}(\mathbb{P}_2(F))$ definido como $T(a + bx + cx^2) = b + ax + cx^2$. Determine $[T]$.

2. Para cada um dos $T \in \mathcal{L}(F^n, F^m)$ abaixo, determine bases para o núcleo e a imagem:

   (a) $T(x, y) = (2x - y, 0)$.
   (b) $T(x, y, z) = (x + 2y, y - z, x + 2z)$.
   (c) $T(x, y) = (x + y, x + y)$.
   (d) $T(x, y, z) = (x + y, y + z)$.
   (e) $T(x, y, z) = (x + z, x - z, y)$.
   (f) $T(x, y, z) = (x + 2z, z)$.

3. Seja $D \in \mathcal{L}(\mathbb{P}_n(F))$ definido como $D(p(x)) = p'(x)$. Determine $[D]$.

4. Seja $T \in \mathcal{L}(F^{2\times 2})$ definido como $T(\mathbf{A}) = \mathbf{AB} - \mathbf{BA}$, onde
$$\mathbf{B} = \mathbf{E}_{11} - \mathbf{E}_{12} - 2\mathbf{E}_{21} + 2\mathbf{E}_{22}.$$
Determine bases para o núcleo e a imagem de T. Generalize para qualquer $\mathbf{B}$, com $b_{11} \neq b_{22}$ e $b_{11} = b_{22}$.

5. Seja $\mathbf{B} \in F^{n\times n}$ nilpotente. Mostre que $\mathbf{BA} = \mathbf{A}$, para algum $\mathbf{A} \in F^{n\times n}$.

6. Seja $T \in \mathcal{L}(\mathbb{P}_2(F), \mathbb{P}_3(F))$ definido como $T(p(x)) = p(x) + x^2 p'(x)$. Determine bases para o núcleo e a imagem de T.

7. Dentre as transformações dos Exercícios de $(1.$ à $4.)$, determine as que são isomorfismos e, para essas, encontre uma regra que defina a inversa.

8. Sejam $S \in \mathcal{L}(F^2, F^3)$ e $S \in \mathcal{L}(F^3, F^2)$ definidas como
$$S(x, y) = (x - y, 3x, y) \text{ e } T(x, y, z) = (2x - y - z, x + y).$$
Determine a representação matricial de $S, T, S \circ T$ e $T \circ S$ com respeito às bases $\alpha = \{(1, 0), (1, 1)\}$ de F^2 e $\beta = \{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}$ de F^3.

9. Seja $T \in \mathcal{L}(\mathbb{P}_2(F), F^3)$ definido como $T(p(x)) = (p(c_1), p(c_2), p(c_3))$, para todos $c_1, c_2, c_3 \in F$ distintos. Mostre que T é um isomorfismo. Generalize.

10. Sejam $c \in F$ e $T \in \mathcal{L}(\mathbb{P}_n(F))$ definido como $T(p(x)) = p(c + x)$. Mostre que T é um isomorfismo. Conclua que $\{1, c + x, \ldots, (c + x)^n\}$ é uma base de $\mathbb{P}_n(F)$. É verdade que $\{1, \frac{1}{1!}(x - c), \ldots, \frac{1}{n!}(x - c)^n\}$ é uma base de $\mathbb{P}_n(F)$?

11. (**Teorema da Alternativa de Fredholm**[2]) Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Então uma e apenas uma das condições ocorre:

    (a) A equação $T(\mathbf{x}) = \mathbf{b}$ possui uma solução, para todo $\mathbf{b} \in V$.

    (b) $\ker T \neq \{\mathbf{0}\}$.

12. Seja $T \in \mathcal{L}(F^2)$ tal que $[T] = -\mathbf{E}_{11} - 2\mathbf{E}_{12} + \mathbf{E}_{22}$.

    (a) Encontre, se possível, vetores $\mathbf{u}$ e $\mathbf{v}$, tais que $T(\mathbf{u}) = \mathbf{u}$ e $T(\mathbf{v}) = -\mathbf{v}$.

    (b) Determine uma base e a dimensão do núcleo e da imagem de T.

[2]Erik Ivar Fredholm, 1866-1927, matemático sueco.

(c) T é um isomorfismo? Caso afirmativo, determine uma matriz que represente T^{-1}, encontrando, também, $T^{-1}(x, y)$.

13. Seja $T \in \mathcal{L}(\mathbb{R}^2)$ tal que $[T] = 3^{-1}(-\mathbf{E}_{11} + 2\mathbf{E}_{12} + 4\mathbf{E}_{21} + \mathbf{E}_{22})$. Determine $[T]_\beta^\beta$, com $\beta = \{(1, 2), (-1, 1)\}$. Qual o significado geométrico de T?

14. Seja V o espaço vetorial das funções $q : \mathbb{R}^2 \to \mathbb{R}$ definida como

$$q(x, y) = ax^2 + bxy + cy^2 + dx + ey + f.$$

Mostre que $T : V \to V$ definido como

$$T(q) = \frac{\partial}{\partial x}\int q(x, y)dy + \frac{\partial}{\partial y}\int q(x, y)dx$$

é linear. Determine $[T]_\alpha^\alpha$, com $\alpha = \{x^2, xy, y^2, x, y, 1\}$.

15. Seja $V = U \oplus W$, com $\dim U = k$ e $\dim W = n - k$. Mostre que as seguintes condições são equivalentes:

(a) E é uma projeção de V sobre U paralela a W;

(b) $V = \operatorname{Im} E \oplus \ker E$ e $\operatorname{Im} E = \{\mathbf{v} \in V : E(\mathbf{v}) = \mathbf{v}\}$;

(c) Existe uma base de V tal que $[E] = \mathbf{I}_k \oplus \mathbf{O}$;

(d) $E^2 = E$.

16. Seja $E \in \mathcal{L}(\mathbb{R}^2)$ a projeção sobre $U = \mathbb{R}[(1, -1)]$ paralela a $W = \mathbb{R}[(1, 2)]$. Determine $E(x, y)$.

17. Seja $E \in \mathcal{L}(\mathbb{R}^3)$ a projeção sobre $\pi = \{(x, y, z) \in \mathbb{R}^3 : 3x + 2y + z = 0\}$.

(a) Determine $E(x, y, z)$.

(b) Determine uma base β de $\mathbb{R}^3$ tal que $[E]_\beta^\beta = \mathbf{I}_2 \oplus \mathbf{O}$.

18. Seja $V = U \oplus W$, com $\dim U = k$ e $\dim W = n - k$. Mostre que as seguintes condições são equivalentes:

(a) R é a reflexão de V em U paralela a W, isto é, $R(\mathbf{u} + \mathbf{w}) = \mathbf{u} - \mathbf{w}$;

(b) $V = \ker(R - I) \oplus \ker(R + I)$;

(c) Existe uma base de V tal que $[R] = \mathbf{I}_k \oplus -\mathbf{I}_{n-k}$;

(d) $R^2 = I$.

19. Seja $R \in \mathcal{L}(\mathbb{R}^3)$ a reflexão em $\pi = \{(x, y, z) \in \mathbb{R}^3 : 3x + 2y + z = 0\}$.

    (a) Determine $R(x, y, z)$.

    (b) Determine uma base β de $\mathbb{R}^3$ tal que $[R]_\beta^\beta = \mathbf{I}_2 \oplus -\mathbf{I}_1$.

20. Determine a rotação de um ângulo θ em torno do eixo dos z. Generalize para uma reta que passa pela origem e possui a direção do vetor $(a, b, c) \in \mathbb{R}^3$, com $a^2 + b^2 + c^2 = 1$.

21. Seja $T_i \in \mathcal{L}(F^3)$, $i = 1, 2, 3$, tal que

$$[T_1] = \begin{pmatrix} 2 & 1 & 2 \\ 1 & 2 & 2 \\ 2 & 2 & 3 \end{pmatrix}, \quad [T_2] = \begin{pmatrix} 1 & -2 & 2 \\ 2 & -1 & 2 \\ 2 & -2 & 3 \end{pmatrix} \text{ e } [T_3] = \begin{pmatrix} -1 & 2 & 2 \\ -2 & 1 & 2 \\ -2 & 2 & 3 \end{pmatrix}.$$

    Mostre que cada T_i preserva ternos $(a, b, c) \in \mathbb{Z}$ tais que $a^2 + b^2 = c^2$, ou seja, preserva triplas de Pitágoras.

22. Geometricamente, um *cisalhamento* é um elemento de $T \in \mathcal{L}(\mathbb{R}^2)$ que satisfaz as seguintes condições: $T(r) \subseteq r$, para toda reta r em $\mathbb{R}^2$ contendo a origem; a reta determinada por P e $T(P)$ é paralela a r, para todo $P \notin r$. Mostre que as seguintes condições são equivalentes:

    (a) T é um cisalhamento;

    (b) Existe uma base β de $\mathbb{R}^2$ e um $a \in \mathbb{R}$ tal que $[T]_\beta^\beta = \mathbf{I} + a\mathbf{E}_{12}$;

    (c) $(T - I)^k = 0$, para algum $k \in \mathbb{N}$.

23. Seja $\mathbf{A} \in F^{n\times n}$. Mostre que $\mathbf{AP} = \mathbf{PA}$, para todo $\mathbf{P} \in F^{n\times n}$ não singular, se, e somente se, $\mathbf{A} = a\mathbf{I}_n$, para algum $a \in F$.

24. Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Mostre que $[T]_\alpha^\alpha = [T]_\beta^\beta$, para todas as bases α e β de V se, e somente se, $T = aI$, para algum $a \in F$.

25. Seja $f \in \mathcal{L}(F^{n\times n}, F)$ tal que $f(\mathbf{AB}) = f(\mathbf{BA})$, para todos $\mathbf{A}, \mathbf{B} \in F^{n\times n}$. Mostre que $f = c\,\mathrm{tr}$, para algum $c \in F$. Conclua que $\mathbf{X} \in F^{n\times n}$ pode ser escrita como $\mathbf{X} = \mathbf{AB} - \mathbf{BA}$ se, e somente se, $\mathrm{tr}(\mathbf{X}) = 0$.

26. Seja V um espaço vetorial sobre F, com $\dim V = n$.

(a) Mostre que $ST - TS \neq I$, para todos $S, T \in \mathcal{L}(V)$.

(b) Mostre, com um exemplo, que a afirmação (a) não é necessariamente verdade se $\dim V$ for infinita.

27. Seja $T : \mathbb{R} \to \mathbb{R}$ uma função aditiva. Mostre que $T \in \mathcal{L}(\mathbb{R})$ se, e somente se, T for contínua.

28. Sejam V um espaço vetorial sobre $\mathbb{R}$ e $J \in \mathcal{L}(V)$ tal que $J^2 = -I$ e chama-se uma *estrutura complexa* de V. Por exemplo, $J(x, y) = (-y, x)$ e $J(x, y) = (y, -x)$. Vamos estender a ação de $\mathbb{R}$ sobre V para uma ação de $\mathbb{C}$ sobre V como: para todo $a + bi \in \mathbb{C}$ e $\mathbf{u} \in V$, $(a + bi) \cdot \mathbf{u} = a\mathbf{u} - bJ(\mathbf{u})$.

   (a) Mostre que V é um espaço vetorial sobre $\mathbb{C}$. Conclua que $\mathbb{R}^3$ não possui uma estrutura complexa.

   (b) Mostre que se $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ for um conjunto LI de V sobre os complexos, então $\{\mathbf{v}_1, \ldots, \mathbf{v}_n, J(\mathbf{v}_1), \ldots, J(\mathbf{v}_n)\}$ é um conjunto LI de V sobre os reais. Conclua, no caso finito, que $\dim_{\mathbb{R}} V = 2 \dim_{\mathbb{C}} V$.

   (c) Mostre que $\mathbf{A} \in \mathbb{R}^{2n \times 2n}$ é tal que $\mathbf{A}^2 = -\mathbf{I}$ se, e somente se, $\mathbf{A}$ é semelhante a uma matriz sob a forma

$$\begin{pmatrix} \mathbf{O} & -\mathbf{I}_n \\ \mathbf{I}_n & \mathbf{O} \end{pmatrix}.$$

4.4 Dualidade

Sejam V e W espaços vetoriais sobre F. Então, pelo Exemplo 3.2, $\mathcal{L}(V, W)$ é um espaço vetorial sobre F. É muito importante, de um ponto de vista didático e teórico, apresentar uma prova direta de que $\dim \mathcal{L}(V, W) = mn$, quando $\dim V = n$ e $\dim W = m$. Para isto, sejam $\alpha = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ uma base de V e $\beta = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ uma base de W. Então, pelo Teorema 4.11, existem únicos elementos $E_{ij} \in \mathcal{L}(V, W)$, com $i = 1, \ldots, n$ e $j = 1, \ldots, m$, definidos como

$$E_{ij}(\mathbf{v}_k) = \delta_{jk}\mathbf{w}_i = \begin{cases} \mathbf{w}_i, & \text{se } j = k \\ \mathbf{0}, & \text{se } j \neq k, \end{cases} \quad k = 1, \ldots, n.$$

Então os E_{ij} agem sobre um vetor $\mathbf{v} = \sum_{k=1}^{n} x_k \mathbf{v}_k$ como

$$E_{ij}(\mathbf{v}) = x_k \sum_{k=1}^{n} E_{ij}(\mathbf{v}_k) = x_k \sum_{k=1}^{n} \delta_{jk}\mathbf{w}_i = x_j \mathbf{w}_i.$$

Afirmação. $\gamma = \{E_{ij} : i = 1, \ldots, n; j = 1, \ldots, m\}$ é uma base de $\mathcal{L}(V, W)$.
De fato. Dado $T \in \mathcal{L}(V, W)$ tal que $T(\mathbf{v}_k) = \sum_{i=1}^{m} a_{ik}\mathbf{w}_i$, $k = 1, \ldots, n$. Assim, pondo

$$S = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} E_{ij} \in \mathcal{L}(V, W),$$

obtemos

$$S(\mathbf{v}_k) = (\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} E_{ij})(\mathbf{v}_k) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}\delta_{jk}\mathbf{w}_i = \sum_{i=1}^{m} a_{ik}\mathbf{w}_i = T(\mathbf{v}_k).$$

Portanto, $T = S$ e $\mathcal{L}(V, W) = F[\gamma]$. Se $S = 0$, então $S(\mathbf{u}_k) = 0$, para todo $k = 1, \ldots, n$. Assim $\sum_{i=1}^{m} a_{ik}\mathbf{v}_i = \mathbf{0}$, para todo $k = 1, \ldots, n$. Como β é LI temos que $a_{ik} = 0$, para todo $i = 1, \ldots, m$ e $k = 1, \ldots, n$. Portanto, γ é LI. Consequentemente, γ é uma base de $\mathcal{L}(V, W)$ e $\dim \mathcal{L}(V, W) = mn$. Neste caso,

$$[T]_{\alpha}^{\beta} = (a_{ij}) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}\mathbf{E}_{ij} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}[E_{ij}]_{\alpha}^{\beta} \in F^{m\times n}.$$

Este processo chama-se *extensão linear* e sempre produz um $T \in \mathcal{L}(V, W)$. Observe que $\delta = \{E_{11}, E_{21}, \ldots, E_{m1}, E_{12}, E_{22}, \ldots, E_{m2}, \ldots\}$ é uma base de $\mathcal{L}(V, W)$ em ordem antilexicográfica (colunas de $[T]_{\alpha}^{\beta}$).

Um caso particular de grande importância destas considerações é quando W é o corpo F visto como um espaço vetorial sobre F, desta forma $V^* = \mathcal{L}(V, F)$ chama-se *espaço dual* (algébrico) de V. Um elemento $f \in V^*$ chama-se *funcional linear* ou *forma linear* sobre V. Um dos exemplos mais importante de um funcional linear é a função traço, confira Exercício (17) da Seção 1.2: $\operatorname{tr} : F^{n\times n} \to F$ definida como $\operatorname{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$, para todo $\mathbf{A} = (a_{ij}) \in F^{n\times n}$. Dados $\mathbf{u} \in V$ e $f \in V^*$ é conveniente, por razões técnicas, escrever

$$f(\mathbf{u}) = \langle \mathbf{u}, f \rangle \tag{4.3}$$

o qual chama-se um *emparelhamento de dualidade* de V e V^*. Neste caso, a função

$E : V \times V^* \to F$ definida como $E(\mathbf{u}, f) = \langle \mathbf{u}, f \rangle$ é linear em cada argumento ou componente, ou seja,

$$\langle \mathbf{u} + a\mathbf{v}, f \rangle = \langle \mathbf{u}, f \rangle + a\langle \mathbf{v}, f \rangle \ \text{ e } \ \langle \mathbf{u}, f + bg \rangle = \langle \mathbf{u}, f \rangle + b\langle \mathbf{u}, g \rangle,$$

para todos $\mathbf{u}, \mathbf{v} \in V$, $f, g \in V^*$ e $a, b \in F$. A função E chama-se *forma bilinear*. Em particular, $E_f \in V^*$, para todo $f \in V^*$.

Exemplo 4.48 *Para cada lista $\mathbf{a} = (a_1, \ldots, a_n)$ sobre F^n, mostre que a função $f_{\mathbf{a}} : F^n \to F$ definida como $f_{\mathbf{a}}(\mathbf{x}) = a_1x_1 + \cdots + a_nx_n$ é um funcional linear. Reciprocamente, qualquer elemento de $(F^n)^*$ é desta forma. Conclua que a função $\sigma : F^n \to (F^n)^*$ definida como $\sigma(\mathbf{a}) = f_{\mathbf{a}}$ é um isomorfismo.*

Solução. Dado $f \in (F^n)^*$, obtemos $f(\mathbf{x}) = \sum_{i=1}^{n} x_i f(\mathbf{e}_i)$. Portanto, basta escolher $a_i = f(\mathbf{e})_i \in F$. ■

Sejam $\alpha = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ uma base de V e $\beta = \{1\}$ a base canônica de F. Então os funcionais lineares $f_i \in V^*$, com $i = 1, \ldots, n$, definidos como

$$f_i(\mathbf{v}_j) = \langle \mathbf{v}_j, f_i \rangle = \delta_{ij}, j = 1, \ldots, n.$$

formam uma base de V^*, a qual chama-se *base dual* de α e denotamos por $\alpha^* = \{f_1, \ldots, f_n\}$. Logo, $\dim V = \dim V^*$. Note que os funcionais lineares f_i agem sobre um vetor $\mathbf{v} = \sum_{j=1}^{n} x_j \mathbf{v}_j$ como

$$f_i(\mathbf{v}) = x_j \sum_{j=1}^{n} f_i(\mathbf{v}_j) = x_j \sum_{j=1}^{n} \delta_{ij} = x_i, i = 1, \ldots, n.$$

Portanto, f_i é a função que associa cada vetor $\mathbf{v}$ a sua i-ésima coordenada x_i em relação à base α e os f_i chamam-se *formas coordenadas* ou *projeções coordenadas* associadas com $\mathbf{v}_i$. Além disso, dados $\mathbf{v} \in V$ e $f \in V^*$, existem únicos $a_i, b_i \in F$ tais que

$$\mathbf{v} = \sum_{i=1}^{n} a_i \mathbf{v}_i \ \text{ e } \ f = \sum_{i=1}^{n} b_i f_i.$$

É fácil verificar que $a_i = p_i(\mathbf{a}) = (f_i \circ L_\alpha)(\mathbf{a}) = f_i(\mathbf{v}) = \langle \mathbf{v}, f_i \rangle$, confira diagrama abaixo, $b_i = f_\alpha(\mathbf{e}_i) = (f_\alpha \circ T_\alpha)(\mathbf{v}_i) = f(\mathbf{v}_i) = \langle \mathbf{v}_i, f\rangle$ e

$$\langle \mathbf{v}, f\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j \langle \mathbf{v}_i, f_j\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j \delta_{ij} = \sum_{i=1}^{n} a_i b_i$$

é semelhante ao "*produto escalar*" usual de vetores, pois $V \neq V^*$. Quando $\langle \mathbf{v}, f\rangle = 0$ diremos que $\mathbf{v}$ é *anulado* por f ou "ortogonal" a f. Por isto, os vetores de V chamam-se *contravariantes* e os de V^* *covariantes*.

$$\begin{tikzcd} F^n \arrow[r, "L_\alpha"] \arrow[dr, dashed, "p_i"'] & V \arrow[d, dashed, "f_i"] \\ & F \end{tikzcd} \qquad \begin{tikzcd} V \arrow[r, "T_\alpha"] \arrow[dr, dashed, "f"'] & F^n \arrow[d, dashed, "f_\alpha"] \\ & F \end{tikzcd}$$

(a) (b)

Neste caso, f_α chama-se a *representação* de f em relação à base α.

Teorema 4.49 *Seja $\alpha = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ uma base de V. Enão existe uma única base $\alpha^* = \{f_1, \ldots, f_n\}$ de V^* dual a α. Além disso,*

1. $\mathbf{v} = \sum_{i=1}^{n} f_i(\mathbf{v})\mathbf{v}_i$*, para todo* $\mathbf{v} \in V$*, com* $[\mathbf{v}]_\alpha = (f_1(\mathbf{v}), \ldots, f_n(\mathbf{v}))^t$.
2. $f = \sum_{i=1}^{n} f(\mathbf{v}_i) f_i$*, para todo* $f \in V^*$*, com* $[f]_{\alpha^*} = (f(\mathbf{v}_1), \ldots, f(\mathbf{v}_n))^t$.
3. $\dim V = \dim V^*$, $\sigma : V \to V^*, \sigma(\mathbf{v}_i) = f_i$.

Prova. Fica como um exercício. ■

É muito importante, de um ponto de vista teórico e didático, observar o seguinte: sejam $\alpha = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ uma base qualquer de V e $\beta = \{f_1, \ldots, f_n\}$ uma base qualquer de V^*. Então a "*matriz de Gram*[3]" $\mathbf{G} = (\langle \mathbf{v}_i, f_j\rangle) = (g_{ij})$ é não singular e

$$\langle \mathbf{v}, f\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \langle \mathbf{v}_i, f_j\rangle a_i b_j = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij} a_i b_j.$$

Exemplo 4.50 *Seja $\alpha = \{(3, 0-3), (-1, 1, 2), (2, 1, 1)\}$ uma base de F^3. Determine a base dual α^*.*

[3] Jørgen Pedersen Gram, 1850-1916, matemático dinamarquês.

Solução. Pelo item (1) do Teorema 4.49, basta determinar $[\mathbf{v}]_\alpha = [(x, y, z)]_\alpha$. Mas, isto é equivalente a escalonar a matriz

$$\left(\begin{array}{rrr|r} 3 & -1 & 2 & x \\ 0 & 1 & 1 & y \\ -3 & 2 & 1 & z \end{array}\right) \to \cdots \to \left(\begin{array}{ccc|l} 1 & 0 & 0 & \frac{1}{6}(-x+5y-3z) \\ 0 & 1 & 0 & \frac{1}{2}(-x+3y-z) \\ 0 & 0 & 1 & \frac{1}{2}(x-y+z) \end{array}\right).$$

Portanto,

$$f_1(\mathbf{v}) = \frac{1}{6}(-x+5y-3z),\ f_2(\mathbf{v}) = \frac{1}{2}(-x+3y-z) \text{ e } f_3(\mathbf{v}) = \frac{1}{2}(x-y+z)$$

e $\alpha^* = \{f_1, f_2, f_3\}$ é a base desejada. ■

Exemplo 4.51 *Para cada* $\mathbf{A} \in F^{n\times n}$ *fixada, mostre que* $f_{\mathbf{A}} : F^{n\times n} \to F$ *definida como* $f_{\mathbf{A}}(\mathbf{X}) = \operatorname{tr}(\mathbf{A}^t\mathbf{X})$ *é um funcional linear. Reciprocamente, qualquer elemento de* $(F^{n\times n})^*$ *é desta forma. Conclua que a função* $\sigma : F^{n\times n} \to (F^{n\times n})^*$ *definida como* $\sigma(\mathbf{A}) = f_{\mathbf{A}}$ *é um isomorfismo.*

Solução. Observe que $\operatorname{tr}(\mathbf{A}^t\mathbf{E}_{ij}) = \operatorname{tr}(\sum_{k=1}^n a_{kj}\mathbf{E}_{ki}) = a_{ij}$, para cada $\mathbf{A} \in F^{n\times n}$. Sejam $f \in (F^{n\times n})^*$, $\alpha = \{\mathbf{E}_{ij} : i = 1, \dots, n; j = 1, \dots, n\}$ a base canônica de $F^{n\times n}$ e $\alpha^* = \{E_{ij} : i = 1, \dots, n; j = 1, \dots, n\}$ sua base dual. Então

$$E_{ij}(\mathbf{X}) = \langle \mathbf{X}, E_{ij}\rangle = x_{ij} \text{ e } f = \sum_{i=1}^n \sum_{j=1}^n a_{ij}E_{ij},$$

de modo que

$$\begin{aligned} f(\mathbf{X}) &= \textstyle\sum_{i=1}^n \sum_{j=1}^n a_{ij}E_{ij}(\mathbf{X}) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_{ij} \\ &= \textstyle\sum_{i=1}^n \sum_{j=1}^n x_{ij}\operatorname{tr}(\mathbf{A}^t\mathbf{E}_{ij}) = \operatorname{tr}(\sum_{i=1}^n \sum_{j=1}^n x_{ij}\mathbf{A}^t\mathbf{E}_{ij}) \\ &= \operatorname{tr}(\mathbf{A}^t\mathbf{X}). \end{aligned}$$

Portanto, $f(\mathbf{X}) = \operatorname{tr}(\mathbf{A}^t\mathbf{X})$. Em particular, $\operatorname{tr} = \sum_{i=1}^n E_{ii}$. ■

Sejam $V = \mathbb{P}_n(F)$ e $t \in F$ fixado. Então a função $P_t : V \to F$ definida como $P_t(p(x)) = p(t)$ é claramente um funcional linear, ou seja, $P_t \in V^*$, para todo $t \in F$. Assim, se $t_0, t_1, \dots, t_n \in F$ forem distintos, então $\alpha^* = \{P_0, \dots, P_n\}$ é uma base de V^*. De fato. Se

$$P = c_0P_0 + \cdots + c_nP_n = 0,$$

então $P(p(x)) = 0$, para todo $p(x) \in V$. Em particular, avaliando em $1, x, \ldots, x^n$, obtemos o sistema homogêneo

$$c_0 t_0^i + c_1 t_1^i + \cdots + c_n t_n^i = 0, i = 0, 1, \ldots, n, \Leftrightarrow \mathbf{V}_{n+1}\mathbf{C} = \mathbf{O}.$$

Assim, $c_i = 0$, pois o determinante da matriz de Vandermonde $\mathbf{V}_{n+1} = (a_{ij})$, com $a_{ij} = t_j^i, i, j = 0, 1, \ldots, n$, é não nulo. Seja $\alpha = \{p_0(x), \ldots, p_n(x)\}$ a base de V, a qual possui α^* como dual. Então α é determinada pelas relações

$$P_i(p_j(x)) = p_j(t_i) = \delta_{ij} \Leftrightarrow \mathbf{V}_{n+1}\mathbf{X}_j = \mathbf{E}_j,$$

com $\mathbf{X}_j = (a_{0j}, \ldots, a_{nj})^t$ os coeficientes de $p_j(x)$. Neste caso, depois de alguns cálculos,

$$p_j(x) = \frac{1}{\widehat{p}_j(t_j)}\widehat{p}_j(x), \text{ com } \widehat{p}_j(x) = (x - t_0)\cdots(x - t_{j-1})(x - t_{j+1})\cdots(x - t_n).$$

Portanto, existe exatamente um polinômio

$$p(x) = c_0 p_0(x) + \cdots + c_n p_n(x) \in V$$

tal que $p(t_i) = c_i$. Esta expressão chama-se *Fórmula de Interpolação de Lagrange*. Logo, podemos concluir que qualquer polinômio de grau n é completamente determinado a partir de seus valores em $n + 1$ pontos distintos, isto é, a função $T : \mathbb{P}_n(F) \to F^{n+1}$ definida como $T(p(x)) = (p(t_0), \ldots, p(t_n))$ é um isomorfismo linear.

Teorema 4.52 *Sejam* $\alpha = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}, \beta = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ *bases de* V *e* $\alpha^* = \{f_1, \ldots, f_n\}, \beta^* = \{g_1, \ldots, g_n\}$ *as correspondentes bases duais. Se* $\mathbf{P} = [I]_\alpha^\beta$*, então* $[I]_{\alpha^*}^{\beta^*} = (\mathbf{P}^{-1})^t$.

Prova. Seja $\mathbf{Q} = [I]_{\alpha^*}^{\beta^*}$. Então $\mathbf{w}_i = \sum_{j=1}^n p_{ji}\mathbf{v}_j$ e $\mathbf{g}_i = \sum_{j=1}^n q_{ji} f_j$. Então

$$\delta_{ij} = g_i(\mathbf{w}_j) = \sum_{k=1}^n q_{ki} f_k(\sum_{i=1}^n p_{ij}\mathbf{v}_i) = \sum_{k=1}^n q_{ki}(\sum_{i=1}^n p_{ij}\delta_{ik}) = \sum_{k=1}^n q_{ki} p_{kj}.$$

Portanto, $\mathbf{Q}^t\mathbf{P} = \mathbf{I}$ e $\mathbf{Q} = (\mathbf{P}^{-1})^t$. ∎

O Teorema da Transformação de Coordenadas 4.52 nos fornece um método alternativo para obter bases duais em F^n. Seja $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ uma base qualquer de F^n.

Então $\mathbf{P} = [I]_\alpha^\beta$ é uma matriz cuja j-ésima coluna $\mathbf{C}_j$ é as coordenadas de $\mathbf{v}_j$ em relação à base canônica de F^n. Observe, pelo Exemplo 4.48, que cada linha $\mathbf{L}_i = (b_{i1}, \ldots, b_{in})$ de $\mathbf{P}^{-1}$ induz um único funcional linear $f_i : F^n \to F$ definido como

$$f_i(\mathbf{x}) = b_{i1}x_1 + \cdots + b_{in}x_n.$$

Como $f_i(\mathbf{v}_j) = \mathbf{L}_i\mathbf{C}_j = \delta_{ij}$ temos que $\alpha^* = \{f_1, \ldots, f_n\}$. Consideremos novamente o Exemplo 4.50. Assim, basta escalonar a matriz

$$\left(\begin{array}{rrr|rrr} 3 & -1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ -3 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \to \cdots \to \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & -\frac{1}{6} & \frac{5}{6} & -\frac{3}{6} \\ 0 & 1 & 0 & -\frac{1}{2} & \frac{3}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{array}\right).$$

Portanto, $\alpha^* = \{f_1, f_2, f_3\}$, com $f_1 = \frac{1}{6}[-1, 5, -3]$ etc.

É muito importante, de um ponto de vista teórico e didático, termos uma noção geral de dualidade de um espaço vetorial qualquer. Para isto, seja S um conjunto infinito. Então já vimos que

$$V = F^{(S)} = \{\mathbf{x} \in \mathcal{F}(S, F) : \operatorname{supp}(\mathbf{x}) \text{ é finito}\}$$

era um subespaço de $\mathcal{F}(S, F)$ e que $\alpha = \{\mathbf{e}_s : s \in S\}$ era a base canônica de V. Note que a função $\lambda : S \to V$ definida como $\lambda(s) = \mathbf{e}_s$ é claramente injetora. Pelo caso geral do Teorema 4.11, para cada $g \in \mathcal{F}(S, F)$, existe um único $f \in V^*$ tal que $g = f \circ \lambda$, ou seja, $g(s) = f(\mathbf{e}_s)$. Assim, a função $\sigma : V^* \to \mathcal{F}(S, F)$ definida como $\sigma(f) = g$ é um isomorfismo. Portanto, V não é isomorfo a V^*, ou seja, $\dim V < \dim V^*$, pois $h \in \mathcal{F}(S, F)$ definida como $h(s) = 1$, para todo $s \in S$, implica que $\alpha \cup \{h\}$ é LI.

O processo de obter o espaço dual pode ser repetido, pois se V for um espaço vetorial sobre F, então V^* é um espaço vetorial sobre F e $V^{**} = (V^*)^*$ é um espaço vetorial sobre F e chama-se *espaço bidual* de V.

Vamos apresentar um método alternativo de expressar a conexão entre um espaço vetorial e seu dual. Para isto, consideremos a forma bilinear $E : V \times V^* \to F$ definida como $E(\mathbf{v}, f) = \langle \mathbf{v}, f \rangle$. Então, cada $\mathbf{v} \in V$ fixado, induz um funcional linear $E_\mathbf{v} : V^* \to F$ definido como $E_\mathbf{v}(f) = E(\mathbf{v}, f)$, ou seja, $E_\mathbf{v} \in V^{**}$, para todo $\mathbf{v} \in V$. Assim, obtemos a transformação linear $\phi_V : V \to V^{**}$ definida como $\phi_V(\mathbf{v}) = E_\mathbf{v}$ e denotamos por $\phi = \phi_V$. Note que $\langle f, \phi(\mathbf{v}) \rangle = \phi(\mathbf{v})(f) = \langle \mathbf{v}, f \rangle$, para todo $\mathbf{v} \in V$ e $f \in V^*$. Uma propriedade fundamental de ϕ é que ela é sempre injetora, ou seja, se $\mathbf{v} \neq \mathbf{0}$, então sempre existe um $f \in V^*$ tal que $f(\mathbf{v}) \neq 0$. Por isto, ϕ chama-se *imersão canônica*.

Teorema 4.53 *Para qualquer espaço vetorial V, ϕ é sempre injetora. Conclua que se V for de dimensão finita, então ϕ é um isomorfismo.*

Prova. Dado $\mathbf{v} \in V$, com $\mathbf{v} \neq \mathbf{0}$, temos, pelo Teorema 3.34, que V contém uma base α tal que $\mathbf{v} \in \alpha$. Seja $f : V \to F$ definido como $f(\mathbf{v}) = 1$ e $f(\mathbf{u}) = 0$, para todo $\mathbf{u} \in \alpha - \{\mathbf{v}\}$. Então $f \in V^*$ e $E_{\mathbf{v}}(f) = f(\mathbf{v}) = 1$, de modo que $\phi(\mathbf{v}) = E_{\mathbf{v}} \neq 0$. Portanto, $\ker \phi = \{\mathbf{0}\}$ e ϕ é injetora. Por outro lado, quando V for de dimensão finita, $\dim V = \dim V^* = \dim (V^*)^*$ e ϕ é sobrejetora. ■

Vale ressaltar o seguinte: seja V um espaço vetorial sobre F com uma base $\alpha = \{\mathbf{e}_s : s \in S\}$. Então cada $s \in S$ induz um funcional linear $f_s \in V^*$ definido como $f_s(\mathbf{e}_t) = \delta_{st}$, de modo que o conjunto $\alpha^* = \{f_s : s \in S\}$ é sempre LI, pois dado $n \in \mathbb{N}$, se $s_i \in S, x_{s_i} \in F$, com $i = 1, \ldots, n$, e

$$f = x_{s_1} f_{s_1} + \cdots + x_{s_n} f_{s_n} = 0,$$

então $f(\mathbf{v}) = 0$, para todo $\mathbf{v} \in V$. Em particular, avaliando em $\mathbf{e}_{s_1}, \ldots, \mathbf{e}_{s_n}$, obtemos

$$0 = f(\mathbf{e}_{s_j}) = \sum_{i=1}^{n} x_{s_i} f_{s_i}(\mathbf{e}_{s_j}) = \sum_{i=1}^{n} x_{s_i} \delta_{s_i s_j} = x_{s_j}, j = 1, \ldots, n.$$

Assim, $x_{s_1} = \cdots = x_{s_n} = 0$ e α^* é LI. Portanto, $\dim V \leq \dim V^*$. Note que se S for um conjunto infinito, então existe um $f_\alpha \in V^*$ tal que $f_\alpha(e_s) = 1$, para todo s em S, de modo que $\alpha^* \cup \{f_\alpha\}$ é LI. Neste caso, $\dim V < \dim V^*$. Além disso, quando $\dim V = n$, $\operatorname{Im} \phi = \{E_{\mathbf{v}} : \mathbf{v} \in V\} = V^{**}$. Assim, para cada $L \in V^{**}$, existe um único $\mathbf{v} \in V$ tal que $L = \phi(\mathbf{v}) = E_{\mathbf{v}}$ e $L(f) = f(\mathbf{v})$, para todo $f \in V^*$. Logo, se $\beta = \{f_1, \ldots, f_n\}$ for uma base de V^*, então existe uma base $\alpha = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ de V tal que $f_i(\mathbf{v}_j) = \delta_{ij}$, pois se $\beta^* = \{E_1, \ldots, E_n\}$ for a base dual de β, então existe um único $\mathbf{v}_j \in V$ tal que $f_i(\mathbf{v}_j) = E_j(f_i) = \delta_{ij}$. Portanto, às vezes, é conveniente identificar $\mathbf{v}$ com $E_{\mathbf{v}}$, ou seja, $V = V^{**}$. Além disso, identificar cada $f \in V^*$ com suas projeções coordenadas $[x_1, \ldots, x_n] = (x_1, \ldots, x_n)^t \in F^{n \times 1}$

Sejam V um espaço vetorial sobre F e S um subconjunto de V. O conjunto

$$S^\circ = \{f \in V^* : \langle \mathbf{v}, f \rangle = 0, \;\; \forall \;\; \mathbf{v} \in S\} = \{f \in V^* : f(S) = 0\}$$

é um subespaço de V^* e chama-se *anulador* de S. É muito importante ressaltar que S não precisa ser um subespaço de V, $\{\mathbf{0}\}^\circ = V^*$, $V^\circ = \{0\}$ e $S^\circ \neq V^*$ quando $S \neq \{\mathbf{0}\}$. Por outro lado, se W for um subespaço de V^*, então seu anulador W° é um subespaço de V^{**}. Em particular, se V for de dimensão finita, então $V = V^{**}$ e

$$\phi^{-1}(W^{\circ}) = \{\mathbf{v} \in V : \langle \mathbf{v}, f\rangle = 0, \;\; \forall \; f \in W\} = \bigcap_{f\in W} \ker f$$

é um subespaço de V, ou seja, W é completamente determinado por W°.

Teorema 4.54 *Sejam V espaço vetorial sobre F e U um subespaço de V. Então $(V/U)^*$ é isomorfo a U°.*

Prova. Cada $f \in U^{\circ} \subseteq V^*$ induz um funcional linear $g : V/U \to F$ definido como $g(\mathbf{v} + U) = f(\mathbf{v})$, pois $\mathbf{v} + U = \mathbf{w} + U$ se, e somente se, $\mathbf{w} - \mathbf{v} \in U$, de modo que $f(\mathbf{w} - \mathbf{v}) = 0$ implica que $f(\mathbf{v}) = f(\mathbf{w})$ e $g(\mathbf{v} + U) = g(\mathbf{w} + U)$. Reciprocamente, cada $g \in (V/U)^*$ induz um $f \in U^{\circ} \subseteq V^*$ definido como $f(\mathbf{v}) = g(\mathbf{v} + U)$, pois $f(\mathbf{v}) = 0$, para todo $\mathbf{v} \in U$. Portanto, a função $T : U^{\circ} \to (V/U)^*$ definida como $T(f) = g$ é o isomorfismo desejado, por exemplo, se $f \in \ker T$, então $f(\mathbf{v}) = 0$, para todo $\mathbf{v} \in V$, de modo que $f = 0$, confira exercício (18) a seguir. ■

Corolário 4.55 *Sejam V espaço vetorial sobre F, com $\dim V = n$, e U um subespaço de V. Então $\dim V = \dim U + \dim U^{\circ}$ e $U^{\circ\circ} = (U^{\circ})^{\circ} = U$.*

Prova. Seja $\dim U = k$. Então $\dim(V/U) = n - k$. Assim, pelo Teorema 4.54, $\dim U^{\circ} = \dim(V/U)^* = n - k$. Note que $\dim V^* = \dim U^{\circ} + \dim U^{\circ\circ}$, de modo que $\dim U^{\circ\circ} = k$. Por outro lado, dado $\mathbf{v} \in U$, temos que $\langle \mathbf{v}, f\rangle = 0$, para todo $f \in U^{\circ}$, de modo que $\mathbf{v} \in (U^{\circ})^{\circ}$ e $U \subseteq (U^{\circ})^{\circ}$. Portanto, $U^{\circ\circ} = (U^{\circ})^{\circ} = U$. Uma outra prova, cada $f \in V^*$ induz um funcional linear $g : U \to F$ definido como $g(\mathbf{u}) = f(\mathbf{u})$ e a função $T : V^* \to U^*$ definida como $T(f) = g$ é sobrejetora, confira exercício (17) a seguir. Assim, pelo Teorema 4.30, $V^*/\ker T$ isomorfo a U^* e $\ker T = U^{\circ}$. ■

É importante de um ponto vista, teórica e prático, ressaltar o seguinte: determinar U° é o mesmo que U e vice-versa.

Vamos apresentar uma prova direta do Corolário 4.55. Sabemos, pelo Teorema 3.41, que qualquer base $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ de U é parte de uma base de V, digamos, $\alpha = \{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$, Portanto, $V = U \oplus W$, com $W = F[\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n]$ o subespaço complementar de U. Por outro lado, pelo Teorema 4.49, existe uma única base de $\alpha^* = \{f_1, \ldots, f_n\}$ de V^* dual a α. Assim, $f_i(\mathbf{v}) = 0$, para todo $\mathbf{v} \in U$ e $i = k + 1, \ldots, n$, isto significa que $\mathbf{v} \in \ker f_i$, para todo $\mathbf{v} \in U$ e $i = k + 1, \ldots, n$. Logo, é suficiente provar que $\{f_{k+11}, \ldots, f_n\}$ gera U°. De fato. Dado $f \in U^{\circ} \subseteq V^*$

temos, pelo Teorema 4.49, que

$$f = \sum_{i=1}^{n} f(\mathbf{v}_i) f_i = \sum_{i=k+1}^{n} f(\mathbf{v}_i) f_i.$$

Portanto, $U^\circ = F[f_{k+11}, \ldots, f_n]$ e $\dim U^\circ = n - k = \dim V - \dim U$ é a *codimensão* de U. Antes de continuar vamos lembrar que um sistema de equações lineares homogêneo, com m equações e n incógnitas, foi definido como:

$$f_i(\mathbf{x}) := \sum_{j=1}^{n} a_{ij} x_j \ \text{ e } \ f_i(\mathbf{x}) = 0,$$

para todo $\mathbf{x} = (x_1, \ldots, x_n) \in F^n$, confira equação (2.3). O problema aqui é determinar a dimensão do espaço solução. Para isto, pelo Exemplo 4.48, cada linha $\mathbf{L}_i = (a_{i1}, \ldots, a_{in}) \in F^n$ corresponde a um único funcional linear $f_i \in (F^n)^*$. Pondo $U = F[\mathbf{L}_1, \ldots, \mathbf{L}_m]$ e $f(\mathbf{x}) = c_1 x_1 + \cdots + c_n x_n$, temos $f \in U^\circ$ se, e somente se, $\mathbf{C} = (c_1, \ldots, c_n)^t$ é uma solução do sistema $\mathbf{AX} = \mathbf{O}$, ou seja, $\mathbf{L}_i \mathbf{C} = 0$. Assim, podemos concluir que, o espaço solução do sistema é o anulador das linhas de $\mathbf{A}$ e, portanto, do espaço linha de $\mathbf{A}$. Retornando ao nosso caso, como $U^\circ = F[f_{k+11}, \ldots, f_n]$ temos que

$$\mathbf{v} = x_1 \mathbf{v}_1 + \cdots + x_n \mathbf{v}_n \in U \Leftrightarrow f_{k+1}(\mathbf{v}) = 0, \ldots, f_n(\mathbf{v}) = 0.$$

Mas, isto é equivalente a um sistema homogêneo, com $n - k$ equações e n incógnitas, cujo espaço solução é $(U^\circ)^\circ = U$. Posto de outra forma $U = \cap_{i=k+1}^{n} \ker f_i$. Além disso, vale observar que se $U = F[\mathbf{v}]$, então $U^\circ = \ker E_{\mathbf{v}}$ é um subespaço de V^* e $(U^\circ)^\circ = U$.

Exemplo 4.56 *Sejam* $\mathbf{v}_1 = \mathbf{e}_1 - \mathbf{e}_4, \mathbf{v}_2 = \mathbf{e}_2 - \mathbf{e}_4, \mathbf{v}_3 = \mathbf{e}_3 - \mathbf{e}_4 \in F^4$ *e* $U = F[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$. *Determine uma base de* U°.

Solução. Pelo exposto, basta escalonar a matriz

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix},$$

a qual já está na forma de escada. Como $f = [c_1, c_2, c_3, c_4] \in (F^4)^*$ e (c_1, c_2, c_3, c_4) deve ser uma solução do sistema $\mathbf{AX} = \mathbf{O}$ temos que $f \in U^\circ$ se, e somente se, $f(\mathbf{v}_i) =$

0 se, e somente se, $c_1 = c_2 = c_3 = c_4$. Portanto, os $f \in U^\circ$ são da forma $f_c(\mathbf{x}) = c(x_1 + x_2 + x_3 + x_4)$. Em particular, $f_1 = [1, 1, 1, 1]$ é uma base de U°. ■

Exemplo 4.57 *Determine uma base para o espaço solução do sistema*

$$\left\{ \begin{array}{rcl} x_1 + x_2 + 2x_3 & = & 0 \\ x_1 + 0x_2 + x_3 & = & 0 \\ 2x_1 + x_2 + 3x_3 & = & 0. \end{array} \right.$$

Solução. Sejam $f_1 = [1, 1, 2]$, $f_2 = [1, 0, 1]$ e $f_3 = [2, 1, 3]$. Então devemos encontrar o subespaço U anulado por eles. Para isto, basta escalonar a matriz

$$\begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 3 \end{pmatrix} \longrightarrow \cdots \longrightarrow \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Assim, $g_1(\mathbf{x}) = x_1 + x_3$ e $g_2(\mathbf{x}) = x_2 + x_3$ geram o mesmo subespaço de $(F^3)^*$ como f_1, f_2 e f_3. Logo, $(x_1, x_2, x_3) \in U$ se, e somente se, $x_1 + x_3 = 0$ e $x_2 + x_3 = 0$. Portanto, $U = F[(1, 1, -1)]$. ■

Sejam V um espaço vetorial sobre F e U um subespaço de V. Diremos que U é um *subespaço maximal* de V se U for um subespaço próprio de V e não existir um subespaço W de V tal que $U \subset W \subset V$. Um *hiperplano* de V é qualquer subespaço maximal de V.

Teorema 4.58 *Seja V um espaço vetorial sobre F, com $\dim V = n$. Então as seguintes condições são equivalentes.*

1. *U é um hiperplano de V;*

2. $\dim U = n - 1$;

3. *$U = \ker f$, para algum $f \in V$, com $f \neq 0$.*

Prova. $(1 \Rightarrow 2)$ Suponhamos que U seja um hiperplano de V. Então existe um $\mathbf{v} \in V$, com $\mathbf{v} \neq \mathbf{0}$ e $\mathbf{v} \notin U$. Assim, $W = F[U, \mathbf{v}] = \{\mathbf{u} + x\mathbf{v} : \mathbf{u} \in U, x \in F\}$ é um subespaço W de V, com $U \subseteq W$ e $U \neq W$. Logo, $V = W$. Seja $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ uma base de U. Então claramente $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}\}$ é uma base de V. Portanto, $n = \dim V = k + 1$ e $\dim U = k = n - 1$.

$(2 \Rightarrow 3)$ Suponhamos que U seja um subespaço de V tal que $\dim U = n - 1$. Então existe um $\mathbf{v} \in V$, com $\mathbf{v} \neq \mathbf{0}$ e $\mathbf{v} \notin U$, de modo que $V = U \oplus \{\mathbf{v}\}$. Neste caso, $\{\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}, \mathbf{v}\}$ é uma base de V. Logo, pelo Teorema 4.49, existe um $f = f_n \in V^*$ tal que $f(\mathbf{v}) = 1$ e $f(\mathbf{u}) = 0$, para todo $\mathbf{u} \in U$. É fácil verificar que $\mathbf{u} \in \ker f$ se, e somente se, $\mathbf{u} \in U$. Portanto, $U = \ker f$.

$(3 \Rightarrow 1)$ Suponhamos que $U = \ker f$, para algum $f \in V$, com $f \neq 0$. Então existe um $\mathbf{v} \in V$ tal que $f(\mathbf{v}) \neq 0$, de modo que $\mathbf{v} \notin U$. Assim, $U \subset F[U, \mathbf{v}] \subseteq V$. Por outro lado, pondo $c = f(\mathbf{v})^{-1}$ e dado $\mathbf{u} \in V$, obtemos

$$\mathbf{u} = (\mathbf{u} - cf(\mathbf{u})\mathbf{v}) + cf(\mathbf{u})\mathbf{v} \in F[U, \mathbf{v}],$$

pois $f(\mathbf{u} - cf(\mathbf{u})\mathbf{v}) = f(\mathbf{u}) - f(\mathbf{u})(cf(\mathbf{v})) = f(\mathbf{u}) - f(\mathbf{u}) = 0$. Portanto, $V = F[U, \mathbf{v}]$ e U é um hiperplano de V. ■

Observe, pelo Exemplo 4.48, que qualquer hiperplano de F^n é da forma

$$U_{\mathbf{a}} = \{(x_1, \ldots, x_n) \in F^n : a_1x_1 + \cdots + a_nx_n = 0\}$$

para algum $\mathbf{a} = (a_1, \ldots, a_n) \in F^n$. Neste caso, $U_{\mathbf{a}}^{\circ}$ é uma reta em F^n.

Sejam V e W espaços vetoriais sobre F. Então cada $T \in \mathcal{L}(V, W)$ fixado induz uma função $T^t : \mathcal{L}(W, U) \to \mathcal{L}(V, U)$ definida como $T^t(S) = S \circ T$, para qualquer espaço vetorial U sobre F, confira o diagrama:

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

É fácil verificar que T^t satisfaz as seguintes condições:

1. $T^t \in \mathcal{L}(\mathcal{L}(W, U), \mathcal{L}(V, U))$.

2. $(R + aT)^t = R^t + aT^t$, para todos $R, T \in \mathcal{L}(V, W)$ e $a \in F$.

3. Se I for a identidade sobre V, então I^t é a identidade sobre $\mathcal{L}(V, U)$.

4. $(TR)^t = R^tT^t$, para todo $R \in \mathcal{L}(V, W)$ e $T \in \mathcal{L}(W, Z)$.

5. Se T for um isomorfismo, então T^t também o é e $(T^t)^{-1} = (T^{-1})^t$.

Vamos estudar mais algumas propriedades da transformação linear T^t, pata todo $T \in \mathcal{L}(V, W)$, no caso particular em que $U = F$, e diremos que $T^t : W^* \to V^*$ é a *transformação transposta* ou *adjunta* de T e denotamos por T^*, confira diagrama (a), e é comum denotar $T^*(g) = g \circ T$ pela justaposição gT.

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[dr, dashed, "g\circ T"'] & W \arrow[d, "g"] \\ & F \end{tikzcd} \qquad \begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\phi_V"'] & W \arrow[d, "\phi_W"] \\ V^{**} \arrow[r, "T^{**}"'] & W^{**} \end{tikzcd}$$

(a) (b)

Note, para todo $\mathbf{v} \in V$ e $g \in W^*$, que

$$\langle T(\mathbf{v}), g\rangle = g(T(\mathbf{v})) = (g \circ T)(\mathbf{v}) = T^*(g)(\mathbf{v}) = \langle \mathbf{v}, \mathbf{T}^*(\mathbf{g})\rangle.$$

Teorema 4.59 *Sejam* $\alpha = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}, \beta = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ *bases de* V, W *e* $\alpha^* = \{f_1, \dots, f_n\}, \beta^* = \{g_1, \dots, g_m\}$ *as correspondentes bases duais. Se* $T \in \mathcal{L}(V, W)$ *e* $\mathbf{A} = [T]_\alpha^\beta$*, então* $T^* \in \mathcal{L}(W^*, V^*)$ *e* $[T^*]_{\alpha^*}^{\beta^*} = \mathbf{A}^t$.

Prova. É fácil verificar que $T^* \in \mathcal{L}(W^*, V^*)$. Seja $\mathbf{B} = (b_{ij}) = [T^*]_{\alpha^*}^{\beta^*}$. Então

$$\langle T(\mathbf{v}_i), g_j\rangle = \langle\sum_{k=1}^{m} a_{ki}\mathbf{w}_k, g_j\rangle = \sum_{k=1}^{m} a_{ki} g_j(\mathbf{w}_k) = \sum_{k=1}^{m} a_{ki}\delta_{jk} = a_{ji}.$$

Por outro lado,

$$\langle \mathbf{v}_i, T^*(g_j)\rangle = \langle \mathbf{v}_i, \sum_{k=1}^{m} b_{kj} f_k\rangle = \sum_{i=1}^{m} b_{kj}\langle \mathbf{v}_i, f_k\rangle = \sum_{i=1}^{m} b_{kj}\delta_{ik} = b_{ij}.$$

Portanto, $b_{ij} = a_{ji}$ e $\mathbf{B} = \mathbf{A}^t$. ■

Vamos finalizar esta seção apresentando a conexão entre uma transformação linear e seu adjunto.

Teorema 4.60 *Sejam* $T \in \mathcal{L}(V, W)$*, com* $\dim V = n$ *e* $\dim W = m$*, e* $T^* \in \mathcal{L}(W^*, V^*)$*. Então*

1. $\ker T^* = (\operatorname{Im} T)^\circ$.

2. $\operatorname{Im} T^* = (\ker T)^\circ$.

3. $\dim \operatorname{Im} T = \dim \operatorname{Im} T^* = \dim(\ker T)^\circ$. *Em particular, T^* é injetora se, e somente se, T for sobrejetora e vice-versa.*

Prova. (1) Dado $g \in W^*$, temos que $g \in \ker T^*$ se, e somente se, $\langle T(\mathbf{v}), g\rangle = \langle \mathbf{v}, \mathbf{T}^*(\mathbf{g})\rangle = 0$, para todo $\mathbf{v} \in V$, se, e somente se, $g \in (\operatorname{Im} T)^\circ$. Portanto, $\ker T^* = (\operatorname{Im} T)^\circ$. Observe que não necessitamos da dimensão na prova.

(2) É fácil verificar, em geral, que $\operatorname{Im} T^* \subseteq (\ker T)^\circ$. Por outro lado, pelo Teorema 4.21, pelo item (1) e o Corolário 4.55, obtemos

$$\begin{aligned}
\dim \operatorname{Im} T^* &= \dim W^* - \dim \ker T^* = \dim W - \dim(\operatorname{Im} T)^\circ \\
&= \dim \operatorname{Im} T = \dim V - \dim \ker T \\
&= \dim(\ker T)^\circ.
\end{aligned}$$

Portanto, $\operatorname{Im} T^* = (\ker T)^\circ$.

(3) Note que $\dim(\ker T)^\circ + \dim \ker T = \dim V = \dim \ker T + \dim \operatorname{Im} T$, de modo que $\dim(\ker T)^\circ = \dim \operatorname{Im} T$. Portanto, pelo item (2), $\dim \operatorname{Im} T = \dim \operatorname{Im} T^*$. Observe, pelos itens (1) e (2), que $\ker T^* = \{\mathbf{0}\}$ se, e somente se, $\operatorname{Im} T = W$ e $\ker T = \{\mathbf{0}\}$ se, e somente se, $\operatorname{Im} T^* = V^*$. Portanto, T^* é injetora se, e somente se, $\operatorname{Im} T = W$ se, e somente se, T é sobrejetora. ■

Finalmente, para qualquer $T \in \mathcal{L}(V, W)$, já vimos como construir o adjunto $T^* \in \mathcal{L}(W^*, V^*)$. Assim, de modo similar, construiremos $T^{**} \in \mathcal{L}(V^{**}, W^{**})$. Portanto, o diagrama (b) acima comuta. De fato. Para cada $\mathbf{v} \in V$, temos que $\phi_W(T(\mathbf{v})) = E_{T(\mathbf{v})}$. Logo, para qualquer $g \in W^*$, obtemos

$$E_{T(\mathbf{v})}(g) = \langle T(\mathbf{v}), g\rangle = \langle \mathbf{v}, \mathbf{T}^*(\mathbf{g})\rangle = \langle T^{**}(\mathbf{v}), g\rangle = E_{T^{**}(\mathbf{v})}(g),$$

de modo que $(\phi_W \circ T)(\mathbf{v}) = (T^{**} \circ \phi_V)(\mathbf{v})$. Portanto, $\phi_W \circ T = T^{**} \circ \phi_V$. Neste caso, se $\dim V = n$ e $\dim W = m$, então identificando $V = V^{**}$ e $W = W^{**}$, de modo que $\phi_V = \phi_W$ e $I_V = I_W$. Assim, $T^{**} = T$. Além disso, seja $\mathbf{A}$ a representação de T em relação à bases de V e W. Então, pelo Teorema 4.59, o posto coluna de $\mathbf{A}$ é igual a $\dim \operatorname{Im} T$ e o posto linha de $\mathbf{A}$ (igual ao posto coluna de A^t) é igual a $\dim \operatorname{Im} T^*$. Mas, pelo item (3) do Teorema 4.60, $\dim \operatorname{Im} T = \dim \operatorname{Im} T^*$. Portanto, o posto linha e o posto coluna de $\mathbf{A}$ são iguais.

De posse destes novos conceitos vamos reavaliar a questão: sejam $T \in \mathcal{L}(V)$ e α, β bases de V. Qual a relação entre $\mathbf{A} = [T]_\alpha^\alpha = (a_{ij})$ e $\mathbf{B} = [T]_\beta^\beta = (b_{ij})$? Para isto, consideremos $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ e $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ bases de V. Então

$$T(\mathbf{u}_j) = \sum_{i=1}^{n} a_{ij}\mathbf{u}_i \text{ e } T(\mathbf{v}_j) = \sum_{i=1}^{n} b_{ij}\mathbf{v}_i.$$

Se $\alpha^* = \{f_1, \ldots, f_n\}$ e $\beta^* = \{g_1, \ldots, g_n\}$ são as bases duais, então $a_{ij} = f_i(T(\mathbf{u}_j))$ e $b_{ij} = g_i(T(\mathbf{v}_j))$. Por outro lado, seja $P \in \mathcal{L}(V)$ definido como $P(\mathbf{u}_i) = \mathbf{v}_i$. Então P é não singular e $P^*(g_i)(\mathbf{u}_i) = \delta_{ij} = f_i(\mathbf{u}_i)$, de modo que $P^*(g_i) = f_i$. Assim, $(P^*)^{-1} = (P^{-1})^*$ e $g_i = (P^{-1})^*(f_i)$. Logo, pondo $\mathbf{P} = [P]_\alpha^\beta = [I]_\beta^\alpha$, obtemos

$$b_{ij} = [(P^{-1})^*(f_i)](T(\mathbf{v}_j)) = f_i[P^{-1}T(\mathbf{v}_j)] = f_i[(P^{-1}TP)(\mathbf{u}_j)] = c_{ij}.$$

Portanto. $\mathbf{B} = \mathbf{C} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ é a nossa fómula de mudança de bases.

Exercícios

1. Seja $\alpha = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ a base canônica de F^n. Determine a base dual α^*.

2. Seja $\alpha = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ uma base de F^3, com $\mathbf{u}_1 = \mathbf{e}_1 + \mathbf{e}_3, \mathbf{u}_2 = \mathbf{e}_2 - 2\mathbf{e}_3$ e $\mathbf{u}_3 = -\mathbf{e}_1 - \mathbf{e}_2$.

   (a) Se $f \in (F^3)^*$ satisfaz $f(\mathbf{u}_1) = 1, f(\mathbf{u}_2) = -1$ e $f(\mathbf{u}_3) = 3$, então determine $f(x, y, z)$.

   (b) Exiba pelo menos um $f \in (F^3)^*$ tal que $f(\mathbf{u}_1) = 0, f(\mathbf{u}_2) = 0$ e $f(\mathbf{u}_3) \neq 0$.

   (c) Mostre que se $f \in (F^3)^*$ satisfaz $f(\mathbf{u}_1) = 0, f(\mathbf{u}_2) = 0$ e $f(\mathbf{u}_3) \neq 0$, então $f(2, 3, -1) \neq 0$.

3. Sejam $V = F^3$ e $f_i \in V^*$ definido como $f_1(\mathbf{x}) = x+2y+3z$, $f_2(\mathbf{x}) = 3x+y+2z$ e $f_3(\mathbf{x}) = 2x + 3y + z$. Mostre que $\beta^* = \{f_1, f_2, f_3\}$ é uma base de V^* exibindo a base de V da qual ela é dual.

4. Sejam $V = \mathbb{P}_2(\mathbb{R})$ e $f_i \in V^*$ definido como

   $$f_1(p(x)) = \int_0^1 p(t)dt, \quad f_2(p(x)) = \int_0^2 p(t)dt \text{ e } f_2(p(x)) = \int_0^{-1} p(t)dt.$$

   Mostre que $\beta^* = \{f_1, f_2, f_3\}$ é uma base de V^* exibindo a base de V da qual ela é dual.

5. Sejam $\alpha^* = \{f_1, \ldots, f_n\}$ uma base de V^* e $\alpha = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ qualquer lista de V tal que a matriz $\mathbf{G} = (f_i(\mathbf{v}_j))$ seja não singular. Mostre que α é uma base de V.

6. Sejam $V = \mathbb{P}_n(\mathbb{R})$ e $f_k \in V^*$ definido como $f_k(p(x)) = p^{(k)}(0)$, a k-ésima derivada de $p(x)$ avaliada em 0. Determine a base dual da base canônica.

7. Mostre que $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ é uma base de F^n, em que $\mathbf{u}_{i-1} = \mathbf{e}_i - \mathbf{e}_1$, com $i = 2, \ldots, n$, e $\mathbf{u}_n = \sum_{i=1}^n \mathbf{e}_i$, e determine a base dual α^*.

8. Mostre que qualquer $T \in \mathcal{L}(F^n, F^m)$ é da forma $T(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}))$, onde $f_1, \ldots, f_m \in (F^n)^*$.

9. Seja W o subespaço de F^4 gerado por $\mathbf{u}_1 = (1, 0, -1, 2)$ e $\mathbf{u}_1 = (2, 3, 1, 1)$. Determine W°.

10. Sejam $V = F^{2\times 2}$ e $\mathbf{B} = 2\mathbf{E}_{11} - 2\mathbf{E}_{12} - \mathbf{E}_{21} + \mathbf{E}_{22} \in V$. Para o subespaço $W = \{\mathbf{A} \in V : \mathbf{AB} = \mathbf{O}\}$, se $f \in W^\circ$ for tal que $f(\mathbf{I}) = 0$ e $f(\mathbf{E}_{22}) = 3$, então determine $f(\mathbf{B})$.

11. Sejam $V = F^n$ e $f_k \in V^*$ definido como $f_k(\mathbf{x}) = \sum_{i=1}^n (k - i)x_i$, com $n \geq 2$. Qual a dimensão do subespaço anulado por $f_1, \ldots, f_n$?

12. Sejam V um espaço vetorial sobre F e S, T subconjuntos de V.

    (a) Mostre que se $S \subseteq T$, então $T^\circ \subseteq S^\circ$.

    (b) Mostre que $S^\circ = (F[S])^\circ$.

13. Sejam U e W subespaços de V, com $\dim V = n$.

    (a) Mostre que se $U = W$ se, e somente se, $U^\circ = W^\circ$.

    (b) Mostre que $(U + W)^\circ = U^\circ \cap W^\circ$.

    (c) Mostre que $(U \cap W)^\circ = U^\circ + W^\circ$.

14. Sejam U e W subespaços de V tais que $V = U \oplus W$. Mostre que $V^* = U^\circ \oplus W^\circ$. Além disso, mostre que W^* é isomorfo a U° e U^* é isomorfo a W°.

15. Sejam V um espaço vetorial sobre F, com $\dim V = n$, e $W = V \times V^*$. Mostre que W é isomorfo a W^*.

16. Sejam V e W espaço vetoriais sobre F. Mostre que $(V \times W)^*$ é isomorfo a $V^* \times W^*$.

17. Sejam V um espaço vetorial sobre F, com $\dim V = n$, e W um subespaço de V. Mostre que se $f \in W^*$, então existe um $g \in V^*$ tal que $g(\mathbf{u}) = f(\mathbf{u})$, para todo $\mathbf{u} \in W$, isto é, $g|_W = f$.

18. Seja V um espaço vetorial sobre F. Mostre, para qualquer $f \in V^*$, com $f \neq 0$, que existe pelo menos um $\mathbf{u} \in V$ tal que $f(\mathbf{u}) = 1$. Além disso, mostre que $\mathbf{u} = \mathbf{0}$ se, e somente se, $f(\mathbf{u}) = 0$, para todo $f \in V^*$.

19. Sejam V um espaço vetorial sobre F e $f \in V^*$. Mostre que se $f(\mathbf{v}) \neq \mathbf{0}$, então $\operatorname{Im} f = F$ e $V = F[\mathbf{v}] \oplus \ker f$. Conclua que se $\dim V = n$, então $\dim \ker f = \dim V - \dim \operatorname{Im} f = n - 1$.

20. Sejam V um espaço vetorial sobre F e $f, g \in V^*$, com $f \neq 0$. Mostre que $\ker f = \ker g$ se, e somente se, $g = cf$, para algum $c \in F$, com $c \neq 0$.

21. Sejam V um espaço vetorial sobre F e $f, g \in V^*$ tais que $h : V \to F$ definida como $h(\mathbf{u}) = f(\mathbf{u})g(\mathbf{u})$ seja linear. Mostre que $f = 0$ ou $g = 0$.

22. Seja V um espaço vetorial sobre F, com $\dim V = n$. Mostre que se $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ for qualquer lista de vetores não nulos de V, então existe pelo menos um $f \in V^*$ tal que $f(\mathbf{u}_i) \neq 0$, para $i = 1, \ldots, k$.

23. Seja $V = F^{(\mathbb{N})}$ o conjunto das sequências de suporte finito. Mostre que V^* é isomorfo a $\operatorname{Seq}(F)$.

24. Sejam $V = F^{2\times 2}$, $\mathbf{A} \in V$ fixada e $T \in \mathcal{L}(V)$ definido como $T(\mathbf{X}) = \mathbf{AX}$. Mostre que $\operatorname{tr}(T) = 2\operatorname{tr}(\mathbf{A})$.

25. Sejam V um espaço vetorial sobre F e $f_1, \ldots, f_k, g \in V^*$. Mostre que $\cap_{i=1}^k \ker f_i \subseteq \ker g$ se, e somente se, $g \in F[f_1, \ldots, f_k]$.

26. Seja $f_i \in (F^n)^*$, com $i = 1, \ldots, n$. Mostre que se as projeções de coordenadas $p_i \in F[f_1, \ldots, f_n]$, então o operador $T \in \mathcal{L}(F^n)$ definido como $T(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_n(\mathbf{x}))$ é não singular.

27. Sejam V um espaço vetorial sobre $\mathbb{R}$, com $\dim V = n$, U um hiperplano de V e $L_{\mathbf{uv}} = \{t\mathbf{u}+(1-t)\mathbf{v} : t \in [0,1]\}$, para todos $\mathbf{u}, \mathbf{v} \in V$. Dados $\mathbf{u}, \mathbf{v} \in S = V - U$, definimos $\mathbf{u} \sim \mathbf{v}$ se, e somente se, $U \cap L_{\mathbf{uv}} = \emptyset$. Mostre que "$\sim$" é uma relação de equivalência sobre S e que ela possui exatamente duas classes laterais.

28. Sejam α^* uma lista sobre $(\mathbb{R}^n)^*$ e $S = \{\mathbf{x} \in \mathbb{R}^n : f_i(\mathbf{x}) \geq 0, f_i \in \alpha^*\}$. Mostre que $\mathbf{x} + t\mathbf{y} \in S$, para todos $\mathbf{x}, \mathbf{y} \in S$ e $t \in \mathbb{R}$, com $t \geq 0$. Conclua que se α^* for uma base, então $S = \{x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n : x_i \in \mathbb{R}_+\}$, com $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ uma base de $\mathbb{R}^n$ cuja dual é α^*.

29. Sejam V um espaço vetorial sobre F e $T \in \mathcal{L}(V)$. Mostre que se $T \circ S = S \circ T$, para todo $S \in \mathcal{L}(V)$, então $T = cI_V$, para algum $c \in F$.

30. Seja $U = \{(x_1, \ldots, x_n) \in F^n : x_1 + \cdots + x_n = 0\}$ um subespaço de F^n.

    (a) Mostre que $U^\circ = F[x_1 + \cdots + x_n]$.

    (b) Mostre que U^* pode ser identificado com o conjunto dos funcionais lineares sobre F^n da forma $f_{\mathbf{c}}(\mathbf{x}) = \sum_{i=1}^n c_i x_i$, com $c_1 + \cdots + c_n = 0$.

31. Sejam U um subespaço de V e $\mathbf{v} \in V - U$. Mostre, para cada $f \in U^*$, que existe um $g \in V^*$ tal que $g(v) = 1$ e $g(u) = f(u)$, para todo $\mathbf{u} \in U$.

32. Seja V um espaço vetorial sobre F, com $\dim V = n$, U qualquer subespaço de V. Mostre que se $\{f_1, \ldots, f_k\}$ for uma base de U°, então $U = \cap_{i=1}^k \ker f_i$.

33. Seja U qualquer subespaço de $V = \mathcal{F}(S, F)$, com $\dim U = n$. Mostre que existem $x_1, \ldots, x_n \in S$ e $f_1, \ldots, f_n \in U$ tais que $f_j(x_i) = \delta_{ij}$.

34. Sejam U e W subespaços de V tais que $V = U \oplus W$ e $\dim V = n$. Mostre que se $P \in \mathcal{L}(V)$ for a projeção de V sobre U paralela a W, então $P^* \in \mathcal{L}(V^*)$ é a projeção de V^* sobre W° paralela a U°.

35. Sejam $V = \mathbb{R}[x]$ e $f \in V^*$ definido como

$$f(p(x)) = \int_a^b p(t)dt.$$

Se $D \in \mathcal{L}(V)$ definido como $D(p(x)) = p'(x)$, "o operador diferencial", o que é D^*f?

36. Sejam $V = F^{n\times n}$ e $\mathbf{B} \in V$ fixada. Se $T \in \mathcal{L}(V)$ é definido como $T(\mathbf{X}) = \mathbf{XB} - \mathbf{BX}$ e $f = \operatorname{tr} \in V^*$, o que é T^*f? Conclua que $f \in \ker T^*$.

37. Sejam V um espaço vetorial sobre F, com $\dim V = n$, e $T \in \mathcal{L}(V)$. Mostre que se $c \in F$ e existir um $\mathbf{v} \in V$, com $\mathbf{v} \neq \mathbf{0}$, tal que $T(\mathbf{v}) = c\mathbf{v}$, então existe um $f \in V^*$, com $f \neq 0$, tal que $T^*(f) = cf$.

38. Sejam $V = \mathbb{P}_n(\mathbb{R})$ e $D \in \mathcal{L}(V)$ definido como $D(p(x)) = p'(x)$. Determine uma base de $\ker D^*$.

39. Sejam $T \in \mathcal{L}(V, W)$, com $\dim V = n, \dim W = m$, e $\mathbf{b} \in W$. Então uma e apenas uma das condições ocorre:

    (a) $\mathbf{b} \in \operatorname{Im} T$, ou seja, a equação $T(\mathbf{x}) = \mathbf{b}$ possui uma solução.

    (b) Existe um $g \in W^*$ tal que $T^*(g) = 0$ e $g(\mathbf{b}) = 1$.

    O que significa isto em termos da equação linear $\mathbf{AX} = \mathbf{B}$?

40. Sejam V e W espaços vetoriais sobre F. Mostre que a função $\sigma : \mathcal{L}(V, W) \to \mathcal{L}(W^*, V^*)$ definida como $\sigma(T) = T^*$ é injetora. Conclua que se $\dim V = n$ e $\dim W = m$, então σ é um isomorfismo.

5

Formas Canônicas Elementares

Já estudamos o nosso primeiro problema da álgebra linear, o qual foi entender e estudar como resolver o sistema linear $\mathbf{AX} = \mathbf{B}$ e o que significa a solução. Nosso objetivo neste capítulo é iniciar o estudo do nosso segundo problema: $\mathbf{AX} = \lambda\mathbf{X}$, o qual representa outra ferramenta crucial da álgebra linear, usando ideias e técnicas completamente diferentes das vistas até o presente. Mais precisamente, sejam V um espaço vetorial de dimensão finita sobre F e $T \in \mathcal{L}(V)$. Então vamos determinar uma base de V, em relação à qual, a matriz de T tenha uma forma a mais simples possível.

5.1 Autovalores e Autovetores

Vimos, no Exemplo 4.14, que o operador reflexão R sobre $\mathbb{R}^2$ em torno de uma reta $y = 2x$ satisfazia $R(1,2) = (1,2)$ e $R(-2,1) = -(-2,1)$, ou seja, se $U = \mathbb{R}[(1,2)]$ e $V = \mathbb{R}[(-2,1)]$ são subespaços de $\mathbb{R}^2$, então $R(U) \subseteq U$ e $R(V) \subseteq V$. Isto motiva a seguinte definição.

Sejam U um subespaço de V e $T \in \mathcal{L}(V)$. Diremos que U é *invariante* sob T se $T(U) \subseteq U$, ou seja, $T(\mathbf{u}) \in U$, para todo $\mathbf{u} \in U$ ou, equivalentemente, T induz um operador $S \in \mathcal{L}(U)$ definido como $S(\mathbf{u}) = T(\mathbf{u})$, o qual chama-se *função restrição* e denotamos por $S = T|_U$. É fácil verificar que $\{\mathbf{0}\}, V, \ker T$ e $\operatorname{Im} T$ são subespaços invariantes sob T. Mais tarde nos dedicaremos ao estudo detalhado dos subespaços invariantes. Seja $\mathbf{u} \in V$, com $\mathbf{u} \neq \mathbf{0}$. Então o conjunto $U = F[\mathbf{u}]$ é um subespaço de V, com $\dim U = 1$. Reciprocamente, qualquer subespaço de V de dimensão um é desta forma.

Lema 5.1 *Sejam U um subespaço de V, com $\dim U = 1$, e $T \in \mathcal{L}(V)$. Então U é invariante sob T se, e somente se, para qualquer $\mathbf{u} \in U$, com $\mathbf{u} \neq \mathbf{0}$, existir um $\lambda \in F$ (independente de $\mathbf{u}$) tal que $T(\mathbf{u}) = \lambda\mathbf{u}$.*

Prova. Suponhamos que $U = F[\mathbf{u}_0]$, com $\mathbf{u}_0 \neq \mathbf{0}$, seja invariante sob T. Então para qualquer $\mathbf{u} \in U$, com $\mathbf{u} \neq \mathbf{0}$, temos que $T(\mathbf{u}) \in U$. Assim, existem $c_0, c \in F$, com $c_0 \neq 0$, tais que $\mathbf{u} = c_0\mathbf{u}_0$ e $T(\mathbf{u}) = c\mathbf{u}_0$, de modo que existe um $\lambda = c_0^{-1}c \in F$ tal que $T(\mathbf{u}) = T(c_0\mathbf{u}_0) = c\mathbf{u}_0 = c_0^{-1}c(c_0\mathbf{u}_0) = \lambda\mathbf{u}$. Se $\mathbf{v}$ for outro vetor de U, então $\mathbf{v} = d\mathbf{u}_0$, para algum $d \in F$. Portanto, $T(\mathbf{v}) = dT(\mathbf{u}_0) = d(\lambda\mathbf{u}_0) = \lambda\mathbf{v}$ implica a unicidade de λ. A recíproca é clara. ■

Sejam V um espaço vetorial sobre F e $T \in \mathcal{L}(V)$. Um escalar $\lambda \in F$ é um *autovalor* de T se existir um $\mathbf{v} \in V$, com $\mathbf{v} \neq \mathbf{0}$, tal que

$$T(\mathbf{v}) = \lambda\mathbf{v}. \tag{5.1}$$

O vetor $\mathbf{v}$ chama-se um *autovetor* ou *autovetor à direita* de T associado ao autovalor λ. Portanto, a existência de subespaços invariantes sob T de dimensão 1 é equivalente a existência de autovetores de T. Observe, pelo Lema 5.1, que existe um único autovalor associado a um dado autovetor $\mathbf{v} \neq \mathbf{0}$. Por esta razão o vetor $\mathbf{0}$ nunca será um autovetor. O conjunto

$$V_\lambda = \{\mathbf{v} \in V : T(\mathbf{v}) = \lambda\mathbf{v}\} = \ker(T - \lambda I_V)$$

é um subespaço de V. Mais geralmente, o conjunto

$$V^\lambda = \{\mathbf{v} \in V : (T - \lambda I_V)^k(\mathbf{v}) = \mathbf{0},\ \ \exists\ k \in \mathbb{N}\} = \cup_{n\in\mathbb{N}} \ker(T - \lambda I_V)^n$$

é um subespaço de V contendo V_λ. De fato. Dados $\mathbf{u}, \mathbf{v} \in V^\lambda$ e $c \in F$, existem $k, m \in \mathbb{N}$ tais que $(T - \lambda I)^k(\mathbf{u}) = \mathbf{0}$ e $(T - \lambda I)^m(\mathbf{v}) = \mathbf{0}$. Pondo $n = k + m \in \mathbb{N}$ ou $n = \max\{k, m\} \in \mathbb{N}$, obtemos

$$\begin{aligned}(T - \lambda I)^n(\mathbf{u} + c\mathbf{v}) &= (T - \lambda I)^n(\mathbf{u}) + c(T - \lambda I)^n(\mathbf{v}) \\ &= (T - \lambda I)^m((T - \lambda I)^k(\mathbf{u})) + \\ &\quad\ c(T - \lambda I)^k((T - \lambda I)^m(\mathbf{v})) \\ &= (T - \lambda I)^m(\mathbf{0}) + c(T - \lambda I)^k(\mathbf{0}) = \mathbf{0}.\end{aligned}$$

Portanto, $\mathbf{u}+c\mathbf{v} \in V^{\lambda}$. Observe que se $V_{\lambda} \neq \{\mathbf{0}\}$, então λ é um autovalor de T e diremos que V_{λ} é o *autoespaço* de T associado ao autovalor λ e V^{λ} é o *autoespaço generalizado* de T associado ao autovalor λ. É muito importante observar que $\ker(T-\lambda I)^n \subseteq \ker(T-\lambda I)^{n+1}$, para todo $n \in \mathbb{N}$. Em particular, $\mathbf{v} \in \ker(T-\lambda I)^{n+1}$ se, e somente se, $(T-\lambda I)(\mathbf{v}) \in \ker(T-\lambda I)^n$.

Teorema 5.2 *Sejam V um espaço vetorial sobre $F, T \in \mathcal{L}(V)$ e $\lambda \in F$. Então as seguintes condições são equivalentes*:

1. *λ é um autovalor de T*;
2. *$T-\lambda I_V \in \mathcal{L}(V)$ é singular, isto é, $V_{\lambda} \neq \{\mathbf{0}\}$*;
3. *$V^{\lambda} \neq \{\mathbf{0}\}$.*

Prova. É clara que (1) é equivalente a (2). Como $V_{\lambda} \subseteq V^{\lambda}$ temos que $V^{\lambda} \neq \{\mathbf{0}\}$. Reciprocamente, seja $\mathbf{u} \in V^{\lambda}$, com $\mathbf{u} \neq \mathbf{0}$. Então existe um $m \in \mathbb{N}$ tal que $(T-\lambda I)^m(\mathbf{u}) = \mathbf{0}$ e o conjunto $S = \{p \in \mathbb{N} : (T-\lambda I)^p(\mathbf{u}) = \mathbf{0}\}$ é não vazio. Assim, pelo PBO, existe um $k \in \mathbb{N}$ tal que $(T-\lambda I)^k(\mathbf{u}) = \mathbf{0}$, mas $(T-\lambda I)^{k-1}(\mathbf{u}) \neq \mathbf{0}$ e diremos que k é o *índice* de V^{λ}. Portanto, $\mathbf{v} = (T-\lambda I)^{k-1}(\mathbf{u}) \in V_{\lambda}$. ■

Exemplo 5.3 *Seja $R \in \mathcal{L}(\mathbb{R}^2)$ definido como $R(x,y) = (-y,x)$, o operador rotação por um ângulo de $90°$, confira Exemplo 4.7. Determine, se existirem, os autovalores e autovetores de R.*

Solução. Um escalar $\lambda \in \mathbb{R}$ é um autovalor de R se, e somente se, existir um $\mathbf{v} = (x,y) \in \mathbb{R}^2$, com $\mathbf{v} \neq \mathbf{0}$, tal que $R(\mathbf{v}) = \lambda\mathbf{v}$. Mas, isto é equivalente as equações $-y = \lambda x$ e $x = \lambda y$, de modo que $-y = \lambda^2 y$. Como $y \neq 0$ (por que?) temos que a equação $\lambda^2+1 = 0$ não possui raízes sobre $\mathbb{R}$. Portanto, R não possui autovalores e nem autovetores. Não obstante, essa equação possui duas raízes sobre $\mathbb{C}$, a saber, $\lambda_1 = -i$ e $\lambda_2 = i$. Neste caso, $\mathbf{v}_1 = (1,i)$ é um autovetor de R associado a λ_1 e $\mathbf{v}_2 = (1,-i)$ é um autovetor de R associado a λ_2. Veremos, em geral, que se V for um espaço vetorial de dimensão finita sobre $\mathbb{C}$, então qualquer $T \in \mathcal{L}(V)$ possui pelo menos um autovetor. Portanto, os autovalores dependem do corpo básico F do espaço vetorial V. ■

Sejam $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ uma base de V, $T \in \mathcal{L}(V)$ e $\mathbf{A} = [T]_{\alpha}^{\alpha} = (a_{ij})$. Seja $\mathbf{v} \in V$, com $\mathbf{v} \neq \mathbf{0}$, tal que $T(\mathbf{v}) = \lambda\mathbf{v}$ e $\mathbf{X} = [\mathbf{v}]_{\alpha} = (x_1, \ldots, x_n)^t$. Então, pelo Teorema 4.36,

$$[T(\mathbf{v})]_{\alpha} = [T]_{\alpha}^{\alpha}[\mathbf{v}]_{\alpha},$$

de modo que $\lambda\mathbf{X} = \mathbf{AX}$ se, e somente se, $(\lambda\mathbf{I}_n - \mathbf{A})\mathbf{X} = 0$. Portanto, o problema de determinar os autovalores e autovetores de T é reduzido à de encontrar os λ para os quais a matriz $\lambda\mathbf{I}_n - \mathbf{A}$ seja singular, ou seja, $\det(\lambda\mathbf{I}_n - \mathbf{A}) = 0$. Posto de outra forma: o sistema homogêneo

$$(\lambda - a_{ii})x_i = \sum_{j\neq i} a_{ij}x_j \Leftrightarrow \sum_{j=1}^{n}(\lambda\delta_{ij} - a_{ij})x_j = 0, i = 1, \ldots, n, \tag{5.2}$$

possui uma solução não nula $\mathbf{X} \neq \mathbf{O}$, ou seja, os x_j nem todos são nulos. Observe que $\lambda\delta_{ij} - a_{ij} = -a_{ij}$, se $i \neq j$, e $\lambda\delta_{ij} - a_{ij} = \lambda - a_{ij}$, se $i = j$, implicam que

$$\det(\lambda\mathbf{I}_n - \mathbf{A}) = \prod_{i=1}^{n}(\lambda - a_{ii}) + f_{n-2}(\lambda),$$

é um polinômio de grau n em λ, em que $f_{n-2}(\lambda)$ é um polinômio de grau no máximo $n - 2$ em λ ou em forma expandida: a equação polinomial em λ, a saber,

$$\lambda^n + b_1\lambda^{n-1} + b_2\lambda^{n-2} + \cdots + b_{n-1}\lambda + b_n = 0,$$

em que os coeficientes b_k são determinados em função das entradas de $\mathbf{A}$. Por exemplo, para $n = 3$, é fácil verificar que

$$b_1 = -\operatorname{tr}(\mathbf{A}),\ \ b_2 = \operatorname{tr}(\operatorname{adj}(\mathbf{A})) = \sum_{i=1}^{3}\det(\mathbf{A}_{ii}) \ \text{ e } \ b_3 = -\det\mathbf{A}$$

A matriz $x\mathbf{I} - \mathbf{A}$ chama-se *matriz característica* de $\mathbf{A}$ e $f_{\mathbf{A}}(x) = \det(x\mathbf{I} - \mathbf{A})$ chama-se *polinômio característico* de $\mathbf{A}$. A equação polinomial $f_{\mathbf{A}}(x) = 0$ chama-se *equação característica* de $\mathbf{A}$ e as raízes desta equação, dadas pelo Teorema 1.2, são os autovalores de $\mathbf{A}$. À primeira vista, poderia parecer da definição de polinômio característico de T, que ele fosse dependente da matriz $\mathbf{A}$. Mas, isto não é verdade, pois se β for outra base de V e $\mathbf{B} = [T]_\beta^\beta$, então já sabemos que $\mathbf{B}$ é semelhante a $\mathbf{A}$. Assim, vamos provar o seguinte resultado:

Lema 5.4 *Matrizes semelhantes possuem o mesmo polinômio característico.*

Prova. Sejam $\mathbf{A}$ e $\mathbf{B}$ matrizes semelhantes. Então existe uma matriz não singular $\mathbf{P}$ tal que $\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}$. Assim, pela Fórmula de Binet-Cauchy,

$$\det(x\mathbf{I}_n - \mathbf{B}) = \det[\mathbf{P}^{-1}(x\mathbf{I}_n - \mathbf{A})\mathbf{P}] = \det(x\mathbf{I}_n - \mathbf{A}).$$

Portanto, $f_{\mathbf{B}}(x) = f_{\mathbf{A}}(x)$. ■

Observe que a recíproca do Lema 5.4 é falsa, pois as matrizes $\mathbf{A} = \mathbf{I}_2$ e $\mathbf{B} = \mathbf{I}_2 + \mathbf{E}_{12}$ possuem o mesmo polinômio característico $f(x) = (x-1)^2$. Mas, claramente $\mathbf{B} \neq \mathbf{P}^{-1}\mathbf{AP}$, para toda matriz não singular $\mathbf{P}$. Não obstante, matrizes semelhantes podem possuir autovetores diferentes associados ao mesmo autovalor.

Sejam V um espaço vetorial sobre F, com $\dim V = n$, e $T \in \mathcal{L}(V)$. O *polinômio característico* de T é o polinômio característico de qualquer representação matricial de T em relação a alguma base de V. Neste contexto, seja $\mathbf{A} \in F^{n\times n}$. Diremos que $\lambda \in F$ é um *autovalor* de $\mathbf{A}$ quando λ for um autovalor do operador induzido $T_{\mathbf{A}} \in \mathcal{L}(F^n)$ e $\mathbf{X}$ um *autovetor* de $\mathbf{A}$ associado a λ.

Teorema 5.5 *Sejam α uma base de V, $T \in \mathcal{L}(V)$ e $\mathbf{A} = [T]_\alpha^\alpha = (a_{ij}) \in F^{n\times n}$. Se $f_{\mathbf{A}}(x) = \sum_{i=0}^{n-1} c_i x^i + x^n$ for o polinômio característico de T, então cada coeficiente c_r, com $r = 0, \ldots, n-1$, é igual a $(-1)^{n-r}$ vezes a soma de todos os menores principais de $\mathbf{A}$ de ordem $n-r$. Conclua que $\rho(T) = \max\{n - r : c_r \neq 0\}$.*

Prova. Sejam $\mathbf{C}_1, \ldots, \mathbf{C}_n$ as colunas de $\mathbf{A}$ e $\mathbf{E}_1, \ldots, \mathbf{E}_n$ as colunas de $\mathbf{I}_n$. Então

$$(-1)^n f_{\mathbf{A}}(x) = \det(\mathbf{A} - x\mathbf{I}_n) = \det\begin{pmatrix} \mathbf{C}_1 - x\mathbf{E}_1 & \cdots & \mathbf{C}_n - x\mathbf{E}_n \end{pmatrix}.$$

Assim, aplicando sucessivamente o item (1) do Teorema 1.24, obtemos

$$(-1)^n f_{\mathbf{A}}(x) = \det(-x\mathbf{I}) + \sum_{k=1}^{n} \sum_{1\leq i_1 < i_2 < \cdots < i_k \leq n} \mathbf{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}.$$

Mais explicitamente, o determinante como a soma de 2^n determinantes de matrizes tendo colunas $\mathbf{C}_j$ ou $-x\mathbf{E}_j$. Logo, agrupando os determinantes envolvendo r colunas da forma $-x\mathbf{E}_j$, para algum j, temos que existem exatamente $\binom{n}{k}$ deles, obtidos substituindo-se as r colunas de $\mathbf{A}$ por $(-x)\mathbf{E}_j$, de todas as maneiras possíveis, e as $n-r$ colunas de $\mathbf{A}$ inalteradas. Portanto, pelo Exemplo 1.31, cada um destes determinantes é da forma $(-x)^{n-r}$ vezes um menor principal de $\mathbf{A}$ de ordem $n-r$. Por outro lado, cada menor principal de ordem $n-r$ aparece no somatório e o resultado segue, indutivamente, fazendo $r = 0, 1, \ldots, n-1$. ■

É muito instrutivo comprovar a prova do Teorema 5.5 para valores pequeno de n, digamos $n = 2$,

$$\begin{array}{rcl} f_{\mathbf{A}}(x) & = & \det\begin{pmatrix} \mathbf{C}_1 - x\mathbf{E}_1 & \mathbf{C}_2 - x\mathbf{E}_2 \end{pmatrix} \\ & = & \det\begin{pmatrix} -x\mathbf{E}_1 & -x\mathbf{E}_2 \end{pmatrix} + \det\begin{pmatrix} -x\mathbf{E}_1 & \mathbf{C}_2 \end{pmatrix} \\ & + & \det\begin{pmatrix} \mathbf{C}_1 & -x\mathbf{E}_2 \end{pmatrix} + \det\begin{pmatrix} \mathbf{C}_1 & \mathbf{C}_2 \end{pmatrix} \\ & = & x^2 - \operatorname{tr}(\mathbf{A})x + \det \mathbf{A}. \end{array}$$

Observe, pondo $b_{n-r} = c_r$, com $r = 0, \ldots, n-1$, que

$$f_{\mathbf{A}}(x) = x^n + b_1 x^{n-1} + b_2 x^{n-2} + \cdots + b_{n-1}x + b_n.$$

Em particular,

$$b_1 = -\operatorname{tr}(\mathbf{A}), \;\; b_2 = \sum_{1 \leq i < j \leq n} \det\begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{pmatrix} \;\text{ e }\; b_n = (-1)^n \det(\mathbf{A}).$$

Note que $b_n = f_{\mathbf{A}}(0)$. É muito importante, de um ponto de vista teórico e didático, o seguinte resultado: sejam $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ os autovalores de $T \in \mathcal{L}(V)$, não necessariamente distintos. Então, pelo Lema 1.4,

$$b_k = (-1)^k E_k(\lambda_1, \ldots, \lambda_n), k = 1, \ldots, n.$$

Além disso, vale ressaltar que para determinar o autoespaço V_λ (autovetores) de V associado ao autovalor λ, devemos resolver o sistema homogêno dado pelas equações (5.2). Portanto, um método alternativo de obter o autoespaço é escalonando a matriz:

$$\left(\begin{array}{c|c} (\lambda\mathbf{I}_n - \mathbf{A})^t & \mathbf{I}_n \end{array} \right),$$

confira o Teorema 4.39.

Exemplo 5.6 *Seja $T \in \mathcal{L}(\mathbb{R}^3)$ definido como $T(x, y, z) = (4x + 2y, -x + y, y + 2z)$. Determine os autovalores e autovetores de T.*

Solução. É fácil verificar que

$$\mathbf{A} = [T] = \begin{pmatrix} 4 & 2 & 0 \\ -1 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix} \;\text{ e }\; f_{\mathbf{A}}(x) = x^3 - 7x^2 + 16x - 12 \in \mathbb{R}[x].$$

Logo, pelo Exemplo 1.3, as raízes de $f_{\mathbf{A}}(x)$ são: $\lambda_1 = 2$ e $\lambda_2 = 3$, ou seja, $\lambda_1 = 2$ e $\lambda_2 = 3$ são os autovalores de T. Para obter o autoespaço V_{λ_1}, devemos escalonar a matriz

$$\left(\begin{array}{ccc|ccc} -2 & 1 & 0 & 1 & 0 & 0 \\ -2 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right) \longrightarrow \cdots \longrightarrow \left(\begin{array}{ccc|ccc} 1 & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right).$$

Portanto, $\rho(\mathbf{A} - \lambda_1\mathbf{I}_3) = 2$ e $V_{\lambda_1} = \mathbb{R}[(0,0,1)]$. De modo análogo, temos que $\rho(\mathbf{A} - \lambda_2\mathbf{I}_3) = 2$ e $V_{\lambda_2} = \mathbb{R}[(-2,1,1)]$. ■

Seja $\lambda \in \mathbb{C}$ uma raiz do polinômio $f(x) \in F[x]$. Diremos que λ possui *multiplicidade algébrica* m se $(x-\lambda)^m$ é um fator de $f(x)$, mas $(x-\lambda)^{m+1}$ não e denotamos por $m_a(\lambda) = m$, isto é, $f(x) = (x-\lambda)^m g(x)$, $g(\lambda) \neq 0$. Observe que

$$m = \max\{k \in N : (x-\lambda)^k \text{ é um fator de } f(x)\} \text{ e } m \leq \partial(f).$$

A dimensão do autoespaço $V_\lambda = \ker(T - \lambda I)$ chama-se *multiplicidade geométrica* de λ e denotamos por $m_g(\lambda) = \dim V_\lambda$. Note, no Exemplo 5.6, que $f_{\mathbf{A}}(x) = (x-2)^2(x-3)$, de modo que $m_a(2) = 2$ e $m_a(3) = 1$. Além disso, $m_g(2) = 1$ e $m_g(3) = 1$. Observe que $m_g(2) < m_a(2)$. Por outro lado, se $T = I_V \in \mathcal{L}(V)$, então $T(\mathbf{v}) = 1 \cdot \mathbf{v}$, para todo $\mathbf{v} \in V$. Portanto, 1 é o único autovalor de T, de modo que $V_1 = V$ e $m_g(1) = \dim V = m_a(1)$. Provaremos, em geral, que $m_g(\lambda) \leq m_a(\lambda)$.

Teorema 5.7 *Sejam $T \in \mathcal{L}(V)$, com $\dim V = n$, e $\lambda \in F$. Se λ for um autovalor de T, então $m_g(\lambda) \leq m_a(\lambda) = \dim V^\lambda$.*

Prova. Pelo Teorema 5.5, existe um $\mathbf{v} \in V^\lambda$, com $\mathbf{v} \neq \mathbf{0}$, de modo que podemos escolher um menor $k \in \mathbb{N}$ tal que $(T-\lambda I)^k(\mathbf{v}) = \mathbf{0}$, mas $(T-\lambda I)^{k-1}(\mathbf{v}) \neq \mathbf{0}$. Pondo $\mathbf{v}_j = (T-\lambda I)^{k-j}(\mathbf{v})$, para $j = 1, \ldots, k$, é fácil verificar que $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ é uma base de V^λ, a qual é parte de uma base $\alpha = \{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ de V. Como $(T-\lambda I)(\mathbf{v}_j) = \mathbf{v}_{j-1}$, para $j = 2, \ldots, k$, mais explicitamente,

$$\left\{\begin{array}{rcl} T(\mathbf{v}_1) & = & \lambda\mathbf{v}_1 \\ T(\mathbf{v}_j) & = & \lambda\mathbf{v}_j + \mathbf{v}_{j-1}, j = 2, \ldots, k, \\ T(\mathbf{v}_{k+j}) & = & \sum_{i=1}^n a_{i(k+j)}\mathbf{v}_i, j = 1, \ldots, n-k, \end{array}\right. \tag{5.3}$$

temos que

$$\mathbf{A} = [T]_\alpha^\alpha = \begin{pmatrix} \mathbf{J}_k & \mathbf{B} \\ \mathbf{O} & \mathbf{C} \end{pmatrix}, \text{ com } \mathbf{J}_k = \lambda \mathbf{I}_k + \mathbf{E}_{12} + \cdots + \mathbf{E}_{(k-1)k} \in F^{k\times k},$$

$\mathbf{B} \in F^{k\times(n-k)}$ e $\mathbf{C} \in F^{(n-k)\times(n-k)}$. Logo, pelo item (10) do Teorema 1.24,

$$f_\mathbf{A}(x) = (x-\lambda)^k h(x),$$

em que $h(x) = \det(x\mathbf{I}_{n-k} - \mathbf{C})$ é um polinômio de grau $n-k$. Note que λ é o único autovalor de T que satisfaz as equações (5.3), pois se μ for outro autovalor de T, então $T^m(\mathbf{v}) = \mu^m \mathbf{v}$, para todo $m \in \mathbb{N}$, e

$$\mathbf{0} = (\sum_{j=1}^{k} \binom{k}{j} T^j(-\lambda I)^{k-j})(\mathbf{v}) = \sum_{j=1}^{k} \binom{k}{j} (-\lambda)^{k-j}\mu^j \mathbf{v} = (\mu-\lambda)^k \mathbf{v},$$

de modo que $\lambda = \mu$. Finalmente, $(T-\lambda I)(\mathbf{u}) \neq \mathbf{0}$, para todo $\mathbf{u} \in V - V^\lambda$, pois se $(T-\lambda I)(\mathbf{u}) = \mathbf{0}$, então $(T-\lambda I)(\mathbf{u}) \in V^\lambda$. Assim, existe um $r \in \mathbb{N}$ tal que

$$(T-\lambda I)^{r+1}(\mathbf{u}) = (T-\lambda I)^r(T-\lambda I)(\mathbf{u}) = (T-\lambda I)^r(\mathbf{0}) = \mathbf{0},$$

isto é, $\mathbf{u} \in V^\lambda$, o que é impossível. Portanto, $h(\lambda) \neq 0$ e $m_a(\lambda) = k = \dim V^\lambda$. A lista $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ chama-se *cadeia de Jordan*[1] de comprimento k associada ao autovalor λ ou ao autovetor $\mathbf{v}_1$. ■

Já sabemos, na prova do Teorema 5.7, que a partir de $\mathbf{v}_1 \in V_\lambda$, para obter os autovetores generalizados $\mathbf{v}_2, \ldots, \mathbf{v}_k \in V^\lambda$ devemos, recursivamente, resolver $k-1$ equações:

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{X}_j = \mathbf{X}_{j-1}, j = 2, \ldots, k, \tag{5.4}$$

com $(T - \lambda \mathbf{I})^m(\mathbf{v}_{j-1}) = \mathbf{0}$, mas $(T - \lambda \mathbf{I})^{m-1}(\mathbf{v}_{j-1}) \neq \mathbf{0}$. Vamos apresentar outro modo de obter o autoespaço V_λ e o autoespaço generalizado V^λ quando $\lambda \in \mathbb{R}$ é um autovalor de T. Se λ não for um autovalor de T, então $f_\mathbf{A}(\lambda) \neq 0$ e a equação

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{X} = \mathbf{B}, \text{ com } \mathbf{B} = (f_\mathbf{A}(\lambda), 0, \ldots, 0)^t, \tag{5.5}$$

possui, pela Regra de Cramer, uma única solução $\mathbf{X}(\lambda)$, a qual depende de λ, isto é, suas componentes são polinômios em λ. Observe que se os autovalores λ possuem $m_a(\lambda) =$

[1] Marie Ennemond Camille Jordan, 1838-1922, matemático francês.

1, então os únicos autovetores $\mathbf{v}$ associados a eles são obtidos resolvendo a equação (5.5). Neste caso, $f_{\mathbf{A}}(\lambda)$ pode ocorrer em qualquer uma das n componentes de $\mathbf{B}$, comprove isto no Exemplo 5.8. Por outro lado, se λ for um autovalor, com $m_a(\lambda) = k$ e $\rho(\mathbf{A} - \lambda\mathbf{I}) = n - 1$ ou $m_g(\lambda) = 1$, então, pela prova do Teorema 5.7, existe um $\mathbf{v}_1 \in V_\lambda$ e $k - 1$ vetores generalizados $\mathbf{v}_2, \ldots, \mathbf{v}_k \in V^\lambda$, todos satisfazendo a equação (5.5) e $k - 1$ equações obtidas dela por sua derivação até ordem $k - 1$ em relação a λ. Como $f_{\mathbf{A}}^{(j)}(\lambda) = 0$, para $j = 1, \ldots, k - 1$, temos, depois de alguns cálculos e comparações com as equações (5.4), que

$$[\mathbf{v}_j] = \frac{1}{(j-1)!}\mathbf{X}^{(j-1)}(\lambda), j = 2, \ldots, k,$$

de modo que $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ é uma base de V^λ. Em geral, se $\lambda \in \mathbb{R}$ for um autovalor de T, com $m_a(\lambda) = k$ e $\rho(\mathbf{A} - \lambda\mathbf{I}) = n - m$ ou $m_g(\lambda) = m$, com $m > 1$, então existem $\mathbf{v}_1, \ldots, \mathbf{v}_m \in V_\lambda$ e, a partir de um $\mathbf{v}_i$, existem $k - m$ autovetores generalizados $\mathbf{v}_{m+1}, \ldots, \mathbf{v}_k \in V^\lambda$, de modo que $\{\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_k\}$ é uma base de V^λ. Vamos aplicar esta técnica ao Exemplo 5.6. Já vimos que $f_{\mathbf{A}}(x) = (x-2)^2(x-3)$ e que $\lambda_1 = 2$ e $\lambda_2 = 3$ são os autovalores de T. Note que para obter a solução da equação (5.5) basta escalonar a matriz:

$$\left(\begin{array}{ccc|c} 4-\lambda & 2 & 0 & f_{\mathbf{A}}(\lambda) \\ -1 & 1-\lambda & 0 & 0 \\ 0 & 1 & 2-\lambda & 0 \end{array}\right) \to \cdots \to \left(\begin{array}{ccc|c} 1 & 0 & 0 & -\lambda^2 + 3\lambda - 2 \\ 0 & 1 & 0 & \lambda - 2 \\ 0 & 0 & 1 & 1 \end{array}\right).$$

Assim, $\mathbf{X}(\lambda) = (-\lambda^2 + 3\lambda - 2, \lambda - 2, 1)^t$ é a única solução. Como $m_a(\lambda_1) = 2$ e $\rho(\mathbf{A} - \lambda_1\mathbf{I}) = 2$ ou $m_g(\lambda_1) = 1$ temos que $[\mathbf{v}_1] = \mathbf{X}(\lambda_1) = (0, 0, 1)^t$ é o único autovetor de T. Logo, existe um autovetor generalizado $\mathbf{v}_2 \in V^{\lambda_1}$, a saber,

$$[\mathbf{v}_2] = \mathbf{X}'(\lambda_1) = (-1, 1, 0)^t.$$

Portanto, $\{\mathbf{v}_1, \mathbf{v}_2\}$ é uma base de V^{λ_1} e $[\mathbf{v}_3] = \mathbf{X}(\lambda_2) = (-2, 1, 1)^t$ é o único autovetor de T associado a λ_2. Consequentemente, $\mathbb{R}^3 = V^{\lambda_1} \oplus V^{\lambda_2}$ e

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \mathbf{J}, \text{ com } \mathbf{P} = [P]_\beta^\alpha = \begin{pmatrix} 0 & -1 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

a matriz de transição da base canônica α para base $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

Finalizaremos esta seção apresentando mais um método de obter o autoespaço

V_λ, quando todo autovalor $\lambda \in \mathbb{R}$ de T é tal que $m_a(\lambda) = 1$. Sejam α uma base de V, $T \in \mathcal{L}(V)$, $\mathbf{A} = [T]_\alpha^\alpha = (a_{ij})$ e $\lambda \in \mathbb{R}$ um autovalor de T. Então, pelo Teorema 1.26,

$$(\lambda \mathbf{I}_n - \mathbf{A})\,\text{adj}(\lambda \mathbf{I}_n - \mathbf{A}) = \det(\lambda \mathbf{I}_n - \mathbf{A})\mathbf{I}_n = \mathbf{O}$$

ou $\mathbf{A}\,\text{adj}(\lambda \mathbf{I}_n - \mathbf{A}) = \lambda\,\text{adj}(\lambda \mathbf{I}_n - \mathbf{A})$. Assim, existe uma coluna não nula da matriz $\text{adj}(\lambda \mathbf{I}_n - \mathbf{A})$ tal que $\mathbf{A}\mathbf{C}_j = \lambda \mathbf{C}_j$ ou é um autovetor de T associado ao autovalor λ. Vamos comprovar o resultado para valores pequeno de n, digamos $n = 2$. Neste caso, $f_{\mathbf{A}}(x) = x^2 - \text{tr}(\mathbf{A})x + \det(\mathbf{A})$. Assim, derivando em relaçāo a x, obtemos

$$f'_{\mathbf{A}}(x) = 2x - \text{tr}(\mathbf{A}) = 2x - (a_{11} + a_{22}) = \text{tr}\left(\text{adj}(x\mathbf{I}_2 - \mathbf{A})\right).$$

Por outro lado, como $f_{\mathbf{A}}(x) = (x - \lambda_1)(x - \lambda_2)$ temos que $f'_{\mathbf{A}}(x) = 2x - (\lambda_1 + \lambda_2)$, de modo que $\text{tr}\left(\text{adj}(x\mathbf{I}_2 - \mathbf{A})\right) = 2x - (\lambda_1 + \lambda_2)$. Em particular, se $x = \lambda_2$ e $\lambda_1 \neq \lambda_2$, então

$$(\lambda_2 - a_{11}) + (\lambda_2 - a_{22}) = \lambda_2 - \lambda_1 \neq 0.$$

Portanto, existe um j tal que $\lambda_2 - a_{jj} \neq 0$, isto é, existe uma coluna $\mathbf{C}_j$ não nula de $\text{adj}(\lambda_2 \mathbf{I}_2 - \mathbf{A})$ que é um autovetor de T associado ao autovalor λ_2. Na Seção 5.3 apresentaremos o caso geral. É muito importante observar que não foi simples encontrar os autovetores de T, mas é fácil provar se $\mathbf{v}$ for um autovetor de T, pois basta verificar se $T(\mathbf{v}) = \lambda \mathbf{v}$.

Exemplo 5.8 *Seja $T \in \mathcal{L}(\mathbb{R}^3)$ definido como*

$$T(x, y, z) = (2x - y - z, -x + 2y - z, -x - y).$$

Determine os autovalores e autovetores de T.

Solução. Vamos usar os dois métodos para resolver o problema. Observe que

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 0 \end{pmatrix} \quad \text{e} \quad f_{\mathbf{A}}(x) = x^3 - 4x^2 + x + 6 \in \mathbb{R}[x].$$

É fácil verificar que $\lambda_1 = -1, \lambda_2 = 2$ e $\lambda_3 = 3$ são os autovalores de T. Assim, para obter a solução da equação (5.5), basta escalonar a matriz:

$$\left(\begin{array}{ccc|c} 2-\lambda & -1 & -1 & f_{\mathbf{A}}(\lambda) \\ -1 & 2-\lambda & -1 & 0 \\ -1 & -1 & -\lambda & 0 \end{array}\right) \longrightarrow \cdots \longrightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & -\lambda^2+2\lambda+1 \\ 0 & 1 & 0 & \lambda-1 \\ 0 & 0 & 1 & \lambda-3 \end{array}\right).$$

Logo, $\mathbf{X}(\lambda) = (-\lambda^2+2\lambda+1, \lambda-1, \lambda-3)^t$ é a única solução, de modo que $[\mathbf{v}_1] = \mathbf{X}(\lambda_1) = (1,1,2)^t, \mathbf{X}(\lambda_2) = (1,1,-1)^t$ e $\mathbf{X}(\lambda_3) = (-1,1,0)^t$. Observe que se $f_{\mathbf{A}}(\lambda)$ estiver na segunda linha produz o mesmo resultado. Mas, se $f_{\mathbf{A}}(\lambda)$ estiver na terceira linha, obtemos $\mathbf{X}(\lambda) = (\lambda-3)(1,1,-\lambda+1)^t$ e $\mathbf{X}(\lambda_3) = (0,0,0)^t$, pois não podemos dividir por $\lambda-3$ e então fazer $\lambda = 3$ para $(\lambda-3)^{-1}f_{\mathbf{A}}(\lambda) \neq 0$ quando $\lambda = 3$. Pelo método da adjunta clássica, para $\lambda_1 = -1$, teremos

$$\operatorname{adj}(-\mathbf{I}_3 - \mathbf{A}) = \begin{pmatrix} 2 & 2 & 4 \\ 2 & 2 & 4 \\ 4 & 4 & 8 \end{pmatrix}$$

e $\mathbf{v}_1 = (1,1,2)$ é um autovetor de T associado ao autovalor $\lambda_1 = -1$. De modo análogo, trabalha com os outros autovalores. ■

Exemplo 5.9 *Seja $T \in \mathcal{L}(\mathbb{R}^3)$ definido como*

$$T(x,y,z) = (-x+2y-2z, 4x-3y+4z, 4x-4y+5z).$$

Determine os autovalores e autovetores de T.

Solução. Observe que

$$\mathbf{A} = \begin{pmatrix} -1 & 2 & -2 \\ 4 & -3 & 4 \\ 4 & -4 & 5 \end{pmatrix} \text{ e } f_{\mathbf{A}}(x) = x^3 - x^2 - x + 1 \in \mathbb{R}[x].$$

É fácil verificar que $\lambda_1 = -1$ e $\lambda_2 = 1$ são os autovalores de T. Assim, para obter a solução da equação (5.5), basta escalonar a matriz:

$$\left(\begin{array}{ccc|c} -1-\lambda & 2 & -2 & f_{\mathbf{A}}(\lambda) \\ 4 & -3-\lambda & 4 & 0 \\ 4 & -4 & 5-\lambda & 0 \end{array}\right) \longrightarrow \cdots \longrightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & -(\lambda-1)^2 \\ 0 & 1 & 0 & -4(\lambda-1) \\ 0 & 0 & 1 & -4(\lambda-1) \end{array}\right).$$

Logo, $\mathbf{X}(\lambda) = (\lambda-1)(\lambda-1,4,4)^t$ é a única solução. Não obstante, é muito importante observar que $\mathbf{Y}(\lambda) = (\lambda - 1)^{-1}\mathbf{X}(\lambda) = (\lambda - 1, 4, 4)^t$ satisfaz a equação

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{Y}(\lambda) = (\lambda^2 - 1, 0, 0)^t,$$

de modo que $\mathbf{Y}(\lambda_1) = (-1, 2, 2)^t$, $\mathbf{Y}(\lambda_2) = (0, 1, 1)^t$ e $\mathbf{Y}'(\lambda_2) = (1, 0, 0)^t$. ■

Exercícios

1. Determine, se existirem, os autovalores e autovetores de T em cada caso. Nos casos afirmativos determine o autoespaço generalizado.

   (a) $T(x, y) = (4y, x)$.
   (b) $T(x, y) = (x + y, 4x + y)$.
   (c) $T(x, y) = (-4x - 3y, 3x + 6y)$.
   (d) $T(x, y) = (2x + 3y, -x - 2y)$.
   (e) $T(x, y) = (2x + y, -y)$.
   (f) $T(x, y, z) = (x, 2y, -z)$.
   (g) $T(x, y, z) = (x + y + z, 2y + z, -z)$.
   (h) $T(a + bx + cx^2) = b + ax + cx^2$, para todo $a + bx + cx^2 \in \mathbb{P}_2(F)$.
   (i) $T(p(x)) = p(1 + x)$, para todo $p(x) \in \mathbb{P}_2(F)$.
   (j) $T(\mathbf{A}) = \mathbf{A}^t$, para todo $\mathbf{A} \in F^{2\times 2}$.
   (k) $T(x, y, z) = (3x + 2y - z, -3x - 2y + 2z, -x - y + 2z)$.
   (l) $T(x, y, z) = (9x + 2y - 6z, 4x + 2y - 3z, 12x + 3y - 8z)$.
   (m) $T(x, y, z) = (2x + 2y, x + y + 2z, x + y + 2z)$.
   (n) $T(x, y, z) = (x + y, x - y + 2z, 2x + y - z)$.
   (o) $T(x, y, z) = (-9x + 4y + 4z, -8x + 3y + 4z, -16x + 8y + 7z)$.
   (p) $T(x, y, z) = (x + 3y - 3z, 4y, -3x + 3y + z)$.

2. Qual é o operador $T \in \mathcal{L}(F^2)$ que possui $\lambda_1 = -2$ e $\lambda_2 = 3$ como autovalores associados aos autovetores $(3y, y)$ e $(-2y, y)$, com $y \neq 0$?

3. Seja $T \in \mathcal{L}(V)$ tendo $\lambda = 0$ como um autovalor. Mostre que T é singular.

4. Sejam $T \in \mathcal{L}(V)$ e $\lambda \in F$ um autovalor de T.

(a) Mostre que $\mu + \lambda$ é um autovalor de $T + \mu I$.

(b) Mostre que $\mu\lambda$ é um autovalor de μT.

(c) Mostre que se T for não singular, então λ^{-1} é um autovalor T^{-1}.

O que se pode dizer sobre os autovetores associados?

5. Sejam $T \in \mathcal{L}(F^2)$ definido como $T(x, y) = (ax + by, cx + dy)$ e $\mathbf{A} = [T]$.

(a) Mostre que os autovalores de T são:

$$\lambda_{1,2} = 2^{-1}[(a + d) \mp \sqrt{(a - d)^2 + 4bc}].$$

(b) Mostre que se $\operatorname{tr}(\mathbf{A})^2 - 4 \det \mathbf{A} > 0$ e $b \neq 0$, então $\mathbf{v}_1 = (-b, a - \lambda_1)$ e $\mathbf{v}_2 = (-b, a - \lambda_2)$

(c) Mostre que se $a, b, c, d \in \mathbb{R}$ são positivos, então $\lambda_1, \lambda_2 \in \mathbb{R}$, distintos e pelo menos um deles é positivo.

6. Sejam V um espaço vetorial sobre $\mathbb{R}$ e $T \in \mathcal{L}(V)$. Suponhamos que $b, c \in \mathbb{R}$ são tais que $T^2 + bT + cI = O$. Mostre que T possui um autovalor se, e somente se, $b^2 \geq 4c$.

7. Seja $\mathbf{A} \in F^{n\times n}$. Mostre que $\mathbf{A}$ e $\mathbf{A}^t$ possuem o mesmo polinômio característico, mas podem ter autovetores distintos.

8. Seja $\mathbf{A} \in \mathbb{R}^{2\times 2}$, com $\mathbf{A}^t = \mathbf{A}$, sendo $\lambda_1 = 1$ um autovalor e $\mathbf{X}_1 = (1, 3)^t$ o autovetor associado.

(a) Determine $\mathbf{A} \neq \mathbf{I}$ que satisfaça essas condições.

(b) Se $\lambda_2 = 9$ for outro autovalor de $\mathbf{A}$, determine um autovetor de $\mathbf{A}$ associado a λ_2.

(c) Determine uma matriz $\mathbf{B}$ tal que $\mathbf{B}^2 = \mathbf{A}$.

9. Sejam $\mathbf{A} \in F^{n\times n}$. Mostre que os autovalores da matriz de blocos

$$\mathbf{C} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix}$$

são exatamente os autovalores simultâneos de $\mathbf{A} + \mathbf{B}$ e $\mathbf{A} - \mathbf{B}$.

10. Sejam $\mathbf{D} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n), \mathbf{E} = \operatorname{diag}(\mu_1, \ldots, \mu_n) \in F^{n\times n}$. Mostre que $\mathbf{D}$ é semelhante a $\mathbf{E}$ se, e somente se, existir um $\sigma \in S_n$ tal que $\mu_i = \lambda_{\sigma(i)}$.

11. Sejam $S, T \in \mathcal{L}(V)$, com $\dim V = n$. Mostre que se $f_{\mathbf{A}}(x) = f_{\mathbf{B}}(x)$, então $\det \mathbf{A} = \det \mathbf{B}$. Mostre, com um exemplo, que a recíproca é falsa.

12. Sejam $\mathbf{A} \in F^{k\times n}$ e $\mathbf{B} \in F^{n\times k}$, com $k \leq n$. Mostre que $\det(x\mathbf{I}_n - \mathbf{BA}) = x^{n-k}\det(x\mathbf{I}_k - \mathbf{AB})$. Se $k = n$, então $\mathbf{AB}$ e $\mathbf{BA}$ possuem o mesmo polinômio característico. O que se pode dizer sobre os autovetores?

13. Sejam $T \in \mathcal{L}(V)$ e $\mathbf{u}, \mathbf{v} \in V$ autovetores de T associados aos autovalores distintos λ, μ. Mostre que se $a\mathbf{u} + b\mathbf{v}$ for autovetor de T, então $a = 0$ ou $b = 0$.

14. Seja $T \in \mathcal{L}(V)$ tal que qualquer $\mathbf{v} \in V - \{\mathbf{0}\}$ é um autovetor de T. Mostre que $T = cI$, para algum $c \in F$.

15. Sejam $T \in \mathcal{L}(V)$, com $\dim V = n$, e $\mathbf{A} = [T]_\alpha^\alpha \in \mathbb{R}^{n\times n}$, para alguma base α de V. Mostre que:

    (a) Se n for ímpar e $\det(\mathbf{A}) > 0$, então T possui pelo menos um autovalor positivo.

    (b) Se n for ímpar e $\det(\mathbf{A}) < 0$, então T possui pelo menos um autovalor negativo.

    (c) Se n for par e $\det(\mathbf{A}) < 0$, então T possui pelo menos um autovalor positivo e um negativo.

16. Seja $T \in \mathcal{L}(\mathbb{R}^3)$. Mostre que $T(r) \subseteq r$, para alguma reta r em $\mathbb{R}^3$ que passa pela origem.

5.2 Operadores Diagonalizáveis

Já vimos que as matrizes diagonais eram bastante fáceis de entender e estudar. Nesta seção, vamos estudar as matrizes que são as mais próximas possíveis de serem diagonais, ou seja, as matrizes que são semelhantes a uma matriz diagonal. Vale lembrar que qualquer matriz é equivalente a uma matriz diagonal.

Teorema 5.10 *Sejam $T \in \mathcal{L}(V)$ e $\lambda_i \in F$ autovalores distintos de T, com $i = 1, \ldots, n$. Se $\mathbf{u}_i \in V$ for o autovetor de T associado λ_i, então $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ é LI.*

Prova. Suponhamos, por absurdo, que $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ seja LD. Então, pelo Corolário 3.33, podemos escolher um menor k tal que

$$\mathbf{u}_k = x_1\mathbf{u}_1 + \cdots + x_{k-1}\mathbf{u}_{k-1}, 2 \leq k \leq n. \tag{5.6}$$

Assim, aplicando T a equação (5.6), obtemos

$$\lambda_k\mathbf{u}_k = x_1\lambda_1\mathbf{u}_1 + \cdots + x_{k-1}\lambda_{k-1}\mathbf{u}_{k-1}. \tag{5.7}$$

Logo, multiplicando a equação (5.6) por λ_k e subtraindo da equação (5.7), temos que

$$\mathbf{0} = (\lambda_k - \lambda_1)x_1\mathbf{u}_1 + \cdots + (\lambda_k - \lambda_{k-1})x_{k-1}\mathbf{u}_{k-1}.$$

Pela escolha de k, $\{\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}\}$ é LI, de modo que $(\lambda_k - \lambda_i)x_i = 0$, para cada $i = 1, \ldots, k-1$. Como $\lambda_k - \lambda_i \neq 0$ temos que $x_i = 0$, para $i = 1, \ldots, k-1$. Portanto, $\mathbf{u}_k = \mathbf{0}$, o que é uma contradição. ■

Corolário 5.11 *Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Então T possui no máximo n autovalores distintos. Se T possui exatamente n autovalores, então V possui uma base de autovetores.*

Prova. Como $\partial(f_{\mathbf{A}}) = n$ temos, pelo Teorema 1.2, que esse polinômio possui no máximo n raízes sobre F. ■

Observe que a recíproca do Teorema 5.10 é falsa, isto é, se T possui autovetores linearmente independentes $\mathbf{u}_1, \ldots, \mathbf{u}_n$, então os autovalores associados $\lambda_1, \ldots, \lambda_n$ não necessitam ser distintos. Por exemplo, se $T = I_V \in \mathcal{L}(V)$, então qualquer $\mathbf{v} \in V$, com $\mathbf{v} \neq \mathbf{0}$, é um autovetor de T associado ao único autovalor 1.

Sejam α, β bases de V, $T \in \mathcal{L}(V)$, $\mathbf{A} = [T]_\alpha^\alpha$ e $\mathbf{B} = [T]_\beta^\beta$. Então existe uma matriz não singular $\mathbf{P}$ tal que $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. Se $\mathbf{v} \in V$, com $\mathbf{v} \neq \mathbf{0}$, é um autovetor de T associado ao autovalor $\lambda \in F$ e $\mathbf{X} = [\mathbf{v}]_\alpha$, então $\mathbf{Y} = \mathbf{P}^{-1}\mathbf{X}$ for um autovetor de $\mathbf{B}$ associado ao mesmo autovalor λ, pois $\mathbf{B}\mathbf{Y} = \lambda\mathbf{Y}$.

Teorema 5.12 *Sejam α uma base de V, $T \in \mathcal{L}(V)$, $\mathbf{A} = [T]_\alpha^\alpha$ e $\mathbf{X}_j = [\mathbf{v}]_\alpha$ de um autovetor $\mathbf{v}$ de T associado ao autovalor λ_j, com $j = 1, \ldots, n$. As colunas $\mathbf{X}_1, \ldots, \mathbf{X}_n$ da matriz $\mathbf{P}$ são autovetores de T se, e somente se, $\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{D}$, com $\mathbf{D} = \operatorname{diag}(\lambda_1, \ldots \lambda_n)$. Neste caso, $\lambda\mathbf{E}_j$ é a j-ésima coluna de $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$.*

Prova. Observe que

$$\mathbf{AP} = \mathbf{A}\begin{pmatrix} \mathbf{X}_1 & \cdots & \mathbf{X}_n \end{pmatrix} = \begin{pmatrix} \mathbf{AX}_1 & \cdots & \mathbf{AX}_n \end{pmatrix}$$

e

$$\mathbf{PD} = \mathbf{P}\begin{pmatrix} \lambda_1\mathbf{E}_1 & \cdots & \lambda_n\mathbf{E}_n \end{pmatrix} = \begin{pmatrix} \lambda_1\mathbf{X}_1 & \cdots & \lambda_n\mathbf{X}_n \end{pmatrix}$$

Portanto, $\mathbf{AP} = \mathbf{PD}$ se, e somente se, $\mathbf{AX}_i = \lambda_i\mathbf{X}_i$ se, e somente se, as colunas $\mathbf{X}_1, \ldots, \mathbf{X}_n$ de $\mathbf{P}$ são autovetores de T. ■

Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Diremos que T é *diagonalizável* se existir uma base de V formada de autovetores de T.

É muito interessante, de um ponto de vista prático e teórico, resumir o que já foi estudado em um procedimento (um algoritmo) para decidir sobre a diagonalização ou não de uma matriz $\mathbf{A} \in F^{n\times n}$:

$1.^o$ **Passo**. Determine o polinômio característico de $\mathbf{A}$.

$2.^o$ **Passo**. Determine os autovalores de $\mathbf{A}$.

$3.^o$ **Passo**. Repete (a) e (b) para cada autovalor λ de $\mathbf{A}$:

(a) Resolva a equação (5.2) ou a equação (5.5).
(b) Determine uma base de V_λ e/ou de V^λ.

$4.^o$ **Passo**. Considere o conjunto dos autovetores $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ de $\mathbf{A}$ obtidos no $3.^o$ **Passo**:

(a) Se $m \neq n$ ou $V_\lambda \neq V^\lambda$, para algum autovalor λ de $\mathbf{A}$, então ela não é diagonalizável. Neste caso, determine $\mathbf{P}$, de modo que $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{J}$.
(b) Se $m = n$ ou $V_\lambda = V^\lambda$, para todo autovalor λ de $\mathbf{A}$, então ela é diagonalizável. Neste caso, determine $\mathbf{P}$, de modo que $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$.

De posse do Teorema 5.12, é muito importante, de um ponto de vista teórico e didático, ressaltar o seguinte. Se $\mathbf{Q} = (\mathbf{P}^{-1})^t$ e

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Y}_1 & \cdots & \mathbf{Y}_n \end{pmatrix}$$

então $\mathbf{Q}^t\mathbf{P} = \mathbf{I}$, de modo que $\mathbf{Y}_i^t\mathbf{X}_j = \delta_{ij}$ é uma condição de "ortogonalidade". Por outro lado, $\mathbf{A}^t = \mathbf{QDQ}^{-1}$ ou $\mathbf{A}^t\mathbf{Q} = \mathbf{QD}$, de modo que

$$\mathbf{A}^t\mathbf{Y}_i = \lambda_i\mathbf{Y}_i, i = 1, \ldots, n.$$

Assim, $\mathbf{Y}_i$ é um autovetor de $\mathbf{A}^t$, o qual é conhecido como *autovetor à esquerda* de $\mathbf{A}$, pois $\mathbf{Y}_i^t\mathbf{A} = \lambda_i\mathbf{Y}_i^t$. Portanto,

$$\mathbf{A} = \mathbf{PDQ}^t = \sum_{i=1}^{n} \lambda_i \mathbf{X}_i \mathbf{Y}_i^t \tag{5.8}$$

é uma soma de matrizes de posto um quando $\lambda_i \neq 0$ e zero caso contrário. Note que cada matriz é definida em termos das *propriedades espectrais* de $\mathbf{A}$, ou seja, autovalores, autovetores e autovetores à esquerda e denotamos por $\sigma(T)$ ou $\sigma(\mathbf{A})$ o conjunto dos autovalores de T e chama-se *conjunto espectral* de T. Observe, pelo Teorema 5.16, que $\sigma(T) \neq \emptyset$ e $\rho(\mathbf{A})$ é igual ao número de autovalores não nulos de $\mathbf{A}$, com cada um contado tantas vezes quanto for sua multiplicidade.

Exemplo 5.13 *Seja $T \in \mathcal{L}(\mathbb{R}^3)$ definido como*

$$T(x, y, z) = (x - 3y + 3z, 3x - 5y + 3z, 6x - 6y + 4z).$$

Mostre que $\mathbb{R}^3$ possui uma base de autovetores.

Solução. É fácil verificar que $f_{\mathbf{A}}(x) = (x+2)^2(x-4)$, $V_{-2} = \mathbb{R}[(1,1,0),(1,0,-1)]$ e $V_4 = \mathbb{R}[(1,1,2)]$. Portanto, $\beta = \{(1,1,0),(1,0,-1),(1,1,2)\}$ é uma base de autovetores de $\mathbb{R}^3$ e T é diagonalizável. Note que $m_g(-2) = m_a(-2) = 2$, $m_g(4) = m_a(4) = 1$, $\mathbb{R}^3 = V_{-2} \oplus V_4$ e $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$, com

$$\mathbf{D} = [T]_\beta^\beta = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \text{ e } \mathbf{P} = [P]_\beta^\alpha = [I]_\alpha^\beta = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & -1 & 2 \end{pmatrix}$$

é a matriz de transição da base canônica α para a base β. ■

Exemplo 5.14 *Seja $T \in \mathcal{L}(\mathbb{R}^3)$ definido como*

$$T(x, y, z) = (-3x + y - z, -7x + 5y - z, -6x + 6y - 2z).$$

T é diagonalizável?

Solução. É fácil verificar que $f_{\mathbf{A}}(x) = (x+2)^2(x-4)$. Para obter a solução da equação

(5.5), basta escalonar a matriz:

$$\left(\begin{array}{ccc|c} -3-\lambda & 1 & -1 & 0 \\ -7 & 5-\lambda & -1 & f_{\mathbf{A}}(\lambda) \\ -6 & 6 & -2-\lambda & 0 \end{array}\right) \to \cdots \to \left(\begin{array}{ccc|c} 1 & 0 & 0 & -(\lambda-4) \\ 0 & 1 & 0 & -\lambda(\lambda+5) \\ 0 & 0 & 1 & -6(\lambda+2) \end{array}\right).$$

Assim, $\mathbf{X}(\lambda) = (4-\lambda, -\lambda(\lambda+5), -6(\lambda+2))^t$ é a única solução. Logo, $\mathbf{X}(-2) = (6, 6, 0)^t$ ou $\mathbf{v}_1 = (1, 1, 0)$ é o único autovetor T associado a -2, de modo que $m_g(-2) < m_a(-2) = 2$. Portanto, T não é diagonalizável. Observe que o autovetor generalizado é $\mathbf{X}'(-2) = (-1, -1, -6)^t$ ou $\mathbf{v}_2 = (1, 1, 6)$. Por outro lado, $\mathbf{X}(4) = (0, -36, -36)^t$ ou $\mathbf{v}_3 = (0, 1, 1)$ é o único autovetor T associado a 4. Portanto, $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ é uma base de $\mathbb{R}^3$. Note que $\mathbb{R}^3 = V^{-2} \oplus V^4$ e $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{J}$, com

$$\mathbf{J} = [T]_\beta^\beta = \begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \text{ e } \mathbf{P} = [P]_\beta^\alpha = [I]_\alpha^\beta = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 6 & 0 \end{pmatrix}$$

é a matriz de transição da base canônica α para a base β. ■

Seja $T \in \mathcal{L}(V)$. Vamos lembrar que a sequência de operadores $\{T^n : n \in \mathbb{Z}_+\}$ sobre V foi definida. indutivamente, como

$$T^0 = I_V = I \text{ e } T^n = T^{n-1} \circ T = T^{n-1}T, \ \forall \ n \in \mathbb{N}.$$

Lema 5.15 *Sejam V um espaço vetorial sobre F e $S, T \in \mathcal{L}(V)$.*

1. $T^kT^m = T^{k+m}$, para todos $k, m \in \mathbb{N}$.
2. $(T^k)^m = T^{km}$, para todos $k, m \in \mathbb{N}$.
3. Se $ST = TS$, então $(S+T)^m = \sum_{k=0}^m \binom{m}{k} S^{m-k}T^k$, para todo $m \in \mathbb{N}$.

Prova. Fica como um exercício. ■

Sejam $T \in \mathcal{L}(V)$ e $f(x) = c_mx^m + \cdots + c_1x + c_0 \in F[x]$ um polinômio de grau $\partial(f) = m$.. Então $f(T)$ é um operador sobre V definido como

$$f(\mathbf{T}) = c_m\mathbf{T}^m + \cdots + c_1\mathbf{T} + c_0\mathbf{I} \in \mathcal{L}(V),$$

o qual chama-se *função polinomial* em T. Note que o operador $f(T)$ foi obtido de $f(x)$ substituindo-se a variável x pelo operador T e o escalar c_0 pelo operador escalar c_0I.

Diremos que $f(x)$ é um *polinômio anulador* de T se $f(T) = 0$ e T é um *zero* de $f(x)$. Formalmente, cada $T \in \mathcal{L}(V)$ induz uma função $\sigma : F[x] \to \mathcal{L}(V)$, com $\sigma = \sigma_T$, definida como $\sigma(f(x)) = f(T)$, que goza das seguintes propriedades:

1. Se $f(x) = g(x)$, então $f(T) = g(T)$, ou seja, σ está bem definida.

2. $\sigma(f(x) + cg(x)) = \sigma(f(x)) + c\sigma(g(x))$, para todos $f(x), g(x) \in F[x]$ e $c \in F$, ou seja, σ é linear.

3. $\sigma((f \cdot g)(x)) = \sigma(f(x))\sigma(g(x))$, ou seja, $(f \cdot g)(T) = f(T)g(T)$, para todos $f(x), g(x) \in F[x]$. Em particular, $f(T)g(T) = g(T)f(T)$.

4. $\sigma(1) = I_V$ e $\sigma(x) = T$, ou seja, σ é completamente determinada pelos valores $\sigma(1)$ e $\sigma(x)$.

É importante lembrar que $F[x]$ é algebricamente semelhante a $\mathbb{Z}$. Por exemplo, o **Algoritmo da Divisão** (AD): dados $f(x), g(x) \in F[x]$, com $g(x) \neq 0$, existem únicos $q(x), r(x) \in F[x]$ tais que

$$f(x) = g(x)q(x) + r(x), \text{ com } r(x) = 0 \text{ ou } \partial(r) < \partial(g). \tag{5.9}$$

Quando $r(x) = 0$, diremos que $g(x)$ é um *fator* de $f(x)$ ou que $g(x)$ *divide* $f(x)$.

Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Então, pelo Teorema 4.32, a função $\phi : \mathcal{L}(V) \to F^{n\times n}$ definida como $\phi(T) = \mathbf{A}$, com $\mathbf{A} = [T]_\alpha^\alpha$, para alguma base fixada α de V, é um isomorfismo de álgebras lineares. Assim, $\varphi = \phi \circ \sigma$ é completamente determinada pelos valores $\varphi(1) = \mathbf{I}$ e $\varphi(x) = \mathbf{A}$. Portanto, $f(\mathbf{A}) = [f(T)]_\alpha^\alpha$, para todo $f(x) \in F[x]$, e vice-versa. Neste ponto, é importante ressaltar que: se V for um espaço vetorial sobre $\mathbb{R}$, então, pelo Exercício (12) da Seção 3.1, $V^{\mathbb{C}}$ é um espaço vetorial sobre $\mathbb{C}$. Seja $T \in \mathcal{L}(V)$. Então vamos estender sua ação sobre $V^{\mathbb{C}}$ como

$$T^{\mathbb{C}}(\mathbf{z}) = T(\mathbf{u}) + iT(\mathbf{v}), \ \forall \ \mathbf{z} = \mathbf{u} + i\mathbf{v} \in V^{\mathbb{C}}.$$

Note que T e $T^{\mathbb{C}}$ são os mesmos, pois se α for uma base de V, então α é uma base de $V^{\mathbb{C}}$. Portanto, $f(T) = f(T^{\mathbb{C}})$, para todo $f(x) \in \mathbb{R}[x]$. Além disso, somente os $\lambda \in \mathbb{R}$ são admitidos como autovalores de T.

Observe, para cada $T \in \mathcal{L}(V)$, que σ induz um produto (uma ação) de $F[x]$ sobre V definido como

$$f(x)\mathbf{v} = f(T)(\mathbf{v}) = c_nT^n(\mathbf{v}) + \cdots + c_1T(\mathbf{v}) + c_0\mathbf{v}, \ \forall \ f(x) \in F[x] \text{ e } \mathbf{v} \in V.$$

Note que este produto satisfaz todas as propriedades de um espaço vetorial. Por exemplo,

$$f(x)(\mathbf{u}+\mathbf{v}) = f(x)\mathbf{u} + f(x)\mathbf{v} \text{ e } (f(x)g(x))\mathbf{u} = f(x)(g(x)\mathbf{u}).$$

Não obstante, V não é um espaço vetorial sobre $F[x]$, pois $F[x]$ não é um corpo.

Teorema 5.16 *Seja V um espaço vetorial de dimensão finita sobre $\mathbb{C}$. Então qualquer $T \in \mathcal{L}(V)$ possui pelo menos um autovalor.*

Prova. Sejam $\dim V = n > 1$ e $\mathbf{v} \in V$, com $\mathbf{v} \neq \mathbf{0}$. Então a lista

$$\alpha = (\mathbf{v}, T(\mathbf{v}), \ldots, T^n(\mathbf{v}))$$

sobre V é LD. Assim, pelo Corolário 3.33, podemos escolher um menor k tal que

$$T^k(\mathbf{v}) = c_0\mathbf{v} + \cdots + c_{k-1}T^{k-1}(\mathbf{v}), 2 \leq k \leq n.$$

Isto implica que existe um $g(x) = \sum_{i=0}^{k-1} c_i x^i + x^k \in \mathbb{C}[x]$ tal que $g(T)(\mathbf{v}) = \mathbf{0}$. Por outro lado, pelo Teorema 1.2, existem $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$, não necessariamente distintos, tais que $g(x) = (x-\lambda_1)\cdots(x-\lambda_k)$. Assim,

$$\mathbf{0} = g(T)(\mathbf{v}) = ((T-\lambda_1 I)\cdots(T-\lambda_k I))(\mathbf{v}) = (T-\lambda_j I)(\prod_{i\neq j}(T-\lambda_i I)(\mathbf{v})),$$

para algum j. Mas, isto significa, pela escolha de k, que $(T-\lambda_j I)(\mathbf{u}) \neq \mathbf{0}$, para algum $\mathbf{u} \in V$, com $\mathbf{u} \neq \mathbf{0}$. Portanto, T possui pelo menos um autovalor. ■

Observe que o Teorema 5.16 nos afirma que existe uma base α de V tal que a primeira coluna de $[T]_\alpha^\alpha$ é da forma $\lambda\mathbf{E}_1$.

Sejam V um espaço vetorial sobre F e $W_1, \ldots, W_k$ subespaços de V. Diremos que $W_1, \ldots, W_k$ são *independentes* se a seguinte condição for satisfeita: $\mathbf{u}_i \in W_i$ e $\mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_k = \mathbf{0}$ implica que $\mathbf{u}_i = \mathbf{0}$, para cada $i = 1, \ldots, k$.

Lema 5.17 *Sejam $W_1, \ldots, W_k$ subespaços de V e $W = W_1 + \cdots + W_k$, com $\dim V = n$. Então as seguintes condições são equivalentes*:

1. *$W_1, \ldots, W_k$ são independentes*;
2. *$W_j \cap (W_1 + \cdots + W_{j-1}) = \{\mathbf{0}\}$, para $2 \leq j \leq k$*;
3. *Se α_i for uma base de W_i, então a lista $\alpha = (\alpha_1, \ldots, \alpha_k)$ é uma base de W.*

Prova. É claro que (1) implica (2). $(2 \Rightarrow 3)$ É fácil verificar que $W = F[\alpha]$. Como qualquer relação linear entre os vetores de α terá a forma $\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_k = \mathbf{0}$, em que cada vetor $\mathbf{v}_i$ é alguma combinação linear dos vetores de α_i, temos que $\mathbf{v}_k \in W_k \cap (W_1 + \cdots + W_{k-1}) = \{\mathbf{0}\}$, de modo que $\mathbf{v}_k = \mathbf{0}$. Continuando deste modo, temos que α é LI. É claro que (3) implica (1). ■

Sejam V um espaço vetorial sobre F e $W_1, \ldots, W_k$ subespaços de V. Diremos que V é uma *soma direta interna* de $W_1, \ldots, W_k$ se pelo menos uma (e, portanto, todas) das condições do Lema 5.17 for satisfeita e denotamos por

$$V = W_1 \oplus \cdots \oplus W_k.$$

Lema 5.18 *Seja $T \in \mathcal{L}(V)$ tal que $T(\mathbf{u}) = \lambda\mathbf{u}$, com $\mathbf{u} \neq \mathbf{0}$. Então, para todo $f(x) \in F[x]$, $f(T)(\mathbf{u}) = f(\lambda)\mathbf{u}$ e $\mathbf{u}$ é o autovetor de $f(T)$ associado a $f(\lambda)$.*

Prova. Basta notar, indutivamente, que $T^n(\mathbf{u}) = \lambda^n\mathbf{u}$, para todo $n \in \mathbb{Z}_+$. ■

Teorema 5.19 *Sejam $T \in \mathcal{L}(V)$, com $\dim V = n$, e $V_{\lambda_i} = \ker(T - \lambda_i I)$ associados aos autovalores distintos $\lambda_i \in F$. Então as seguintes condições são equivalentes*:

1. *T é diagonalizável*;

2. *O polinômio característico de T é*

$$f_{\mathbf{A}}(x) = (x - \lambda_1)^{m_1}(x - \lambda_2)^{m_2} \cdots (x - \lambda_k)^{m_k}, \quad \textit{com} \ \ m_i = m_g(\lambda_i);$$

3. $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$.

Prova. $(1 \Rightarrow 2)$ Suponhamos que T seja diagonalizável. Então existe uma base $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ de V tal que $T(\mathbf{u}_i) = \lambda_i\mathbf{u}_i$. Assim, $\mathbf{A} = \operatorname{diag}(\lambda_1\mathbf{I}_1, \ldots, \lambda_k\mathbf{I}_k)$. Portanto, pelo item (10) do Teorema 1.24,

$$f_{\mathbf{A}}(x) = \det(x\mathbf{I} - \mathbf{A}) = (x - \lambda_1)^{m_1}(x - \lambda_2)^{m_2} \cdots (x - \lambda_k)^{m_k}.$$

$(2 \Rightarrow 3)$ Seja $\mathbf{u}_i \in V_{\lambda_i}$ tal que $\mathbf{u}_1 + \cdots + \mathbf{u}_k = \mathbf{0}$. Então, para cada i fixado, pondo $T_j = \prod_{j \neq i}(T - \lambda_j I_j) = \widehat{g}_{\mathbf{A}}(T)$, temos, pelo Lema 5.18, que

$$T_j(\mathbf{u}_i) = \widehat{g}_{\mathbf{A}}(T)(\mathbf{u}_i) = \widehat{g}_{\mathbf{A}}(\lambda_i)\mathbf{u}_i = \prod_{j \neq i}(\lambda_i - \lambda_j)\mathbf{u}_i.$$

Por outro lado, $\mathbf{0} = T_j(\mathbf{u}_1 + \cdots + \mathbf{u}_n) = T_j(\mathbf{u}_i)$ implica que $\prod_{j \neq i}(\lambda_i - \lambda_j)\mathbf{u}_i = \mathbf{0}$, de modo que $\mathbf{u}_i = \mathbf{0}$, pois $\lambda_j \neq \lambda_i$ quando $j \neq i$. $(3 \Rightarrow 1)$ É claro. ■

Já vimos, no teorema 5.16, que qualquer operador sobre um espaço vetorial complexo tinha pelo menos um autovalor. No entanto, pelo Exemplo 5.3, o mesmo não é verdade sobre um espaço vetorial real, ou seja, ele pode não possuir subespaços invariantes de dimensão 1.

Teorema 5.20 *Seja V um espaço vetorial de dimensão finita sobre $\mathbb{R}$. Então qualquer $T \in \mathcal{L}(V)$ possui um subespaço invariante de dimensão 1 e/ou 2.*

Prova. Sejam $\dim V = n > 1$ e $\mathbf{v} \in V$, com $\mathbf{v} \neq \mathbf{0}$. Então a lista

$$\alpha = (\mathbf{v}, T(\mathbf{v}), \ldots, T^n(\mathbf{v}))$$

sobre V é LD. Assim, existem $c_0, \ldots, c_n \in \mathbb{R}$, não todos nulos, tais que

$$c_0\mathbf{v} + c_1T(\mathbf{v}) + \cdots + c_nT^n(\mathbf{v}) = \mathbf{0}.$$

Isto implica que existe um $g(x) = \sum_{i=0}^{n} c_i x^i \in \mathbb{R}[x]$ tal que $g(T)(\mathbf{v}) = \mathbf{0}$. Dividindo pelo coeficiente líder, podemos escrever $g(x) = \sum_{i=0}^{m-1} d_i x^i + x^m$. Por outro lado, pelo equação (1.3),

$$g(x) = \prod_{j=1}^{r}(x - \lambda_j)\prod_{k=1}^{s}(x^2 + b_kx + c_k), \quad \text{onde} \quad r, s \in \mathbb{N} \quad \text{e} \quad r + 2s = m.$$

Assim,

$$\mathbf{0} = g(T)(\mathbf{v}) = [\prod_{j=1}^{r}(T - \lambda_jI)\prod_{k=1}^{s}(T^2 + b_kT + c_kI)](\mathbf{v}).$$

Mas, isto significa que $T - \lambda_jI$ é singular, para algum j, ou $T^2 + b_kT + c_kI$ é singular, para algum k. Se $T - \lambda_jI$ for singular, então existe um $\mathbf{u} \in V$, com $\mathbf{u} \neq \mathbf{0}$, tal que $(T - \lambda_jI)(\mathbf{u}) = \mathbf{0}$. Portanto, T possui um subespaço invariante de dimensão 1. Se $T^2 + b_kT + c_kI$ for singular, então existe um $\mathbf{u} \in V$, com $\mathbf{u} \neq \mathbf{0}$, tal que

$$T^2(\mathbf{u}) + b_kT(\mathbf{u}) + c_k\mathbf{u} = \mathbf{0}.$$

Afirmação. $W = \mathbb{R}[\mathbf{u}, T(\mathbf{u})]$ é invariante sob T, com $\dim W = 1$ ou 2.

De fato. Dado $\mathbf{w} \in W$, existem $a, b \in \mathbb{R}$ tais que $\mathbf{w} = a\mathbf{u} + bT(\mathbf{u})$. Assim,

$$T(\mathbf{w}) = aT(\mathbf{u}) + bT^2(\mathbf{u}) = (-bc_k)\mathbf{u} + (a - bb_k)T(\mathbf{u}) \in W.$$

Portanto, W é invariante sob T. ■

É muito importante, de um ponto de vista teórico e didático, interpretar a prova do Teorema 5.20 a luz da complexificação $V^{\mathbb{C}}$ de V. Seja $\mathbf{v} \in V$, com $\mathbf{v} \neq \mathbf{0}$, um autovetor de T associado ao autovalor λ. Se $\lambda \in \mathbb{R}$, então $T^{\mathbb{C}}(\mathbf{v}) = T(\mathbf{v}) = \lambda\mathbf{v}$, de modo que $T^{\mathbb{C}}$ possui o mesmo autovalor e autovetor de T. Se $\lambda \in \mathbb{C}$, então $T^{\mathbb{C}}(\mathbf{z}) = \lambda\mathbf{z}$, com $\mathbf{z} = \mathbf{u} + i\mathbf{v} \in V^{\mathbb{C}}$ não nulo. Assim, definindo $\bar{\mathbf{z}} = \mathbf{u} - i\mathbf{v}$ em $V^{\mathbb{C}}$, obtemos

$$T^{\mathbb{C}}(\bar{\mathbf{z}}) = T(\mathbf{u}) - iT(\mathbf{v}) = \overline{T(\mathbf{u}) + iT(\mathbf{v})} = \overline{T^{\mathbb{C}}(\mathbf{z})} = \bar{\lambda}\bar{\mathbf{z}}.$$

Logo, $\bar{\mathbf{z}}$ é um autovetor de $T^{\mathbb{Z}}$ associado ao autovalor $\bar{\lambda}$. Como $\lambda \neq \bar{\lambda}$ temos, pelo Teorema 5.10, que $\mathbf{z}$ e $\bar{\mathbf{z}}$ são LI. Portanto, $W = \mathbb{R}[\mathbf{z}, \bar{\mathbf{z}}]$ é invariante sob T com uma base $\{\mathbf{v}, \mathbf{u}\}$, pois $2\mathbf{u} = \mathbf{z} + \bar{\mathbf{z}}, 2\mathbf{v} = i(\bar{\mathbf{z}} - \mathbf{z}) \in W$ e se $\lambda = a + ib \in \mathbb{C}$, então

$$T(\mathbf{v}) = a\mathbf{v} + b\mathbf{u} \text{ e } T(\mathbf{u}) = -b\mathbf{v} + a\mathbf{u} \Rightarrow [T|_W]_\alpha^\alpha = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Note que $T|_W$ não possui autovalores reais. Neste caso, concluímos que se o polinômio característico de $T \in \mathcal{L}(V)$ possui um autovalor complexo, então ele está associada a um subespaço W de V invariante sob T, com $\dim W = 2$.

Exercícios

1. Para cada $T \in \mathcal{L}(V)$ do Exercício (1) da Seção 5.1, identifique os que são diagonalizáveis. Nos casos afirmativos, especifique uma matriz $\mathbf{P}$ tal que $\mathbf{P}^{-1}\mathbf{AP}$ seja diagonal.

2. Seja $T \in \mathcal{L}(\mathbb{R}^2)$ definido como $T(x, y) = (ax + by, cx + dy)$. Mostre que se $a, b, c, d \in \mathbb{R}_+$, então T é diagonalizável.

3. Seja $T \in \mathcal{L}(\mathbb{R}^2)$, com $[T]^t = [T]$. Mostre que os autovalores de T são reais e que T é diagonalizável.

4. Seja $T \in \mathcal{L}(F^3)$ tal que $[T] = \mathbf{U}_3 = (u_{ij})$, com $u_{ij} = 1$. Mostre que T é diagonalizável. Generalize para F^n.

5. Considere as matrizes

$$\mathbf{A} = \begin{pmatrix} -4 & -3 \\ 10 & 7 \end{pmatrix} \text{ e } \mathbf{B} = \begin{pmatrix} 1 & -2 & 3 \\ 0 & -1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$$

Calcule $\mathbf{A}^{10}$ e $\mathbf{B}^{35}$.

6. Seja $T \in \mathcal{L}(F^3)$ definido como

$$T(x, y, z) = (ax + by + bz, bx + ay + bz, bx + by + az).$$

Determine os autovalores e autovetores de T. Além disso, especifique uma matriz $\mathbf{P}$ tal que $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ seja diagonal.

7. Seja $T \in \mathcal{L}(F^3)$ definido como $T(x, y, z) = (ay, bx + bz, ay)$. Determine os autovalores e autovetores de T. Além disso, para que valores de a e b, T é diagonalizável.

8. Seja $\mathbb{C}$ visto como um espaço vetorial sobre $\mathbb{R}$.

   (a) Mostre que $T \in \mathcal{L}(\mathbb{C})$ se, e somente se, existem $a, b \in \mathbf{C}$ tais que $T(z) = az + b\bar{z}$, para todo $z \in \mathbb{C}$. Conclua que T é não singular se, e somente se, $|a| \neq |b|$.

   (b) Mostre que qualquer automorfismo $T \in \mathcal{L}(\mathbb{C})$ é da forma $T = I$ ou $T(z) = \bar{z}$, para todo $z \in \mathbb{C}$.

9. Seja $C \in \mathcal{L}(V^{\mathbb{C}})$ definido como $C(\mathbf{z}) = \bar{\mathbf{z}}$. Mostre que C é um isomorfismo sobre $\mathbb{R}$. Conclua que C é uma involução, ou seja, $C^2 = I$.

10. Sejam $\mathbf{A}, \mathbf{B} \in F^{n\times n}$. Mostre que se $\mathbf{A}$ e $\mathbf{B}$ são semelhantes, então $f(\mathbf{A})$ e $f(\mathbf{B})$ são semelhantes, para todo $f(x) \in F[x]$.

11. Seja $V = \operatorname{Seq}_f(F)$ o espaço das sequências do tipo Fibonacci.

    (a) Mostre que

    $$\begin{pmatrix} a_{n-1} & a_n \\ a_n & a_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n = \mathbf{C}_2^n,$$

    para todo $n \in \mathbb{N}$ e $a_0 = 0$. Conclua que $a_{n+1}a_{n-1} - a_n^2 = (-1)^n$.

    (b) Encontre uma fórmula explícita para a_n, para todo $n \in \mathbb{N}$.

12. Seja $S \in \mathcal{L}(V)$, com $\dim V = n$, diagonalizável e todos os seus autovalores distintos. Mostre que qualquer $T \in \mathcal{L}(V)$ diagonalizável pode ser escrito como um polinômio em S.

13. Sejam $V = \mathcal{C}^{\infty}(\mathbb{R}, \mathbb{R})$ e $\beta = \{e^{a_1 x}, \ldots, e^{a_n x}\}$, onde os $a_i \in \mathbb{R}$ são distintos. Mostre que β é um subconjunto linearmente independente em V.

5.3 Polinômio Minimal

Já vimos que as funções $\det, \operatorname{tr} : F^{n\times n} \to F$ e $\psi : F^{n\times n} \to F[x]$ definida como $\psi(\mathbf{A}) = f_{\mathbf{A}}(x)$ eram invariantes por similaridade. Nesta seção, vamos estudar outro invariante, o qual será muito útil para entender e estudar os subespaços invariantes ou a *abordagem geométrica* de um operador.

Seja V um espaço vetorial sobre F, com $\dim V = n$. Então já sabemos que cada $T \in \mathcal{L}(V)$ (ou $\mathbf{A} \in F^{n\times n}$) induz uma função $\sigma : F[x] \to \mathcal{L}(V)$ definida como $\sigma(f(x)) = f(T)$, a qual é linear. Assim, $\ker \sigma$ é um subespaço não trivial de $F[x]$. De fato. Como $\dim \mathcal{L}(V) = n^2$ temos que a lista $\alpha = (I, T, \ldots, T^{n^2})$ é LD em $\mathcal{L}(V)$. Assim, existem $c_0, \ldots, c_{n^2} \in F$, não todos nulos, tais que

$$c_0 I + c_1 T + \cdots + c_{n^2} T^{n^2} = 0.$$

Isto implica que existe um $g(x) = \sum_{i=0}^{n^2} c_i x^i \in F[x]$ tal que $g(T) = 0$, ou seja, $\ker \sigma \neq \{0\}$. Portanto, pelo Teorema 4.30,

$$F[x]/\ker \sigma \simeq F[T] = \{f(T) : f(x) \in F[x]\}.$$

Observe que dividindo pelo coeficiente líder podemos, sem perda de generalidade, escrever $g(x) = \sum_{i=0}^{m-1} d_i x^i + x^m$ como um polinômio mônico. Neste caso, o conjunto

$$S = \{m \in \mathbb{N} : \exists\ f(x) \in F[x] \ \text{ tal que } \ \partial(f) = m \ \text{ e } \ f(T) = 0\}$$

é não vazio. Logo, pelo PBO, S contém um menor elemento, digamos $k \in S$, e denotamos por $m(x) \in F[x]$ o polinômio mônico de grau k tal que $m(T) = 0$.

Afirmação. $\ker \sigma = m(x)F[x] = \{m(x)q(x) : q(x) \in F[x]\}$.
De fato. Dado $q(x) \in F[x]$, obtemos $m(T)q(T) = 0\dot{q}(T) = 0$, de modo que $m(x)F[x] \subseteq \ker \sigma$. Por outro lado, para qualquer $f(x) \in \ker \sigma$, existem únicos $q(x), r(x) \in F[x]$

tais que

$$f(x) = m(x)q(x) + r(x), \text{ com } r(x) = 0 \text{ ou } \partial(r) < \partial(m).$$

Assim, $r(T) = 0$ implica, pela minimalidade do grau de $m(x)$, que $r(x) = 0$. Logo, $f(x) = m(x)q(x) \in m(x)F[x]$ e $\ker \sigma \subseteq m(x)F[x]$. Portanto, $f(x)g(x) \in \ker \sigma$, para todo $g(x) \in \ker \sigma$ e $f(x) \in F[x]$ e diremos que $\ker \sigma$ é um ideal na álgebra linear $F[x]$, de modo que podemos concluir: se $f(x) \in F[x]$ e $f(T) = 0$, então $m(x)$ é um fator de $f(x)$. Por isto, o polinômio $m(x)$ chama-se *polinômio minimal* de T. Posto de outro forma, $m(x)$ é o único polinômio mônico que satisfaz as seguintes condições:

1. $m(T) = 0$.
2. $m(x)$ é o polinômio de menor grau dentre aqueles que anulam T.
3. Se $f(x) \in F[x]$ e $f(T) = 0$, então $m(x)$ é um fator de $f(x)$.

Note que o polinômio minimal $m(x)$ não necessita ser irredutível sobre F, confira o Exemplo 5.21.

Sejam V um espaço vetorial sobre F, com $\dim V = n$, α uma base de V e $T \in \mathcal{L}(V)$. Então, é fácil verificar que, o polinômio minimal de T é o mesmo de $\mathbf{A} = [T]_\alpha^\alpha$.

Exemplo 5.21 *Seja $T \in \mathcal{L}(F^2)$ tal que $\mathbf{A} = [T] = \mathbf{E}_{12}$. Determine o polinômio minimal de T.*

Solução. É claro que o polinômio característico de T é $f_\mathbf{A}(x) = x^2$. Se $\partial(m) = 1$, então $m(x) = x + b$. Assim,

$$m(\mathbf{A}) = \begin{pmatrix} b & 1 \\ 0 & b \end{pmatrix} \neq \mathbf{O},$$

Logo, $\partial(m) \geq 2$. Como $\mathbf{A}^2 = \mathbf{O}$ temos que $\partial(m) = 2$. Portanto, $m(x) = x^2$ não é irredutível sobre F. ■

Lema 5.22 *Matrizes semelhantes possuem o mesmo polinômio minimal.*

Prova. Sejam $\mathbf{A}$ e $\mathbf{B}$ matrizes semelhantes. Então existe uma matriz não singular $\mathbf{P}$ tal que $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. É fácil verificar, indutivamente, que $\mathbf{B}^m = \mathbf{P}^{-1}\mathbf{A}^m\mathbf{P}$, para todo $m \in \mathbb{N}$. Assim, $f(\mathbf{B}) = \mathbf{P}^{-1}f(\mathbf{A})\mathbf{P}$, para todo $f(x) \in F[x]$. Em particular,

$m_{\mathbf{A}}(\mathbf{A}) = \mathbf{O}$ implica que $m_{\mathbf{A}}(\mathbf{B}) = \mathbf{O}$ e $m_{\mathbf{B}}(\mathbf{B}) = \mathbf{O}$ implica que $m_{\mathbf{B}}(\mathbf{A}) = \mathbf{O}$. Portanto, $m_{\mathbf{B}} = m_{\mathbf{A}}$, pois ambos são mônicos. ■

Observe que a recíproca do Lema 5.22 é falsa, pois as matrizes $\mathbf{A} = \mathbf{E}_{12}$ e $\mathbf{B} = \mathbf{E}_{12} + \mathbf{E}_{34}$ possuem o mesmo polinômio minimal $m(x) = x^2$ (prove isto!). Mas, $\mathbf{B} \neq \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, para toda matriz não singular $\mathbf{P}$, pois $\rho(\mathbf{A}) = 1$ e $\rho(\mathbf{B}) = 2$.

Vamos dar um nova olhada no espaço das sequências $V = \mathrm{Seq}_r(F)$ do tipo recorrência. Dados $a_0, \ldots, a_{k-1} \in F$, devemos encontrar as soluções $(x)_{n\in\mathbb{Z}_+} \in V$ da equação

$$x_{n+k} = a_0 x_n + \cdots + a_{k-1}x_{n+k-1}, \ \ \forall \ k, n \in \mathbb{Z}_+.$$

Observe que tais soluções são sequências infinitas:

$$(x_0, \ldots, x_{k-1}, x_k, \ldots, x_n, \ldots).$$

Assim, a função $T : V \to F^k$ definida como $T(\mathbf{x}) = (x_0, \ldots, x_{k-1})$ é um isomorfismo, pois qualquer solução é unicamente determinada por $(x_0, \ldots, x_{k-1})$. Neste caso, pondo $\mathbf{X}_n = (x_n, \ldots, x_{n+k-1})^t$. Então $\mathbf{X}_n$ satisfaz o *sistema companheiro* $\mathbf{X}_{n+1} = \mathbf{C}_k\mathbf{X}_n$, para todo $n \in \mathbb{Z}_+$, em que

$$\mathbf{C}_k = \left(\begin{array}{c|ccc} \mathbf{O} & & \mathbf{I}_{k-1} & \\ \hline -a_0 & -a_1 & \cdots & -a_{k-1} \end{array}\right)$$

é a matriz companheira. Então já vimos, pelo Exercício (12) da Seção 1.3, que o polinômio característico de $\mathbf{C}_k$ é

$$f_{\mathbf{C}_k}(x) = x^k + a_{k-1}x^{k-1} + \cdots + a_1 x + a_0,$$

de modo que $\mathbf{C}_k$ é não singular se, e somente se, $a_0 \neq 0$. Suponhamos que λ seja um autovalor de $\mathbf{C}_k$. Então $\mathbf{X} = (1, \lambda, \ldots, \lambda^{k-1})^t$ é um autovetor de $\mathbf{C}_k$ associado a λ, com $m_g(\lambda) = 1$. De fato. Seja $\mathbf{X} = (y_0, \ldots, y_{k-1})^t$ um autovetor de $\mathbf{C}_k$ associado a λ. Então a equação $\mathbf{C}_k\mathbf{X} = \lambda\mathbf{X}$ implica que $y_i = \lambda y_{i-1}$, com $i = 1, \ldots, k-1$, e

$$-(a_0 y_0 + \cdots + a_{k-1}y_{k-1}) = \lambda y_{k-1}.$$

Observe que $y_0 \neq 0$, pois se $y_0 = 0$, então $y_0 = y_1 = \cdots = y_{k-1} = 0$ e $\mathbf{X} = \mathbf{O}$. Logo, $y_i = \lambda^i y_0$, com $i = 1, \ldots, k-1$, e

$$-(a_0 + \cdots + a_{k-1}\lambda^{k-1})y_0 = \lambda^k y_0,$$

de modo que $f_{\mathbf{C}_k}(\lambda) = 0$. Portanto, escolhendo $y_0 = 1$, $\mathbf{X} = (1, \lambda, \ldots, \lambda^{k-1})^t$ é o autovetor de $\mathbf{C}_k$ associado a λ. Como $\rho(\mu\mathbf{I} - \mathbf{C}_k) = k - 1$, para todo $\mu \in F$, temos que $m_g(\lambda) = 1$. Consequentemente, $\mathbf{C}_k$ é diagonalizável se, e somente se, os seus autovalores são simples. Note, indutivamente, que $\mathbf{X}_n = \mathbf{C}_k^n \mathbf{X}_0$, para todo $n \in \mathbb{Z}_+$, nos fornece a solução. Para completar a nossa discussão sobre $\mathbf{C}_k$. Seja $T \in \mathcal{L}(F^k)$ o operador linear tal que $[T] = \mathbf{C}_k^t$. Então, é fácil verificar que, $T(\mathbf{e}_i) = \mathbf{e}_{i+1}$, com $i = 1, \ldots, k-1$, e

$$T(\mathbf{e}_k) = -(a_0\mathbf{e}_1 + \cdots + a_{k-1}\mathbf{e}_k).$$

Como $T^i(\mathbf{e}_1) = \mathbf{e}_{i+1}$, para cada $i = 1, \ldots, k-1$, temos que

$$T^k(\mathbf{e}_1) = -(a_0 + a_1 T + \cdots + a_{k-1}T^{k-1})(\mathbf{e}_1).$$

Assim, $f_{\mathbf{C}_n}(T)(\mathbf{e}_1) = \mathbf{0}$, de modo que $f_{\mathbf{C}_n}(T)(\mathbf{e}_{i+1}) = f_{\mathbf{C}_n}(T)(T^i(\mathbf{e}_1)) = \mathbf{0}$. Logo, $f_{\mathbf{C}_n}(T) = 0$. Se $g(x) = c_0 + \cdots + +c_m x^m$, com $\partial(g) < \partial(f_{\mathbf{C}_n})$, então

$$\mathbf{0} = g(T)(\mathbf{e}_1) = c_0\mathbf{e}_1 + \cdots + c_m\mathbf{e}_{m+1}$$

implica que $c_0 = \cdots = c_{m+1} = 0$ ou $g(x) = 0$. Portanto, $m(x) = f_{\mathbf{C}_n}(x)$.

Antes de continuarmos com as comprações entre o polinômio característico e minimal, vamos apresentar alguns conceitos sobre matrizes com entradas em $F[x]$, as quais chamam-se x-*matrizes*. Por exemplo, dado $\mathbf{A} \in F^{n\times n}$, já sabemos que a matriz característica $x\mathbf{I} - \mathbf{A}$ de $\mathbf{A}$ é uma x-matriz, pois as entradas diagonais $x - a_{ii}$ são polinômios lineares e as a_{ij}, com $i \neq j$, são polinômios constantes. De modo análogo, $\mathbf{A}, x\mathbf{A}, x^2\mathbf{A}, \ldots$ são x-matrizes. Seja $\mathbf{B} = \mathbf{B}(x) = (f_{ij}(x)) \in F[x]^{n\times n}$ qualquer x-matriz e $\mathbf{B}_m = \mathrm{co}_m(\mathbf{B}) = (\mathrm{co}_m(f_{ij})) \in F^{n\times n}$, com $m = 0, \ldots, k = \max\{\partial(f_{ij})\}$. Então $\mathbf{B}$ pode ser escrita sob a forma

$$\mathbf{B} = \mathbf{B}_0 + \mathbf{B}_1 x + \cdots + \mathbf{B}_k x^k \in F^{n\times n}[x], \text{ com } \mathbf{B}_k \neq \mathbf{O},$$

e diremos que $\mathbf{B}$ possui grau k, isto significa que a função $\phi : F[x]^{n\times n} \to F^{n\times n}[x]$ definida como $\phi(\mathbf{B}) = \sum_{m=0}^{k} \mathbf{B}_m x^m$ é um isomorfismo de álgebras. Pode ser visto, com cautela, que as x-matrizes herdam as propriedades de $F^{n\times n}$. Por exemplo, $\mathbf{B}$ é não singular ou *unimodular* se, e somente se, existir um $\mathbf{C}$ tal que $\mathbf{BC} = \mathbf{CB} = \mathbf{I}$, ou seja, $\mathbf{C} = \mathbf{B}^{-1}$ existe e é uma x-matriz. Em particular, $\mathbf{B} \in F^{n\times n}$ quando as entradas são avaliadas em $\lambda \in F$. Em relação as propriedades de $F[x]$ tem restrições. Por exemplo o AD. Dadas $\mathbf{A}$ e $\mathbf{B}$, com $\mathbf{B}$ unimodular. Então existem únicas x-matrizes $\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{R}_1$ e $\mathbf{R}_2$ tais que $\mathbf{A} - \mathbf{R}_1 = \mathbf{BQ}_1$ e $\mathbf{A} - \mathbf{R}_2 = \mathbf{Q}_2\mathbf{B}$, com o grau de $\mathbf{R}_1$ e $\mathbf{R}_2$ menor do que o grau de $\mathbf{B}$. Neste caso, temos a divisão à esquerda e à direita.

Lema 5.23 *Seja* $\mathbf{B}$ *qualquer* x*-matriz. Então* $\mathbf{B}$ *é não singular se, e somente se,* $\det \mathbf{B} = c \in F$*, com* $c \neq 0$.

Prova. Note que o cofator $c_{ij} = (-1)^{i+j} \det(\mathbf{B}_{ij})$ é claramente um elemento de $F[x]$. Assim, $\operatorname{adj} \mathbf{B}$ é uma x-matriz. Se $\det \mathbf{B} = c \in F$, com $c \neq 0$, então $\mathbf{B}^{-1} = c^{-1} \operatorname{adj} \mathbf{B}$ existe e é uma x-matriz. A recíproca é clara. ∎

Por exemplo, a x-matriz de ordem 2 e grau 2,

$$\mathbf{B} = \begin{pmatrix} (x+1)^2 & x \\ x+2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} x + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} x^2,$$

é não singular, pois $\det \mathbf{B} = 1 \neq 0$.

Dado $a \in F$, é bastante conhecida em $F[x]$, a fatoração $x^2 - a^2 = (x+a)(x-a)$ e, indutivamente, obtemos

$$x^k - a^k = (x^{k-1} + x^{k-2}a + \cdots + xa^{k-2} + a^{k-1})(x-a), \quad \forall\ k \in \mathbb{N}. \tag{5.10}$$

Para cada $\mathbf{A} \in F^{n\times n}$, temos que $(x\mathbf{I})\mathbf{A} = \mathbf{A}(x\mathbf{I})$ e, de modo análogo,

$$x^k\mathbf{I} - \mathbf{A}^k = (x^{k-1}\mathbf{I} + x^{k-2}\mathbf{A} + \cdots + x\mathbf{A}^{k-2} + \mathbf{A}^{k-1})(x\mathbf{I} - \mathbf{A}), \quad \forall\ k \in \mathbb{N}.$$

Isto pode ser provado direto desenvolvendo o lada direito (soma telescópica). Pondo

$$\mathbf{L}_k = \mathbf{L}_k(x\mathbf{I}, \mathbf{A}) = x^{k-1}\mathbf{I} + x^{k-2}\mathbf{A} + \cdots + x\mathbf{A}^{k-2} + \mathbf{A}^{k-1},$$

com $\mathbf{L}_0 = \mathbf{O}$ e $\mathbf{L}_1 = \mathbf{I}$, obtemos

$$x^k\mathbf{I} - \mathbf{A}^k = \mathbf{L}_k(x\mathbf{I} - \mathbf{A}), \quad \forall\ k \in \mathbb{Z}_+. \tag{5.11}$$

Teorema 5.24 (Teorema de Cayley-Hamilton) *Seja* $\mathbf{A} \in F^{n\times n}$*. Se* $f_{\mathbf{A}}(x)$ *for o polinômio característico de* $\mathbf{A}$*, então* $f_{\mathbf{A}}(\mathbf{A}) = \mathbf{O}$.

Prova. Seja

$$\det(x\mathbf{I} - \mathbf{A}) = f_{\mathbf{A}}(x) = c_0 + c_1 x + \cdots + c_{n-1} + x^n \in F[x]$$

o polinômio característico de $\mathbf{A}$. Como os cofatores $c_{ij} = (-1)^{i+j}\det(x\mathbf{I} - \mathbf{A})_{ij}$ são claramente polinômios de grau no máximo $n-1$ temos que $\mathbf{B} = \operatorname{adj}(x\mathbf{I} - \mathbf{A})$ é uma

x-matriz. Por outro lado, pelo Teorema 1.26,

$$\mathbf{B}(x\mathbf{I}-\mathbf{A}) = f_{\mathbf{A}}(x)\mathbf{I} \text{ e } (x\mathbf{I}-\mathbf{A})\mathbf{B} = f_{\mathbf{A}}(x)\mathbf{I}, \tag{5.12}$$

ou seja, $f_{\mathbf{A}}(x)\mathbf{I}$ é divisível à direita e à esquerda por $x\mathbf{I}-\mathbf{A}$. Como

$$f_{\mathbf{A}}(x)\mathbf{I} - f_{\mathbf{A}}(\mathbf{A}) = \sum_{k=0}^{n} c_k(x^k\mathbf{I}-\mathbf{A}^k)$$

temos, pela equação (5.11), que

$$f_{\mathbf{A}}(\mathbf{A}) = f_{\mathbf{A}}(x)\mathbf{I} - \sum_{k=0}^{n} c_k\mathbf{L}_k(x\mathbf{I}-\mathbf{A}).$$

Assim, pela equação (5.12), $f_{\mathbf{A}}(\mathbf{A}) = \mathbf{C}(x\mathbf{I}-\mathbf{A})$, com $\mathbf{C} = \mathbf{B} - \sum_{k=0}^{n} c_k\mathbf{L}_k$ uma x-matriz. Suponhamos, por absurdo, que $\mathbf{C} \neq \mathbf{O}$. Então $\mathbf{C}$ possui grau $m \leq n-1$. Logo, expandindo a equação, obtemos

$$f_{\mathbf{A}}(\mathbf{A}) = (\mathbf{C}_0 + \mathbf{C}_1x + \cdots + \mathbf{C}_mx^m)(x\mathbf{I}-\mathbf{A}),$$

com $\mathbf{C}_m \neq \mathbf{O}$. Efetuando a multiplicação do lado direito, temos que o coeficiente de x^{m+1} é $\mathbf{C}_m$, o que é uma contradição, pois $f_{\mathbf{A}}(\mathbf{A})$ é uma matriz de entradas constantes (os coeficientes em x são constantes). Portanto, $f_{\mathbf{A}}(\mathbf{A}) = \mathbf{0}$. ■

Corolário 5.25 *Seja $\mathbf{A} \in F^{n\times n}$. Se $f_{\mathbf{A}}(x)$ e $m(x)$ são os polinômios característico e minimal de $\mathbf{A}$. Então $m(x)$ é um fator de $f_{\mathbf{A}}(x)$.*

Prova. Fica como um exercício. ■

De posse do Teorema de Cayley-Hamilton, vamos apresentar um procedimento para obter a adjunta clássica, a inversa e o autoespaço de uma matriz $\mathbf{A} \in \mathbb{R}^{n\times n}$. Para isto, sejam $F(x) = \mathbf{B} = \operatorname{adj}(x\mathbf{I}-\mathbf{A})$ e

$$f(x) = f_{\mathbf{A}}(x) = x^n + b_1x^{n-1} + b_2x^{n-2} + \cdots + b_{n-1}x + b_n \in F[x].$$

Então, para qualquer $a \in \mathbb{R}$, temos, pela equação (5.10), que

$$f(x) - f(a) = \delta(x,a)(x-a), \tag{5.13}$$

em que $\delta(x, a)$ é um polinômio de grau $n-1$ em x e a, ou seja,

$$\delta(x, a) = x^{n-1} + c_1 x^{n-2} + c_2 x^{n-3} + \cdots + c_{n-1},$$

com $c_1 = a + b_1, c_2 = a^2 + b_1 a + b_2, \ldots, c_{n-1} = a^{n-1} + \sum_{k=1}^{n-1} b_k a^{n-k-1}$. Note que

$$f'(a) = \lim_{x \to a} \delta(x, a).$$

Como $(x\mathbf{I})\mathbf{A} = \mathbf{A}(x\mathbf{I})$ temos, pela a equação (5.13), que

$$f(x)\mathbf{I} = \delta(x\mathbf{I}, \mathbf{A})(x\mathbf{I} - \mathbf{A}).$$

Assim, pela a equação (5.12) e a unicidade do resto, $\mathbf{B} = \delta(x\mathbf{I}, \mathbf{A})$. Portanto,

$$F(x) = \mathbf{B} = x^{n-1}\mathbf{I} + \mathbf{B}_1 x^{n-2} + \mathbf{B}_2 x^{n-3} + \cdots + \mathbf{B}_{n-1},$$

em que $\mathbf{B}_0 = \mathbf{I}$ e

$$\mathbf{B}_k = \mathbf{A}^k + b_1 \mathbf{A}^{k-1} + b_2 \mathbf{A}^{k-2} + \cdots + b_{k-1}\mathbf{A} + b_k \mathbf{I}, k = 1, 2, \ldots, n-1.$$

Observe que $\mathbf{B}_n = f(\mathbf{A}) = \mathbf{O}$ e, indutivamente,

$$\mathbf{B}_k = \mathbf{B}_{k-1}\mathbf{A} + b_k \mathbf{I} \text{ e } \mathbf{B}_{n-1}\mathbf{A} + b_n \mathbf{I} = \mathbf{0}, k = 1, \ldots, n-1.$$

Pondo $g(x) = x\mathbf{I} - \mathbf{A}$, obtemos

$$\frac{d^{(p)}(Fg)}{dx}(x) = \frac{d^{(p)}(gF)}{dx}(x) = f^{(p)}(x)\mathbf{I}, \ \forall \ p \in \mathbb{Z}_+.$$

(1.) Se $\mathbf{A}$ for não singular, então $b_n = (-1)^n \det \mathbf{A} \neq 0$ e $\mathbf{A}^{-1} = -b_n^{-1}\mathbf{B}_{n-1}$.
(2.) Se $\lambda \in \mathbb{R}$ for um autovalor de $\mathbf{A}$, então $f(\lambda) = 0$ e

$$g(\lambda)F(\lambda) = (\lambda\mathbf{I} - \mathbf{A})\mathbf{B} = \mathbf{O} \Leftrightarrow g(\lambda)F(\lambda) = \lambda F(\lambda).$$

Se $m_a(\lambda) = 1$, então $\rho(g(\lambda)) = n - 1$ e $\delta(\lambda, a) = f'(\lambda) \neq 0$, de modo que $F(\lambda) = \mathbf{B} \neq 0$. Assim, pela Desigualdade de Sylvester, $\rho(\mathbf{B}) = 1$, ou seja, quaisquer duas colunas de $\mathbf{B}$ são LD. Portanto, qualquer coluna não nula $\mathbf{C}_j$ de $\mathbf{B}$ é tal que $\mathbf{A}\mathbf{C}_j = \lambda\mathbf{C}_j$, isto é, $\mathbf{C}_j$ é um autovetor de $\mathbf{A}$ associado a λ. Em geral, se $m_a(\lambda) = m$, então $\rho(\lambda\mathbf{I} - \mathbf{A}) = n - m_a(\lambda) \geq n - m$, de modo que $F(\lambda) = \mathbf{B} = \mathbf{0}$ e, pela Desigualdade de Sylvester, $\rho(\mathbf{B}) \leq m$. Como $f^{(i)}(\lambda) = 0$, para $i = 0, \ldots, m-1$, e $f^{(m)}(\lambda) \neq 0$

temos, pelo Teorema 5.5, que $c_m \neq 0$. Assim, existe pelo menos um menor de ordem $n-m$ não nulo, ou seja, todos os menores de ordem r, com $r < n-m$, possuem um fator $(x-\lambda)$. Por exemplo, se $m=2$, então

$$\delta^{(1)}(x,a) = -\frac{f(x)-f(a)}{(x-a)^2} - \frac{f^{(1)}(x)}{(x-a)}$$

implica que $\delta^{(1)}(\lambda,a) = Q(\lambda) \neq 0$. Logo, $F(x) = (x-\lambda)^k Q(x)$, com $k \leq m$. Portanto,

$$g(x)Q(x) = Q(x)g(x) = q(x)\mathbf{I}.$$

Por outro lado, pelo exposto a seguir, podemos escolher $q(x) = m(x)$ o polinômio minimal de $\mathbf{A}$. Assim, temos o caso já discutido:

$$g(\lambda)F^{(m-1)}(\lambda) = \mathbf{O} \Leftrightarrow g(\lambda)F^{(m-1)}(\lambda) = \lambda F^{(m-1)}(\lambda).$$

Note que se $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ são as raízes de $f(x) = f_{\mathbf{A}}(x)$, não necessariamente distintas, então

$$F(\lambda_j) = \delta(\lambda_j, a) = \prod_{i \neq j} (\lambda_i - a)$$

Vamos resumir o que foi visto: se os coeficientes de $f_{\mathbf{A}}(x)$ forem conhecidos, então $\mathbf{B} = \delta(x\mathbf{I}, \mathbf{A}) = f(x)(x\mathbf{I} - \mathbf{A})^{-1}, \mathbf{A}^{-1} = -b_n^{-1}\mathbf{B}^{n-1}$ e as colunas não nulas de $\mathbf{B}$ e/ou de suas derivadas são os autovetores de $\mathbf{A}$ associados a λ.

Exemplo 5.26 *Seja $T \in \mathcal{L}(\mathbb{R}^3)$ definido como*

$$T(x,y,z) = (-x+2y-2z, 4x-3y+4z, 4x-4y+5z).$$

Determine as propriedades espectrais T.

Solução. Pelo Exemplo 5.9, temos

$$\mathbf{A} = \begin{pmatrix} -1 & 2 & -2 \\ 4 & -3 & 4 \\ 4 & -4 & 5 \end{pmatrix} \text{ e } f(x) = f_{\mathbf{A}}(x) = x^3 - x^2 - x + 1 \in \mathbb{R}[x].$$

Como $\delta(x,x_0) = x^2 + (x_0 - 1)x + (x_0^2 - x_0 - 1)$ temos que

$$F(x) = \mathbf{B} = \delta(x\mathbf{I}, \mathbf{A}) = \mathbf{I}x^2 + (\mathbf{A} - \mathbf{I})x + (\mathbf{A}^2 - \mathbf{A} - \mathbf{I}).$$

Mas,

$$\mathbf{B}_1 = \mathbf{A} - \mathbf{I} = \begin{pmatrix} -2 & 2 & -2 \\ 4 & -4 & 4 \\ 4 & -4 & 4 \end{pmatrix} \text{ e } \mathbf{B}_2 = \mathbf{A}^2 - \mathbf{A} - \mathbf{I} = \begin{pmatrix} 1 & -2 & 2 \\ -4 & 3 & -4 \\ -4 & 4 & -5 \end{pmatrix}$$

implicam que

$$F(x) = \begin{pmatrix} (x-1)^2 & 2(x-1) & -2(x-1) \\ 4(x-1) & (x-1)(x-3) & 4(x-1) \\ 4(x-1) & -4((x-1)) & (x-1)(x+5) \end{pmatrix}.$$

Portanto, $\det \mathbf{A} = 1$ e $\mathbf{A}^{-1} = -\mathbf{B}_2$. É fácil verificar que $f(x) = (x+1)(x-1)^2$, de modo que $\lambda_1 = -1$ e $\lambda_2 = 1$ são os autovalores de T. Para $\lambda_1 = -1$ temos, pela primeira coluna de $F(x)$, que $\mathbf{v}_1 = (-1, 2, 2)$ é o único autovetor de T associado a λ_1. Para $\lambda_2 = 1$, obtemos $F(x) = \mathbf{O}$. Neste caso, devemos usar a matriz $F^{(1)}(x)$, para obter $\mathbf{v}_2 = (0, 1, 1)$ o único autovetor de T associado a λ_2. ■

Exemplo 5.27 *Seja $T \in \mathcal{L}(\mathbb{R}^3)$ definido como*

$$T(x, y, z) = (2x - 2y + 3z, 10x - 4y + 5z, 5x - 4y + 6z).$$

Determine as propriedades espectrais T.

Solução. É fácil verificar que

$$g(x) = \begin{pmatrix} x-2 & 2 & -3 \\ -10 & x+4 & -5 \\ -5 & 4 & x-6 \end{pmatrix} \quad \text{e} \quad f(x) = (x-1)^2(x-2) \in \mathbb{R}[x].$$

Assim, $F(1) = g(1)g(2)$, $F^{(1)}(1) = g(2)$ e $F(2) = g(1)^2$:

$$\begin{pmatrix} -5 & -2 & 5 \\ -25 & -10 & 25 \\ -15 & -6 & 15 \end{pmatrix}, \begin{pmatrix} 0 & 2 & -3 \\ -10 & 6 & -5 \\ -5 & 4 & -4 \end{pmatrix} \text{ e } \begin{pmatrix} -4 & -4 & 8 \\ -15 & -15 & 30 \\ -10 & -10 & 20 \end{pmatrix}.$$

Observe que $F(1)$ e $F^{(1)}(1)$ juntas devem possuir duas colunas LI: $\rho(F(1)) = 1$ e $\mathbf{v}_1 = (1, 5, 3)$ é o único autovetor de T associado a 1. De modo análogo, a primeira coluna de $F^{(1)}(1)$ produz outro autovetor $\mathbf{v}_2 = (0, 2, 1)$ de T associado a 1. Em particular, é fácil verificar que as colunas $\mathbf{C}_2$ e $\mathbf{C}_3$ de $F^{(1)}(1)$ são combinações lineares de $\mathbf{v}_1$ e $\mathbf{v}_2$.

Em geral, se $m_a(\lambda) = m$, então existem exatamente m colunas não nulas LI entre as matrizes $F^{(i)}(\lambda)$, para $i = 0, 1, \ldots, m-1$. Finalmente, como $\rho(F(2)) = 1$ temos que $\mathbf{v}_3 = (4, 15, 10)$ é o único autovetor de T associado a 2. Portanto, T é diagonalizável. ■

Teorema 5.28 *Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Então os polinômios característico e minimal de T possuem as mesmas raízes, a menos de multiplicidades.*

Prova. Sejam $m(x)$ o polinômio minimal de T e $\lambda \in F$. Devemos provar que $m(\lambda) = 0$ se, e somente se, λ é um autovalor de T. Suponhamos que $m(\lambda) = 0$. Então, pelo AD, existe um $q(x) \in F[x]$ tal que $m(x) = (x - \lambda)q(x)$. Como $\partial(q) < \partial(m)$ temos que $q(T) \neq 0$. Assim, existe um $\mathbf{w} \in V$, $\mathbf{w} \neq \mathbf{0}$, tal que $\mathbf{u} = q(T)(\mathbf{w}) \neq \mathbf{0}$. Logo,

$$\mathbf{0} = m(T)(\mathbf{w}) = (T - \lambda I)q(T)(\mathbf{w}) = (T - \lambda I)(\mathbf{u}).$$

Portanto, λ é um autovalor de T e $\mathbf{u}$ é o autovetor associado a λ. Reciprocamente, suponhamos que λ seja um autovalor de T. Então existe um $\mathbf{u} \in V$, com $\mathbf{u} \neq \mathbf{0}$, tal que $T(\mathbf{u}) = \lambda\mathbf{u}$. Assim, pelo Lema 5.22, obtemos $m(\lambda)\mathbf{u} = m(T)(\mathbf{u}) = \mathbf{0}$. Como $\mathbf{u} \neq \mathbf{0}$ temos que $m(\lambda) = 0$. Uma outra prova. $m(x) = (x - \lambda)q(x) + r$, onde $r \in F$. Se $r \neq 0$, então $(T - \lambda I)(-r^{-1}q(T)) = I$, de modo que $T - \lambda I$ é não singular, o que é impossível. ■

Exemplo 5.29 *Seja $T \in \mathcal{L}(V)$, com $f(x) = (x-3)^2(x-1)^3(x+5)$ o polinômio característico. Determine a $\dim V$ e os candidatos a polinômio minimal de T.*

Solução. É claro que $\dim V = 6$. Pelo Coroário 5.25, os candidatos a polinômio minimal de T são:

$$\begin{aligned}
m(x) &= (x-3)(x-1)(x+5) \\
m(x) &= (x-3)^2(x-1)(x+5) \\
m(x) &= (x-3)(x-1)^2(x+5) \\
m(x) &= (x-3)(x-1)^3(x+5) \\
m(x) &= (x-3)^2(x-1)^2(x+5)
\end{aligned}$$

e o próprio polinômio característico, ou seja, $m(x) = f(x)$. ■

Observe que o Teorema 5.28 afirma que o polinômio minimal possui as mesmas raízes do polinômio característico, as quais não podem serem facilmente calculadas. Por isto, apresentamos um método alternativo para determinar o polinômio minimal sem precisar conhecer, a priori, suas raízes. Para isto, devemos generalizar as operações elementares sobre uma matriz $\mathbf{A} \in F^{n \times n}$ para uma x-matriz, o leitor interessado em mais detalhes pode consultar Hoffman-Kunze [10]. As operações elementares sobre as linhas (as colunas) de uma x-matriz $\mathbf{A}$ são:

1. Permutação das i-ésima e j-ésima linhas ou $\mathbf{P}_{ij} = \mathbf{I} - \mathbf{E}_{ii} + \mathbf{E}_{ij} + \mathbf{E}_{ji} - \mathbf{E}_{jj}$.

2. Multiplicação da i-ésima linha por um escalar não nulo c ou $\mathbf{D}_i(c) = \mathbf{I} + (c - 1)\mathbf{E}_{ii}$.

3. Substituição da i-ésima linha pela i-ésima linha mais $f(x)$ vezes a j-ésima linha, $i \neq j$, ou $\mathbf{T}_{ij}(f(x)) = \mathbf{I} + f(x)\mathbf{E}_{ij}$.

É pertinente ressaltar que estas operações são idênticas as sobre F, exceto que em (2) só é permitido os elementos unidades (polinômios constantes não nulos) de F.

Exemplo 5.30 *Determine uma matriz equivalente a x-matriz*

$$\mathbf{A} = \begin{pmatrix} a_{11}(x) & a_{12}(x) \\ a_{21}(x) & a_{22}(x) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Solução. Se $\mathbf{A} = \mathbf{O}$, nada há para ser provado. Se $\mathbf{A} \neq \mathbf{O}$, então existe um a_{ij} em $\mathbf{A}$ tal que $a_{ij} \neq 0$. Então, permutando a i-ésima linha com a primeira linha ($L_i \leftrightarrow L_1$) e a j-ésima coluna com a primeira coluna ($C_j \leftrightarrow C_1$), podemos supor que $a_{11} \neq 0$. Seja $d = \operatorname{mdc}(a_{11}, a_{21})$. Então existem $b_{11}, b_{21} \in F[x]$ tais que $d = a_{11}b_{11} + a_{21}b_{21}$. Pondo $a_{11} = c_{11}d$ e $a_{21} = c_{21}d$, obtemos $b_{11}c_{11} + b_{21}c_{21} = 1$ e

$$\begin{pmatrix} b_{11} & b_{21} \\ -c_{21} & c_{11} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} d & a_{12}b_{11} + a_{22}b_{21} \\ a_{21}c_{11} - a_{11}c_{21} & a_{22}c_{11} - a_{12}c_{21} \end{pmatrix}.$$

Como d é um fator de $a_{21}c_{11} - a_{11}c_{21}$ temos que $a_{21}c_{11} - a_{11}c_{21} = qd$, para algum $q \in F[x]$. Assim, efetuando a operação $L_2 \to L_2 - qL_1$, obtemos

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \longrightarrow \cdots \longrightarrow \begin{pmatrix} d & d_{12} \\ 0 & d_{22} \end{pmatrix}.$$

Seja $f = \operatorname{mdc}(d, d_{12})$. Então seguindo as mesma etapas acima à direita, teremos

$$\begin{pmatrix} d & d_{12} \\ 0 & d_{22} \end{pmatrix} \longrightarrow \cdots \longrightarrow \begin{pmatrix} d_1 & 0 \\ 0 & f \end{pmatrix}.$$

Sendo $\partial(d_1) \leq \partial(f)$ ou $\partial(d_1) \geq \partial(f)$. Então permutando, se necessário, as linhas e colunas podemos supor que $\partial(d_1) \leq \partial(f)$. Logo, pelo AD, existem $q, r \in F[x]$ tais que

$f = qd_1 + r$, com $r = 0$ ou $\partial(r) < \partial(d_1)$. Se $r = 0$, para. Se $r \neq 0$, então aplicando sucessivamente o AD até chegar em $f = q_1 d_1$. Portanto,

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \longrightarrow \cdots \longrightarrow \mathbf{D} = \begin{pmatrix} d_1 & 0 \\ 0 & q_1 d_1 \end{pmatrix} = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix},$$

com d_1 um fator de d_2. Consequentemente, existem x-matrizes unimodulares $\mathbf{P}$ e $\mathbf{Q}$ tais que $\mathbf{PAQ} = \mathbf{D}$ e $\mathbf{A}$ é unimodular se, e somente se, $d_1 d_2 \neq 0$. ■

É importante ressaltar, na prova do Exemplo 5.30, que os polinômios d_1 e d_2 são mônicos e se d_1 for um polinômio constante não nulo, então ele deve ser substituído por 1.

Teorema 5.31 *Qualquer x-matriz $\mathbf{A}$ é equivalente a uma matriz sob a forma*

$$\mathbf{D}_r = \operatorname{diag}(a_1(x), \ldots, a_r(x), 0, \ldots, 0),$$

com os $a_i(x)$ mônicos e $a_i(x)$ um fator de $a_{i+1}(x)$, para $i = 1, \ldots, r-1$.

Prova. Confira o Exemplo 5.30. ■

Seja $\mathbf{A}$ qualquer x-matriz. Então o posto "determinante" de $\mathbf{A}$, $r = \rho(\mathbf{A})$, é igual a ordem de seu maior menor não nulo. Por exemplo, $r = \rho(\mathbf{D}_r)$. Além disso, pelo Lema 5.23, $n = \rho(\mathbf{A})$ se, e somente se, $\mathbf{A}$ for unimodular. Portanto, usando o Corolário 1.30 (a Fórmula de Binet-Cauchy), podemos provar que se $\mathbf{A}$ for equivalente a $\mathbf{B}$, então $a_i(x) = b_i(x)$, para $i = 1, \ldots, n$.

Seja $\mathbf{A}$ qualquer x-matriz, com $r = \rho(\mathbf{A})$, e $d_i(x)$ igual ao máximo divisor comum de todos os menores de ordem i de $\mathbf{A}$, com $i = 1, \ldots, r$, existem $\binom{n}{i}^2$ deles. Então é bem conhecido que qualquer $d_i(x)$, com $i \geq 2$, é uma combinação linear dos $d_{i-1}(x)$, de modo que $d_{i-1}(x)$ é um fator de $d_i(x)$. Portanto, pondo $d_0(x) = 1$, obtemos a sequência

$$d_0(x), d_1(x), \ldots, d_r(x)$$

tal que $d_{i-1}(x)$ é um fator de $d_i(x)$, para $i = 1, \ldots, r$. Em particular, quando $\mathbf{A} = \mathbf{D}_r$ temos, com uma escolha adequada de elementos unidades em F, que podemos escrever

$$d_1(x) = a_1(x), \ldots, d_r(x) = a_1(x) \cdots a_r(x).$$

e $d_{r+j}(x) = 0$, para $j = 1, \ldots, n-r$. Portanto, como acima, os $d_i(x)$ são invariantes por equivalência. Neste caso, os quocientes

$$q_i(x) = \frac{d_i(x)}{d_{i-1}(x)} \in F[x], i = 1, \ldots, r.$$

chamam-se *polinômios invariantes* de $\mathbf{A}$ e

$$d_i(x) = q_1(x) \cdots q_i(x), i = 1, \ldots, r.$$

Além disso, $q_{i-1}(x)$ é um fator de $q_i(x)$ e $d_i^2(x)$ é um fator de $d_{i-1}(x)d_{i+1}(x)$, para $i = 2, \ldots, r$. O próximo resultado é um dos mais importante em Álgebra Linear.

Teorema 5.32 (Forma Normal de Smith) [2] *Qualquer x-matriz $\mathbf{A}$ de posto r é equivalente a uma matriz sob a forma $\mathbf{N}_r = \operatorname{diag}(q_1(x), \ldots, q_r(x), 0, \ldots, 0)$, com os $q_i(x)$ mônicos e $q_i(x)$ um fator de $q_{i+1}(x)$, para $i = 1, \ldots, r-1$.*

Prova. Confira o exposto acima e/ou a Bibliografia. ■

Lema 5.33 *Sejam $\mathbf{A}, \mathbf{B} \in F^{n \times n}$. Então $\mathbf{A}$ é semelhante a $\mathbf{B}$ se, e somente se, $x\mathbf{I} - \mathbf{A}$ for equivalente a $x\mathbf{I} - \mathbf{B}$.*

Prova. Suponhamos que $\mathbf{A}$ seja semelhante a $\mathbf{B}$. Então existe uma matriz não singular $\mathbf{P}$ tal que $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. Assim,

$$x\mathbf{I} - \mathbf{B} = \mathbf{P}^{-1}(x\mathbf{I} - \mathbf{A})\mathbf{P}.$$

Portanto, $x\mathbf{I} - \mathbf{A}$ é equivalente a $x\mathbf{I} - \mathbf{B}$. Reciprocamente, suponhamos que $x\mathbf{I} - \mathbf{A}$ seja equivalente a $x\mathbf{I} - \mathbf{B}$. Então existem matrizes unimodulares $\mathbf{P}$ e $\mathbf{Q}$ tais que

$$x\mathbf{I} - \mathbf{B} = \mathbf{P}(x\mathbf{I} - \mathbf{A})\mathbf{Q} \Leftrightarrow \mathbf{T}(x\mathbf{I} - \mathbf{B}) = (x\mathbf{I} - \mathbf{A})\mathbf{Q},$$

em que $\mathbf{T} = \mathbf{P}^{-1}$. Assim, pelo AD, existem x-matrizes $\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{R}_1$ e $\mathbf{R}_2$ tais que

$$\mathbf{T} = (x\mathbf{I} - \mathbf{A})\mathbf{Q}_1 + \mathbf{R}_1 \ \text{ e } \ \mathbf{Q} = \mathbf{Q}_2(x\mathbf{I} - \mathbf{B}) + \mathbf{R}_2,$$

com $\mathbf{R}_1$ e $\mathbf{R}_2$ independentes de x, pois $x\mathbf{I} - \mathbf{A}$ possui grau 1. Assim, substituindo e comparando, obtemos $\mathbf{Q}_1 = \mathbf{Q}_2$. Logo,

$$\mathbf{R}_1(x\mathbf{I} - \mathbf{B}) = (x\mathbf{I} - \mathbf{A})\mathbf{R}_2,$$

[2] Henry John Stephen Smith, 1826-1883, matemático inglês.

de modo que $\mathbf{R}_1 = \mathbf{R}_2, \mathbf{R}_1\mathbf{B} = \mathbf{A}\mathbf{R}_2$ e $\mathbf{R}_1\mathbf{B} = \mathbf{A}\mathbf{R}_1$. Portanto, basta provar que $\mathbf{R}_1$ é não singular. Novamente, pelo AD, existem x-matrizes $\mathbf{Q}_3$ e $\mathbf{R}_3$ tais que

$$\mathbf{P} = (x\mathbf{I} - \mathbf{B})\mathbf{Q}_3 + \mathbf{R}_3,$$

com $\mathbf{R}_3$ independente de x. Assim, depois de alguns cálculos, teremos

$$\mathbf{I} = \mathbf{T}\mathbf{P} = (x\mathbf{I} - \mathbf{A})[\mathbf{Q}\mathbf{Q}_3 + \mathbf{Q}_1\mathbf{R}_3] + \mathbf{R}_1\mathbf{R}_3.$$

Portanto, $\mathbf{R}_1\mathbf{R}_3 = \mathbf{I}$ e $\det(\mathbf{R}_1) = c \neq 0$. ■

Sejam $T \in \mathcal{L}(V)$, com $\dim V = n$, e $\mathbf{A} = [T]_\alpha^\alpha$, para alguma base α de V. Então $\rho(x\mathbf{I} - \mathbf{A}) = n$, pois as entradas diagonais de $x\mathbf{I} - \mathbf{A}$ são todas polinômios lineares, e $x\mathbf{I}_n - \mathbf{A}$ é equivalente a $\mathbf{N}_r = \operatorname{diag}(q_1(x), \ldots, q_n(x))$, com os $q_i(x)$ mônicos e $q_i(x)$ um fator de $q_{i+1}(x)$, para $i = 1, \ldots, n-1$. Portanto,

$$f_\mathbf{A}(x) = \det(x\mathbf{I}_n - \mathbf{A}) = q_1(x) \cdots q_n(x) = d_{n-1}(x)q_n(x),$$

de modo que o polinômio minimal de T é $m(x) = q_n(x)$.

Exemplo 5.34 *Seja $T \in \mathcal{L}(\mathbb{R}^3)$ definido como*

$$T(x, y, z) = (-x + 2y - 2z, 4x - 3y + 4z, 4x - 4y + 5z).$$

Determine o polinômio minimal de T.

Solução. Pelo exposto, basta escalonar a x-matriz:

$$\begin{pmatrix} x+1 & -2 & 2 \\ -4 & x+3 & -4 \\ -4 & 4 & x-5 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (x+1)(x-1)^2 \end{pmatrix}.$$

Portanto, $m(x) = f_\mathbf{A}(x) = (x+1)(x-1)^2$. ■

Exemplo 5.35 *Seja $\mathbf{C}_n$ uma matriz companheira. Ache seu polinômio minimal.*

Solução. É claro que $x\mathbf{I} - \mathbf{C}_n$ possui menores não nulos de ordem i, com $i = 1, \ldots, n-1$, os quais são independentes de x. Assim, $d_1(x) = \cdots = d_{n-1}(x) = 1$. Como $d_n(x) = \det(x\mathbf{I} - \mathbf{C}_n)$ temos que $q_1(x) = \cdots = q_{n-1}(x) = 1$. Portanto, $m(x) = d_n(x) = f_\mathbf{A}(x)$. Consequentemente, $\mathbf{C}_n$ é equivalente a uma matriz diagonal sob a forma $\mathbf{D} = \operatorname{diag}(1, \ldots, 1, m(x))$. ■

Finalizamos esta seção com mais alguns procedimentos para discutir se uma dada matriz é diagonalizável ou não. Para isto, vamos lembrar alguns conceitos e resultados. Já vimos, no Exercício (36) da Seção 4.2, que se $V = W_1 \oplus W_2$, então a função $E_1 \in \mathcal{L}(V)$ definida como $E_1(\mathbf{v}) = E_1(\mathbf{u}+\mathbf{w}) = \mathbf{u}$ era a *projeção* de V sobre W_1 paralela a W_2, confira Figura 5.1. Neste caso, $E_2 = I_V - E_1$ é a projeção de V sobre W_2 paralela a W_1 e $\ker E_2 = \operatorname{Im} E_1$.

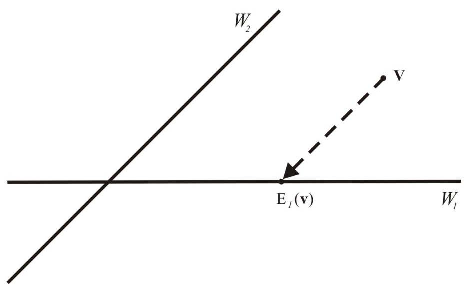


Figura 5.1: Representação gráfica da projeção.

Exemplo 5.36 *Seja $E_1 \in \mathcal{L}(\mathbb{R}^3)$ a projeção sobre o plano W_1, $3x+y-2z=0$, na direção da reta $W_2 = \mathbb{R}[(1,1,1)]$. Determine $[E_1]$ e $[E_1]_\alpha^\alpha$, em que $\alpha = \{(1,-1,0),(2,0,1),(1,2,3)$*

Solução. Note que $\beta = \{(1,-3,0),(0,2,1),(1,1,1)\}$ é uma base de $\mathbb{R}^3$, pois $(1,1,1) \notin W_1 = \mathbb{R}[(1,-3,0),(0,2,1)]$. Assim, é fácil verificar que

$$E_1(x,y,z) = 2^{-1}(-x-y+2z, -3x+y+2z, -3x-y+4z).$$

Portanto,

$$[E_1] = \frac{1}{2}\begin{pmatrix} -1 & -1 & 2 \\ -3 & 1 & 2 \\ -3 & -1 & 4 \end{pmatrix} \text{ e } [E_1]_\alpha^\alpha = \frac{1}{6}\begin{pmatrix} 12 & 12 & 8 \\ -6 & -6 & 9 \\ 0 & 0 & 6 \end{pmatrix}$$

são as matrizes desejadas. ■

Lema 5.37 *Seja $E \in \mathcal{L}(V)$, com $\dim V = n$. Então as seguintes condições são equivalentes*:

1. *E é uma projeção,*

2. *$V = \operatorname{Im} E \oplus \ker E$ e $\mathbf{w} \in \operatorname{Im} E$ se, e somente se, $E(\mathbf{w}) = \mathbf{w}$,*

3. *Existe uma base* α *de* V *tal que* $[E]_\alpha^\alpha = \mathbf{I}_k \oplus \mathbf{O}$. *Em particular,* E *é diagonalizável e* $\operatorname{tr}([E]_\alpha^\alpha) = \rho([E]_\alpha^\alpha) = k$.

4. $E^2 = E$.

Prova. Confira a solução do Exercício (15) da Seção 4.3. ■

Lema 5.38 *Seja* $T \in \mathcal{L}(V)$, *com* $\dim V = n$. *Se* T *for diagonalizável e existir um* $\mathbf{u} \in V$ *tal que* $T^2(\mathbf{u}) = \mathbf{0}$, *então* $T(\mathbf{u}) = \mathbf{0}$.

Prova. Suponhamos que $\mathbf{A} = [T]$ seja diagonalizável. Então existe uma matriz não singular $\mathbf{P}$ tal que $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Consideremos

$$U = \{\mathbf{X} \in F^n : \mathbf{A}\mathbf{X} = \mathbf{O}\} \text{ e } W = \{\mathbf{X} \in F^n : \mathbf{A}^2\mathbf{X} = \mathbf{O}\}.$$

Então $U \subseteq W$. Por outro lado, como $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$ e $\mathbf{P}^{-1}\mathbf{A}^2\mathbf{P} = \mathbf{D}^2$ temos que $\lambda_i = 0$ se, e somente se, $\lambda_i^2 = 0$. Logo, $\rho(\mathbf{A}) = \rho(\mathbf{A}^2)$. Portanto, $U = W$, ou seja, $\mathbf{A}^2\mathbf{X} = \mathbf{O}$ implica que $\mathbf{A}\mathbf{X} = \mathbf{O}$. ■

Teorema 5.39 *Seja* $T \in \mathcal{L}(V)$, *com* $\dim V = n$. *Então as seguintes condições são equivalentes*:

1. T *é diagonalizável*;

2. *Para cada* $\mathbf{u} \in V$ *e* $\lambda \in F$, *se* $(T - \lambda I)^2(\mathbf{u}) = \mathbf{0}$, *então* $(T - \lambda I)(\mathbf{u}) = \mathbf{0}$;

3. *Se* $\mathbf{v}_0 \in V$ *for um autovetor de* T *associado a* $\lambda_0 \in F$, *então* $(T - \lambda_0 I)(\mathbf{u}) \neq \mathbf{v}_0$, *para todo* $\mathbf{u} \in V$, *ou seja,* $V_{\lambda_0} = V^{\lambda_0}$;

4. *As raízes* $\lambda \in F$ *do polinômio minimal de* T *são todas distintas e* $m_a(\lambda) = 1$;

5. *Existe um* $k \in \mathbb{N}$, *escalares distintos* $\lambda_1, \ldots, \lambda_k \in F$ *e* $E_i \in \mathcal{L}(V)$ *não nulos, para* $i = 1, \ldots, k$, *tal que*

$$T = \lambda_1 E_1 + \cdots + \lambda_k E_k,\; E_1 + \cdots + E_k = I \;\text{ e }\; E_i E_j = 0, \;\text{ se }\; i \neq j.$$

Prova. $(1 \Rightarrow 2)$ Seja $\mathbf{A} = [T]_\alpha^\alpha$, para alguma base α de V. Suponhamos que $\mathbf{A}$ seja diagonalizável. Então existe uma matriz não singular $\mathbf{P}$ tal que $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$. Pondo $\mathbf{Y} = \mathbf{P}^{-1}\mathbf{X}$, em que $\mathbf{X} = [\mathbf{u}]_\alpha$, para todo $\mathbf{u} \in V$. Então

$$\mathbf{0} = (\mathbf{A} - \lambda\mathbf{I})^2\mathbf{P}\mathbf{Y} = (\mathbf{P}\mathbf{D}\mathbf{P}^{-1} - \lambda\mathbf{P}\mathbf{P}^{-1})^2\mathbf{P}\mathbf{Y} = \mathbf{P}(\mathbf{D} - \lambda\mathbf{I}_n)^2\mathbf{Y},$$

de modo que $(\mathbf{D} - \lambda\mathbf{I}_n)^2\mathbf{Y} = \mathbf{0}$, pois $\mathbf{P}$ é não singular. Assim, pelo Lema 5.38, $(\mathbf{D} - \lambda\mathbf{I})\mathbf{Y} = \mathbf{0}$. Portanto, $(\mathbf{D} - \lambda\mathbf{I})\mathbf{X} = \mathbf{0}$, ou seja, $(T - \lambda I)(\mathbf{u}) = \mathbf{0}$.

$(2 \Rightarrow 3)$ Suponhamos que $T(\mathbf{v}_0) = \lambda_0\mathbf{v}_0$, com $\mathbf{v}_0 \neq \mathbf{0}$. Se existir um $\mathbf{u} \in V$ tal que $(T - \lambda_0 I)(\mathbf{u}) = \mathbf{v}_0$, então

$$(T - \lambda_0 I)^2(\mathbf{u}) = (T - \lambda_0 I)(\mathbf{v}_0) = \mathbf{0}.$$

Assim, $\mathbf{v}_0 = (T - \lambda_0 I)(\mathbf{u}) = \mathbf{0}$, o que é impossível.

$(3 \Rightarrow 4)$ Suponhamos, por absurdo, que $m(x) = (x - \lambda_0)^2 q(x)$. Então, como $\partial((x - \lambda_0)q) < \partial(m)$, podemos escolher um $\mathbf{w} \in V$, com $\mathbf{w} \neq \mathbf{0}$, tal que $\mathbf{v}_0 = (T - \lambda_0 I)q(T)(\mathbf{w}) \neq \mathbf{0}$. Assim, $(T - \lambda_0 I)(\mathbf{v}_0) = m(T)(\mathbf{v}_0) = \mathbf{0}$ ou $T(\mathbf{v}_0) = \lambda_0\mathbf{v}_0$, de modo que a equação $(T - \lambda_0 I)(\mathbf{u}) = \mathbf{v}_0$ possui uma solução $\mathbf{u} = q(T)(\mathbf{w})$, o que é uma contradição.

$(4 \Rightarrow 5)$ Suponhamos que $m(x) = (x - \lambda_1)\cdots(x - \lambda_k)$, onde $\lambda_1, \ldots, \lambda_k \in F$ são os autovalores distintos de T. Definimos $p_i(x) \in F[x]$ pelas relações

$$m(x) = (x - \lambda_i)p_i(x), i = 1, \ldots, k.$$

Então $p_i(\lambda_i) \neq 0$ e $p_i(\lambda_j) = 0$ se, e somente se, $i \neq j$. Consideramos o polinômio

$$g(x) = 1 - \sum_{i=1}^{k}(p_i(\lambda_i))^{-1}p_i(x).$$

Então $\partial(g) < \partial(m) = k$ e $g(\lambda_i) = 0$, para $i = 1, \ldots, k$. Assim, $g(x) = 0$, de modo que $g(S) = 0$, para todo $S \in \mathcal{L}(V)$. Portanto, os *operadores espectrais*

$$E_i = (p_i(\lambda_i))^{-1}p_i(T) \in \mathcal{L}(V), i = 1, \ldots, k,$$

satisfazem $E_1 + \cdots + E_k = I$, $Tp_i(T) = \lambda_i p_i(T)$ etc.

$(5 \Rightarrow 1)$ Seja $\mathbf{v} \in V$, com $\mathbf{v} \neq \mathbf{0}$. Então é fácil verificar que $\mathbf{u}_i = E_i(\mathbf{v})$ é tal que $T(\mathbf{u}_i) = \lambda_i\mathbf{u}_i$ e $\mathbf{v} = I(\mathbf{v}) = \sum_{i=1}^{k} E_i(\mathbf{v}) = \sum_{i=1}^{k} \mathbf{u}_i$, de modo que todo vetor é combinação linear dos autovetores. Portanto, $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ é uma base de V formada de autovetores de T, isto é, T é diagonalizável. ■

É muito importante ressaltar, pelo item (5) do Teorema 5.39, que se a matriz $\mathbf{A} = \sum_{i=1}^{k} \lambda_i\mathbf{E}_i$, então $\mathbf{A}\mathbf{E}_i = \lambda_i\mathbf{E}_i$ implica que $\lambda_i\mathbf{I} - \mathbf{A}$ é singular, de modo que λ_i é um autovalor de $\mathbf{A}$. Portanto, qualquer vetor não nulo sob a forma $\mathbf{E}_i\mathbf{X}$ é um autovetor de $\mathbf{A}$ associado a λ_i. Consequentemente, V_{λ_i} é gerado pelos vetores colunas da matriz

$\mathbf{E}_i$. Além disso, vale lembrar que se $\mathbf{A}$ for diagonalizável, então, pelo Teorema 5.12, obtemos

$$\mathbf{A} = \lambda_1 \mathbf{X}_1 \mathbf{Y}_1^t + \cdots + \lambda_n \mathbf{X}_n \mathbf{Y}_n^t,$$

com $\mathbf{X}_i$ um autovetor de $\mathbf{A}$, $\mathbf{Y}_i$ um autovetor de $\mathbf{A}^t$ associado a λ_i e $\mathbf{Y}_i^t \mathbf{X}_j = \delta_{ij}$. Em particular, se $\mathbf{A}$ for idempotente, então $\mathbf{A} = \mathbf{X}_1 \mathbf{Y}_1^t + \cdots + \mathbf{X}_k \mathbf{Y}_k^t$, com $\rho(\mathbf{A}) = k$.

Exemplo 5.40 *Seja $T \in \mathcal{L}(F^3)$ definido como*

$$T(x, y, z) = (2x + 2y - 5z, 3x + 7y - 15z, x + 2y - 4z).$$

T é diagonalizável?

Solução. Pelo exposto, basta escalonar a x-matriz:

$$\begin{pmatrix} x-2 & -2 & 5 \\ -3 & x-7 & 15 \\ -1 & -2 & x+4 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & x-1 & 0 \\ 0 & 0 & (x-1)(x-3) \end{pmatrix}.$$

Assim, $m(x) = (x-1)(x-3)$ é o polinômio minimal de T. Portanto, T é diagonalizável. Neste caso, $p_1(x) = x - 3$ e $E_1 = -2^{-1}(T - 3I)$; $p_2(x) = x - 1$ e $E_2 = 2^{-1}(T - I)$, de modo que $E_1 E_2 = 0$, $E_1 + E_2 = I$, $E_1^2 = E_1$ e $E_2^2 = E_2$ ou na forma matricial

$$\mathbf{E}_1 = \frac{1}{2}\begin{pmatrix} 1 & -2 & 5 \\ -3 & -4 & 15 \\ -1 & -2 & 7 \end{pmatrix} \text{ e } \mathbf{E}_2 = -\frac{1}{2}\begin{pmatrix} -1 & -2 & 5 \\ -3 & -6 & 15 \\ -1 & -2 & 5 \end{pmatrix}$$

implicam que $\mathbf{A} = \mathbf{E}_1 + 3\mathbf{E}_2$. Observe que V_1 é gerado pelos vetores colunas da matriz $\mathbf{E}_1$ e V_3 é gerado pelos vetores colunas da matriz $\mathbf{E}_2$, pois é fácil verificar que $V_1 = F[(-2, 1, 0), (5, 0, 1)]$ e $V_3 = F[(1, 3, 1)]$. ■

É muito importante, de posse destas informações, dar uma aproximação sucessiva da raiz quadrada de um número real $a > 0$. Para isto, seja $(z_n)_{n \in \mathbb{N}} \in \mathrm{Seq}_r(\mathbb{R})$, em que

$$z_{n+1} = \frac{x_{n+1}}{y_{n+1}}, x_1 = y_1 = 1, x_{n+1} = x_n + a y_n \text{ e } y_{n+1} = x_n + y_n.$$

Então, pondo $\mathbf{X}_{n+1} = (x_{n+1}, y_{n+1})^t$, obtemos

$$\mathbf{X}_{n+1} = \mathbf{A}\mathbf{X}_n, \quad \forall \ n \in \mathbb{N}, \quad \text{e} \quad \mathbf{A} = \mathbf{E}_{11} + a\mathbf{E}_{12} + \mathbf{E}_{21} + \mathbf{E}_{22}.$$

Assim, recursivamente, teremos $\mathbf{X}_n = \mathbf{A}^{n-1}\mathbf{X}_1$, para todo $n \in \mathbb{N}$. Neste caso, basta conhecer $\mathbf{A}$ para obter as sequências $(x_n)_{n\in\mathbb{N}}$ e $(y_n)_{n\in\mathbb{N}}$. É fácil verificar que $f_{\mathbf{A}}(x) = x^2 - 2x + 1 - a$ e $\lambda_1 = 1 - \sqrt{a}, \lambda_2 = 1 + \sqrt{a}$, com autovetores associados $\mathbf{v}_1 = (-\sqrt{a}, 1)$, $\mathbf{v}_2 = (\sqrt{a}, 1)$. Logo, existe uma matriz não singular $\mathbf{P}$, cujas colunas são $\mathbf{v}_1$ e $\mathbf{v}_2$, tal que $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \operatorname{diag}(\lambda_1, \lambda_2)$. Pondo

$$\mathbf{Y}_1 = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \mathbf{P}^{-1}\mathbf{X}_1 = \begin{pmatrix} -\frac{1}{2\sqrt{a}} & \frac{1}{2} \\ \frac{1}{2\sqrt{a}} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} - \frac{1}{2\sqrt{a}} \\ \frac{1}{2} + \frac{1}{2\sqrt{a}} \end{pmatrix},$$

teremos

$$\mathbf{X}_n = \mathbf{A}^{n-1}\mathbf{X}_1 = \mathbf{P}\mathbf{D}^{n-1}\mathbf{Y}_1 = \begin{pmatrix} \sqrt{a}(\lambda_2^{n-1}y_2 - \lambda_1^{n-1}y_1) \\ \lambda_2^{n-1}y_2 + \lambda_1^{n-1}y_1 \end{pmatrix},$$

de modo que $x_n = \sqrt{a}(\lambda_2^{n-1}y_2 - \lambda_1^{n-1}y_1)$ e $y_n = \lambda_2^{n-1}y_2 + \lambda_1^{n-1}y_1$, para todo $n \in \mathbb{N}$. Por outro lado, depois de alguns cálculos, obtemos

$$z_n = \frac{x_n}{y_n} = \sqrt{a}\left(\frac{1 - y_2^{-1}y_1(\lambda_2^{-1}\lambda_1)^{n-1}}{1 + y_2^{-1}y_1(\lambda_2^{-1}\lambda_1)^{n-1}}\right).$$

Como $0 < |\lambda_2^{-1}\lambda_1| < 1$ temos que $\lim_{n\to\infty} z_n = \sqrt{a}$. Portanto, qualquer número real $a > 0$ possui uma raiz quadrada real.

Exercícios

1. Para cada $T \in \mathcal{L}(V)$ do Exercício (1) da Seção 5.1, determine o polinômio minimal e identifique os que são diagonalizáveis.

2. Seja $T \in \mathcal{L}(V)$, com $f(x) = (x-1)^2(x-4)^3(x+2)$ o polinômio característico.

   (a) Determine a $\dim V$.

   (b) Quais são as possibilidades para o polinômio minimal de T?

   (c) Se T é diagonalizável, qual é o seu polinômio minimal?

3. Seja $T \in \mathcal{L}(V)$, com $\dim V = 5$ e $m(x) = (x-1)^2(x-2)$ o polinômio minimal.

   (a) Quais são as possibilidades para o polinômio característico de T?

   (b) T é diagonalizável?

4. Seja $T \in \mathcal{L}(\mathbb{R}^2)$ definido como $T(x, y) = (ax + by, cx + dy)$. Determine condições necessárias e suficientes em a, b, c e d, de modo que T não seja semelhante a uma matriz diagonal.

5. Seja $\mathbf{A} \in F^{4\times 4}$, com autovalores distintos $\lambda = 1$ e $\mu = -1$.

   (a) Escreva todas as possibilidades para o polinômio característico de $\mathbf{A}$.

   (b) Para cada possibilidade do polinômio característico de $\mathbf{A}$, escreva os possíveis polinômios minimais de $\mathbf{A}$.

6. Seja $T \in \mathcal{L}(V)$. Mostre que se λ e μ são autovalores distintos de T, então $V_\lambda \cap V_\mu = \{\mathbf{0}\}$.

7. Seja $T \in \mathcal{L}(V)$, com $f(x) = (x - 4)^2(x + 2)^4$ o polinômio característico.

   (a) Quais são as possibilidades para $\dim V_4$ e $\dim V_{-2}$?

   (b) Se T for diagonalizável, qual a dimensão de V_4 e a de V_{-2}?

8. Seja $T \in \mathcal{L}(V)$, com $\dim V = 6$ e autovalores distintos λ e μ. Se $\dim V_\lambda = 3$ e $\dim V_\mu = 1$.

   (a) Quais são as possibilidades para o polinômio característico de T?

   (b) O polinômio minimal de T pode ser $m(x) = (x - \lambda)(x - \mu)$?

9. Seja $T \in \mathcal{L}(V)$ não diagonalizável, com $f(x) = (x + 1)(x - 3)^3$ o polinômio característico. Quais são as possibilidades para $\dim V_{-1}$ e $\dim V_3$?

10. Seja $T \in \mathcal{L}(V)$ não singular, com $\dim V = n$. Mostre que T^{-1} é um polinômio em T de grau no máximo $n - 1$.

11. Sejam $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ uma base V e $E_{ij} \in \mathcal{L}(V)$ definido como $E_{ij}(\mathbf{u}_k) = \delta_{ik}\mathbf{u}_j$, determine o conjunto espectral $\sigma(E_{ij})$.

12. Seja $T \in \mathcal{L}(F^3)$ definido como $T(x, y, z) = (\lambda x + ay, \lambda y + az, \lambda z)$. Determine o polinômio minimal de T quando $a = 0$ e $a \neq 0$. Generalize para F^n.

13. Seja $T \in \mathcal{L}(\mathbb{R}^2)$ o operador rotação por um ângulo θ. Mostre que se θ for um múltiplo inteiro de π, então o autovalor de T será $\lambda = 1$ ou $\lambda = -1$.

14. Determine a projeção E de $\mathbb{R}^2$ sobre $W_1 = \mathbb{R}[(1, -1)]$ na direção da reta $y = 2x$.

15. Sejam E_1 a projeção de V sobre W_1 paralela a W_2 e E_2 a projeção de V sobre U_1 paralela a U_2. Mostre que $E_1 + E_2$ é uma projeção se, e somente se, $E_1E_2 = E_2E_1 = 0$. Se $E_1 + E_2$ é uma projeção, determine $\ker(E_1 + E_2)$ e $\operatorname{Im}(E_1 + E_2)$.

16. Seja $W_1 = F[(1, 1, 0)]$ um subespaço de F^3. Determine um subespaço W_2 de F^3 tal que $F^3 = W_1 \oplus W_2$. Além disso, determine projeções $E_1, E_2 \in \mathcal{L}(F^3)$ tais que $\ker E_1 = W_2, \operatorname{Im} E_1 = W_1, \ker E_2 = W_1$ e $\operatorname{Im} E_2 = W_2$.

17. Sejam $E \in \mathcal{L}(V)$ uma projeção e $f(x) \in F[x]$. Mostre que $f(E) = aI + bE$.

18. Seja $E \in \mathcal{L}(V)$ uma projeção. Mostre que $I + E$ é não singular exibindo sua inversa $(I + E)^{-1}$.

19. Sejam $E_1, \ldots, E_k \in \mathcal{L}(V)$ projeções, com $\dim V = n$. Mostre que se $E_1 + \cdots + E_k = I$, então $E_iE_j = 0$, quando $i \neq j$, “divisores de zero”.

20. Seja $T \in \mathcal{L}(F^3)$ definido como

$$T(x, y, z) = (5x - 6y - 6z, -x + 4y + 2z, 3x - 6y - 4z).$$

Determine matrizes $\mathbf{E}_1$ e $\mathbf{E}_2$ tais que $\mathbf{A} = \lambda_1\mathbf{E}_1 + \lambda_2\mathbf{E}_2, \mathbf{E}_1 + \mathbf{E}_2 = \mathbf{I}$ e $\mathbf{E}_1\mathbf{E}_2 = \mathbf{O}$.

21. Seja $T \in \mathcal{L}(F^4)$ definido como $T(x, y, z, t) = (y + t, x + z, y + t, x + z)$. Determine matrizes $\mathbf{E}_1, \mathbf{E}_2$ e $\mathbf{E}_3$ tais que $\mathbf{A} = \lambda_1\mathbf{E}_1 + \lambda_2\mathbf{E}_2 + \lambda_3\mathbf{E}_3$, $\mathbf{E}_1 + \mathbf{E}_2 + \mathbf{E}_3 = \mathbf{I}$ e $\mathbf{E}_1\mathbf{E}_2 = \mathbf{E}_1\mathbf{E}_3 = \mathbf{E}_2\mathbf{E}_3 = \mathbf{O}$.

22. Seja $T \in \mathcal{L}(V)$ tal que $T^2 = I$.

    (a) Mostre que $V = W_1 \oplus W_2$, em que

    $$W_1 = \{\mathbf{u} \in V : T(\mathbf{u}) = \mathbf{u}\} \text{ e } W_2 = \{\mathbf{u} \in V : T(\mathbf{u}) = -\mathbf{u}\}$$

    (b) Mostre que se $\dim V = n$, então existe um subespaço H de V, com $\dim H = n - 1$, e um vetor não nulo $\mathbf{v} \in V - H$ tal que $T(\mathbf{u}) = \mathbf{u}$, para todo $\mathbf{u} \in H$, e $T(\mathbf{v}) = -\mathbf{v}$.

    (c) Determine W_1 e W_2, para $T \in \mathcal{L}(F^{n\times n})$ definido como $T(\mathbf{A}) = \mathbf{A}^t$.

23. Seja $T \in \mathcal{L}(V)$. Mostre que as seguintes condições são equivalentes:

    (a) $T^2 = I$;

(b) Se $E_1 = \frac{1}{2}(I - T)$ e $E_2 = \frac{1}{2}(I + T)$, então $E_1^2 = E_1$, $E_2^2 = E_2$ e $E_1 + E_2 = I$;

(c) $\ker(T + I) = \operatorname{Im}(T - I)$;

(d) $\ker(T - I) = \operatorname{Im}(T + I)$;

(e) T é uma reflexão.

24. Seja $T \in \mathcal{L}(V)$, com $\dim V = n$, tal que $T^k = 0$, para algum $k \in \mathbb{N}$. Mostre que o polinômio característico de T é x^n.

25. Seja $\mathbf{A} \in F^{n\times n}$. Mostre que $\mathbf{A}$ e $\mathbf{A}^t$ possuem o mesmo polinômio minimal.

26. Sejam $\mathbf{A}, \mathbf{B} \in F^{n\times n}$. Mostre que o polinômio minimal de $\mathbf{C} = \mathbf{A} \oplus \mathbf{B}$ é o mínimo múltiplo comum dos polinômios minimais de $\mathbf{A}$ e $\mathbf{B}$. Generalize.

27. Sejam $\mathbf{B} \in V = F^{n\times n}$ fixada e $T \in \mathcal{L}(V)$ definido como $T(\mathbf{X}) = \mathbf{BX}$. Mostre que o polinômio minimal de T é o polinômio minimal de $\mathbf{B}$.

28. Sejam $\mathbf{B} \in V = F^{n\times n}$ fixada, com $\mathbf{B} \neq \mathbf{O}$, e $T \in \mathcal{L}(V)$ definido como $T(\mathbf{X}) = \mathbf{BX} - \mathbf{XB}$.

    (a) Mostre que que 0 é um autovalor de T.

    (b) Mostre que se $\mathbf{B}^k = \mathbf{O}$, então $T^{2k} = 0$.

    (c) $T(\mathbf{XY}) = \mathbf{X}T(\mathbf{Y}) + T(\mathbf{X})\mathbf{Y}$.

    (d) Mostre que se $\mathbf{B}$ for diagonalizável, então T também o é.

29. Sejam $\mathbf{A}, \mathbf{B} \in F^{n\times n}$. As matrizes $\mathbf{AB}$ e $\mathbf{BA}$ possuem o mesmo polinômio minimal?

30. Sejam $\mathbf{A} \in \mathbb{R}^{n\times n}$. Mostre que o polinômio minimal de $\mathbf{A}$ em $\mathbb{R}$ é o polinômio minimal de $\mathbf{A}$ em $\mathbb{C}$.

31. Seja $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{3\times 3}$ um quadrado mágico, com $a_{11} = a + b, a_{22} = a$ e $a_{31} = a + c$. Determine condições necessárias e suficientes sobre a, b e c, para que $\mathbf{A}$ seja diagonalizável.

32. Sejam $\mathbf{T}_n \in F^{n\times n}$ a matriz tridiagonal, com $a_1 = 1, b_1 = -4$ e $c_1 = 5$, e $T_n = \det \mathbf{T}_n$. Determine a solução da sequência $(T_n)_{n\in\mathbb{N}}$.

33. (**Teorema de Sylvester**) Seja $\mathbf{A} \in \mathbb{C}^{n\times n}$, com $f(x) = (x-\lambda_1)\cdots(x-\lambda_n)$ o polinômio característico. Mostre que $g(x) = (x-p(\lambda_1))\cdots(x-p(\lambda_n))$ é o polinômio característico de $p(\mathbf{A})$, para todo $p(x) \in F[x]$. Conclua que qualquer autovalor de $p(\mathbf{A})$ é igual a $p(\lambda_j)$, para algum $j = 1, \ldots, n$.

34. Sejam $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$, com $g(x)$ o polinômio característico de $\mathbf{B}$. Mostre que $\mathbf{C} = g(\mathbf{A})$ é não singular se, e somente se, $\mathbf{A}$ e $\mathbf{B}$ não possuem autovalores comuns.

35. Seja $\mathbf{A} \in F^{n\times n}$.

    (a) Mostre que se $\lambda \in F$, com $\lambda \neq 0$, for um autovalor de $\mathbf{A}$, então $\lambda^{-1}\det\mathbf{A}$ é um autovalor de $\operatorname{adj}(\mathbf{A})$.

    (b) Mostre que se $\mathbf{X} \in F^{n\times 1}$ for um autovetor de $\mathbf{A}$, então $\mathbf{X}$ é um autovetor de $\operatorname{adj}(\mathbf{A})$.

36. Seja $\mathbf{A} \in F^{n\times n}$, com autovalores $\lambda_1, \ldots, \lambda_n \in F$ não nulos. Mostre que os autovalores de $\operatorname{adj}(\mathbf{A})$ são: $\mu_j = \prod_{i\neq j}\lambda_i$, com $j = 1, \ldots, n$.

37. Seja $\mathbf{A} \in F^{n\times n}$, com autovalores $\lambda_1, \ldots, \lambda_n \in F$, não necessariamente distintos. Mostre que os autovalores do menor $f_{ij}(x)$ de ordem $(n-1)$ de $f_{\mathbf{A}}(x) = \det(x\mathbf{I} - \mathbf{A})$ são: $\mu_j(x-\lambda_i) = f_{\mathbf{A}}(x)$, com $j = 1, \ldots, n$.

38. Sejam $\mathbf{A} \in F^{n\times n}$ e $\lambda \notin \sigma(\mathbf{A})$. A matriz $\mathbf{R}_\lambda = (\lambda\mathbf{I} - \mathbf{A})^{-1}$ chama-se *resolvente* de $\mathbf{A}$ e $f(\lambda)\mathbf{R}_\lambda = \operatorname{adj}(\lambda\mathbf{I} - \mathbf{A})$.

    (a) Mostre que $\mathbf{R}_\lambda - \mathbf{R}_\mu = (\mu - \lambda)\mathbf{R}_\lambda\mathbf{R}_\mu$.

    (b) Mostre que se $\mathbf{A}$ for diagonalizável e $\mathbf{G}_i = \mathbf{X}_i\mathbf{Y}_i^t$, com $i = 1, \ldots, n$, confira equação 5.8, então $\mathbf{R}_\lambda = \sum_{i=1}^n (\lambda - \lambda_i)^{-1}\mathbf{G}_i$.

    (c) Mostre que se $\lambda_1, \ldots, \lambda_n \in F$ são os autovalores de $\mathbf{A}$, então $\operatorname{tr}(\mathbf{R}_\lambda) = \sum_{i=1}^n (\lambda - \lambda_i)^{-1} = \frac{f'(\lambda)}{f(\lambda)}$.

39. Seja $\mathbf{R}_n = (a_{ij}) \in F^{n\times n}$, com $a_{ij} = 1$, se $j = n - (i-1)$, e $a_{ij} = 0$, caso contrário. Mostre que $\mathbf{R}_n$ é diagonalizável. A matriz $\mathbf{R}_n$ chama-se *identidade reversa*.

40. Seja $\mathbf{C}_n \in F^{n\times n}$ uma matriz companheira. Mostre que se $\mathbf{C}_n$ possuir m autovalores distintos em F, então ela possui m autovetores LI.

41. Sejam $\mathbf{C}_n \in F^{n\times n}$ uma matriz companheira e $\mathbf{C}_n^s$ sua antitransposta. Mostre que $\mathbf{C}_n^s$ é semelhante a $\mathbf{C}_n$.

42. Seja $T \in \mathcal{L}(\mathbb{C}^n)$ definido como $T(x_1, x_2, \dots, x_n) = (x_2, x_3, \dots, x_n, x_1)$. Mostre que T é diagonalizável. O que podemos afirmar sobre o *operador diferença* $D = T - I$?

43. Seja $\mathbf{C} = (c_{ij}) \in \mathbb{C}^{n\times n}$, com $c_{ij} = c_{i-j}$ e $i \equiv j \pmod{n}$. Mostre que $\mathbf{C}$ é diagonalizável. A matriz $\mathbf{C}$ chama-se *circulante*.

44. Seja $\mathbf{A} = a\mathbf{E}_{11} + b\mathbf{E}_{12} + c\mathbf{E}_{21} + d\mathbf{E}_{22} \in \mathbb{C}^{2\times 2}$, com $\det \mathbf{A} = 1$.

    (a) Determine $\mathbf{A}^{-1}$.

    (b) Escreva $\mathbf{A}$ como um produto de $\mathbf{T}_{12}(z) = \mathbf{I} + z\mathbf{E}_{12}$ e $\mathbf{T}_{21}(z) = \mathbf{I} + z\mathbf{E}_{21}$, para algum $z \in \mathbb{C}$.

    (c) Mostre que se $|\operatorname{tr}(\mathbf{A})| > 2$, então $\mathbf{A}$ é semelhante a $\operatorname{diag}(\lambda, \lambda^{-1})$, onde $\lambda \notin \{-1, 0, 1\}$.

    (d) Mostre que se $|\operatorname{tr}(\mathbf{A})| < 2$, então $\mathbf{A}$ é semelhante a $\operatorname{diag}(\lambda, \lambda^{-1})$, onde $\lambda \notin \mathbb{R} \cup (i\mathbb{R})$.

    (e) Mostre que se $|\operatorname{tr}(\mathbf{A})| = 2$, então $\mathbf{A} \in \mathbb{R}^{2\times 2}$. Quais as possíveis matrizes que são semelhantes a $\mathbf{A}$?

    (f) Mostre que se $|\operatorname{tr}(\mathbf{A})| \neq 2$, então $\mathbf{A}$ é semelhante a $2^{-1}\operatorname{tr}(\mathbf{A})\mathbf{E}_{11} + z\mathbf{E}_{12} + z\mathbf{E}_{21} + 2^{-1}\operatorname{tr}(\mathbf{A})\mathbf{E}_{22}$, para algum $z \in \mathbb{C}$.

    (g) O item (f) permanece verdade quando $|\operatorname{tr}(\mathbf{A})| = 2$?

6

Forma Canônica de Jordan

Sejam V um espaço vetorial de dimensão finita sobre F e $T \in \mathcal{L}(V)$. Já vimos que a matriz $\mathbf{A} = [T]_\alpha^\alpha$ em relação a alguma base α de V era semelhante a uma matriz diagonal se, e somente se, V tinha uma base formada de autovetores de T. Nosso objetivo neste capítulo é o seguinte: se T não pode ser diagonalizável, então determinar uma base de V em relação à qual a matriz de T tenha uma forma tão próximo quanto possível da matriz diagonal.

6.1 Teorema da Decomposição Primária

Já observamos que $F[x]$ era algebricamente semelhante a $\mathbb{Z}$, de modo que os conceitos de números primos e compostos, mdc e mmc, etc podem ser definidos em $F[x]$ de modo natural. Dado $f(x) \in F[x]$, diremos $f(x)$ é *redutível* sobre F se existirem $g(x), h(x) \in F[x]$ tais que

$$f(x) = g(x)h(x), \text{ com } 1 \leq \partial(g), \partial(h) < \partial(f).$$

Caso contrário, diremos que ele é *irredutível* sobre F. Por exemplo, o polinômio $f(x) = 1 + x^2 \in \mathbb{R}[x]$ é irredutível sobre $\mathbb{R}$ (prove isto!).

Sejam $f_1(x), \ldots, f_k(x) \in F[x]$. Diremos que $f_1(x), \ldots, f_k(x)$ são *relativamente* *primos* se $\text{mdc}(f_1(x), \ldots, f_k(x)) = 1$ ou, equivalentemente, podemos determinar polinômios $g_i(x) \in F[x]$ tais que

$$g_1(x)f_1(x) + \cdots + g_k(x)f_k(x) = 1. \tag{6.1}$$

A equação (6.1) chama-se *Identidade de Bezout*.[1]

Vamos relembrar o seguinte: sejam V um espaço vetorial sobre F, $T \in \mathcal{L}(V)$ e W um subespaço de V. Diremos que W é *invariante* sob T se $T(W) \subseteq W$. Diremos que W é *reduzido* sob T se ele for invariante sob T e existir um subespaço U de V invariante sob T tal que $V = W \oplus U$. Por exemplo, $\{\mathbf{0}\}, V, \operatorname{Im} T$ e o autoespaço generalizado V^λ, para todo autovalor $\lambda \in F$ de T, são invariantes sob T, pois para qualquer $\mathbf{u} \in V$, temos que $T(\mathbf{u}) \in \operatorname{Im} T$ e, em particular, quando $\mathbf{u} \in \operatorname{Im} T$. Por outro lado, dado $\mathbf{u} \in V^\lambda$, existe um $m \in \mathbb{N}$ tal que $(T - \lambda I)^m(\mathbf{u}) = \mathbf{0}$. Assim,

$$(T - \lambda I)^m(T(\mathbf{u})) = T((T - \lambda I)^m(\mathbf{u})) = T(\mathbf{0}) = \mathbf{0},$$

de modo que $T(\mathbf{u}) \in V^\lambda$. Portanto, V^λ é invariante sob T. Em particular, o autoespaço V_λ é invariante sob T. É muito importante ressaltar o seguinte: se $\dim V = n$, então, pelo Exercício (34) da Seção 4.2, V possui um subespaço invariante maximal, um minimal e é soma direta.

Exemplo 6.1 *Sejam V um espaço vetorial sobre F e $T \in \mathcal{L}(V)$ tal que $TS = ST$, para todo $S \in \mathcal{L}(V)$. Então $\ker S$ e $\operatorname{Im} S$ são invariantes sob T.*

Solução. Dado $\mathbf{v} \in \ker S$, obtemos

$$S(T(\mathbf{v})) = ST(\mathbf{v}) = TS(\mathbf{v}) = T(S(\mathbf{v})) = T(\mathbf{0}) = \mathbf{0}.$$

Portanto, $T(\mathbf{v}) \in \ker S$ e $\ker S$ é invariante sob T. Por outro lado, se $\mathbf{v} \in \operatorname{Im} S$, então existe um $\mathbf{u} \in V$ tal que $\mathbf{v} = S(\mathbf{v})$. Assim,

$$T(\mathbf{v}) = T(S(\mathbf{u})) = TS(\mathbf{u}) = ST(\mathbf{u}) = S(T(\mathbf{u}).$$

Portanto, $T(\mathbf{v}) \in \operatorname{Im} S$ e $\operatorname{Im} S$ é invariante sob T. ■

Teorema 6.2 *Sejam $T \in \mathcal{L}(V)$ e $f(x) \in F[x]$, com fatoração $f(x) = \prod_{i=1}^{k} p_i(x)$, onde os $p_i(x) \in F[x]$ são relativamente primos.*

[1] Étienne Bézout, 1730-1783, matemático francês.

1. *$W_i = \ker p_i(T)$ é invariante sob T e $\ker f(T) = W_1 \oplus \cdots \oplus W_k$.*

2. *A projeção $E_i \in \mathcal{L}(V)$ associada ao item (1) é um polinômio em T. Além disso, se $S \in \mathcal{L}(V)$ for tal que $TS = ST$, então $E_iS = SE_i$.*

3. *Se $W \subseteq \ker f(T)$ for um subespaço invariante sob T, então*

$$W = (W \cap W_1) \oplus \cdots \oplus (W \cap W_k).$$

Prova. (1) Seja $f_i(x) = \prod_{j\neq i} p_j(x)$ ou $f(x) = p_i(x)f_i(x)$, para $i = 1, \ldots, k$. Então é fácil verificar que os $f_1(x), \ldots, f_k(x)$ são relativamente primos e que $f(x)$ divide $f_i(x)f_j(x)$, se $i \neq j$. Dado $\mathbf{u}_i \in W_i = \ker p_i(T)$ tal que $\sum_{i=1}^{k} \mathbf{u}_i = \mathbf{0}$. Então $f_i(T)(\mathbf{u}_j) = \mathbf{0}$, se $i \neq j$, pois $f_i(T)$ contém o fator $p_j(T)$. Por outro lado, $\mathrm{mdc}(f_i(x), p_i(x)) = 1$ implica que $g_i(x)f_i(x) + h_i(x)p_i(x) = 1$, para alguns $g_i(x), h_i(x) \in F[x]$. Assim,

$$\begin{aligned} \mathbf{u}_i &= I(\mathbf{u}_i) = (g_i(T)f_i(T) + h_i(T)p_i(T))(\mathbf{u}_i) = g_i(T)f_i(T)(\mathbf{u}_i) \\ &= g_i(T)f_i(T)(-\textstyle\sum_{j\neq i} \mathbf{u}_j) = -\textstyle\sum_{j\neq i}(g_i(T)f_i(T)(\mathbf{u}_j)) = \mathbf{0}. \end{aligned}$$

Portanto, os W_i são independentes. Por outro lado, existem $g_1(x), \ldots, g_k(x) \in F[x]$ tais que

$$g_1(x)f_1(x) + \cdots + g_k(x)f_k(x) = 1.$$

Se $\mathbf{u} \in \ker f(T)$, então $\mathbf{u}_i = g_i(T)f_i(T)(\mathbf{u}) \in W_i$, pois

$$p_i(T)(\mathbf{u}_i) = p_i(T)g_i(T)f_i(T)(\mathbf{u}) = g_i(T)f(T)(\mathbf{u}) = \mathbf{0}.$$

Logo, $\mathbf{u} \in \sum_{i=1}^{k} W_i$, isto é, $\ker f(T) \subseteq W_1 \oplus \cdots \oplus W_k$. Como $f(T)$ contém cada fator $p_i(T)$ temos que $W_1 \oplus \cdots \oplus W_k \subseteq \ker f(T)$. (2) Ponha $E_i = g_i(T)f_i(T)$ e $W_i = \mathrm{Im}\, E_i, i = 1, \ldots, k$. (3) Seja $\mathbf{u} \in W \subseteq \ker f(T)$. Então $\mathbf{u} = \sum_{i=1}^{k} \mathbf{u}_i$, onde $\mathbf{u}_i \in W_i$, e, pelo Exemplo 6.1, $\mathbf{u}_i = E_i(\mathbf{u}) = g_i(T)f_i(T)(\mathbf{u}) \in W$. Portanto,

$$W = (W \cap W_1) \oplus \cdots \oplus (W \cap W_k).$$

Note que qualquer complemento direto U_i' de $U_i = W \cap W_i$ em W_i é invariante sob T. Pondo $W' = \sum_{i=1}^{k} U_i'$, obtemos $\ker f(T) = W \oplus W'$. ■

Seja $T \in \mathcal{L}(V)$. Diremos que T é *cíclico* se existir um $\mathbf{v} \in V$, com $\mathbf{v} \neq \mathbf{0}$, tal que $V = F[T^m(\mathbf{v}) : m \in \mathbb{Z}_+]$. Neste caso, $\mathbf{v}$ chama-se um *vetor cíclico* de T.

Exemplo 6.3 *Seja $T \in \mathcal{L}(F^2)$. Mostre que T é cíclico se, e somente se, $T \neq cI$, para algum $c \in F$.*

Solução. Suponhamos que T seja cíclico. Então existe um $\mathbf{v} \in F^2$, com $\mathbf{v} \neq \mathbf{0}$, tal que $T(\mathbf{v}) \notin F[\mathbf{v}]$. Assim, $\alpha = \{T(\mathbf{v}), \mathbf{v}\}$ é uma base de F^2. Portanto, $T \neq cI$, para algum $c \in F$. Reciprocamente, se T não fosse cíclico, então $T(\mathbf{v}) \in F[\mathbf{v}]$, para todo $\mathbf{v} \in F^2$, com $\mathbf{v} \neq \mathbf{0}$, e, pelo Exercício (14) da Seção 5.1, o resultado segue. É muito importante observar o seguinte: como $T^2(\mathbf{v}) \in F^2$ temos que existem únicos $c_0, c_1 \in F$ tais que $T^2(\mathbf{v}) = c_1T(\mathbf{v}) + c_0\mathbf{v}$, de modo que $(T^2 - c_1T - c_0I)(\mathbf{v}) = \mathbf{0}$ e $(T^2 - c_1T - c_0I)(T(\mathbf{v})) = \mathbf{0}$. Portanto, $f(x) = x^2 - c_1x - c_0$ é o polinômio característico e minimal de T e $\mathbf{C}_2 = [T]_\alpha^\alpha$ é a matriz companheira. ■

Teorema 6.4 (Teorema da Decomposição Primária) *Sejam $T \in \mathcal{L}(V)$ e o polinômio minimal de T $m(x) = \prod_{i=1}^k p_i^{r_i}(x)$, onde os $p_i(x) \in F[x]$ são distintos, irredutíveis e mônicos, com $\dim V = n$.*

1. *Os $W_i = \ker p_i^{r_i}(T)$ são invariantes sob T e $V = W_1 \oplus \cdots \oplus W_k$.*
2. *Se $T_i = T|_{W_i} \in \mathcal{L}(W_i)$, então o polinômio minimal $m_i(x)$ de T_i é igual a $p_i^{r_i}(x)$. Além disso, $T = T_1 \oplus \cdots \oplus T_k$.*
3. *Se $\mathbf{A}_i = [W_i]_{\alpha_i}^{\alpha_i}$, para alguma base α_i de W_i, então $\mathbf{A} = [T] = \mathbf{A}_1 \oplus \cdots \oplus \mathbf{A}_k$.*

Prova. (1) Como $m(T) = 0$ temos que $\ker m(T) = V$. Assim, pelo item (1) do Teorema 6.2, os W_i são invariantes sob T e $V = W_1 \oplus \cdots \oplus W_k$.

(2) Seja $m_i(x)$ o polinômio minimal de T_i. Então $m_i(x)$ é um fator de $p_i^{r_i}(x)$, pois $p_i^{r_i}(T)(\mathbf{u}) = \mathbf{0}$, para todo $\mathbf{u} \in W_i$. Por outro lado, como $m_i(T_i) = 0$ temos que $(m_i(x)f_i(x))(T) = m_i(T)f_i(T) = 0$. Assim, $m(x)$ é um fator de $m_i(x)f_i(x)$, isto é, $p_i^{r_i}(x)f_i(x)$ é um fator de $m_i(x)f_i(x)$. Logo, $p_i^{r_i}(x)$ é um fator de $m_i(x)$. Portanto, $m_i(x) = p_i^{r_i}(x)$, pois ambos são mônicos. ■

Lema 6.5 *Sejam $S, T \in \mathcal{L}(V)$, com $\dim V = n$, diagonalizáveis e $ST = TS$. Então S e T são simultaneamente diagonalizáveis, ou seja, existe uma base β de V tal que $[S]_\beta^\beta$ e $[T]_\beta^\beta$ são diagonais.*

Prova. Como S e T são diagonalizáveis temos, pelo Teorema 5.19, que

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k} \text{ e } V = V_{\mu_1} \oplus \cdots \oplus V_{\mu_m}.$$

Para cada μ_j fixado, dado $\mathbf{u} \in V_{\mu_j}$, existe um $\mathbf{v}_i \in V_{\lambda_i}$, com $i = 1, \ldots, k$, tal que $\mathbf{u} = \sum_{i=1}^{k} \mathbf{v}_i$. Assim, $\sum_{i=1}^{k} T(\mathbf{v}_i) = T(\mathbf{u}) = \mu_j \mathbf{u} = \sum_{i=1}^{k} \mu_j \mathbf{v}_i$. Por outro lado, os V_{λ_i} são invariantes sob T e a soma direta implicam que $T(\mathbf{v}_i) = \mu_j \mathbf{v}_i$, com $i = 1, \ldots, k$, de modo que $\mathbf{v}_i \in V_{\mu_j}$. Logo, pelo item (3) do Teorema 6.2,

$$V_{\mu_j} = (V_{\mu_j} \cap V_{\lambda_1}) \oplus \cdots \oplus (V_{\mu_j} \cap V_{\lambda_k}).$$

Portanto, $V = \sum_{j=1}^{m} \sum_{i=1}^{k} (V_{\mu_j} \cap V_{\lambda_i})$, de modo que escolhendo uma base para cada $V_{\mu_j} \cap V_{\lambda_i}$, obtemos uma base de V formada de autovetores de ambos S e T. ■

Exemplo 6.6 *Sejam $S, T \in \mathcal{L}(F^2)$ tais que $\mathbf{A} = \mathbf{E}_{11} + 2\mathbf{E}_{12} + 2\mathbf{E}_{22}$ e $\mathbf{B} = 3\mathbf{E}_{11} - 8\mathbf{E}_{12} - \mathbf{E}_{22}$. Mostre que existe uma matriz não singular $\mathbf{P}$ tal que $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ e $\mathbf{P}^{-1}\mathbf{B}\mathbf{P}$ sejam ambas diagonalizáveis.*

Solução. É fácil verificar que S e T são diagonalizáveis e $ST = TS$, com

$$V_1 = F[(1,0)], \;\; V_2 = F[(2,1)] \;\text{ e }\; V_{-1} = F[(2,1)], \;\; V_3 = F[(1,0)].$$

Assim, $F^2 = F[(1,0)] \oplus F[(2,1)]$. Portanto, pondo $\mathbf{P} = \mathbf{E}_{11} + 2\mathbf{E}_{12} + \mathbf{E}_{22}$, obtemos $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \operatorname{diag}(1,2)$ e $\mathbf{P}^{-1}\mathbf{B}\mathbf{P} = \operatorname{diag}(-1,3)$. ■

Seja $T \in \mathcal{L}(V)$. Diremos que T é *nilpotente* se existir um $r \in \mathbb{N}$ tal que $T^r = 0$. O menor $k \in \mathbb{N}$ tal que $T^k = 0$ chama-se *índice de nilpotência* de T.

Lema 6.7 *Sejam $S, T \in \mathcal{L}(V)$, com $\dim V = n$, tais que $ST = TS$.*

1. *Se S e T forem diagonalizáveis, então $S + T$ também o é.*
2. *Se S e T forem nilpotentes, então $S + T$ também o é.*

Prova. (1) Segue do Lema 6.5. (2) Suponhamos que $S^k = 0$ e $T^m = 0$. Então, pondo $r = k + m - 1$, obtemos

$$(S+T)^r = \sum_{j=0}^{m} \binom{r}{j} S^{r-j} T^j + \sum_{j=m+1}^{r} \binom{r}{j} S^{r-j} T^j = 0 + 0 = 0,$$

pois $r - j = k + (m - 1 - j) \geq k$. Portanto, $S + T$ é nilpotente. ■

Teorema 6.8 *Sejam $T \in \mathcal{L}(V)$, com $\dim V = n$ e $m(x) = \prod_{i=1}^{k} (x - \lambda_i)^{r_i}$ o polinômio minimal de T, onde os $\lambda_i \in F$ são distintos.*

1. *Existem $D, N \in \mathcal{L}(V)$, com D diagonalizável e N nilpotente, tais que (a) $T = D + N$. (b) $DN = ND$. Além disso, D e N são determinados de modo único por (a) e (b), e são polinômios em T.*

2. *$m_a(\lambda_i) = r_i$ e $V^{\lambda_i} = \ker(T - \lambda_i I)^{r_i}$, para cada $i = 1, \ldots, k$.*

Prova. (1) Pelo item (2) do Teorema 6.2, $\operatorname{Im} E_i = \ker(T - \lambda_i I)^{r_i}$, com $i = 1, \ldots, k$. Pondo $D = \sum_{i=1}^{k} \lambda_i E_i$ e $\sum_{i=1}^{k} TE_i = T$ temos, pelo item (5) do Teorema 5.39, que D é diagonalizável. Seja $N = T - D$. Então N é nilpotente, pois, indutivamente,

$$N^r = \sum_{i=1}^{k} (T - \lambda_i I)^r E_i, \ \ \forall \ r \in \mathbb{N}.$$

Assim, existe um $r \geq \max\{r_1, \ldots, r_r\}$ tal que $N^r = 0$. Portanto, $T = D + N$ e $DN = ND$. Suponhamos que $T = D_1 + N_1$ seja outra decomposição. Então é fácil verificar que $TD_1 = D_1T$ e $TN_1 = N_1T$. Logo, $f(T)D_1 = D_1 f(T)$ e $f(T)N_1 = N_1 f(T)$, para todo $f(x) \in F[x]$. Em particular, $DD_1 = D_1D$, $ND_1 = D_1N, DN_1 = N_1D$ e $NN_1 = N_1N$. Pelo Lema 6.7, temos que $D - D_1$ é diagonalizável e $N_1 - N$ é nilpotente. Como $D - D_1 = N_1 - N$ temos que $D - D_1$ é nilpotente. Logo, pelo Exercício (24) da Seção 5.3, o polinômio minimal de $D - D_1$ é $m(x) = x^k$, com $k \leq \dim V$. Mas, sendo $D - D_1$ diagonalizável, devemos ter $k = 1$, isto é, $m(x) = x$. Portanto, $D - D_1 = 0$, ou seja, $D = D_1$ e $N = N_1$.

(2) É claro que $\ker(T - \lambda_i I)^{r_i} \subseteq V^{\lambda_i}$. Suponhamos, por absurdo, que exista um $\mathbf{u} \in V^{\lambda_i}$ tal que $\mathbf{v} = (T - \lambda_i I)^{r_i}(\mathbf{u}) \neq \mathbf{0}$. Então existe um $s \in \mathbb{N}$, com $s > r_i$, tal que $(T - \lambda_i I)^s(\mathbf{u}) = \mathbf{0}$. Escrevendo $m(x) = q(x)(x - \lambda_i)^{r_i}$, temos que $q(x)$ e $(x - \lambda_i)^{s - r_i}$ são relativamente primos, de modo que existem $g(x), h(x) \in F[x]$ tais que $g(x)q(x) + h(x)(x - \lambda_i)^{s - r_i} = 1$. Assim,

$$\mathbf{v} = g(T)m(T)(\mathbf{u}) + h(T)(T - \lambda_i I)^s(\mathbf{u}) = \mathbf{0} + \mathbf{0} = \mathbf{0},$$

o que é uma contradição. Portanto, $V^{\lambda_i} = \ker(T - \lambda_i I)^{r_i}$, para $i = 1, \ldots, k$. ■

Exemplo 6.9 *Seja $T \in \mathcal{L}(F^3)$ definido como*

$$T(x, y, z) = (3x + y - z, 2x + 2y - z, 2x + 2y).$$

Mostre que existem $D, N \in \mathcal{L}(F^3)$, com D diagonalizável e N nilpotente, tais que $T = D + N$ e $DN = ND$. Determine as matrizes de D e N em relação à base canônica de F^3.

Solução. É fácil verificar que

$$\mathbf{A} = \begin{pmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{pmatrix}$$

e $f(x) = (x-1)(x-2)^2$ o polinômio característico e minimal de T. Então $p_1(x) = x-1$ e $p_2(x) = x - 2$ são distintos, irredutíveis e mônicos. Como

$$(T - I)(x, y, z) = (2x + y - z, 2x + y - z, 2x + 2y - z)$$

temos que $\alpha_1 = \{(1, 0, 2)\}$ é uma base de $W_1 = \ker p_1(T)$. Sendo

$$(T - 2I)^2(x, y, z) = (x - y, 0, 2x - 2y)$$

temos que $\alpha_2 = \{(1, 1, 0), (0, 0, 1)\}$ é uma base de $W_2 = \ker p_2(T)^2$. Assim, $F^3 = W_1 \oplus W_2$ e

$$\mathbf{P}^{-1}\mathbf{AP} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & -1 \\ 0 & 4 & 0 \end{pmatrix}, \text{ com } \mathbf{P} = [P]_\beta^\alpha = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

a matriz de transição da base canônica α para $\beta = \{(1, 0, 2), (1, 1, 0), (0, 0, 1)\}$. Pondo $f_1(x) = x^2 - 4x + 4$ e $f_2(x) = x - 1$. Então $f_1(x)$ e $f_2(x)$ são relativamente primos. Logo, aplicando sucessivamente o AD, existem polinômios $g_1(x) = 1$ e $g_2(x) = -x+3$ tais que $g_1(x)f_1(x) + g_2(x)f_2(x) = 1$. Fazendo

$$E_1 = g_1(T)f_1(T) = T^2 - 4T + 4I \text{ e } E_2 = g_2(T)f_2(T) = -T^2 + 4T - 3I,$$

obtemos $W_1 = \operatorname{Im} E_1$ e $W_2 = \operatorname{Im} E_2$. Portanto, existem

$$D = E_1 + 2E_2 = -T^2 + 4T - 2I \text{ e } N = T - D = T^2 - 3T + 2I$$

tais que $T = D + N$ e $DN = ND$. É fácil verificar que

$$[D] = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ -2 & 2 & 2 \end{pmatrix} \text{ e } [N] = \begin{pmatrix} 2 & 0 & -1 \\ 2 & 0 & -1 \\ 4 & 0 & -2 \end{pmatrix}.$$

Observe que o autovetor $\mathbf{v}_1 = (1, 0, 2)$ é uma coluna de $\mathbf{E}_1$ e os autovetores generalizados $\mathbf{v}_2 = (1, 1, 2), \mathbf{v}_3 = (0, 0, 1)$ são as colunas LI de $\mathbf{E}_2$.

Exercícios

1. Sejam $T \in \mathcal{L}(V)$, U e W subespaços invariantes sob T. Mostre que $U \cap W$, $U + W$ e $\ker f(T)$ são invariantes sob T, para todo $f(x) \in F[x]$.

2. Sejam $S, T \in \mathcal{L}(F^2)$ definidos como $S(x, y) = (x + y, x + y)$ e $T(x, y) = (x + ay, ax + y)$. Mostre que existe uma matriz não singular $\mathbf{P}$ tal que $\mathbf{P}^{-1}\mathbf{AP}$ e $\mathbf{P}^{-1}\mathbf{BP}$ sejam ambas diagonalizáveis.

3. Seja $T \in \mathcal{L}(F^3)$ definido como
$$T(x, y, z) = (6x - 3y - 2z, 4x - y - 2z, 10x - 5y - 3z).$$
Escreva o polinômio minimal de T sob a forma $m(x) = p_1(x)p_2(x)$, em que $p_1(x)$ e $p_2(x)$ são distintos, irredutíveis e mônicos sobre F. Sejam $W_1 = \ker p_1(T)$ e $W_2 = \ker p_2(T)$. Determine bases α_1 e α_2 de W_1 e W_2. Se $T_i = T|_{W_i}$, determine a matriz de T_i em relação à base α_i.

4. Sejam $T \in \mathcal{L}(V)$, com $\dim V = n$, e $\mathbf{v} \in V$. Mostre que o *subespaço cíclico* $W = F[T^m(\mathbf{v}) : m \in \mathbb{Z}_+] = F[T^{n-1}(\mathbf{v}), \ldots, T(\mathbf{v}), \mathbf{v}]$. Conclua que W é o menor subespaço invariante sob T que contém $\mathbf{v}$.

5. Sejam $S, T \in \mathcal{L}(F^3)$ definidos como
$$S(x, y, z) = (x + y, y + z, x) \text{ e } T(x, y, z) = (y, x, z).$$
Mostre que S é cíclico, mas T não.

6. Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Mostre que T é cíclico se, e somente se, o polinômio minimal e característico de T forem idênticos.

7. Sejam $T \in \mathcal{L}(V)$ e W um subespaço de V. Mostre que W é invariante sob T se, e somente se, W for invariante sob $aT + bI$, para alguns $a, b \in F$, com $a \neq 0$.

8. Seja $T \in \mathcal{L}(F^2)$ definido como $T(x, y) = (x + y, y)$. Mostre que é impossível escrever $F^2 = U \oplus W$, com U e W subespaços invariantes sob T.

9. Sejam V um espaço vetorial de dimensão finita sobre $\mathbb{C}, T \in \mathcal{L}(V)$ e D a parte diagonal de T. Mostre que a parte diagonal de $g(T)$ é $D(T)$, para todo $g(x) \in \mathbb{C}[x]$.

10. Seja $T \in \mathcal{L}(V)$, com $\dim V = n$, tal que $\rho(T) = 1$. Mostre que T é diagonalizável ou nilpotente, não ambos.

11. Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Mostre que se T comuta com todo operador linear diagonalizável sobre V, então $T = cI$, para algum $c \in F$.

12. Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Mostre que T é diagonalizável se, e somente se, o polinômio minimal de T for um produto de fatores lineares distintos.

13. Seja $\mathbf{A} \in \mathbb{R}^{n\times n}$, com $\mathbf{A} \neq \mathbf{I}$ e $\mathbf{A}^3 = \mathbf{I}$. Determine se a matriz $\mathbf{A}$ é ou não diagonalizável.

6.2 Operadores Nilpotentes

Nesta seção faremos um estudo mais detalhado de operadores nilpotentes seguindo H. F. Trotter.

Lema 6.10 *Sejam $T \in \mathcal{L}(V)$, com $\dim V = n$ e $\mathbf{u} \in V$ tal que $T^k(\mathbf{u}) = \mathbf{0}$, mas $T^{k-1}(\mathbf{u}) \neq \mathbf{0}$.*

1. *O conjunto $\alpha = \{T^{k-1}(\mathbf{u}), \ldots, T(\mathbf{u}), \mathbf{u}\}$ é LI.*

2. *$W = F[\alpha]$ é invariante sob T.*

3. *$S = T|_W$ é nilpotente de índice k e cíclico.*

4. *Pondo $\mathbf{u}_i = T^{k-i}(\mathbf{u})$, para $i = 1, \ldots, k$, temos que $\beta = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ é uma base de W tal que $S(\mathbf{u}_1) = \mathbf{0}$ e $S(\mathbf{u}_i) = \mathbf{u}_{i-1}$, para $i = 1, \ldots, k$, de modo que $[S]_\beta^\beta = \mathbf{E}_{12} + \cdots + \mathbf{E}_{(k-1)k} = (e_{ij})$, com $e_{ij} = 1$, se $i = j - 1$ e $e_{ij} = 0$, se $i \neq j - 1$.*

Prova. Com objetivos didáticos apresentaremos a prova do item (1). É fácil verificar, indutivamente, que $T^{k+m}(\mathbf{u}) = \mathbf{0}$, para todo $m \in \mathbb{N}$. Suponhamos que

$$c_1\mathbf{u} + c_2T(\mathbf{u}) + \cdots + c_kT^{k-1}(\mathbf{u}) = \mathbf{0}.$$

Então, aplicando T^{k-1}, obtemos $c_1 T^{k-1}(\mathbf{u}) = \mathbf{0}$, de modo que $c_1 = 0$, pois $T^{k-1}(\mathbf{u}) \neq \mathbf{0}$. Continuando assim, temos $c_2 = \cdots = c_k = 0$. Portanto, α é um conjunto LI de V. ■

Lema 6.11 *Sejam $T \in \mathcal{L}(V)$ e $W_i = \ker T^i$, onde $i \in \mathbb{Z}_+$. Então:*

1. *$W_i \subseteq W_{i+1}$.*
2. *$T(W_{i+1}) \subseteq W_i$.*
3. *Se $\alpha_i, \alpha_{i+1} = \alpha_i \cup \{\mathbf{v}_1, \ldots, \mathbf{v}_l\}$ e $\alpha_{i+2} = \alpha_{i+1} \cup \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ são bases de W_i, W_{i+1} e W_{i+2}, então $\alpha = \alpha_i \cup \{T(\mathbf{w}_1), \ldots, T(\mathbf{w}_m)\}$ é LI em W_{i+1}.*

Prova. Vamos provar apenas o item (3). Seja $\alpha_i = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ e suponhamos, por absurdo, que α seja LD. Então existem $a_1, \ldots, a_k, b_1, \ldots, b_m \in F$, não todos nulos, tais que $\sum_{i=1}^{k} a_i \mathbf{u}_i + \sum_{j=1}^{m} b_j T(\mathbf{w}_j) = \mathbf{0}$. Assim,

$$b_1 T(\mathbf{w}_1) + \cdots + b_m T(\mathbf{w}_m) = -(a_1 \mathbf{u}_1 + \cdots + a_k \mathbf{u}_k) \in W_i,$$

de modo que $T^i(b_1 T(\mathbf{w}_1) + \cdots + b_m T(\mathbf{w}_m)) = \mathbf{0}$. Logo, $T^{i+1}(\sum_{j=1}^{m} b_j \mathbf{w}_j) = \mathbf{0}$, ou seja, $\sum_{j=1}^{m} b_j \mathbf{w}_j \in W_{i+1}$. Como $W_{i+1} = F[\alpha_{i+1}]$ temos que existem $c_i, d_j \in F$ tais que

$$c_1 \mathbf{u}_1 + \cdots + c_k \mathbf{u}_k + d_1 \mathbf{v}_1 + \cdots + d_l \mathbf{v}_l + (-b_1)\mathbf{w}_1 + \cdots + (-b_m)\mathbf{w}_m = \mathbf{0},$$

o que é uma contradição, pois α_{i+2} é LI. ■

Lema 6.12 *Seja $T \in \mathcal{L}(V)$, com $\dim V = n$, tal que $T^k = 0$, mas $T^{k-1} \neq 0$. Então T admite uma representação matricial em bloco $\mathbf{J}$ cujos elementos diagonais possuem a forma $\mathbf{N} = \mathbf{E}_{12} + \cdots + \mathbf{E}_{(m-1)m}$, com $m \leq k$. Além disso:*

1. *Existe pelo menos um bloco $\mathbf{N}$ de ordem k e todos os outros são de ordem menor do que ou igual a k.*
2. *O número de blocos $\mathbf{N}$ de cada ordem possível é determinado de modo único por T.*
3. *O número total de blocos $\mathbf{N}$ de todas as ordens é igual a $\nu(T) = \dim \ker T$.*

Prova. Sejam $W_i = \ker T^i$ e $n_i = \dim W_i$, para $i = 1, \ldots, k$. Então $V = W_k$, $W_{k-1} \subset V$ e $n_{k-1} < n_k = n$, pois $T^k = 0$, mas $T^{k-1} \neq 0$. Assim, pelo item (1) do Lema 6.11, temos que $\{\mathbf{0}\} = W_0 \subset W_1 \subset \cdots \subset W_{k-1} \subset W_k = V$. Logo, indutivamente, obtemos uma base $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ de V tal que

$$\alpha_i = \{\mathbf{u}_1, \ldots, \mathbf{u}_{n_i}\}$$

seja uma base de W_i, para $i = 1, \ldots, k$. A partir desta base vamos escolher uma nova base de V em relação à qual T tenha a forma desejada. Pondo

$$\mathbf{v}_{(i,k)} = \mathbf{u}_{n_{(k-1)}+i} \text{ e } \mathbf{v}_{(i,k-1)} = T(\mathbf{v}_{(i,k)}),\ \ i = 1, \ldots, n_k - n_{k-1},$$

temos, pelo item (3) do Lema 6.11, que

$$\beta_1 = \{\mathbf{u}_1, \ldots, \mathbf{u}_{n_{k-2}}, \mathbf{v}_{(1,k-1)}, \ldots, \mathbf{v}_{(n_k-n_{k-1},k-1)}\}$$

é LI em W_{k-1}. Assim, estendendo β_1, se necessário, a uma base de W_{k-1} acrescentando elementos $\mathbf{v}_{(n_k-n_{(k-1)}+j,k-1)}$, com $j = 1, \ldots, 2n_{k-1} - n_k - n_{k-2}$. Novamente, pondo

$$\mathbf{v}_{(i,k-2)} = T(\mathbf{v}_{(i,k-1)}),\ \ i = 1, \ldots, n_{k-1} - n_{k-2},$$

temos, pelo item (3) do Lema 6.11, que

$$\beta_2 = \{\mathbf{u}_1, \ldots, \mathbf{u}_{n_{k-3}}, \mathbf{v}_{(1,k-2)}, \ldots, \mathbf{v}_{(n_{k-1}-n_{k-2},k-2)}\}$$

é LI em W_{k-2}, que pode ser estendido a uma base de W_{k-2} acrescentando elemen-tos $\mathbf{v}_{(n_{k-1}-n_{(k-2)}+j,k-1)}$, com $j = 1, \ldots, 2n_{k-2} - n_{k-1} - n_{k-3}$. Continuando assim, obtemos uma nova base de V, confira Tabela 6.1.

Note que a última linha da Tabela 6.1 forma a base de W_1, as duas últimas linhas da Tabela 6.1 formam a base de W_2 e, assim por diante. Pela construção, temos que

$$T(\mathbf{v}_{(i,j)}) = \begin{cases} \mathbf{v}_{(i,j-1)} & \text{se } j > 1, \\ \mathbf{0} & \text{se } j = 1. \end{cases} \tag{6.2}$$

Assim, pelo item (4) do Lema 6.10, T terá a forma desejada se os $\mathbf{v}_{(i,j)}$ são ordenados de maneira lexicográfica, confira Tabela 6.1. Pela equação (6.2), obtemos

$$T^m(\mathbf{v}_{(i,j)}) = \mathbf{v}_{(i,j-m)},\ \ \forall\ 1 \leq m < j.$$

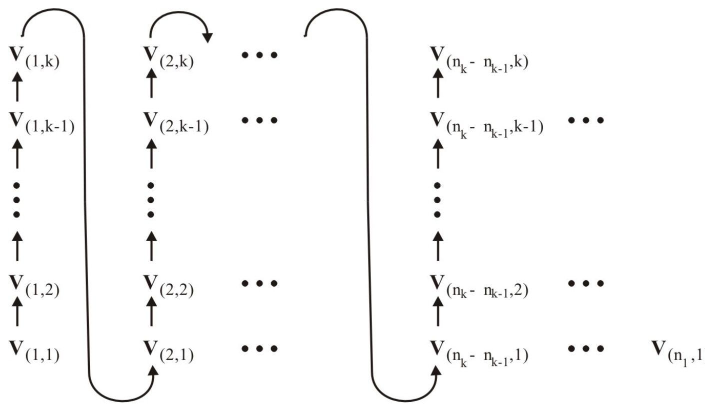


Figura 6.1: Base de Jordan.

Além disso, (1) haverá exatamente (blocos diagonais)

$$
\begin{array}{ll}
n_k - n_{k-1} & \text{de ordem } \; k \\
n_{k-1} - n_{k-2} - (n_k - n_{k-1}) = 2n_{k-1} - n_k - n_{k-2} & \text{de ordem } \; k-1 \\
\vdots & \vdots \\
2n_2 - n_3 - n_1 & \text{de ordem } \; 2 \\
2n_1 - n_2 & \text{de ordem } \; 1.
\end{array}
$$

(2) Como os números $n_1, \ldots, n_k$ são determinados de modo único por T temos que o número de elementos diagonais de cada ordem é determinado de modo único por T. (3) Como (soma telescópica)

$$n_1 = (n_k - n_{k-1}) + (2n_{k-1} - n_k - n_{k-2}) + \cdots + (2n_2 - n_3 - n_1) + (2n_1 - n_2)$$

temos que o número total de blocos diagonais é $n_1 = \dim W_1 = \dim \ker T$. ■

Teorema 6.13 *Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Então as seguintes condições são equivalentes*:

1. *T é nilpotente*;

2. *Existe uma base de V em relação à qual T admite uma representação matricial em bloco $\mathbf{J}$ cujos elementos diagonais possuem a forma $\mathbf{N} = \sum_{i=1}^{k-1} \mathbf{E}_{i(i+1)}$, com $k \leq n$.*

3. *Existe uma base de V em relação à qual T é representado por uma matriz triangular superior com zeros na diagonal*;

4. $T^n = 0$.

Prova. A implicação $(1 \Rightarrow 2)$ segue do Lema 6.12. É fácil verificar as implicações $(2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 1)$. ■

Exemplo 6.14 *Seja $T \in \mathcal{L}(F^4)$ definido como*

$$T(x, y, z, t) = (-x + y + t, 0, -x + y + t, -x + y + t)$$

Verifique se T é nilpotente. Caso afirmativo:

1. *Determine a matriz* $\mathbf{J}$ *em forma canônica que seja semelhante a* $\mathbf{A}$.

2. *Determine uma matriz não singular* $\mathbf{P}$ *tal que* $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{J}$.

Solução. (1) É fácil verificar que $T^2 = \mathbf{0}$. Portanto, T é nilpotente de índice 2. (a) Como o índice de nilpotência de T é igual a 2 temos que $\mathbf{J}$ contém pelo menos um bloco de ordem 2 e todos os outros de ordem menor do que ou igual a 2. (b) Pelo Teorema 4.39, basta escalonar a matriz

$$\left(\begin{array}{rrrr|rrrr} -1 & 0 & -1 & -1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \to \cdots \to \left(\begin{array}{rrrr|rrrr} 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right).$$

Portanto, $W_1 = F[(1,0,0,1),(0,1,0,-1),(0,0,1,0)]$ e $n_1 = \dim W_1 = 3$, isto é, o número total de blocos diagonais de $\mathbf{J}$ é igual a 3. (c) Como $k = 2, n_1 = 3$ e $n_2 = 4$ temos que $n_2 - n_1 = 1$ e $2n_1 - n_2 = 2$, isto é, $\mathbf{J}$ tem um bloco diagonal de ordem 2 e dois de ordem 1. Portanto, $\mathbf{J} = \mathbf{E}_{12}$.

(2) Pela forma de $\mathbf{J}$ basta escolher $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4 \in F^4$ tais que $T(\mathbf{u}_2) = \mathbf{u}_1$ e $T(\mathbf{u}_i) = \mathbf{0}$, para $i = 1, 3, 4$. Como $\operatorname{Im} T = F[(1,0,1,1)]$ e $\mathbf{u}_1 \in \operatorname{Im} T$ temos que $\mathbf{u}_1 = (1,0,1,1)$. Desde que $T^2 = 0$ temos que $\mathbf{u}_1 \in W_1 = \ker T$. Assim, podemos escolher $\mathbf{u}_2$ como qualquer solução da equação $T(\mathbf{u}_2) = \mathbf{u}_1$. Pelo Teorema 2.14, basta

escalonar a matriz

$$\left(\begin{array}{cccc|c} -1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 1 & 1 \\ -1 & 1 & 0 & 1 & 1 \end{array}\right) \to \cdots \to \left(\begin{array}{cccc|c} 1 & -1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right),$$

de modo que $\mathbf{u}_2 = (x, y, z, t)$ sujeito à condição $x = y + t - 1$, para todos $y, z, t \in F$, digamos $\mathbf{u}_2 = (0, 1, 0, 0)$. Portanto, pondo $\mathbf{u}_3 = (1, 0, 0, 1)$ e $\mathbf{u}_4 = (0, 1, 0, -1)$, obtemos

$$\mathbf{P}^{-1}\mathbf{AP} = \mathbf{J}, \text{ com } \mathbf{P} = [P]_\beta^\alpha = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & -1 \end{pmatrix}$$

a matriz de transição da base canônica α para $\beta = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ de F^4. ■

Exercícios

1. Seja $T \in \mathcal{L}(F^3)$ definido como $T(x, y, z) = (2y + z, 3z, 0)$. Determine a matriz nilpotente $\mathbf{J}$ em forma canônica que seja semelhante a $\mathbf{A}$. Além disso, determine uma matriz $\mathbf{P}$ tal que $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{J}$.

2. Seja $T \in \mathcal{L}(F^4)$ definido como $T(x, y, z, t) = (-z + t, z, 0, 0)$. Determine a matriz nilpotente $\mathbf{J}$ em forma canônica que seja semelhante a $\mathbf{A}$. Além disso, determine uma matriz $\mathbf{P}$ tal que $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{J}$.

3. Seja $T \in \mathcal{L}(F^5)$ definido como $T(x, y, z, t, u) = (y+u, t+u, t, u, 0)$. Determine a matriz nilpotente $\mathbf{J}$ em forma canônica que seja semelhante a $\mathbf{A}$. Além disso, determine uma matriz $\mathbf{P}$ tal que $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{J}$.

4. Seja $\mathbf{N} = \mathbf{E}_{12} + \mathbf{E}_{23} + \mathbf{E}_{34}$.

   (a) Mostre que $\mathbf{AN} = \mathbf{NA}$ se, e somente se, $\mathbf{A}$ for da forma

$$\mathbf{A} = \begin{pmatrix} a & b & c & d \\ 0 & a & b & c \\ 0 & 0 & a & b \\ 0 & 0 & 0 & a \end{pmatrix}.$$

(b) Mostre que se $b \neq 0$, então $\dim V_\lambda = 1$, para todo autovalor λ de $\mathbf{A}$.

(c) Mostre que se $b = 0$ e $c \neq 0$, então $\dim V_\lambda = 2$, para todo autovalor λ de $\mathbf{A}$.

(d) Mostre que se $b = c = 0$ e $d \neq 0$, então $\dim V_\lambda = 3$, para todo autovalor λ de $\mathbf{A}$.

(e) Generalize para qualquer matriz quadrada $\mathbf{N}$.

5. Seja $T \in \mathcal{L}(V)$ tal que $T^k = 0$, mas $T^{k-1} \neq 0$. Mostre que qualquer operador linear semelhante a T é nilpotente de índice k.

6. Seja $T \in \mathcal{L}(V)$, com $\dim V = n$, tal que $T^k = 0$, mas $T^{k-1} \neq 0$. Mostre que $\operatorname{Im} T^{k-i} \subseteq \ker T^i$, para $i = 1, \ldots, k-1$.

7. Seja $T \in \mathcal{L}(V)$, com $\dim V = n$ tal que $T^k = 0$, mas $T^{k-1} \neq 0$. Mostre que $I - T$ e $I + T$ são não singulares. Além disso, $\operatorname{tr}(T) = 0$ e T não é diagonalizável.

8. Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Mostre que T é nilpotente se, e somente se, todos os autovalores de T são nulos. Mostre, com um exemplo, que uma das implicações da afirmação é falsa se V for um espaço vetorial de dimensão infinita sobre $\mathbb{R}$.

9. Seja $T \in \mathcal{L}(V)$, com $\dim V = n$.

   (a) Mostre que se T for cíclico, então qualquer autovalor de T está associado a um único autovetor, a menos de multiplicação por escalar.

   (b) Mostre que se T possuir n autovalores distintos, então T é cíclico.

10. Sejam $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n \times n}$. Mostre que se $\mathbf{AB} + \mathbf{BA} = \mathbf{0}$ e $\mathbf{B}$ não for nilpotente, então a equação matricial $\mathbf{AX} + \mathbf{XA} = \mathbf{B}$ não possui solução.

6.3 Forma Canônica de Jordan

Nesta seção provaremos que qualquer$T \in \mathcal{L}(V)$, com $\dim V = n$, pode ser decomposto como soma de um operador diagonalizável com um operador nilpotente. Veremos que a forma de Jordan é muito útil na resolução de problemas teóricos e computacionais.

Lema 6.15 *Sejam $T \in \mathcal{L}(V)$, com $\dim V = n$, $f(x) = \prod_{i=1}^{k}(x - \lambda_i)^{d_i}$ e $m(x) = \prod_{i=1}^{k}(x - \lambda_i)^{r_i}$ os polinômios característico e minimal de T, onde os $\lambda_1, \ldots, \lambda_k \in F$*

são distintos e $1 \leq r_i \leq d_i$. Então T admite uma representação matricial em bloco $\mathbf{J}$ *cujos elementos diagonais possuem a forma* $\mathbf{J}_{ij} = \lambda_i \mathbf{I}_m + \sum_{p=1}^{m-1} \mathbf{E}_{p(p+1)}$, *com* $m \leq r_i$ *e* $i = 1, \ldots, k$. *Além disso, para cada* λ_i, $i = 1, \ldots, k$, *os blocos* $\mathbf{J}_{ij}$ *possuem as seguintes propriedades*:

1. *Existe pelo menos um bloco* $\mathbf{J}_{ij}$ *de ordem* r_i *e todos os outros são de ordem menor do que ou igual a* r_i.

2. *A soma dos blocos* $\mathbf{J}_{ij}$ *é igual a* $d_i = m_a(\lambda_i)$.

3. *O número dos blocos* $\mathbf{J}_{ij}$ *é igual a* $m_g(\lambda_i)$. *Neste caso,* T *possui* $\sum_{i=1}^{k} m_g(\lambda_i)$ *autovetores* LI.

4. *O número dos blocos* $\mathbf{J}_{ij}$ *de cada ordem possível é determinado de modo único por* T.

Prova. Como $m(x) = \prod_{i=1}^{k}(x - \lambda_i)^{r_i}$ temos, pelo Teorema 6.4, que

$$T = T_1 \oplus \cdots \oplus T_k \text{ e } V = W_1 \oplus \cdots \oplus W_k,$$

em que $W_i = \ker(T - \lambda_i I)^{r_i}$, $i = 1, \ldots, k$. Note que o polinômio minimal de T_i, $m_i(x) = (x - \lambda_i)^{r_i}$, implica que $(T_i - \lambda_i I)^{r_i} = 0$ é nilpotente de índice r_i, para $i = 1, \ldots, k$. Pondo $N_i = T_i - \lambda_i I$, obtemos $T_i = \lambda_i I + N_i$ e $N_i^{r_i} = 0$, para $i = 1, \ldots, k$, isto é, T_i é a soma de um operador diagonalizável $\lambda_i I$ e de um operador nilpotente N_i de índice r_i. Assim, pelo Lema 6.12, podemos escolher uma base de W_i em relação à qual N_i esteja na forma canônica. Nesta base $T_i = \lambda_i I + N_i$ é representado por uma matriz diagonal de bloco $\mathbf{J}_i$ cujos elementos diagonais são as matrizes $\mathbf{J}_{ij}$. Portanto,

$$\mathbf{J} = \mathbf{J}_1 \oplus \cdots \oplus \mathbf{J}_k$$

está na forma canônica e é a representação matricial de T. Além disso, (1) Como $N_i^{r_i} = 0$, para $i = 1, \ldots, k$, temos que existe, pelo menos, um $\mathbf{J}_{ij}$ de ordem r_i e todos os outros de ordem menor do que ou igual a r_i.

(2) Como T e $\mathbf{J}$ possuem o mesmo polinômio característico $f(x)$ temos que a soma das ordens dos $\mathbf{J}_{ij}$ é igual a $d_i = m_a(\lambda_i)$.

(3) Como $N_i = T_i - \lambda_i I$ e $m_g(\lambda_i) = \dim\ker(T_i - \lambda_i I)^{r_i} = \nu(N_i)$, temos que o número dos $\mathbf{J}_{ij}$ é igual a $m_g(\lambda_i)$. (4) Segue do item (2). do Lema 6.12. ■

A matriz $\mathbf{J}$ chama-se *forma "canônica" de Jordan* de T. Um bloco diagonal $\mathbf{J}_{ij}$ chama-se *bloco "elementar" de Jordan* associado ao autovalor λ_i. Note que **se** V **for**

um espaço vetorial de dimensão finita sobre $\mathbb{C}$, então qualquer $T \in \mathcal{L}(V)$ admite uma representação matricial na forma canônica de Jordan.

Teorema 6.16 *Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Então as seguintes condições são equivalentes*:

1. *O polinômio característico de T fatora-se, $f(x) = \prod_{i=1}^{k}(x - \lambda_i)^{d_i}$*;
2. *Existe uma base de V em relação à qual T admite uma representação matricial na forma de Jordan*;
3. *T é triangularizável, isto é, existe uma base de V em relação à qual T é representado por uma matriz triangular superior da forma $[T] = (a_{ij})$, com $a_{ij} = 0$, se $i > j$, e $a_{ii} = \lambda_i$.*
4. *Existem subespaços $W_0, W_1, \ldots, W_k$ de V invariantes sob T tais que*
$$\{\mathbf{0}\} = W_0 \subset W_1 \subset \cdots \subset W_{k-1} \subset W_k = V;$$
5. *O corpo F contém n autovalores de T (contando as multiplicidades)*;
6. *$V = V^{\lambda_1} \oplus \cdots \oplus V^{\lambda_k}$.*

Prova. A implicação $(1 \Rightarrow 2)$ segue do Lema 6.15. É fácil verificar as implicações $(2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 5 \Rightarrow 6 \Rightarrow 1)$. A implicação $(1 \Rightarrow 6)$ segue do item (2) do Teorema 6.8. ■

Já vimos, no Lema 6.16, que para obter a cadeia de Jordan e/ou a forma de Jordan de T ou de $\mathbf{A}$ deveríamos, recursivamente, resolver as equações:

$$\mathbf{AX}_i = \lambda_i \mathbf{X}_i \text{ ou } \mathbf{AX}_i = \lambda_i \mathbf{X}_i + \mathbf{X}_{i-1}, i = 2, \ldots, k. \tag{6.3}$$

Portanto, podemos usar o procedimento dado após o Teorema 5.7, para obter o autoespaço V_{λ_i} e/ou o autoespaço generalizado V^{λ_i} e, como consequentemente, a forma de Jordan.

Exemplo 6.17 *Seja $T \in \mathcal{L}(F^4)$ definido como*

$$T(x, y, z, t) = (y + t, y, -x + y + z + t, -x + y + 2t)$$

Determine a forma de Jordan de T.

Solução. É fácil verificar que

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

de modo que $f(x) = (x-1)^4$ e $m(x) = (x-1)^2$. Assim, pelo Teorema 6.4, $T = T_1$ e $V = \ker(T-I)^2$. Pondo $\mathbf{N} = \mathbf{A} - \mathbf{I}$, obtemos $\mathbf{N}^2 = \mathbf{0}$ e, pelo Exemplo 6.14, $\mathbf{M} = \mathbf{P}^{-1}\mathbf{N}\mathbf{P} = \mathbf{P}^{-1}(\mathbf{A}-\mathbf{I})\mathbf{P} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} - \mathbf{I}$, isto é,

$$\mathbf{J} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{I} + \mathbf{M} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

é a forma de Jordan de T. ■

Exemplo 6.18 *Seja $T \in \mathcal{L}(F^3)$ definido como $T(x,y,z) = (y, z, x - 3y + 3z)$. Determine a forma Jordan de T.*

Solução. Note que

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 3 \end{pmatrix}$$

é uma matriz companheira. Assim, $f(x) = m(x) = x^3 - 3x^2 + 3x - 1 = (x-1)^3$. Logo, $\rho(T - I) = 2$. Portanto, $\dim V_1 = \dim\ker(T-I) = 1$ e a forma de Jordan de T é $\mathbf{J} = \mathbf{I} + \mathbf{E}_{12} + \mathbf{E}_{23}$. ■

Exemplo 6.19 *Seja $T \in \mathcal{L}(V)$, com $f(x) = (x-3)^2(x-1)^3(x+5)$ o polinômio característico. Determine as possíveis formas de Jordan de T.*

Solução. É claro que $\dim V = 6$ e os candidatos a polinômio minimal de T foram dados no Exemplo 5.29. Portanto, pondo $\mathbf{D} = \operatorname{diag}(3,3,1,1,1,-5)$, as possíveis formas de Jordan de T são: $\mathbf{D}, \mathbf{D}+\mathbf{E}_{12}, \mathbf{D}+\mathbf{E}_{34}, \mathbf{D}+\mathbf{E}_{34}+\mathbf{E}_{45}$, $\mathbf{D}+\mathbf{E}_{12}+\mathbf{E}_{34}$ e $\mathbf{D}+\mathbf{E}_{12}+\mathbf{E}_{34}+\mathbf{E}_{45}$. ■

Exemplo 6.20 (Teorema de Cayley-Hamilton) *Seja $\mathbf{A} \in F^{n\times n}$. Mostre que se $f(x) = \prod_{i=1}^{k}(x-\lambda_i)^{d_i}$ for o polinômio característico de $\mathbf{A}$, então $f(\mathbf{A}) = \mathbf{0}$.*

Solução. Seja $\mathbf{J}$ a forma de Jordan de $\mathbf{A}$. Então $f(\mathbf{J}) = \prod_{i=1}^{n} \mathbf{J}_i$, com $\mathbf{J}_i = \mathbf{J} - \lambda_i \mathbf{I}$, $i = 1, \ldots, n$. Como a n-ésima linha de $\mathbf{J}_n$ é uma linha de zeros temos que: as duas últimas linhas de $\mathbf{J}_{n-1}\mathbf{J}_n$ são de zeros, as três últimas linhas de $\mathbf{J}_{n-2}\mathbf{J}_{n-1}\mathbf{J}_n$ são de zeros, e assim por diante. Portanto, $f(\mathbf{J}) = \mathbf{0}$. Como $f(\mathbf{A})$ e $f(\mathbf{J})$ são semelhantes temos que $f(\mathbf{A}) = \mathbf{0}$. ■

Exemplo 6.21 *Seja $T \in \mathcal{L}(F^3)$ definido como $T(x, y, z) = (6x + 2y + 2z, -2x + 2y, 2z)$. Determine a forma de Jordan de T.*

Solução. Vamos fazer uma prova direta. É fácil verificar que $f(x) = (x-2)(x-4)^2$ é o polinômio característico e minimal de T. Note que $\mathbf{v}_1 = (0, -1, 1)$ é o único autovetor de T associado a $\lambda_1 = 2$ e $\mathbf{v}_2 = (2, -2, 0)$ é o único autovetor de T associado a $\lambda_2 = 4$. Assim, resolvendo a equação $(T - 4I)(\mathbf{v}_3) = \mathbf{v}_2$, digamos $\mathbf{v}_3 = (1, 0, 0)$ é o autovetor generalizado de T associado a $\lambda_2 = 4$. Portanto, $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ é a base de Jordan de F^3 e $\mathbf{J} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ é a forma de Jordan de T, com $\mathbf{P} = [P]_\beta^\alpha$ e α a base canônica de F^3. ■

Exemplo 6.22 *Sejam $\mathbf{M}, \mathbf{N} \in F^{3\times 3}$ nilpotentes. Mostre que $\mathbf{M}$ e $\mathbf{N}$ são semelhantes se, e somente se, $\mathbf{M}$ e $\mathbf{N}$ possuem o mesmo polinômio minimal.*

Solução. Sejam $f(x), g(x)$ e $m(x), n(x)$ os polinômios característicos e minimais de $\mathbf{M}$ e $\mathbf{N}$. Então $m(x) = n(x)$ e $f(x) = g(x)$. Como $m(x)$ e $f(x)$ possuem as mesma raízes temos que a forma canônica de Jordan de $\mathbf{M}$ é uma das matrizes: $\mathbf{O}, \mathbf{E}_{12}$ e $\mathbf{E}_{12} + \mathbf{E}_{23}$. Portanto, em qualquer caso, $\mathbf{M}$ e $\mathbf{N}$ são semelhantes. A recíproca é clara. ■

Exemplo 6.23 *Sejam V um espaço vetorial de dimensão finita sobre $\mathbb{C}$ e $T \in \mathcal{L}(V)$. Determine todos os subespaços de dimensão 2 e invariantes sob T.*

Solução. Seja W qualquer subespaço invariante sob T, com $\dim W = 2$. Então, pelo Teorema 5.16, existe um $\mathbf{v}_1 \in W$, com $\mathbf{v}_1 \neq \mathbf{0}$, e $\lambda_1 \in \mathbb{C}$ tal que $T(\mathbf{v}_1) = \lambda_1 \mathbf{v}_1$. Assim, podemos escolher um $\mathbf{v}_2 \in W$ tal que $W = \mathbb{C}[\mathbf{v}_1, \mathbf{v}_2]$. Como W é invariante sob T temos que $T(\mathbf{v}_2) \in W$, de modo que existem $c_1, c_2 \in \mathbb{C}$ tais que

$$T(\mathbf{v}_2) = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2.$$

Se $c_1 = 0$, então $\mathbf{v}_2$ é o autovetor de T associado a $\lambda_2 = c_2$. Neste caso, a matriz $[T|_W]$ é semelhante a $\mathbf{D} = \operatorname{diag}(\lambda_1, \lambda_2)$. Se $c_1 \neq 0$ e $c_2 \neq \lambda_1$, então é fácil verificar que $\mathbf{v}_3 = -c_1\mathbf{v}_1 + (\lambda_1 - c_2)\mathbf{v}_2$ é o autovetor de T associado a $\lambda_2 = c_2$ e $W = \mathbb{C}[\mathbf{v}_1, \mathbf{v}_3]$. Neste caso, a matriz $[T|_W]$ é semelhante a $\mathbf{D} = \operatorname{diag}(\lambda_1, \lambda_2)$. Finalmente, se $c_1 \neq 0$ e

$c_2 = \lambda_1$, então $\mathbf{v}_1$ é o único autovetor em W de T associado a λ_1 (prove isto!). Logo, resolvendo a equação $(T - \lambda_1 I)(\mathbf{v}_4) = \mathbf{v}_1$ em W, temos que $\mathbf{v}_4 = c_1^{-1}\mathbf{v}_2$ é o autovetor generalizado em W de T associado a λ_1 e $W = \mathbb{C}[\mathbf{v}_1, \mathbf{v}_4]$. Portanto, a matriz $[T|_W]$ é semelhante a forma de Jordan $\lambda_1\mathbf{I} + \mathbf{E}_{12}$. ■

Exercícios

1. $T \in \mathcal{L}(F^3)$ definido como $T(x, y, z) = (2y + z, 3z, 0)$. Determine a forma de Jordan de T.

2. Determine se as matrizes
$$\mathbf{A} = \begin{pmatrix} 1 & 4 \\ -1 & -3 \end{pmatrix} \text{ e } \mathbf{B} = \begin{pmatrix} -1 & 1 \\ c & -1 \end{pmatrix}$$
são ou não semelhantes, onde $c \in F$.

3. Seja $T \in \mathcal{L}(V)$, com polinômio característico $f(x) = (x-2)^3(x+7)^2$ e minimal $m(x) = (x-2)^3(x+7)$. Determine a forma de Jordan de T.

4. Seja $T \in \mathcal{L}(V)$, com polinômio característico $f(x) = (x+2)^4(x-1)^2$. Determine as possíveis formas de Jordan de T.

5. Seja $T \in \mathcal{L}(\mathbb{R}^2)$ definido como $T(x, y, z) = (ax+by, cx+dy)$. Determine condições necessárias e suficientes sobre a, b, c e d, de modo que T não seja semelhante a uma matriz diagonal.

6. Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Mostre que se $\lambda_1, \ldots, \lambda_n$ forem os autovalores de T, então $\det T = \lambda_1 \cdots \lambda_n$.

7. Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Mostre que T é não singular se, e somente se, todos os autovalores de T são não nulos.

8. Sejam $\mathbf{A}, \mathbf{B} \in F^{n\times n}$. Mostre que se $\mathbf{A}$ e $\mathbf{B}$ possuem o mesmo polinômio característico e minimal $f(x) = \prod_{i=1}^{k}(x - \lambda_i)^{d_i}$, com $d_i \leq 3$ e $i = 1, \ldots, k$, então $\mathbf{A}$ e $\mathbf{B}$ são semelhantes.

9. Sejam $\mathbf{A}, \mathbf{B} \in F^{n\times n}$. Mostre que se $\mathbf{A}$ for semelhante a $\mathbf{B}$ em $\mathbb{C}^{n\times n}$, então $\mathbf{A}$ é semelhante a $\mathbf{B}$ em $F^{n\times n}$. Conclua que $\mathbf{A}$ é semelhante a $\mathbf{A}^t$.

10. Seja $\mathbf{J} \in \mathbb{C}^{n\times n}$ uma matriz de Jordan não singular. Mostre que $\mathbf{J}^{-1} = (a_{ij})$, com $a_{ij} = 0$, se $j \notin \{i, i+1\}$. Mostre, com um exemplo, que $\mathbf{J}^{-1}$ não necessita ser de Jordan.

11. Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Mostre que se $\operatorname{tr}(T^i) = 0$, para $i = 1, \ldots, n$, então T é nilpotente.

12. Sejam $\mathbf{A} \in \mathbb{C}^{n\times n}$ e $f(x) = x^n + \sum_{i=1}^{n} c_i x^{n-i}$ seu polinômio característico. Mostre que $c_1 = -\operatorname{tr}(\mathbf{A})$ e $kc_k = -(\sum_{i=1}^{k} c_{k-i} \operatorname{tr}(\mathbf{A}^i))$, com $k = 2, \ldots, n$ e $c_0 = 1$. Estas equações nos fornece um método alternativo de determinar $\mathbf{A}^{-1}$: (a) Calculando $\mathbf{A}, \ldots, \mathbf{A}^{n-1}$ e os elementos diagonais de $\mathbf{A}^n$. (b) Calculando os $\operatorname{tr}(\mathbf{A}^i)$, para determinar os c_i, com $i = 1, \ldots, n-1$. (c) Usar a fórmula

$$\mathbf{A}^{-1} = -c_n^{-1}(\mathbf{A}^{n-1} + c_1\mathbf{A}^{n-1} + \cdots + c_{n-1}\mathbf{I}).$$

O algoritmo de Csanky para obter $\mathbf{A}^{-1}$: se $\lambda_1, \ldots, \lambda_n \in F$ forem os autovalores de $\mathbf{A}$ e $s_k = \operatorname{tr}(\mathbf{A}^k) = \sum_{i=1}^{n} \lambda_i^k$ as somas de Newton[2], então os c_i e s_i satisfazem o sistema $\mathbf{SX} = \mathbf{B}$, em que

$$\mathbf{S} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ s_1 & 2 & 0 & \cdots & 0 \\ s_2 & s_1 & 3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{n-1} & s_{n-2} & \cdots & s_1 & n \end{pmatrix}, \mathbf{X} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_n \end{pmatrix} \text{ e } \mathbf{B} = \begin{pmatrix} -s_1 \\ -s_2 \\ -s_3 \\ \vdots \\ -s_n \end{pmatrix}.$$

13. Sejam V um espaço vetorial de dimensão finita sobre $F, \mathcal{E}(V)$ o conjunto de todas as projeções sobre V e $\mathcal{I}(V)$ o conjunto de todas as involuções sobre V. Mostre que a função $f : \mathcal{E}(V) \to \mathcal{I}(V)$ definida como $f(E) = 2E - I$ é bijetora.

14. Seja $T \in \mathcal{L}(V)$, com $\dim V = n$. Mostre que existem (únicos) subespaços U e W de V tais que as seguintes condições são satisfeitas:

    (a) $V = U \oplus W$.

    (b) U e W são invariantes sob T.

    (c) $T|_U$ é nilpotente.

    (d) $T|_W$ é não singular.

[2] Sir Isaac Newton, 1643-1727, físico e matemático inglês.

7

Espaços com Produto Interno

Neste capítulo, veremos o efeito de adicionar mais estrutura geométrica à estrutura algébica de um espaço vetorial real ou complexo, como definido no Capítulo 3, de modo que tenha sentido falar do "comprimento", "distância" e "ângulo".

7.1 Produto Interno

Sejam F o corpo dos reais $\mathbb{R}$ ou dos complexos $\mathbb{C}$ e V um espaço vetorial sobre F. Uma função de $V \times V$ em F, denotada por $\langle \cdot, \cdot \rangle$, é um *produto interno* sobre V se as seguintes condições são satisfeitas:

1. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$, para todos $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ (aditividade).
2. $\langle a\mathbf{u}, \mathbf{v} \rangle = a \langle \mathbf{u}, \mathbf{v} \rangle$, para todos $\mathbf{u}, \mathbf{v} \in V$ e $a \in F$ (homogeneidade).
3. $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$, para todos $\mathbf{u}, \mathbf{v} \in V$ (simétrica Hermitiana).
4. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$, para todo $\mathbf{u} \in V$, e $\langle \mathbf{u}, \mathbf{u} \rangle = 0 \Leftrightarrow \mathbf{u} = \mathbf{0}$ (positividade).

Note que a condição (4) é equivalente a: se $\mathbf{u} \neq \mathbf{0}$, então $\langle \mathbf{u}, \mathbf{u} \rangle > 0$. As condições $(1), (2)$ e (3) desta definição afirmam que um produto interno é uma *função bilinear* *conjugada* ou *sesquilinear*. Enquanto, a condição (4) afirma que ela é *definida positiva* e, em geral, a mais difícil de ser verificada. Quando $F = \mathbb{R}$, a condição (3) reduz à afirmação de simetria: $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$. Neste caso, um produto interno é uma *função bilinear, simétrica e definida positiva*. É importante ressaltar que, em qualquer caso, a condição (3) nos garante que $\langle \mathbf{u}, \mathbf{u} \rangle$ é sempre um número real, de modo que desigualdade da condição (4) tenha sentido.

Exemplo 7.1 *Seja $V = F^3$. Mostre que $\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{j=1}^{3} x_j \bar{y}_j$ é um produto interno sobre V, o qual chama-se de* produto interno usual (canônico)*. Em forma matricial $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{Y}^* \mathbf{X}$, com $\mathbf{X} = [\mathbf{u}]$ e $\mathbf{Y} = [\mathbf{v}]$. Em geral,*

$$f(\mathbf{X}, \mathbf{Y}) = \mathbf{Y}^* \mathbf{Q}^* \mathbf{Q} \mathbf{X} = (\mathbf{Q}\mathbf{Y})^* (\mathbf{Q}\mathbf{X}) = \langle \mathbf{Q}\mathbf{X}, \mathbf{Q}\mathbf{Y} \rangle,$$

para alguma matriz não singular $\mathbf{Q} \in F^{n \times n}$, é um produto interno sobre $F^{n \times 1}$.

Solução. Vamos provar apenas a condição (4). Dado $\mathbf{u} = (x_1, x_2, x_3) \in V$, é claro que $\langle \mathbf{u}, \mathbf{u} \rangle = \sum_{j=1}^{3} |x_j|^2 \geq 0$ e se $\mathbf{u} = \mathbf{0}$, então $\langle \mathbf{u}, \mathbf{u} \rangle = 0$. Por outro lado, se $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ e $\mathbf{u} \neq \mathbf{0}$, digamos $x_1 \neq 0$, então

$$|x_1|^2 + |x_2|^2 + |x_3|^2 = 0 \Leftrightarrow |x_1^{-1} x_2|^2 + |x_1^{-1} x_3|^2 = -1,$$

o que é impossível, pois o lado esquerdo da última equação é positivo enquanto o lado direito é negativo. ■

Exemplo 7.2 *Sejam $V = \mathbb{P}_1(\mathbb{R})$ e $f(x) = a_0 + a_1 x, g(x) = b_0 + b_1 x \in V$. Mostre que*

$$\langle f(x), g(x) \rangle = \int_0^1 f(t) g(t) dt = \sum_{i=0}^{1} \sum_{j=0}^{1} \frac{a_i b_j}{i + j + 1}$$

é um produto interno sobre V.

Solução. Vamos provar apenas a condição (4). Como

$$\langle f(x), f(x) \rangle = a_0^2 + a_0 a_1 + \frac{1}{3} a_1^2 = (a_0 + \frac{a_1}{2})^2 + \left(\frac{a_1}{\sqrt{12}} \right)^2$$

temos que $\langle f(x), f(x) \rangle \geq 0$ e $\langle f(x), f(x) \rangle = 0$ se, e somente se, $f(x) = 0$. ■

Exemplo 7.3 *Sejam $V = \mathbb{R}^2$. Mostre que $\langle \mathbf{u}, \mathbf{v}\rangle = x_1y_1 - x_1y_2 - x_2y_1 + 5x_2y_2$ é um produto interno sobre V. Note que $\langle \mathbf{u}, \mathbf{v}\rangle = \mathbf{Y}^t\mathbf{AX}$, com $\mathbf{X} = [\mathbf{u}]$, $\mathbf{Y} = [\mathbf{v}]$ e*

$$\mathbf{A} = \mathbf{E}_{11} - \mathbf{E}_{12} - \mathbf{E}_{21} + 5\mathbf{E}_{22} \in F^{2\times 2}.$$

Solução. Vamos provar apenas a condição (4). Como

$$\langle \mathbf{u}, \mathbf{u}\rangle = x_1^2 - 2x_1x_2 + 5x_2^2 = (x_1 - x_2)^2 + (2x_2)^2$$

temos que $\langle \mathbf{u}, \mathbf{u}\rangle \geq 0$ e $\langle \mathbf{u}, \mathbf{u}\rangle = 0$ se, e somente se, $\mathbf{u} = 0$. Note que $\mathbf{A}^t = \mathbf{A}$ e existe uma matriz não singular $\mathbf{P}$ tal que $\mathbf{P}^t\mathbf{AP} = \mathbf{D}$ é diagonal, pois aplicando as mesmas operações de linhas e colunas, $\mathbf{T}_{12}(1)\mathbf{AT}_{12}(1) = \mathbf{D}$ ou, de outro modo,

$$\left(\begin{array}{c|c} \mathbf{A} & \mathbf{I} \end{array}\right) \to \cdots \to \left(\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 0 & 4 & 1 & 1 \end{array}\right) = \left(\begin{array}{c|c} \mathbf{D} & \mathbf{P}^t \end{array}\right).$$

Observe que os elementos diagonais de $\mathbf{D}$ são positivos. Portanto, veremos que a função $\langle \mathbf{u}, \mathbf{v}\rangle = \mathbf{Y}^t\mathbf{AX}$ define um produto interno se, e somente se, $\mathbf{A}$ for uma matriz simétrica e definida positiva. ■

Exemplo 7.4 *Seja $V = l^2 = \{(x_n)_{n\in\mathbb{N}} \in \mathrm{Seq}(\mathbb{R}) : \sum_{n=1}^{\infty} x_n^2 < \infty\}$ o espaço das sequências de quadrado somável. Mostre que $\langle \mathbf{x}, \mathbf{y}\rangle = \sum_{n=1}^{\infty} x_ny_n$ é um produto interno sobre V.*

Solução. Vamos provar apenas que $\langle \mathbf{x}, \mathbf{y}\rangle$ está bem definida. Basta lembar que $2|x_ny_n| \leq |x_n|+|y_n|$ implica que a sequência $(x_ny_n)_{n\in\mathbb{N}}$ é convergente, de modo que $(x_ny_n)_{n\in\mathbb{N}} \in \mathrm{Seq}(\mathbb{R})$. ■

Um *espaço com produto interno* é qualquer espaço vetorial V sobre F munido de um produto interno, ou seja, um par $(V, \langle\cdot,\cdot\rangle)$. Quando $F = \mathbb{R}$, diremos que V é um *espaço Euclidiano* e quando $F = \mathbb{C}$, diremos que V é um *espaço unitário*. Note que qualquer subespaço W de V é também um espaço com produto interno sob a restrição do produto interno sobre V a W.

Lema 7.5 *Sejam V um espaço com produto interno. Então $\langle \mathbf{x}, \mathbf{0}\rangle = 0 = \langle \mathbf{0}, \mathbf{x}\rangle$, para todo $\mathbf{x} \in V$. Se $\langle \mathbf{u}, \mathbf{x}\rangle = \langle \mathbf{v}, \mathbf{x}\rangle$, para todo $\mathbf{x} \in V$, então $\mathbf{u} = \mathbf{v}$.*

Prova. Como $\langle \mathbf{u} - \mathbf{v}, \mathbf{x}\rangle = \langle \mathbf{u}, \mathbf{x}\rangle - \langle \mathbf{v}, \mathbf{x}\rangle = 0$, para todo $\mathbf{x} \in V$, temos, em particular, que $\langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v}\rangle = 0$. Portanto, $\mathbf{u} - \mathbf{v} = \mathbf{0}$ ou $\mathbf{u} = \mathbf{v}$. ■

O próximo resultado ressalta uma das principais diferença entre o espaço Euclidiano e o unitário.

Teorema 7.6 *Sejam V um espaço com produto interno e $T \in \mathcal{L}(V)$.*

1. *Se $\langle T(\mathbf{u}), \mathbf{v}\rangle = 0$, para todos $\mathbf{u}, \mathbf{v} \in V$, então $T = 0$.*
2. *Se $F = \mathbb{C}$ e $\langle T(\mathbf{u}), \mathbf{u}\rangle = 0$, para todo $\mathbf{u} \in V$, então $T = 0$. Conclua, com um exemplo, que isto é falso quando $F = \mathbb{R}$.*

Prova. Vamos provar apenas o item (2). Pondo $\mathbf{u} = \mathbf{x} + c\mathbf{y}$, para todos $\mathbf{x}, \mathbf{y} \in V$ e $c \in F$, é fácil verificar que

$$0 = \langle T(\mathbf{x} + c\mathbf{y}), \mathbf{x} + c\mathbf{y}\rangle = c\langle T(\mathbf{y}), \mathbf{x}\rangle + \bar{c}\langle T(\mathbf{x}), \mathbf{y}\rangle.$$

Escolhendo $c = 1$ e $c = i$, obtemos $\langle T(\mathbf{x}), \mathbf{y}\rangle = 0$, para todos $\mathbf{x}, \mathbf{y} \in V$. Assim, pelo item (1), $T = 0$. Note que se $F = \mathbb{R}$, então $\langle T(\mathbf{x}), \mathbf{y}\rangle = -\langle T(\mathbf{y}), \mathbf{x}\rangle$, para todos $\mathbf{x}, \mathbf{y} \in V$. Em particular, seja $T \in \mathcal{L}(\mathbb{R}^2)$ definido como $T(x, y) = (-y, x)$, o operador rotação por um ângulo de $90°$. Então $\langle T(\mathbf{u}), \mathbf{u}\rangle = 0$, para todo $\mathbf{u} \in \mathbb{R}^2$, mas $T \neq 0$. ■

Vamos finalizar esta seção considerando um espaço com produto interno V quando $\dim V = n$. Seja $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ uma base de V. Então dados $\mathbf{u}, \mathbf{v} \in V$, existem únicos $x_i, y_j \in F$ tais que $\mathbf{u} = \sum_{i=1}^n x_i\mathbf{u}_i$ e $\mathbf{v} = \sum_{j=1}^n y_j\mathbf{u}_j$. Assim, depois de alguns cálculos,

$$\langle \mathbf{u}, \mathbf{v}\rangle = \sum_{i=1}^n \sum_{j=1}^n x_i\bar{y}_j\langle \mathbf{u}_i, \mathbf{u}_j\rangle = \mathbf{Y}^*\mathbf{G}\mathbf{X}, \tag{7.1}$$

em que $\mathbf{X} = [\mathbf{u}]_\alpha, \mathbf{Y} = [\mathbf{v}]_\alpha$ e $\mathbf{G} = (\langle \mathbf{u}_i, \mathbf{u}_j\rangle) \in F^{n\times n}$ chama-se a *matriz do produto interno* na base α. Note que $\mathbf{G}^* = \mathbf{G}$ é Hermitiana e $\mathbf{X}^*\mathbf{G}\mathbf{X} > 0$, para todo $\mathbf{X} \neq \mathbf{O}$, implica que o sistema $\mathbf{G}\mathbf{X} = \mathbf{O}$ possui apenas a solução nula, de modo que $\mathbf{G}$ é não singular ($\det \mathbf{G} > 0$). Reciprocamente, seja $\mathbf{G} \in F^{n\times n}$ tal que $\mathbf{G}^* = \mathbf{G}$ e $\mathbf{X}^*\mathbf{G}\mathbf{X} > 0$, para todo $\mathbf{X} \neq \mathbf{O}$. Então a forma bilinear conjugada dada pela equação (7.1) induz um produto interno sobre V em relação à base α. Isto motiva a seguinte definição. Seja $\mathbf{A} \in F^{n\times n}$. Diremos que $\mathbf{A}$ é *definida positiva* se existir uma matriz não singular $\mathbf{Q} \in F^{n\times n}$ tal que $\mathbf{A} = \mathbf{Q}^*\mathbf{Q}$. Portanto, $\mathbf{A}$ é definida positiva se, e somente se, $\mathbf{A}^* = \mathbf{A}$ e $\mathbf{X}^*\mathbf{A}\mathbf{X} > 0$, para todo $\mathbf{X} \neq \mathbf{O}$.

Exercícios

1. Sejam $V = F^2$ e $f(\mathbf{u}, \mathbf{v}) = 2x_1\bar{y}_1 + x_1\bar{y}_2 + x_2\bar{y}_1 + x_2\bar{y}_2$. Mostre que f é um produto interno sobre V.

2. Sejam $V = F^2$ e $f(\mathbf{u}, \mathbf{v}) = x_1\bar{y}_1 x_2\bar{y}_2$. Verifique se f é um produto interno sobre V.

3. Sejam $V = F^3$ e $f(\mathbf{u}, \mathbf{v}) = x_1\bar{y}_1 + x_2\bar{y}_2 - x_3\bar{y}_3$. Verifique se f é um produto interno sobre V.

4. Sejam $V = \mathbb{R}^{2\times 2}$ e $f(\mathbf{A}, \mathbf{B}) = a_{11}b_{11} + 2a_{12}b_{12} + 3a_{21}b_{21} + a_{22}b_{22}$. Mostre que f é um produto interno sobre V.

5. Sejam $V = F^2$ e $f(\mathbf{u}, \mathbf{v}) = |y_1 - x_1| + |y_2 - x_2|$. Verifique se f é um produto interno sobre V..

6. Sejam $V = \mathbb{C}$ visto como um espaço vetorial sobre $\mathbb{R}$ e $f(\mathbf{z}, \mathbf{w}) = \operatorname{Re}(z\bar{w})$. Mostre que f é um produto interno sobre V.

7. Seja $T \in \mathcal{L}(\mathbb{R}^2)$ o operador rotação por um ângulo de $90°$. Determine todos os produtos internos f sobre $\mathbb{R}^2$ tais que $f(\mathbf{u}, T(\mathbf{u})) = 0$, para cada $\mathbf{u} \in \mathbb{R}^2$.

8. Sejam V um espaço com produto interno e $T \in \mathcal{L}(V)$. Que condições T deve satisfazer para que $g(\mathbf{u}, \mathbf{v}) = \langle T(\mathbf{u}), T(\mathbf{v})\rangle$ seja um produto interno sobre V?

9. Sejam $\mathbf{A} \in F^{m\times n}$, $\mathbf{X} \in F^{n\times 1}$ e $\mathbf{Y} \in F^{m\times 1}$. Mostre que $\langle \mathbf{AX}, \mathbf{Y}\rangle = \langle \mathbf{X}, \mathbf{A}^*\mathbf{Y}\rangle$.

10. Sejam f e g produtos internos sobre V e W. Mostre que $h((\mathbf{u}, \mathbf{w}), (\mathbf{v}, \mathbf{z})) = f(\mathbf{u}, \mathbf{v}) + g(\mathbf{w}, \mathbf{z})$ é um produto interno sobre $V \times W$.

11. Para cada $\mathbf{A} \in \mathbb{R}^{2\times 2}$, definimos $f(\mathbf{X}, \mathbf{Y}) = \mathbf{Y}^t\mathbf{AX}$, para todos $\mathbf{X}, \mathbf{Y}$ em $V = \mathbb{R}^{2\times 1}$. Mostre que f é um produto interno sobre V se, e somente se, $\mathbf{A} = \mathbf{A}^t, a_{11} > 0, a_{22} > 0$ e $\det(\mathbf{A}) > 0$.

12. Seja V um espaço vetorial sobre F. Mostre que a soma de dois produtos internos sobre V é um produto interno sobre V. A diferença de dois produtos internos sobre V é um produto interno sobre V? Mostre que um múltiplo positivo de um produto interno sobre V é um produto interno sobre V.

13. Descreva explicitamente todos os produtos internos sobre $\mathbb{R}$ e sobre $\mathbb{C}$.

14. Mostre que $f(\mathbf{A}, \mathbf{B}) = \operatorname{tr}(\mathbf{B}^* \mathbf{A})$ é um produto interno sobre $F^{n \times n}$.

15. Sejam V um espaço com produto interno e $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ uma base de V, $(c_1, \ldots, c_n)$ qualquer lista em F^n. Mostre que existe um único $\mathbf{u} \in V$ tal que $\langle \mathbf{u}_j, \mathbf{u} \rangle = c_j$, com $j = 1, \ldots, n$.

16. Mostre que qualquer espaço vetorial de dimensão finita sobre F pode ser munido de um produto interno.

17. Seja $V = \mathcal{C}([0, 1], \mathbb{R})$. Mostre que
$$\langle f, g \rangle = \int_0^1 f(t)g(t)dt$$
é um produto interno sobre V.

18. Seja $V = \mathcal{C}^1([0, 1], \mathbb{R})$ o espaço das funções de classe 1. Mostre que
$$\langle f, g \rangle = \int_0^1 f'(t)g'(t)dt + \int_0^1 f(t)g(t)dt$$
é um produto interno sobre V, o qual chama-se *produto interno de Sobolev*[1].

7.2 Norma e Distância

Sejam V um espaço unitário e $\mathbf{u}, \mathbf{v} \in V$. Então
$$\langle \mathbf{u}, \mathbf{v} \rangle = \operatorname{Re}\langle \mathbf{u}, \mathbf{v} \rangle + i \operatorname{Im}\langle \mathbf{u}, \mathbf{v} \rangle.$$
e $f(\mathbf{u}, \mathbf{v}) = \operatorname{Re}\langle \mathbf{u}, \mathbf{v} \rangle$ é um produto Euclidiano. Como $\operatorname{Im} z = \operatorname{Re}(-iz)$ temos que $\operatorname{Im}\langle \mathbf{u}, \mathbf{v} \rangle = \operatorname{Re}\langle \mathbf{u}, i\mathbf{v} \rangle$. Portanto,
$$\langle \mathbf{u}, \mathbf{v} \rangle = \operatorname{Re}\langle \mathbf{u}, \mathbf{v} \rangle + i \operatorname{Re}\langle \mathbf{u}, i\mathbf{v} \rangle$$
é completamente determinado pela sua "parte real". Além disso, se (V, f) for um espaço Euclidiano, então $(\tilde{V}, g)$ é o único espaço unitário, com g definida como
$$g(\mathbf{u} + i\mathbf{v}, \mathbf{w} + i\mathbf{z}) = f(\mathbf{u}, \mathbf{w}) + f(\mathbf{v}, \mathbf{z}) + i(f(\mathbf{v}, \mathbf{w}) - f(\mathbf{u}, \mathbf{z}))$$

[1] Sergei Lvovich Sobolev, 1908-1989, matemático russo.

Seja V um espaço com produto interno. A *norma* de um vetor $\mathbf{u} \in V$, induzida pelo produto interno $\langle \cdot, \cdot \rangle$, é definida como $\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$. Note que esta definição é possível, pois $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$, para todo $\mathbf{u} \in V$. A *forma quadrática* associada ao produto interno $\langle \cdot, \cdot \rangle$ é a função $q : V \to F$ definida como $q(\mathbf{u}) = \|\mathbf{u}\|^2$. Neste caso, depois de alguns cálculos, obtemos

$$q(\mathbf{u} \pm \mathbf{v}) = q(\mathbf{u}) \pm 2\,\mathrm{Re}(\langle \mathbf{u}, \mathbf{v} \rangle) + q(\mathbf{u}).$$

Como

$$\mathrm{Re}(\langle \mathbf{u}, \mathbf{v} \rangle) = \frac{1}{4}(q(\mathbf{u}+\mathbf{v}) - q(\mathbf{u}-\mathbf{v})) \text{ e } \mathrm{Re}(\langle \mathbf{u}, i\mathbf{v} \rangle) = \frac{1}{4}(q(\mathbf{u}+i\mathbf{v}) - q(\mathbf{u}-i\mathbf{v}))$$

temos que

$$\langle \mathbf{u}, \mathbf{v} \rangle) = \frac{1}{4}(\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2) + \frac{i}{4}(\|\mathbf{u}+i\mathbf{v}\|^2 - \|\mathbf{u}-i\mathbf{v}\|^2). \tag{7.2}$$

A equação (7.2) chama-se *identidade de polarização*. Diremos que $\mathbf{u} \in V$ é um *vetor unitário* se $\|\mathbf{u}\| = 1$. Se $\mathbf{v}$ for um vetor não nulo qualquer, então $\mathbf{u} = \frac{1}{\|\mathbf{v}\|}\mathbf{v}$ é um vetor unitário tal que $F[\mathbf{v}] = F[\mathbf{u}]$. Neste caso, diremos que $\mathbf{u}$ é a *normalização* de $\mathbf{v}$.

Teorema 7.7 (Desigualdade de Cauchy-Schwarz) *Seja V um espaço com produto interno. Então $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\|\|\mathbf{v}\|$, para todos $\mathbf{u}, \mathbf{v} \in V$. Conclua que a igualdade ocorre se, e somente se, os vetores $\mathbf{u}$ e $\mathbf{v}$ são LD.*

Prova. Se $\mathbf{v} = \mathbf{0}$, nada há para ser provado. Se $\mathbf{v} \neq \mathbf{0}$, então $a = \frac{1}{\|\mathbf{v}\|}$, $\mathbf{w} = a\mathbf{v}$ e $b = \langle \mathbf{u}, \mathbf{w} \rangle$ estão bem definidos. Como $\|\mathbf{w}\| = 1$ e $\langle \mathbf{w}, \mathbf{u} \rangle = \bar{b}$ temos, depois de alguns cálculos, que

$$0 \leq \|\mathbf{u} - b\mathbf{w}\|^2 = \|\mathbf{u}\|^2 - b\bar{b} - \bar{b}b + b\bar{b} = (\|\mathbf{u}\| + |b|)(\|\mathbf{u}\| - |b|).$$

Mas, $\|\mathbf{u}\| \geq 0$ e $|b| \geq 0$ implicam que

$$\|\mathbf{u}\| \geq |b| = |\langle \mathbf{u}, \mathbf{w} \rangle| = \frac{1}{\|\mathbf{v}\|}|\langle \mathbf{u}, \mathbf{v} \rangle| \Leftrightarrow |\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\|\|\mathbf{v}\|.$$

Note que se vale a igualdade, então $\mathbf{u} = b\mathbf{w} = ab\mathbf{v}$.

Teorema 7.8 *Seja V um espaço com produto interno. Então*:

1. $\|\mathbf{u}\| \geq 0$, *para todo* $\mathbf{u} \in V$, *e* $\|\mathbf{u}\| = 0$ *se, e somente se,* $\mathbf{u} = \mathbf{0}$.

2. $\|a\mathbf{u}\| = |a|\|\mathbf{u}\|$, *para todo* $\mathbf{u} \in V$ *e* $a \in F$.

3. $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$, *para todos* $\mathbf{u}, \mathbf{v} \in V$ (***Desigualdade Triangular***)

4. $\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 = 2(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2)$, *para todos* $\mathbf{u}, \mathbf{v} \in V$. (***Identidade do Paralelogramo***)

Prova. Fica como um exercício. ■

É muito importante, de um ponto de vista teórico e didático, fazermos uma abordagem axiomática da norma como fizemos com o produto interno. Uma função de V em F, denotada por $\|\cdot\|$, é uma *norma* sobre V se ela satisfaz as condições (1), (2) e (3) do Teorema 7.8. Neste caso, um *espaço vetorial normado* é qualquer espaço vetorial V sobre F munido de uma norma, ou seja, um par $(V, \|\cdot\|)$.

Exemplo 7.9 *Seja* $V = \mathbb{R}^2$. *Mostre que as funções* $\|\mathbf{u}\|_1 = |x_1| + |x_2|$ *e* $\|\mathbf{u}\|_\infty = \max\{|x_1|, |x_2|\}$ *são normas sobre* V.

Solução. Vamos provar apenas que $\|\cdot\|_1$ é uma norma sobre V. É claro que $\|\mathbf{u}\|_1 \geq 0$. Se $\|\mathbf{u}\|_1 = 0$, mas $\mathbf{u} \neq \mathbf{0}$, digamos $x_1 \neq 0$, então $|x_1| + |x_2| = 0$ se, e somente se, $|x_1^{-1}x_2| = -1$, o que é impossível em $\mathbb{R}$. Portanto, $\mathbf{u} = \mathbf{0}$. As outras condições são fáceis de serem verificadas. Pondo $\mathbf{u} = (1, 1)$ e $\mathbf{v} = (-1, 1)$, obtemos

$$\|\mathbf{u}+\mathbf{v}\|_1^2 + \|\mathbf{u}-\mathbf{v}\|_1^2 = 8 \text{ e } 2(\|\mathbf{u}\|_1^2 + \|\mathbf{v}\|_1^2) = 16.$$

Assim, não vale, em geral, a identidade do paralelogramo em um espaço vetorial normado qualquer. ■

Seja $(V, \|\cdot\|)$ um espaço vetorial normado. A função $d : V \times V \to F$ definida como $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u}-\mathbf{v}\|$ satisfaz as seguintes condições:

1. $d(\mathbf{u}, \mathbf{v}) \geq 0$, para todos $\mathbf{u}, \mathbf{v} \in V$, e $d(\mathbf{u}, \mathbf{v}) = 0$ se, e somente se, $\mathbf{u} = \mathbf{v}$.

2. $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$, para todos $\mathbf{u}, \mathbf{v} \in V$.

3. $d(\mathbf{u}, \mathbf{v}) \leq d(\mathbf{u}, \mathbf{w}) + d(\mathbf{v}, \mathbf{w})$, para todos $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.

A função d chama-se *distância* sobre V induzida pela norma $\|\cdot\|$. O par (V, d) chama-se um *espaço vetorial métrico*. Dado $\mathbf{u} \in V$ e $(\mathbf{u}_n)_{n\in\mathbb{N}} \in \mathrm{Seq}(V)$. Diremos que $(\mathbf{u}_n)_{n\in\mathbb{N}}$ *converge* para $\mathbf{u}$ em V se $\lim_{n\to\infty} d(\mathbf{u}_n, \mathbf{u}) = 0$ em F. Portanto. d e $\|\cdot\|$ são funções contínuas sobre V, isto é, se $\lim_{n\to\infty} x_n = x$ e $\mathbf{v} \in V$, então $\lim_{n\to\infty} \|x_n\mathbf{u}\| = \|x\mathbf{u}\|$ (prove isto!).

Proposição 7.10 *Seja $(V, \|\cdot\|)$ um espaço vetorial normado sobre $\mathbb{R}$. Enão $\|\cdot\|$ provém de um produto interno sobre V se, e somente se, $\|\cdot\|$ satisfaz a identidade do paralelogramo.*

Prova. A ida segue do Teorema 7.8. Reciprocamente, a identidade de polarização sugere que consideremos a função $f : V \times V \to \mathbb{R}$ definida como

$$f(\mathbf{u}, \mathbf{v}) = \frac{1}{4}(\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2) = \frac{1}{2}(\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2).$$

É claro que $f(\mathbf{u}, \mathbf{u}) = \|\mathbf{u}\|^2 \geq 0$, para todo $\mathbf{u} \in V$, e $f(\mathbf{u}, \mathbf{u}) = 0$ se, e somente se, $\mathbf{u} = \mathbf{0}$. Além disso, $f(\mathbf{u}, \mathbf{v}) = f(\mathbf{v}, \mathbf{u})$, para todos $\mathbf{u}, \mathbf{v} \in V$. Dados $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, usando a identidade do paralelogramo e algumas manipulações, obtemos

$$\begin{array}{rcl} 8(f(\mathbf{u}, \mathbf{w}) + f(\mathbf{v}, \mathbf{w})) & = & 2(\|\mathbf{u}+\mathbf{w}\|^2 - \|\mathbf{u}-\mathbf{w}\|^2 + \|\mathbf{v}+\mathbf{w}\|^2 - \|\mathbf{v}-\mathbf{w}\|^2) \\ & = & 2(\|\mathbf{u}+\mathbf{w}\|^2 + \|\mathbf{v}+\mathbf{w}\|^2) - 2(\|\mathbf{v}-\mathbf{w}\|^2 \\ & + & \|\mathbf{u}-\mathbf{w}\|^2) \\ & = & \|\mathbf{u}+\mathbf{v}+2\mathbf{w}\|^2 - \|\mathbf{u}+\mathbf{v}-2\mathbf{w}\|^2 \\ & = & 4f(\mathbf{u}+\mathbf{v}, 2\mathbf{w}). \end{array}$$

Em particular, se $\mathbf{v} = \mathbf{0}$, então $2f(\mathbf{u}, \mathbf{w}) = f(\mathbf{u}, 2\mathbf{w})$. Assim,

$$2f(\mathbf{u}+\mathbf{v}, \mathbf{w}) = f(\mathbf{u}+\mathbf{v}, 2\mathbf{w}) = 2f(\mathbf{u}, \mathbf{w}) + 2f(\mathbf{v}, \mathbf{w}).$$

É fácil ver, pelo Exemplo 4.10, que $f(r\mathbf{u}, \mathbf{v}) = rf(\mathbf{u}, \mathbf{v})$, para todo $r \in \mathbb{Q}$. Dado $a \in \mathbb{R}$, existe uma sequência $(r_n)_{n\in\mathbb{N}} \in \mathrm{Seq}(\mathbb{Q})$ tal que $\lim_{n\to\infty} r_n = a$, de modo que $(r_n\mathbf{u})_{n\in\mathbb{N}} \in \mathrm{Seq}(V)$ e $\lim_{n\to\infty} r_n\mathbf{u} = a\mathbf{u}$. Como $\|\cdot\|$ é contínua temos que

$$f(a\mathbf{u}, \mathbf{v}) = f(\lim_{n\to\infty} r_n\mathbf{u}, \mathbf{v}) = \lim_{n\to\infty} f(r_n\mathbf{u}, \mathbf{v}) = (\lim_{n\to\infty} r_n)f(\mathbf{u}, \mathbf{v}) = af(\mathbf{u}, \mathbf{v}).$$

Portanto, f é um produto internos sobre V. Finalmente, $4f(\mathbf{u}, \mathbf{u}) = \|\mathbf{u}+\mathbf{u}\|^2 = 4\|\mathbf{u}\|^2$, isto é, $\|\mathbf{u}\| = \sqrt{f(\mathbf{u}, \mathbf{u})}$. Vale observar, com um pouco mais de trabalho, que esta prova pode ser estendida para $F = \mathbb{C}$. ■

Tendo definido o conceito de comprimento e distância em um espaço com produto interno qualquer, é natural perguntar: se o conceito de ângulo pode ser generalizado? A resposta é verdadeira se nosso corpo é os reais $\mathbb{R}$, mas é falsa no corpo dos números complexos $\mathbb{C}$. Seja V um espaço Euclidiano. Para quaisquer $\mathbf{u}, \mathbf{v} \in V - \{\mathbf{0}\}$, o *ângulo* entre $\mathbf{u}$ e $\mathbf{v}$ é definido como o ângulo θ tal que $0 \leq \theta \leq \pi$ e $\langle \mathbf{u}, \mathbf{v} \rangle = \|\mathbf{u}\|\|\mathbf{v}\| \cos\theta$.

Note que o produto interno foi usada para definir a noção de direção. Pela desigualdade de Cauchy-Schwarz,

$$-1 \leq \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\|\mathbf{v}\|} \leq 1$$

e, assim, o ângulo θ sempre existe e é único, pois $\cos\theta$ não muda quando $\mathbf{u}$ e $\mathbf{v}$ são substituídos por $a\mathbf{u}$ e $b\mathbf{v}$, para todos $a, b \in \mathbb{R}_+^\times$.

Exercícios

1. Seja V um espaço vetorial sobre F. Mostre que d é uma distância sobre V se, e somente se, (a) $d(\mathbf{u}, \mathbf{v}) = 0$ se, e somente se, $\mathbf{u} = \mathbf{v}$. (b) Para quaisquer $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, $d(\mathbf{u}, \mathbf{w}) \leq d(\mathbf{u}, \mathbf{v}) + d(\mathbf{w}, \mathbf{v})$.

2. Seja (V, d) um espaço vetorial métrico. Mostre que as seguintes funções, são distâncias sobre V.

   (a) $d_1(\mathbf{u}, \mathbf{v}) = \min\{1, d(\mathbf{u}, \mathbf{v})\}$, para todos $\mathbf{u}, \mathbf{v} \in V$.

   (b) $d_2(\mathbf{u}, \mathbf{v}) = \sqrt{d(\mathbf{u}, \mathbf{v})}$, para todos $\mathbf{u}, \mathbf{v} \in V$.

   (c) $d_3(\mathbf{u}, \mathbf{v}) = \frac{d(\mathbf{u}, \mathbf{v})}{1+d(\mathbf{u}, \mathbf{v})}$, para todos $\mathbf{u}, \mathbf{v} \in V$.

3. Sejam (V, d) um espaço vetorial métrico e $f : V \to V$ qualquer função injetora. Mostre que $d_1(\mathbf{u}, \mathbf{v}) = d(f(\mathbf{u}), f(\mathbf{v}))$, para todos $\mathbf{u}, \mathbf{v} \in V$, é uma distância sobre V.

4. Seja (V, d) um espaço vetorial métrico. Mostre, para quaisquer $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, que $|d(\mathbf{u}, \mathbf{v}) - d(\mathbf{v}, \mathbf{w})| \leq d(\mathbf{v}, \mathbf{w})$.

5. Seja $(V, \|\cdot\|)$ um espaço vetorial normado. Mostre que se $(\mathbf{u}_n)_{n\in\mathbb{N}}$ converge para $\mathbf{u}$ em V, então $(\|\mathbf{u}_n\|)_{n\in\mathbb{N}}$ converge para $\|\mathbf{u}\|$ em F.

6. Sejam $V = \mathbb{R}^2$ e $f(\mathbf{u}, \mathbf{v}) = 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2$. Calcule o ângulo entre os vetores $\mathbf{u} = (1, 1)$ e $\mathbf{v} = (1, -1)$.

7. Sejam $V = \mathbb{R}^2, \mathbf{u}, \mathbf{v} \in V - \{\mathbf{0}\}$ e θ o ângulo entre $\mathbf{u}$ e $\mathbf{v}$. Mostre que $\langle \mathbf{u}, \mathbf{v} \rangle = \|\mathbf{u}\|\|\mathbf{v}\| \cos\theta$ é equivalente a *Lei dos Cossenos*. Faça o mesmo para $V = \mathbb{R}^3$.

8. Seja V um espaço com produto interno. Mostre que $\langle \mathbf{u} + \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle = \|\mathbf{u}\|^2 - 2i \operatorname{Im}\langle \mathbf{u}, \mathbf{v} \rangle - \|\mathbf{v}\|^2$, para todos $\mathbf{u}, \mathbf{v} \in V$. Mostre que se $\operatorname{Im}\langle \mathbf{u}, \mathbf{v} \rangle = 0$, então $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ se, e somente se, $\|\mathbf{u} - \mathbf{v}\| = \|\mathbf{u} + \mathbf{v}\|$.

9. (**Identidade de Apolônios**[2]) Seja V um espaço Euclidiano. Mostre que
$$\|\mathbf{w}-\mathbf{u}\|^2+\|\mathbf{w}-\mathbf{v}\|^2=\frac{1}{2}\|\mathbf{u}-\mathbf{v}\|^2+2\|\mathbf{w}-\frac{1}{2}(\mathbf{u}+\mathbf{v})\|^2,$$
para todos $\mathbf{u},\mathbf{v},\mathbf{w}\in V$.

10. Seja V um espaço vetorial com os produtos internos f e g. Mostre que se $\|\mathbf{u}\|_f=\|\mathbf{u}\|_g$, para todo $\mathbf{u}\in V$, então $f=g$.

11. Seja $(V,\|\cdot\|)$ um espaço vetorial normado.
    (a) Mostre que $f:F\times V\to V$ definida como $f(x,\mathbf{u})=x\mathbf{u}$ é contínua.
    (b) Mostre que $g:V\times V\to V$ definida como $g(\mathbf{u},\mathbf{v})=\mathbf{u}+\mathbf{v}$ é contínua.

12. Sejam V um espaço Euclidiano e $\mathbf{u},\mathbf{v}\in V$, com $\mathbf{v}\neq\mathbf{0}$. Mostre que a função $f(x)=\|\mathbf{u}+x\mathbf{v}\|^2$, para todo $x\in\mathbb{R}$, possui um ponto de mínimo.

13. Seja V espaço com produto interno. Mostre que $\|\mathbf{u}+\mathbf{v}\|=\|\mathbf{u}\|+\|\mathbf{v}\|$ se, e somente se, $\mathbf{u}$ e $\mathbf{v}$ são LD.

14. Seja V um espaço vetorial com os produtos internos f,g e $\dim V=n$. Mostre que existe uma constante c independente de $\mathbf{u}$ tal que $\|\mathbf{u}\|_g\leq c\|\mathbf{u}\|_f$, para todo $\mathbf{u}\in V$.

15. (**Desigualdade de Ptolomeu**[3]) Seja V um espaço Euclidiano. Mostre que
$$\|\mathbf{u}-\mathbf{v}\|\|\mathbf{w}\|\leq\|\mathbf{v}-\mathbf{w}\|\|\mathbf{u}\|+\|\mathbf{w}-\mathbf{u}\|\mathbf{v}\|,\ \ \forall\ \mathbf{u},\mathbf{v},\mathbf{w}\in V.$$
Conclua que a igualdade ocorre se, e somente se, $\mathbf{0}\in\{\mathbf{u},\mathbf{v},\mathbf{w}\}$ ou os vetores $\mathbf{u},\mathbf{v},\mathbf{w},\mathbf{0}$ estão sob um círculo em um subespaço de dimensão 2 em V e os pares $(\mathbf{u},\mathbf{v})$ e $(\mathbf{w},\mathbf{0})$ separam-se.

16. Seja $(V,\|\cdot\|)$ um espaço vetorial normado sobre $\mathbb{R}$, com $\dim V=n$. Mostre que $\|\cdot\|$ provém de um produto interno sobre V se, e somente se, o conjunto $E=\{\mathbf{u}\in V:\|\mathbf{u}\|=1\}$ representa um elipsoide.

[2]Apollonius de Perga, 282-190 a.C., matemático e astrônomo grego.
[3]Claudius Ptolemy, 100-170, matemático e astrônomo grego.

7.3 Ortogonalidade

Sejam V um espaço com produto interno, $\mathbf{u} \in V$ e α, β subconjuntos de V, denotamos por

$$\langle \alpha, \beta \rangle = \{\langle \mathbf{u}, \mathbf{v} \rangle : \mathbf{u} \in \alpha, \mathbf{v} \in \beta\} \text{ e } \langle \mathbf{u}, \beta \rangle = \{\langle \mathbf{u}, \mathbf{v} \rangle : \mathbf{v} \in \beta\}.$$

Sejam $\mathbf{u}, \mathbf{v} \in V$. Diremos que $\mathbf{u}$ e $\mathbf{v}$ são *ortogonais* se $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ e denotamos por $\mathbf{u} \perp \mathbf{v}$. Diremos que α e β são *ortogonais* se $\langle \alpha, \beta \rangle = \{\mathbf{0}\}$ e denotamos por $\alpha \perp \beta$. Em particular, $\mathbf{u} \perp \beta$ quando $\langle \mathbf{u}, \beta \rangle = \{\mathbf{0}\}$. É muito importante ressaltar que ortogonalidade não é preservada por isomorfismo, ou seja, $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ não garante que $\langle T(\mathbf{u}), T(\mathbf{v}) \rangle = 0$, confira o Exemplo 7.15 a seguir..

Proposição 7.11 *Seja V um espaço com produto interno. Então*:

1. $\mathbf{0} \perp \mathbf{u}$, *para todo* $\mathbf{u} \in V$.
2. *Se* $\mathbf{u} \perp \mathbf{v}$, *então* $\mathbf{v} \perp \mathbf{u}$, *para todos* $\mathbf{u}, \mathbf{v} \in V$.
3. *Se* $\mathbf{u} \perp \mathbf{v}$, *para todo* $\mathbf{v} \in V$, *então* $\mathbf{u} = \mathbf{0}$.
4. *Se* $\mathbf{u} \perp \mathbf{w}$ *e* $\mathbf{v} \perp \mathbf{w}$, *então* $(\mathbf{u} + \mathbf{v}) \perp \mathbf{w}$, *para todos* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.
5. *Se* $\mathbf{u} \perp \mathbf{v}$, *então* $(a\mathbf{u}) \perp \mathbf{v}$, *para todos* $\mathbf{u}, \mathbf{v} \in V$ *e* $a \in F$.

Prova. Vamos provar apenas o item (1). Como, para cada $\mathbf{u} \in V$ fixado, a função $f : V \to F$ definida como $f(\mathbf{v}) = \langle \mathbf{v}, \mathbf{u} \rangle$ é linear temos que $f(\mathbf{0}) = 0$, ou seja, $\langle \mathbf{0}, \mathbf{u} \rangle = 0$. ■

Sejam V um espaço com produto interno e $\alpha = (\mathbf{u}_s)_{s \in S}$ uma família sobre V. Diremos que α é um *sistema ortonormal* se $\langle \mathbf{u}_s, \mathbf{u}_t \rangle = \delta_{st}$, para todos $s, t \in S$, e α é um *sistema ortonormal completo* se $\mathbf{0}$ é o único vetor ortogonal a qualquer vetor de α. Quando S for finito, diremos que α é um *sistema de coordenadas cartesianas*. Por exemplo, F^n possuiu um sistema de coordenadas cartesianas $\alpha = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$.

Teorema 7.12 *Seja V um espaço com produto interno. Se α for um sistema ortonormal sobre V, então α é LI.*

Prova. Sejam $\mathbf{u}_{s_1}, \ldots, \mathbf{u}_{s_n}$ vetores distintos de α e $x_1, \ldots, x_n \in F$ tais que

$$x_1\mathbf{u}_{s_1} + \cdots + x_n\mathbf{u}_{s_n} = \mathbf{0}.$$

Então $0 = \langle \mathbf{0}, \mathbf{u}_{s_j} \rangle = \sum_{i=1}^{n} x_i \langle \mathbf{u}_{s_i}, \mathbf{u}_{s_j} \rangle = x_j$. Assim, $x_j = 0$, para $j = 1, \ldots, n$. Portanto, α é LI. ■

Sejam V um espaço com produto interno e $\alpha = (\mathbf{u}_s)_{s \in S}$ uma família sobre V. Diremos que α é *ortogonal* se $\mathbf{u}_s \perp \mathbf{u}_t$, para todos $s, t \in S$.

Corolário 7.13 *Seja V um espaço com produto interno. com $\dim V = n$. Se $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ for um sistema ortonormal sobre V, então α é uma base ortogonal de V.*

Prova. Fica como um exercício. ■

Exemplo 7.14 *Seja $V = F^2$ com o produto interno do Exemplo 7.3. Mostre que $\alpha = \{(2, 1), (-3, 1)\}$ é uma base ortogonal de V.*

Solução. Como $\langle (2, 1), (-3, 1) \rangle = 2(-3) - 2 \cdot 1 - 1(-3) + 5 \cdot 1 \cdot 1 = 0$ temos que os vetores $(2, 1)$ e $(-3, 1)$ são LI. Portanto, β é uma base ortogonal de V. ■

Exemplo 7.15 *Seja $V = \mathbb{P}_1(F)$ com o produto interno do Exemplo 7.2. Mostre que $\alpha = \{1, 1 - 2x\}$ é uma base ortogonal de V.*

Solução. Como $\langle 1, 1 - 2x \rangle = 0$ temos que os vetores 1 e $1 - 2x$ são LI. Portanto, α é uma base ortogonal de V. Observe que se F^2 for o espaço do Exemplo 7.14, então a função $T : F^2 \to \mathbb{P}_1(F)$ definida como $T(a, b) = a + bx$ é claramente um isomorfismo. Assim, $\langle (2, 1), (-3, 1) \rangle = 0$, mas $\langle T(2, 1), T(-3, 1) \rangle = -\frac{37}{6}$, ou seja, T não preserva ortogonalidade. ■

Exemplo 7.16 *Seja $V = l^2$ com o produto interno do Exemplo 7.4. Mostre que $\alpha = \{\mathbf{e}_1, \ldots, \mathbf{e}_n, \ldots\}$, com $\mathbf{e}_n = (\delta_{kn})_{k \in \mathbb{N}} \in V$, é um sistema ortonormal sobre V, mas não é uma base ortogonal de V.*

Solução. É claro que $\langle \mathbf{e}_m, \mathbf{e}_n \rangle = \delta_{mn}$, isto é, α é um sistema ortonormal de V. Assim, α é LI, mas $V \neq \mathbb{R}[\alpha]$, pois $\mathbf{u} = (n^{-1})_{n \in \mathbb{N}} \in V$ e $\mathbf{u} \notin \mathbb{R}[\alpha]$. ■

Sejam V um espaço com produto interno e $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ uma base ortogonal de V. Então é fácil verificar que $\mathbf{u} = \sum_{i=1}^{n} x_i \mathbf{u}_i$, para todo $\mathbf{u} \in V$, com $x_i = \langle \mathbf{u}, \mathbf{u}_i \rangle \|\mathbf{u}_i\|^{-2} \in F$. Neste caso, $[\mathbf{u}]_\alpha = (x_1, \ldots, x_n)^t$. Os escalares x_i chamam-se *coeficientes de Fourier*[4] de $\mathbf{u}$ em relação à base α ou *as componentes de* $\mathbf{u}$ *na direção de* $\mathbf{u}_i$ e a expressão para $\mathbf{u}$ no lado direito chama-se *expansão de Fourier* de $\mathbf{u}$ em relação à

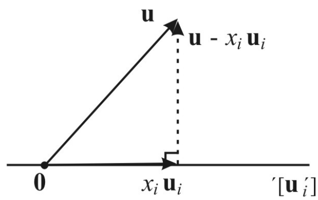


Figura 7.1: Projeção de $\mathbf{u}$ sobre $F[\mathbf{u}_i]$.

base α. Os $x_i\mathbf{u}_i$ são os *vetores projeções* de $\mathbf{u}$ sobre $F[\mathbf{u}_i]$. Neste caso, $(\mathbf{u}-x_i\mathbf{u}_i) \perp \mathbf{u}_i$, para $i = 1, \ldots, n$, confira Figura 7.1.

Exemplo 7.17 *Sejam $V = \mathbb{R}^2$ com o produto interno usual e $\alpha = \{(1,1),(-1,1)\}$. Mostre que α é uma base ortogonal de V. Calcule $[(2,3)]_\alpha$.*

Solução. É claro que $\langle (1,1),(-1,1)\rangle = 0$. Assim, α é uma base ortogonal de V. como $x_1 = \frac{5}{2}$ e $x_2 = \frac{1}{2}$ temos que $[(2,3)]_\alpha = \frac{1}{2}(5,1)^t$. ■

Exemplo 7.18 *Sejam V um espaço com produto interno e $\mathbf{u} \in V$. Mostre que $\mathbf{u} \perp \mathbf{v}$ se, e somente se, $\|\mathbf{u}\| \leq \|\mathbf{u} - a\mathbf{v}\|$, para todo $a \in F$, ou se, e somente se, $\|\mathbf{u} + a\mathbf{v}\| = \|\mathbf{u} - a\mathbf{v}\|$, para todo $a \in F$.*

Solução. Se $\mathbf{u} \perp \mathbf{v}$, então $\|\mathbf{u} - a\mathbf{v}\|^2 = \|\mathbf{u}\|^2 - 2\operatorname{Re}\langle \mathbf{u}, a\mathbf{v}\rangle + \|a\mathbf{v}\|^2$ implica que $\|\mathbf{u} - a\mathbf{v}\|^2 \geq \|\mathbf{u}\|^2$, pois $\|a\mathbf{v}\|^2 \geq 0$, para todo $a \in F$. Reciprocamente, se $\mathbf{v} = \mathbf{0}$, nada há para ser provado. Se $\mathbf{v} \neq \mathbf{0}$ e $\|\mathbf{u}\| \leq \|\mathbf{u} - a\mathbf{v}\|$, para todo $a \in F$, então

$$-2\operatorname{Re}\langle \mathbf{u}, a\mathbf{v}\rangle + \|a\mathbf{v}\|^2 = \|\mathbf{u} - a\mathbf{v}\|^2 - \|\mathbf{u}\|^2 \geq 0.$$

Vamos escolher $a = \langle \mathbf{u}, \mathbf{v}\rangle \|\mathbf{v}\|^{-2}$, de modo que $\langle \mathbf{u}, a\mathbf{v}\rangle$ seja um número real. Assim, $-|\langle \mathbf{u}, \mathbf{v}\rangle|^2 \|\mathbf{v}\|^{-2} \geq 0$. Portanto, $\langle \mathbf{u}, \mathbf{v}\rangle = 0$ ou $\mathbf{u} \perp \mathbf{v}$. ■

Exercícios

1. Sejam $V = F^n$ um espaço com produto interno e $\mathbf{A} \in F^{n\times n}$. Mostre que $\mathbf{A}^*\mathbf{A}$ é diagonal se, e somente se, $\mathbf{C}_i \perp \mathbf{C}_j = 0$, se $i \neq j$, em V.

[4]Jean-Baptiste Joseph Fourier, 1768-1830, matemático e físico francês.

2. Seja $V = F^4$ com o produto interno usual. Mostre que

$$\alpha = \{(1, 1, 0, -1), (1, 2, 1, 3), (1, 1, -9, 2), (16, -13, 1, 3)\}$$

é uma base ortogonal de V. Calcule $[(a, b, c, d)]_\alpha$.

3. Sejam V um espaço com produto interno e $x, y, z, t \in F$. Mostre que se $\mathbf{u}, \mathbf{v} \in V$ forem ortogonais e unitários, então

$$(x\mathbf{u} + y\mathbf{v}) \perp (z\mathbf{u} + t\mathbf{v}) \Leftrightarrow xz + yt = 0.$$

4. (**Teorema de Pitágoras**) Sejam V um espaço com produto interno e vetores $\mathbf{u}_1, \ldots, \mathbf{u}_n \in V$. Mostre que se $\mathbf{u}_i \perp \mathbf{u}_j$, quando $i \neq j$, então

$$\|\mathbf{u}_1 + \cdots + \mathbf{u}_n\|^2 = \|\mathbf{u}_1\|^2 + \cdots + \|\mathbf{u}_n\|^2.$$

5. Sejam V um espaço Euclidiano, $S = \{\mathbf{v} \in V : \|\mathbf{v}\| = 1\}$ e $f : S \to \mathbb{R}$ a função definida como $f(\mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$, para qualquer vetor unitário $\mathbf{u} \in V$.

   (a) Mostre que o valor máximo de f é 2 e o valor mínimo de f é 0.

   (b) Mostre que $f(\mathbf{v}) = \sqrt{2}$ se, e somente se, $\mathbf{u} \perp \mathbf{v}$.

6. Sejam $V = \mathbb{P}_n(\mathbb{R})$ e $c_0, \ldots, c_n \in \mathbb{R}$ distintos aos pares. Mostre que a função $f(p(x), q(x)) = \sum_{i=0}^{n} p(c_i)q(c_i)$ é um produto interno sobre V. Conclua que $L_i(x) = \prod_{i \neq k}(c_i - c_k)^{-1}(x - c_k)$ formam uma base ortonormal de V.

7.4 Processo de Ortogonalização de Gram-Schmidt

Sejam V um espaço com produto interno e $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ qualquer lista de vetores LI em V. Então podemos construir vetores ortogonais $\mathbf{v}_1, \ldots, \mathbf{v}_n$ em V tal que a lista $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ seja uma base de $W = F[\mathbf{u}_1, \ldots, \mathbf{u}_k]$, para $k = 1, \ldots, n$, como segue: para iniciar o processo vamos escolher $\mathbf{v}_1$ como qualquer um dos vetores da lista α, digamos $\mathbf{v}_1 = \mathbf{u}_1$. Já vimos, na Figura 7.1. que o vetor

$$\mathbf{v}_2 = \mathbf{u}_2 - \langle \mathbf{u}_2, \mathbf{v}_1 \rangle \|\mathbf{v}_1\|^{-2}\mathbf{v}_1$$

é ortogonal ao vetor $\mathbf{v}_1$ e é claro que $F[\mathbf{v}_1, \mathbf{v}_2] = F[\mathbf{u}_1, \mathbf{u}_2]$. Como os vetores de $F[\mathbf{v}_1, \mathbf{v}_2]$ são da forma $x_1\mathbf{v}_1 + x_2\mathbf{v}_2$ e $\mathbf{u}_3 \notin F[\mathbf{v}_1, \mathbf{v}_2]$ temos que

$$\langle \mathbf{u}_3 - (x_1\mathbf{v}_1 + x_2\mathbf{v}_2), \mathbf{v}_1\rangle = 0 \Leftrightarrow x_1 = \langle \mathbf{u}_3, \mathbf{v}_1\rangle \|\mathbf{v}_1\|^{-2}.$$

De modo análogo, $x_2 = \langle \mathbf{u}_3, \mathbf{v}_2\rangle \|\mathbf{v}_2\|^{-2}$. Logo, o vetor

$$\mathbf{v}_3 = \mathbf{u}_3 - \langle \mathbf{u}_3, \mathbf{v}_1\rangle \|\mathbf{v}_1\|^{-2}\mathbf{v}_1 - \langle \mathbf{u}_3, \mathbf{v}_2\rangle \|\mathbf{v}_2\|^{-2}\mathbf{v}_2$$

é tal que $\mathbf{v}_1 \perp \mathbf{v}_3, \mathbf{v}_2 \perp \mathbf{v}_3$ e $F[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3] = F[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3]$, confira Figura 7.2. Continuando desta maneira, obtemos uma base ortogonal $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ de $F[\mathbf{u}_1, \ldots, \mathbf{u}_k]$, em que

$$\mathbf{v}_k = \mathbf{u}_k - \sum_{i=1}^{k-1} \frac{\langle \mathbf{u}_k, \mathbf{v}_i\rangle}{\|\mathbf{v}_i\|^2}\mathbf{v}_i, k = 1, \ldots, n.$$

Note que existem 2^n tais bases. Este processo de ortogonalização é conhecido como o *processo de ortogonalização de Gram-Schmidt.*[5]

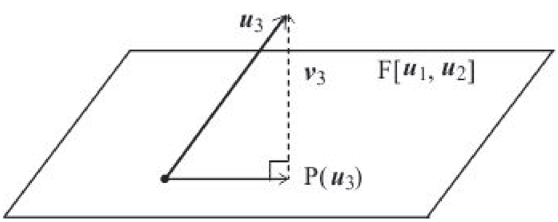


Figura 7.2: Projeção de $\mathbf{u}_3$ sobre $F[\mathbf{u}_1, \mathbf{u}_2]$.

Proposição 7.19 *Qualquer espaço com produto interno e dimensão finita possui uma base ortonormal.*

Prova. Basta escolher $\mathbf{w}_k = \|\mathbf{v}_k\|^{-1}\mathbf{v}_k$, com $k = 1, \ldots, n = \dim V$. ■

Exemplo 7.20 *Sejam* $V = \mathbb{R}^3$ *com o produto interno usual e* $\alpha = \{\mathbf{v}_1, \mathbf{u}_2, \mathbf{u}_3\}$, *com* $\mathbf{u}_1 = (3, 0, -3), \mathbf{u}_2 = (-1, 1, 2)$ *e* $\mathbf{u}_3 = (2, 1, 1)$ *uma base de* V. *Determine a partir de* α *uma base ortonormal de* V.

[5] Erhard Schmidt, 1876-1959, matemático alemão.

Solução. Pelo exposto, vamos escolher um vetor inicial $\mathbf{w}_1$, digamos

$$\mathbf{w}_1 = \|\mathbf{u}_1\|^{-1}\mathbf{u}_1 = 2^{-\frac{1}{2}}(1, 0, -1).$$

Assim, $\mathbf{v}_2 = \mathbf{u}_2 - \langle \mathbf{u}_2, \mathbf{w}_1 \rangle \mathbf{w}_1 = 2^{-1}(1, 2, 1)$ e $\mathbf{w}_2 = \|\mathbf{v}_2\|^{-1}\mathbf{v}_2 = 6^{-\frac{1}{2}}(1, 2, 1)$. Finalmente,

$$\mathbf{v}_3 = \mathbf{u}_3 - \langle \mathbf{u}_3, \mathbf{w}_1 \rangle \mathbf{w}_1 - \langle \mathbf{u}_3, \mathbf{w}_2 \rangle \mathbf{w}_2 = 3^{-1}(2, -2, 2)$$

e $\mathbf{w}_3 = \|\mathbf{v}_3\|^{-1}\mathbf{v}_3 = 3^{-\frac{1}{2}}(1, -1, 1)$. Portanto, $\beta = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ é uma base ortonormal de V. Observe, para qualquer $\mathbf{u} = (x_1, x_2, x_3) \in V$, que

$$\mathbf{u} = \sum_{i=1}^{3} \langle \mathbf{u}, \mathbf{w}_i \rangle \mathbf{w}_i = \frac{x_1 - x_3}{2}\mathbf{w}_1 + \frac{x_1 + 2x_2 + x_3}{6}\mathbf{w}_2 + \frac{x_1 - x_2 + x_3}{3}\mathbf{w}_3.$$

Portanto, a base $\beta^* = \{f_1, f_2, f_3\}$ de V^*, a qual é dual a β, é definida explicitamente como os coeficientes de Fourier $f_i(\mathbf{u}) = \langle \mathbf{u}, \mathbf{w}_i \rangle$ em relação à base β. ■

Teorema 7.21 *Sejam V um espaço com produto interno, W qualquer subespaço de V e $\mathbf{u} \in V$.*

1. *Se $\mathbf{w} \in W$, então $(\mathbf{u} - \mathbf{w}) \perp \mathbf{z}$, para todo $\mathbf{z} \in W$ se, e somente se, $\|\mathbf{u} - \mathbf{w}\|^2 = \inf\{\|\mathbf{u} - \mathbf{v}\|^2 : \mathbf{v} \in W\}$.*

2. *Se a* melhor aproximação $\mathbf{w}$ *de* $\mathbf{u}$ *por vetores de W existir, é única.*

3. *Se $\dim W = n$ e $\alpha = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ for qualquer base ortogonal de W, então*

$$\mathbf{w} = \sum_{i=1}^{n} \langle \mathbf{u}, \mathbf{v}_i \rangle \|\mathbf{v}_i\|^{-2}\mathbf{v}_i$$

é a melhor aproximação de $\mathbf{u}$ por vetores de W e $d(\mathbf{u}, W) = \|\mathbf{u} - \mathbf{w}\|$.

Prova. Primeiro confira o Exemplo 7.18 com $W = F[\mathbf{v}]$ e, para qualquer $\mathbf{v} \in W$, $\mathbf{u} - \mathbf{v} = (\mathbf{u} - \mathbf{w}) + (\mathbf{w} - \mathbf{v})$ implica que

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u} - \mathbf{w}\|^2 + 2\,\mathrm{Re}\langle \mathbf{u} - \mathbf{w}, \mathbf{w} - \mathbf{v} \rangle + \|\mathbf{w} - \mathbf{v}\|^2 \tag{7.3}$$

Se $(\mathbf{u} - \mathbf{w}) \perp \mathbf{z}$, para todo $\mathbf{z} \in W$, então $\|\mathbf{u} - \mathbf{v}\|^2 \geq \|\mathbf{u} - \mathbf{w}\|^2$, pois $\mathbf{w} - \mathbf{v} \in W$, $\|\mathbf{w} - \mathbf{v}\|^2 \geq 0$ e o mínimo ocorre quando $\mathbf{v} = \mathbf{w}$. Reciprocamente, se $\mathbf{v} = \mathbf{w}$, nada há para ser provado. Se $\mathbf{v} \neq \mathbf{w}$ e $\|\mathbf{u}-\mathbf{w}\|^2 \leq \|\mathbf{u}-\mathbf{v}\|^2$, para todo $\mathbf{v} \in W$, então, pela equação (7.3),

$$2\,\mathrm{Re}\langle \mathbf{u}-\mathbf{w}, \mathbf{w}-\mathbf{v}\rangle + \|\mathbf{v}-\mathbf{w}\|^2 = \|\mathbf{u}-\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{w}\|^2 \geq 0,$$

de modo que $2\,\mathrm{Re}\langle \mathbf{u}-\mathbf{w}, \mathbf{z}\rangle + \|\mathbf{z}\|^2 \geq 0$, para todo $\mathbf{z} = \mathbf{w}-\mathbf{v} \in W$. Vamos escolher $\mathbf{z} = -\langle \mathbf{u}-\mathbf{w}, \mathbf{w}-\mathbf{v}\rangle \|\mathbf{w}-\mathbf{v}\|^{-2}(\mathbf{w}-\mathbf{v})$, para que $\langle \mathbf{u}-\mathbf{w}, \mathbf{z}\rangle$ seja um número real. Assim,

$$-|\langle \mathbf{u}-\mathbf{w}, \mathbf{w}-\mathbf{v}\rangle|^2 \|\mathbf{w}-\mathbf{v}\|^{-2} \geq 0 \Leftrightarrow \langle \mathbf{u}-\mathbf{w}, \mathbf{w}-\mathbf{v}\rangle = 0.$$

Portanto, $(\mathbf{u}-\mathbf{w}) \perp \mathbf{z}$, para todo $\mathbf{z} \in W$. (2) Seja $\mathbf{w}_1 \in W$ outra melhor aproximação de $\mathbf{u}$ por vetores de W. Então $\|\mathbf{u}-\mathbf{w}\|^2 \leq \|\mathbf{u}-\mathbf{w}_1\|^2$ e $\|\mathbf{u}-\mathbf{w}_1\|^2 \leq \|\mathbf{u}-\mathbf{w}\|^2$, de modo que $\|\mathbf{u}-\mathbf{w}\|^2 = \|\mathbf{u}-\mathbf{w}_1\|^2$. Logo, pela equação (7.3),

$$\|\mathbf{u}-\mathbf{w}\|^2 = \|(\mathbf{u}-\mathbf{w}_1)+(\mathbf{w}_1-\mathbf{w})\|^2 = \|\mathbf{u}-\mathbf{w}_1\|^2 + \|\mathbf{w}_1-\mathbf{w}\|^2$$

implica que $\mathbf{w}_1 = \mathbf{w}$. (3) Como $\mathbf{z} = \sum_{i=1}^{n} \langle \mathbf{z}, \mathbf{v}_i\rangle \|\mathbf{v}_i\|^{-2}\mathbf{v}_i$ temos, depois de alguns cálculos, que

$$\langle \mathbf{u}-\mathbf{w}, \mathbf{z}\rangle = \sum_{i=1}^{n} \overline{\langle \mathbf{z}, \mathbf{v}_i\rangle}\langle \mathbf{u}, \mathbf{v}_i\rangle \|\mathbf{v}_i\|^{-2} - \sum_{i=1}^{n} \langle \mathbf{u}, \mathbf{v}_i\rangle \overline{\langle \mathbf{z}, \mathbf{v}_i\rangle} \|\mathbf{v}_i\|^{-2} = 0.$$

Portanto, $\mathbf{w}$ é o ponto de mínimo da função $f : W \to F$ definida como $f(\mathbf{x}) = \|\mathbf{x}-\mathbf{u}\|^2$. Mais explicitamente, se $S = \{\mathbf{x} \in W : \|\mathbf{x}-\mathbf{u}\| \leq \|\mathbf{u}\|\}$, então S é "compacto" e $f : S \to F$ definida como $f(\mathbf{x}) = \|\mathbf{x}-\mathbf{u}\|$ possui o valor mínimo $d(\mathbf{u}, W) = f(\mathbf{w}) = \|\mathbf{u}-\mathbf{w}\|$. Isto é uma generalização da distância de um ponto a um plano em $\mathbf{R}^3$. ■

Vamos finalizar esta seção com a formulação matricial do processo de ortogonalização de Gram-Schmidt. Seja $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ qualquer base de V com um produto interno. Então já vimos que os vetores

$$\mathbf{v}_k = \mathbf{u}_k - \sum_{i=1}^{k-1} \langle \mathbf{u}_k, \mathbf{w}_i\rangle \mathbf{w}_i \ \text{ e } \ \mathbf{w}_k = \|\mathbf{v}_k\|^{-1}\mathbf{v}_k, k = 1, \ldots, n,$$

formavam uma base ortogonal e uma base ortonormal de V. Pondo $t_{jj} = \|\mathbf{v}_j\|$ e $t_{ij} =$

$\langle \mathbf{u}_j, \mathbf{w}_i \rangle = \overline{\langle \mathbf{w}_i, \mathbf{u}_j \rangle}$, com $i < j$, obtemos $\mathbf{v}_j = t_{jj}\mathbf{w}_j$ e

$$\mathbf{u}_1 = t_{11}\mathbf{w}_1, \mathbf{u}_2 = t_{12}\mathbf{w}_1 + t_{22}\mathbf{w}_2, \ldots, \mathbf{u}_n = \sum_{i=1}^{n-1} t_{in}\mathbf{w}_i + t_{nn}\mathbf{w}_n.$$

Assim, considerando as matrizes $\mathbf{A}$, $\mathbf{Q}$ e $\mathbf{R}$ cujas colunas são as coordenadas dos vetores $\mathbf{u}_j, \mathbf{w}_j$ e $\mathbf{t}_j = (t_{1j}, \ldots, t_{nj})$, respectivamente, de modo que $\mathbf{A} = \mathbf{QR}$. Portanto, para cada $\mathbf{A} \in \mathbb{R}^{n\times n}$, cujas colunas são LI, existe uma matriz ortogonal $\mathbf{Q}$ e uma matriz triangular superior $\mathbf{R}$ tal que $\mathbf{A} = \mathbf{QR}$, ou seja, a matriz $\mathbf{A}$ possui uma *fatoração QR* em uma matriz ortogonal $\mathbf{Q}$ e uma matriz triangular superior $\mathbf{R}$. Pelo Exemplo 7.20, as matrizes $\mathbf{A}, \mathbf{Q}$ e $\mathbf{R}$ são:

$$\begin{pmatrix} 3 & -1 & 2 \\ 0 & 1 & 1 \\ -3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix} \text{ e } \begin{pmatrix} 3\sqrt{2} & -\frac{3}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{\sqrt{6}}{2} & \frac{5}{\sqrt{6}} \\ 0 & 0 & \frac{2\sqrt{3}}{3} \end{pmatrix}.$$

Observe que o sistema $\mathbf{AX} = \mathbf{B}$ é equivalente ao sistema $\mathbf{RX} = \mathbf{Q}^t\mathbf{B}$.

Exercícios

1. Seja V um espaço com produto interno. Mostre que se $(x\mathbf{u} + y\mathbf{v}) \perp \mathbf{w}$, então não é verdade, em geral, que $\mathbf{u} \perp \mathbf{w}$ e $\mathbf{v} \perp \mathbf{w}$.

2. Sejam $V = F^2$ e $f(\mathbf{u}, \mathbf{v}) = 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2$. Determine uma base ortonormal de V a partir da base $\alpha = \{(-1, 1), (1, 1)\}$.

3. Seja V um espaço com produto interno e $\dim V = n$. Mostre que V é isomorfo a F^n, como um espaço com produto interno.

4. Sejam $V = \mathbb{R}[x]$ e
$$\langle f(x), g(x) \rangle = \int_{-1}^{1} f(t)g(t)dt$$
Determine uma base ortonormal de V a partir da base canônica de V.

5. Seja $V = \mathbb{R}^3$ com o produto interno usual. Determine uma base ortonormal para o subespaço $W = \{(x, y, z) \in V : x - y + z = 0\}$ de V e a distância $d(\mathbf{u}, W)$, com $\mathbf{u} = (1, 2, 3)$.

7.5 Complementar Ortogonal

Sejam V um espaço com produto interno e β um subconjunto não vazio de V. O *complementar ortogonal* de β em V é o conjunto

$$\beta^{\perp} = \{\mathbf{v} \in V : \langle \mathbf{v}, \mathbf{u} \rangle = 0, \ \forall \ \mathbf{u} \in \beta\} = \bigcap_{\mathbf{u} \in \beta} \{\mathbf{v} \in V : \langle \mathbf{v}, \mathbf{u} \rangle = 0\}.$$

Observe, pelos itens (1), (4) e (5) da Proposição 7.11, que $\beta^{\perp}$ é um subespaço de V se β for um subespaço ou não de V. Note, também, pelos itens (1) e (3) da Proposição 7.11, que $\{\mathbf{0}\}^{\perp} = V$ e $V^{\perp} = \{\mathbf{0}\}$.

Exemplo 7.22 *Seja $W = F[(1,0,1,0),(1,1,0,0)]$ um subespaço de F^4 munido com o produto interno usual. Determine $W^{\perp}$.*

Solução. Pelo exposto, $\mathbf{u} = (x, y, z, t) \in W^{\perp}$ se, e somente se,

$$\langle \mathbf{u}, (1,0,1,0) \rangle = 0 \ \text{ e } \ \langle \mathbf{u}, (1,1,0,0) \rangle = 0,$$

isto é, resolver o sistema $x + z = 0$ e $x + y = 0$. Assim, $x = -z$, $y = z$ e z, t quaisquer. Portanto, $W^{\perp} = F[(-1,1,1,0),(0,0,0,1)]$. Note que $W \cap W^{\perp} = \{(0,0,0,0)\}$ e $F^4 = W \oplus W^{\perp}$. ■

Teorema 7.23 (Teorema da Projeção) *Sejam V um espaço com produto interno e W um subespaço de V, com $\dim W = k$. Então $V = W \oplus W^{\perp}$.*

Prova. Como $\dim W = k$ temos, pelo processo de ortogonalização de Gram-Schmidt, que W contém uma base ortonormal $\beta = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$. Assim, para cada $\mathbf{v} \in V$ existe, pelo item (3) do Teorema 7.21, um único vetor

$$\mathbf{w} = \langle \mathbf{v}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \cdots + \langle \mathbf{v}, \mathbf{u}_k \rangle \mathbf{u}_k \in W,$$

tal que $\mathbf{v} - \mathbf{w} \in W^{\perp}$. Logo, $\mathbf{v} = \mathbf{w} + (\mathbf{v} - \mathbf{w}) \in W + W^{\perp}$. Portanto, $V = W + W^{\perp}$, de modo que $V = W \oplus W^{\perp}$, pois $W \cap W^{\perp} = \{\mathbf{0}\}$. Vale lembrar que a melhor aproximação $\mathbf{w}$ é projeção de $\mathbf{v}$ sobre W ou a expansão de Fourier de $\mathbf{v}$ em relação à base β. ■

É muito importante ressaltar o seguinte: se a dimensão de W, no Teorema 7.23, for infinita o resultado é, em geral, falso. Por exemplo, consideremos o espaço $V = l^2$

do Exemplo 7.16 e $W = \mathbb{R}[\alpha]$. Então

$$W^{\perp} = \{\mathbf{v} \in V : \langle \mathbf{v}, \mathbf{u} \rangle = 0, \ \forall \ \mathbf{u} \in W\} = \{\mathbf{0}\},$$

pois se $\mathbf{v} = (y_n) \in W^{\perp}$, então $y_n = \langle \mathbf{v}, \mathbf{e}_n \rangle = 0$, para todo $n \in \mathbb{N}$, de modo que $\mathbf{v} = \mathbf{0}$. Portanto, $W \oplus W^{\perp} = W \neq V$. Note que α é um conjunto ortonormal maximal (completo), pois não existe $\mathbf{v} = (y_n) \in W^{\perp}$ diferente do vetor nulo.

Teorema 7.24 *Sejam V um espaço com produto interno e U, W subespaços de V. Então as seguintes condições são equivalentes*:

1. $V = U \oplus W$ *e* $U \perp W$ (***Soma Direta Ortogonal***);
2. $V = U \oplus W$ *e* $W = U^{\perp}$;
3. $V = U \oplus W$ *e* $W \subseteq U^{\perp}$.

Prova. $(1 \Rightarrow 2)$ Como $U \perp W$ temos que $W \subseteq U^{\perp}$. Por outro lado, se $\mathbf{v} \in U^{\perp}$, então existe um $\mathbf{u} \in U$ e um $\mathbf{w} \in W$ tal que $\mathbf{v} = \mathbf{u} + \mathbf{w}$, de modo que

$$0 = \langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle$$

implica que $\mathbf{u} = \mathbf{0}$ e $\mathbf{v} = \mathbf{w} \in W$. Assim, $U^{\perp} \subseteq W$. Portanto, $W = U^{\perp}$. A implicação $(2 \Rightarrow 3)$ é clara. $(3 \Rightarrow 1)$ Como $W \subseteq U^{\perp}$ temos que $U \perp W$. ■

Proposição 7.25 *Sejam V um espaço com produto interno, com $\dim V = n$, e W um subespaço de V. Então $V = W \oplus W^{\perp}$ se, e somente se, existir um $E \in \mathcal{L}(V)$ tal que $\operatorname{Im} E = W$, $\ker E = W^{\perp}$ e $E(\mathbf{w}) = \mathbf{w}$, para todo $\mathbf{w} \in W$.*

Prova. Se $V = W \oplus W^{\perp}$, basta definir $E : V \to V$ como $E(\mathbf{w} + \mathbf{u}) = \mathbf{w}$, para todo $\mathbf{w} \in W$ e $\mathbf{u} \in W^{\perp}$. Reciprocamente, cada vetor $\mathbf{v} \in V$ pode ser escrito sob a forma

$$\mathbf{v} = E(\mathbf{v}) + (\mathbf{v} - E(\mathbf{v})),$$

onde $E(\mathbf{v}) \in W$ e $\mathbf{v} - E(\mathbf{v}) \in W^{\perp}$. Assim, $V = W + W^{\perp}$. É fácil verificar que $W \cap W^{\perp} = \{\mathbf{0}\}$. Portanto, $V = W \oplus W^{\perp}$. ■

O operador E definido na Proposição 7.25 chama-se *projeção ortogonal* de V sobre W. Em particular, se $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ for uma base ortonormal de W, então

$$E(\mathbf{v}) = \langle \mathbf{v}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \cdots + \langle \mathbf{v}, \mathbf{u}_k \rangle \mathbf{u}_k,$$

$\|E(\mathbf{v})\| \leq \|\mathbf{v}\|$ e $d(\mathbf{v}, W) = \|\mathbf{v} - E(\mathbf{v})\|$, para todo $\mathbf{v} \in V$. Além disso, $F \in \mathcal{L}(V)$ definido como $F(\mathbf{v}) = \mathbf{v} - F(\mathbf{v})$ é a projeção ortogonal de V sobre $W^{\perp}$. Neste caso, o processo de Gram-Schmidt pode ser descrito geometricamente como segue: sejam V um espaço com produto interno e $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ qualquer lista de vetores LI em V. Então definimos $E_1 = I$ e E_k a projeção ortogonal de V sobre $F[\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}]^{\perp}$, para $k = 2, \ldots, n$. Portanto, os vetores

$$\mathbf{v}_k = E(\mathbf{u}_k), k = 1, \ldots, n,$$

formam uma base ortogonal de $W = F[\alpha]$.

Exemplo 7.26 *Seja $W = \mathbb{R}[(1,1,1,1), (1,-3,4,-2)]$ um subespaço de $V = \mathbb{R}^4$ com o produto interno usual. Determine a melhor aproximação de $\mathbf{v} = (1,3,5,7)$ sobre W e $d(\mathbf{v}, W)$.*

Solução. Como $\langle(1,1,1,1), (1,-3,4,-2)\rangle = 0$ temos que eles formam uma base ortogonal β de W. Então os coeficientes de Fourier $\mathbf{v}$ em relação a β são: $x_1 = 4$ e $x_2 = -\frac{1}{15}$. Assim,

$$P(\mathbf{v}) = x_1(1,1,1,1) + x_2(1,-3,4,-2) = \frac{1}{15}(59, 63, 56, 62).$$

Portanto, $d(\mathbf{v}, W) = \|\mathbf{v} - P(\mathbf{v})\| = \frac{1}{15}\sqrt{4470}$. ■

Exemplo 7.27 *Seja $\mathbb{R}^3$ com o produto interno usual. Determine a solução do sistema*

$$\begin{cases} x + 2y - 2z & = & 1 \\ 2x + y - 2z & = & 6 \\ x + 8y - 6z & = & -7 \end{cases}$$

mais próxima do vetor $\mathbf{0} = (0,0,0) \in \mathbb{R}^3$.

Solução. Pelo Teorema 2.21, basta escalonar a matriz

$$\left(\begin{array}{rrr|rrr} 1 & 2 & 1 & 1 & 0 & 0 \\ 2 & 1 & 8 & 0 & 1 & 0 \\ -2 & -2 & -6 & 0 & 0 & 1 \\ \hline -1 & -6 & 7 & 0 & 0 & 0 \end{array}\right) \to \cdots \to \left(\begin{array}{rrr|rrr} 1 & 0 & 5 & \frac{1}{3} & -\frac{2}{3} & 0 \\ 0 & 1 & -2 & \frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{2}{3} & \frac{2}{3} & 1 \\ \hline 0 & 0 & 0 & \frac{11}{3} & -\frac{4}{3} & 0 \end{array}\right).$$

Portanto, $\mathbf{x} = \frac{1}{3}(11, -4.0) + t(2,2,3)$, para todo $t \in \mathbb{R}$, é a solução geral do sistema.

Seja $W = \mathbb{R}[(2,2,3)]$ o espaço solução do sistema homogêneo. Observe que $W = L^{\perp}$ do espaço linha da matriz dos coeficientes $\mathbf{A}$. Então o nosso problema é encontrar $\mathbf{v} \in W$ que minimize $\|\mathbf{x} - \mathbf{x}_p\|$, ou seja, $d(\mathbf{x}_p, W)$. Uma base ortonormal de W é $\{\mathbf{u} = \frac{1}{\sqrt{17}}(2,2,3)\}$ e $P(\mathbf{x}_p) = \langle \mathbf{x}_p, \mathbf{u}\rangle \mathbf{u} = \frac{14}{51}(2,2,3)$. Logo,

$$\mathbf{x}_0 = \mathbf{x}_p + P(\mathbf{x}_p) = \frac{1}{3}(11,-4,0) + \frac{14}{51}(2,2,3) = \frac{1}{51}(215,-40,42)$$

é a solução mais próxima do vetor $\mathbf{0} = (0,0,0)$. ■

Vamos finalizar esta seção apresentando um método alternativo para obter a matriz da projeção ortogonal se $V = F^n$ for um espaço com produto interno.

Lema 7.28 *Seja $\mathbf{A} \in F^{n\times k}$ cujas colunas são LI. Então $\mathbf{A}^*\mathbf{A}$ é não singular.*

Prova. Pelo Exemplo 4.38, temos que $k = \rho(\mathbf{A}) = \rho(\mathbf{A}^*\mathbf{A})$. Portanto, $\mathbf{A}^*\mathbf{A}$ é não singular. ■

Teorema 7.29 *Sejam W um subespaço de V e $\beta = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ qualquer base de W. Se $\mathbf{A} \in F^{n\times k}$ for a matriz cujas colunas são os vetores $\mathbf{u}_i$, então $\mathbf{P} = \mathbf{A}(\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*$ é a matriz da projeção ortogonal $E \in \mathcal{L}(V)$, a qual chama-se* matriz da projeção ortogonal *para W.*

Prova. Suponhamos que $\mathbf{A} = (\mathbf{u}_1, \ldots, \mathbf{u}_k) \in F^{n\times k}$. Então $\mathbf{AX}$ é um elemento do espaço linha de $\mathbf{A}$, para algum $\mathbf{X} \in F^{k\times 1}$. Assim, para cada $\mathbf{Y} \in F^{n\times 1}$ temos, pelo Lema 7.28, que

$$\langle \mathbf{Y} - \mathbf{AX}, [\mathbf{u}_j]\rangle = 0 \Leftrightarrow \mathbf{A}^*(\mathbf{Y} - \mathbf{AX}) = \mathbf{O} \Leftrightarrow \mathbf{X} = (\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*\mathbf{Y}$$

Portanto, $\mathbf{EY} = \mathbf{AX} = \mathbf{A}(\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*\mathbf{Y}$, de modo que $\mathbf{E} = \mathbf{A}(\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*$. ■

É muito importante ressaltar que a matriz da projeção ortogonal para W não depende da base β, pois se $\mathbf{B} \in F^{n\times k}$ for outra matriz, então existe uma matriz não singular $\mathbf{Q} \in F^{k\times k}$ tal que $\mathbf{A} = \mathbf{BQ}$. Assim, depois de alguns cálculos,

$$\mathbf{E} = \mathbf{A}(\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^* = \mathbf{BQ}((\mathbf{BQ})^*\mathbf{BQ})^{-1}\mathbf{BQ}^* = \mathbf{B}(\mathbf{B}^*\mathbf{B})^{-1}\mathbf{B}^*.$$

Lembre-se que $\mathbf{A}^*\mathbf{A}$ é a matiz de Gram induzida pelo produto interno. Em particular, se escolhermos uma base ortonormal de W, então $\mathbf{B}^*\mathbf{B} = \mathbf{I}_k$ e

$$\mathbf{E} = \mathbf{B}\mathbf{B}^* = \sum_{j=1}^{k} \mathbf{C}_j\mathbf{C}_j^* \text{ e } \rho(\mathbf{E}) = \sum_{j=1}^{k} \rho(\mathbf{C}_j\mathbf{C}_j^*).$$

Exemplo 7.30 *Seja $W = \{(x, y, z) \in \mathbb{R}^3 : x - y + z = 0\}$ um subespaço de $V = \mathbb{R}^3$ com o produto interno usual. Determine a melhor aproximação de $\mathbf{v} = (1, 2, 3)$ sobre W e $d(\mathbf{v}, W)$.*

Solução. É fácil verificar que $\beta = \{(1, 1, 0), (0, 1, 1)\}$ é uma base de W. Então

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}, \mathbf{A}^*\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \text{ e } (\mathbf{A}^*\mathbf{A})^{-1} = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

Assim,

$$\mathbf{P} = \mathbf{A}(\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^* = \frac{1}{3}\begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}.$$

Pondo $\mathbf{Y} = [\mathbf{v}]$, obtemos $\mathbf{P}\mathbf{Y} = \frac{1}{3}(1, 8, 7)^t$. Portanto, $\mathbf{w} = P(\mathbf{v}) = \frac{1}{3}(1, 8, 7)$ e $d(\mathbf{v}, W) = \|\mathbf{v} - \mathbf{w}\| = \frac{2}{3}\sqrt{3}$. ■

Exercícios

1. Sejam $T \in \mathcal{L}(F^3)$ definido como $T(x, y, z) = (z, x - y, -z)$ e $W = \ker T$. Encontre uma base ortonormal de $W^\perp$:

   (a) Em relação ao produto interno usual.

   (b) Em relação ao produto interno $\langle \mathbf{u}, \mathbf{v} \rangle = 2x_1x_2 + y_1y_2 + 4z_1z_2$.

2. Seja $W = F[(1, 1, 0), (0, 1, 1)]$ um subespaço de $V = F^3$ com o produto interno usual. Determine $W^\perp$ e $T \in \mathcal{L}(V)$ tal que $\operatorname{Im} T = W$ e $\ker T = W^\perp$.

3. Sejam $V = F^3$, com $\langle \mathbf{u}, \mathbf{v} \rangle = 2x_1x_2 + 4y_1y_2 + 3z_1z_2$, $T \in \mathcal{L}(V)$ definido como $T(x, y, z) = (x - y, 0, z)$ e $W = \ker T$. Encontre bases ortogonais de W e $W^\perp$. Use estas bases para determinar uma base ortogonal de V.

4. Seja $W = F[(1,0,-1),(0,1,1)]$ um subespaço de $V = F^3$ com o produto interno usual. Determine $W^\perp$ e $T \in \mathcal{L}(V)$ diagonalizável tal que $\operatorname{Im} T = W$ e $\ker T = W^\perp$.

5. Seja $\beta = \{(1,0,1),(1,1,0),(2,1,1)\}$ um subconjunto de $V = F^3$ com o produto interno usual. Determine $\beta^\perp$. Se $\beta = [(1,0,1),(1,1,0),(0,1,1)]$, o que seria $\beta^\perp$? Encontre bases ortogonais de β e $\beta^\perp$.

6. Sejam $\mathbf{u}_1 = (1,0,-2,1), \mathbf{u}_2 = (4,3,0,-1), \mathbf{u}_3 = (0,-3,-8,5) \in F^4$, com o produto interno usual, e $W = F[\mathbf{u}_i]$. Ache bases ortonormais de W e $W^\perp$.

7. Seja $\{\mathbf{u}_1,\dots,\mathbf{u}_k\}$ uma base ortogonal de V, com um produto interno. Mostre que $\{k_1\mathbf{u}_1,\dots,k_n\mathbf{u}_k\}$ é uma base ortogonal de V, para todo $k_i \in F^\times$.

8. Seja $W = F[(1,1,1,1),(1,-1,2,2),(1,2,-3,-4)]$ um subespaço de F^4, com o produto interno usual. Determine a melhor aproximação de $\mathbf{v} = (1,2,-3,4) \in F^4$ sobre W.

9. Seja $V = \mathcal{C}([-1,1],\mathbb{R})$ o espaço das funções contínuas, com
$$\langle f, g\rangle = \int_{-1}^{1} f(t)g(t)dt$$
e $W = \mathbb{R}[1,x,x^2,x^3]$ um subespaço de V. Determine a melhor aproximação de $f(x) = e^x$ sobre W.

10. Sejam $V = F^{2\times 2}$, com $f(\mathbf{A},\mathbf{B}) = \operatorname{tr}(\mathbf{B}^*\mathbf{A})$ e $W = \{\mathbf{A} \in V : \mathbf{A}^* = \mathbf{A}\}$. Determine uma base ortogonal de $W^\perp$.

11. Seja $W = \{f \in V : f(-x) = -f(x), \text{ para todo } x \in [-1,1]\}$ um subespaço de V do Exercício (9). Determine $W^\perp$.

12. Sejam V um espaço com produto interno, $\dim V = n$ e U, W subespaços de V. Mostre que:

    (a) $U^{\perp\perp} = U$ e $U \subseteq W$ se, e somente se, $W^\perp \subseteq U^\perp$.

    (b) $(U+W)^\perp = U^\perp \cap W^\perp$ e $(U\cap W)^\perp = U^\perp + W^\perp$.

    (c) Se $\langle \mathbf{v}, U + U^\perp\rangle = 0$, então $\mathbf{v} = \mathbf{0}$.

13. Sejam V um espaço com produto interno e $\mathbf{u}, \mathbf{v} \in V$ vetores unitários. Mostre, sem usar a desigualdade de Cauchy-Schwarz, que $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq 1$. Use isto, para deduzir uma outra prova da desigualdade de Cauchy-Schwarz.

14. Sejam V um espaço com produto interno, W um subespaço de V e $\mathbf{u} \in V$ satisfaça $\langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle \leq \langle \mathbf{w}, \mathbf{w} \rangle$, para todo $\mathbf{w} \in W$. Mostre que $\mathbf{u} \in W^{\perp}$.

15. Sejam V um espaço com produto interno, com $\dim V = n$, e H um subespaço de V, com $\dim H = n - 1$.

    (a) Mostre que existe um vetor unitário $\mathbf{n} \in V$ tal que $H = \{\mathbf{x} \in V : \langle \mathbf{x}, \mathbf{n} \rangle = 0\}$.

    (b) Defina $\mathbf{w} = \mathbf{x} - 2\langle \mathbf{x}, \mathbf{n} \rangle \mathbf{n}$, para cada $\mathbf{x} \in V$. Mostre que $\mathbf{x} - \mathbf{w} \in H^{\perp}$ e $\frac{1}{2}(\mathbf{x} + \mathbf{w}) \in H$.

    (c) Mostre que $R : V \to V$ definida como $R(\mathbf{x}) = \mathbf{w}$ é linear e $R^2 = I$, ou seja, R é uma reflexão de V em H paralela a $\mathbf{n}$.

16. Sejam V um espaço Euclidiano e $S = \{\mathbf{x} \in V : \|\mathbf{x}\| = \|\mathbf{u}\|\}$, para cada $\mathbf{u} \in V$. Mostre que $\langle \mathbf{u}, \mathbf{w} \rangle = 0$, para algum $\mathbf{w} \in V$, se, e somente se, $S \cap r = \{\mathbf{u}\}$, em que $r = \{\mathbf{u} + t\mathbf{w} : t \in \mathbb{R}\}$.

17. Sejam V um espaço Euclidiano e $S = \{\mathbf{x} \in V : (\mathbf{x} - \mathbf{u}) \perp (\mathbf{x} - \mathbf{v})\}$, para vetores $\mathbf{u}, \mathbf{v} \in V$ distintos. Mostre que S é uma esfera, isto é, $S = \{\mathbf{x} \in V : \|\mathbf{x} - \mathbf{x}_0\| = R\}$, com $\mathbf{x}_0 \in V$ e $R > 0$.

18. Sejam $\mathbf{A} \in F^{2\times 2}$ e $T_{\mathbf{A}} \in \mathcal{L}(V)$, com $V = F^{2\times 1}$ um espaço Euclidiano. Mostre que se $\langle T(\mathbf{X}), \mathbf{X} \rangle = 0$, para todo $\mathbf{X} \in V$, então $\mathbf{A} = \lambda(\mathbf{E}_{21} - \mathbf{E}_{12})$, para algum $\lambda \in F$.

19. Sejam V um espaço com produto interno e $T \in \mathcal{L}(V)$. Mostre que se $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ for uma base ortonormal de V, então $[T]_{\alpha}^{\alpha} = (\langle T(\mathbf{u}_j), \mathbf{u}_i \rangle)$.

20. Sejam V um espaço com produto interno e $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ uma lista em V e $\mathbf{G}_{\alpha} = (\langle \mathbf{u}_i, \mathbf{u}_j \rangle)$. Mostre que α é LD se, e somente se, $\det \mathbf{G}_{\alpha} = 0$.

21. Sejam $V = F^n$ um espaço com produto interno e $T \in \mathcal{L}(V)$ tal que $T(\mathbf{e}_i) = \mathbf{u}_i$, para $i = 1, \ldots, n$, em que $\beta = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ é uma base ortonormal de V. Mostre que $\|T(\mathbf{u}) - T(\mathbf{v})\| = \|\mathbf{u} - \mathbf{v}\|$, para todos $\mathbf{u}, \mathbf{v} \in V$. Conclua que T é um isomorfismo.

22. (**Desigualdade de Bessel**[6]) Sejam V um espaço com produto interno e $\beta = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ um sistema ortonormal de V. Mostre que

$$\sum_{i=1}^{n} |\langle \mathbf{v}, \mathbf{u}_i \rangle|^2 \leq \|\mathbf{v}\|^2, \ \ \forall \ \mathbf{v} \in V.$$

Além disso, se $W = F[\beta]$, então as seguintes condições são equivalentes:

(a) $\mathbf{w} \in W$;

(b) $\sum_{i=1}^{n} |\langle \mathbf{w}, \mathbf{u}_i \rangle|^2 = \|\mathbf{w}\|^2$; (**Identidade de Bessel**)

(c) $\mathbf{w} = \sum_{i=1}^{n} \langle \mathbf{w}, \mathbf{u}_i \rangle \mathbf{u}_i$;

(d) $\langle \mathbf{w}, \mathbf{v} \rangle = \sum_{i=1}^{n} \langle \mathbf{w}, \mathbf{u}_i \rangle \langle \mathbf{u}_i, \mathbf{v} \rangle$, para todo $\mathbf{v} \in V$.

23. (**Identidade de Parseval**[7]) Sejam V um espaço com produto interno e $\beta = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ uma base ortonormal de V. Mostre que

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{u}_1 \rangle \langle \mathbf{w}, \mathbf{u}_1 \rangle + \cdots + \langle \mathbf{v}, \mathbf{u}_n \rangle \langle \mathbf{w}, \mathbf{u}_n \rangle, \ \ \forall \ \mathbf{v}, \mathbf{w} \in V.$$

24. Sejam $V = \mathbb{R}^2$ um espaço Euclidiano e $\theta, \phi \in [0, \pi]$. Mostre que

$$\cos(\theta - \phi) = \cos\theta \cos\phi + \operatorname{sen}\theta \operatorname{sen}\phi.$$

25. Sejam V um espaço Euclidiano, $\mathbf{u} \in V$ e W um subespaço de V, com uma base $\alpha = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Mostre que $d(\mathbf{u}, W)^2 = G(\alpha)^{-1} G(\{\mathbf{u}\} \cup \alpha)$, em que $G(\alpha) = \det(\langle \mathbf{u}_i, \mathbf{u}_j \rangle)$ é o *determinante de Gram*.

26. Sejam V um espaço com produto interno e $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ uma base de V. Mostre que existem únicos $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ tais que $\langle \mathbf{u}_i, \mathbf{v}_j \rangle = \delta_{ij}$. Conclua que as bases ortonormais de V são precisamente as *bases autoduais*.

27. Sejam $a_1, \ldots, a_n \in \mathbb{R}_+^\times$ e $\lambda \in \mathbb{R}$. Mostre que

$$(\sum_{i=1}^{n} \sqrt{a_i})^2 \leq \sum_{i=1}^{n} a_i^{\lambda} \sum_{i=1}^{n} a_i^{1-\lambda}.$$

[6]Friedrich Wilhelm Bessel, 1784-1846, matemático, físico e astrônomo alemão.

[7]Marc-Antoine Parseval des Chênes, 1755-1836, matemático francês.

Conclua que a igualdade ocorre se, e somente se, $\lambda = 2^{-1}$ ou $\lambda \neq 2^{-1}$ e $a_1 = \cdots = a_n$.

28. Sejam V um espaço com produto interno e $\beta = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ uma base ortonormal de V. Mostre que

$$f(S,T) = \sum_{i=1}^{n} \langle S(\mathbf{u}_i), T(\mathbf{u}_1) \rangle, \ \ \forall \ S, T \in \mathcal{L}(V),$$

é um produto interno sobre $\mathcal{L}(V)$.

29. Sejam V um espaço com produto interno e $T \in \mathcal{L}(V)$. Mostre que se $\mathbf{v} \in V$ for qualquer vetor unitário, então

$$\langle T(\mathbf{v}), \mathbf{v} \rangle \langle \mathbf{v}, T(\mathbf{v}) \rangle \leq \langle T(\mathbf{v}), T(\mathbf{v}) \rangle.$$

Conclua que se $\mathbf{A} \in \mathbb{C}^{n \times n}$ e $\mathbf{X} \in \mathbb{C}^{n \times 1}$, com $\|\mathbf{X}\| = 1$, então $|\mathbf{X}^* \mathbf{A} \mathbf{X}| \leq \|\mathbf{A}\mathbf{X}\|$ e a igualdade ocorre se, e somente se, $\mathbf{X}$ for um autovetor de $\mathbf{A}$.

8

Operadores em Espaços com Produto Internos

O objetivo deste capítulo é estudar a estrutura de certos tipos especiais de operadores lineares em espaços com produto internos reais e complexos de dimensão finita. Para definir esses operadores, introduzimos outro tipo de adjunto, diferente do operador adjunto do Capítulo 4. Além disso, vamos classificar as cônicas e as superfícies quaádricas.

8.1 Operador Adjunto

Vamos iniciar, esta seção, tratando os funcionais lineares em um espaço com produto interno e sua relação com o produto interno. Usaremos esse resultado para provar a existência do "adjunto" de um operador linear T sobre V.

Sejam V um espaço com produto interno e $\mathbf{v} \in V$ fixado. Então a função $f_{\mathbf{v}} : V \to F$ definida como

$$f_{\mathbf{v}}(\mathbf{u}) = \langle \mathbf{u}, \mathbf{v} \rangle, \ \ \forall \ \ \mathbf{u} \in V,$$

é um funcional linear (prove isto!).

Teorema 8.1 (Teorema da Representação de Riesz) [1] *Sejam V um espaço com pro-*

[1]Frigyes Riesz, 1880-1956, matemático húngaro.

duto interno e $\dim V = n$*. Então* $f \in V^*$ *se, e somente se, existir um único* $\mathbf{v}_f \in V$ *tal que* $f(\mathbf{u}) = \langle \mathbf{u}, \mathbf{v}_f \rangle$*, para todo* $\mathbf{u} \in V$*.*

Prova. Se $f = 0$, então existe um $\mathbf{v}_f = \mathbf{0} \in V$. Se $f \neq 0$, então $W = \ker f$ é um subespaço de V, com $\dim W = n - 1$. Assim, $V = W \oplus F[\mathbf{v}]$, para algum $\mathbf{v} \in W^\perp$. Logo, pela Proposição 7.25, existe uma projeção ortogonal P de V sobre $F[\mathbf{v}]$. Portanto, $f(\mathbf{u}) = f(P(\mathbf{u}))$, para todo $\mathbf{u} \in V$. Por outro lado, para qualquer vetor não nulo $\mathbf{w} \in W^\perp$, obtemos

$$P(\mathbf{u}) = \frac{\langle \mathbf{u}, \mathbf{w} \rangle}{\|\mathbf{w}\|^2}\mathbf{w} \ \text{ e } \ f(\mathbf{u}) = \frac{\langle \mathbf{u}, \mathbf{w} \rangle}{\|\mathbf{w}\|^2} f(\mathbf{w}),$$

para todo $\mathbf{u} \in V$. Escolha $\mathbf{v}_f = \|\mathbf{v}\|^{-2}\overline{f(\mathbf{v})}\mathbf{v} \in V$. ■

É muito importante ressaltar que o Teorema 8.1 prova que a função $T : V \to V^*$ definida como $T(\mathbf{v}) = f_{\mathbf{v}} = \langle \cdot, \mathbf{v} \rangle$, para todo $\mathbf{v} \in V$ é um isomorfismo conjugado. Em particular, se $V = \mathbb{R}^n$ com o produto interno usual e $f \in V^*$. Então existe um único $\mathbf{v} = (c_1, \ldots, c_n) \in V$ tal que $f(x_1, \ldots, x_n) = c_1x_1 + \cdots + c_nx_n$. Quando $n = 3, \mathbf{u}, \mathbf{v} \in V$ fixados e $f \in V^*$ definido como $f(\mathbf{w}) = \det([\mathbf{u}]\,[\mathbf{v}]\,[\mathbf{w}])$, para todo $\mathbf{w} \in V$. Então existe um único produto vetorial $\mathbf{u} \times \mathbf{v} \in V$ tal que

$$f(\mathbf{w}) = \langle \mathbf{u} \times \mathbf{v}, \mathbf{w} \rangle, \ \ \forall \ \ \mathbf{w} \in V.$$

Exemplo 8.2 *Sejam* $V = \mathbb{P}_2(\mathbb{R})$*, com o produto interno do Exemplo* 7.2*, e* $t = 1$*, determine* $g_t(x) \in V$ *tal que* $\langle f(x), g_t(x) \rangle = f(t)$*, para todo* $f(x) \in V$*.*

Solução. Seja $g_t(x) = a_0 + a_1x + a_2x^2 \in V$. Então

$$\begin{array}{rcl} 1 & = & \langle 1, g_t(x) \rangle = a_0 + \frac{1}{2}a_1 + \frac{1}{3}a_2 \\ t & = & \langle x, g_t(x) \rangle = \frac{1}{2}a_0 + \frac{1}{3}a_1 + \frac{1}{4}a_2 \\ t^2 & = & \langle x^2, g_t(x) \rangle = \frac{1}{3}a_0 + \frac{1}{4}a_1 + \frac{1}{5}a_2. \end{array}$$

Assim, resolvendo o sistema, obtemos $a_0 = 3, a_1 = -24$ e $a_2 = 30$. Portanto, $g_t(x) = 3 - 24x + 30x^2$. ■

É muito importante, de um ponto de vista teórica e didática, o seguinte: se a dimensão de V, no Teorema 8.1, for infinita o resultado é, em geral, falso, confira exemplo a seguir.

Exemplo 8.3 *Seja $V = \mathbb{C}[x]$ munido com o produto interno*

$$\langle f(x), g(x)\rangle = \int_0^1 f(t)\overline{g(t)}dt = \sum_{i=0}^{m}\sum_{j=0}^{n}\frac{a_i\bar{b}_j}{i+j+1},$$

para todos $f(x) = \sum_{i=0}^{m} a_i x^i, g(x) = \sum_{j=0}^{n} b_j x^j \in V$. Cada $c \in \mathbb{C}$ fixado, induz um $L_c \in V^*$ definido como $L_c(f(x)) = f(c)$. Mostre que não existe um $g(x) \in V$ tal que $L_c(f(x)) = \langle f(x), g(x)\rangle$, para todo $f(x) \in V$.

Solução. Suponhamos, por absurdo, que um tal $g(x)$ exista. Então

$$f(c) = \int_0^1 f(t)\overline{g(t)}dt.$$

Pondo $h(x) = x - c \in V$, de modo que $(h \cdot f)(c) = 0$ e

$$0 = \int_0^1 h(t)f(t)\overline{g(t)}dt,$$

para todo $f(x) \in V$. Em particular, para $f(x) = \overline{h(x)}g(x)$, obtemos

$$\int_0^1 |h(t)g(t)|^2 dt = 0.$$

Isto implica que $h(x)g(x) = 0$. Como $h(x) \neq 0$ temos que $g(x) = 0$, o que é uma contradição, pois $L_c \neq 0$. ■

Teorema 8.4 *Sejam V um espaço com produto interno,* $\dim V = n$ *e $T \in \mathcal{L}(V)$. Então existe um único $T^* \in \mathcal{L}(V)$ tal que $\langle T(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, T^*(\mathbf{v})\rangle$, para todos $\mathbf{u}, \mathbf{v} \in V$.*

Prova. Cada $\mathbf{v} \in V$ fixado, induz um $f_{\mathbf{v}} \in V^*$ definida como $f_{\mathbf{v}}(\mathbf{u}) = \langle T(\mathbf{u}), \mathbf{v}\rangle$, para todo $\mathbf{u} \in V$. Assim, pelo Teorema 8.1, existe um um único $\mathbf{w} \in V$ tal que $f_{\mathbf{v}}(\mathbf{u}) = \langle \mathbf{u}, \mathbf{w}\rangle$, para todo $\mathbf{u} \in V$. Vamos definir $T^* : V \to V$ como $T^*(\mathbf{v}) = \mathbf{w}$, de modo que $\langle T(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, T^*(\mathbf{v})\rangle$, para todos $\mathbf{u}, \mathbf{v} \in V$. É claro que T^* está bem definido e é único. Dados $\mathbf{v}, \mathbf{w} \in V$ e $a \in F$, obtemos

$$\begin{aligned}\langle \mathbf{u}, T^*(\mathbf{v} + a\mathbf{w})\rangle &= \langle T(\mathbf{u}), \mathbf{v} + a\mathbf{w}\rangle = \langle T(\mathbf{u}), \mathbf{v}\rangle + \bar{a}\langle T(\mathbf{u}), \mathbf{w}\rangle \\ &= \langle \mathbf{u}, T^*(\mathbf{v})\rangle + \bar{a}\langle \mathbf{u}, T^*(\mathbf{w})\rangle = \langle \mathbf{u}, T^*(\mathbf{v}) + aT^*(\mathbf{w})\rangle,\end{aligned}$$

para todo $\mathbf{u} \in V$. Portanto, $T^*(\mathbf{v} + a\mathbf{w}) = T^*(\mathbf{v}) + aT^*(\mathbf{w})$ e $T^* \in \mathcal{L}(V)$. ■

Sejam V um espaço com produto interno, $\beta = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ qualquer base ortonormal de V e $T \in \mathcal{L}(V)$. Então, pelo Exercício (20) da Seção 7.5,

$$[T]_\beta^\beta = (\langle T(\mathbf{u}_j), \mathbf{u}_i \rangle),$$

pois $\mathbf{u} = \sum_{i=1}^n \langle \mathbf{u}, \mathbf{u}_i \rangle \mathbf{u}_i$, para todo $\mathbf{u} \in V$, e $\operatorname{tr}(T) = \sum_{i=1}^n \langle T(\mathbf{u}_i), \mathbf{u}_i \rangle$..

Proposição 8.5 *Sejam V um espaço com produto interno,* $\dim V = n$ *e* $T \in \mathcal{L}(V)$. *Então* $[T^*] = \mathbf{A}^*$, *em que* $\mathbf{A} = [T]$ *é a matriz de T em relação à qualquer base ortonormal de V.*

Prova. Fica como um exercício. ■

Sejam V um espaço com produto interno e $T \in \mathcal{L}(V)$. Diremos que T possui um *operador adjunto* sobre V se existir um $T^* \in \mathcal{L}(V)$ tal que

$$\langle T(\mathbf{u}), \mathbf{v} \rangle = \langle \mathbf{u}, T^*(\mathbf{v}) \rangle, \ \ \forall \ \ \mathbf{u}, \mathbf{v} \in V.$$

Quando $\dim V = n$ o operador adjunto sempre existe. É muito importante observar que T^* depende somente de T, mas não do produto interno sobre V.

Exemplo 8.6 *Sejam $V = F^2$ com o produto interno usual e $T \in \mathcal{L}(V)$ definido como $T(x, y) = (x + 2y, y)$. Determine T^*.*

Solução. Um modo de fazer isto é calculando $\langle \mathbf{u}, T^*(\mathbf{v}) \rangle$. Neste caso, é melhor usar a Proposição 8.5.

$$[T] = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \text{ e } [T]^* = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}.$$

Portanto, $T^*(x, y) = (x, 2x + y)$. ■

Exemplo 8.7 *Sejam $V = \mathbb{C}[x]$ com o produto interno do Exemplo 8.3, $h(x) \in V$ fixado e $T_h \in \mathcal{L}(V)$ definido como $T_h(f(x)) = h(x)f(x)$. Determine T_h^*.*

Solução. Sejam $h(x) = \sum_{i=0}^n a_i x^i$ e $\bar{h}(x) = \sum_{i=0}^n \bar{a}_i x^i$. Então $\bar{h}(t) = \overline{h(t)}$, para todo $t \in \mathbb{R}$. Assim,

$$\langle T_h(f(x)), g(x) \rangle = \int_0^1 h(t)f(t)\overline{g(t)}dt = \int_0^1 f(t)\overline{\overline{h(t)}g(t)}dt = \langle f(x), T_{\bar{h}}(g(x)) \rangle.$$

Portanto, $T_h^* = T_{\bar{h}}$. ■

Teorema 8.8 *Sejam V um espaço com produto interno e $\dim V = n$. Então a função $\psi : \mathcal{L}(V) \to \mathcal{L}(V)$ definida como $\psi(T) = T^*$ satisfaz as seguintes condições:*

1. $(S+T)^* = S^* + T^*$ *e* $(aT)^* = \bar{a}T^*$.
2. $(TS)^* = S^*T^*$ *e* $T^{**} = (T^*)^* = T$.
3. *Se T for não singular, então T^* é não singular e* $(T^*)^{-1} = (T^{-1})^*$.

Prova. Vamos provar apenas o item (2). Como

$$\langle \mathbf{u}, (TS)^*(\mathbf{v})\rangle = \langle TS(\mathbf{u}), \mathbf{v}\rangle = \langle S(\mathbf{u}), T^*(\mathbf{v})\rangle = \langle \mathbf{u}, S^*T^*(\mathbf{v})\rangle,$$

para todos $\mathbf{u}, \mathbf{v} \in V$, temos que $(TS)^* = S^*T^*$. ■

É muito importante o seguinte. Sejam V um espaço com produto interno e $\dim V = n$. Então qualquer $T \in \mathcal{L}(V)$ pode ser escrito de modo único sob a forma

$$T = \frac{1}{2}(T + T^*) + i\frac{1}{2i}(T - T^*) = T_1 + iT_2.$$

Note que $T_1^* = T_1$ e $T_2^* = T_2$. Um operador $T \in \mathcal{L}(V)$ tal que $T^* = T$ chama-se *autoadjunto* ou *Hermitiano*. Neste caso, T é autoadjunto se, e somente se, $[T]_\alpha^\alpha$ for uma matriz Hermitiana, para qualquer base ortonormal α de V.

Teorema 8.9 *Sejam V um espaço com produto interno, $\dim V = n$ e $T \in \mathcal{L}(V)$.*

1. $\det(T^*) = \overline{\det(T)}$.
2. $\ker T^* = (\operatorname{Im} T)^\perp$ *e* $\operatorname{Im} T^* = (\ker T)^\perp$. *Conclua que* $V = \operatorname{Im} T^* \oplus \ker T$ *e* $V = \operatorname{Im} T \oplus \ker T^*$. *Em particular, a equação $T(\mathbf{u}) = \mathbf{b}$ possui uma solução se, e somente se, $\mathbf{b} \perp \mathbf{v}$, para todo $\mathbf{v} \in \ker T^*$.*
3. *W é um subespaço invariante sob T se, e somente se, $W^\perp$ for um subespaço invariante sob T^*.*
4. *λ é um autovalor de T se, e somente se, $\bar{\lambda}$ for um autovalor de T^*.*
5. *Sejam $\mathbf{u}_1$ e $\mathbf{u}_2$ autovetores de T e T^* associados com autovalores distintos, $\mu \neq \bar{\lambda}$. Então $\mathbf{u}_1 \perp \mathbf{u}_2$.*
6. $\ker T^*T = \ker T$ *e* $\operatorname{Im} T^*T = \operatorname{Im} T^*$. *Conclua que* $\rho(T^*T) = \rho(T^*) = \rho(T)$.

Prova. Vamos provar apenas os itens (2) e (6). Dado $\mathbf{v} \in V$, obtemos

$$\mathbf{v} \in (\operatorname{Im} T)^{\perp} \Leftrightarrow 0 = \langle T(\mathbf{u}), \mathbf{v} \rangle = \langle \mathbf{u}, T^{*}(\mathbf{v}) \rangle \Leftrightarrow \mathbf{v} \in \ker T^{*}.$$

Assim, $\ker T^{*} = (\operatorname{Im} T)^{\perp}$. Para o resto basta substituir T por T^{*} e tomar o perpendicular. (6) É claro $\ker T \subseteq \ker T^{*}T$. Por outro lado, $\mathbf{u} \in \ker T^{*}T$ implica que $0 = \langle \mathbf{u}, \mathbf{0} \rangle = \langle \mathbf{u}, T^{*}T(\mathbf{u}) \rangle = \langle T(\mathbf{u}), T(\mathbf{u}) \rangle = \|T(\mathbf{u})\|^{2}$, de modo que $T(\mathbf{u}) = \mathbf{0}$, isto é, $\mathbf{u} \in \ker T$. ■

Vamos finalizar esta seção apresentando um método alternativo para obter o melhor $\mathbf{X}$ que resolva a equação $\mathbf{AX} = \mathbf{B}$ mesmo que ela não possua solução.

Proposição 8.10 *Sejam V, W espaços com produto internos de dimensões finitas e $T \in \mathcal{L}(V, W)$. Então para cada $\mathbf{b} \in W$, existe um $\mathbf{v}_0 \in V$ tal que*

$$\|T(\mathbf{v}_0) - \mathbf{b}\| \leq \|T(\mathbf{v}) - \mathbf{b}\|, \tag{8.1}$$

para todo $\mathbf{v} \in V$. Conclua que $\mathbf{v}_0 \in V$ é uma solução do problema de minimização se, e somente se, $\mathbf{v}_0$ for uma solução do sistema normal $T^{*}T(\mathbf{v}) = T^{*}\mathbf{b}$.

Prova. Pelo Teorema 7.21, existe um $T(\mathbf{v}_0) \in \operatorname{Im} T$ que satisfaz (8.1). Isto ocorre se, e somente se, $T(\mathbf{v}_0) - \mathbf{b} \perp T(\mathbf{v})$, para todo $\mathbf{v} \in V$. Assim, a solução é caracterizada como $\langle T(\mathbf{v}_0) - \mathbf{b}, T(\mathbf{v}) \rangle = 0$, para todo $\mathbf{v} \in V$, de modo que $\langle T^{*}T(\mathbf{v}_0) - T^{*}(\mathbf{b}), \mathbf{v} \rangle = 0$, para todo $\mathbf{v} \in V$. Portanto, obtemos $\mathbf{v}_0$ resolvendo o sistema $T^{*}T(\mathbf{v}_0) = T^{*}(\mathbf{b})$ e $e = \|T(\mathbf{v}_0) - \mathbf{b}\|$ é o erro. ■

Exemplo 8.11 *Seja $\mathbb{R}^2$ com o produto interno usual. Determine o ajuste linear de mínimos quadrados aos pontos $P_1 = (x_1, y_1), \ldots, P_n = (x_n, y_n) \in \mathbb{R}^2$, não todos sobre uma reta vertical.*

Solução. Seja $y = mx + b$ a equação da reta desejada. Então $\mathbf{AX} = \mathbf{B}$, em que

$$\mathbf{A} = \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix}, \mathbf{X} = \begin{pmatrix} m \\ b \end{pmatrix} \text{ e } \mathbf{B} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Assim, o nosso problema é escolher $b, m \in \mathbb{R}$ que minimize $\|\mathbf{AX} - \mathbf{B}\|^2$. Para isto, pela Proposição 8.10, basta conhecer:

$$\mathbf{A}^*\mathbf{A} = \begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{pmatrix}, \mathbf{A}^*\mathbf{B} = \begin{pmatrix} \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i \end{pmatrix}$$

e

$$(\mathbf{A}^*\mathbf{A})^{-1} = \frac{1}{\det \mathbf{A}^*\mathbf{A}} \begin{pmatrix} n & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{pmatrix}.$$

Note que $\det \mathbf{A}^*\mathbf{A} \neq 0$, pois existem i, j, com $i \neq j$, tais que $x_i \neq x_j$. Portanto, resolvendo a equação $\mathbf{X} = (\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*\mathbf{B}$, obtemos m e b, ou seja,

$$m = -\frac{ce - dn}{\det \mathbf{A}^*\mathbf{A}} \text{ e } b = -\frac{cd - ae}{\det \mathbf{A}^*\mathbf{A}},$$

em que $a = \sum_{i=1}^n x_i^2, c = \sum_{i=1}^n x_i, d = \sum_{i=1}^n x_i y_i$ e $e = \sum_{i=1}^n y_i$. ■

Exercícios

1. Mostre todas as afirmações deixadas nesta seção.

2. Sejam $V = \mathbb{C}^2$ com o produto interno usual e $T \in \mathcal{L}(V)$ definido como $T(x, y) = (x + iy, -2x - y)$. Determine T^*.

3. Sejam V um espaço com produto interno e $T \in V^*$ definido como $T(\mathbf{x}) = \langle \mathbf{x}, \mathbf{v} \rangle$, para cada $\mathbf{v} \in V$ fixado. Determine $T^*(a)$, para todo $a \in F$.

4. Sejam $V = F^n$ com o produto interno usual e $T \in \mathcal{L}(V)$ definido como $T(x_1, \ldots, x_n) = (0, x_1, \ldots, x_{n-1})$. Determine T^*.

5. Sejam V um espaço com produto interno, $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ qualquer base ortonormal em V e $T \in \mathcal{L}(V)$. Mostre que $T^*(\mathbf{v}) = \sum_{i=1}^n \langle \mathbf{v}, T(\mathbf{u}_i) \rangle \mathbf{u}_i$, para todo $\mathbf{v} \in V$.

6. Sejam $V = F^3$ com o produto interno usual e $T \in \mathcal{L}(V)$ definido como $T(x, y, z) = (x + y + z, 3x + y + z, x + 3y + 3z)$. Mostre que a equação $T(\mathbf{x}) = (3, 10, 1)$ não possui solução, mostrando que $(3, 10, 1) \notin (\ker T^*)^\perp$.

7. Sejam $V = \mathbb{C}[x]$ com o produto interno do Exemplo 8.3 e $D \in \mathcal{L}(V)$ definido como $D(f(x)) = f'(x)$. Mostre que D^* não existe.

8. Sejam V um espaço com produto interno, $\dim V = n$ e $T \in \mathcal{L}(V)$.

(a) Mostre que T é injetora se, e somente se, T^* for sobrejetora.

(b) Mostre que T é sobrejetora se, e somente se, T^* for injetora.

9. Sejam V um espaço com produto interno, $\dim V = n$ e W um subespaço de V. Mostre que E é a projeção ortogonal de V sobre W se, e somente se, $E^2 = E^* = E$

10. Sejam V um espaço com produto interno, $\dim V = n$ e $T \in \mathcal{L}(V)$ tal que $T^2 = T$. Mostre que as seguintes condições são equivalentes:

    (a) T é uma projeção ortogonal;

    (b) $T^* = T$ ou $\langle T(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, T(\mathbf{v})\rangle$, para todos $\mathbf{u}, \mathbf{v} \in V$;

    (c) $\|T(\mathbf{u})\| \leq \|\mathbf{u}\|$, para todo $\mathbf{u} \in V$.

11. Sejam V um espaço unitário (sobre $F = \mathbb{C}$) e $T \in \mathcal{L}(V)$. Mostre que T é autoadjunto se, e somente se, $\langle T(\mathbf{v}), \mathbf{v}\rangle \in \mathbb{R}$, para todo $\mathbf{v} \in V$. O que se pode dizer sobre T quando $F = \mathbb{R}$?

12. Sejam V um espaço com produto interno e $T \in \mathcal{L}(V)$ autoadjunto. Mostre que se $\langle T(\mathbf{v}), \mathbf{v}\rangle = 0$, para todo $\mathbf{v} \in V$, então $T = O$.

13. Sejam V um espaço com produto interno e $T \in \mathcal{L}(V)$ autoadjunto.

    (a) Mostre que qualquer autovalor $\lambda \in F$ de T é real.

    (b) Mostre que se $\dim V = n$ e $F = \mathbb{R}$, então T possui um autovalor real.

    (c) Mostre que autovetores de T associados com autovalores distintos são ortogonais.

    (d) Mostre que $V_\lambda^\perp$ é invariante sob T, para todo autovalor $\lambda \in F$ de T.

    (e) Mostre que $p(T)$ é autoadjunto, para todo $p(x) \in \mathbb{R}[x]$.

    (f) Mostre que $\ker T^k = \ker T$, para todo $k \in \mathbb{N}$.

    (g) Mostre que se $\dim V = n$, então o polinômio minimal de T é um produto de fatores lineares distintos (T é diagonalizável).

14. Sejam V um espaço com produto interno e $T_{\mathbf{u},\mathbf{v}} : V \to V$ definido como $T_{\mathbf{u},\mathbf{v}}(\mathbf{x}) = \langle \mathbf{x}, \mathbf{v}\rangle \mathbf{u}$, para todos $\mathbf{u}, \mathbf{v} \in V$.

(a) Mostre que $T_{\mathbf{u},\mathbf{v}} \in \mathcal{L}(V)$, o qual chama-se *produto tensorial* dos vetores $\mathbf{u}$ e $\mathbf{v}$ e denotamos por $\mathbf{u} \otimes \mathbf{v}$. Conclua que $\mathbf{u} \otimes \mathbf{v} = 0$ se, e somente se, $\mathbf{u} = \mathbf{0}$ ou $\mathbf{v} = \mathbf{0}$.

(b) $T_{\mathbf{u},\mathbf{v}} \circ T_{\mathbf{v},\mathbf{w}} = \|\mathbf{v}\|^2 T_{\mathbf{u},\mathbf{w}}$, para todos $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.

(c) $T^*_{\mathbf{u},\mathbf{v}} = T_{\mathbf{v},\mathbf{u}}$, para todos $\mathbf{u}, \mathbf{v} \in V$.

(d) Se $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ for qualquer base ortonormal de V, então o conjunto $\beta = \{\mathbf{u}_i \otimes \mathbf{u}_j : i, j = 1, \ldots, n\}$ é uma base de $\mathcal{L}(V)$.

(e) $T_{\mathbf{u},\mathbf{v}}$ é autoadjunto se, e somente se, $\mathbf{u} = \lambda\mathbf{v}$, para algum $\lambda \in \mathbb{R}$.

(f) Seja $V = \mathbb{C}^n$, com o produto interno usual, $\mathbf{u} = (y_1, \ldots, y_n)$ e $\mathbf{v} = (z_1, \ldots, z_n)$. Qual é a entrada a_{ij} da matriz $\mathbf{A} = [T_{\mathbf{u},\mathbf{v}}]$? Qual é o $\rho(\mathbf{A})$?

15. Sejam V um espaço Euclidiano, com $\dim V = n \geq 3$, $\alpha = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ qualquer base ortonormal de V e $\det(\mathbf{v}_1, \ldots, \mathbf{v}_n) = \det_\alpha(\mathbf{v}_1, \ldots, \mathbf{v}_n)$, para todos $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$.

(a) Mostre que existe um único $\mathbf{w} \in V$ tal que $\det(\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}, \mathbf{x}) = \langle \mathbf{x}, \mathbf{w} \rangle$, para todo $\mathbf{x} \in V$. O vetor $\mathbf{w}$ chama-se *produto cruzado* de $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ e denotamos por $\mathbf{w} = \mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_{n-1}$.

(b) Mostre que $\langle \mathbf{v}_i, \mathbf{w} \rangle = 0$, para $i = 1, \ldots, n-1$.

(c) Mostre que $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ são LD se, e somente se, $\mathbf{w} = \mathbf{0}$.

(d) Sejam $\mathbf{v}_j = \sum_{i=1}^n a_{ij}\mathbf{e}_i, \mathbf{A} = (a_{ij})$ e $\mathbf{x} = \mathbf{e}_i$ no item (a). Mostre que $\mathbf{w} = \sum_{i=1}^n (-1)^{n+i} \Delta_i \cdot \mathbf{e}_i$, com Δ_i o determinante da submatriz de $\mathbf{A}$ obtida eliminando-se a i-ésima linha.

(e) Mostre que se $\beta = \{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ for uma base ortonormal de V, então $\det(\mathbf{f}_1, \ldots, \mathbf{f}_n) \in \{-1, 1\}$. Diremos que β é *positiva* em relação a α se $\det(\mathbf{f}_1, \ldots, \mathbf{f}_n) = 1$.

(f) Mostre que $\{\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}, \mathbf{w}\}$ for uma base ortonormal positiva, para cada sistema ortonormal $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ de V.

16. Sejam V um espaço Euclidiano, com $\dim V = n \geq 3$, $\alpha = (\mathbf{v}_1, \ldots, \mathbf{v}_{n-1})$ uma lista LI em V e $H = \mathbb{R}[\alpha]$.

(a) Mostre que $E \in \mathcal{L}(V)$ definida como $E(\mathbf{x}) = \mathbf{x} - \frac{\langle \mathbf{w}, \mathbf{x} \rangle}{\|\mathbf{w}\|^2}\mathbf{w}$ é a projeção ortogonal de V sobre H e $d(\mathbf{x}, H) = \frac{|\langle \mathbf{w}, \mathbf{x} \rangle|}{\|\mathbf{w}\|}$, com $\mathbf{w} = \mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_{n-1}$.

(b) Mostre que $H = \{\mathbf{x} \in V : \langle \mathbf{x}, \mathbf{w} \rangle = 0\}$

8.2 Operadores Unitários

Sejam $V = F^2$ com o produto interno usual e $T \in \mathcal{L}(V)$ definido como $T(x, y) = (y, -x)$. Então

$$\begin{aligned}\langle (x, y), T^*(z, w)\rangle &= \langle T(x, y), (z, w)\rangle = \langle (y, -x), (z, w)\rangle \\ &= yz - xw = \langle (x, y), (-w, z)\rangle.\end{aligned}$$

Portanto, $T^*(z, w) = (-w, z)$. Observe que $(T^*T)(x, y) = T^*(y, -x) = (x, y)$, de modo que $T^*T = I$ e, de modo análogo, $TT^* = I$. Isto motiva a seguinte definição.

Sejam V um espaço com produto interno e $T \in \mathcal{L}(V)$. Diremos que T é *unitário* se $TT^* = T^*T = I$, isto é, T é não singular, com $T^{-1} = T^*$. Quando $F = \mathbb{R}$, diremos que T é *ortogonal* e $T^* = T^t$.

Teorema 8.12 *Sejam V um espaço com produto interno,* $\dim V = n$ *e* $T \in \mathcal{L}(V)$. *Então as seguintes condição são equivalentes*:

1. *T é unitário*;
2. $\langle T(\mathbf{u}), T(\mathbf{v})\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$, *para todos* $\mathbf{u}, \mathbf{v} \in V$;
3. $\|T(\mathbf{u})\| = \|\mathbf{u}\|$, *para todo* $\mathbf{u} \in V$;
4. *T leva qualquer base ortonormal de V sobre uma base ortonormal de V*;
5. *T leva alguma base ortonormal de V sobre uma base ortonormal de V.*

Prova. ($1 \Leftrightarrow 2$) Basta observar que $\langle T(\mathbf{u}), T(\mathbf{v})\rangle = \langle \mathbf{u}, T^*T(\mathbf{v})\rangle$, para todos $\mathbf{u}, \mathbf{v} \in V$. ($2 \Rightarrow 3$) É claro. ($3 \Rightarrow 4$) Seja $\beta = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ qualquer base ortonormal de V. Então, pela identidade de polarização,

$$\begin{aligned}\langle T(\mathbf{u}_i), T(\mathbf{u}_j)\rangle &= \tfrac{1}{4}\textstyle\sum_{k=1}^{4} i^k\|T(\mathbf{u}_i) + i^kT(\mathbf{u}_j)\|^2 \\ &= \tfrac{1}{4}\textstyle\sum_{k=1}^{4} i^k\|T(\mathbf{u}_i + i^k\mathbf{u}_j)\|^2 \\ &= \tfrac{1}{4}\textstyle\sum_{k=1}^{4} i^k\|\mathbf{u}_i + i^k\mathbf{u}_j\|^2 = \langle \mathbf{u}_i, \mathbf{u}_j\rangle = \delta_{ij}.\end{aligned}$$

Portanto, $T(\beta) = \{T(\mathbf{u}_1), \dots, T(\mathbf{u}_n)\}$ é uma base ortonormal de V. ($4 \Rightarrow 5$) É claro. ($5 \Rightarrow 2$) Seja $\beta = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ uma base ortonormal de V. Então $T(\beta) = \{T(\mathbf{u}_1), \dots, T(\mathbf{u}_n)\}$ é uma base ortonormal de V. Assim, dados $\mathbf{u}, \mathbf{v} \in V$, existem únicos $x_i, y_j \in F$ tais que $\mathbf{u} = \sum_{i=1}^{n} x_i\mathbf{u}_i$ e $\mathbf{v} = \sum_{i=1}^{n} y_i\mathbf{u}_i$. Portanto, depois de alguns cálculos,

$$\langle T(\mathbf{u}), T(\mathbf{v})\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j \langle T(\mathbf{u}_i), T(\mathbf{u}_j)\rangle = \langle \mathbf{u}, \mathbf{v}\rangle,$$

ou seja, $\langle T(\mathbf{u}), T(\mathbf{v})\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$, para todos $\mathbf{u}, \mathbf{v} \in V$. ■

Um operador que satisfaz uma (e, portanto, todas) as condições do Teorema 8.12 chama-se uma *isometria*. Seja $\mathbf{A} = (a_{ij}) \in F^{n\times n}$. Vamos lembrar que $\mathbf{A}$ é uma matriz unitário se $\mathbf{A}\mathbf{A}^* = \mathbf{A}^*\mathbf{A} = \mathbf{I}$, em que $\mathbf{A}^* = (\bar{a}_{ji})$, de modo que $\sum_{k=1}^{n} \bar{a}_{ik}a_{kj} = \delta_{ij}$, ou seja, as colunas de $\mathbf{A}$ formam uma base ortonormal de $F^{n\times 1}$. Portanto, $T_{\mathbf{A}} \in \mathcal{L}(F^{n\times 1})$ é um operador unitário. Quando $F = \mathbb{R}$, a matriz $\mathbf{A}$ é ortogonal. Consequentemente, as matrizes unitárias são os análogos complexos de matrizes ortogonais reais.

Proposição 8.13 *Sejam V um espaço com produto interno, $\dim V = n$ e $T \in \mathcal{L}(V)$. Então T é unitário se, e somente se $\mathbf{A} = [T]_\beta^\beta$ for unitária, para qualquer base ortonormal β de V.*

Prova. Seja $\beta = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ uma base ortonormal de V. Então $T(\mathbf{u}_j) = \sum_{k=1}^{n} a_{kj}\mathbf{u}_k$, de modo que $a_{ij} = \langle T(\mathbf{u}_j), \mathbf{u}_i\rangle$. Assim,

$$\bar{a}_{ji} = \overline{\langle T(\mathbf{u}_i), \mathbf{u}_j\rangle} = \langle \mathbf{u}_j, T(\mathbf{u}_i)\rangle = \langle T^*(\mathbf{u}_j), \mathbf{u}_i\rangle.$$

Portanto, $\mathbf{A}^* = [T^*]_\beta^\beta$. Neste caso, $TT^* = I$ se, e somente se, $\mathbf{A}\mathbf{A}^* = \mathbf{I}$. ■

Exemplo 8.14 *Seja $V = \mathbb{C}^2$ com o produto interno usual. Determine todas as isometrias sobre V.*

Solução. Já sabemos que qualquer $T \in \mathcal{L}(V)$ é representado, em relação à base canônica (ortonormal), pela matriz:

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

Assim, $\mathbf{A}$ é unitária se, e somente se,

$$\begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{pmatrix} = \mathbf{A}^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Como $|\det \mathbf{A}| = 1$ temos que $\det \mathbf{A} = e^{i\theta}$, para algum $\theta \in [0, 2\pi)$. Portanto, $\mathbf{A}$ é unitária se, e somente se,

$$\mathbf{A} = \begin{pmatrix} a & -be^{i\theta} \\ \bar{b} & \bar{a}e^{i\theta} \end{pmatrix} = \begin{pmatrix} a & -b \\ \bar{b} & \bar{a} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix},$$

onde $a, b \in \mathbb{C}$, com $|a|^2 + |b|^2 = 1$. Observe que quando $\det \mathbf{A} = 1$, estas matrizes podem ser identificadas com os pontos de $S^3 = \{\mathbf{x} \in \mathbb{R}^4 : \|\mathbf{x}\| = 1\}$. Quando $V = \mathbb{R}^2$, $\mathbf{A}$ é ortogonal se, e somente se,

$$\mathbf{A} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \text{ ou } \mathbf{A} = \begin{pmatrix} a & b \\ b & -a \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

onde $a, b \in \mathbb{R}$, com $a^2 + b^2 = 1$. Neste caso, já sabemos que $a = \cos\theta$ e $b = \operatorname{sen}\theta$, para algum $\theta \in [0, 2\pi)$. Portanto, se $\det \mathbf{A} = 1$, então $\mathbf{A}$ é uma rotação ou se $\det \mathbf{A} = -1$, então $\mathbf{A}$ é uma reflexão seguida de uma rotação, ou seja, qualquer operador ortogonal sobre V é uma *rotação própria* ou uma *rotação imprópria*. É muito importante observar que $f(x) = x^2 - 2ax + 1$ ($f(x) = x^2 - 1$) é o polinômio característico de $\mathbf{A}$. Neste caso, $\mathbf{X}_1 = (i, 1)^t$ e $\mathbf{X}_2 = (1, i)^t$ são os autovetores de $\mathbf{A}$ associados aos autovalores $a \pm bi = e^{\pm\theta} \in \mathbb{C}$. ■

Teorema 8.15 *Sejam V um espaço com produto interno e $T \in \mathcal{L}(V)$ unitário.*

1. *Qualquer autovalor $\lambda \in F$ de T satisfaz $|\lambda| = 1$, ou seja, estão sobre um círculo unitário.*

2. *Se $T(\mathbf{v}) = \lambda\mathbf{v}$, então $T^*(\mathbf{v}) = \bar{\lambda}\mathbf{v}$.*

3. *Autovetores de T associados com autovalores distintos são ortogonais.*

4. *$V_\lambda^\perp$ é invariante sob T, para todo autovalor $\lambda \in F$ de T.*

5. *$V^\lambda = V_\lambda$, para todo autovalor $\lambda \in F$ de T.*

6. *Se $\dim V = n$, então o polinômio minimal de T é um produto de fatores lineares distintos sobre F, ou seja, T é diagonalizável.*

Prova. Vamos provar apenas os itens $(1), (2), (5)$ e (6): (1) Seja $\lambda \in F$ um autovalor de T. Então existe um $\mathbf{v} \in V$, com $\mathbf{v} \neq \mathbf{0}$, tal que $T(\mathbf{v}) = \lambda\mathbf{v}$. Assim,

$$\lambda\bar{\lambda}\langle \mathbf{v}, \mathbf{v}\rangle = \langle \lambda\mathbf{v}, \lambda\mathbf{v}\rangle = \langle T(\mathbf{v}), T(\mathbf{v})\rangle = \langle \mathbf{v}, T^*T(\mathbf{v})\rangle = \langle \mathbf{v}, \mathbf{v}\rangle$$

implica que $\lambda\bar{\lambda} = 1$. Neste caso, $\lambda = e^{i\theta} = \cos\theta + \operatorname{sen}\theta i$, para algum $\theta \in \mathbb{R}$. (2) Observe, depois de alguns cálculos, que

$$\|(T^* - \bar{\lambda})\mathbf{v}\|^2 = \|\mathbf{v}\|^2 - 2|\lambda|^2\|\mathbf{v}\|^2 + \|\mathbf{v}\|^2 = 0.$$

Portanto, $T^*(\mathbf{v}) = \bar{\lambda}\mathbf{v}$. (5) É claro que $V_\lambda \subseteq V^\lambda$. Por outro lado, se $\mathbf{v} \in V^\lambda$, então $(T - \lambda I)^{2^k}(\mathbf{v}) = \mathbf{0}$, para algum $k \in \mathbb{N}$. Como $T - \lambda I$ e $T^* - \bar{\lambda} I$ comutam temos que

$$\mathbf{0} = (T^* - \bar{\lambda}I)^{2^k}(T - \lambda I)^{2^k}(\mathbf{v}) = [(T^* - \bar{\lambda}I)^{2^{k-1}}(T - \lambda I)^{2^{k-1}}]^2(\mathbf{v}).$$

Assim,

$$0 = \langle [(T^* - \bar{\lambda}I)^{2^{k-1}}(T - \lambda I)^{2^{k-1}}]^2(\mathbf{v}), \mathbf{v}\rangle = \|(T^* - \bar{\lambda}I)^{2^{k-1}}(T - \lambda I)^{2^{k-1}}(\mathbf{v})\|^2,$$

de modo que $\|(T - \lambda I)^{2^{k-1}}(\mathbf{v})\|^2 = 0$ implica que $(T - \lambda I)^{2^{k-1}}(\mathbf{v}) = \mathbf{0}$. Portanto, recursivamente, $(T - \lambda I)(\mathbf{v}) = \mathbf{0}$ e $\mathbf{v} \in V_\lambda$. (6) Se $m(x) = (x - \lambda)^2 q(x)$, então existe um $\mathbf{u} \in V$ tal que $\mathbf{v} = (T - \lambda I)(\mathbf{w}) \neq \mathbf{0}$, com $\mathbf{w} = q(T)(\mathbf{u})$, implica que $(T - \lambda I)(\mathbf{v}) = \mathbf{0}$ e use o item (5). Uma outra prova construtiva é usando indução sobre n. Se $n = 1$, nada há para ser provado. Suponhamos que o resultado seja válido para todo k, com $1 \leq k \leq n-1$ e $n > 1$. Então, pelo Teorema 5.16, T possui um autovalor $\lambda \in F$. Assim, pelo processo de Gram-Schmidt, V_λ possui uma base ortonormal $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$, com $1 \leq m \leq n$. Como $\dim V_\lambda^\perp < n$ e V_λ é invariante sob T temos que $S = T|_{V_\lambda^\perp}$ é unitáro (prove isto!), de modo que $V_\lambda^\perp$ contém uma base ortonormal $\{\mathbf{u}_{m+1}, \ldots, \mathbf{u}_n\}$ formada de autovetores de T. Portanto, $\{\mathbf{u}_1, \ldots, \mathbf{u}_m, \mathbf{u}_{m+1}, \ldots, \mathbf{u}_n\}$ é uma base ortonormal formada de autovetores de T, ou seja, T é diagonalizável. ■

Observe que o Teorema 8.15 ressalta: os operadores unitários desempenham, na álgebra $\mathcal{L}(V)$, o mesmo papel dos números complexos de valor absoluto um em $\mathbb{C}$.

Vamos finalizar esta seção com alguns comentários sobre mudança de coordenadas em um espaço com produto interno. Sejam V um espaço com produto interno, $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ e $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ bases ortonormais de V. Então já vimos que existe uma única matriz não singular $\mathbf{P} \in F^{n\times n}$ tal que

$$[\mathbf{u}]_\beta = \mathbf{P}^{-1}[\mathbf{u}]_\alpha, \tag{8.2}$$

para todo $\mathbf{u} \in V$. Se $P \in \mathcal{L}(V)$ for definido como $P(\mathbf{u}_j) = \mathbf{v}_j$, então P é unitário e $[P]_\alpha^\beta = \mathbf{P}$. Portanto, para qualquer $T \in \mathcal{L}(V)$, existe uma matriz não singular $\mathbf{P} \in$

$F^{n\times n}$ tal que

$$[T]_\beta^\beta = \mathbf{P}^{-1}[T]_\alpha^\alpha\mathbf{P} = \mathbf{P}^*[T]_\alpha^\alpha\mathbf{P}.$$

Isto motiva a seguinte definição. Sejam $\mathbf{A}, \mathbf{B}$. Diremos que $\mathbf{B}$ é *unitariamente equivalente* a $\mathbf{A}$ se existir uma matriz unitária $\mathbf{P} \in F^{n\times n}$ tal que

$$\mathbf{B} = \mathbf{P}^*\mathbf{A}\mathbf{P}.$$

Neste contexto, se $T \in \mathcal{L}(V)$, diremos que T é *unitariamente diagonalizável* se V possuir uma base ortonormal formada de autovetores de V. Quando $F = \mathbb{R}$, diremos que T é *ortogonalmente diagonalizável*. Por simplicidade na exposição, usaremos o termo unitariamente diagonalizável para ambos os casos. Portanto, pelo Teorema 8.15, qualquer operador unitário é unitariamente diagonalizável.

Teorema 8.16 (Teorema de Schur) *Qualquer* $\mathbf{A} \in \mathbb{C}^{n\times n}$ *é unitariamente equivalente a uma matriz triangular superior* $\mathbf{U}$. *Conclua que os autovalores de* $\mathbf{A}$ *são as entradas diagonais de* $\mathbf{U}$.

Prova. Vamos usar indução sobre n. Se $n = 1$ ou $n = 2$, nada há para ser provado. Suponhamos que o resultado seja válido para todo k, com $1 \leq k \leq n-1$ e $n > 2$. Seja $\lambda \in \mathbb{C}$ um autovalor de $\mathbf{A}$ e $\mathbf{X} \in \mathbb{C}^{n\times 1}$ um autovetor associado, com $\|\mathbf{X}\| = 1$. Então estendendo o conjunto $\{\mathbf{X}\}$ para uma base de $\mathbb{C}^{n\times 1}$ e via o Processo de Gram-Schmidt, obtemos uma base ortonormal $\beta = \{\mathbf{X}, \mathbf{Y}_2, \ldots, \mathbf{Y}_n\}$ de $\mathbb{C}^{n\times 1}$. Pondo

$$\mathbf{Q} = \begin{pmatrix} \mathbf{X} & \mathbf{Y}_2 & \cdots & \mathbf{Y}_n \end{pmatrix} \in \mathbb{C}^{n\times n},$$

é fácil verificar que $\mathbf{Q}^*\mathbf{Q} = \mathbf{I}$ e, pelo Teorema 5.12,

$$\mathbf{Q}^*\mathbf{A}\mathbf{Q} = \begin{pmatrix} \lambda & \mathbf{C} \\ \mathbf{O} & \mathbf{B} \end{pmatrix},$$

onde $\mathbf{B} \in \mathbb{C}^{(n-1)\times(n-1)}$. Assim, existe uma matriz unitária $\mathbf{P}_1$ e uma matriz triangular superior $\mathbf{U}_1$ tal que $\mathbf{P}_1^*\mathbf{B}\mathbf{P}_1 = \mathbf{U}_1$. Consideremos $\mathbf{R} = 1 \oplus \mathbf{P}_1$, de modo que $\mathbf{R}^*\mathbf{R} = \mathbf{I}$ e $\mathbf{P} = \mathbf{Q}\mathbf{R}$ é unitária. Logo,

$$\mathbf{P}^*\mathbf{A}\mathbf{P} = \mathbf{R}^*(\mathbf{Q}^*\mathbf{A}\mathbf{Q})\mathbf{R} = \begin{pmatrix} \lambda & \mathbf{C}\mathbf{P}_1 \\ \mathbf{O} & \mathbf{U}_1 \end{pmatrix} = \mathbf{U},$$

ou seja, $\mathbf{A}$ é unitariamente equivalente a $\mathbf{U}$. ■

Exemplo 8.17 *Determine* $\mathbf{P}$ *tal que* $\mathbf{P}^*\mathbf{AP}$ *seja uma matriz triangular superior* $\mathbf{U}$*, com*

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ -6 & 2 & 1 \\ 2 & -3 & 2 \end{pmatrix}.$$

Solução. É fácil verificar que $f(x) = x(x-3)^2$ e $\mathbf{X}_1 = (1,2,2)^t$ é o único autovetor de $\mathbf{A}$ associado ao autovalor 0. Vamos usar um processo direto de construir uma base ortonormal de $V = \mathbb{R}^3$ a partir de um dado vetor. Assim, devemos encontrar $\mathbf{v}_2 = (x_2, y_2, z_2)$ que satisfaça a equação $x + 2y + 2z = 0$, digamos $\mathbf{X}_2 = (2, 1, -2)^t$. Segundo devemos encontrar $\mathbf{v}_3 = (x_3, y_3, z_3)$ que satisfaça as equações $x+2y+2z = 0$ e $2x+y-2z = 0$, a saber, $\mathbf{X}_3 = (2,-2,1)^t$. Portanto, $\beta = \{3^{-1}\mathbf{X}_1, 3^{-1}\mathbf{X}_2, 3^{-1}\mathbf{X}_3\}$ é uma base ortonormal de V. Seja $\mathbf{Q}$ a matriz cujas colunas são os elementos de β. Então,

$$\mathbf{Q}^*\mathbf{AQ} = \begin{pmatrix} 0 & 3 & 0 \\ 0 & 0 & -3 \\ 0 & 3 & 6 \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{C} \\ \mathbf{O} & \mathbf{B} \end{pmatrix}.$$

Note que a matriz $\mathbf{B}$ possui o mesmo autovalor 3 de $\mathbf{Q}^*\mathbf{AQ}$, com autovetor associado $\sqrt{2}\mathbf{Y}_1 = (1,-1)^t$ e, de modo similar, obtemos uma base ortonormal de $\mathbb{R}^2$, $\gamma = \{\mathbf{Y}_1, \mathbf{Y}_2\}$, em que $\sqrt{2}\mathbf{Y}_2 = (1,1)^t$. Assim,

$$\mathbf{P}_1^*\mathbf{BP}_1 = \begin{pmatrix} 3 & -6 \\ 0 & 3 \end{pmatrix} = \mathbf{U}_1 \text{ e } \mathbf{R} = \begin{pmatrix} 1 & \mathbf{O} \\ \mathbf{O} & \mathbf{P}_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Portanto, pondo $\mathbf{P} = \mathbf{QR}$, teremos

$$\mathbf{P}^*\mathbf{AP} = \mathbf{R}^*(\mathbf{Q}^*\mathbf{AQ})\mathbf{R} = \begin{pmatrix} 0 & \frac{3}{\sqrt{2}} & -\frac{3}{\sqrt{2}} \\ 0 & 3 & -6 \\ 0 & 0 & 3 \end{pmatrix} = \mathbf{U},$$

ou seja, $\mathbf{A}$ é unitariamente equivalente a $\mathbf{U}$. ■

Exemplo 8.18 *Seja* $V = \mathbb{R}^3$ *com o produto interno usual. Determine todas as isometrias sobre* V.

Solução. Sejam $T \in \mathcal{L}(V)$ e $\mathbf{A} \in \mathbb{R}^{3\times 3}$ a representação matricial de T em relação à base canônica (ortonormal) de V. Como $\dim V = 3$ temos que $\mathbf{A}$ possui um autovalor real, digamos λ. Assim, existe um $\mathbf{v} \in V$, com $\mathbf{v} \neq \mathbf{0}$, tal que $T(\mathbf{v}) = \lambda\mathbf{v}$ ou $\mathbf{AX} = \lambda\mathbf{X}$, com $\mathbf{X} = [\mathbf{v}]$. Sendo $S = T|_{V_\lambda^\perp}$ ortogonal e $\dim V_\lambda^\perp = 2$ temos, pelo Exemplo 8.14, que $V = V_\lambda \oplus V_\lambda^\perp$ e

$$\mathbf{A} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \cos\theta & -\operatorname{sen}\theta \\ 0 & \operatorname{sen}\theta & \cos\theta \end{pmatrix} \text{ ou } \mathbf{A} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \cos\theta & \operatorname{sen}\theta \\ 0 & \operatorname{sen}\theta & -\cos\theta \end{pmatrix},$$

para algum $\theta \in [0, 2\pi)$. Além disso,

$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \cos\theta & \operatorname{sen}\theta \\ 0 & \operatorname{sen}\theta & -\cos\theta \end{pmatrix} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \cos\theta & -\operatorname{sen}\theta \\ 0 & \operatorname{sen}\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Por outro lado, $\det \mathbf{A} = \pm 1$ implica que $\lambda = 1$ ou $\lambda = -1$. Se $\lambda = 1$ e $\mathbf{A}^t \neq \mathbf{A}$, então T é uma rotação em torno da reta $r = \mathbb{R}[\mathbf{v}]$ ($\dim V_\lambda = 1$) ou se $\lambda = 1$ e $\mathbf{A}^t = \mathbf{A}$, então T é uma reflexão sobre um plano seguida de uma rotação em torno de uma reta normal ao plano ($\dim V_\lambda = 2$). Se $\lambda = -1$, então $\mathbf{A} = (-\mathbf{A})(-\mathbf{I})$, de modo que T é uma reflexão sobre a origem seguida da discussão acima aplicada a $-\mathbf{A}$. É muito importante observar que: $1 + 2\cos\theta = \operatorname{tr}(\mathbf{A})$ e o sinal de $\operatorname{sen}\theta$ é determinado por $\operatorname{sen}\theta = \det \mathbf{B}$, em que $\mathbf{B}$ é a matriz cujas colunas são os vetores $\mathbf{X}, \mathbf{X}_2$ e $\mathbf{AX}_2$, com $\langle \mathbf{X}, \mathbf{X}_2 \rangle = 0$. ■

Exercícios

1. Dar um exemplo de uma matriz unitária que não seja ortogonal e vice-versa.

2. Seja $V = \mathbb{C}$ visto como um espaço vetorial sobre $\mathbb{R}$, munido com o produto interno $\langle z, w \rangle = \operatorname{Re}(z\bar{w})$.

   (a) Mostre que V é isomorfo a $\mathbb{R}^2$, com o produto interno usual.

   (b) Para cada $z \in V$, defina $T_z \in \mathcal{L}(V)$ como $T_z(x) = zx$. Mostre que $T_z^* = T_{\bar{z}}$.

   (c) Para quais z, T_z é autoadjunto?

   (d) Para quais z, T_z é unitário?

   (e) Para quais z, T_z é positivo?

   (f) Qual é o $\det T_z$?

   (g) Determine $[T_z]_\alpha^\alpha$, com $\alpha = \{1, i\}$.

(h) Seja $T \in \mathcal{L}(V)$, encontre condições necessárias e suficientes tal que $T = T_z$, para algum $z \in V$.

(i) Encontre $T \in \mathcal{L}(V)$ unitário tal que $T \neq T_z$, para todo $z \in V$.

3. Seja $\mathbf{Q} \in \mathbb{R}^{3\times 3}$ dada no Exemplo 8.17. Descreva a isometria de $\mathbb{R}^{3\times 1}$ induzida por $\mathbf{Q}$.

4. Descreva a isometria de $\mathbb{R}^{3\times 1}$ induzida por

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

5. Sejam V um espaço com produto interno e $E \in \mathcal{L}(V)$. Mostre que $T = 2E - I$ é uma isometria e $T^2 = I$ se, e somente se, E é uma projeção ortogonal.

6. Sejam V um espaço Euclidiano, $\mathbf{v} \in V$, com $\|\mathbf{v}\| = 1$, e a função $T : V \to V$ definida como $T(\mathbf{x}) = \mathbf{x} + k\langle \mathbf{x}, \mathbf{v}\rangle \mathbf{v}$, para todo $k \in \mathbb{R}$.

   (a) Mostre que T é linear.

   (b) Para quais k, T é simétrico ($T^t = T$)?

   (c) Para quais k, T é ortogonal?

7. Sejam V um espaço com produto interno e $T : V \to V$ uma função sobrejetora tal que $\langle T(\mathbf{u}), T(\mathbf{v})\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$, para todos $\mathbf{u}, \mathbf{v} \in V$. Mostre que T é injetora e linear, ou seja, T é uma isometria.

8. Sejam V um espaço unitário e $J : V \to V$ uma função sobrejetora tal que $J^2 = I$ e $\langle J(\mathbf{u}), J(\mathbf{v})\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$, para todos $\mathbf{u}, \mathbf{v} \in V$, a qual chama-se *conjugação*.

   (a) Dar um exemplo de uma conjugação.

   (b) Mostre que $\langle J(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, J(\mathbf{v})\rangle$, para todos $\mathbf{u}, \mathbf{v} \in V$.

   (c) Mostre que $J(\mathbf{u} + a\mathbf{v}) = J(\mathbf{u}) + \bar{a}J(\mathbf{v})$, para todos $\mathbf{u}, \mathbf{v} \in V$ e $a \in \mathbb{C}$.

9. Seja $\mathbf{A} \in \mathbb{C}^{n\times n}$ tal que $\mathbf{A}^* = \mathbf{A}$. Mostre que $\mathbf{A}$ é unitariamente diagonalizável. Conclua que todos os autovalores de $\mathbf{A}$ são reais.

10. Sejam V um espaço com produto interno, $\dim V = n$, $T \in \mathcal{L}(V)$ uma isometria e W um subespaço invariante sob T. Mostre que $T(W) = W$ e $T(W^\perp) = W^\perp$.

11. Sejam V um espaço Euclidiano e $T : V \to V$ uma função tal que $T(\mathbf{0}) = \mathbf{0}$ e $\|T(\mathbf{u}) - T(\mathbf{v})\| = \|\mathbf{u} - \mathbf{v}\|$, para todos $\mathbf{u}, \mathbf{v} \in V$. Mostre que T é linear, ou seja, T é uma isometria.

12. Sejam V um espaço Euclidiano e $T : V \to V$ uma função. Diremos que T é um *movimento rígido* quando $\|T(\mathbf{u}) - T(\mathbf{v})\| = \|\mathbf{u} - \mathbf{v}\|$, para todos $\mathbf{u}, \mathbf{v} \in V$. Mostre que qualquer movimento rígido é uma isometria seguida de uma translação.

13. Sejam V um espaço Euclidiano, com $\dim V = n$, e $T \in \mathcal{L}(V)$ não singular. Mostre que as seguintes condições são equivalentes:

    (a) $T = kU$, em que U é um operador ortogonal e $k \in \mathbb{R}_+^\times$;

    (b) T preserva ângulo, isto é,
    $$\frac{\langle T(\mathbf{u}), T(\mathbf{v})\rangle}{\|T(\mathbf{u})\|\|T(\mathbf{v})\|} = \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{u}\|\|\mathbf{v}\|},$$
    para todos $\mathbf{u}, \mathbf{v} \in V - \{\mathbf{0}\}$;

    (c) Se $\langle \mathbf{u}, \mathbf{v}\rangle = 0$, então $\langle T(\mathbf{u}), T(\mathbf{v})\rangle = 0$, para todos $\mathbf{u}, \mathbf{v} \in V$;

    (d) Se $\|\mathbf{u}\| = \|\mathbf{v}\|$, então $\|T(\mathbf{u})\| = \|T(\mathbf{v})\|$, para todos $\mathbf{u}, \mathbf{v} \in V$.

14. Sejam V um espaço unitário e $T \in \mathcal{L}(V)$ autoadjunto. Mostre que:

    (a) $\|\mathbf{u} + iT(\mathbf{u})\|^2 = \|\mathbf{u} - iT(\mathbf{u})\|^2 = \|\mathbf{u}\|^2 + \|T(\mathbf{u})\|^2$, para todo $\mathbf{u} \in V$.

    (b) Os operadores $I + iT$ e $I - iT$ são ambos injetores.

    (c) Se $\dim V = n$, então o operador $U = (I - iT)(I + iT)^{-1}$ é unitário e 1 não é autovalor de U, o qual chama-se *transformação de Cayley* de T.

    (d) Se $\dim V = n, U \in \mathcal{L}(V)$ for unitário e 1 não for autovalor de U, então $T = i(I + U)(I - U)^{-1}$ é autoadjunto, o qual chama-se *transformação inversa de Cayley*.

8.3 Operadores Normais

Sejam $V = F^2$ com o produto interno usual e $T \in \mathcal{L}(V)$ definido como $T(x, y) = (2x - 3y, 3x + 2y)$. Então

$$\begin{aligned}
\langle (x, y), T^*(z, w)\rangle &= \langle T(x, y), (z, w)\rangle = \langle (2x - 3y, 3x + 2y), (z, w)\rangle \\
&= 2xz - 3yz + 3xw + 2yw = x(2z + 3w) + y(-3z + 2w) \\
&= \langle (x, y), (2z + 3w, -3z + 2w)\rangle.
\end{aligned}$$

Portanto, $T^*(z,w) = (2z+3w, -3z+2w)$. Note que $(T^*T)(x,y) = (13x, 13y)$ e $(TT^*)(x,y) = (13x, 13y)$. Portanto, $TT^* = T^*T$. Isto motiva a seguinte definição.

Sejam V um espaço com produto interno e $T \in \mathcal{L}(V)$. Diremos que T é *normal* se $TT^* = T^*T$. É claro que qualquer operador autoadjunto (unitário) é normal. Mas, a recíproca é falsa, confira o exemplo acima.

Teorema 8.19 *Sejam V um espaço com produto interno e $T \in \mathcal{L}(V)$. Então as seguintes condição são equivalentes*:

1. *T é normal*;
2. $\langle T(\mathbf{u}), T(\mathbf{v})\rangle = \langle T^*(\mathbf{u}), T^*(\mathbf{v})\rangle$, *para todos* $\mathbf{u}, \mathbf{v} \in V$;
3. $\|T(\mathbf{u})\| = \|T^*(\mathbf{u})\|$, *para todo* $\mathbf{u} \in V$;
4. $\langle (TT^* - T^*T)(\mathbf{u}), \mathbf{u}\rangle$, *para todo* $\mathbf{u} \in V$;
5. *$T(\mathbf{v}) = \lambda\mathbf{v}$ se, e somente se, $T(\mathbf{v}) = \bar{\lambda}\mathbf{v}$, ou seja, T e T^* possuem os mesmos autovetores.*

Prova. $(1 \Rightarrow 2)$ Suponhamos que T seja normal. Então

$$\langle T(\mathbf{u}), T(\mathbf{v})\rangle = \langle \mathbf{u}, T^*T(\mathbf{v})\rangle = \langle \mathbf{u}, TT^*(\mathbf{v})\rangle = \langle T^*(\mathbf{u}), T^*(\mathbf{v})\rangle,$$

para todos $\mathbf{u}, \mathbf{v} \in V$. $(2 \Rightarrow 3)$ É clara. $(3 \Rightarrow 4)$ Observe que

$$\begin{aligned}\langle (TT^* - T^*T)(\mathbf{u}), \mathbf{u}\rangle &= \langle TT^*(\mathbf{u}), \mathbf{u}\rangle - \langle T^*T(\mathbf{u}), \mathbf{u}\rangle \\ &= \langle T^*(\mathbf{u}), T^*(\mathbf{u})\rangle - \langle T(\mathbf{u}), T(\mathbf{u})\rangle = 0,\end{aligned}$$

para todo $\mathbf{u} \in V$. $(4 \Rightarrow 1)$ Seja $S = TT^* - T^*T$. Então $S^* = S$. Se $F = \mathbb{C}$, então, pelo Teorema 7.6, $S = O$. Se $F = \mathbb{R}$, enttão $0 = \langle S(\mathbf{u}+\mathbf{v}), \mathbf{u}+\mathbf{v}\rangle = 2\langle S(\mathbf{u}), \mathbf{v}\rangle$, de modo que $\|S(\mathbf{u})\|^2 = 0$, para todo $\mathbf{u} \in V$. Portanto, $S = O$ e T é normal. $(1 \Leftrightarrow 5)$ Basta notar que $\|(T - \lambda I)(\mathbf{u})\| = \|(T^* - \bar{\lambda} I)(\mathbf{u})\|$. ■

Proposição 8.20 *Sejam V um espaço com produto interno, $\dim V = n$ e $T \in \mathcal{L}(V)$. Então T é normal se, e somente se $\mathbf{A} = [T]_\beta^\beta$ for normal, para qualquer base ortonormal β de V.*

Prova. Seja $\beta = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ uma base ortonormal de V. Então $T(\mathbf{u}_j) = \sum_{k=1}^n a_{kj}\mathbf{u}_k$, de modo que $a_{ij} = \langle T(\mathbf{u}_j), \mathbf{u}_i\rangle$. Assim,

$$\bar{a}_{ji} = \overline{\langle T(\mathbf{u}_i), \mathbf{u}_j\rangle} = \langle \mathbf{u}_j, T(\mathbf{u}_i)\rangle = \langle T^*(\mathbf{u}_j), \mathbf{u}_i\rangle.$$

Portanto, $\mathbf{A}^* = [T^*]_\beta^\beta$. Neste caso, $TT^* = T^*T$ se, e somente se, $\mathbf{AA}^* = \mathbf{A}^*\mathbf{A}$. ■

Seja $\mathbf{A} \in F^{n\times n}$. Diremos que $\mathbf{A}$ é uma *matriz normal* se $\mathbf{AA}^* = \mathbf{A}^*\mathbf{A}$. Portanto, $T_\mathbf{A} \in \mathcal{L}(F^{n\times 1})$ é um operador normal. É muito importante ressaltar que: se $T \in \mathcal{L}(V)$ for normal e β for uma base ortonormal de V, então

$$[T]_\beta^\beta([T]_\beta^\beta)^* = [T]_\beta^\beta[T^*]_\beta^\beta = [TT^*]_\beta^\beta \text{ e } ([T]_\beta^\beta)^*[T]_\beta^\beta = [T^*]_\beta^\beta[T]_\beta^\beta = [T^*T]_\beta^\beta.$$

Portanto, $T \in \mathcal{L}(V)$ é normal se, e somente se $\mathbf{A} = [T]_\beta^\beta$ for normal, para qualquer base ortonormal β de V. Observe que isso não vale para bases que não sejam ortonormais. Lembramos que se U e W são subespaços de V, o par (U, W) *reduz* T se $V = U \oplus W$ e U, W são invariantes sob T.

Exemplo 8.21 *Seja $V = \mathbb{C}^2$ com o produto interno usual. Determine todas as matrizes normais sobre V.*

Solução. Já sabemos que qualquer $T \in \mathcal{L}(V)$ é representado, em relação à base canônica (ortonormal), pela matriz:

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

Vamos supor que $bc \neq 0$. Assim, $\mathbf{A}$ é normal se, e somente se,

$$\begin{pmatrix} a\bar{a} + b\bar{b} & a\bar{c} + b\bar{d} \\ \bar{a}c + \bar{b}d & c\bar{c} + d\bar{d} \end{pmatrix} = \mathbf{AA}^t = \mathbf{A}^t\mathbf{A} = \begin{pmatrix} a\bar{a} + c\bar{c} & \bar{a}c + \bar{b}d \\ a\bar{c} + b\bar{d} & b\bar{b} + d\bar{d} \end{pmatrix}.$$

Logo, $|b| = |c|$ e $(a-d)\bar{c} = (\bar{a} - \bar{d})b$. Portanto, $\mathbf{A}$ é normal se, e somente se, $c = be^{i\theta}$ e $a - d = (\bar{a} - \bar{d})\frac{b}{\bar{c}}e^{i\theta}$, para algum $\theta \in [0, 2\pi)$. Em particular, para a normalidade de $\mathbf{A}$ é necessário que $|b| = |c|$. Pode ser provado que: se $b = |b|e^{i\phi}$, para algum $\phi \in [0, 2\pi)$, então $e^{-\frac{i(\theta+\phi)}{2}}(\mathbf{A} - a\mathbf{I})$ é autoadjunta. Reciprocamente, se $\mathbf{A} - \lambda\mathbf{I}$ for autoadjunta, para algum $\lambda \in \mathbb{C}$, então $\mathbf{A}$ é normal. Quando $F = \mathbb{R}$, $\mathbf{A}$ é normal se, e somente se, $b^2 = c^2$ e $(a-d)(b-c) = 0$. Portanto, $\mathbf{A}$ é normal se, e somente se,

$$\mathbf{A} = \begin{pmatrix} a & b \\ b & d \end{pmatrix} \text{ ou } \mathbf{A} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

onde $a, b, d \in \mathbb{R}$, com $b \neq 0$. Note que $\mathbf{A}^t = \mathbf{A}$ ou $\mathbf{AA}^t = (a^2 + b^2)\mathbf{I}$ ou $\mathbf{A}^t = -\mathbf{A}$, se $a = 0$. ■

Teorema 8.22 *Sejam V um espaço com produto interno,* $\dim V = n$ *e* $T \in \mathcal{L}(V)$ *normal.*

1. *Autovetores de T associados com autovalores distintos são ortogonais.*
2. $\ker T = \ker T^*$ *e* $\operatorname{Im} T = \operatorname{Im} T^*$. *Conclua que* $V = \ker T \oplus \operatorname{Im} T$ *é uma soma direta ortogonal e* $T|_{\operatorname{Im} T}$ *é bijetora.*
3. *Se $F = \mathbb{C}$, então o polinômio minimal de T é um produto de fatores lineares distintos sobre F, ou seja, T é diagonalizável.*
4. *Se $F = \mathbb{C}$, então $T^* = p(T)$, para algum $p(x) \in F[x]$ e vice-versa.*

Prova. (1) Se $T(\mathbf{u}) = \lambda\mathbf{u}$ e $T(\mathbf{v}) = \mu\mathbf{v}$, com $\lambda \neq \mu$, então

$$\lambda\langle \mathbf{u}, \mathbf{v}\rangle = \langle T(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, T^*(\mathbf{v})\rangle = \langle \mathbf{u}, \bar{\mu}\mathbf{v}\rangle = \mu\langle \mathbf{u}, \mathbf{v}\rangle$$

implica que $\langle \mathbf{u}, \mathbf{v}\rangle = 0$. (2) Pelo item (3) do Teorema 8.19, $\ker T = \ker T^*$, de modo que $\operatorname{Im} T = (\ker T^*)^\perp = (\ker T)^\perp = \operatorname{Im} T^*$. (3) Suponhamos, por absurdo, que $m(x) = (x - \lambda)^2 q(x)$. Então existe um $\mathbf{u} \in V$ tal que $\mathbf{v} = (T - \lambda I)q(T)(\mathbf{u}) \neq \mathbf{0}$. Pondo $\mathbf{w} = q(T)(\mathbf{u})$, obtemos $T(\mathbf{w}) = \lambda\mathbf{w} + \mathbf{v}$ e

$$\lambda\langle \mathbf{w}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{v}\rangle = \langle T(\mathbf{w}), \mathbf{v}\rangle = \langle \mathbf{w}, T^*(\mathbf{v})\rangle = \langle \mathbf{w}, \bar{\lambda}\mathbf{v}\rangle = \lambda\langle \mathbf{w}, \mathbf{v}\rangle,$$

de modo que $\|\mathbf{v}\| = 0$ ou $\mathbf{v} = \mathbf{0}$, o que é uma contradição. Uma outra prova: seja $\mathbf{A} \in \mathbb{C}^{n\times n}$ a representação matricial de T em relação a alguma base ortonormal de V. Então, pelo Teorema 8.16, existe uma matriz unitária $\mathbf{P}$ tal que $\mathbf{P}^*\mathbf{A}\mathbf{P} = \mathbf{U}$, com $\mathbf{U} = (u_{ij})$ uma matriz triangular superior. Assim,

$$\mathbf{U}\mathbf{U}^* = \begin{pmatrix} u_{11} & \mathbf{b}^* \\ \mathbf{O} & \mathbf{U}_1 \end{pmatrix}\begin{pmatrix} \bar{u}_{11} & \mathbf{O}^* \\ \mathbf{b} & \mathbf{U}_1^* \end{pmatrix} = \begin{pmatrix} \bar{u}_{11} & \mathbf{O}^* \\ \mathbf{b} & \mathbf{U}_1^* \end{pmatrix}\begin{pmatrix} u_{11} & \mathbf{b}^* \\ \mathbf{O} & \mathbf{U}_1 \end{pmatrix} = \mathbf{U}^*\mathbf{U},$$

ou seja, $\mathbf{U}$ é normal, de modo que $\mathbf{U}_1$ também o é. Logo, efetuando a multiplicação em bloco, obtemos $\mathbf{b} = \mathbf{0}$. Portanto, recursivamente, teremos $u_{ij} = 0$, se $i \neq j$, e $\mathbf{U}$ é diagonal. (4) Pelo item (3), existe uma base ortonormal $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ de V consistindo de autovetores de T. Então $T(\mathbf{v}_i) = \lambda_i\mathbf{v}_i$ e $T^*(\mathbf{v}_i) = \bar{\lambda}_i\mathbf{v}_i$. Por outro lado, pela fórmula de interpolação de Lagrange, existe um único $p(x) \in F[x]$ tal que $p(\lambda_i) = \bar{\lambda}_i$ ou ($q(\bar{\lambda}_i) = \lambda_i$). Assim,

$$p(T)(\mathbf{v}_i) = p(\lambda_i)\mathbf{v}_i = \bar{\lambda}_i\mathbf{v}_i = T^*(\mathbf{v}_i), i = 1, \ldots, n.$$

Portanto, $T^* = p(T)$ ou $(T = q(T^*))$. ■

Proposição 8.23 *Sejam V um espaço unitário, $\dim V = n, W$ um subespaço de V e $T \in \mathcal{L}(V)$ normal.*

1. *W é invariante sob T se, e somente se, W for invariante sob T^*.*
2. *O par $(W, W^{\perp})$ reduz T. Vale a recíproca.*
3. *Se W for invariante sob T, então $S = T|_W$ é normal e $T^*|_W = S^*$.*

Prova. (1) Observe que se W for invariante sob T, então W é invariante sob $p(T)$ para todo $p(x) \in F[x]$, pois se $\mathbf{v} \in W$, então, indutivamente, $T^m(\mathbf{v}) \in W$, para todo $m \in \mathbb{N}$. Assim, o resultado segue do item (4) do Teorema 8.22. (2) Consequência direta do item (1). (3) Dados $\mathbf{u}, \mathbf{v} \in W$,

$$\langle S(\mathbf{u}), \mathbf{v} \rangle = \langle T(\mathbf{u}), \mathbf{v} \rangle = \langle \mathbf{u}, T^*(\mathbf{v}) \rangle = \langle \mathbf{u}, S^*(\mathbf{v}) \rangle,$$

ou $S^* = T^*|_W$ e $SS^* = T^*|_W (T^*|_W)^* = (T^*|_W)^* T^*|_W = S^* S$. Portanto, S é normal. ■

Exemplo 8.24 *Sejam V um espaço com o produto interno do Exercício (17) da Seção 7.1 e $J \in \mathcal{L}(V)$ definido como*

$$J(f)(x) = \int_0^1 xtf(t)dt.$$

1. *Mostre que $J^* = J$.*
2. *Mostre que $f_n(x) = x^n - 2(n+2)^{-1}$, para todo $n \in \mathbb{N}$, são autofunções de J associadas ao autovalor 0. Conclua que existem pelo menos duas autofunções ortogonais associadas a ele.*
3. *Mostre que existe um único autovalor não nulo de J.*

Solução. (1) Basta observar que

$$\langle J(f), g \rangle = \int_0^1 J(f)(x)g(x)dx = \int_0^1 tf(t)dt \int_0^1 xg(x)dx = \langle f, J(g) \rangle,$$

pois $xt = tx$. (2) É fácil verificar que $J(f_n) = 0 \cdot f_n$, para todo $n \in \mathbb{N}$. Além disso, pelo processo de Gram-Schmidt, as autofunções $g_1 = f_1$ e $g_2(x) = x^2 - x + 6^{-1}$ possuem as propriedades desejadas. Portanto, a recíproca do item (1) do Teorema 8.22 não vale em geral. (3) Se $J(f) = \lambda f$, com $\lambda \neq 0$, então

$$\lambda f(x) = x \int_0^1 t f(t) dt$$

implica que f deve ser da forma $f(x) = ax$, para alguma constante a, de modo que $\lambda = 3^{-1}$. Portanto, $f(x) = x$ é a autofunção associada. ■

Teorema 8.25 *Seja $\mathbf{A} \in \mathbb{C}^{n \times n}$ uma matriz normal.*

1. *$\mathbf{A}$ é Hermitiana se, e somente se, todos os seus autovalores são reais.*
2. *$\mathbf{A}$ é anti-Hermitiana se, e somente se, todos os seus autovalores são imaginários puros.*
3. *$\mathbf{A}$ é unitária se, e somente se, todos os seus autovalores são de módulo unitário.*
4. *Se $\mathbf{A} \in \mathbb{R}^{n \times n}$ e $\mathbf{A}^t \neq \mathbf{A}$, então $\mathbf{A}$ possui pelo menos um par de autovalores complexos.*

Prova. Vamos provar apenas os itens $(1), (2)$ e (4): (1) Suponhamos que todos os autovalores de $\mathbf{A}$ sejam reais. Então, pelo Teorema 8.22, existe uma matriz unitária $\mathbf{P}$ tal que $\mathbf{P}^*\mathbf{A}\mathbf{P} = \mathbf{D}$, com $\mathbf{D}$ diagonal e entradas os autovalores de $\mathbf{A}$. Assim,

$$\mathbf{P}^*\mathbf{A}\mathbf{P} = \mathbf{D} = \mathbf{D}^* = (\mathbf{P}^*\mathbf{A}\mathbf{P})^* = \mathbf{P}^*\mathbf{A}^*\mathbf{P}.$$

Logo, $\mathbf{A}^* = \mathbf{A}$. Portanto, $\mathbf{A}$ é Hermitiana. (2) Basta observar que $i\mathbf{A}$ é Hermitiana e usar o item (1). (4) Suponhamos, por absurdo, que todos os autovalores de $\mathbf{A}$ sejam reais. Então, pelo item (1), $\mathbf{A}$ é Hermitiana ($\mathbf{A}^t = \mathbf{A}$), o que é uma contradição. Portanto, $\mathbf{A}$ possui pelo menos um par de autovalores complexos. ■

É muito importante, de um ponto de vista teórico e didático, apresentar uma prova alternativa (para os que possuem alguma experiência com cálculo) de que um operador autoadjunto (simétrico) sobre um espaço Euclidiano de dimensão finita V possui um autovetor. Observe que se W for invariante sob T, então $W^\perp$ é invariante sob T e vice-versa, pois dado $\mathbf{v} \in W^\perp$ e $\mathbf{u} \in W$, obtemos $T(\mathbf{u}) \in W$ e

$$\langle T(\mathbf{v}), \mathbf{u} \rangle = \langle \mathbf{v}, T(\mathbf{u}) \rangle = 0.$$

Em particular, $S = T|_{W^\perp}$ é autoadjunto. A função quociente de Rayleigh[2] $F : V - \{\mathbf{0}\} \to \mathbb{R}$ definida como

$$F(\mathbf{u}) = \frac{\langle \mathbf{u}, T(\mathbf{u})\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}$$

é claramente contínua e homogênea de grau zero, pois é um quociente de funções contínuas e $F(c\mathbf{u}) = F(\mathbf{u})$. Portanto, o problema de maximizar (minimizar) a função F é equivalente a maximizar (minimizar) a função $f : V \to \mathbb{R}$ definida como $f(\mathbf{u}) = \langle \mathbf{u}, T(\mathbf{u})\rangle$ sujeita ao vinculo $\langle \mathbf{u}, \mathbf{u}\rangle = 1$, pois se $F(\mathbf{u}) \leq F(\mathbf{u}_0)$, para todo $\mathbf{u} \in S = \{\mathbf{v} : \|\mathbf{v}\| = 1 \in V\}$, então, dado $\mathbf{w} \in V$, com $\mathbf{w} \neq \mathbf{0}$, de modo que $\mathbf{w} = \|\mathbf{w}\|\mathbf{w}_0$, onde $\mathbf{w}_0 \in S$, e $F(\mathbf{w}) = F(\mathbf{w}_0) \leq F(\mathbf{u}_0)$. Para responder esta questão vamos considerar a matriz $\mathbf{A} = [T]_\alpha^\alpha$, para qualquer base ortonormal α de V. Assim, devemos maximizar a função $f : \mathbb{R}^{n\times 1} \to \mathbb{R}$ definida como $f(\mathbf{X}) = \mathbf{X}^t\mathbf{A}\mathbf{X}$ sujeita ao vinculo $\mathbf{X}^t\mathbf{X} = 1$. Neste caso, vamos considerar a *função de Lagrange*

$$L(\mathbf{X}, \lambda) = \mathbf{X}^t\mathbf{A}\mathbf{X} - \lambda(1 - \mathbf{X}^t\mathbf{X}),$$

em que λ é o *multiplicador de Lagrange*. É bem conhecido que uma condição necessária é dada pela solução do sistema:

$$\left\{\begin{array}{rcl} \frac{\partial L}{\partial \mathbf{X}}(\mathbf{X}_0, \lambda_0) & = & 2\mathbf{X}_0^t\mathbf{A} - 2\lambda_0\mathbf{X}_0^t = \mathbf{O} \\ \frac{\partial L}{\partial \lambda}(\mathbf{X}_0, \lambda_0) & = & 1 - \mathbf{X}_0^t\mathbf{X}_0 = \mathbf{O}. \end{array}\right.$$

Logo, $\mathbf{A}\mathbf{X}_0 = \lambda_0\mathbf{X}_0$ e $\mathbf{X}_0^t\mathbf{X}_0 = 1$, de modo que $\mathbf{X}_0$ é um autovetor de $\mathbf{A}$ associado ao autovalor λ_0. Portanto, concluímos que o máximo de $\mathbf{X}^t\mathbf{A}\mathbf{X}$ sujeito ao vinculo $\mathbf{X}^t\mathbf{X} = 1$ é o autovetor de $\mathbf{A}$ associado ao maior autovalor, de modo que

$$\frac{\langle \mathbf{u}, T(\mathbf{u})\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle} \leq \lambda_0 = \max\{\langle \mathbf{u}, T(\mathbf{u})\rangle : \|\mathbf{u}\| = 1\},$$

para todo $\mathbf{u} \in V$, com $\mathbf{u} \neq \mathbf{0}$. Neste caso, a igualdade ocorre se, e somente se, $T(\mathbf{u}_0) = \lambda_0\mathbf{u}_0$. Como $S = T|_{W^\perp}$, com $W = \mathbb{R}[\mathbf{u}_0]$, é autoadjunto temos, via um argumento indutivo, que existe uma base ortonormal de V consistindo de autovetores $\mathbf{u}_i$ de T associados com os autovalores $\lambda_1 \leq \cdots \leq \lambda_n = \lambda_0$ e $V_{\lambda_i} \perp V_{\lambda_j}$ quando $\lambda_i \neq \lambda_j$. Observe que a função de Lagrange pode ser aplicado no caso geral do vinculo $\mathbf{X}^t\mathbf{P}\mathbf{X} = 1$, com $\mathbf{P}$ qualquer matriz não singular e simétrica.

Teorema 8.26 *Seja* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *normal. Então existe uma matriz ortogonal* $\mathbf{Q}$ *tal que*

[2] John William Strutt (Lord Rayleigh), 1843-1919, físico e matemático inglês.

$\mathbf{Q}^t\mathbf{A}\mathbf{Q} = \mathbf{D}$, *em que*

$$\mathbf{D} = \operatorname{diag}\left(\lambda_1, \ldots, \lambda_k, \begin{pmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{pmatrix}, \ldots, \begin{pmatrix} a_m & -b_m \\ b_m & a_m \end{pmatrix}\right),$$

onde $\lambda_i, a_j, b_j \in \mathbb{R}$ *e* $b_j \neq 0$, *para* $i = 1, \ldots, k$ *e* $j = 1, \ldots, m$.

Prova. Vamos usar indução sobre n. Sejam $V = \mathbb{R}^{n\times 1}$ e $T = T_{\mathbf{A}} \in \mathcal{L}(V)$. Se $n = 1$ ou $n = 2$, então segue do Exemplo 8.21. Suponhamos que o resultado seja válido para todo k, com $1 \leq k < n$ e $n > 2$. Então, pelo Teorema 5.20, V possui um subespaço W invariante sob T, com $\dim W = 1$ e/ou 2. Assim, $\dim W^\perp < n$ e, pela Proposição 8.23, $S = T|_{W^\perp}$ é normal. Logo, $W^\perp$ pode ser escrito como uma soma direta de subespaço invariante sob T, com dimensão 1 e/ou 2. Portanto, segue de $V = W \oplus W^\perp$ e do Exemplo 8.21. ■

Exercícios

1. Mostre todas as afirmações deixadas nesta seção

2. Para qualquer $\mathbf{A} \in F^{n\times n}$, existe uma matriz diagonal $\mathbf{J} = \operatorname{diag}(\pm 1, \ldots, \pm 1)$ tal que $\det(\mathbf{A} + \mathbf{J}) \neq 0$.

3. Seja $\mathbf{A} = \mathbf{P} + i\mathbf{Q} \in \mathbb{C}^{n\times n}$. Mostre que se $\mathbf{A}$ for não singular, então existem $a, b \in \mathbb{R}$ tais que $a\mathbf{P} + b\mathbf{Q}$ seja não singular. Conclua que se $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ forem semelhantes sobre $\mathbb{C}^{n\times n}$, então também o são sobre $\mathbb{R}^{n\times n}$.

4. Sejam $\mathbf{A} \in \mathbb{R}^{n\times n}$ ortogonal e $V = \mathbb{R}^{n\times 1}$. Então V possui um subespaço invariante sob $\mathbf{A}$ de dimensão 1 e/ou 2.

5. Seja $\mathbf{A} \in \mathbb{R}^{n\times n}$ ortogonal. Mostre que existe uma matriz ortogonal $\mathbf{Q}$ tal que $\mathbf{Q}^t\mathbf{A}\mathbf{Q} = \mathbf{D}$, em que

$$\mathbf{D} = \operatorname{diag}\left(1, \ldots, 1, -1, \ldots, -1, \begin{pmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{pmatrix}, \ldots, \begin{pmatrix} a_m & -b_m \\ b_m & a_m \end{pmatrix}\right),$$

   com $a_j = \cos\theta_j$ e $b_j = \operatorname{sen}\theta_j$, onde $\theta_j \in [0, 2\pi) - \{\pi\}$, para $j = 1, \ldots, m$.

6. Sejam V um espaço unitário e $T \in \mathcal{L}(V)$.

   (a) Mostre que T pode ser escrito de modo único sob a forma $T = T_1 + iT_2$, com $T_1^* = T_1$ e $T_2^* = T_2$.

(b) Mostre que T é normal se, e somente se, $T_1T_2 = T_2T_1$.

7. Sejam V um espaço unitário, $\dim V = n$ e $T \in \mathcal{L}(V)$ normal. Mostre que $\ker T^m = \ker T$ e $\operatorname{Im} T^m = \operatorname{Im} T$, para todo $m \in \mathbb{N}$.

8. Sejam V um espaço unitário, com $\dim V = n$ e $T \in \mathcal{L}(V)$. Mostre que se T for normal e possui exatamente k autovalores zero, então $\rho(T) = n - k$.

9. Sejam V um espaço unitário e $T \in \mathcal{L}(V)$ normal.

   (a) Mostre que $p(T)$ é normal, para todo $p(x) \in F[x]$.

   (b) Mostre que $p(T)(\mathbf{u}) = \mathbf{0}$ se, e somente se, $\bar{p}(T^*)(\mathbf{u}) = \mathbf{0}$, para todo $p(x) \in F[x]$ e $\mathbf{u} \in V$.

10. Sejam V um espaço unitário, com $\dim V = n$ e $T \in \mathcal{L}(V)$. Mostre que T é normal se, e somente se, $T^* = TU$, para algum operador unitário $U \in \mathcal{L}(V)$.

11. Sejam $V = \mathbb{R}^{n\times 1}$ e $\mathbf{A} \in \mathbb{R}^{n\times n}$ simétrica, com autovalores $\lambda_1 \leq \cdots \leq \lambda_n$ e base ortonormal de autovetores $\alpha = \{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$.

   (a) Mostre, para cada $\mathbf{X} = (x_1, \ldots, x_n)^t \in V$, que o quociente de Rayleigh é dado por:

   $$F(\mathbf{X}) = \frac{x_1^2\lambda_1 + \cdots + x_n^2\lambda_n}{\|\mathbf{X}\|^2}.$$

   (b) Mostre que se $S = \{\mathbf{X}^t\mathbf{A}\mathbf{X} : \|\mathbf{X}\| = 1\}$, então

   $$\lambda_1 = \min_{\mathbf{X}\neq\mathbf{O}} F(\mathbf{X}) = \min S \ \text{ e } \ \lambda_n = \max_{\mathbf{X}\neq\mathbf{O}} F(\mathbf{X}) = \max S.$$

12. Sejam V um espaço unitário, com $\dim V = n, T \in \mathcal{L}(V)$ e $\lambda_1, \ldots, \lambda_n$ os autovalores de T, não necessariamente distintos. Mostre que

   $$\sum_{i=1}^{n} |\lambda_i|^2 \leq \operatorname{tr}(T^*T).$$

   Conclua que a igualdade ocorre se, e somente se, T for normal.

13. Seja $f(x) = x^3 - bx^2 - (b+3)x - 1 \in \mathbb{R}[x]$. Mostre que $f(x)$ possui todas as raízes reais, para todo $b \in \mathbb{R}$.

8.4 Formas Quadráticas

Formas quadráticas ocorrem em vários contextos, por exemplo, as equações de uma cônica no plano e uma superfície quaádrica no espaço envolvem formas quadráticas.

Vamos retornar ao estudo de um produto interno qualquer sobre $V = \mathbb{R}^{n\times 1}$. Sejam $\mathbf{X} = (x_1, \ldots, x_n)^t$ e $\mathbf{Y} = (y_1, \ldots, y_n)^t$ vetores de V. Então, depois de alguns cálculos,

$$\langle \mathbf{X}, \mathbf{Y} \rangle = \sum_{i=1}^{n} x_i y_i \langle \mathbf{E}_i, \mathbf{E}_j \rangle.$$

Portanto,

$$\langle \mathbf{X}, \mathbf{Y} \rangle = \mathbf{Y}^t \mathbf{A} \mathbf{X}, \tag{8.3}$$

em que $\mathbf{A} = (a_{ij})$ e $a_{ij} = \langle \mathbf{E}_i, \mathbf{E}_j \rangle$. Observe que as condições (3) e (4) da definição de produto interno impõem certas propriedades sobre a matriz do produto interno $\mathbf{A}$. A condição (3) implica que $\mathbf{A}^t = \mathbf{A}$, pois

$$a_{ij} = \langle \mathbf{E}_i, \mathbf{E}_j \rangle = \langle \mathbf{E}_j, \mathbf{E}_i \rangle = a_{ji}, i, j = 1, \ldots, n.$$

Reciprocamente, a condição $\mathbf{A}^t = \mathbf{A}$ garante a simetria do produto interno, pois

$$\langle \mathbf{X}, \mathbf{Y} \rangle = \mathbf{Y}^t \mathbf{A} \mathbf{X} = (\mathbf{Y}^t \mathbf{A} \mathbf{X})^t = \mathbf{X}^t \mathbf{A}^t \mathbf{Y} = \mathbf{X}^t \mathbf{A} \mathbf{Y} = \langle \mathbf{Y}, \mathbf{X} \rangle.$$

Enquanto, a condição (4) implica que

$$\|\mathbf{X}\|^2 = \langle \mathbf{X}, \mathbf{X} \rangle = \mathbf{X}^t \mathbf{A} \mathbf{X} \geq 0, \ \ \forall \ \mathbf{X} \in V,$$

com igualdade se, e somente se, $\mathbf{X} = \mathbf{0}$. Isto motiva a seguinte definição.

Seja $\mathbf{A} \in \mathbb{R}^{n\times n}$. Diremos que $\mathbf{A}$ é *definida positiva* se $\mathbf{X}^t \mathbf{A} \mathbf{X} > 0$, para todo $\mathbf{X} \in V$, com $\mathbf{X} \neq \mathbf{0}$. Diremos que $\mathbf{A}$ é *semidefinida positiva* ou simplesmente *positiva* se $\mathbf{X}^t \mathbf{A} \mathbf{X} \geq 0$, para todo $\mathbf{X} \in V$ e $\mathbf{X}^t \mathbf{A} \mathbf{X} > 0$, para algum $\mathbf{X} \in V$, com $\mathbf{X} \neq \mathbf{0}$. Caso contrário, diremos que $\mathbf{A}$ é *indefinida*, ou seja, existem $\mathbf{X}, \mathbf{Y} \in V$, não nulos e distintos, tais que $\mathbf{X}^t \mathbf{A} \mathbf{X} < 0$ e $\mathbf{Y}^t \mathbf{A} \mathbf{Y} > 0$.

Lema 8.27 *Qualquer produto interno sobre $V = \mathbb{R}^{n\times 1}$ é dado pela forma bilinear* (8.3)*, com $\mathbf{A}$ uma matriz simétrica e definida positiva.*

Prova. Fica como um exercício. ■

Devido a sua importância nesta seção vamos provar novamente uma propriedade fundamental de matrizes simétricas reais: **seus autovalores complexos são realmente reais**. Seja λ um autovalor de $\mathbf{A}$. Então, vendo $\mathbf{A}$ como um elemento de $\mathbb{C}^{n\times n}$, existe um $\mathbf{Z} = \mathbf{X} + i\mathbf{Y} \in \mathbb{C}^{n\times 1}$ não nulo tal que $\mathbf{AZ} = \lambda\mathbf{Z}$. Assim, pondo $\lambda = a + ib$ e desenvolvendo, obtemos $\mathbf{AX} = a\mathbf{X} - b\mathbf{Y}$ e $\mathbf{AY} = b\mathbf{X} + a\mathbf{Y}$. Como $\langle\mathbf{AX}, \mathbf{Y}\rangle = \langle\mathbf{X}, \mathbf{AY}\rangle$ temos que

$$a\langle\mathbf{X}, \mathbf{Y}\rangle - b\|\mathbf{Y}\|^2 = a\langle\mathbf{X}, \mathbf{Y}\rangle + b\|\mathbf{X}\|^2,$$

de modo que $b(\|\mathbf{X}\|^2 + \|\mathbf{Y}\|^2) = 0$ implica que $b = 0$ e λ real, pois $\mathbf{X} \neq \mathbf{O}$ ou $\mathbf{Y} \neq \mathbf{O}$ ou ambos. Neste caso, $\mathbf{X} \in \mathbb{R}^{n\times 1}$ ou $\mathbf{Y} \in \mathbb{R}^{n\times 1}$ é o autovetor de $\mathbf{A}$ associado a $\lambda \in \mathbb{R}$.

Teorema 8.28 (Teorema do Eixo Principal) *Seja* $\mathbf{A} \in \mathbb{R}^{n\times n}$. *Então as seguintes condições são equivalentes*:

1. $\mathbf{A}$ *é simétrica*;
2. $\mathbf{A}$ *possui uma base ortonormal de autovetores*;
3. $\mathbf{A}$ *é ortogonalmente diagonalizável.*

Prova. Fica como um exercício. ■

A base ortonormal de autovetores da matriz simétrica $\mathbf{A}$ dada no item (2) do Teorema 8.28 chama-se *conjunto de eixos principais* de $\mathbf{A}$.

Sejam $V = \mathbb{R}^{n\times 1}$ e $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{n\times n}$ simétrica. A função $q : V \to \mathbb{R}$ definida como

$$q(\mathbf{X}) = \mathbf{X}^t\mathbf{AX} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{1\leq i<j\leq n} a_{ij}x_ix_j$$

chama-se *forma quadrática* sobre V associada a forma bilinear ou a $\mathbf{A}$. Reciprocamente, qualquer forma quadrática $q : V \to \mathbb{R}$ induz uma forma bilinear sobre V via a identidade de polarização (*forma polar* de q):

$$\langle\mathbf{X}, \mathbf{Y}\rangle = \frac{1}{4}(q(\mathbf{X} + \mathbf{Y}) - q(\mathbf{X} - \mathbf{Y})) = \mathbf{Y}^t\mathbf{AX},$$

para todos $\mathbf{X}, \mathbf{Y} \in V$. Portanto, q é definida positiva se, e somente se, $\mathbf{A}$ também o é. Observe que $q(\mathbf{X})$ é um polinômio homogêneo de grau 2, ou seja, $q(t\mathbf{X}) = t^2q(\mathbf{X})$, para todo $t \in \mathbb{R}$. Sejam α e β bases de V. Qual a relação entre $\mathbf{A} = [q]_\alpha$ e $\mathbf{B} = [q]_\beta$?

Para responder esta pergunta, vamos considerar $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ e $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Então $\mathbf{X} = \mathbf{PY}$ ou $\mathbf{Y} = \mathbf{P}^{-1}\mathbf{X}$, para alguma matriz não singular $\mathbf{P}$. Assim,

$$q(\mathbf{X}) = \mathbf{X}^t\mathbf{AX} = \mathbf{PY}^t\mathbf{APY} = \mathbf{Y}^t(\mathbf{P}^t\mathbf{AP})\mathbf{Y}.$$

Portanto, pela unicidade, $\mathbf{B} = \mathbf{P}^t\mathbf{AP}$. Sejam p e q formas quadráticas sobre V associadas as matrizes $\mathbf{A}$ e $\mathbf{B}$. Diremos que elas são *congruentes* se existir uma matriz não singular $\mathbf{P}$ tal que $\mathbf{B} = \mathbf{P}^t\mathbf{AP}$, confira Exemplo 7.3. Neste caso, $\det \mathbf{A}$ chama-se *discriminante* de q e $\det \mathbf{B} = (\det \mathbf{P})^2 \det \mathbf{A}$.

Teorema 8.29 *Sejam $V = \mathbb{R}^{n\times 1}$ com o produto interno usual e $\mathbf{A} \in \mathbb{R}^{n\times n}$ simétrica. Então as seguintes condições são equivalentes*:

1. *$\mathbf{A}$ é semidefinida positiva*;
2. *Cada autovalor de $\mathbf{A}$ é positivo*;
3. *Existe uma matriz simétrica $\mathbf{B} \in \mathbb{R}^{n\times n}$ tal que $\mathbf{A} = \mathbf{B}^2$*;
4. *Existe uma matriz $\mathbf{C} \in \mathbb{R}^{n\times n}$ tal que $\mathbf{A} = \mathbf{C}^t\mathbf{C}$.*

Prova. $(1 \Rightarrow 2)$ Suponhamos que $\mathbf{A}$ seja semidefinida e $\lambda \in \mathbb{R}$ um autovalor de $\mathbf{A}$ com autovetor $\mathbf{X} \in V$. Então $\lambda\|\mathbf{X}\|^2 = \langle \mathbf{AX}, \mathbf{X}\rangle = \mathbf{X}^t\mathbf{AX} \geq 0$. Portanto, $\lambda \geq 0$. $(2 \Rightarrow 3)$ Suponhamos que $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ sejam os autovalores de $\mathbf{A}$, não necessariamente distintos, e $\lambda_i \geq 0$, para $i = 1, \ldots, n$. Então existem $\mu_i \in \mathbb{R}$ tais que $\mu_i^2 = \lambda_i$, para $i = 1, \ldots, n$. Por outro lado, pelo Teorema 8.28, existe uma matriz ortogonal $\mathbf{Q}$ tal que $\mathbf{Q}^t\mathbf{AQ} = \mathbf{D} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Pondo $\mathbf{E} = \operatorname{diag}(\mu_1, \ldots, \mu_n)$ e $\mathbf{B} = \mathbf{QEQ}^t$, obtemos $\mathbf{B}^t = \mathbf{B}$ e

$$\mathbf{B}^2 = (\mathbf{QEQ}^t)^2 = \mathbf{QE}^2\mathbf{Q}^t = \mathbf{QDQ}^t = \mathbf{A}.$$

$(3 \Rightarrow 4)$ Ponha $\mathbf{C} = \sqrt{\mathbf{B}}$. $(4 \Rightarrow 1)$ Suponhamos que $\mathbf{A} = \mathbf{C}^t\mathbf{C}$, para alguma matriz $\mathbf{B} \in \mathbb{R}^{n\times n}$. Então

$$\mathbf{X}^t\mathbf{AX} = \langle \mathbf{AX}, \mathbf{X}\rangle = \langle \mathbf{C}^t\mathbf{CX}, \mathbf{X}\rangle = \langle \mathbf{CX}, \mathbf{CX}\rangle = \|\mathbf{CX}\|^2 \geq 0,$$

para todo $\mathbf{X} \in V$. Observe que $\mathbf{A}$ é definida positiva se, e somente se, $\mathbf{B}$ for não singular. ■

Vamos interpretar os Teoremas 8.28 e 8.29 em termos de formas quadráticas. Para qualquer forma quadrática $q : V \to \mathbb{R}$, existe uma base ortonormal de V consistindo de autovetores $\mathbf{X}_1, \ldots, \mathbf{X}_n$ de $\mathbf{A}$, ou seja, $\mathbf{A}\mathbf{X}_i = \lambda_i \mathbf{X}_i$. Então existe uma matriz ortogonal $\mathbf{Q}$, cujas colunas são os vetores $\mathbf{X}_i$, tal que $\mathbf{Q}^t\mathbf{A}\mathbf{Q} = \mathbf{D}$, com $\mathbf{D} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, ou seja, $\mathbf{A}$ é congruente a $\mathbf{D}$. Assim,

$$q(\mathbf{X}) = \mathbf{X}^t\mathbf{A}\mathbf{X} = \mathbf{X}^t\mathbf{Q}\mathbf{D}\mathbf{Q}^t\mathbf{X} = \mathbf{Y}^t\mathbf{D}\mathbf{Y} = \sum_{i=1}^{n} \lambda_i y_i^2, \tag{8.4}$$

em que $\mathbf{Y} = \mathbf{Q}^t\mathbf{X} = \mathbf{Q}^{-1}\mathbf{X}$ é a mudança de coordenadas (novas variáveis) e $\mathbf{Q}$ é a matriz de transição da base α para a base canônica de V. É muito importante, de um ponto de vista teórico e didático, fazer mais algumas observações sobre a equação (8.4). Podemos supor que $\lambda_i \neq 0$, para $i = 1, \ldots, k \leq n$. Caso contrário, eliminamos os termos $\lambda_i y_i^2$ e, se necessário, reordenamos. Suponhamos que $\lambda_i > 0$, se $1 \leq i \leq r$, e $\lambda_j < 0$, se $r + 1 \leq j \leq s$. Então

$$q(\mathbf{X}) = \sum_{i=1}^{r} (\sqrt{\lambda_i} y_i)^2 - \sum_{j=r+1}^{s} (\sqrt{-\lambda_j} y_j)^2 = \sum_{i=1}^{r} z_i^2 - \sum_{j=r+1}^{s} z_j^2,$$

com $r + s = k \leq n$. Como os vetores $\mathbf{Y}_1, \ldots, \mathbf{Y}_k$ são LI temos que os vetores

$$\mathbf{Z}_i = \sqrt{\lambda_i}\mathbf{Y}_i \ \text{ e } \ \mathbf{Z}_j = \sqrt{-\lambda_{j+r}}\mathbf{Y}_{j+r}$$

também o são. Portanto, $\beta = \{\mathbf{Z}_1, \ldots, \mathbf{Z}_k, \mathbf{Y}_{k+1}, \ldots, \mathbf{Y}_n\}$ é uma base ortogonal de V. Neste caso, $\mathbf{A}$ é congruente a $\mathbf{I}_r \oplus (-\mathbf{I}_s) \oplus \mathbf{O_t}$, com $r + s + t = n$. Pode ser provado que o par (r, s) é unicamente determinado por q, o qual chama-se *assinatura de Sylvester* de q, r chama-se o *índice* de q e $k = r + s$ chama-se o *posto* de q. Portanto, podemos concluir que:

1. q é não degenerada se, e somente se, $k = n$. Enquanto, q é degenerada se, e somente se, $k < n$.

2. q é definida positiva se, e somente se, $s = 0$ e $r = k = n$.

3. q é semidefinida positiva se, e somente se, $r = k < n$ e $m_a(0) = n - k$.

4. q é indefinida se, e somente se, $0 < r < k \leq n$.

5. p e q são congruentes se, e somente se, elas possuem o mesmo posto e índice.

Exemplo 8.30 *Sejam $V = \mathbb{R}^3$ e $q : V \to \mathbb{R}$ definida como*

$$q(x, y, z) = 5x^2 + 5y^2 + 5z^2 - 2xy - 2xz - 2yz$$

uma forma quadrática sobre V. Determine a assinatura de Sylvester de q.

Solução. Neste exemplo vamos apresentar um procedimento geral (algoritmo) para resolver este tipo de problema. A forma matricial de q é: $q(\mathbf{X}) = \mathbf{X}^t\mathbf{A}\mathbf{X}$, onde $\mathbf{X} = [\mathbf{x}] = (x, y, z)^t \in \mathbb{R}^{3\times 1}$ e

$$\mathbf{A} = \begin{pmatrix} 5 & -1 & -1 \\ -1 & 5 & -1 \\ -1 & -1 & 5 \end{pmatrix}.$$

Note que as entradas não nulas de $\mathbf{A}$ correspondem aos termos cruzados de q. É fácil verificar que $f(x) = (x - 3)(x - 6)^2$ é o polinômio característico de $\mathbf{A}$ e $\mathbf{X}_1 = (1, -1, 0)^t, \mathbf{X}_2 = (1, 0, -1)^t, \mathbf{X}_3 = (1, 1, 1)^t$ autovetores associados aos autovalores 6 e 3. Aplicando o processo de Gram-Schmidt em $\mathbf{X}_1$ e $\mathbf{X}_2$, obtemos uma base ortonormal

$$\mathbf{Y}_1 = \frac{1}{\sqrt{2}}(1, -1, 0)^t, \mathbf{Y}_2 = \frac{1}{\sqrt{6}}(1, 1, -2)^t, \mathbf{Y}_3 = \frac{1}{\sqrt{3}}(1, 1, 1)^t$$

de V. Seja $\mathbf{Q}$ a matriz cujas colunas são os vetores $\mathbf{Y}_1, \mathbf{Y}_2$ e $\mathbf{Y}_3$. Então $\mathbf{Q}^t = \mathbf{Q}^{-1}$ e $\mathbf{Q}^t\mathbf{A}\mathbf{Q} = \operatorname{diag}(6, 6, 3)$ ou $\mathbf{A} = \mathbf{Q}\operatorname{diag}(6, 6, 3)\mathbf{Q}^t$. Pondo $\mathbf{Y} = \mathbf{Q}^t\mathbf{X} = (u, v, w)^t$, obtemos

$$q(\mathbf{X}) = 6u^2 + 6v^2 + 3w^2 = (\sqrt{6}u)^2 + (\sqrt{6}v)^2 + (\sqrt{3}w)^2.$$

Portanto, $(3, 0)$ é a assinatura de Sylvester de q e q é definida positiva. ■

Seja V um espaço Euclidiano, com $\dim V = n$. Se f for um produto interno sobre V, então, pelo Teorema 8.1, existe um único $T \in \mathcal{L}(V)$ tal que $T^t = T$ e

$$f(\mathbf{u}, \mathbf{v}) = \langle T(\mathbf{u}), \mathbf{v}\rangle,$$

para todos $\mathbf{u}, \mathbf{v} \in V$. Isto motiva a seguinte definição:

Sejam V um espaço Euclidiano, com $\dim V = n$ e $T \in \mathcal{L}(V)$. Diremos que T é *definido positivo* se $\langle T(\mathbf{u}), \mathbf{u}\rangle > 0$, para todo $\mathbf{u} \in V$, com $\mathbf{u} \neq \mathbf{0}$. Já sabemos que $T^t = T$ se, e somente se, $\mathbf{A} = [T]_\alpha^\alpha$ é simétrica, para qualquer base ortonormal de V. Portanto, T é definido positivo se, e somente se, $\mathbf{A}$ também o é, e assim por diante.

Seja $V = \mathbb{R}^n$ com o produto interno usual. Uma função $f : V \to \mathbb{R}$ definida como

$$f(x_1,\dots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j+\sum_{k=1}^{n}b_kx_k+c, \tag{8.5}$$

com pelo menos um $a_{ij}\neq 0$, chama-se *função quadrática* sobre V nas variáveis não homogêneas $x_1,\dots,x_n$. A função definida pela equação (8.5) pode ser escrita sob a forma matricial

$$f(\mathbf{X})=\mathbf{X}^t\mathbf{A}\mathbf{X}+\mathbf{B}^t\mathbf{X}+c=q(\mathbf{X})+\mathbf{B}^t\mathbf{X}+c,$$

com q a forma quadrática associada a f,

$$\mathbf{A}=(a_{ij})\in\mathbb{R}^{n\times n}\ \text{ e }\ \mathbf{B}=(b_1,\dots,b_n)^t\in\mathbb{R}^{n\times 1}.$$

Observe que podemos supor, sem perda de generalidade, que $\mathbf{A}$ seja uma matriz simétrica, pois a função quadrática f permanece inalterada quando substituímos $\mathbf{A}$ pela matriz $\mathbf{B}=\frac{1}{2}(\mathbf{A}+\mathbf{A}^t)$, que é uma matriz simétrica, pois se $i\neq j$, então $a_{ij}+a_{ji}$ é o coeficiente de x_ix_j. Neste caso, cada função quadrática f sobre V determina um *lugar geométrico* (conjunto) S_f de todos os vetores $\mathbf{x}=(x_1,\dots,x_n)\in V$ tais que $f(\mathbf{x})=0$, ou seja, $S_f=\{\mathbf{x}\in V: f(\mathbf{x})=0\}$. Se $n=2$, diremos que S_f é uma *seção cônica* ou simplesmente uma *cônica*. Se $n=3$, diremos que S_f é uma *superfície quádrica*. Finalmente, se $\mathbf{B}=\mathbf{O}$, diremos que S_f é uma *quádrica central*. Em particular, o lugar geométrico de $\{(x,y,z)\in V: ax^2+by^2-cz^2=0\}$ representa um *cone* "assintótico", pois ele está relacionado com as quaádricas centrais (superfícies quádricas) $ax^2+by^2-cz^2\pm 1=0$.

Estamos pronto para completar o procedimento de classificar o lugar geométrico S_f determinado pela função quadrática f. Como $\mathbf{A}^t=\mathbf{A}$ temos, pelo Teorema 8.28, que existe uma matriz ortogonal $\mathbf{Q}$ tal que $\mathbf{Q}^t\mathbf{A}\mathbf{Q}=\mathbf{D}=\operatorname{diag}(\lambda_1,\dots,\lambda_n)$. Pondo $\mathbf{Y}=\mathbf{Q}^{-1}\mathbf{X}$ e $\mathbf{D}=\mathbf{Q}^{-1}\mathbf{B}$, obtemos

$$\mathbf{Y}^t\mathbf{D}\mathbf{Y}+\mathbf{D}^t\mathbf{Y}+c=0\Leftrightarrow\sum_{i=1}^{n}\lambda_iy_i^2+\sum_{i=1}^{n}d_iy_i+c=0.$$

Assim, reordenando, se necessário, podemos supor que $\lambda_i\neq 0$, se $i=1,\dots,k$, e $\lambda_j=0$, se $j=k+1,\dots,n$. Logo, completando os quadrados, teremos

$$\sum_{i=1}^{k}\lambda_i(y_i+(2\lambda_i)^{-1}d_i)^2+\sum_{j=k+1}^{n}d_iy_i+c-\sum_{i=1}^{k}((2\lambda_i)^{-1}d_i)^2=0.$$

Considere a translação $T : V \to V$ definida como $T(\mathbf{y}) = \mathbf{y} + \mathbf{p} = \mathbf{z}$, com

$$\mathbf{p} = ((2\lambda_1)^{-1}d_1, \ldots, (2\lambda_k)^{-1}d_k, 0, \ldots, 0) \text{ e } e = c - \sum_{i=1}^{k}((2\lambda_i)^{-1}d_i)^2.$$

Portanto,

$$\sum_{i=1}^{k}\lambda_i z_i^2 + \sum_{j=k+1}^{n} d_j z_j + e = 0.$$

Observe que uma translação muda apenas a localização da forma quadrática, mas não altera sua forma (dimensões). Além disso, a natureza de S_f é unicamente determinada por q.

Exemplo 8.31 *Sejam $V = \mathbb{R}^3$ e $f : V \to \mathbb{R}$ definida como*

$$f(x, y, z) = 6x^2 + 7y^2 + 5z^2 - 4xy + 4xz - 12x + 6y - 16z - 18$$

uma função quadrática sobre V. Determine o lugar geométrico de f.

Solução. Note que $q(\mathbf{X}) = \mathbf{X}^t\mathbf{A}\mathbf{X}$, onde $\mathbf{X} = [\mathbf{x}] = (x, y, z)^t \in \mathbb{R}^{3\times 1}$ e

$$\mathbf{A} = \begin{pmatrix} 6 & -2 & 2 \\ -2 & 7 & 0 \\ 2 & 0 & 5 \end{pmatrix}.$$

É fácil verificar que $f(x) = (x-3)(x-6)(x-9)$ é o polinômio característico de $\mathbf{A}$ e

$$\mathbf{X}_1 = \frac{1}{3}(2, 1, -2)^t, \mathbf{X}_2 = \frac{1}{3}(1, 2, 2)^t \text{ e } \mathbf{X}_3 = \frac{1}{3}(2, -2, 1)^t$$

autovetores associados aos autovalores $3, 6$ e 9. Seja $\mathbf{Q}$ a matriz cujas colunas são os vetores $\mathbf{X}_1, \mathbf{X}_2$ e $\mathbf{X}_3$. Então $\mathbf{Q}^t = \mathbf{Q}^{-1}$ e $\mathbf{Q}^t\mathbf{A}\mathbf{Q} = \operatorname{diag}(3, 6, 9)$. Pondo $\mathbf{Y} = \mathbf{Q}^t\mathbf{X} = (u, v, w)^t$, obtemos

$$3u^2 + 6v^2 + 9w^2 + 6u - 12v - 18w - 18 = 0.$$

Logo, completando os quadrados, teremos

$$\frac{(u+1)^2}{12} + \frac{(v-1)^2}{6} + \frac{(w-1)^2}{2} = 1.$$

A translação $T : V \to V$, $T(u, v, w) = (u + 1, v - 1, w - 1) = (a, b, c)$ implica que

$$\frac{a^2}{12} + \frac{b^2}{6} + \frac{c^2}{2} = 1.$$

Portanto, S_f representa um elipsoide de centro $C = (-1, 1, 1)$. Observe que $(3, 0)$ é a assinatura de Sylvester de q e q é definida positiva. ■

Teorema 8.32 *Sejam $V = \mathbb{R}^3$ e $f : V \to \mathbb{R}$ definida como*

$$f(x, y, z) = ax^2 + by^2 + cz^2 - 1 = q(x, y, z) - 1$$

uma quaádrica central sobre V, com assinatura de Sylvester (r, s) e $k = r + s \leq 3$.

1. *Se $k = 3$ e $(3, 0)$, então S_f representa um elipsoide.*
2. *Se $k = 3$ e $(2, 1)$, então S_f representa um hiperboloide de uma folha.*
3. *Se $k = 3$ e $(1, 2)$, então S_f representa um hiperboloide de duas folhas.*
4. *Se $k = 3$ e $(0, 3)$, então $S_f = \emptyset$.*
5. *Se $k = 2$ e $(2, 0)$, então S_f representa um cilindro elíptico.*
6. *Se $k = 2$ e $(1, 1)$, então S_f representa um cilindro "hiperbólico".*

Prova. Fica como um exercício. Observe que os itens (5) e (6), etc. são os casos degenerados. ■

Exercícios

1. Identifique as seguintes cônicas:

   (a) $12x^2 + 24xy + 9y^2 = 5$.
   (b) $5x^2 - 8xy + 5y^2 = 9$.
   (c) $2x^2 + 2\sqrt{3}xy = 1$.
   (d) $23x^2 - 72xy + 2y^2 + 30x + 40y = 0$.
   (e) $3x^2 + 3y^2 - 2xy + 6x - 2y - 3 = 0$.

2. Identifique as seguintes superfícies quaádricas:

(a) $12x^2 + 12y^2 + 12z^2 + 16xy + 12yz = 2$.
(b) $x^2 + y^2 + 4z^2 + 8xy + 2xz + 2yz = 3$.
(c) $y^2 - z^2 + 4xy - 6x + 4y + 2z + 8 = 0$.
(d) $-x^2 - y^2 - z^2 + 2xy + 2xz + 2yz = 2$.
(e) $x^2 + 3y^2 - 3z^2 + 4xy - 2xz + 2y - 4z + 2 = 0$.
(f) $x^2 + 2\sqrt{2}yz = 0$.

3. Identifique a quadrática $5x^2 + 6y^2 + 7z^2 - 4xy + 4yz = 0$. Conclua que

$$5x^2 + 6y^2 + 7z^2 > 4xy - 4yz, \ \forall \ (x, y, z) \in \mathbb{R}^3 - \{(0, 0, 0)\}.$$

4. Dar uma prova direta que: $q(x, y) = ax^2 + 2bxy + cy^2$ é definida positiva se, e somente se, $a > 0$ e $ac > b^2$.

5. Seja $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{n\times n}$ simétrica. Mostre que se $\mathbf{A}$ for definida positiva, então (a) $a_{ii} > 0$. (b) det $\mathbf{A} > 0$. (c) $a_{ii}a_{jj} > a_{ij}^2$, para $i, j = 1, \dots, n$.

6. Seja $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{n\times n}$ simétrica. Mostre que as seguintes condições são equivalentes:

(a) As submatrizes principais $\mathbf{A}_k$, para $k = 1, \dots, n$, possuem determinantes positivos;
(b) $\mathbf{A} = \mathbf{B}^t\mathbf{B}$, para alguma matriz triangular superior $\mathbf{B} = (b_{ij}) \in \mathbb{R}^{n\times n}$, com $b_{ii} > 0$;
(c) $\mathbf{A}$ é definida positiva.
(d) Comprove isto para $n = 2$.

7. Seja $S^2 = \mathbb{C} \cup \{\infty\}$. Cada $\mathbf{A} = a\mathbf{E}_{11} + b\mathbf{E}_{12} + c\mathbf{E}_{21} + d\mathbf{E}_{22} \in \mathrm{GL}_2(\mathbb{R})$ induz uma função $h_{\mathbf{A}} : S^2 \to S^2$ definida como

$$h_{\mathbf{A}}(z) = \frac{az + b}{cz + d}, \ \ c \neq 0 \Rightarrow h_{\mathbf{A}}(-\frac{d}{c}) = \infty \ \text{ e } \ h_{\mathbf{A}}(\infty) = \frac{a}{c},$$

e $c = 0$ implica que $h_{\mathbf{A}}(\infty) = \infty$.

(a) Mostre que $h_{\mathbf{A}} \in P(S^2)$, o *grupo de permutações* de S^2.

(b) Mostre que $h : \mathrm{GL}_2(\mathbb{R}) \to P(S^2)$ definida como $h(\mathbf{A}) = h_{\mathbf{A}}$ preserva as operações, ou seja, $h(\mathbf{AB}) = h(\mathbf{A}) \circ h(\mathbf{B})$. O que é $\ker h$?

(c) Seja $\mathbb{H} = \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$. Determine $\operatorname{Im} h_{\mathbf{A}}(z)$ em termos de $\operatorname{Im} z$. Conclua que se $G^+ = \{\mathbf{A} \in \mathrm{GL}_2(\mathbb{R}) : \det \mathbf{A} > 0\}$, então $h|_{G^+}$ satisfaz o item (b). Neste caso, diremos que G^+ *age* sobre $\mathbb{H}$.

(d) Mostre que $\mathrm{PSL}_2(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R})/\{\pm\mathbf{I}_2\}$ age sobre $\mathbb{H}$. O subgrupo $\mathrm{SL}_2(\mathbb{R}) = \{\mathbf{A} \in \mathrm{GL}_2(\mathbb{R}) : \det \mathbf{A} = 1\}$ chama-se *grupo modular*.

(e) Dado $\mathbf{A} \in \mathrm{SL}_2(\mathbb{R})$. Determine, em termos de $\mathrm{tr}(\mathbf{A})$, o número de pontos fixos de $h_{\mathbf{A}}$ sobre $\mathbb{H}$.

Bibliografia

[1] **Andrade, L. N. de**, *Introdução à Computação Algébrica com o Maple*, SBM, 2004.

[2] **Barone Júnior, M.**, *Álgebra Linear*, Vol. I e II, 3.^{a} Ed., São Paulo, 1988.

[3] **Baudry, W. C. et al**, *Linear Algebra with Maple*, John Wiley, 1995.

[4] **Boudrini, J. L. et al**, *Álgebra Linear*, 3.^{a} Edição, Ed. Harbra Ltda, 1986.

[5] **Deeba, E. and Gynawardena, A.**, *Interactive Linear Algebra with Maple V*, Springer-Verlag, 1998.

[6] **Finkbeiner, D. T.**, *Introdução às Matrizes e Transformações Lineares*, Ed. LTC, Rio de Janeiro, 1970.

[7] **Gantmather, F. R.**, *Theory of Matrices*, Vol. 1, Chelsea, New York, 1959

[8] **Goldberg, J. L.**, *Matrix Theory with Applications*, McGraw-Hill, 1991.

[9] **Halmos, P. R.**, *Espaços Vetoriais de Dimensão Finita*, Ed. Campus Ltda, 1978.

[10] **Hoffman, K. e Kunze, R.**, *Álgebra Linear*, 2.^{a} Edição, Ed. LTC, Rio de Janeiro, 1979.

[11] **Jacobson, N.**, *Lectures in Abstract Algebra*, Vol. II, Springer-Verlag, 1975.

[12] **Lang, S.**, *Álgebra Linear*, Ed. Edgard Blücher Ltda, 1971.

[13] **Lipschutz, S.**, *Álgebra Linear*, Makron Books (Coleção Schaum), 1994.

[14] **Medeiros, A. S. de**, "Transformações Lineares e Escalonamento de Matrizes," *Matemática Universitária*, $N.^o$ 30 - Junho 2001, pp. 131-133.

[15] **Splinder, K.**, *Abstract Algebra with Applications*, Vol. 1 Marcel Dekker, Inc. 1994.

[16] **Zhang, F.** *Linear Algebra-Challenging Problems for Students*, 2nd Edition, The Johns Hopkins University Press, 2009.

Respostas e Sugestões

Capítulo 1

Seção 1.1

1. $b = 1 \cdot b = (a^{-1}a)b = a^{-1}(ab) = a^{-1}(ac) = (a^{-1}a)c = 1 \cdot c = c$. 2. (c) Basta notar $\Re(z) \leq |\Re(z)| \leq |z|$, para todo $z \in \mathbb{C}$, e desenvolver $|z+w|^2$. 3. Não, pois se F fosse um subcorpo finito de $\mathbb{C}$, então $1, 1+1, 1+1+1, \ldots$ seria infinitos elementos de F. 4. Basta lembrar que as raízes complexas ocorrem aos pares.

Seção 1.2

2. Basta verificar que f é bijetora e preserva as operações, para concluir que $M_2(\mathbb{R})$ é um corpo. 3. Tome

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ e } \mathbf{B} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

4. e 5. Use indução sobre m. 6. Seja $\mathbf{A} = (a_{ij})$. Como $\mathbf{AX} = \mathbf{X}$, para todo $\mathbf{X} \in F^{n\times 1}$, temos, em particular, que $\mathbf{AE}_j = \mathbf{E}_j$. Assim, $c_{ij} = \sum_{k=1}^{n} a_{ik}\delta_{kj} = \delta_{ij}$ implica que $a_{ij} = \delta_{ij}$. Portanto, $\mathbf{A} = \mathbf{I}$. 7. (a)

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & 1-a \end{pmatrix}, \ \forall\, a, b \in \mathbb{R}^{\times}, \ \text{ com } \ bc = a - a^2.$$

Note que $\mathbf{A} \sim \mathbf{E}_{11}$. (b) $\mathbf{A}^2 = (\mathbf{AB})\mathbf{A} = \mathbf{A}(\mathbf{BA}) = \mathbf{AB} = \mathbf{A}$. 8. (a) Considerando os casos $a + d \neq 0$ e $a + d = 0$, com $a \neq 0$, obtemos $\mathbf{A} = \mathbf{O}$ e $\mathbf{A} = c\mathbf{E}_{21}$ ou

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}, \text{ com } bc = -a^2 = -d^2.$$

Note que $\mathbf{A} \sim \mathbf{E}_{21}$. (b) Observe que $\mathbf{I} = \mathbf{I} - \mathbf{A}^k = (\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A} + \cdots + \mathbf{A}^{k-1})$. 9. (a)

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}, \forall\, a, b \in \mathbb{R}, \text{ com } b \neq 0, \ bc = 1 - a^2.$$

Note que $\mathbf{A} \sim \mathbf{E}_{11} - \mathbf{E}_{22}$. 10. (a)

$$\mathbf{A} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}, a = \pm 1, \text{ ou } \mathbf{A} = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}, \forall\, a, b \in \mathbb{R}, \text{ com } a^2 + bc = 1.$$

11. Desenvolva $(\mathbf{I} - \mathbf{BA})(\mathbf{I} + \mathbf{B}(\mathbf{I} - \mathbf{AB})^{-1}\mathbf{A})$. 12. Análogo a (11). 13. $\mathbf{X} = \mathbf{CA}^{-1}$ e $\mathbf{Y} = \mathbf{A}^{-1}\mathbf{B}$. 14. Sejam $\mathbf{A} = (a_{ij})$ e $\mathbf{B} = (b_{ij})$. Então $\mathbf{AB} = \mathbf{BA}$ se, e somente se, $\sum_{k=1}^{n} a_{ik}b_{kj} = \sum_{k=1}^{n} b_{ik}a_{kj}$. Em particular, para $b_{rs} = \delta_{pr}\delta_{sq}$, obtemos

$$\sum_{k=1}^{n} a_{ik}\delta_{pk}\delta_{jq} = \sum_{k=1}^{n} \delta_{pi}\delta_{kq}a_{kj} \Rightarrow a_{ip}\delta_{jq} = a_{qj}\delta_{pi}.$$

Assim, $a_{ip} = 0$ se $i \neq p$ e $j = q$. Enquanto, $a_{ii} = a_{jj} = a$ se $i = p$ e $j = q$. Portanto, $\mathbf{A} = a\mathbf{I}$. 15. Observe que $\mathbf{D} = \mathbf{A}(\mathbf{I} - \mathbf{D})$. 16. É fácil verificar que se $\mathbf{X} = u\mathbf{A} + v\mathbf{I}$, então $\mathbf{AX} = \mathbf{XA}$. Reciprocamente, sejam

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \mathbf{X} = \begin{pmatrix} x & y \\ z & w \end{pmatrix} \in F^{2\times 2} \text{ e } \mathbf{A}^2 = (a + d)\mathbf{A} + (bc - ad)\mathbf{I}.$$

Então $\mathbf{AX} = \mathbf{XA}$ implica que $cy = bz$ e $cx = cw + (a - d)z$. Se $c = 0$, então $z = 0$ ou $a = d$. Assim, $\mathbf{X} = u\mathbf{A} + v\mathbf{I}$, com $u \neq 0$ e $v = 0$. Se $c \neq 0$, então $z \neq 0$, $y = bu$ e $x = w + (a - d)u$, com $z = cu$, de modo que $\mathbf{X} = u\mathbf{A} + (w - du)\mathbf{I}$. 17. (c) Sejam $\mathbf{A} = (a_{ij}), \mathbf{B} = (b_{ij}), \mathbf{AB} = (c_{ij})$ e $\mathbf{BA} = (d_{ij})$. Então

$$\begin{aligned} \operatorname{tr}(\mathbf{AB}) &= \textstyle\sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{ik}b_{ki}\right) = \sum_{k=1}^{n} \left(\sum_{i=1}^{n} b_{ki}a_{ik}\right) \\ &= \textstyle\sum_{k=1}^{n} d_{kk} = \operatorname{tr}(\mathbf{BA}). \end{aligned}$$

(d) Note que se $\mathbf{A} = (a_{ij})$ e $\mathbf{A}\mathbf{A}^t = (c_{ij})$, então $c_{ij} = \sum_{k=1}^n a_{ik}a_{jk}$, de modo que $c_{ii} = \sum_{k=1}^n a_{ik}^2$. 18. Basta observar que $\mathbf{U}_n^2 = n\mathbf{U}_n$.

Seção 1.3

2. Dado $\phi \in I$, existe um $\sigma = \phi\tau \in P$ tal que $f(\sigma) = (\phi\tau)\tau = \phi$, ou seja, f é sobrejetora. Se $f(\sigma) = f(\phi)$, então

$$\sigma = \sigma \circ I = \sigma(\tau^2) = (\sigma\tau)\tau = (\phi\tau)\tau = \phi(\tau^2) = \phi \circ I = \phi,$$

ou seja, f é injetora. 3. Aplique recursivamente o item (2) do Teorema 1.24. 4. Como $\mathbf{A}^2 = \mathbf{A}$ temos que $(\det \mathbf{A})^2 = \det \mathbf{A}$, de modo que $\det \mathbf{A} = 0$ ou $\det \mathbf{A} = 1$. 5. Como $\mathbf{A}^k = \mathbf{O}$ temos que $(\det \mathbf{A})^k = 0$, de modo que $\det \mathbf{A} = 0$. 6. Basta aplicar a fórmula ou o Teorema de Laplace em relação às linhas $1, 2, \ldots, n$ e $n+1, \ldots, 2n$. 7. Pelo Exercício (16) da Seção 1.2, existem $p, q, r, s \in F$ tais que $\mathbf{C} = p\mathbf{A} + q\mathbf{I}$ e $\mathbf{C} = r\mathbf{B} + s\mathbf{I}$. Se $p \neq 0$, então $\mathbf{A} = (p^{-1}r)\mathbf{B} + p^{-1}(s-q)\mathbf{I}$, de modo que $\mathbf{AB} = \mathbf{BA}$. Se $p = 0$ e $r \neq 0$, então $\mathbf{B} = r^{-1}(q-s)\mathbf{I}$, de modo que $\mathbf{AB} = \mathbf{BA}$. Se $p = r = 0$, então $q = s$ e $\mathbf{C} = q\mathbf{I}$, de modo que $1 = \det \mathbf{C} = q^2$ implica que $q = 1$ ou $q = -1$. Portanto, $\mathbf{AB} = \mathbf{BA}$ ou $\mathbf{AB} = -\mathbf{BA}$. 8. De modo análogo ao Exercício (7), existem $p, q, r, s \in F$ tais que $\mathbf{C} = p\mathbf{A} + q\mathbf{I}$ e $\mathbf{C} = r\mathbf{B} + s\mathbf{I}$. Se $p = r = 0$, então $q = s$ e $\mathbf{C} = q\mathbf{I}$, de modo que $0 = \operatorname{tr}(\mathbf{C}) = 2q$ implica que $q = 0$. Portanto, $\mathbf{AB} = \mathbf{BA}$. 9. Sejam

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \mathbf{X} = \begin{pmatrix} x & y \\ z & w \end{pmatrix} \in F^{2\times 2}.$$

Então $\mathbf{AX} = -\mathbf{XA}$ implica que $2ax + cy + bz = 0$, $bx + (a+d)y + bw = 0$, $cx + (a+d)z + cw = 0$ e $cy + bz + 2dw = 0$. Mas a existência de $\mathbf{X} \neq \mathbf{O}$ é equivalente ao determinante da matriz dos coeficientes do "sistema homogêneo" ser zero, ou seja, $4(a+d)^2(ad-bc) = 0$, de modo que $ad - bc = 0$ ou $a + d = 0$. Portanto, $\det \mathbf{A} = 0$ ou $\operatorname{tr}(\mathbf{A}) = 0$. 10. Como

$$\det(x\mathbf{I} + \mathbf{A}) = (x + a_{11}) \cdots (x + a_{nn}) + p_{n-2}(x)$$

e os fatores $x + a_{ii}$ ou $x + a_{jj}$ não aparecem em $p_{n-2}(x)$, quando $i \neq j$, temos que $p_{n-2}(x)$ é um polinômio de grau no máximo $n-2$. Portanto, $\det(x\mathbf{I} + \mathbf{A})$ é um polinômio de grau n. Pelo Teorema 1.2, esse polinômio possui no máximo n raízes sobre F. Assim, existe um $c \in F$ tal que $\det(c\mathbf{I} + \mathbf{A}) \neq 0$, pois F é infinito. Portanto, $c\mathbf{I} + \mathbf{A}$ é não singular. 11. Confira o Exercício (10) 12. Vamos usar indução sobre n. Se $n = 2$, então é claro que $\det(x\mathbf{I} - \mathbf{C}_2) = x^2 + a_1x + a_0$. Suponhamos que o resultado seja válido para todo k, com $1 \leq k \leq n-1$ e $n > 2$. Então, pela fórmula de Laplace em relação à primeira coluna, obtemos

$$\det(x\mathbf{I} - \mathbf{C}_n) = x\det(x\mathbf{I} - \mathbf{C}_{n-1}) + (-1)^{n+1}a_0(-1)^{n-1}.$$

Como

$$x\mathbf{I} - \mathbf{C}_{n-1} = \left(\begin{array}{c|c} x\mathbf{I}_{n-2} & -\mathbf{I}_{n-2} \\ \hline a_1 & a_2 \quad \cdots \quad x + a_{n-1} \end{array}\right)$$

temos, pela hipótese de indução, que

$$p(x) = \det(x\mathbf{I}_n - \mathbf{C}_n) = x^n + a_{n-1}x^{n-1} + \cdots + a_2x^2 + a_1x + a_0.$$

Outra prova é aplicando a sequência de operações elementares

$$\mathbf{C}_j \to \mathbf{C}_j + x\mathbf{C}_{j+1}, \quad j = n-1, \ldots, 2, 1,$$

13. Note que $\mathbf{B} = \mathbf{V}_n\mathbf{A}$. 14. Note que $\mathbf{V}_{n+1}\mathbf{X} = \mathbf{O}$. 15. Como $\mathbf{A}\operatorname{adj}\mathbf{A} = \det\mathbf{A}\cdot\mathbf{I}$ temos que $\det\mathbf{A}\cdot\det(\operatorname{adj}\mathbf{A}) = \det(\mathbf{A}\operatorname{adj}\mathbf{A}) = (\det\mathbf{A})^n$, de modo que $\det(\operatorname{adj}\mathbf{A}) = (\det\mathbf{A})^{n-1}$. 16. Como $1 = \det\mathbf{I} = \det(\mathbf{A}\mathbf{A}^t) = (\det\mathbf{A})^2$ temos que $\det\mathbf{A} = \pm 1$. 17. (b) Pelo Exercício (16) $|\det\mathbf{A}| = 1$, de modo que $\det\mathbf{A} = e^{i\theta}$, para algum $\theta \in \mathbb{R}$. 18. $f(x)$ e $g(x)$ possuem uma raiz em comum λ se, e somente se, existem $p(x) = ax+b$ e $q(x) = cx+d$, com $a \neq 0$ e $c \neq 0$, tais que $f(x) = (x-\lambda)p(x)$ e $g(x) = (x-\lambda)q(x)$. Mas, isto é equivalente a: $q(x)f(x) - p(x)g(x) = 0$ e, depois de alguns cálculos, obtemos

$$\begin{aligned}(ca_0 - ab_0)x^3 + (ca_1 + da_0 - ab_1 - bb_0)x^2 &\quad + \\ (ca_2 + da_1 - ab_2 - bb_1)x + da_2 - b_2b &= 0,\end{aligned}$$

ou seja, o sistema homogêneo

$$\begin{pmatrix} a_0 & 0 & b_0 & 0 \\ a_1 & a_0 & b_1 & b_0 \\ a_2 & a_1 & b_2 & b_1 \\ 0 & a_2 & 0 & b_2 \end{pmatrix} \begin{pmatrix} c \\ d \\ -a \\ -b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

possui uma solução não nula ou, equivalentemente, $\det\operatorname{res}(f, g) = 0$. 19. Basta notar que

$$\det \mathbf{H} = \det \begin{pmatrix} 0 & a^2 & c^2 & 1 \\ a^2 & -a^2 & b^2-c^2 & 0 \\ c^2 & b^2-a^2 & -c^2 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} = -\det \begin{pmatrix} a^2 & -a^2 & b^2-c2 \\ c^2 & b^2-a^2 & -c^2 \\ 1 & 1 & 1 \end{pmatrix}$$

20. Sejam $\mathbf{A}$ a matriz cujas linhas são $(a_1, \ldots, a_n)$ e $(b_1, \ldots, b_n)$ e $\mathbf{B}$ a matriz cujas colunas são $(c_1, \ldots, c_n)$ e $(d_1, \ldots, b_n)$. Então $\mathbf{AB} \in F^{2\times 2}$ e, pelo Corolário 1.30,

$$\det(\mathbf{AB}) = \sum_{1\leq i<j\leq n} \det \begin{pmatrix} a_i & a_j \\ b_i & b_j \end{pmatrix} \det \begin{pmatrix} c_i & c_j \\ d_i & d_j \end{pmatrix}.$$

21. Faça $c_i = a_i$ e $d_i = b_i$ no Exercício (19). 22. Note que o determinante à direita da identidade de Cauchy é positivo e $b_i = \lambda a_i$ se, e somente se, o determinante $a_ib_j - a_jb_i = 0$. 23. Faça $b_i = 1$ no Exercício (22). 24. Imite a prova do Teorema 1.13 usando operações elementares de linhas e colunas. 25. Primeiro vamos provar que f satisfaz as condições do Exercício (24). Consideremos as operações elementares $\mathbf{B} = \mathbf{L}_j(c^{-1})\mathbf{T}_{ij}(1)\mathbf{L}_j(c)\mathbf{A}$, obtemos

$$f(\mathbf{B}) = c^{-1}f(\mathbf{T}_{ij}(1)\mathbf{L}_j(c)\mathbf{A}) = c^{-1}f(\mathbf{L}_j(c)\mathbf{A}) = c^{-1} \cdot cf(\mathbf{A}) = f(\mathbf{A}).$$

Note que $\mathbf{B} = \mathbf{T}_{ij}(c)\mathbf{A}$. Novamente, $\mathbf{C} = \mathbf{T}_{ij}(1)\mathbf{T}_{ji}(-1)\mathbf{T}_{ij}(1)\mathbf{A}$ implica que

$$f(\mathbf{C}) = f(\mathbf{T}_{ij}(-1)\mathbf{T}_{ij}(1)\mathbf{A}) = f(\mathbf{T}_{ij}(1)\mathbf{A}) = f(\mathbf{A}).$$

Observe que $\mathbf{C} = \mathbf{L}_j(-1)\mathbf{P}_{ij}\mathbf{A}$. Se $\mathbf{A}$ for não singular, então ela pode ser transformada em uma matriz diagonal $\mathbf{D} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, com $\lambda_i \neq 0$. Assim,

$$f(\mathbf{A}) = f(\mathbf{D}) = \lambda_1 \cdots \lambda_n f(\mathbf{I}) = \lambda_1 \cdots \lambda_n = \det \mathbf{A}.$$

Se $\mathbf{A}$ for singular, então ela possui uma linha que é combinação linear das outras, a qual pode ser transformada em uma linha nula, de modo que $f(\mathbf{A}) = f(\mathbf{E})$, em que $\mathbf{E}$ possui a linha $\mathbf{L}_j$ nula. Logo, $f(\mathbf{F}) = f(\mathbf{L}_j(0)\mathbf{E}) = 0$. Portanto, $f(\mathbf{A}) = 0 = \det \mathbf{A}$. 26. Note que $f(\mathbf{O}) = f(\mathbf{OY}) = f(\mathbf{O})f(\mathbf{Y})$ implica que $f(\mathbf{O}) = 0$ e $f(\mathbf{X}) = f(\mathbf{IX}) = f(\mathbf{I})f(\mathbf{X})$ implica que $f(\mathbf{I}) = 1$. Como $\mathbf{P}_{12}^2 = \mathbf{I}$, $\mathbf{T}_{12}(c)\mathbf{T}_{12}(-c) = \mathbf{I}$ e $\mathbf{D}_1(c)\mathbf{D}_1(c^{-1}) = \mathbf{I}$ temos que $f(\mathbf{T}_{12}(c))f(\mathbf{T}_{12}(-c)) = 1$, de modo que $f(\mathbf{T}_{12}(c))$ é independente de c. Assim, $f(\mathbf{T}_{ij}(c)) = 1$, pois vale com $c = 0$. Note que $\mathbf{P}_{12}\mathbf{D}_1(c)\mathbf{P}_{12} = \mathbf{D}_2(c)$ implica que $f(\mathbf{D}_1(c)) = f(\mathbf{D}_2(c))$. Logo, $\mathbf{D}_1(c)$ e $\mathbf{D}_1(c^{-1})$ são monômios de mesmo grau m em c e c^{-1}, digamos a_mc^m, de modo que $a_m = 1$ e $m = 1$ é o menor valor. Portanto, $f(\mathbf{D}_i(c)) = c$. Finalmente, $\mathbf{T}_{12}(-1)\mathbf{T}_{21}(1)\mathbf{T}_{12}(-1)\mathbf{P}_{12} = \mathbf{D}_1(-1)$ implica que $f(\mathbf{P}_{12}) = -1$. Consequentemente, $f(\mathbf{EA}) = cf(\mathbf{A})$, com $c \neq 0$, garante que f é uma função com as propriedades de um determinante e o resultado segue. 27. Note que $\Delta = (\det \mathbf{V}_n)^2 = \det(\mathbf{V}_n\mathbf{V}_n^t)$. 28. Use o Exercício (10).

Capítulo 2

Seção 2.1

1. Basta notar que $\rho(\mathbf{R}) < n$ ou $\rho(\mathbf{R}) = n$. 2. $a = 4, b = 0$ e $c = 2$. 3. Consistente e determinado, $\left(\frac{8}{3}, \frac{1}{3}, 0\right)$; Consistente e indeterminado, $(\frac{5}{2} + \frac{7}{2}t, -2 - 3t, t)$, para todo $t \in F$. 4. $a = 3$ e $b = 11$. 5. $b_2 = 2b_1 + b_3$. 6. $\lambda \neq 9$. 7. $\det \mathbf{H} = \frac{(\prod_{k=1}^{n-1} k!)^4}{\prod_{k=1}^{2n-1} k!}$, 8. e 9.

$$\mathbf{H}^{-1} = \begin{pmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{pmatrix}, \frac{1}{10}\begin{pmatrix} 16 & -20 \\ -7 & 20 \end{pmatrix} \text{ e } \frac{1}{8}\begin{pmatrix} 3 & -2 & 3 \\ 2 & 0 & -2 \\ 1 & 2 & 1 \end{pmatrix}.$$

10. Aplique simultaneamente operações de linhas e colunas a matriz

$$\left(\begin{array}{rrr|rrr} 1 & 2 & -3 & 1 & 0 & 0 \\ 2 & 5 & -4 & 0 & 1 & 0 \\ -3 & -4 & 8 & 0 & 0 & 1 \end{array}\right) \to \cdots \to \left(\begin{array}{c|c} \mathbf{D} & \mathbf{P}^t \end{array}\right).$$

11. Se $t = 0$, então $U \cap V \cap W = \emptyset$. Se $t \neq 0$, então

$$U \cap V \cap W = \left(\frac{t^2 + 5t - 2}{2t}, \frac{-2 + 3t}{2t}, \frac{2 - t}{2t}\right).$$

12. Basta notar que $\mathbf{AX} = \mathbf{B}$ é equivalente a n sistemas. (a) $\mathbf{AXC} = \mathbf{BC} = \mathbf{AD}$.

Seção 2.2

1. Consistente e indeterminado, $(\frac{11}{3} + t\frac{2}{3}, -\frac{4}{3} + t\frac{2}{3}, t$, para todo $t \in F$; Inconsistente. 2. Observe que a matriz escalonada $\mathbf{R}$ de $\mathbf{A}$ possui pelo menos $n - r$ linhas nulas. Portanto, $\rho(\mathbf{A}) \leq r$. 3. Suponhamos que $\rho(\mathbf{A}) = r$ e que $\mathbf{R}$ seja a matriz escalonada de $\mathbf{A}$. Então existe uma matriz não singular $\mathbf{P}$ tal que $\mathbf{PA} = \mathbf{R}$. Assim, $\mathbf{P}(\mathbf{AB}) = \mathbf{RB}$, de modo que $\rho(\mathbf{AB}) = \rho(\mathbf{RB})$. Observe que se a i-ésima linha de $\mathbf{R}$ for nula, então a i-ésima linha de $\mathbf{RB}$ é nula. Como $\mathbf{R}$ possui exatamente $n - r$ linhas nulas temos que

$\mathbf{RB}$ possui pelo menos $n-r$ linhas nulas. Logo, $\rho(\mathbf{RB}) \leq r$. Portanto, pelo Exercício (2), $\rho(\mathbf{AB}) \leq r = \rho(\mathbf{A})$. Por outro lado,

$$\rho(\mathbf{AB}) = \rho((\mathbf{AB})^t) = \rho(\mathbf{B^t A^t}) \leq \rho(\mathbf{B^t}) = \rho(\mathbf{B}).$$

Portanto, $\rho(\mathbf{AB}) \leq \min\{\rho(\mathbf{A}), \rho(\mathbf{B})\}$. Se $\mathbf{B}$ for não singular, então $\mathbf{BA}$ é equivalente por linha a $\mathbf{A}$, de modo que $\rho(\mathbf{BA}) = \rho(\mathbf{A})$. Para a outra igualdade, como $\rho((\mathbf{AB})\mathbf{B}^{-1}) = \rho(\mathbf{A})$ temos que $\rho(\mathbf{A}) = \rho((\mathbf{AB})\mathbf{B}^{-1}) \leq \rho(\mathbf{AB}) \leq \rho(\mathbf{A})$. Portanto, $\rho(\mathbf{A}) = \rho(\mathbf{AB})$. 4. Observe que $\rho(\mathbf{B}) + \rho(\mathbf{C}) \leq m + \rho(\mathbf{C}) = \rho(\mathbf{E})$, pois

$$\begin{pmatrix} \mathbf{I}_m & \mathbf{O} \\ -\mathbf{CA^{-1}} & \mathbf{I}_n \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B} \end{pmatrix}.$$

5. Como

$$\begin{pmatrix} \mathbf{I}_n & \mathbf{O} \\ -\mathbf{A} & \mathbf{I}_m \end{pmatrix} \begin{pmatrix} \mathbf{I}_n & \mathbf{B} \\ \mathbf{A} & \mathbf{O} \end{pmatrix} \begin{pmatrix} \mathbf{I}_n & -\mathbf{B} \\ \mathbf{O} & \mathbf{I}_m \end{pmatrix} = \begin{pmatrix} \mathbf{I}_n & \mathbf{O} \\ \mathbf{O} & -\mathbf{AB} \end{pmatrix}.$$

temos que

$$\rho\begin{pmatrix} \mathbf{I}_n & \mathbf{B} \\ \mathbf{A} & \mathbf{O} \end{pmatrix} = \rho\begin{pmatrix} \mathbf{I}_n & \mathbf{O} \\ \mathbf{O} & -\mathbf{AB} \end{pmatrix}.$$

Assim, pelo Exercício (4), $\rho(\mathbf{A}) + \rho(\mathbf{B}) \leq n + \rho(\mathbf{AB})$. 6. $a = 1$ e $b = c = d = e = 0$. 7. (c) Note, pelo item (b), que $1 + b_2 + b_3 = 3b_2$ e $2 + b_3 + 4 = 3b_2$ implicam que $2b_2 - b_3 = 1$ e $3b_2 - b_3 = 6$, de modo que $b_2 = 5$ e $b_3 = 9$. Assim, $s = 15$, continue! 8. Basta notar que $y_1 = x_1$ e $y_2 = -x_2$ e resolver o sistema. 9. O sistema possui soluções inteiras se, e somente se, $a + 2b = c$ e $a + b$ é um múltiplo (divisível por) de 3. 10. Observe que se a decomposição Cartesiana $\mathbf{Z}_0 = \mathbf{X}_0 + i\mathbf{Y}_0 \in \mathbb{C}^{n\times 1}$ for uma solução, então $\bar{\mathbf{Z}}_0 = \mathbf{X}_0 - i\mathbf{Y}_0$ também o é, pois $\mathbf{A}\bar{\mathbf{Z}}_0 = \overline{\mathbf{AZ}_0} = \bar{\mathbf{B}} = \mathbf{B}$. Portanto, $\mathbf{X}_0 = \frac{1}{2}(\mathbf{Z}_0 + \bar{\mathbf{Z}}_0)$ e $\mathbf{Y}_0 = \frac{1}{2i}(\mathbf{Z}_0 - \bar{\mathbf{Z}}_0)$são soluções reais. 11. Seja $f(x) = a_1 + a_2x + \cdots + a_nx^{n-1}$ o polinômio desejado. Então $\mathbf{X} = (\mathbf{V}_n^t\mathbf{V}_n)^{-1}\mathbf{V}_n^t\mathbf{Y}$, com $\mathbf{X} = (a_1 \cdots a_n)^t$ e $\mathbf{Y} = (y_1 \cdots y_n)^t$ ou aplique a Regra de Cramer ao sistema $\mathbf{V}_n\mathbf{X} = \mathbf{Y}$, para obter os coeficientes a_i. 12. (a) Vamos usar indução sobre n. Efetuando a operação elementar

$\mathbf{L}_1 \to \mathbf{L}_1 - \mathbf{L}_2$, obtemos

$$\mathbf{D}_n \to \begin{pmatrix} x_1 - a & a & a & a & \cdots & a & a \\ b - x_2 & x_2 & a & a & \cdots & a & a \\ 0 & b & x_3 & a & \cdots & a & a \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & b & b & b & \ldots & x_{n-1} & a \\ 0 & b & b & b & \ldots & b & x_n \end{pmatrix}.$$

Assim, pela fórmula de Laplace em relação à primeira coluna, teremos

$$\det(\mathbf{D}_n) = (x_1 - a)\det(\mathbf{D}_{n-1}) + a(x_2 - b)\cdots(x_n - b)$$

Logo, pela hipótese de indução,

$$\det(\mathbf{D}_n) = (x_1 - a)\frac{bg(a) - ag(b)}{b - a} + a(x_2 - b)\cdots(x_n - b),$$

com $g(x) = (x_2 - x)\cdots(x_n - x)$. Portanto, depois de alguns cálculos,

$$\det(\mathbf{D}_n) = \frac{bf(a) - af(b)}{b - a}.$$

(b) Note, recursivamente, que

$$\begin{aligned} \det(\mathbf{D}_n) &= (x_1 - a)\det(\mathbf{D}_{n-1}) + af_1(a) \\ &= (x_1 - a)[(x_2 - a)\det(\mathbf{D}_{n-2}) + ag_2(a)] + af_1(a) \\ &= (x_1 - a)(x_2 - a)\det(\mathbf{D}_{n-2}) + af_2(a) + af_1(a)\cdots \\ &= (x_1 - a)\cdots(x_{n-2} - a)\det(\mathbf{D}_2) + a\textstyle\sum_{k=2}^{n-1} f_{n-k}(a). \end{aligned}$$

Como $\det(\mathbf{D}_2) = x_{n-1}x_n - a^2 = x_n(x_{n-1} - a) + a(x_n - a)$, o resultado segue.

Capítulo 3

Seção 3.1

2. Não, pois $\mathbf{u} = (1, 2) \in V$ implica que $1 \odot \mathbf{u} = (6, -1) \neq \mathbf{u}$. 3. Não, pois $a = 2, b = 3$ e $\mathbf{u} = (1, 2) \in V$ implicam que $(a + b) \odot \mathbf{u} = (25, 50)$ e $a \odot \mathbf{u} + b \odot \mathbf{u} = (13, 26)$, de modo que $(a + b) \odot \mathbf{u} \neq a \odot \mathbf{u} + b \odot \mathbf{u}$. 4. Não, pois $\mathbf{u} = (1, 2) \in V$ implica que

$1 \odot \mathbf{u} = (5, 10) \neq \mathbf{u}$. 5. Não, pois $\mathbf{u} = (1, 2) \in V$ implica que $1 \odot \mathbf{u} = (1, 0) \neq \mathbf{u}$. 6. Não, pois $\mathbf{u} = (1, 0, \ldots, 0) \in V$ implica que $1 \odot \mathbf{u} = (0, 0, \ldots, 0) \neq \mathbf{u}$. 7. Sim. 8. Não, se $\mathbf{v} \neq \mathbf{0}$, então $\mathbf{0} \oplus \mathbf{v} \neq \mathbf{v}$. 9. Não, se $\mathbf{u} = (1, 2), \mathbf{v} = (3, 4), \mathbf{w} = (4, 2) \in V$, então $\mathbf{u} \oplus (\mathbf{v} \oplus \mathbf{w}) = (2, 4)$ e $(\mathbf{u} \oplus \mathbf{v}) \oplus \mathbf{w} = (-6, -4)$, de modo que $\mathbf{u} \oplus (\mathbf{v} \oplus \mathbf{w}) \neq (\mathbf{u} \oplus \mathbf{v}) \oplus \mathbf{w}$. 10. Desenvola $(1 + 1)(\mathbf{u} + \mathbf{v})$ de duas maneiras. 11. Se existir um $\mathbf{u} \in V$, com $\mathbf{u} \neq \mathbf{0}$, então a função $f : F \to V$ definida como $f(a) = a\mathbf{u}$ é claramente injetora, de modo que V possui infinitos elementos. 11. Basta notar que a *função de complexificação* $\lambda : V \to \tilde{V}$ definida como $\lambda(\mathbf{u}) = (\mathbf{u}, \mathbf{0})$ é injetora e preserva as operações, de modo que podemos identificar $\mathbf{u}$ com sua complexificação $(\mathbf{u}, \mathbf{0})$. Portanto, $(\mathbf{u}, \mathbf{v}) = (\mathbf{u}, \mathbf{0}) \oplus (\mathbf{0}, \mathbf{v}) = (\mathbf{u}, \mathbf{0}) \oplus i \odot (\mathbf{v}, \mathbf{0}) = \mathbf{u} + i\mathbf{v}$.

Seção 3.2

1. (a) Sim. (b) Não, pois $\mathbf{u} = (0, 1, 2) \in W$, mas $-1 \cdot \mathbf{u} = (0, -1, -2) \notin W$. (c) Sim. (d) Não, pois $\mathbf{u} = (1, 2, 3) \in W$, mas $\frac{1}{2} \cdot \mathbf{u} \notin W$. (e) Não, pois $\mathbf{u} = (1, 0, 0), \mathbf{v} = (0, 1, 0) \in W$, mas $\mathbf{u} + \mathbf{v} = (1, 1, 0) \notin W$. (f) Não. (g) Não, pois $\mathbf{u} = (1, 0, 1), \mathbf{v} = (0, 1, 1) \in W$, mas $\mathbf{u} + \mathbf{v} = (1, 1, 2) \notin W$. (h) Não. 2. (a) Sim. (b) Não, pois $\mathbf{A} \in W$, com entradas $a = d = -1$ e $b = c = 1$, mas $-1 \cdot \mathbf{A} \notin W$. (c) Sim. (d) Não, pois $\mathbf{I} \in W$, mas $2\mathbf{I} \notin W$. (e) Não, pois $\mathbf{O} \notin W$. (f) Não, pois $\mathbf{A} \in W$, com entradas $a = 1$ e $b = c = d = 0$; $\mathbf{B} \in W$, com entradas $d = 1$ e $a = b = c = 0$, mas $\mathbf{A} + \mathbf{B} \notin W$. 3. (a) Sim. (b) Sim. (c) Sim. (d) Não, $\mathbf{0} \notin W$. (e) Sim. 4. (a) Não, pois $\mathbf{0} \notin W$. $(b), (c), (d), (e)$ e (f) São. 5. Note que

$$(x, y, z) = (x, y, x) + (0, 0, z - x) \in W_1 + W_2 \Rightarrow W_1 + W_2 = V.$$

De modo análogo, $W_1 + W_3 = V = W_2 + W_3$; $W_1 \oplus W_2 = F^3 = W_2 \oplus W_3$, mas $W_1 \cap W_3 = \{(x, -2x, x) \in V : x \in F\}$. 6. Se $\mathbf{v} \in V$, então existe um $\mathbf{u}_1 \in W_1$ e $\mathbf{u}_2 \in W_2$ tal que $\mathbf{v} = \mathbf{u}_1 + \mathbf{u}_2$. Seja $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$, onde $\mathbf{v}_1 \in W_1$ e $\mathbf{v}_2 \in W_2$, outra forma de escrever. Então $\mathbf{u}_1 + \mathbf{u}_2 = \mathbf{v}_1 + \mathbf{v}_2$, de modo que

$$\mathbf{u}_1 - \mathbf{v}_1 = \mathbf{v}_2 - \mathbf{u}_2 \in W_1 \cap W_2 = \{\mathbf{0}\}.$$

Assim, $\mathbf{u}_1 = \mathbf{v}_1$ e $\mathbf{u}_2 = \mathbf{v}_2$. Reciprocamente. é claro que $V = W_1 + W_2$. Se $\mathbf{v} \in W_1 \cap W_2$, então $\mathbf{v} \in W_1$ e $\mathbf{v} \in W_2$. Como $\mathbf{v} + \mathbf{0} = \mathbf{v} = \mathbf{0} + \mathbf{v}$ temos, pela unicidade, que $\mathbf{v} = \mathbf{0}$. Portanto, $W_1 \cap W_2 = \{\mathbf{0}\}$ e $V = W_1 \oplus W_2$. 7. $W_2 = \{(x, y) \in F^2 : y = mx\}$, para todo $m \in F - \{1\}$. 8. Confira Exemplo 3.20. 9. Confira Exemplo 3.2. 10. Suponhamos que $W_1 \cup W_2$ seja um subespaço de V, mas $W_1 \nsubseteq W_2$ e $W_2 \nsubseteq W_1$. Então existe um $\mathbf{u} \in W_1$ tal que $\mathbf{u} \notin W_2$ e existe um $\mathbf{v} \in W_2$ tal que $\mathbf{v} \notin W_1$. Como $\mathbf{u}, \mathbf{v} \in W_1 \cup W_2$ temos que $\mathbf{u} + \mathbf{v} \in W_1 \cup W_2$. Assim, $\mathbf{u} + \mathbf{v} \in W_1$ ou $\mathbf{u} + \mathbf{v} \in W_2$.

Se $\mathbf{u}+\mathbf{v} \in W_1$, então $\mathbf{v} = -\mathbf{u} + (\mathbf{u}+\mathbf{v}) \in W_1$, o que é impossível. Se $\mathbf{u}+\mathbf{v} \in W_2$, então $\mathbf{u} = -\mathbf{v} + (\mathbf{u}+\mathbf{v}) \in W_2$, o que é impossível. Portanto, $W_1 \subseteq W_2$ ou $W_2 \subseteq W_1$. Reciprocamente, se $W_1 \subseteq W_2$, então $W_1 \cup W_2 = W_2$ é um subespaço de V. 11. (a) Basta notar que $(W_1 \cap W_2) \subseteq W_1$ e $(W_1 \cap W_3) \subseteq W_2 + W_3$. (c) Pelo Exercício (6), $W_1 + W_2 = \{(0,0,0)\}$, $W_1 \cap W_3 = \{(x,-2x,x) \in V : x \in F\}$ e $W_2 + W_3 = V$ implicam que $(W_1 \cap W_2) + (W_1 \cap W_3) = \{(x,-2x,x) \in V : x \in F\} \subset W_1 = W_1 \cap (W_2 + W_3)$. (d) Se $W_3 \subseteq W_1$, então $W_2 + W_3 \subseteq W_2 + W_1$, de modo que $W_1 \cap (W_2 + W_3) \subseteq (W_1 \cap W_2) + (W_1 \cap W_3)$. Portanto, $(W_1 \cap W_2) + (W_1 \cap W_3) = W_1 \cap (W_2 + W_3)$. 12. (a) Use $U = \{(x,y,z) \in V : y = 0\}$ no Exemplo 3.19. (b) Suponhamos que $W_1 \subseteq U$. Dado $\mathbf{v} \in U \subseteq V$, temos, pelo Exercício (7), que existe único $\mathbf{u}_1 \in W_1$ e $\mathbf{u}_2 \in W_2$ tal que $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$. Assim, $\mathbf{u}_2 = \mathbf{u} - \mathbf{u}_1 \in U$. Como $\mathbf{u}_1 \in U \cap W_1$ e $\mathbf{u}_2 \in U \cap W_2$ temos que $\mathbf{u} \in (U \cap W_1) \oplus (U \cap W_2)$.

Seção 3.3

1. Note que $(x,y) = \frac{y}{2}(1,2) + \frac{2x-y}{10}(5,0)$ e $F^2 = F[(1,2),(5,0)]$. 2. Observe que $f(x)x_1 + g(x)x_2 = p(x)$ é equivalente a escalonar a matriz

$$\left(\begin{array}{rr|rr} 2 & -8 & -26 & 1 \\ -3 & 5 & 11 & 1 \\ 5 & -2 & 7 & 1 \end{array}\right) \to \cdots \to \left(\begin{array}{cc|cc} 1 & 0 & 0 & \frac{13}{152} \\ 0 & 1 & 0 & \frac{7}{38} \\ 0 & 0 & 1 & \frac{23}{152} \end{array}\right).$$

Portanto, nenhum dos vetores é combinação linear. 3. Nenhum dos vetores é combinação linear. 4. Como

$$(a,b,c) = 14^{-1}(4a+2b-3c)\mathbf{u}_1 + 7^{-1}(a+b+c)\mathbf{u}_2 + 14^{-1}(10a-4b+3c)\mathbf{u}_3$$

temos que $V = F[\alpha]$. Portanto, $(4,-5,k)$ pertence, para todo $k \in F$. 5. Basta notar que $V = F[\alpha]$. 6. Observe que $\mathbf{A} = -3\mathbf{A}_1 + 2\mathbf{A}_2 - \mathbf{A}_3$. Mas, os vetores $\mathbf{B}$ e $\mathbf{C}$ não são combinações lineares dos vetores da lista α, pois

$$\left(\begin{array}{rrr|rrr} 1 & 3 & -1 & 4 & 3 & -1 \\ 1 & 0 & 5 & -5 & 1 & 1 \\ -2 & 4 & 2 & 9 & -4 & -2 \\ 1 & -1 & 2 & -7 & 4 & 1 \end{array}\right) \to \cdots \to \left(\begin{array}{ccc|rcr} 1 & 0 & 0 & -3 & 0 & \frac{107}{129} \\ 0 & 1 & 0 & 2 & 0 & -\frac{18}{43} \\ 0 & 0 & 1 & -1 & 0 & \frac{20}{129} \\ 0 & 0 & 0 & 0 & 1 & \frac{6}{43} \end{array}\right).$$

7. (a) $W_1 = F[(1,1,0),(0,0,1)]$. (b) $W_2 = F[(2,1,-2)$. (c) $W_3 = F[(-2,1,0),(3,0,1)]$. (d) Determinar $W_1 \cap W_2$ é equivalente a resolver o sistema de equações lineares $x - y = 0, x + z = 0$ e $x - 2y = 0$. Portanto, $W_1 \cap W_2 = \{(0,0,0)\}$. (e) Determinar $W_1 + W_3 = [W_1 \cup W_3] = F[(2,1,-2),(-2,1,0),(3,0,1)]$ é equivalente a

escalonar a matriz

$$\begin{pmatrix} 2 & -2 & 3 \\ 1 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Portanto, $W_1 + W_2 = F[(2,1,-2),(-2,1,0),(3,0,1)] = F^3$ 8. Como $W = F[(-2,1,0,0),(2,0,1,0)]$, basta escalonar a matriz

$$\left(\begin{array}{cc|ccc} -2 & 2 & -2 & 6 & -2 \\ 1 & 0 & 4 & 2 & 1 \\ 0 & 1 & 3 & 4 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right) \to \cdots \to \left(\begin{array}{cc|ccc} 1 & 0 & 4 & 0 & 1 \\ 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Portanto, $\mathbf{u} = (-2,4,3,0), \mathbf{w} = (-2,1,0,0) \in W$, mas $\mathbf{v} = (6,2,4,1) \notin W$. 9. É fácil verificar que $J(\mathbf{z} + c\mathbf{w}) = J(\mathbf{z}) + \bar{c}J(\mathbf{w})$ e $J^2(\mathbf{z}) = \mathbf{z}$, para todos $\mathbf{z}, \mathbf{w} \in \tilde{V}$ e $c \in \mathbb{C}$. Além disso, $J(V) \subseteq V$. Se $W = \tilde{U}$ e $\mathbf{z} \in W$, então existem $\mathbf{u}, \mathbf{v} \in U$ tais que $\mathbf{z} = \mathbf{u} + i\mathbf{v}$. Assim, $\mathbf{w} = \mathbf{u} - i\mathbf{v} \in W$ e $J(\mathbf{z}) = \mathbf{w} \in W$, de modo que $J(W) \subseteq W$. Reciprocamente, se $J(W) \subseteq W$, então $J_1 : W \to W$ definida como $J_1(\mathbf{z}) = J(\mathbf{z})$ é uma conjugação. Portanto, $U = \{\mathbf{z} \in W : J_1(\mathbf{z}) = \mathbf{z}\}$ possui as propriedades desejadas, por exemplo, para qualquer $\mathbf{z} \in W$, obtemos

$$\mathbf{z} = 2^{-1}(\mathbf{z} + J_1(\mathbf{z})) + i(2i)^{-1}(\mathbf{z} - J_1(\mathbf{z})) \in \tilde{U},$$

pois $J_1(2^{-1}(\mathbf{z} + J_1(\mathbf{z}))) = 2^{-1}(\mathbf{z} + J_1(\mathbf{z}))$ etc. 10. Seja W um subespaço minimal de V. Então $W \neq V$ e existe um $\mathbf{u} \in W$, com $\mathbf{u} \neq \mathbf{0}$. Assim, $\{\mathbf{0}\} \subset F[\mathbf{u}] \subseteq W$. Portanto, $W = F[\mathbf{u}]$.

Seção 3.4

1. Confira a Desigualdade de Cauchy-Schwarz. 2. (a) Note que discutir se β é LI é equivalente escalonar a matriz

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ -2 & -1 & 1 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Portanto, β é LI. (b) Como $2(\mathbf{v}+\mathbf{w}) = -(\mathbf{u}+\mathbf{v}-3\mathbf{w})+(\mathbf{u}+3\mathbf{v}-\mathbf{w})$ temos que γ é LD. 3. Confira Exercício (1). 4. Como $2+x+2x^2$ é uma combinação linear dos precedentes temos que α é LD. Podemos afirmar que é LI. 5. O vetor $x(1,2,3) + y(1,-1,1)$

pertence a W_1 se, e somente se, $3x + y = 0$. Portanto, $\mathbf{u} = (1, -3, 0)$. 6. $k^2 - k \neq 0$. 7. Para resolver este problema confira primeiro o Exemplo 3.28. (a) Sim. (b) Não. (c) Sim. (d) Não. (e) Não. (f) Sim. (g) Note que a equação $a \operatorname{sen} x + b \cos x + c \operatorname{tg} x = \mathbf{0}$ vale para todo $x \in \mathbb{R}$. Em particular, para $x = 0, x = \frac{\pi}{4}$ e $x = \frac{\pi}{3}$ implicam que $b = 0, \sqrt{2}a + 2c = 0$ e $a + 2c = 0$, de modo que $a = b = c = 0$. Portanto, a lista é LI. 8. Já vimos que $F[\alpha] = F[\beta]$ e

$$\sum_{i \neq k} y_i \mathbf{u}_i + y_k \mathbf{u}_k = \sum_{i \neq k} (y_i - x_i x_k^{-1} y_k) \mathbf{u}_i + (x_k^{-1} y_k) \mathbf{u}.$$

9. Vamos usar indução sobre m. Se $m = 1$, nada há para ser provado. Suponhamos que o resultado seja válido para todo k, com $1 \leq k < m$. Então

$$\mathbf{v}_m = \sum_{i<m} x_i \mathbf{v}_i + \sum_{i \geq m} x_i \mathbf{u}_i.$$

Assim, $x_i \neq 0$, para algum $i = 1, \ldots, n$, pois $\mathbf{v}_m$ é LI. Como os $\mathbf{v}_1, \ldots, \mathbf{v}_m$ são LI temos que $x_i \neq 0$, para algum $i \geq m$, digamos $i = m$. Portanto, $\mathbf{u}_m$ pode ser substituído por $\mathbf{v}_m$.

Seção 3.5

1. $VVFF$. 2. Não, confira o Exemplo 3.46. 3. Note que $\dim W_1 = 1$ e $W_1 \not\subseteq W_2$ implicam que $W_1 \cap W_2 = \{(0, 0, 0)\}$, de modo que $\dim(W_1 + W_2) = 3$. Portanto, $V = W_1 + W_2$ e $V = W_1 \oplus W_2$. 4. Como $\dim(W_1 \cap W_2) \leq \min\{4, 5\}$ e $7 \geq \dim(W_1 + W_2) = 4 + 5 - \dim(W_1 \cap W_2)$ temos que $2 \leq \dim(W_1 \cap W_2) \leq 4$. Portanto, $\dim(W_1 \cap W_2) = 2, 3$ ou 4. 5. Confira o Exemplo 3.37. 6. Confira o Exemplo 3.45. 7. Confira o Exemplo 3.45. 8. Se $p(x) = a + bx + cx^2 + dx^3$, então $p'(x) = b + 2cx + 3dx^2 = 0$ implica que $b = c = d = 0$ e $a \in F$ qualquer. Portanto, $\alpha = \{1\}$ é uma base de W e $\dim W = 1$. 9. Pelo Exercício (3) da Seção , se $(1 - a)^2 = (1 + a)^2$, de modo que $a = 0$. 10. Como $p'(x) = 4x - 3$ e $p''(x) = 4$ temos que $ap(x) + bp'(x) + cp''(x) = 0$ implica que $a = b = c = 0$. Portanto, α é uma base de V. 11. Derivando três vezes a equação

$$\begin{array}{rcl} a(1-x)^3 + b(1-x)^2 + c(1-x) + d & = & 0 \\ -3a(1-x)^2 - 2b(1-x) - c & = & 0 \\ 6a(1-x) + 2b & = & 0 \\ -6a & = & 0 \end{array}$$

temos, pela substituição reversa, que $a = b = c = d = 0$. Portanto, α é uma base de

V. 12. Confira o Exemplo 3.37. 13. Confira o Exemplo 3.43. 14. (a) É claro que se $\mathbf{Q}, \mathbf{R} \in W$ e $a \in F$, então $\mathbf{Q} + a\mathbf{R} \in W$. Observe que

$$a\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} + b\begin{pmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} + c\begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} = \mathbf{Q},$$

com

$$\mathbf{Q} = \begin{pmatrix} a+b & a-b+c & a-c \\ a-b-c & a & a+b+c \\ a+c & a+b-c & a-b \end{pmatrix} \in W.$$

Portanto, $W = F[\alpha]$. (b) Pelo item (a), o sistema $a+b = a_1, a-b+c = a_2$ e $a-c = a_3$ possui uma solução, para todos $a_1, a_2, a_3 \in F$. 15. Note que

$$\dim(W_1 + W_2 + W_3) = \dim(W_1 + W_2) + \dim W_3 - \dim((W_1 + W_2) \cap W_3)$$

e use o Exercício (12) da Seção 3.2. 16. Se $a\mathbf{u} + bi\mathbf{u} = \mathbf{0}$, então $(a + bi)\mathbf{u} = \mathbf{0}$, de modo que $a + bi = 0$, ou seja, $a = b = 0$. 17. Seja $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ uma base de V. Então devemos provar que sua complexificação

$$\lambda(\alpha) = \{\mathbf{u}_1 + i\mathbf{0}, \ldots, \mathbf{u}_n + i\mathbf{0}\}$$

é uma base de $\tilde{V}$. Dado $\mathbf{z} \in \tilde{V}$ existem, pelo Exercício (12) da Seção 3.1, $\mathbf{u}, \mathbf{v} \in V$ tais que $\mathbf{z} = \mathbf{u} + i\mathbf{v}$. Como $\mathbf{u}, \mathbf{v} \in V$ temos que existem únicos x_j, y_k tais que $\mathbf{u} = \sum_{j=1}^{n} x_j\mathbf{u}_j$ e $\mathbf{v} = \sum_{k=1}^{n} y_k\mathbf{u}_k$. Assim, $\mathbf{z} = \mathbf{u} + i\mathbf{v} = \sum_{j=1}^{n}(x_j + iy_j)\mathbf{u}_j$. Portanto, $\lambda(\alpha)$ é uma base de $\tilde{V}$ sobre $\mathbb{C}$ e $\dim \tilde{V} = \dim V$. Note que $\beta = \alpha \cup (i\alpha)$ é uma base de $\tilde{V}$ sobre $\mathbb{R}$ e $\dim \tilde{V} = 2\dim V$. 18. $(a \Rightarrow b)$ Suponhamos que $V \neq F[\alpha]$. Então existiria um $\mathbf{u} \in V$ tal que $\mathbf{u} \notin F[\alpha]$. Assim, de modo análogo a prova do Lema 3.40, o conjunto $\beta = \alpha \cup \{\mathbf{u}\}$ seria LI, com $\alpha \subset \beta$, o que é impossível. $(b \Rightarrow c)$ Suponhamos que α não seja LI. Então existiria um conjunto $\beta = \alpha - \{\mathbf{u}\}$ tal que $V = F[\beta]$, com $\beta \subset \alpha$, o que é impossível. $(c \Rightarrow a)$ Tente provar! 19. (b) Por exemplo,

$$\begin{aligned} (\mathbf{u} + W) \oplus [(\mathbf{v} + W) \oplus (\mathbf{w} + W)] &= \mathbf{u} + W \oplus (\mathbf{v} + \mathbf{w} + W) \\ &= \mathbf{u} + (\mathbf{v} + \mathbf{w}) + W \\ &\overset{\text{em } V}{=} (\mathbf{u} + \mathbf{v}) + \mathbf{w} + W \\ &= (\mathbf{u} + \mathbf{v} + W) \oplus (\mathbf{w} + W) \\ &= [(\mathbf{u} + W) \oplus (\mathbf{v} + W)] \oplus (\mathbf{w} + W). \end{aligned}$$

(c) Se $\mathbf{u} \in \alpha \subseteq W$, então $\mathbf{u} + W = \mathbf{0} + W = W$ é o vetor nulo de $\overline{V}$, de modo que $\mathbf{u} + W \notin \gamma$. Assim, $\mathbf{u} \notin \beta$ e $\alpha \cap \beta = \emptyset$. Dado $\mathbf{v} \in V$, temos que $\mathbf{v} + W \in \overline{V}$. Logo, existem vetores distintos $\mathbf{v}_1 + W, \ldots, \mathbf{v}_n + W \in \gamma$ e $y_1, \ldots, y_n \in F$ tais que $\mathbf{v} + W = \sum_{j=1}^{n} y_j \mathbf{v}_j + W$ se, e somente se, $\mathbf{v} - \sum_{j=1}^{n} y_j \mathbf{v}_j \in W$, de modo que existem vetores distintos $\mathbf{u}_1, \ldots, \mathbf{u}_m \in \alpha$ e $x_1, \ldots, x_m \in F$ tais que $\mathbf{v} - \sum_{j=1}^{n} y_j \mathbf{v}_j = \sum_{i=1}^{m} x_i \mathbf{u}_i$, ou seja, $\mathbf{v} = \sum_{i=1}^{m} x_i \mathbf{u}_i + \sum_{j=1}^{n} y_j \mathbf{v}_j$. Portanto, $V = F[\alpha \cup \beta]$. Observe que $\sum_{i=1}^{m} x_i \mathbf{u}_i + \sum_{j=1}^{n} y_j \mathbf{v}_j = \mathbf{0}$ implica que $W = \sum_{j=1}^{n} y_j \mathbf{v}_j + W$. Assim, $y_1 = \cdots = y_n = 0$, pois γ é LI. Logo, $\sum_{i=1}^{m} x_i \mathbf{u}_i = \mathbf{0}$ implica que $x_1 = \cdots = x_m = 0$, pois α é LI. Consequentemente, $\alpha \cup \beta$ é LI. (d) Primeiro note que $V = W \cup (V - W)$. Assim, para cada $\mathbf{v} \in V$, podemos escolher vetores distintos $\mathbf{u}_1, \ldots, \mathbf{u}_m \in \alpha$, $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \beta - \alpha$ e $x_1, \ldots, x_m, y_1, \ldots, y_n \in F$ tais que $\mathbf{v} = \sum_{i=1}^{m} x_i \mathbf{u}_i + \sum_{j=1}^{n} y_j \mathbf{v}_j$. Logo, $\mathbf{v} + W = \sum_{j=1}^{n} y_j \mathbf{v}_j + W$. Portanto, $\overline{V} = F[\gamma]$. Prove que γ é LI! (e) Basta notar que $\dim V = |\alpha \cup \beta| = |\alpha| + |\beta| = \dim W + \dim \overline{V}$. (f) Dado $\mathbf{v} + U \in \overline{W_1 + W_2}$, existem $\mathbf{w}_1 \in W_1$ e $\mathbf{w}_2 \in W_2$ tais que $\mathbf{v} + U = \mathbf{w}_1 + \mathbf{w}_2 + U = (\mathbf{w}_1 + U) \oplus (\mathbf{w}_2 + U) \in \overline{W_1} + \overline{W_2}$, ou seja, $\overline{W_1 + W_2} = \overline{W_1} + \overline{W_2}$. Se $\mathbf{v} + U \in \overline{W_1} \cap \overline{W_2}$, então existem $\mathbf{w}_1 \in W_1$ e $\mathbf{w}_2 \in W_2$ tais que $\mathbf{v} + U = \mathbf{w}_1 + U$ e $\mathbf{v} + U = \mathbf{w}_2 + U$. Assim, $\mathbf{w}_2 - \mathbf{w}_1 = (\mathbf{w}_2 - \mathbf{v}) + (\mathbf{w}_1 - \mathbf{v}) \in U$, de modo que $\mathbf{w}_2 = (\mathbf{w}_2 - \mathbf{w}_1) + \mathbf{w}_1 \in W_1$ e $\mathbf{v} + U = (\mathbf{v} - \mathbf{w}_2) + \mathbf{w}_2 + U = U$. Portanto, $\overline{W_1} \cap \overline{W_2} = \{U\}$. Consequentemente, $\dim(\overline{W_1 + W_2}) = \dim \overline{W_1} + \dim \overline{W_2}$.

Seção 3.6

1. Confira o Exemplo 3.47, $[(x, y)]_\alpha = 3^{-1}(x + y, x - y)^t$. 2. Primeiro confira o Exemplo 3.47. (a) $[(x, y)]_\alpha = \frac{1}{3}(x+y, x-y)^t$. (b) $[(x, y)]_\alpha = (\frac{x}{2}, -y)^t$. (c) $[(x, y)]_\alpha = (x, y)^t$. (d) $[(x, y)]_\alpha = \frac{1}{3}((2x - y, -x + 2y)$. 3. Observe que os vetores $\mathbf{u} = (a, b)$ e $\mathbf{v} = (c, d)$ estão sobre o círculo $x^2 + y^2 = 1$. Assim, existem $\theta, \phi \in [0, 2\pi)$ tais que $a = \cos\theta, b = \operatorname{sen}\theta$ e $c = \cos\phi, d = \operatorname{sen}\phi$. Logo,

$$\cos(\phi - \theta) = \cos\phi\cos\theta + \operatorname{sen}\phi\operatorname{sen}\theta = ac + bd = 0,$$

de modo que $\phi - \theta = \frac{\pi}{2}$ ou $\phi - \theta = \frac{3\pi}{2}$. Portanto, $c = -\operatorname{sen}\theta$ e $d = \cos\theta$ ou $c = \operatorname{sen}\theta$ e $d = -\cos\theta$. Consequentemente, $\beta = \{\mathbf{u}, \mathbf{v}\}$ é uma base de V, pois $ad - bc = 1$. Se $\mathbf{u}$ estiver no primeiro quadrante, então $\mathbf{u} = \cos\theta\mathbf{e}_1 + \operatorname{sen}\theta\mathbf{e}_2$ e $\mathbf{v} = -\operatorname{sen}\theta\mathbf{e}_1 + \cos\theta\mathbf{e}_2$. Assim,

$$[I]_\alpha^\beta = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \text{ e } ([I]_\alpha^\beta)^{-1} = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix},$$

com α a base canônica de V. Portanto,

$$[(x,y)]_\beta = ([I]_\alpha^\beta)^{-1}[(x,y)]_\alpha = \begin{pmatrix} x\cos\theta + y\,\mathrm{sen}\,\theta \\ -x\,\mathrm{sen}\,\theta + y\cos\theta \end{pmatrix}.$$

Compare este exercício com o Exemplo 3.50. 4. Confira o Exemplo 3.48. 5. e 6. Confira o Exemplo 3.49. 7. e 8. Confira o Exemplo 3.51. 9. e 10. Note que

$$[I]_\alpha^\beta = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix} \text{ e } [I]_\beta^\alpha = ([I]_\alpha^\beta)^{-1}.$$

11. Confira o Exercício (7). 12., 13. e 14. Confira o Exercício (9). 15. $[I]_\alpha^\alpha = \mathbf{I}$. 16. Confira o Exercício (9) identificando $\mathbf{E}_{11} \leftrightarrow (1,0,0,0)$ etc. 17. Se $\beta = (\mathbf{u}_1\mathbf{A}, \ldots \mathbf{u}_n\mathbf{A})$ for uma base de $F^{1\times n}$, então os $\mathbf{u}_i$ são combinações lineares dos $\mathbf{u}_i\mathbf{A}$ e vice-versa, digamos $\alpha^t = \beta^t\mathbf{B}$, Assim, $\beta^t = \alpha^t\mathbf{A} = (\beta^t\mathbf{B})\mathbf{A}$, de modo que $\mathbf{BA} = \mathbf{I}$. Portanto, $\mathbf{A}$ é não singular. Reciprocamente, se $\mathbf{A}$ for não singular, então $\beta^t = \alpha^t\mathbf{A}$ implica que $\alpha^t = \beta^t\mathbf{A}^{-1}$, de modo que $F^{1\times n} = F[\beta]$. Portanto, β é uma base de $F^{1\times n}$. 18. Dados $\mathbf{u}, \mathbf{v} \in W$ e $a \in F$, obtemos $\mathbf{u}, \mathbf{v} \in W_n$, para todo $n \in \mathbb{N}$. Assim, $\mathbf{u} + a\mathbf{v} \in W_n$, para todo $n \in \mathbb{N}$. Portanto, $\mathbf{u} + a\mathbf{v} \in W$ e W é um subespaço de V. 19. Dados $\mathbf{u}, \mathbf{v} \in W$ e $a \in F$, existem $m, n \in \mathbb{N}$ tais que $\mathbf{u} \in W_m$ e $\mathbf{v} \in W_n$. Se $m \leq n$, então $\mathbf{u}, \mathbf{v} \in W_n$. Assim, $\mathbf{u} + a\mathbf{v} \in W_n$, para algum $n \in \mathbb{N}$. Portanto, $\mathbf{u} + a\mathbf{v} \in W$ e W é um subespaço de V. 20. (a) $\alpha = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2)$. (b) Suponhamos, por absurdo, que $x_i = 0$, para algum i, digamos $x_1 = 0$. Então $x_2\mathbf{u}_2 + \cdots + x_n\mathbf{u}_n = \mathbf{0}$, de modo que $x_2 = \cdots = x_n = 0$, o que contradiz o fota de α ser LD. Portanto, $x_i \neq 0$, para todo i. (c) Podemos supor que $y_1 \neq 0$ e, pelo item (b), $\mathbf{u}_1 = (-x_1^{-1}x_2)\mathbf{u}_2 + \cdots + (-x_1^{-1}x_n)\mathbf{u}_n$. Assim, substituindo na equação, obtemos $(y_2 - y_1x_1^{-1}x_2)\mathbf{u}_2 + \cdots + (y_n - y_1x_1^{-1}x_n)\mathbf{u}_n = \mathbf{0}$. Logo, por hipótese, $y_i = ax_i$, onde $a = x_1^{-1}y_1 \in F^\times$ e $i = 1, \ldots n$. (d) É fácil verificar que W é um subespaço de F^n e, pelo item (c), $W = F[x_1, \ldots, x_n]$. É importante observar que $W = L_\alpha^{-1}(\mathbf{0})$ e $\operatorname{Im} L_\alpha = F[\alpha - \{\mathbf{u}_i\}]$, para algum $i = 1, \ldots n$. 21. Suponhamos que

$$\mathbf{B} = \sum_{i=1}^n \sum_{j=1}^n a_{ij}\mathbf{F}_{ij} = \sum_{i=1}^n \sum_{j=1}^n a_{ij}\mathbf{E}_{ij} + \sum_{i=1}^n \sum_{j=1}^n a_{ij}\mathbf{I} = \mathbf{O}.$$

Se $p \neq q$, então $b_{pq} = a_{pq} + 0 = 0$. Se $p = q$, então $b_{pp} = a_{pp} + \sum_{i=1}^n \sum_{j=1}^n a_{ij} = 0$. Como $a_{ij} = 0$, quando $i \neq j$, temos que $a_{pp} + \sum_{i=1}^n a_{ii} = 0$, de modo que $a_{pp} = 0$, pois os a_{pp} são todo iguais. Portanto, α é uma base de V. 22. É fácil verificar que $F_f(x)$ é um subespaço de $\mathcal{F}(F,F)$ e que a lista

$$\alpha = \left(\frac{1}{(x-a)^m(x-b)^n}, \frac{x}{(x-a)^m(x-b)^n}, \ldots, \frac{x^{m+n-1}}{(x-a)^m(x-b)^n} \right)$$

é uma base de $F_f(x)$ (Use derivação, se necessário, para o caso LI).

Capítulo 4

Seção 4.1

1. (a) Sim. (b) Não, pois $T(0,0,0) = (-1,0) \neq (0,0)$. (c) Sim. (d) Não, pois $T(2\mathbf{e}_1 + \mathbf{e}_2) = (1,8) \neq (1.2) = 2T(\mathbf{e}_1) + T(\mathbf{e}_2)$. (e) Sim. Conclua que qualquer operador linear sobre F^2 é dessa forma. Por exemplo, $T(\mathbf{e}_1) = a\mathbf{e}_1 + c\mathbf{e}_2 = (a,c)$, onde $a, b \in F$. 2. (a) Sim. (b) Sim. (c) Não. pois $T(\mathbf{O}) = \mathbf{B} \neq \mathbf{O}$. (d) Sim. (e)
Sim. 3. Como um exemplo vamos provar apenas o item (d) Dados $f, g \in V$ e $a \in \mathbb{R}$, obtemos

$$[T(f+cg)](x) = \frac{1}{2}[(f+cg)(x+\frac{h}{2}) - (f+cg)(x-\frac{h}{2})] = (Tf + aTg)(x).$$

(4) Direto da definição. 5. $T(x,y) = (x+ay, y)$ e $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida como $T(1,0) = (1,a)$ e $T(0,1) = (0,1)$, onde $a \in \mathbb{R}^\times$. 6. $T(x,y) = (-x+y, x)$. 7. $T(x,y) = (ax+cy, bx+dy)$. 8. (a) Dados $p(x), q(x) \in V$ e $a \in \mathbb{R}$, obtemos

$$T(p(x) + aq(x)) = x(p(x) + aq(x)) = xp(x) + c(xq(x)) = T(p(x)) + aT(q(x)).$$

9. Dados $\mathbf{u}, \mathbf{v} \in V$ e $c \in F$, obtemos

$$(T+S)(\mathbf{u} + c\mathbf{v}) = T(\mathbf{u} + c\mathbf{v}) + S(\mathbf{u} + c\mathbf{v}) = (T+S)(\mathbf{u}) + c(T+S)(\mathbf{v}).$$

Portanto, $T+S, aT \in \mathcal{L}(V,W)$ e é um espaço vetorial sobre F. 10. Sejam $\alpha = \{\mathbf{v}_1, \mathbf{v}_2\}$ uma base de V e $\beta = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ uma base de W. Então, pelo Teorema 4.11, existem únicas transformações lineares $E_{ij} : V \to W$, com $i = 1,2,3$ e $j = 1,2$, definidas como $E_{ij}(\mathbf{v}_k) = \delta_{jk}\mathbf{w}_i$, para $k = 1,2$. Então os E_{ij} agem sobre um vetor $\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2$ como $E_{ij}(\mathbf{v}) = x_1E_{ij}(\mathbf{v}_1) + x_2E_{ij}(\mathbf{v}_2) = x_1\delta_{j1}\mathbf{w}_i + x_2\delta_{j2}\mathbf{w}_i = x_j\mathbf{w}_i$. **Afirmação**. $\gamma = \{E_{11}, E_{12}, E_{21}, E_{22}, E_{31}, E_{32}\}$ é uma base de $\mathcal{L}(V,W)$. De fato. Dado $T \in \mathcal{L}(V,W)$ tal que $T(\mathbf{v}_k) = \sum_{i=1}^{3} a_{ik}\mathbf{w}_i$, $k = 1,2$. Então, considerando

$S = \sum_{i=1}^{3} \sum_{j=1}^{2} a_{ij} E_{ij} \in \mathcal{L}(V, W)$, obtemos

$$S(\mathbf{v}_k) = (\sum_{i=1}^{3} \sum_{j=1}^{2} a_{ij} E_{ij})(\mathbf{v}_k) = \sum_{i=1}^{3} \sum_{j=1}^{2} a_{ij} \delta_{jk} \mathbf{w}_i = \sum_{i=1}^{3} a_{ik} \mathbf{w}_i = T(\mathbf{v}_k).$$

Portanto, $T = S$ e $\mathcal{L}(V, W) = F[\gamma]$. Note que $\dim L(V, W) = 6$. 11. e 12. Direto das operações usuais e da composição de funções. 13. Como

$$\begin{aligned} (DM - MD)(p(x)) &= D(M(p(x))) - M(D(p(x))) \\ &= p(x) + xp'(x) - xp'(x) = I(p(x)), \ \forall\, p(x) \in \mathbb{R}[x], \end{aligned}$$

temos que $DM - MD = I$. Note que $DM = D(DM - MD)M$. 14. $(a \Rightarrow b)$ Sejam $\mathbf{x} = f(\mathbf{0}) \in W$ e $T : V \to W$ definida como $T(\mathbf{y}) = f(\mathbf{y}) - \mathbf{x}$. Então T é linear. De fato. Como $\mathbf{y} - c\mathbf{y} = (c-1)(\mathbf{0} - \mathbf{y})$ temos que

$$T(\mathbf{y}) - T(c\mathbf{y}) = f(\mathbf{y}) - f(c\mathbf{y}) = (c-1)[f(\mathbf{0}) - f(\mathbf{y})] = (c-1)(-T(\mathbf{y})).$$

Assim, $T(c\mathbf{y}) = cT(\mathbf{y})$, para todo $\mathbf{y} \in V$. Sendo $2\mathbf{z} - (\mathbf{y} + \mathbf{z}) = \mathbf{z} - \mathbf{y} = -2^{-1}(2\mathbf{y} - 2\mathbf{z})$, obtemos

$$\begin{aligned} 2T(\mathbf{z}) - T(\mathbf{y} + \mathbf{z}) &= T(2\mathbf{z}) - T(\mathbf{y} + \mathbf{z}) = f(2\mathbf{z}) - f(\mathbf{y} + \mathbf{z}) \\ &= -2^{-1}[f(2\mathbf{y}) - f(2\mathbf{z})] = -[T(\mathbf{y}) - T(\mathbf{z})]. \end{aligned}$$

Portanto, $T(\mathbf{y} + \mathbf{z}) = T(\mathbf{y}) + T(\mathbf{z})$, para todos $\mathbf{y}, \mathbf{z} \in V$. $(b \Rightarrow c)$ Direto da linearidade de T. $(c \Rightarrow a)$. Basta notar que $\mathbf{w} - \mathbf{u} = c(\mathbf{v} - \mathbf{w})$ implica que $\mathbf{u} = (-c)\mathbf{v} + (1 + c)\mathbf{w}$. 15. (a) Suponhamos que

$$x_0 \mathbf{u} + x_1 T(\mathbf{u}) + \cdots + x_{k-1} T^{k-1}(\mathbf{u}) = \mathbf{0}.$$

Como $T^{n+k} = 0$, para todo $n \in \mathbb{Z}_+$ temos, aplicando recursivamente T^{k-i}, com $i = 1, \ldots, k-1$, na equação, que $x_0 = x_1 = \cdots = x_{k-1} = 0$. Portanto α é LI.

Seção 4.2

1. Dados $\mathbf{w}, \mathbf{z} \in T(U)$ e $a \in F$, existem $\mathbf{u}, \mathbf{v} \in U$ tais que $\mathbf{w} = T(\mathbf{u})$ e $\mathbf{z} = T(\mathbf{v})$. Assim,

$$\mathbf{w} + a\mathbf{z} = T(\mathbf{u}) + aT(\mathbf{v}) = T(\mathbf{u} + a\mathbf{v}) \in T(U),$$

pois $\mathbf{u} + a\mathbf{v} \in U$. Portanto, $T(U)$ é um subespaço de W. 2. Dados $\mathbf{u}, \mathbf{v} \in T^{-1}(Z)$ e

$a \in F$, obtemos $T(\mathbf{u}), T(\mathbf{v}) \in Z$, de modo que $T(\mathbf{u}) + aT(\mathbf{v}) \in Z$ implica que

$$T(\mathbf{u} + a\mathbf{v}) = T(\mathbf{u}) + aT(\mathbf{v}) \in Z.$$

Logo, $\mathbf{u} + a\mathbf{v} \in T^{-1}(Z)$. Portanto, $T^{-1}(Z)$ é um subespaço de V. 3. Note que $(r, s) \in T(A)$ se, e somente se, existir um $(x, y) \in A$ tal que $T(x, y) = (r, s)$ se, e somente se, $x = r - s$ e $y = s$. Assim,

$$x^2 + y^2 = 1 \Leftrightarrow (r - s)^2 + s^2 = 1 \Leftrightarrow r^2 + 2s^2 - 2rs = 1$$

Portanto, $T(A) = \{(r, s) \in \mathbb{R}^2 : r^2 + 2s^2 - 2rs = 1\}$ representa uma elipse, isto é, T transforma círculo em elipse e vice-versa. 4. (a) $\ker T = \{(0, 0)\}$ e $\operatorname{Im} T = F[(-1, 0, 5), (1, 0, 0)]$. (b) $\ker T = F[(1, -1, 0)]$ e $\operatorname{Im} T = F[(1, 0), (1, 1)]$. 5. É claro que $F[T(\alpha)] \subseteq \operatorname{Im} T$. Por outro lado, dado $\mathbf{w} \in \operatorname{Im} T$, existe um $\mathbf{u} \in V$ tal que $\mathbf{w} = T(\mathbf{u})$. Como $V = F[\alpha]$ e $\mathbf{u} \in V$ temos que existem $x_1, \ldots, x_n \in F$ tais que $\mathbf{u} = \sum_{i=1}^n x_i \mathbf{u}_i$. Assim, $\mathbf{w} = T(\mathbf{u}) = \sum_{i=1}^n x_i T(\mathbf{u}_i) \in F[T(\alpha)]$, ou seja, $\operatorname{Im} T \subseteq F[T(\alpha)]$. 6. Note, pelo Exercício (5), que

$$\operatorname{Im} T = F[T(\mathbf{e}_1), T(\mathbf{e}_2), T(\mathbf{e}_3)] = F[(1, 2, -1), (-1, 1, -2), (2, 0, 2)].$$

Assim, discutir se $(a, b, c) \in \operatorname{Im} T$ é equivalente a escalonar a matriz

$$\left(\begin{array}{rrr|c} 1 & -1 & 2 & a \\ 2 & 1 & 0 & b \\ -1 & -2 & 2 & c \end{array}\right) \to \cdots \to \left(\begin{array}{rrr|c} 1 & 0 & \frac{2}{3} & \frac{1}{3}(b-a) \\ 0 & 1 & -\frac{4}{3} & \frac{1}{3}(b-2a) \\ 0 & 0 & 0 & -a+b+c \end{array}\right).$$

Portanto, $(a, b, c) \in \operatorname{Im} T$ se, e somente se, $a = b + c$, de modo que $\operatorname{Im} T = F[(1, 1, 0), (1, 0, 1)]$ e $\rho(T) = 2$. (b) Pelo item (a), $(a, b, c) \in \ker T$ se, e somente se, $a = -\frac{2c}{3}$ e $b = \frac{4c}{3}$, de modo que $\ker T = F[(-2, 4, 3)]$ e $\nu(T) = 1$. 7., 8. e 9. Confira os Exemplos 4.18 e 4.25, pois T ser sobrejetora e

$$T(1, 1, 0) = T(0, 0, 1) \Leftrightarrow T(1, 1, -1) = (0, 0, 0)$$

significa que $\ker T = F[(1, 1, -1)]$. 10. Sim, pois $\{(1, -1, 1), (1, 1, 1)\}$ é parte de uma base de F^3. 11. Não, pois $(-3, 2) = -(1, -1) - (2, -1)$ implica que

$$(1, 1) = T(-3, 2) = -T(1, -1) - T(2, -1) = -(1, 1),$$

o que é impossível. 12. Suponhamos que existam escalares $x_1, \ldots, x_k \in F$ tais que

$\sum_{i=1}^{k} x_i\mathbf{u}_i = \mathbf{0}$. Então

$$\sum_{i=1}^{k} x_i\mathbf{w}_i = \sum_{i=1}^{k} x_iT(\mathbf{u}_i) = T(\sum_{i=1}^{k} x_i\mathbf{u}_i) = T(\mathbf{0}) = \mathbf{0},$$

de modo que $x_1 = \cdots = x_n = 0$. Portanto, α é LI em V. Para cada $\mathbf{u} \in V$, sempre teremos $T(\mathbf{u}) \in \operatorname{Im} T$. Assim, existem escalares $y_1, \ldots, y_k \in F$ tais que $T(\mathbf{u}) = T(\sum_{i=1}^{k} y_i\mathbf{u}_i)$, de modo que $T(\mathbf{u} - \sum_{i=1}^{k} y_i\mathbf{u}_i) = \mathbf{0}$ implica que $\mathbf{u} - \sum_{i=1}^{k} y_i\mathbf{u}_i \in \ker T$. Logo, $V = F[\alpha] + \ker T$. É fácil verificar que $F[\alpha] \cap \ker T = \{\mathbf{0}\}$. Portanto, $V = F[\alpha] \oplus \ker T$. 13. Determine $\ker T$ resolvendo o sistema e use o Exercício (12). 14. (a) Discutir se $\mathbf{u}_1 \in F[\mathbf{u}_2, \mathbf{u}_3]$ é equivalente a saber se $(1, 0, 1) \in F[(-2, 1, 0), (-1, 1, 1)]$. 15. Dado $\mathbf{u} \in \ker(T \circ S)$, obtemos $T(S(\mathbf{u})) = \mathbf{0}$. Assim, $S(\mathbf{u}) \in \ker T = \{\mathbf{0}\}$, de modo que $S(\mathbf{u}) = \mathbf{0}$. Logo, $\mathbf{u} \in \ker S = \{\mathbf{0}\}$. Portanto, $\mathbf{u} = \mathbf{0}$ e $\ker(T \circ S) = \{\mathbf{0}\}$, ou seja, $\nu(T \circ S) = 0$. 16. (a) Dado $\mathbf{w} \in \operatorname{Im}(T \circ S)$, existe um $\mathbf{u} \in U$ tal que $\mathbf{w} = (T \circ S)(\mathbf{u}) = T(S(\mathbf{u}))$. Assim, $\mathbf{w} \in \operatorname{Im} T$, pois $S(\mathbf{u}) \in V$, de modo que $\operatorname{Im}(T \circ S) \subseteq \operatorname{Im} T$. Como $\operatorname{Im}(T \circ S)$ e $\operatorname{Im} T$ são subespaços de W temos, pelo Corolário 3.42, que

$$\rho(T \circ S) = \dim \operatorname{Im}(T \circ S) \leq \dim \operatorname{Im} T = \rho(T).$$

(b) Dado $\mathbf{u} \in \ker S$, obtemos $S(\mathbf{u}) = \mathbf{0}$, de modo que $(T \circ S)(\mathbf{u}) = \mathbf{0}$, ou seja, $\mathbf{u} \in \ker(T \circ S)$. Assim, de modo análogo ao item (a), $\nu(S) \leq \nu(S \circ T)$. 17. Como $\operatorname{Im} S \subseteq V$ temos que $m = \nu(S) + \rho(S)$ implica que $\nu(S) = m - \rho(S) \geq m - n > 0$. Assim, pelo item (b) do Exercício (16), $\nu(S \circ T) > 0$. Portanto, $T \circ S$ não é injetora. 18. Suponhamos que exista um isomorfismo $T \in \mathcal{L}(V, W)$. Então $\ker T = \{\mathbf{0}\}$ e $\operatorname{Im} T = W$. Assim, pelo Teorema 4.21, $\dim V = \dim W$, de modo que $m = n$. Reciprocamente, suponhamos que $m = n$. Então, pelo Teorema 4.28, existem isomorfismos $E_\alpha : F^n \to V$ e $E_\beta : F^n \to W$. Portanto, $T = E_\beta \circ E_\alpha^{-1} : V \to W$ é um isomorfismo, de modo que V e W são isomorfos. 19. (a) Se T fosse sobrejetora, então $\operatorname{Im} T = W$. Assim, pelo Teorema 4.21, $n = \dim \ker T + \dim \operatorname{Im} T \geq m$, o que é impossível. 20. Basta observar que $S \circ T = I_V$ implica que T é não singular. Portanto, T é bijetora. 21. Note que $T^2 - T + I = 0$ é equivalente a: $(I - T) \circ T = I$ e use o Exercício (20). 22. (a) Suponhamos que $\operatorname{Im} S \subseteq \operatorname{Im} T$. Então podemos escolher uma base $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ de U tal que $\{\mathbf{u}_{r+1}, \ldots, \mathbf{u}_k\}$ seja uma base de $\ker S$. Assim, $\{S(\mathbf{u}_1), \ldots, S(\mathbf{u}_r)\}$ é LI em W, mas isto significa que: existem $\mathbf{v}_1, \ldots, \mathbf{v}_r \in V$, necessariamente LI, tais que $S(\mathbf{u}_i) = T(\mathbf{v}_i)$, pois $\operatorname{Im} S \subseteq \operatorname{Im} T$. Logo, estendendo obtemos uma base $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ de V. Então a função $R : U \to V$ definida como

$R(\mathbf{u}_i) = \mathbf{v}_i$, com $i = 1, \ldots, r$, e $R(\mathbf{u}_j) = \mathbf{0}$, com $j = r+1, \ldots, n$, é tal que $S = T \circ R$. A recíproca é clara. (b) Suponhamos que $\ker T \subseteq \ker S$. Então podemos escolher uma base $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ de U tal que $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ seja uma base de $\ker S$ e $\{\mathbf{u}_1, \ldots, \mathbf{u}_s\}$ seja uma base de $\ker T$, com $s \leq r$. Assim, $\{T(\mathbf{u}_{s+1}), \ldots, T(\mathbf{u}_k)\}$ é LI em V. Logo, estendendo obtemos uma base $\{T(\mathbf{u}_{s+1}), \ldots, T(\mathbf{u}_k), \mathbf{v}_{k+1}, \ldots, \mathbf{v}_m\}$ de V. Então a função $R : V \to W$ definida como $R(T(\mathbf{u}_i)) = S(\mathbf{u}_i)$, com $i = s+1, \ldots, k$, e $R(\mathbf{v}_j) = \mathbf{0}$, com $j = k+1, \ldots, m$, é tal que $S = R \circ T$. A recíproca é clara. 23. (a) Como $\operatorname{Im}(T+S)$ é subespaço de V e $\operatorname{Im}(T+S) \subseteq \operatorname{Im} T + \operatorname{Im} S$ temos, pelo Corolário 3.42, que $\dim \operatorname{Im}(T+S) \leq \dim(\operatorname{Im} T + \operatorname{Im} S)$. Portanto, pelo Teorema 3.44, $\dim \operatorname{Im}(T+S) \leq \dim \operatorname{Im} T + \dim \operatorname{Im} S$, ou seja, $\rho(T+S) \leq \rho(S) + \rho(T)$. (b) Use o item (a) e $n = \nu(S+T) + \rho(S+T)$. (c) e (d) Confira o Exercício (13) e os itens anteriores. 24. Podemos imitar a prova do item (b) do Exercíco (22). Mas, vamos apresentar uma prova direta. Suponhamos que T seja injetora. Então $\ker T = \{\mathbf{0}\}$. Assim, pelo Teorema 4.21, $\dim V = \dim \operatorname{Im} T$, de modo que $\dim V = n \leq m$. Logo, para cada base $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ de V, $T(\alpha) = \{T(\mathbf{u}_1), \ldots, T(\mathbf{u}_n)\}$ é parte de uma base $\beta = \{T(\mathbf{u}_1), \ldots, T(\mathbf{u}_n), \mathbf{w}_{n+1} \ldots, \mathbf{w}_m\}$ de W. Portanto, pelo Teorema 4.11, existe uma única transformação linear $S : W \to V$ tal que $S(T(\mathbf{u}_i)) = \mathbf{u}_i$, se $i = 1, \ldots, n$, e $S(\mathbf{w}_i) = \mathbf{0}$, se $i = n+1, \ldots, m$. Consequentemente, $S \circ T = I_V$. Reciprocamente, suponhamos que $S \circ T = I_V$. Então, dado $\mathbf{u} \in \ker T$, obtemos $T(\mathbf{u}) = \mathbf{0}$, de modo que $\mathbf{u} = (S \circ T)(\mathbf{u}) = S(T(\mathbf{u})) = S(\mathbf{0}) = \mathbf{0}$. Portanto, $\ker T = \{\mathbf{0}\}$ e T é injetora. Observe que se T for não injetora, então, pelo Corolário 4.22, existem bases tais que $T(\mathbf{v}_i) = \mathbf{w}_i$, com $i = 1, \ldots, k$, e $T(\mathbf{v}_j) = \mathbf{0}$, com $j = i+1, \ldots, m$. Portanto, basta definir $R : V \to V$ como $R(\mathbf{w}_j) = \mathbf{v}_j$, com $j = i+1, \ldots, m$, e $R(\mathbf{w}_i) = \mathbf{0}$, com $i = 1, \ldots, k$. 25. Podemos imitar a prova do item (a) do Exercíco (22). Mas, vamos apresentar uma prova direta. Suponhamos que T seja sobrejetora. Então $\operatorname{Im} T = W$. Assim, pelo Teorema 4.21, $n = \dim \ker + T \dim \operatorname{Im} T$, de modo que $\dim W = m \leq n$. Assim, para cada base $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ de W, existem $\mathbf{u}_i \in V$ tais que $\mathbf{w}_i = T(\mathbf{u}_i)$, com $i = 1, \ldots, m$. Por outro lado, pelo Teorema 4.11, existe uma única transformação linear $S : W \to V$ tal que $S(\mathbf{w}_i) = \mathbf{u}_i$, com $i = 1, \ldots, m$. Portanto, $(T \circ S)(\mathbf{w}_i) = T(S(\mathbf{w}_i)) = T(\mathbf{u}_i) = \mathbf{w}_i$, com $i = 1, \ldots, m$, implica que $T \circ S = I_W$. Reciprocamente, suponhamos que $T \circ S = I_W$. Então, dado $\mathbf{w} \in W$, obtemos $\mathbf{w} = (T \circ S)(\mathbf{w}) = T(S(\mathbf{w}))$. Assim, existe um $\mathbf{u} = S(\mathbf{w}) \in V$ tal que $\mathbf{w} = T(\mathbf{u})$. Portanto, T é sobrejetora. 26. Suponhamos que S seja sobrejetora e T seja injetora. Então, pelos Exercícios (24) e (25), existe um $S_1 \in \mathcal{L}(V, U)$ tal que $S \circ S_1 = I_V$ e existe um $T_1 \in \mathcal{L}(W, V)$ tal que $T_1 \circ T = I_V$. Portanto, pondo $R = S_1 \circ T_1 \in \mathcal{L}(W, U)$, obtemos $S \circ R \circ T = I_V$. A recíproca é análoga aos Exercícios (24) e (25). 27. Suponhamos que $\operatorname{Im} T = \ker T$. Então, pelo Teorema

4.21, $n = \dim\ker T + \dim \operatorname{Im} T = 2\dim \operatorname{Im} T$, de modo que $T \neq 0$. Enquanto, $T(\mathbf{u}) \in \operatorname{Im} T = \ker T$, para todo $\mathbf{u} \in V$, implica que $T^2(\mathbf{u}) = T(T(\mathbf{u})) = \mathbf{0}$. Portanto, $T^2 = 0$. Reciprocamente, $T \neq 0$ implica que existe um $\mathbf{u} \in V$ tal que $T(\mathbf{u}) \neq \mathbf{0}$, de modo que $\operatorname{Im} T \neq \{\mathbf{0}\}$ e $\ker T \neq V$. Enquanto, $T^2 = 0$ implica que $T(\mathbf{u}) \in \ker T$, para todo $\mathbf{u} \in V$, de modo que $\operatorname{Im} T \subseteq \ker T$. Por outro lado, $2\dim \operatorname{Im} T = n = \dim\ker T + \dim \operatorname{Im} T$ implica que $\dim \operatorname{Im} T = \dim\ker t$. Portanto, $\operatorname{Im} T = \ker T$. Confira o Exercício (8) da Seção 1.2 para determinar vários $T \in \mathcal{L}(\mathbb{R}^2)$. 28. Basta observar que se $\mathbf{u}_1 \in V$ for outra solução, então $T(\mathbf{u}_1 - \mathbf{u}_0) = \mathbf{0}$. 29. (a) Dado $p(x) = a_0 + a_0 x + \cdots + a_n x^n \in \ker E$, obtemos

$$\mathbf{0} = E(p(x)) = a_0 x + \frac{a_1}{2}x^2 + \cdots + \frac{a_n}{n+1}x^{n+1},$$

de modo que $a_0 = a_1 = \cdots = a_n = 0$ e $p(x) = \mathbf{0}$. Portanto, $\ker E = \{\mathbf{0}\}$ e E é injetora. Como $a \notin \operatorname{Im} E$, para todo $a \in \mathbb{R}^{\times}$, temos que E não é sobrejetora. 30. Suponhamos que S e T sejam não singulares. Então $\ker S = \{\mathbf{0}\}$ e $\ker T = \{\mathbf{0}\}$. Dado $\mathbf{u} \in \ker(S \circ T)$, obtemos $S(T(\mathbf{u})) = \mathbf{0}$, de modo que $T(\mathbf{u}) \in \ker S = \{\mathbf{0}\}$ e $T(\mathbf{u}) = \mathbf{0}$. Assim, $\mathbf{u} \in \ker T = \{\mathbf{0}\}$ e $\mathbf{u} = \mathbf{0}$. Portanto, $\ker(S \circ T) = \{\mathbf{0}\}$ e $S \circ T$ é não singular. A recíproca segue do item (b) do Exercício (16). 31. Basta observar o sistema na forma matricial:

$$\begin{pmatrix} S_1 & S_2 \\ S_2 & S_1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} T_1 \\ T_2 \end{pmatrix}$$

e usar a Regra de Cramer. 32. Se $T \in \ker f$, então $f(T) = 0$ ou $S^{-1} \circ T \circ S = 0$. Assim,

$$T = (S \circ S^{-1}) \circ T \circ (S \circ S^{-1}) = S \circ (S^{-1} \circ T \circ S) \circ S^{-1} = 0.$$

Portanto, $\ker f = \{0\}$ e f é injetora. 33. (b) Suponhamos que $W_1 \cap W_2 \neq \{\mathbf{0}\}$. Então existe um $\mathbf{u} \in W_1 \cap W_2$ tal que $\mathbf{u} \neq \mathbf{0}$. Assim, $T(\mathbf{u}, \mathbf{0}) = \mathbf{u} = T(\mathbf{0}, \mathbf{u})$, mas $(\mathbf{u}, \mathbf{0}) \neq (\mathbf{0}, \mathbf{u})$, de modo que T não é injetora. Reciprocamente, se $(\mathbf{w}_1, \mathbf{w}_2) \in \ker T$, então $\mathbf{w}_1 + \mathbf{w}_2 = \mathbf{0}$. Logo, $\mathbf{w}_2 = -\mathbf{w}_1 \in W_1$ e $\mathbf{w}_1 = -\mathbf{w}_2 \in W_2$, de modo que $\mathbf{w}_1 = \mathbf{0}$ e $\mathbf{w}_2 = \mathbf{0}$. Portanto, $\ker T = \{\mathbf{0}, \mathbf{0}\}$ e T é injetora. (d) Note que $\ker T = \{(\mathbf{u}, -\mathbf{u}) : \mathbf{u} \in W_1 \cap W_2\}$ e $\operatorname{Im} T = \mathbf{W}_1 + \mathbf{W}_2$. 34. Note, pelo Exercício (16), que $\operatorname{Im}(T^p \circ T) \subseteq \operatorname{Im} T^p$ e $\ker T^q \subseteq \ker(T \circ T^q)$, para todos $p, q \in \mathbb{N}$. Como $\dim V = n$ temos, pelo Corolário 3.42, que a cadeia

$$\{\mathbf{0}\} \subseteq \ker T \subseteq \cdots \subseteq \ker T^q \cdots \subseteq \cdots \subseteq V$$

é finita, digamos $\ker T^{k+1} = \ker T^k$, para algum $k \in \mathbb{N}$. Portanto, $\ker T^{k+m} = \ker T^k$, para todo $m \in \mathbb{N}$. Caso contrário, existiria um $m > k$ tal que $\ker(T^m) \neq \ker(T^{m+1})$. Isto significa que existe um $\mathbf{u} \in V$ tal que $T^{m+1}(\mathbf{u}) = \mathbf{0}$, mas $T^m(\mathbf{u}) \neq \mathbf{0}$. Pondo $m = k + i$, obtemos $T^{k+1}(T^i(\mathbf{u})) = \mathbf{0}$, enquanto $T^k(T^i(\mathbf{u})) \neq \mathbf{0}$, o que é impossível. Usando que $\dim \operatorname{Im} T^k + \dim \ker T^k = \dim V$, deduzimos o resultado para a cadeia das imagens. Dado $\mathbf{v} \in \operatorname{Im} T^k \cap \ker T^k$, obtemos $T^k(\mathbf{v}) = \mathbf{0}$ e $\mathbf{v} = T^k(\mathbf{u})$, para algum $\mathbf{u} \in V$. Assim, $\mathbf{0} = T^k(\mathbf{v}) = T^{2k}(\mathbf{u})$, de modo que $\mathbf{u} \in \ker T^{2k} = \ker T^k$ implica que $\mathbf{v} = T^k(\mathbf{u}) = \mathbf{0}$. Portanto, $\operatorname{Im} T^k \cap \ker T^k = \{\mathbf{0}\}$. Conclua que $V = \operatorname{Im} T^k \oplus \ker T^k$. 35. Use o Exercício (12) com $U = \operatorname{Im} S$. 36. (a) É fácil verificar que P é linear, $\operatorname{Im} P = U$ e $\ker P = W$, de modo que $V = \ker P \oplus \operatorname{Im} P$. Note que $\mathbf{u} = T(\mathbf{u}) - (I_V - T)(\mathbf{u})$, para todo $\mathbf{u} \in V$, implica que $\ker Q = \operatorname{Im} P$ e $\operatorname{Im} Q = \ker P$. 37. Seja U um subespaço de V tal que $V = U \oplus \ker T$. Como $\ker(I_V - P) = \operatorname{Im} P = \ker T$ e $\operatorname{Im}(I_W - Q) = \ker Q = \operatorname{Im} T$, o resultado segue do Exercício (22). 38. Seja $k = \rho(T)$ e $\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ uma base de $\operatorname{Im} T$, a qual pode ser estendida para uma base $\beta = \{\mathbf{w}_1, \dots, \mathbf{w}_k, \mathbf{w}_{k+1} \dots, \mathbf{w}_n\}$ de V. Como $\mathbf{w}_i \in \operatorname{Im} T$, para $i = 1, \dots, k$, temos que existe um $\mathbf{u}_i \in V$ tal que $\mathbf{w}_i = T(\mathbf{u}_i)$. Assim, podemos escolher uma base $\{\mathbf{u}_{k+1} \dots, \mathbf{u}_n\}$ de $\ker T$ tal que $\alpha = \{\mathbf{u}_1, \dots \mathbf{u}_k, \mathbf{u}_{k+1} \dots, \mathbf{u}_n\}$ é uma base de V. Logo, pelo Teorema 4.11, existe um único $S \in \mathcal{L}(V)$ tal que $S(\mathbf{w}_i) = \mathbf{u}_i$, com $i = 1, \dots, n$. Portanto, $S \circ T = P$, com P uma projeção de V. 39. (c) Dado $\mathbf{v} \in \operatorname{Im} S \cap \ker T$, obtemos $T(\mathbf{v}) = \mathbf{0}$ e $\mathbf{v} = S(\mathbf{u})$, para algum $\mathbf{u} \in V$. Assim, $\mathbf{u} = (T \circ S)(\mathbf{u}) = T(S(\mathbf{u})) = T(\mathbf{v}) = \mathbf{0}$, de modo que $\mathbf{v} = S(\mathbf{u}) = \mathbf{0}$. Portanto, $\operatorname{Im} S \cap \ker T = \{\mathbf{0}\}$. 40. ($a \Rightarrow b$) Pelos Teoremas 3.44 e 4.21, obtemos $\dim(\ker T \cap \operatorname{Im} T) = 0$, de modo que $\ker T \cap \operatorname{Im} T = \{\mathbf{0}\}$. Portanto, $V = \ker T \oplus \operatorname{Im} T$. ($b \Rightarrow c$) É claro. ($c \Rightarrow d$) Pelo item ($b$) do Exercício (16), obtemos $\ker T \subseteq \ker T^2$. Por outro lado, dado $\mathbf{u} \in \ker T^2$, teremos $T(\mathbf{u}) \in \ker T \cap \operatorname{Im} T = \{\mathbf{0}\}$, de modo que $T(\mathbf{u}) = \mathbf{0}$, ou seja, $\mathbf{u} \in \ker T$. Assim, $\ker T^2 \subseteq \ker T$. Portanto, $\ker T^2 = \ker T$. ($d \Rightarrow e$) Pelo item (a) do Exercício (16), obtemos $\operatorname{Im} T^2 \subseteq \operatorname{Im} T$. Por outro lado, como $\dim \ker T + \dim \operatorname{Im} T = n = \dim \ker T^2 + \dim \operatorname{Im} T^2$ temos que $\dim \operatorname{Im} T = \dim \operatorname{Im} T^2$. Portanto, $\operatorname{Im} T^2 = \operatorname{Im} T$. ($e \Rightarrow f$) Use o Exercício (38). ($f \Rightarrow a$) Observe que $\mathbf{u} = (\mathbf{u} - (S \circ T)(\mathbf{u})) + (S \circ T)(\mathbf{u}) \in \ker T + \operatorname{Im} T$. Portanto, $V = \ker T + \operatorname{Im} T$. Finalmente, $S = c^{-1} I_V$ é tal que $T \circ S \circ T = T$ e $\operatorname{Im}(S \circ T) = \operatorname{Im} T$. Seja $\{(1, 2)\}$ uma base de $\operatorname{Im} T$ e estenda para uma base $\beta = \{(1, 2), (0, 1)\}$ de $\mathbb{R}^2$. Como $(1, 2) \in \operatorname{Im} T$ temos que existe um $(3, 4) \in \mathbb{R}^2$ tal que $T(3, 4) = (1, 2)$. Assim, $\{(3, 4)\}$ é parte de uma base $\alpha = \{(3, 4), (1, 0)\}$ de $\mathbb{R}^2$, com $\{(1, 0)\}$ base de $\ker T$. Consideremos $S \in \mathcal{L}(\mathbb{R}^2)$ definida como $S(1, 2) = (3, 4)$ e $S(0, 1) = (1, 0)$. Portanto, $P = S \circ T$ ou $T = S^{-1} \circ P$. Note que $P(3, 4) = (3, 4)$ e $P(1, 0) = (0, 0)$ e, depois de alguns cálculos, $(x, y) = x(1, 2) + (y - 2x)(0, 1)$, $S(x, y) = (3x, 2x + y)$ e $P(x, y) = \frac{1}{4}(3y, 4y)$ e/ou use a matriz $\mathbf{A} = \mathbf{X}\mathbf{X}^t \in \mathbb{R}^{2 \times 2}$, com $c = \mathbf{X}^t\mathbf{X}$ e $\mathbf{X} = (a,b)^t \neq (0,0)^t$, para determinar vários $T \in \mathcal{L}(\mathbb{R}^2)$. 41. É fácil verificar que $\mathcal{F}_d(\mathbb{I}_1, \mathbb{R})$ é um subespaço de $\mathcal{F}(\mathbb{R}, \mathbb{R})$. Consideremos as funções $e_k : \mathbb{I}_1 \to \mathbb{R}$ definidas como

$$e_k(x) = \begin{cases} 0, & \text{se } 0 \leq x < k \\ 1, & \text{se } k \leq x < 1. \end{cases}$$

Então $e_k \in \mathcal{F}_d(\mathbb{I}_1, \mathbb{R})$ e $\alpha = \{e_k : k \in \mathbb{I}_1\}$ é um base. De fato. Dado $f \in \mathcal{F}_d(\mathbb{I}_1, \mathbb{R})$, existe uma função $f_i \in \mathcal{F}_d(\mathbb{I}_1, \mathbb{R})$ definida como

$$f_i(x) = \begin{cases} f(x), & \text{se } x \in [a_i, a_{i+1}) \\ 0, & \text{se } x \notin [a_i, a_{i+1}) \end{cases}$$

tal que $f_i(x) = f(a_i)(e_{a_i}(x) - e_{a_{i+1}}(x))$. Assim, α gera $\mathcal{F}_d(\mathbb{I}_1, \mathbb{R})$, pois é claro que $f = f_0 + \cdots + f_n$. 42. É claro que $\mathcal{F}_l(\mathbb{I}_1, \mathbb{R}) = \mathcal{F}_d(\mathbb{I}_1, \mathbb{R}) \oplus \mathcal{C}(\mathbb{I}_1, \mathbb{R})$. Consideremos as funções $g_k : \mathbb{I}_1 \to \mathbb{R}$ definidas como

$$g_k(x) = \begin{cases} 0, & \text{se } 0 \leq x < k \\ x - k, & \text{se } k \leq x < 1. \end{cases}$$

Então $g_k \in \mathcal{C}(\mathbb{I}_1, \mathbb{R})$ e $\beta = \{g_k : k \in \mathbb{I}_1\}$ é um base. De fato. Note primeiro que $(g_{a_i} - g_{a_{i+1}})(x) = x - a_i$, para todo $x \in [a_i, a_{i+1}]$, e 0 caso contrário. Assim, dado $f \in \mathcal{F}_l(\mathbb{I}_1, \mathbb{R})$, existe uma função $f_i \in \mathcal{F}_l(\mathbb{I}_1, \mathbb{R})$ definida como

$$f_i(x) = \begin{cases} f(x), & \text{se } x \in [a_i, a_{i+1}) \\ 0, & \text{se } x \notin [a_i, a_{i+1}). \end{cases}$$

Logo, pondo $d_i = f(a_i) + b_i(a_{i+1} - a_i)$, obtemos

$$f_i(x) = b_i(g_{a_i} - g_{a_{i+1}})(x) + f(a_i)e_{a_i}(x) - d_i e_{a_{i+1}}(x).$$

Portanto, $f = g + e$, onde $g \in \mathcal{C}(\mathbb{I}_1, \mathbb{R})$ e $e \in \mathcal{F}_d(\mathbb{I}_1, \mathbb{R})$; $f = f_0 + \cdots + f_n$. Consequentemente, β é uma base de $\mathcal{C}(\mathbb{I}_1, \mathbb{R})$. Finalmente, basta observar que $T(e_k) = g_k$.

Seção 4.3

1. Basta notar que $T(1) = x = 0 \cdot 1 + 1 \cdot x + 0 \cdot x^2$ etc. 2. (*b*) Como $T(\mathbf{e}_1) = (1,0,1), T(\mathbf{e}_2) = (2,1,0)$ e $T(\mathbf{e}_3) = (0,-1,2)$ temos, pelo Teorema 4.39, que basta

escalonar a matriz

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{array}\right) \to \cdots \to \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & -2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & -\frac{1}{2} & -\frac{1}{2} \end{array}\right).$$

Portanto, $\operatorname{Im} T = F[(1,0,1),(2,1,0)]$ e $\ker T = F[(2,-1,-1)]$. 3. Basta notar que $D(x^k) = kx^{k-1} = 0 \cdot 1 + \cdots + k \cdot x^{k-1} + \cdots + 0 \cdot x^n$, com $k = 0, \ldots, n$. 4. T é claramente linear. Assim, procedendo como no Exemplo 4.40, obtemos

$$\operatorname{Im} T = F[\mathbf{E}_{12} + 2\mathbf{E}_{21}, -2\mathbf{E}_{11} + \mathbf{E}_{12} + 2\mathbf{E}_{22}] \text{ e } \ker T = F[\mathbf{I}, \mathbf{E}_{12} + 2\mathbf{E}_{21} - \mathbf{E}_{22}].$$

Note que $\ker T = \{a\mathbf{I} + b\mathbf{B} : a, b \in F, b \neq 0\}$, pois $\mathbf{B} = \mathbf{I} - (\mathbf{E}_{12} + 2\mathbf{E}_{21} - \mathbf{E}_{22})$, compare com o Exercício (16) da Seção 1.2. 5. A função $T : F^{n\times n} \to F^{n\times n}$ definida como $T(\mathbf{X}) = \mathbf{BX} - \mathbf{X}$ é claramente linear e, pelo item (b) do Exercício (8) da Seção 1.2, T é injetora. Portanto, existe um $\mathbf{A} \in F^{n\times n}$ tal que $T(\mathbf{A}) = \mathbf{O}$ ou $\mathbf{BA} = \mathbf{A}$. 6. Basta notar que $T(1) = 1 = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 + 0 \cdot x^3$ etc. e use o Exemplo 4.40. 8. Confira o Teorema 4.41. 9. Como $T(1) = (1,1,1), T(x) = (c_1, c_2, c_3)$ e $T(x^2) = (c_1^2, c_2^2, c_3^2)$ temos que $[T] = \mathbf{V}_3$ é a matriz de Vandermonde. Portanto, T é um isomorfismo. 10. Confira o Exemplo 4.43 e use o Teorema Binomial de Newton para obter $[T]$. 11. Basta notar que $\ker = \{\mathbf{0}\}$ ou $\ker T \neq \{\mathbf{0}\}$ e não ambos. 12. (a) Note que $\ker = \{\mathbf{0}\}$. Então, pelo Exercício (11), as equações possuem soluções, digamos $\mathbf{u} \neq \mathbf{0}$, e $\mathbf{v} \neq \mathbf{0}$. Assim, $(T - I)(\mathbf{u}) = \mathbf{0}$ significa que $\mathbf{u} \in \ker(T - I)$ e, depois de alguns cálculos, obtemos $\mathbf{u} = (-1, 1)$. 13. Confira o Exemplo 4.45 e/ou determine $\mathbf{P} \in \mathbb{R}^{2\times 2}$ não singular tal que $[T]_\beta^\beta = \mathbf{P}^{-1}[T]_\alpha^\alpha \mathbf{P}$. Neste caso, $\mathbf{P} = \mathbf{E}_{11} - \mathbf{E}_{12} + 2\mathbf{E}_{21} + \mathbf{E}_{22}$. Considere pontos e/ou retas que passam pela origem para obter o significado geométrico (esboço). Por exemplo, se $P = (-2, 3)$, então $T(P) = 3^{-1}(8, -5)$, de modo que T é uma rotação seguida de uma semelhança (contração). 14. Desenvolva as integrais e/ou direto. 15. $(a \Rightarrow b)$ Basta notar que $U = \operatorname{Im} E$ e $W = ker E$. $(b \Rightarrow c)$ Seja $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ uma base $\operatorname{Im} E$. Então basta estender para uma base $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ de V, com $\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ uma base de $\ker E$. $(c \Rightarrow d)$ É claro. $(d \Rightarrow b)$ Basta notar que $\mathbf{v} = (\mathbf{v} - E(\mathbf{v})) + E(\mathbf{v})$. 16. Note que $\operatorname{Im} E = U$ e $\ker E = W$. É clara que $\beta = \{(1,-1),(1,2)\}$ é uma base de $\mathbb{R}^2$. Assim, pelo exemplo 4.13, para determinar $E(x,y)$ basta encontrar $[(x,y)]_\beta$ ou, equivalentemente, escalonar a matriz

$$\left(\begin{array}{cc|c} 1 & 1 & x \\ -1 & 2 & y \end{array}\right) \to \cdots \to \left(\begin{array}{cc|c} 1 & 0 & \frac{2x-y}{3} \\ 0 & 1 & \frac{x+y}{3} \end{array}\right).$$

Portanto, $E(x,y) = 3^{-1}(2x-y, -2x+y)$. Observe que $[E]_\beta^\beta = \mathbf{I}_1 \oplus \mathbf{O}$. 17. Note que devemos impor à condição: $\operatorname{Im} E = \mathbb{R}[(1,0,-3),(0,1,-2)]$ e $\ker E = \mathbb{R}[\mathbf{u}_3]$, de modo que $\beta = \{(1,0,-3),(0,1,-2),\mathbf{u}_3\}$ seja uma base de $\mathbb{R}^3$. Por exemplo, $\mathbf{u}_3 = (3,2,1)$. Assim, pelo Exercício (16), para determinar $E(x,y,z)$ basta escalonar a matriz

$$\left(\begin{array}{rrr|c} 1 & 0 & 3 & x \\ 0 & 1 & 2 & y \\ -3 & -2 & 1 & z \end{array}\right) \to \cdots \to \left(\begin{array}{ccc|c} 1 & 0 & 0 & \frac{5x-6y-3z}{14} \\ 0 & 1 & 0 & \frac{-6x+10y-2z}{14} \\ 0 & 0 & 1 & \frac{3x+2y+z}{14} \end{array}\right).$$

Portanto,

$$E(x,y,z) = 14^{-1}(5x-6y-3z, -6x+10y-2z, -3x-2y+13z).$$

18. $(a \Rightarrow b)$ Como $R(\mathbf{u}+\mathbf{w}) = \mathbf{u}-\mathbf{w}$ temos que

$$\ker(R-I) = \{\mathbf{u}+\mathbf{w} \in V : (R-I)(\mathbf{u}+\mathbf{w}) = \mathbf{0}\} = U$$

e $\ker(R+I) = W$. $(b \Rightarrow c)$ Seja $\{\mathbf{u}_1,\ldots,\mathbf{u}_k\}$ uma base $\ker(R-I)$. Então basta estender para uma base $\{\mathbf{u}_1,\ldots,\mathbf{u}_k,\mathbf{u}_{k+1},\ldots,\mathbf{u}_n\}$ de V, com $\{\mathbf{u}_{k+1},\ldots,\mathbf{u}_n\}$ uma base de $\ker(R+I)$. $(c \Rightarrow d)$ É claro. $(d \Rightarrow b)$ Basta notar que

$$\mathbf{v} = 2^{-1}(\mathbf{v}+R(\mathbf{v})) + 2^{-1}(\mathbf{v}-R(\mathbf{v})) \in \operatorname{Im}(R+I) + \operatorname{Im}(R-I)$$

e $\operatorname{Im}(R \pm I) \subseteq \ker(R \mp I)$. 19. Pelo Exercício (18), devemos impor à condição: $\ker(R-I) = \mathbb{R}[(1,0,-3),(0,1,-2)]$ e $\ker(R+I) = \mathbb{R}[\mathbf{u}_3]$, de tal modo que $\mathbb{R}^3 = \ker(R-I) \oplus \ker(R+I)$. Por exemplo, $\mathbf{u}_3 = (3,2,1)$. Assim, basta escalonar a matriz

$$\left(\begin{array}{rrr|c} 1 & 0 & 3 & x \\ 0 & 1 & 2 & y \\ -3 & -2 & 1 & z \end{array}\right) \to \cdots \to \left(\begin{array}{ccc|c} 1 & 0 & 0 & \frac{5x-6y-3z}{14} \\ 0 & 1 & 0 & \frac{-6x+10y-2z}{14} \\ 0 & 0 & 1 & \frac{3x+2y+z}{14} \end{array}\right).$$

Portanto, $R(x,y,z) = 14^{-1}(-4x-12y-6z, -12x+6y-4z, -6x-4y+12z)$. 20. Seja $R_\theta \in \mathcal{L}(\mathbb{R}^3)$ a rotação desejada. Então $R_\theta(\mathbf{e}_3) = \mathbf{e}_3$. Por outro lado, aplicando o Exercício (17) ao plano $\pi = \{(x,y,z) \in \mathbb{R}^3 : z = 0\}$, obtemos a projeção "ortogonal" $P(x,y,z) = (x,y,0)$, em que $U = \operatorname{Im} P = \mathbb{R}[\mathbf{e}_1,\mathbf{e}_2]$ e $W = \ker P = \mathbb{R}[\mathbf{e}_3]$ o eixo dos z. Note que a restrição de R_θ a U, $R_\theta|_U \in \mathcal{L}(U)$. Assim, pelo Exemplo 3.50, $R_\theta|_U(x,y,0) = (x\cos\theta - y\operatorname{sen}\theta, x\operatorname{sen}\theta + y\cos\theta, 0)$. Logo, $R_\theta(\mathbf{e}_1) =$

$(\cos\theta, \operatorname{sen}\theta, 0), R_\theta(\mathbf{e}_2) = (-\operatorname{sen}\theta, \cos\theta, 0)$ e $R_\theta(\mathbf{e}_3) = \mathbf{e}_3$. Portanto,

$$[R_\theta] = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Observe que o vetor "normal" $\mathbf{u}_3 = (a, b, c)$ determina o plano

$$\pi = \{(x, y, z) \in \mathbb{R}^3 : ax + by + cz = 0\}.$$

Neste caso, $\pi = \mathbb{R}[(c, 0, -a), (-ab, 1 - b^2, -bc)]$, de modo que os vetores $\mathbf{u}_1 = (c, 0, -a), \mathbf{u}_2 = (-ab, 1 - b^2, -bc)$ e $\mathbf{u}_3$ formam uma base "ortogonal" β de $\mathbb{R}^3$. Assim, existe um $S \in \mathcal{L}(\mathbb{R}^3)$ não singular tal que $S(\mathbf{u}_1) = \mathbf{e}_1, S(\mathbf{u}_2) = \mathbf{e}_2$ e $S(\mathbf{u}_3) = \mathbf{e}_3$. Logo, $R = S^{-1} \circ R_\theta \circ S$ é a rotação desejada. Explicitamente,

$$[S]_\beta^\alpha = \begin{pmatrix} \frac{c}{1-b^2} & 0 & -\frac{a}{1-b^2} \\ -\frac{ab}{1-b^2} & 1 & -\frac{bc}{1-b^2} \\ a & b & c \end{pmatrix} \text{ e } [S^{-1}]_\alpha^\beta = \begin{pmatrix} c & -ab & a \\ 0 & 1-b^2 & b \\ -a & -bc & c \end{pmatrix}.$$

Portanto, depois de alguns cálculos e simplificações, $[R]_\beta^\beta = [S^{-1}]_\alpha^\beta [R_\theta]_\alpha^\alpha [S]_\beta^\alpha$ é igual a:

$$\begin{pmatrix} a^2(1-\cos\theta)+\cos\theta & ab(1-\cos\theta)-c\operatorname{sen}\theta & ac(1-\cos\theta)+b\operatorname{sen}\theta \\ ab(1-cos\theta)+c\operatorname{sen}\theta & b^2(1-\cos\theta)+\cos\theta & bc(1-\cos\theta)-a\operatorname{sen}\theta \\ ac(1-\cos\theta)-b\operatorname{sen}\theta & bc(1-cos\theta)+a\operatorname{sen}\theta & c^2(1-\cos\theta)+\cos\theta \end{pmatrix}.$$

21. Como $\det(T_i) \neq 0$ temos que T_i é um isomorfismo. Assim, dado $(x, y, z) \in \mathbb{R}^3$ existe um único $(a, b, c) \in \mathbb{R}^3$ tal que $T_i(a, b, c) = (x, y, z)$. Logo, devemos provar que $a^2 + b^2 = c^2$ implica que $x^2 + y^2 = z^2$. Para $i = 1$, obtemos

$$\begin{aligned} (2a+b+2c)^2 + (a+2b+2c)^2 &= 5a^2 + 5b^2 + 8c^2 + 8ab + 12ac + 12bc \\ &= 4a^2 + 4b^2 + 9c^2 + 8ab + 12ac + 12bc \\ &= (2a+2b+3c)^2. \end{aligned}$$

22. $(a \Rightarrow b)$ Sejam $\mathbf{u} = (1, m)$ a direção de r e $\mathbf{v} \neq k\mathbf{u}$, para todo $k \in \mathbb{R}$. Então $\beta = \{\mathbf{u}, \mathbf{v}\}$ é uma base de $\mathbb{R}^2$. Assim, $T(\mathbf{u}) = \mathbf{u}$ e $T(\mathbf{v}) = \mathbf{v} + a\mathbf{u}$, para algum $a \in \mathbb{R}$, pois a reta que passa por $\mathbf{v}$ e $T(\mathbf{v})$ deve ser paralela a r. Portanto, $[T]_\beta^\beta = \mathbf{I} + a\mathbf{E}_{12}$. $(b \Rightarrow c)$ É claro que $(T - I)^2 = 0$. $(c \Rightarrow a)$. Se $T = I$, nada há para ser provado. Se $S = T - I \neq 0$ e $S^2 = 0$, então existe um $\mathbf{u} \in \mathbb{R}^2$ tal que $S(\mathbf{u}) \neq \mathbf{0}$, mas $S^2(\mathbf{u}) = \mathbf{0}$.

Assim, $\alpha = \{S(\mathbf{u}), \mathbf{u}\}$ é claramente uma base de $\mathbb{R}^2$ e $[S]_\alpha^\alpha = \mathbf{E}_{12}$. Logo, se r for a reta $\mathbf{x} = tS(\mathbf{u})$, então $S(\mathbf{x}) = \mathbf{0}$ e $\mathbf{x} \in \ker S = \operatorname{Im} S = r$, ou seja, $r \subseteq T(r)$. Por outro lado, a reta $\mathbf{x} = \mathbf{u} + tS(\mathbf{u})$, que passa por $\mathbf{u}$ e $T(\mathbf{u})$, é paralela a r. Portanto, T é um cisalhamento. 23. Note que $\mathbf{A}(\mathbf{I}_n + \mathbf{E}_{ij}) = (\mathbf{I}_n + \mathbf{E}_{ij})\mathbf{A}$, se $i \neq j$, e use o Exercício (14) da Seção 1.2. 24. Confira o Exercício (23). 25. É fácil verificar que se $i \neq j$, então $\mathbf{E}_{ij} = \mathbf{E}_{ik}\mathbf{E}_{kj} - \mathbf{E}_{kj}\mathbf{E}_{ik}$ e $\mathbf{E}_{ii} - \mathbf{E}_{11} = \mathbf{E}_{i1}\mathbf{E}_{1i} - \mathbf{E}_{1i}\mathbf{E}_{i1}$, com $i = 2, \ldots, n$. Assim, $f(\mathbf{E}_{ij}) = 0$, se $i \neq j$, e $f(\mathbf{E}_{ii}) = f(\mathbf{E}_{11}) = c$, se $i = 2, \ldots, n$. Por outro lado, como $\mathbf{A} = \sum_{i=1}^n \sum_{j=1}^n a_{ij}\mathbf{E}_{ij}$ temos que $f(\mathbf{A}) = c(\sum_{i=1}^n a_{ii}) = c\operatorname{tr}(\mathbf{A})$, para todo $\mathbf{A} \in F^{n\times n}$. Portanto, $f = c\operatorname{tr}$, para algum $c \in F$. Observe que as matrizes $\mathbf{E}_{ij}$, com $i \neq j$, formam uma base de $\ker f$ e $\dim\ker f = n^2 - 1$. 26. (a) Se existissem $S, T \in \mathcal{L}(V)$ tais que $ST - TS = I$, então $[S][T] - [T][S] = \mathbf{I}$. Assim, $n = \operatorname{tr}(\mathbf{I}) = \operatorname{tr}([S][T] - [T][S]) = 0$, o que é impossível. (b) Confira o Exercício (29) da Seção 4.2. 27. Confira o Exemplo 4.3. Reciprocamente, para cada $x \in \mathbb{R}$, podo ser provado, que existe uma sequência (r_n) em $\mathbb{Q}$ tal que $\lim_{n\to\infty} r_n = x$. Como T é contínua temos, pelo Exemplo 4.10, que $T(x) = T(\lim_{n\to\infty} r_n) = \lim_{n\to\infty} T(r_n) = a\lim_{n\to\infty} r_n = ax$. Portanto, $T \in \mathcal{L}(\mathbb{R})$. 28. (a) Suponhamos que $\mathbb{R}^3$ possua uma estrutura complexa J. Então J é completamente determinada por $J(\mathbf{e}_1), J(\mathbf{e}_2)$ e $J(\mathbf{e}_3)$, os quais são LI. Pondo $J(\mathbf{e}_3) = a\mathbf{e}_1 + b\mathbf{e}_2 + c\mathbf{e}_3$, obtemos $-\mathbf{e}_3 = J^2(\mathbf{e}_3) = a(c\mathbf{e}_1 + J(\mathbf{e}_1)) + b(c\mathbf{e}_2 + J(\mathbf{e}_2)) + c^2\mathbf{e}_3$, de modo que $c^2 = -1$, o que é impossível. (b) Basta observar que

$$\mathbf{0} = \sum_{j=1}^n x_j\mathbf{v}_j + \sum_{j=1}^n y_j J(\mathbf{v}_j) = \sum_{j=1}^n (x_j - y_j i)\mathbf{v}_j$$

(c) Vendo V como um espaço vetorial sobre $\mathbb{R}$ e $\beta = \{\mathbf{v}_i, J(\mathbf{v}_i) : i = 1, \ldots, n\}$ uma base para V. Então $J^2 = -I$ implica que a matriz de T em relação à base β possui a forma desejada, a qual é semelhante a $\mathbf{A}$.

Seção 4.4

1. Como $f_j(\mathbf{e}_i) = \delta_{ij}$ temos que

$$f_j(\mathbf{x}) = f_j(\sum_{i=1}^n x_i\mathbf{e}_i) = \sum_{i=1}^n x_i f_j(\mathbf{e}_i) = \sum_{i=1}^n x_i\delta_{ij} = x_j.$$

Portanto. $\alpha^* = \{p_1, \ldots, p_n\}$ é o conjunto das projeções de coordenadas de F^n sobre F.

2. Note, pelo Teorema 4.52, que para determinar α^*, basta escalonar a matriz

$$\left(\begin{array}{rrr|rrr} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 1 & -2 & 0 & 0 & 0 & 1 \end{array}\right) \to \cdots \to \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & 2 & -2 & -1 \\ 0 & 1 & 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 1 & -2 & -1 \end{array}\right).$$

Portanto, $\alpha^* = \{f_1, f_2, f_3\}$, com $f_1(\mathbf{x}) = 2x - 2y - z$, $f_2(\mathbf{x}) = x - y - z$ e $f_3(\mathbf{x}) = x - 2y - z$. (a) Observe, pelo item (2) do Teorema 4.49, que qualquer $f \in (F^3)^*$ pode ser escrito sob a forma $f = f(\mathbf{u}_1)f_1 + f(\mathbf{u}_2)f_2 + f(\mathbf{u}_3)f_3$. Assim, $f(x, y, z) = 4x - 7y - 3z$. (b) $f = cf_3$, para todo $c \in F$, com $c \neq 0$. (c) Pelo item (b), $f(2, 3, -1) = -4c \neq 0$. 3. Observe, pelo Teorema 4.52, que para determinar β, basta escalonar a matriz (ou sua transposta)

$$\left(\begin{array}{rrr|rrr} 1 & 3 & 2 & 1 & 0 & 0 \\ 2 & 1 & 3 & 0 & 1 & 0 \\ 3 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \to \cdots \to \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & -\frac{5}{18} & \frac{1}{18} & \frac{7}{18} \\ 0 & 1 & 0 & \frac{7}{18} & -\frac{5}{18} & \frac{1}{18} \\ 0 & 0 & 1 & \frac{1}{18} & \frac{7}{18} & -\frac{5}{18} \end{array}\right).$$

Portanto, $\beta = \{18^{-1}(-5, 1, 7), 18^{-1}(7, -5, 1), 18^{-1}(1, 7, -5)\}$. 4. Já vimos que V é isomorfo a $\mathbb{R}^3$, de modo que V^* é isomorfo a $(\mathbb{R}^3)^*$. Assim, podemos associar $\beta^* = \{f_1, f_2, f_3\}$ em V^* com

$$\alpha^* = \{x + \frac{1}{2}y + \frac{1}{3}z, 2x + 2y + \frac{8}{3}z, -x + \frac{1}{2}y - \frac{1}{3}z\}$$

em $(\mathbb{R}^3)^*$. Logo, pelo Teorema 4.52, para determinar α, basta escalonar a matriz

$$\left(\begin{array}{rrr|rrr} 1 & 2 & -1 & 1 & 0 & 0 \\ \frac{1}{2} & 2 & \frac{1}{2} & 0 & 1 & 0 \\ \frac{1}{3} & \frac{8}{3} & -\frac{1}{3} & 0 & 0 & 1 \end{array}\right) \to \cdots \to \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & 1 & 1 & -\frac{3}{2} \\ 0 & 1 & 0 & -\frac{1}{6} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{3} & 1 & -\frac{1}{2} \end{array}\right).$$

Portanto, $\alpha = \{(1, 1, -\frac{3}{2}), (-\frac{1}{6}, 0, \frac{1}{2}), (-\frac{1}{3}, 1, -\frac{1}{2})\}$ e $\beta = \{p_1(x), p_2(x), p_3(x)\}$, com $p_1(x) = 1 + x - \frac{3}{2}x^2$, $p_2(x) = -\frac{1}{6} + \frac{1}{2}x^2$ e $p_3(x) = -\frac{1}{3} + x - \frac{1}{2}x^2$, é a base desejada. 5. Suponhamos que α não seja LI. Então existem $x_1, \ldots, x_n \in F$, não todos nulos, tais que $x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n = \mathbf{0}$. Assim, $0 = f_i(\mathbf{0}) = \sum_{j=1}^{n} x_j f_i(\mathbf{v}_j)$, com $i = 1, \ldots, n$, de modo que $\mathbf{GX} = \mathbf{O}$, com $\mathbf{X} = (x_1, \ldots, x_n)^t \neq \mathbf{O}$, implica que $\mathbf{G}$ é singular. Portanto, α é uma base de V. 6. Seja $\alpha = \{1, x, \ldots, x^n\}$ a base canônica de V. Como $f_j(x^i) = \delta_{ij}$ temos que $f_j(p(x)) = a_j = \frac{1}{j!}p^{(j)}(0)$, para todo $p(x) = a_0 + a_1x + \cdots + a_nx^n \in V$. Portanto. $\alpha^* = \{f_0, \ldots, f_n\}$. 7. É fácil verificar que α é uma base de F^n e $\alpha^* = \{f_1, \ldots, f_n\}$, em que $f_n = n^{-1}\sum_{i=1}^{n} p_i$ e $f_{i-1} = p_i - f_n$, com $i =$

$2, \ldots, n$. 8. É claro que $T : F^n \to F^m$ definida como $T(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}))$, onde $f_1, \ldots, f_m \in (F^m)^*$, é linear. Reciprocamente, dado $T \in \mathcal{L}(F^n, F^m)$, obtemos $T(\mathbf{e}_j) \in F^m$. Assim, existem únicos $a_{ij} \in F$ tais que $T(\mathbf{e}_j) = (a_{1j}, \ldots, a_{mj})$. Logo,

$$T(\mathbf{x}) = \sum_{j=1}^{n} x_j T(\mathbf{e}_j) = (\sum_{j=1}^{n} x_j a_{1j}, \ldots, \sum_{j=1}^{n} x_j a_{mj}) = (f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})).$$

Uma outra prova. Seja $\alpha^* = \{p_1, \ldots, p_m\}$ a base dual da base canônica de F^m. Defina $f_i : F^n \to F$ como $f_i(\mathbf{x}) = (p_i \circ T)(\mathbf{x}) = T^t(p_i(\mathbf{x}))$, para $i = 1, \ldots, n$. 9. Pelo Exemplo 4.56, para cada $f = [c_1, c_2, c_3, c_4] \in (F^4)^*$, (c_1, c_2, c_3, c_4) deve ser uma solução do sistema $\mathbf{AX} = \mathbf{O}$, com $\mathbf{A}$ a matriz cujas linhas são os vetores $\mathbf{u}_1$ e $\mathbf{u}_2$. Assim, $c_1 - c_3 + 2c_4 = 0$ e $2c_1 + 3c_2 + c_3 + c_4 = 0$, de modo que $c_3 = c_1 + 2c_4$ e $c_2 = -(c_1 + c_4)$. Pondo $c_1 = a$ e $c_4 = b$, temos que W° é o conjunto dos funcionais lineares sob a forma $f(\mathbf{x}) = ax_1 - (a + b)x_2 + (a + 2b)x_3 + bx_4$. Em particular, fazendo $a = 1, b = 0$ e $a = 0, b = 1$, obtemos a base $\{f_1, f_2\}$ de W°, com $f_1(\mathbf{x}) = x_1 - x_2 + x_3$ e $f_2(\mathbf{x}) = -x_2 + 2x_3 + x_4$. 10. É fácil verificar que $W = F[\mathbf{E}_{11} + 2\mathbf{E}_{12}, \mathbf{E}_{21} + 2\mathbf{E}_{22}]$. Como, pelo Exemplo 4.51, cada $f \in V^*$ é da forma $f(\mathbf{X}) = \operatorname{tr}(\mathbf{A}^t\mathbf{X}) = a_{11}x_1 + a_{12}x_2 + a_{21}x_3 + a_{22}x_4$ temos que $a_{11} + a_{22} = 0$ e $a_{22} = 3$. Por outro lado, $a_{11} + 2a_{12} = 0$ e $a_{21} + 2a_{22} = 0$. Portanto, $f(\mathbf{X}) = 2^{-1}(-6x_1+3x_2-12x_3+6x_4)$ e $f(\mathbf{B}) = 0$. 11. Como $f_{k+2} = (k+1)f_2 - kf_1$, para $k = 1, \ldots, n-2$, temos que o subespaço anulado por eles U é o mesmo de f_1 e f_2. Assim, pelo Exemplo 4.57 e/ou pelo Corolário 4.55, $\dim U = n - 2$. 12. (a) Se $f \in T^\circ$, então $f(\mathbf{v}) = 0$, para todo $\mathbf{v} \in T$. Em particular, $f(\mathbf{v}) = 0$, para todo $\mathbf{v} \in S$. Portanto, $f \in S^\circ$ e $T^\circ \subseteq S^\circ$. (b) Como $S \subseteq F[S]$ temos que $(F[S])^\circ \subseteq S^\circ$. Por outro lado, se $f \in S^\circ$, $n \in \mathbb{N}, \mathbf{v}_i \in S$ e $x_i \in F$, então $f(\sum_{i=1}^n x_i\mathbf{v}_i) = \sum_{i=1}^n x_i f(\mathbf{v}_i) = 0$. Logo, $f \in (F[S])^\circ$ e $S^\circ \subseteq (F[S])^\circ$. Portanto, $S^\circ = (F[S])^\circ$. 13. (a) Segue do item (a) do Exercício (12). (b) Como $U, W \subseteq U + W$ temos, pelo item (a) do Exercício (12), que $(U + W)^\circ \subseteq U^\circ, W^\circ$, de modo que $(U + W)^\circ \subseteq U^\circ \cap W^\circ$. Por outro lado, é claro que $U^\circ \cap W^\circ \subseteq (U + W)^\circ$. Portanto, $(U + W)^\circ = U^\circ \cap W^\circ$. Observe que não necessitamos da dimensão na prova. (c) Como $U \cap W \subseteq U, W$ temos, pelo item (a) do Exercício (12), que $U^\circ, W^\circ \subseteq (U \cap W)^\circ$, de modo que $U^\circ + W^\circ \subseteq (U \cap W)^\circ$. Por outro lado,

$$\begin{aligned}
\dim(U^\circ + W^\circ) &= \dim U^\circ + \dim W^\circ - \dim(U^\circ \cap W^\circ) \\
&= \dim V - \dim U - \dim W + \dim V - \dim(U + W)^\circ \\
&= \dim V - \dim U - \dim W + \dim(U + W) \\
&= \dim V - \dim(U \cap W) = \dim(U \cap W)^\circ;
\end{aligned}$$

Portanto, $(U \cap W)^\circ = U^\circ + W^\circ$. 14. É claro, pelo item (b) do Exercício (13), que $\{0\} = V^\circ = (U + W)^\circ = U^\circ \cap W^\circ$. Como $U^\circ, W^\circ \subseteq V^*$ temos que $U^\circ + W^\circ \subseteq V^*$. Por outro lado, dados $h \in V^*$ e $\mathbf{v} \in V$, existem $\mathbf{u} \in U$ e $\mathbf{w} \in W$ tais que $\mathbf{v} = \mathbf{u} + \mathbf{w}$. Assim, $f, g \in V^*$ definidos como $f(\mathbf{v}) = h(\mathbf{w})$ e $g(\mathbf{v}) = h(\mathbf{u})$ são tais que $f \in U^\circ$ e $g \in W^\circ$, de modo que $h = f + g$, ou seja, $V^* = U^\circ + W^\circ$. Portanto, $V^* = U^\circ \oplus W^\circ$. Observe que cada $f \in U^\circ \subseteq V^*$ induz um funcional linear $g : W \to F$ definido como $g(\mathbf{w}) = f(\mathbf{w})$. Reciprocamente, cada $g \in W^*$ induz um funcional linear $f : V \to F$ definida como $f(\mathbf{u} + \mathbf{w}) = g(\mathbf{w})$, de modo que $f \in U^\circ$. Portanto, a função $\sigma : U^\circ \to W^*$ definida como $\sigma(f) = g$ é um isomorfismo. 15. Basta considerar a função $\sigma : W \to W^*$ definida como $\sigma(\mathbf{v}, f) = (f, \mathbf{v})$. 16. A função $\sigma : (V \times W)^* \to V^* \times W^*$ definida como $\sigma(f) = (f \circ \lambda_1, f \circ \lambda_2)$ possui inversa a função $\varphi : V^* \times W^* \to (V \times W)^*$ definida como $\varphi(f, g) = f \circ p_1 + g \circ p_2$. 17. Seja $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ uma base de W. Então, já sabemos que, podemos estendê-la para uma base $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ de V. Assim, pelo Teorema 4.49, existe uma única base $\alpha^* = \{f_1, \ldots, f_n\}$ de V^* dual a α. Portanto, $g = f + c_{k+1} f_{k+1} + \cdots + c_n f_n$ possui as propriedades desejadas. 18. Como $f \neq 0$ temos que existe um $\mathbf{v} \in V$ tal que $f(\mathbf{v}) \neq 0$. Pondo $c = f(\mathbf{v})$ e $\mathbf{u} = c^{-1}\mathbf{v}$, obtemos $f(\mathbf{u}) = 1$. Observe que não necessitamos da dimensão na prova. 19. Se $\mathbf{u} \in F[\mathbf{v}] \cap \ker f$ e $\mathbf{u} \neq \mathbf{0}$, então existe um $a \in F$, com $a \neq 0$, tal que $\mathbf{u} = a\mathbf{v}$ e $f(\mathbf{u}) = 0$, de modo que $0 = f(\mathbf{u}) = af(\mathbf{v})$, o que é impossível. Assim, $F[\mathbf{v}] \cap \ker f = \{\mathbf{0}\}$ e $F[\mathbf{v}] \oplus \ker f$ existe. Dado $\mathbf{u} \in V$, é fácil verificar que

$$\mathbf{u} = f(\mathbf{v})^{-1} f(\mathbf{u})\mathbf{v} + (\mathbf{u} - f(\mathbf{v})^{-1} f(\mathbf{u})\mathbf{v}) \in F[\mathbf{v}] + \ker f.$$

Portanto, $V = F[\mathbf{v}] \oplus \ker f$ 20. Se $U = \ker f = \ker g$, então existe um $\mathbf{v} \in V - U$ tal que $f(\mathbf{v}) \neq 0$. Assim, pelo Exercício (19), $V = F[\mathbf{v}] \oplus \ker f$. Portanto, $c = f(\mathbf{v})^{-1} g(\mathbf{v})$ é tal que $g = cf$. A recíproca é clara. Uma outra prova. Suponhamos, por absurdo, que $g \neq cf$, para todo $c \in F$. Então existem $\mathbf{u}, \mathbf{v}$ tais que $f(\mathbf{u})g(\mathbf{v}) \neq f(\mathbf{v})g(\mathbf{u})$, de modo que o sistema

$$\begin{cases} f(\mathbf{u})x + f(\mathbf{v})y & = & 1 \\ g(\mathbf{u})x + g(\mathbf{v})y & = & 0 \end{cases}$$

possui uma solução (a, b) tal que $f(a\mathbf{u} + b\mathbf{v}) = 1$, mas $g(a\mathbf{u} + b\mathbf{v}) = 0$, o que é uma contradição. 21. Suponhamos, por absurdo, que $f \neq 0$ e $g \neq 0$. Então, pelo Exercício (18), existem $\mathbf{u}, \mathbf{v} \in V$ tais que $f(\mathbf{u}) = 1$ e $g(\mathbf{v}) = 1$. Assim, $h(\mathbf{u}) \pm h(\mathbf{v}) = g(\mathbf{u}) \pm f(\mathbf{v})$ e $h(\mathbf{u} \pm \mathbf{v}) = 1 + g(\mathbf{u}) \pm f(\mathbf{v}) \pm f(\mathbf{v})g(\mathbf{u})$ implicam que $0 = 2$, o que é uma contradição. 22. Podemos supor, sem perda de generalidade, que α seja LI.

Então imite a solução do Exercício (14). 23. É fácil verificar que $\alpha = \{\mathbf{e}_n : n \in \mathbf{N}\}$, com $\mathbf{e}_n(k) = \delta_{kn}$, por exemplo, $\mathbf{e}_1 = (1, 0, 0, \ldots)$, é uma base de V. Então a função $\sigma : V^* \to \mathrm{Seq}(F)$ definida como $\sigma(f) = f \circ \lambda$, com $\lambda(n) = \mathbf{e}_n$, é o isomorfismo desejado. Observe que $\alpha^* = \{f_n : n \in \mathbf{N}\}$ é sempre LI e existe um $g \in V^*$ tal que $g(\mathbf{e}_n) = 1$, para todo $n \in \mathbb{N}$, então $g \notin \alpha^*$, de modo que α^* não gera V^* e $\dim V < \dim V^*$. 24. Tomando a base $\beta = \{\mathbf{E}_{11}, \mathbf{E}_{21}, \mathbf{E}_{12}, \mathbf{E}_{22}\}$ e $\mathbf{A}\mathbf{E}_{ij} = a_{1i}\mathbf{E}_{1j} + a_{2i}\mathbf{E}_{2j}$ implica que $[T]_\beta^\beta = \mathbf{A} \oplus \mathbf{A}$. 25. Suponhamos que $\cap_{i=1}^k \ker f_i \subseteq \ker g$ e $T : V \to F^k$ definida como $T(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_k(\mathbf{x}))$. Então T é linear e $\ker T = \cap_{i=1}^k \ker f_i$. Assim, pelo item (b) do Exercício (22) da Seção,4.2, existe um $f_{\mathbf{a}} \in (F^k)^*$ tal que $g = f_{\mathbf{a}} \circ T$, isto é, $g(\mathbf{x}) = a_1 f_1(\mathbf{x}) + \cdots + a_k f_k(\mathbf{x}) = (a_1 f_1 + \cdots + a_k f_k)(\mathbf{x})$. A recíproca é clara. 26. Note que $\ker T = \cap_{i=1}^n \ker f_i$ e, pelo Exercício (25), $\ker T \subseteq \ker p_i$, com $i = 1, \ldots, n$. Assim, se $\mathbf{x} \in \ker T$, então $0 = p_i(\mathbf{x}) = x_i$, para $i = 1, \ldots, n$. Portanto, $\mathbf{x} = \mathbf{0}$ e $\ker T = \{\mathbf{0}\}$, de modo que T é não singular. 27. Como $U = \ker f$, para algum $f \in V^*$, com $f \neq 0$, temos que $U^+ = \{\mathbf{u} \in V : f(\mathbf{x}) > 0\}$ e $U^- = \{\mathbf{u} \in V : f(\mathbf{x}) < 0\}$ são as classes laterais. 28. Seja $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ uma base de $\mathbb{R}^n$ cuja dual é $\alpha^* = \{f_1, \ldots, f_n\}$. Então $x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n \in S$ se, e somente se $x_j = f_j(x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n) \geq 0$. 29. Observe que dados $f \in V^*$ e $\mathbf{v} \in V$, a função $S : V \to V$ definida como $S(\mathbf{x}) = f(\mathbf{x})\mathbf{v}$ é claramente linear. Então, para cada $\mathbf{v} \in V$, $T \circ S = S \circ T$ implica que $f(\mathbf{v})T(\mathbf{v}) = f(T(\mathbf{v}))\mathbf{v}$, de modo que $T(\mathbf{v}) = c\mathbf{v}$, com $c = f(T(f(\mathbf{v})^{-1}\mathbf{v}))$. Note que se $\mathbf{w} \in V$, com $\mathbf{w} \neq x\mathbf{v}$, para todo $x \in F$, então $T(\mathbf{w}) = a\mathbf{w}$. Assim, $T(\mathbf{v} + \mathbf{w}) = b(\mathbf{v} + \mathbf{w})$. Por outro lado, $c\mathbf{v} + a\mathbf{w} = T(\mathbf{v} + \mathbf{w})$ implica que $(c - b)\mathbf{v} + (a - b)\mathbf{w} = \mathbf{0}$. Portanto, $c = b = a$ e c não depende do vetor $\mathbf{v}$. Portanto, $T = cI_V$, para algum $c \in F$. 30. (a) É fácil verificar que os vetores $\mathbf{u}_{i-1} = \mathbf{e}_i - \mathbf{e}_1$, com $i = 2, \ldots, n$, formam uma base de U. Assim, U é um hiperplano e $U = \ker f$, para algum $f = [a_1, \ldots, a_n] \in (F^n)^*$. Logo, $f(\mathbf{u}_j) = 0$ se, e somente se, $a_1 = \cdots = a_n = c$. Portanto, $f \in U^\circ$ e $U^\circ = F[x_1 + \cdots + x_n]$. (b) Como $F^n = U \oplus F[\mathbf{v}]$, para algum $\mathbf{v} \notin U$, temos, pelo Exercício (14), que U^* é isomorfo a $F[\mathbf{v}]^\circ$. Por outro lado, $F[\mathbf{v}]^\circ = \ker E_{\mathbf{v}}$ implica que cada $f_{\mathbf{c}} \in U^*$ pode ser identificado com $\mathbf{c} \in U$. 31. Pela prova do Teorema 4.53, podemos construir $g \in V^*$ tal que $g(\mathbf{v}) = 1$ e $g(\mathbf{u}) = f(\mathbf{u})$, para todo $\mathbf{u} \in U$. 32. Note que $g = \sum_{i=1}^k c_i f_i \in U^\circ$ implica, pelo Exercício (25), que $\ker g = \cap_{i=1 i}^k f_i$, pois $f_i \neq 0$, para $i = 1, \ldots, k$. Assim, $\mathbf{v} \in U$ se, e somente se, $\mathbf{v} \in \ker g$. Portanto, $U = \cap_{i=1}^k \ker f_i$. 33. Note que cada $x_i \in S$, com $i = 1, \ldots, n$, induz uma função $g_i : U \to F$ definida como $g_i(f) = f(x_i)$. Então, podemos construir uma base $\beta = \{g_1, \ldots, g_n\}$ de U^* que goze desta propriedade. Assim, existe uma única base dual $\beta^* = \{E_1, \ldots, E_n\}$ de β em U^{**}. Portanto, pelo Teorema 4.53, existe um único $f_j \in U$ tal que $f_j(x_i) = g_i(f_j) = E_j(g_i) = \delta_{ij}$. 34. É fácil verificar que $P^* \in \mathcal{L}(V^*)$. Já vimos que $U = \operatorname{Im} P$ e $W = \ker P$. Então, pelo Teorema 4.60, $\ker P^* = (\operatorname{Im} P)^\circ = U^\circ$ e $\operatorname{Im} P^* = (\ker P)^\circ = W^\circ$. Portanto, P^* é a projeção de V^* sobre W° paralela a U°. 35. Note que $g \in \operatorname{Im} D^*$ significa que existe um $f \in V^*$ tal que $g(p(x)) = f(D(p(x)))$, para todo $p(x) \in V$. Assim, $\operatorname{Im} D^* = (\ker D)^*$ e $\ker D = \{xp(x) : p(x) \in V\}$, de modo que $D^* f = p(b) - p(a)$. 36. $T^* g = 0$, para todo $g \in V^*$. 37. Seja $S = T - cI_V$. Então $S \in \mathcal{L}(V)$ e não é sobrejetora, pois $\mathbf{v} \in \ker S$, com $\mathbf{v} \neq \mathbf{0}$. Assim, pelo item (3) do Teorema 4.60, S^* não é injetora. Portanto, existe um $f \in \ker S^*$, com $f \neq 0$, tal que $S^*(f) = 0$, ou seja, $T^*(f) = cf$. 38. Observe que cada $f \in V^*$ é da forma $f(p(x)) = \sum_{i=0}^n a_i x_i$, com $p(x) = \sum_{i=0}^n a_i x^i \in V$ e $x_i = f(x^i)$. É fácil verificar que $\operatorname{Im} D = \mathbb{P}_{n-1}(\mathbb{R})$. Como $\ker D^* = (\operatorname{Im} D)^\circ$ temos que $f \in \ker D^*$ se, e somente se, $f(x^i) = 0$, para $i = 0, \ldots, n-1$. Neste caso, $a_i = 0$, para $i = 0, \ldots, n-1$. Portanto, $\{[0, \ldots, 0, 1]\}$ é uma base de $\ker D^*$. 39. Como $W = U \cup (W - U)$, com $U = \operatorname{Im} T$, temos que (a) $\mathbf{b} \in U$ ou (b) $\mathbf{b} \notin U = (\ker T^*)^\circ$. (a) Significa que o sistema possui solução. (b) Significa que existe uma linha $\mathbf{L}$ tal que $\mathbf{LA} = \mathbf{O}$ e $\mathbf{LB} = 1$. 40. Note que σ é claramente linear. Dado $T \in \ker \sigma$, obtemos $f(T(\mathbf{v})) = 0$, para todo $f \in W^*$ e $\mathbf{v} \in V$. Assim, pelo Exercício (18), $T(\mathbf{v}) = \mathbf{0}$, para todo $\mathbf{v} \in V$. Portanto, $T = 0$ e σ é injetora. Quando $\dim V = n$ e $\dim W = m$ temos, pelo Teorema 4.59, que $\dim \mathcal{L}(V, W) = mn = \dim \mathcal{L}(W^*, V^*)$. Logo, σ é sobrejetora, de modo que é um isomorfismo.

Capítulo 5

Seção 5.1

1. (a) $\lambda_1 = -2, \mathbf{v}_1 = (-2, 1)$ e $\lambda_2 = 2, \mathbf{v}_2 = (2, 1)$. (b) $\lambda_1 = -1, \mathbf{v}_1 = (-1, 2)$ e $\lambda_2 = 3, \mathbf{v}_2 = (1, 2)$. (c) $\lambda_1 = -3, \mathbf{v}_1 = (-3, 1)$ e $\lambda_2 = 5, \mathbf{v}_2 = (-1, 3)$. (d) $\lambda_1 = -1, \mathbf{v}_1 = (-1, 1)$ e $\lambda_2 = 1, \mathbf{v}_2 = (-3, 1)$. (e) $\lambda_1 = -1, \mathbf{v}_1 = (-1, 3)$ e $\lambda_2 = 2, \mathbf{v}_2 = (1, 0)$. (f) $\lambda_1 = -1, \mathbf{e}_3$, $\lambda_2 = 1, \mathbf{e}_1$ e $\lambda_3 = 2, \mathbf{e}_2$. (g) $\lambda_1 = -1, \mathbf{v}_1 = (-1, -1, 3)$, $\lambda_2 = 1, \mathbf{v}_2 = \mathbf{e}_1$ e $\lambda_3 = 2, \mathbf{v}_3 = (1, 1, 0)$. Observe que este operador é semelhante ao do item (f). Mas, eles possuem autovetores distintos associados ao mesmo autovalor. (h) $\lambda_1 = -1, p_1(x) = -1 + x$ e $\lambda_2 = 1, p_2(x) = 1 + x, p_3(x) = x^2$. (i) $\lambda = 1$ e $p(x) = 1$ ou $\mathbf{X}(\lambda) = (\lambda + 1, 2(\lambda - 1), (\lambda - 1)^2)^t$ é a única solução, de modo que $p_1(x) = 2$ e $p_2(x) = 1 + 2x, p_3(x) = x^2$ autovetores generalizados. (j) $\lambda_1 = -1, -\mathbf{E}_{12} + \mathbf{E}_{21}$ e $\lambda_2 = 1, \mathbf{E}_{11}, \mathbf{E}_{12} + \mathbf{E}_{21}, \mathbf{E}_{22}$. (k) $\lambda = 1$ e $\mathbf{X}(\lambda) = (\lambda^2 - 2, 4 - 3\lambda, 1 - \lambda)^t$ é a única solução, de modo que $\mathbf{v}_1 = (-1, 1, 0)$ e $\mathbf{v}_2 = (2, -3, -1), \mathbf{v}_3 = (1, 0, 0)$ autovetores generalizados. (l) Note que $\lambda = 1$ e $m_g = 2$. Por outro lado, $\mathbf{X}_1(\lambda) = (\lambda - 1)(\lambda + 7, 4, 12)^t$ ou $\mathbf{Y}_1(\lambda) = (\lambda + 7, 4, 12)^t$, $\mathbf{X}_2(\lambda) =$

$(\lambda-1)(2,\lambda,3)^t$ ou $\mathbf{Y}_2(\lambda)=(2,\lambda,3)^t$ e/ou $\mathbf{X}_3(\lambda)=(\lambda-1)(3,6,10-\lambda)^t$ ou $\mathbf{Y}_3(\lambda)=(3,6,10-\lambda)^t$, produz apenas o autovalor $\mathbf{Y}_1(1)=4(2,1,3)^t$, $\mathbf{Y}_2(1)=(2,1,3)^t$ e/ou $\mathbf{Y}_3(1)=3(2,1,3)^t$. Como $m_g=2$ temos que deve existir outro autovetor $\mathbf{v}_3$ e um autovetor generalizado $\mathbf{v}_2$ correspondendo a $\mathbf{v}_1=(2,1,3)$, o qual é dado por $[\mathbf{v}_{21}]=\frac{1}{4}\mathbf{Y}_1'(1)=(\frac{1}{4},0,0)$, $\mathbf{v}_{22}=(0,1,0)$ e/ou $[\mathbf{v}_{23}]=\frac{1}{3}\mathbf{Y}_3'(1)=(0,0,-\frac{1}{3})$. O outro autovetor é obtido resolvendo a equação (5.4). Como $(\mathbf{A}-\mathbf{I})\mathbf{Y}_{21}=\mathbf{Y}_1$ e $(\mathbf{A}-\mathbf{I})\mathbf{Y}_{22}=\mathbf{Y}_1$ temos que $(\mathbf{A}-\mathbf{I})(\mathbf{Y}_{21}-\mathbf{Y}_{22})=\mathbf{O}$, de modo que o outro autovetor é $\mathbf{v}_3=(1,-4,0)$, as outras possibilidade nos leva ao mesmo resultado. A prova direta é resolver $(\mathbf{A}-\mathbf{I})^2\mathbf{X}=\mathbf{O}$. (m) $\lambda_1=0,\mathbf{v}_1=(-1,1,0)$, $\lambda_2=1,\mathbf{v}_2=(-2,1,1)$ e $\lambda_3=4,\mathbf{v}_3=(1,1,1)$. (n) $\lambda_1=-1,\mathbf{v}_1=(-4,4,1)$, $\lambda_2=-2,\mathbf{v}_2=(1,-3,1)$ e $\lambda_3=2,\mathbf{v}_3=(1,1,1)$. (o) $\lambda_1=-1,\mathbf{v}_1=(1,2,0),\mathbf{v}_2=(1,0,2)$ e $\lambda_2=3,\mathbf{v}_3=(1,1,2)$. (p) $\lambda_1=-2,\mathbf{v}_1=(1,0,1)$ e $\lambda_2=4,\mathbf{v}_2=(1,1,0),\mathbf{v}_3=(-1,0,1)$. 2. O operador $T\in\mathcal{L}(F^2)$ é tal que $T(3,1)=(-6,-2)$ e $T(-2,1)=(-6,3)$. Como $\{(3,1),(-2,1)\}$ é uma base de F^2 temos que $T(x,y)=(-6y,-x+y)$. 3. $\ker T=V_\lambda\neq\{\mathbf{0}\}$ 4. Seja $\mathbf{v}\in V$, com $\mathbf{v}\neq\mathbf{0}$, um autovetor de T associado a λ. (a) $(T+\mu I)(\mathbf{v})=T(\mathbf{v})+\mu\mathbf{v}=(\lambda+\mu)\mathbf{v}$. (b) $(\mu T)(\mathbf{v})=\mu T(\mathbf{v})=(\mu\lambda)\mathbf{v}$. (c) $\mathbf{v}=I(\mathbf{v})=T^{-1}(T(\mathbf{v}))=\lambda T^{-1}(\mathbf{v})$, de modo que $\lambda\neq 0$ e $T^{-1}(\mathbf{v})=\lambda^{-1}\mathbf{v}$. Observe que todos possuem o mesmo autovetor $\mathbf{v}$. 5. Basta lembrar que se $\mathbf{A}=[T]$, então $f_{\mathbf{A}}(x)=x^2-\operatorname{tr}(\mathbf{A})x+\det\mathbf{A}$. 6. $\lambda\in\mathbb{R}$ é um autovalor de T se, e somente se, a equação $x^2+bx+c=0$ possui raízes em $\mathbb{R}$. 7. Note que $\det(x\mathbf{I}-\mathbf{A})=\det(x\mathbf{I}-\mathbf{A})^t=\det(x\mathbf{I}-\mathbf{A}^t)$. Escolhendo $a=b=1$ e $c=d=2$ no Exercício (5), obtemos $\mathbf{u}_2=(1,2)$ para $\mathbf{A}$ e $\mathbf{v}_2=(1,1)$ para $\mathbf{A}^t$ associado a $\lambda_2=3$. 8. Sejam $a,b,d\in\mathbb{R}$ as entradas de $\mathbf{A}$. (a) Use a prova dada no item (b) e/ou como $a\neq d$ ou $b\neq 0$ devemos ter $(a-d)^2+4b^2>0$ e, pelo item (b) do Exercício (5), $(1,3)=(-b,a-1)$, de modo que $a=4,b=-1$ e $d=\frac{4}{3}$. (b) Pelo item (a) do Exercício (5), $a+d=\operatorname{tr}(\mathbf{A})=10$. Por outro lado, $a+3b=1$ e $b+3d=3$, de modo que $a-9d=-8$. Assim, $a+d=10$ e $a-9d=-8$ implicam que $a=\frac{41}{5},d=\frac{9}{5}$ e $b=-\frac{12}{5}$. Portanto, $\mathbf{X}_2=(-3,1)^t$ é o autovetor associado a $\lambda_2=9$. (c) Seja $\mathbf{P}$ a matriz cujas colunas são $\mathbf{X}_1$ e $\mathbf{X}_2$. Então $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}=\mathbf{D}$, com $\mathbf{D}=\operatorname{diag}(1,9)$. Como $(\mathbf{P}^{-1}\mathbf{B}\mathbf{P})^2=\mathbf{D}$ e $\sqrt{\mathbf{D}}=\operatorname{diag}(1,3)$ temos que $\mathbf{B}=\mathbf{P}\sqrt{\mathbf{D}}\mathbf{P}^{-1}$. (9) Basta observar a semelhança elementar $\mathbf{T}_{21}(-1)\mathbf{C}\mathbf{T}_{21}(1)$ ou

$$\begin{pmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{B} & \mathbf{A}\end{pmatrix}\to\cdots\to\begin{pmatrix}\mathbf{A}+\mathbf{B} & \mathbf{B}\\ \mathbf{O} & \mathbf{A}-\mathbf{B}\end{pmatrix}=\begin{pmatrix}\mathbf{I} & \mathbf{B}\\ \mathbf{O} & \mathbf{A}-\mathbf{B}\end{pmatrix}\begin{pmatrix}\mathbf{A}+\mathbf{B} & \mathbf{O}\\ \mathbf{O} & \mathbf{I}\end{pmatrix}$$

ou usar o Exercício (6) da Seção 1.3. 10. Se $\prod_{i=1}^n(x-\lambda_i)=f_{\mathbf{D}}(x)=f_{\mathbf{E}}(x)=\prod_{i=1}^n(x-\mu_i)$, então é claro que existe um $\sigma\in S_n$ tal que $\mu_i=\lambda_{\sigma(i)}$. Reciprocamente, pelo Teorema 1.17, basta prova para uma transposição, digamos $\mu_1=\lambda_2,\mu_2=\lambda_1$

e $\mu_i = \lambda_i$. Então, pondo $\mathbf{P} = \mathbf{P}_{12}$, obtemos $\mathbf{E} = \mathbf{PDP}$. Assim, pelo Lema 5.4, $f_{\mathbf{E}}(x) = f_{\mathbf{D}}(x)$. 11. Escolha $S(x, y, z) = (x, 2y, 3z)$ e $T(x, y, z) = (x, y, 6z)$. 12. Basta observar que

$$\begin{pmatrix} \mathbf{I} & -\mathbf{A} \\ \mathbf{O} & \mathbf{I} \end{pmatrix}\begin{pmatrix} \mathbf{AB} & \mathbf{O} \\ \mathbf{B} & \mathbf{O} \end{pmatrix}\begin{pmatrix} \mathbf{I} & \mathbf{A} \\ \mathbf{O} & \mathbf{I} \end{pmatrix} = \begin{pmatrix} \mathbf{O} & \mathbf{O} \\ \mathbf{B} & \mathbf{BA} \end{pmatrix}.$$

Uma outra prova quando $k = n$. Se $\mathbf{A}$ for não singular, então $\mathbf{BA} = \mathbf{A}^{-1}(\mathbf{AB})\mathbf{A}$ e o resultado segue do Lema 5.4. Se $\mathbf{A}$ for singular, então, pelo Exercício (10) da Seção 1.3 ou pelo Teorema 1.2, $f_{\mathbf{A}}(x) \neq 0$, para todo exceto uma quantidade finita de $x \in F$. Assim, o resultado vale para $(x\mathbf{I} - \mathbf{A})\mathbf{B}$ e $\mathbf{B}(x\mathbf{I} - \mathbf{A})$, de modo que $\det(\lambda\mathbf{I} - (x\mathbf{I} - \mathbf{A})\mathbf{B}) = \det(\lambda\mathbf{I} - \mathbf{B}(x\mathbf{I} - \mathbf{A}))$, para todos $\lambda, x \in F$. Logo, para um λ fixado, temos, pelo Exercício (14) da Seção 1.3, um igualdade de polinômios, para todo $x \in F$. Portanto, para $x = 0$, obtemos $\det(\lambda\mathbf{I} - \mathbf{AB}) = \det(\lambda\mathbf{I} - \mathbf{BA})$. Autovetores distintos, tome $\mathbf{A} = \mathbf{E}_{11}$ e $\mathbf{B} = \mathbf{E}_{12}$. 13. É fácil verificar que $\mathbf{u}$ e $\mathbf{v}$ são LI. Se $a\mathbf{u}+b\mathbf{v}$ for autovetor de T, então existe um $\nu \in F$ tal que $T(a\mathbf{u} + b\mathbf{v}) = \nu(a\mathbf{u} + b\mathbf{v})$, de modo que $(\lambda - \nu)a = 0$ e $(\mu - \nu)b = 0$. Se $\lambda = \nu$, então $b = 0$. Se $\lambda \neq \nu$, então $a = 0$. 14. Use o exercício (13), com dois vetores LI e $a = b = 1$. 15. Lembrando que se $z = a + bi$, com $b \neq 0$, for uma raiz de $f_{\mathbf{A}}(x)$, então $\overline{z} = a - bi$ também o é, onde $a, b \in \mathbb{R}$ e $z\overline{z} = a^2 + b^2 > 0$, ou seja, as raízes complexas ocorre aos pares. (a) Se n for ímpar, então $f_{\mathbf{A}}(x)$ possui pelo menos uma raiz real. Como $\det \mathbf{A} = \lambda_1 \cdots \lambda_n > 0$ temos que pelo menos uma raiz real é positiva. 16. Basta observar que $f_{\mathbf{A}}(x)$ possui pelo menos um autovalor real.

Seção 5.2

1. São aqueles com $V_\lambda = V^\lambda$ e $\mathbf{P}$ é a matriz cujas colunas são os autovetores. 2. Confira o Exercício (6) da Seção 5.1. 3. Confira o Exercício (2). 4. Confira o Exercício (18) da Seção 1.2 e/ou como $[T]$ é singular temos que $\lambda_1 = 0$ é um autovalor de T. Por outro lado, sendo $\rho(\lambda_1\mathbf{I} - [T]) = 1$, obtemos $m_g(\lambda_1) = n - 1$, de modo que $\mathbf{v}_i = \mathbf{e}_{i+1} - \mathbf{e}_1$, com $i = 1, \ldots, n - 1$, é uma base de V_{λ_1}. Neste caso, $f_{[T]}(x) = x^{n-1}(x - \lambda_n)$ e $n = \operatorname{tr}([T]) = 0 + \cdots + 0 + \lambda_n$. Logo, $\lambda_n = n$ é o outro autovalor de T com autovetor $\mathbf{v}_n = \sum_{i=1}^{n} \mathbf{e}_i$. Portanto, F^n possui uma base de autovetores e T é diagonalizável. 5. Observe que $\lambda_1 = 1, \mathbf{v}_1 = (-3, 5)$ e $\lambda_2 = 2, \mathbf{v}_2 = (-1, 2)$, de modo que $\mathbf{A}$ é diagonalizável. Assim, $\mathbf{A} = \mathbf{PDP}^{-1}$, com $\mathbf{D} = \operatorname{diag}(1, 2)$. Portanto, basta calcular $\mathbf{A}^{10} = \mathbf{PD}^{10}\mathbf{P}^{-1}$. 6. Observe que $\lambda_1 = a-b, \mathbf{v}_1 = (-1, 1, 0), \mathbf{v}_2 = (-1, 0, 1)$ e $\lambda_2 = 2a+b, \mathbf{v}_3 = (1, 1, 1)$, de modo que T é diagonalizável. 7. Note que $\lambda_1 = -\sqrt{2ab}, \mathbf{v}_1 = (a, -\sqrt{2ab}, a), \lambda_2 = \sqrt{2ab}, \mathbf{v}_2 = (a, \sqrt{2ab}, a)$ e $\lambda_3 = 0, \mathbf{v}_3 = (-1, 0, 1)$, de modo que T é diagonalizável quando $ab > 0$. 8. (a) Qualquer $T \in \mathcal{L}(\mathbb{C})$ é completamente

determinado pelos valores $T(1)$ e $T(i)$. Assim,

$$T(z) = xT(1) + yT(i) = \frac{z+\bar{z}}{2}T(1) + \frac{z-\bar{z}}{2i}T(i) = az + b\bar{z},$$

em que $a = 2^{-1}(T(1) - T(i)i)$ e $b = 2^{-1}(T(1) + T(i)i)$. A recíproca é clara. (b) Como $T(1) = T(1 \cdot 1) = T(1)^2$ temos que $T(1) = 1$, pois necessariamente $T(1) \neq 0$. Por outro lado, $-1 = T(-1) = T(i^2) = T(i)^2$ implica que $T(i) = \pm i$. Portanto, $T(z) = T(a + bi) = a + bT(i)$, ou seja, $T = I$ ou $T(z) = \bar{z}$. 9. Direto. 10. Confira o Exercício (4) da Seção 1.2. 11. (a) Vamos usar indução sobre n. É claro para $n = 1$ e note que

$$\begin{pmatrix} a_n & a_{n+1} \\ a_{n+1} & a_{n+2} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a_{n-1} & a_n \\ a_n & a_{n+1} \end{pmatrix}.$$

(b) $f_{\mathbf{C}_2}(x) = x^2 - x - 1 \in F[x]$, $\lambda_1 = 2^{-1}(1 - \sqrt{5}), \mathbf{v}_1 = 2^{-1}(1 + \sqrt{5}, -2)$ e $\lambda_2 = 2^{-1}(1 + \sqrt{5}), \mathbf{v}_2 = 2^{-1}(1 - \sqrt{5}, -2)$. Assim, $\mathbf{C}_2$ é diagonalizável. Portanto, $\mathbf{C}_2^n = \mathbf{P}\operatorname{diag}(\lambda_1^n, \lambda_2^n)\mathbf{P}^{-1}$ e compare as matrizes para obter a_n. 12. $[S]$ é semelhante a $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ e $[T]$ é semelhante a $\operatorname{diag}(\mu_1, \ldots, \mu_n)$. Como $\lambda_1, \ldots, \lambda_n$ são distintos temos, pela Fórmula de Interpolação de Lagrange, que existe um único $p(x) \in F[x]$ tal que $p(\lambda_i) = \mu_i$, com $i = 1, \ldots, n$. Logo, $p([S])$ é semelhante a $\operatorname{diag}(\mu_1, \ldots, \mu_n)$. Portanto, T pode ser escrito como um polinômio em S. 13. Seja $D \in \mathcal{L}(V)$ definido como $D(f(x)) = f'(x)$. Então $D(e^{a_i x}) = a_i e^{a_i x}$. Assim, $e^{a_i x}$ é um autovetor de D associado ao autovalor a_i. Como os a_i são distintos temos, pelo Teorema 5.10, que β é LI em V.

Seção 5.3

1. Confira o Exemplo 5.40. 2. (a) $\dim V = 2 + 3 + 1 = 6$. (b) $\partial(m) = 3, 4, 5$ ou 6. (c) $m(x) = (x-1)(x-4)(x+2)$. 3. Como $\partial(f) = \dim V$, devemos ter as possibilidades para $f(x)$: $(x-1)^4(x-2), (x-1)^3(x-2)^2$ ou $(x-1)^2(x-2)^3$. (b) Não. 4. Confira o Exercício (2) da Seção 5.2. 5. Como $\partial(f) = 4$, devemos ter as possibilidades para $f(x)$: $(x+1)^3(x-1), (x+1)^2(x-1)^2$ ou $(x+1)(x-1)^3$. (b) Confira o item (b) do Exercício (2). 6. Se $\mathbf{v} \in V_\lambda \cap V_\mu$, então $\lambda\mathbf{v} = T(\mathbf{v}) = \mu\mathbf{v}$, de modo que $(\lambda - \mu)\mathbf{v} = \mathbf{0}$ implica que $\mathbf{v} = \mathbf{0}$. 7. (a) $1 \leq \dim V_4 = m_g(4) \leq 2$ implica que $\dim V_4 = 1$ ou $\dim V_4 = 2$. (b) $\dim V_4 = 2$ e $\dim V_{-2} = 4$. 8. (a) Como $3 = m_g(\lambda) \leq m_a(\lambda)$ e $1 = m_g(\mu) \leq m_a(\mu)$, devemos ter as possibilidades para $f(x)$: $(x-\lambda)^5(x-\mu), (x-\lambda)^4(x-\mu)^2$ ou $(x-\lambda)^3(x-\mu)^3$. (b) Não. 9. $\dim V_{-1} = 1$ e $\dim V_3 = 1$ ou $\dim V_3 = 2$. 10. Seja $f(x) = \sum_{i=0}^{n-1} c_i x^i + x^n$ o

polinômio característico de T. Então $c_0 \neq 0$ e

$$T^{-1} = (-c_0^{-1})(c_1 + c_2T + \cdots + c_{n-1}T^{n-2} + T^{n-1})$$

e $f^*(x) = c_0^{-1}x^n f(x^{-1}) = c_0^{-1}(\sum_{i=0}^{n-1} c_i x^{n-i} + c_0)$ é o polinômio característico de T^{-1}. 11. Note que $[E_{ij}]_\alpha^\alpha = \mathbf{E}_{ij}$, de modo que $\mathbf{E}_{ii}^2 = \mathbf{E}_{ii}$ e $\mathbf{E}_{ij}^2 = \mathbf{O}$ quando $i \neq j$. Assim, $\sigma(E_{ii}) = \{0, 1\}$ e $\sigma(E_{ij}) = \{0\}$. 12. $m(x) = x - \lambda$ e $m(x) = (x - \lambda)^3$. 13. Note que $[T] = \mathbf{I}$ ou $[T] = -\mathbf{I}$. 14. $E(x, y) = 3^{-1}(2x - y, y - 2x)$. 15. Se $(E_1 + E_2)^2 = E_1 + E_2$, então $E_1E_2 + E_2E_1 = 0$, de modo que $E_1E_2 + E_1E_2E_1 = 0$ e $E_1E_2E_1 + E_2E_1 = 0$. Assim, $E_1E_2 - E_2E_1 = 0$. Portanto, $E_1E_2 = 0$. $\ker(E_1 + E_2) = \ker E_1 \cap \ker E_2$ e $\operatorname{Im}(E_1 + E_2) = \operatorname{Im} E_1 \oplus \operatorname{Im} E_2$. 16. Escolhendo $W_2 = F[\mathbf{e}_2, \mathbf{e}_3]$. Então $(x, y, z) = (x, x, 0) + (0, y - x, z)$, de modo que $F^3 = W_1 \oplus W_2$. Assim, devemos definir $E_1, E_2 \in \mathcal{L}(F^3)$ como $E_1(x, y, z) = (x, x, 0)$ e $E_2(x, y, z) = (0, y - x, z)$. 17. Basta provar, indutivamente, que $E^n = E$, para todo $n \in \mathbb{N}$, com $n \geq 2$. 18. Pelo Exercício (17), basta encontrar $a, b \in F$ tais que $(I + E)(aI + bE) = I$. 19. Note, pelo Lema 5.37, que $\operatorname{tr}(E_i) = \dim \operatorname{Im} E_i$. Assim, $\dim V = \dim E_1 + \cdots + \dim E_k$. Por outro lado, para cada $\mathbf{v} \in V$, obtemos $\mathbf{v} = \sum_{i=1}^k E_i(\mathbf{v}) \in \sum_{i=1}^k \operatorname{Im} E_i$, de modo que $V = \operatorname{Im} E_1 \oplus \cdots \oplus \operatorname{Im} E_k$. Observe que $E_j(\mathbf{v}) = (\sum_{i=1}^k E_i)E_j((\mathbf{v}) = \sum_{i=1}^k E_i(E_j(\mathbf{v})))$ implica que $E_j(\mathbf{v}) \in \operatorname{Im} E_j$ e $E_iE_j(\mathbf{v}) \in \operatorname{Im} E_i$, ou seja, $E_iE_j(\mathbf{v}) = \mathbf{0}$, quando $i \neq j$. 20. e 21. Confira o Exemplo 5.40. 22. (a) Basta notar que

$$\mathbf{v} = \frac{1}{2}\mathbf{v} + \frac{1}{2}\mathbf{v} = \frac{1}{2}(\mathbf{v} + T(\mathbf{v})) + \frac{1}{2}(\mathbf{v} - T(\mathbf{v})) \in W_1 + W_2.$$

(c) Matrizes simétricas e antissimétricas. 23. Confira a solução do Exercício (18) da Seção 4.3. 24. Por definição, $m(x) = x^m$, com $m \leq k$. Por outro lado, $\partial(f) = n$, $m(x)$ e $f(x)$ devem possuir as mesmas raízes. Portanto, $f(x) = x^n$. Uma outro prova $xI = xI - T^k = (xI - T)(x^{k-1}I + \cdots + T^k)$ implica que $x^n = \det(xI - T)\det(*)$. 25. Basta notar que $m(\mathbf{A})$ é um fator de $m(\mathbf{A}^t)$ e, vice-versa. 26. Sejam $m_1(x), m_2(x)$ e $m(x)$ os polinômios minimais de $\mathbf{A}, \mathbf{B}$ e $\mathbf{C}$. Como $\mathbf{C}^m = \mathbf{A}^m \oplus \mathbf{B}^m$, para todo $m \in \mathbb{N}$, temos que $m(\mathbf{C}) = \mathbf{O}$ implica que $m(\mathbf{A}) = \mathbf{O}$ e $m(\mathbf{B}) = \mathbf{O}$, de modo que $m_1(x)$ é um fator de $m(x)$ e $m_2(x)$ é um fator de $m(x)$. Por outro lado, se $m_1(x)$ for um fator de $g(x)$ e $m_1(x)$ for um fator de $g(x)$, então existem $q_1(x), q_2(x) \in F[x]$ tais que $g(x) = q_1(x)m_1(x)$ e $g(x) = q_2(x)m_2(x)$. Como $g(\mathbf{A}) = g(\mathbf{B}) = \mathbf{O}$ temos que $g(\mathbf{C}) = \mathbf{O}$, de modo que $m(x)$ é um fator de $g(x)$. Portanto, $m(x) = \operatorname{mmc}(m_1(x), m_2(x))$. 27. Considerando a base $\beta = \{\mathbf{E}_{ij}\}$ ordenada antilexicográfica e $T(\mathbf{E}_{ij}) = \sum_{k=1}^n b_{ki}\mathbf{E}_{kj}$ implicam que $[T]_\beta^\beta = \mathbf{B} \oplus \cdots \oplus \mathbf{B}$. Portanto, o resultado segue do Exercício (26). 28.

(a) Note que $\mathbf{B}$ é o autovetor de T associado a 0. (b) Observe que

$$T^2(\mathbf{X}) = \mathbf{B}(\mathbf{BX} - \mathbf{XB}) - (\mathbf{BX} - \mathbf{XB})\mathbf{B} = \mathbf{B}^2\mathbf{X} - 2\mathbf{BXB} + \mathbf{XB}^2$$

e, recursivamente, $T^m(\mathbf{X}) = \sum_{i=0}^{m} \binom{m}{i}(-1)^i \mathbf{B}^{m-i}\mathbf{XB}^i$ e note que cada termo de $T^{2k}(\mathbf{X})$ contém um fator $\mathbf{B}^m$, com $m \geq k$. (c) Calculação direta. (d) Suponhamos que $\mathbf{B}$ seja diagonalizável. Então existe uma matriz não singular $\mathbf{P}$ tal que $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, ou seja, $\mathbf{BP}_i = \lambda_i \mathbf{P}_i$, com $i = 1, \ldots, n$. Consideremos as matrizes $\mathbf{A}_{ij} = \mathbf{PE}_{ij}$ cuja j-ésima coluna é $\mathbf{P}_i$ e as demais zeros. Então $\alpha = \{\mathbf{A}_{ij}\}$ é uma base de V, com a ordem do Exercício (27). Assim, depois de alguns cálculos, $[T]_\alpha^\alpha = (\lambda_1\mathbf{I} - \mathbf{B}^t) \oplus \cdots \oplus (\lambda_n\mathbf{I} - \mathbf{B}^t)$. 29. Falso, tome $\mathbf{A} = \mathbf{E}_{11}$ e $\mathbf{B} = \mathbf{E}_{12}$. Em geral, se $f(\mathbf{AB}) = \mathbf{O}$ e $g(x) = xf(x)$, então $g(\mathbf{BA}) = \mathbf{O}$, pois $(\mathbf{BA})^{m+1} = \mathbf{BA}(\mathbf{BA})^m = \mathbf{B}(\mathbf{AB})^m\mathbf{A}$. 30. Sejam $m_1(x)$ o polinômio minimal de $\mathbf{A}$ em $\mathbb{R}$ e $m_2(x)$ o polinômio minimal de $\mathbf{A}$ em $\mathbb{C}$. Então claramente $m_2(x)$ é um fator de $m_1(x)$. Por outro lado, se $\partial(m_1) = k$ e $\partial(m_2) = m$, então

$$c_0\mathbf{I} + c_1\mathbf{A} + \cdots + c_m\mathbf{A}^m = \mathbf{O}, \text{ onde } c_i \in \mathbb{C},$$

é equivalente a n^2 sistemas homogêneos nas variáveis $c_0, c_1, \ldots, c_m$. Como os coeficientes (entradas de $\mathbf{A}^j$) estão em $\mathbb{R}$ temos, pelo Exercício (8) da Seção 2.2, que ele possui uma solução em $\mathbb{R}$. Portanto, $k \leq m$. 31. Seja $T \in \mathcal{L}(\mathbb{C}^3)$ tal que $[T] = \mathbf{A}$. Então, pelo Exercício (14) da Seção 3.5, é fácil verificar que $\mathbf{v}_1 = (1, 1, 1)$ é um autovetor de T associado a $3a$. Use a base $\alpha = \{\mathbf{v}_1, \mathbf{e}_2, \mathbf{e}_3\}$ para obter $[T]_\alpha^\alpha = \mathbf{B}$. Assim, $\det(x\mathbf{I} - \mathbf{B}) = (x - 3a)(x^2 - 3(b^2 - c^2))$. A condição é a qualquer e $|c| < |b|$. 32. Pondo $\mathbf{X}_{n+1} = (T_n, T_{n+1})^t$, obtemos $\mathbf{X}_{n+1} = \mathbf{AX}_n$, para todo $n \in \mathbb{N}$, e $\mathbf{A} = \mathbf{E}_{11} + \mathbf{E}_{21} + 20\mathbf{E}_{22}$. Assim, recursivamente, teremos $\mathbf{X}_n = \mathbf{A}^{n-1}\mathbf{X}_1$, para todo $n \in \mathbb{N}$. Logo, depois de alguns cálculos, $T_n = 9^{-1}(5^{n+1} - (-4)^{n+1})$. 33. Confira o Lema 5.18. Seja μ um autovalor de $p(\mathbf{A})$. Então $p(\mathbf{A})\mathbf{X} = \mu\mathbf{X}$, para algum $\mathbf{X} \neq \mathbf{O}$. Podemos supor, sem perda de generalidade, que $p(x)$ seja mônico e ponha $h(x) = p(x) - \mu$. Então, pelo Teorema 1.2, $h(x) = (x - \nu_1) \cdots (x - \nu_k)$. Assim, $h(\mathbf{A})\mathbf{X} = \mathbf{O}$, de modo que $\mathbf{A} - \nu_i\mathbf{I}$ é singular, para algum $i = 1, \ldots, k$. Portanto, $0 = h(\lambda_j) = p(\lambda_j) - \mu$, para algum $j = 1, \ldots, n$. 34. Pelo Exercíco (33), $\mathbf{C} = g(\mathbf{A})$ é não singular se, e somente se, $g(\lambda_j) \neq 0$, para todos os autovalores λ_j de $\mathbf{A}$. 35. (a) Se $\mathbf{AX} = \lambda\mathbf{X}$, então $\det \mathbf{AX} = \operatorname{adj}(\mathbf{A})\mathbf{AX} = \lambda \operatorname{adj}(\mathbf{A})\mathbf{X}$. (b) Se $\mathbf{AX} = \lambda\mathbf{X}$ e $\lambda \neq 0$, então o resultado segue do item (a). Se $\lambda = 0$ e $\rho(\mathbf{A}) \leq n - 2$, então $\operatorname{adj}(\mathbf{A}) = \mathbf{O}$, de modo que $\operatorname{adj}(\mathbf{A})\mathbf{X} = \mathbf{O}$. Se $\rho(\mathbf{A}) = n - 1$, então $m_g(\lambda) = 1$, de modo que $\mathbf{A}(\operatorname{adj}(\mathbf{A})\mathbf{X}) = \mathbf{O}$. Assim, $\operatorname{adj}(\mathbf{A})\mathbf{X}$ é um autovetor de $\mathbf{A}$. Portanto, $\operatorname{adj}(\mathbf{A})\mathbf{X} = \mu\mathbf{X}$, para algum $\mu \in F$. 36. Confira o item (a) do Exercício (35). 37.

Como $f_{\mathbf{A}}(x) = (x-a_{ii})f_{ii}(x)+\sum_{i\neq j}(-1)^{i+j}a_{ij}f_{ij}(x)$ e, pela equação 5.2, temos que $(x-\lambda)\sum_{i=1}^{n}x_i f_{ij}(x) = x_j f_{\mathbf{A}}(x)$, com $j = 1,\ldots,n$. 38. (a) Seja $f(x)$ o polinômio característico de $\mathbf{A}$. Então, pondo $\mathbf{B} = \operatorname{adj}(\lambda\mathbf{I}-\mathbf{A})$ e $\mathbf{C} = \operatorname{adj}(\mu\mathbf{I}-\mathbf{A})$, obtemos

$$\mathbf{R}_\lambda - \mathbf{R}_\mu = \frac{\mathbf{B}f(\mu)-\mathbf{C}f(\lambda)}{f(\lambda)f(\mu)} = \frac{\mathbf{BC}(\mu\mathbf{I}-\mathbf{A})-\mathbf{BC}(\lambda\mathbf{I}-\mathbf{A})}{f(\lambda)f(\mu)} = (\mu-\lambda)\mathbf{R}_\lambda\mathbf{R}_\mu.$$

(b) Observe, pelo Teorema 5.12, que $\mathbf{R}_\lambda = \mathbf{P}(\lambda\mathbf{I}-\mathbf{D})^{-1}\mathbf{Q}^t$, com $\mathbf{Q} = (\mathbf{P}^{-1})^t$. (c) Note que $\operatorname{tr}(\mathbf{R}_\lambda) = \operatorname{tr}(\lambda\mathbf{I}-\mathbf{D})^{-1}$. 39. Note que $\mathbf{R}_n^t = \mathbf{R}_n$ e $\mathbf{R}_n^2 = \mathbf{I}$. Assim, depois de alguns cálculos, obtemos $m_n(x) = (x+1)(x-1)$ o polinômio minimal e $f_n(x) = (x^2-1)f_{n-2}(x)$, para todo $n\in\mathbb{N}$, com $n\geq 3$. Portanto, $f_n(x) = [(x+1)(x-1)]^{\frac{n}{2}}$, se n é par e $f_n(x) = (x+1)^{\frac{n-1}{2}}(x-1)^{\frac{n+1}{2}}$, se n é ímpar. Consequentemente, $\mathbf{R}_n$ é diagonalizável. 40. Seja $\mathbf{C}_n\mathbf{X}_i = \lambda_i\mathbf{X}_i$, com $i = 1,\ldots,m$, e $\sum_{i=1}^{m}c_i\mathbf{X}_i = \mathbf{0}$. Então, aplicando sucessivamente $\mathbf{C}_n$, obtemos o sistema

$$\left\{\begin{array}{rcl} c_1\mathbf{X}_1+\cdots+c_m\mathbf{X}_m & = & \mathbf{0} \\ \lambda_1(c_1\mathbf{X}_1)+\cdots+\lambda_m(c_m\mathbf{X}_m) & = & \mathbf{0} \\ & \vdots & \\ \lambda_1^{m-1}(c_1\mathbf{X}_1)+\cdots+\lambda_m m-1(c_m\mathbf{X}_m) & = & \mathbf{0} \end{array}\right.$$

Note que existem n tais sistemas, pois $\mathbf{X}_i = (1,\lambda_i,\ldots,\lambda_i^{n-1})^t$. Em particular, $\mathbf{V}_m\mathbf{Y} = \mathbf{O}$, com $\mathbf{Y} = (c_1,\ldots,c_m)^t$. Portanto, $\mathbf{Y} = \mathbf{O}$ e os autovetores são LI. 41. Pelo Exercício (38), $\mathbf{C}_n^s = \mathbf{R}_n\mathbf{C}_n\mathbf{R}_n$. 42. Como $T(\mathbf{e}_i) = \mathbf{e}_{i-1}$, para $i = 2,\ldots,n$, e $T(\mathbf{e}_1) = \mathbf{e}_n$ temos que $[T] = \mathbf{P}_n = \mathbf{E}_{n1}+\mathbf{E}_{11}+\cdots+\mathbf{E}_{(n-1)(n-1)}$, a qual chama-se *matriz de permutação*. Observe que $\mathbf{P}_n$ é a matriz companheira do polinômio $p(x) = x^n-1$, tendo raízes complexas distintas $\lambda_k = \omega^k$, em que $\omega = \cos(\frac{2\pi}{n})+\operatorname{sen}(\frac{2\pi}{n})i$, para $k = 0,\ldots,n-1$, de modo que $\mathbf{X}_{k+1} = (1,\omega^k,\ldots,\omega^{k(n-1)})^t$ são os autovetores associados. Portanto, T é diagonalizável. 43. Basta observar que $\mathbf{C} = g(\mathbf{P}_n)$, com $g(x) = c_0+c_1x+\cdots+c_{n-1}x^{n-1}$ o *polinômio representante* de $\mathbf{C}$, e use o Exercício (33). 44. (a) Direto. (b) Se $c\neq 0$, então $\mathbf{A} = \mathbf{T}_{12}(c^{-1}(a-1))\mathbf{T}_{21}(c)\mathbf{T}_{12}(c^{-1}(d-1))$. Se $c = 0$, então $a\neq 0$, de modo que podemos aplicar o caso anterior a $\mathbf{B} = \mathbf{T}_{21}(1)\mathbf{A}$. (c) Sejam $\lambda,\mu\in\mathbb{C}$ os autovalores de $\mathbf{A}$. Então $\lambda\mu = 1$, de modo que $\mu = \lambda^{-1}$. Observe que $\lambda\neq\mu$. Caso contrário, $|\lambda| = |\mu| = 1$ e $2\geq|\lambda+\mu| = |\operatorname{tr}(\mathbf{A})| > 2$. o que é impossível. Portanto, $\mathbf{A}$ é semelhante a $\operatorname{diag}(\lambda,\lambda^{-1})$, onde $\lambda\notin\{-1,0,1\}$. (d) Se $|\operatorname{tr}(\mathbf{A})| < 2$, então

$$|\lambda+\lambda^{-1}| = |\lambda+\mu| = |\operatorname{tr}(\mathbf{A})| < 2.$$

de modo que $\lambda \notin \mathbb{R}$ ou $\lambda \notin i\mathbb{R}$. (e) Se $|\operatorname{tr}(\mathbf{A})| = 2$, então os possíveis autovalores reais de $\mathbf{A}$ são 1 ou -1. Portanto, as possíveis matrizes semelhantes a $\mathbf{A}$ são: $\mathbf{I}$, $-\mathbf{I}$, $\mathbf{T}_{21}(1)$ e $-\mathbf{T}_{21}(-1)$. (f) Se $|\operatorname{tr}(\mathbf{A})| \neq 2$, então $\mathbf{A}$ possui dois autovalores distintos $\lambda, \mu \in \mathbb{C}$, os quais também o são da matriz

$$\begin{pmatrix} \frac{\lambda+\mu}{2} & \frac{\lambda-\mu}{2} \\ \frac{\lambda-\mu}{2} & \frac{\lambda+\mu}{2} \end{pmatrix}.$$

(g) Falso, tome $\mathbf{A} = \mathbf{I} + \mathbf{E}_{12}$.

Capítulo 6

Seção 6.1

1. $f(T)(\mathbf{v}) = \mathbf{0}$, para $\mathbf{v} \in \ker f(T)$. Assim, $f(T)(T(\mathbf{v})) = T(f(T)(\mathbf{v})) = T(\mathbf{0}) = \mathbf{0}$. Portanto, $T(\mathbf{v}) \in \ker f(T)$ e $\ker f(T)$ é invariante sob T. 2. Imite o Exemplo 6.6. 3. Imite o Exemplo 6.9. 4. Seja $f(x) = \sum_{i=0}^{n-1} c_i x^i + x^n$. Então, pelo Teorema de Cayley-Hamilton, $f(T) = 0$, de modo que $T^k(\mathbf{v}) \in F[T^{n-1}(\mathbf{v}), \ldots, T(\mathbf{v}), \mathbf{v}]$, para todo $k \in \mathbb{N}$, com $k \geq n$. Portanto, $W = F[T^{n-1}(\mathbf{v}), \ldots, T(\mathbf{v}), \mathbf{v}]$. 5. Note que $\{S^2(\mathbf{e}_1), S(\mathbf{e}_1), \mathbf{e}_1\}$ é uma base de F^3. Observe que $m(x) = (x+1)(x-1)$ é o polinômio minimal de T. 6. Pelo Exercício (4), basta provar que conjunto $\alpha = \{T^{n-1}(\mathbf{v}), \ldots, T(\mathbf{v}), \mathbf{v}\}$ é LI. Suponhamos, por absurdo, que α seja LD. Então, pelo Corolário 3.33, podemos escolher um $k < n$ tal que

$$T^k(\mathbf{v}) \in W = F[T^{k-1}(\mathbf{v}), \ldots, T(\mathbf{v}), \mathbf{v}].$$

Assim, aplicado T, obtemos $T^{k+1}(\mathbf{v}) \in F[T^k(\mathbf{v}), \ldots, T^2(\mathbf{v}), T(\mathbf{v})] \subset W$. Logo, aplicado recursivamente T, teremos $T^{k+m}(\mathbf{v}) \in W$, para todo $m \in \mathbb{N}$. Portanto, $V \subset W$, de modo que $\dim V \leq k < n$, o que é uma contradição. Note que $T^n(\mathbf{v}) \in V$ implica que existem $c_0, c_1, \ldots, c_{n-1} \in F$ tais que

$$T^n(\mathbf{v}) = c_{n-1}T^{n-1}(\mathbf{v}) + \cdots + c_1 T(\mathbf{v}) + c_0\mathbf{v}$$

Assim, $f(x) = -\sum_{i=0}^{n-1} c_i x^i + x^n$ é o polinômio minimal e característico de T. A recíproca segue do Exercício (4). 7. Suponhamos que existam $a, b \in F$, com $a \neq 0$, tais que W seja invariante sob $aT + bI$. Como

$$T = a^{-1}(aT + bI) - a^{-1}bI$$

temos $T(\mathbf{u}) \in W$, para todo $\mathbf{u} \in W$. Portanto, W é invariante sob T. A recíproca é clara. 8. Note, pelo Teorema 1.2, que existem $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$, não necessariamente distintos, tais que $m(x) = \prod_{i=1}^{k}(x - \lambda_i)^{r_i}$ é o polinômio minimal de T. Assim, pelo Teorema 6.8, $T = \sum_{i=1}^{k} TE_i$ e existem $D, N \in \mathcal{L}(V)$, com $D = \sum_{i=1}^{k} \lambda_i E_i$ diagonalizável e N nilpotente, tais que $T = D + N$ e $DN = ND$. Observe, indutivamente, que $g(T) = \sum_{i=1}^{k} g(T)E_i$ e $g(D) = \sum_{i=1}^{k} g(\lambda_i)E_i$, de modo que $g(D)$ é diagonalizável. Como

$$g(T) = g(D + N) = \sum_{q=0}^{m} c_q(D + N)^q = \sum_{q=0}^{m} c_q \left(\sum_{p=0}^{q} \binom{q}{p} D^{q-p}N^p \right)$$

temos que $g(T) = g(D)+N(*)$. 9. Note que $U = F[\mathbf{e}_1]$ é o único subespaço invariante sob T. Não obstante, existe um $W = F[\mathbf{e}_2]$ tal que $F^2 = U \oplus W$. 10. Suponhamos, por absurdo, que T seja diagonalizável e nilpotente. Então já vimos que $m(x) = x^k$, com $k \leq n$, era o polinômio minimal de T. Por outro lado, sendo T diagonalizável, devemos ter $k = 1$. Portanto, $T = m(T) = 0$, o que é uma contradião, pois $\rho(T) = 1$. 11. Seja $\alpha = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ uma base de V. Então é fácil verificar que os $E_j \in \mathcal{L}(V)$ definidos como $E_j(\mathbf{u}_i) = \delta_{ij}\mathbf{u}_i$ são diagonalizável. Assim, $E_jT = TE_j$, para $j = 1, \ldots, n$. Logo, $T(\mathbf{u}_j) = T(E_j(\mathbf{u}_j)) = TE_j(\mathbf{u}_j) = E_jT(\mathbf{u}_j)$, ou seja, $T(\mathbf{u}_j)$ é um autovetor de E_j associado a $\lambda_j = 1$. Portanto, $V_{\lambda_j} = \operatorname{Im} E_j = F[\mathbf{u}_j]$, de modo que $T(\mathbf{u}_j) \in V_{\lambda_j}$, ou seja, $T(\mathbf{u}_j) = c\mathbf{u}_j$, para algum $c \in F$. 12. Suponhamos que T seja diagonalizável e $\lambda_1, \ldots, \lambda_k \in F$ seus autovalores distintos. Então $p(x) = \prod_{i=1}^{k}(x - \lambda_i)$ é tal que $p(T)(\mathbf{v}) = \mathbf{0}$, para todo $\mathbf{v} \in V$. Assim, $p(T) = 0$ e $m(x)$ é um fator de $p(x)$. Como todo autovalor de T é uma raiz de $m(x)$ temos que $m(x) = p(x)$. Reciprocamente, se $m(x) = \prod_{i=1}^{k}(x - \lambda_i)$, então, pelo Teorema 6.4, $V = \sum_{i=1}^{k} W_i$, com $W_i = \ker(T - \lambda_i I)$. Dado $\mathbf{v} \in W_i$, obtemos $T(\mathbf{v}) = \lambda_i\mathbf{v}$, de modo que $\mathbf{v}$ é um autovetor de T. Portanto, pelo Teorema 5.19, T é diagonalizável. 13. Não, pois $m(x) = (x - 1)(x^2 + x + 1)$ é o polinômio minimal de $\mathbf{A}$.

Seção 6.2

1., 2. e 3. Confira o Exemplo 6.14. 4. (a) Primeiro lembre-se que $\mathbf{AE}_j = \mathbf{C}_j$ e $\mathbf{e}_i\mathbf{A} = \mathbf{L}_i$. Assim,

$$\mathbf{AN} = \begin{pmatrix} \mathbf{O} & \mathbf{AE}_1 & \mathbf{AE}_2 & \mathbf{AE}_3 \end{pmatrix} = \begin{pmatrix} \mathbf{e}_2\mathbf{A} & \mathbf{e}_3\mathbf{A} & \mathbf{e}_4\mathbf{A} & \mathbf{O} \end{pmatrix}^t = \mathbf{NA}$$

se, e somente se, $a_{i(j-1)} = a_{(i+1)j}$, se $i < 4$ e $j > 1$, $a_{4j} = 0$, se $j < 4$, e $a_{i1} = 0$, se $i > 1$, de modo que $a_{ii} = a$, $a_{i(i+1)} = b, a_{i(i+2)} = c$, $a_{14} = d$ e $a_{ij} = 0$ quando

$i > j$. (b) Note que $\mathbf{A} = a\mathbf{I} + \mathbf{N}_1$. Então, pelo item (a) do Exercício (4) da Seção 5.1, os autovalores de $\mathbf{A}$ são da forma $\lambda = a + \mu$, com μ um autovalor de $\mathbf{N}_1$. Se $b \neq 0$, então $\mathbf{N}_1$ possui índice de nilpotência igual a 4 e $\rho(\mathbf{N}_1) = 3$, de modo que $\dim V_\lambda = 1$, pois $\mathbf{A}$ e $\mathbf{N}_1$ possuem os mesmos autovetores. 5. Seja $S \in \mathcal{L}(V)$ semelhante a T. Então existe um $P \in \mathcal{L}(V)$ não singular tal que $S = P^{-1}TP$. Assim, indutivamente, $S^m = P^{-1}T^mP$, para todo $m \in \mathbb{N}$. Portanto, T é nilpotente de índice k se, e somente se, S é nilpotente de índice k. 6. Se $\mathbf{v} \in \operatorname{Im} T^{k-i}$, então existe um $\mathbf{u} \in V$ tal que $\mathbf{v} = T^{k-i}(\mathbf{u})$. Assim, $T^i(\mathbf{v}) = T^i(T^{k-i}(\mathbf{u})) = T^k(\mathbf{u}) = \mathbf{0}$, de modo que $\mathbf{v} \in \ker T^i$. 7. Note que $(-T)^k = (-1)^kT^k = 0$ implica que $I = I - T^k = (I-T)(I+T+\cdots+T^{k-1})$, de modo que $I + T = I - (-T)$ é não singular. 8. Se $T^k = 0$, para algum $k \in \mathbb{N}$, e $\lambda \neq 0$ for um autovalor de T, então $\lambda^k \neq 0$ é um autovalor de $T^k = 0$ ou $\lambda^k\mathbf{v} = \mathbf{0}$, o que é impossível. Reciprocamente, se todos os autovalores de T são zeros, então $f(x) = x^n$ é o polinômio característico T. Assim,, pelo Teorema de Cayley-Hamilton, $T^n = f(T) = 0$. Considere $T \in \mathcal{L}(\mathbb{R}[x])$ definido como $T(f(x)) = f'(x)$. Então 0 é o único autovalor de T, mas T não é nilpotente. 9. (a) Se T é cíclico, então existe um $\mathbf{v} \in V$, com $\mathbf{v} \neq \mathbf{0}$, tal que $\alpha = \{T^{n-1}(\mathbf{v}), \ldots, T(\mathbf{v}), \mathbf{v}\}$ é uma base de V. Assim, $\mathbf{A} = [T]_\alpha^\alpha = \sum_{i=1}^{n-1} \mathbf{E}_{i(i+1)}$. Logo, $\rho(\lambda\mathbf{I} - \mathbf{A}) = n - 1$, para todo autovalor λ de T, e $\dim\ker(\lambda\mathbf{I} - \mathbf{A}) = 1$. Portanto, os autovalores associados a λ são múltiplos um do outro. (b) Se $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ é a base de autovetores de V associada aos autovalores distintos $\lambda_1, \ldots, \lambda_n$, então, pondo $\mathbf{u} = \sum_{i=1}^n \mathbf{v}_i$, obtemos $T^m(\mathbf{u}) = \sum_{i=1}^n \lambda_i^n \mathbf{v}_i$. Como $[T^{n-k}(\mathbf{u})]_\beta$, para $k = 1, \ldots, n$, são as colunas de uma matriz de Vandermonde, temos que o resultado segue. 10. Note que $\mathbf{AB} = -\mathbf{BA}$ implica que $\mathbf{AB}^{2k-1} = -\mathbf{B}^{2k-1}\mathbf{A}$, para todo $k \in \mathbb{N}$. Suponhamos, por absurdo, que $\mathbf{AX} + \mathbf{XA} = \mathbf{B}$ possua uma solução. Então

$$\mathbf{B}^{2k} = (\mathbf{AX} + \mathbf{XA})\mathbf{B}^{2k-1} = \mathbf{A}(\mathbf{XB}^{2k-1}) - (\mathbf{XB}^{2k-1})\mathbf{A},$$

de modo que $\operatorname{tr}(\mathbf{B}^{2k}) = 0$. Assim, $\mathbf{B}^{2k} = \mathbf{O}$ e $\mathbf{B}^2$ é nilpotente. Portanto, $\mathbf{B}$ é nilpotente, o que é uma contradição.

Seção 6.3

1. Confira o Exemplo 6.21. 2. Note que $\mathbf{J} = \mathbf{I} + \mathbf{J}_{12}$ é a forma de Jordan de $\mathbf{A}$. Portanto, $\mathbf{A}$ e $\mathbf{B}$ são semelhantes se, e somente se, $c = 0$. 3. e 4. Confira o Exemplo 6.19. 5. Condições $(a - d)^2 + 4bc = 0$ e $b^2 + c^2 \neq 0$. 6. Basta observar que a forma de Jordan de T é uma matriz triangular com os autovalores em sua diagonal. 7. Confira o Exercício (6). 8. Confira o Exemplo 6.22. 9. É fácil verificar que $V = \{\mathbf{X} \in \mathbb{C}^{n\times n} : \mathbf{AX} = \mathbf{XB}\}$ é um espaço vetorial de dimensão finita sobre $\mathbb{C}$, com uma base $\alpha = \{\mathbf{X}_1, \ldots, \mathbf{X}_m\}$ e $W = F[\mathbf{X}_1, \ldots, \mathbf{X}_m]$ um espaço vetorial sobre F. Assim, $\det : V \to \mathbb{C}$ é sobrejetora, pois existe um $\mathbf{P} \in V$ tal que $\det \mathbf{P} \neq 0$. Logo, usando

indução sobre m, $\det|_W \neq 0$. Portanto, existe uma matriz não singular $\mathbf{Q} \in F^{n\times n}$ tal que $\mathbf{B} = \mathbf{Q}^{-1}\mathbf{AQ}$. Finalmente, sejam $\mathbf{J}$ um bloco de Jordan de $\mathbf{A}$ e $\mathbf{R}_k$ a identidade reversa. Então é fácil verificar que $\mathbf{R}_k^{-1}\mathbf{J}\mathbf{R}_k = \mathbf{J}^t$. Como $\mathbf{J}^t$ é um bloco de Jordan de $\mathbf{A}^t$ temos que $\mathbf{A}$ é semelhante a $\mathbf{A}^t$. 10. Basta considerar apenas um bloco de Jordan. Assim, $\mathbf{J}^{-1}\mathbf{J} = \mathbf{I}$ se, e somente se, $a_{ij} = 0$, quando $j \notin \{i, i+1\}$. Tome $\mathbf{J} = \mathbf{I} + \mathbf{E}_{12} + \mathbf{E}_{23}$, de modo que $\mathbf{J}^{-1}$ não está na forma de Jordan. 11. Sejam $\lambda_1, \ldots, \lambda_n \in F$ os autovalores de T. Enão, pelo Exercício (5) da Seção 5.1, os autovalores de T^m são λ_i^m. Se $\mathbf{J}$ for a forma de Jordan de T, então $\operatorname{tr}(T^i) = 0$ é equivalente ao sistema $\lambda_1^i + \cdots + \lambda_n^i = 0$, para $i = 1, \ldots, n$. Se os λ_i são todos iguais, o resultado segue. Assim, podemos supor, sem perda de generalidade, que os λ_i são todos distintos. Logo, $\mathbf{V}_n\mathbf{X} = \mathbf{O}$, com $\mathbf{X} = (\lambda_1, \ldots, \lambda_n)^t$. Como $\det(\mathbf{V}_n) \neq 0$ temos que $\mathbf{X} = \mathbf{O}$. 12. Sejam $\mathbf{R}_x = (x\mathbf{I} - \mathbf{A})^{-1}, \mathbf{J}$ a forma de Jordan de $\mathbf{A}$ e $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ os autovalores de $\mathbf{A}$. Então, imite o Exercício (38) da Seção 5.3, para obter

$$f(x)\mathbf{R}_x = x^{n-1}\mathbf{I} + \sum_{i=1}^{n-1}\mathbf{B}_i x^{n-i-1} \text{ e } \operatorname{tr}(\mathbf{R}_x) = \sum_{i=1}^{n}(x-\lambda_i)^{-1} = \frac{f'(x)}{f(x)}.$$

Como $f(x)\operatorname{tr}(\mathbf{R}_x) = nx^{n-1} + \sum_{i=1}^{n-1}\operatorname{tr}(\mathbf{B}_i)x^{n-i-1}$ temos, comparando os coeficientes, que $(n-i)c_i = \operatorname{tr}(\mathbf{B}_i)$, para $i = 1, \ldots, n-1$. Finalmente, use

$$\mathbf{B}_i = \mathbf{B}_{i-1}\mathbf{A} + c_i\mathbf{I} \text{ e } \mathbf{B}_{n-1}\mathbf{A} + c_n\mathbf{I} = \mathbf{0}, i = 1, \ldots, n-1.$$

13. Note que $f(E) = 2E - I$ e $f^{-1}(U) = 2^{-1}(U + I)$ 14. Tome $U = W_1$ e $W = W_2 \oplus \cdots \oplus W_k$ no Teorema 6.4.

Capítulo 7

Seção 7.1

1. Confira o Exemplo 7.2. 2. Não, pois $\mathbf{u} = (1,0) \neq \mathbf{0}$, mas $f(\mathbf{u},\mathbf{u}) = 0$. 3. Não, pois se $\mathbf{u} = (0,0,1)$, então $f(\mathbf{u},\mathbf{u}) = -1 < 0$. 4. Confira o Exemplo 7.2. 5. Não, pois $\mathbf{u} = (1,0) \neq \mathbf{0}$, mas $f(\mathbf{u},\mathbf{u}) = 0$. 6. Identifique V com $\mathbb{R}^2$. 7. Sejam $\mathbf{u} = (x_1, x_2)$ e $\mathbf{v} = (y_1, y_2)$. Então, depois de alguns cálculos,

$$f(\mathbf{u},\mathbf{v}) = x_1y_1f(\mathbf{e}_1,\mathbf{e}_1) + x_1y_2f(\mathbf{e}_1,\mathbf{e}_2) + x_2y_1f(\mathbf{e}_2,\mathbf{e}_1) + x_2y_2f(\mathbf{e}_2,\mathbf{e}_2).$$

Como $T(\mathbf{e}_1) = \mathbf{e}_2$ e $T(\mathbf{e}_2) = -\mathbf{e}_1$ temos que

$$f(\mathbf{u}, \mathbf{v}) = x_1 y_1 f(\mathbf{e}_1, \mathbf{e}_1) + x_2 y_2 f(\mathbf{e}_2, \mathbf{e}_2).$$

Em particular, para $\mathbf{u} = (1, 1)$ e $\mathbf{v} = T(\mathbf{u}) = (-1, 1)$, temos que $f(\mathbf{e}_1, \mathbf{e}_1) = f(\mathbf{e}_2, \mathbf{e}_2) = c > 0$. Portanto, $f(\mathbf{u}, \mathbf{v}) = c(x_1 y_1 + x_2 y_2)$. 8. (a) T não singular. (b) Claramente g satisfaz as condições $(1), (2)$ e (3). Assim, resta provar que $g(\mathbf{u}, \mathbf{u}) = 0$ implica que $\mathbf{u} = \mathbf{0}$. Para isto, $\langle T(\mathbf{u}), T(\mathbf{u})\rangle = 0$ se, e somente se, $T(\mathbf{u}) = \mathbf{0}$. Portanto, a condição é que T seja não singular. 9. Basta notar que $\mathbf{Y}^*\mathbf{A}\mathbf{X} = (\mathbf{A}^*\mathbf{Y})^*\mathbf{X}$. 10. Se $(\mathbf{u}, \mathbf{w}) \neq (\mathbf{0}, \mathbf{0})$, digamos $\mathbf{u} \neq \mathbf{0}$, então $h((\mathbf{u}, \mathbf{w}), (\mathbf{u}, \mathbf{w})) = f(\mathbf{u}, \mathbf{u}) + g(\mathbf{w}, \mathbf{w}) > 0$. 11. Suponhamos que f seja um produto interno sobre V. Então $f(\mathbf{E}_1, \mathbf{E}_2) = f(\mathbf{E}_2, \mathbf{E}_1)$ implica que $\mathbf{A} = \mathbf{A}^t$. Como $f(\mathbf{X}, \mathbf{X}) > 0$, se $\mathbf{X} \neq \mathbf{O}$, temos que $a_{11} = f(\mathbf{E}_1, \mathbf{E}_1) > 0$ e $a_{22} = f(\mathbf{E}_2, \mathbf{E}_2) > 0$. Em geral,

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 = a_{11}\left(x + \frac{a_{12}}{a_{11}}y\right)^2 + \left(\frac{a_{11}a_{22} - a_{12}^2}{a_{11}}\right)y^2 > 0$$

implica que $f(a_{12}\mathbf{E}_1, -a_{11}\mathbf{E}_2) = a_{11} \det \mathbf{A} > 0$. A recíproca segue da equação. 12. A diferença de dois produtos internos sobre V não é um produto interno sobre V. O resto é direto. 13. Note que $\alpha = \{1\}$ é uma base de F. Então $\langle x, y\rangle = xy$ é um produto interno sobre F. Em geral, pelo Exemplo 4.3, qualquer $T \in \mathcal{L}(F)$ é da forma $T(x) = ax$, para algum $a \in F$. Portando, se $a, b \neq 0$, então $\langle x, y\rangle = T(x)T(y) = (ab)xy$ é um produto interno sobre F. 14. Confira o Exercício (17) da Seção 1.2. 15. Dado $\mathbf{u} \in V$, existem únicos $a_i \in F$ tais que $\mathbf{u} = \sum_{i=1}^n a_i \mathbf{u}_i$ e $\langle \mathbf{u}_j, \mathbf{u}\rangle = \sum_{i=1}^n \bar{a}_i \langle \mathbf{u}_j, \mathbf{u}_i\rangle$. Assim, é suficiente provar que o sistema $\sum_{i=1}^n \langle \mathbf{u}_j, \mathbf{u}_i\rangle x_i = c_j$, com $j = 1, \ldots, n$, possui uma única solução. 16. Seja $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ uma base de V. Então, já sabemos que a função $L_\alpha : V \to F^n$ definida como $L_\alpha(\mathbf{u}_i) = \mathbf{e}_i$, era um isomorfismo. Assim, $f(\mathbf{u}, \mathbf{v}) = \langle L_\alpha(\mathbf{u}), L_\alpha(\mathbf{v})\rangle$ é um produto interno sobre V, com $\langle \cdot, \cdot\rangle$ qualquer produto interno sobre F^n. 17. Vamos provar apenas a condição (4). É claro que $\langle f, f\rangle \geq 0$, pois $f(x)^2 \geq 0$. Logo, resta provar que se $\langle f, f\rangle = 0$, então $f(x) = 0$, para todo $x \in [0, 1]$. Suponhamos, por absurdo, que $g(x) = f(x)^2$ e $g(x_0) \neq 0$, para algum $x_0 \in [0, 1]$. Então $g(x_0) > 0$. Como g é contínua temos que existe um $\delta > 0$ tal que $(x_0 - \delta, x_0 + \delta) \subseteq [0, 1]$ e $|x - x_0| < \delta$ implica que $|g(x) - g(x_0)| < \frac{1}{2}g(x_0)$. Em particular, $\frac{1}{2}g(x_0) < g(x)$. Definido $h(x) = \frac{1}{2}g(x_0)$, se $x \in (x_0 - \delta, x_0 + \delta)$, e $h(x) = 0$, caso contrário, obtemos $h(x) \leq g(x)$ e

$$\int_0^1 g(t)dt \geq \int_0^1 h(t)dt = \frac{1}{2}g(x_0)(2\delta) = g(x_0)\delta > 0,$$

o que é uma contradição. 18. Confira o Exercício (12) e use o Exercício (17).

Seção 7.2

1. (2) $d(\mathbf{u},\mathbf{v}) \leq d(\mathbf{u},\mathbf{u}) + d(\mathbf{v},\mathbf{u}) = d(\mathbf{v},\mathbf{u})$. 2. (b) $(\sqrt{d(\mathbf{u},\mathbf{v})} + \sqrt{d(\mathbf{v},\mathbf{w})})^2$. (c) Basta verificar que a função $f : [0,\infty) \to F$ definida como $f(x) = \frac{x}{1+x}$ é crescente. 3. Se $d_1(\mathbf{u},\mathbf{v}) = 0$, então $f(\mathbf{u}) = f(\mathbf{v})$, de modo que $\mathbf{u} = \mathbf{v}$. 4. Note que $d(\mathbf{u},\mathbf{v}) \leq d(\mathbf{u},\mathbf{w}) + d(\mathbf{w},\mathbf{v})$ e $d(\mathbf{w},\mathbf{v}) \leq d(\mathbf{w},\mathbf{u}) + d(\mathbf{u},\mathbf{v})$ implicam que $-d(\mathbf{u},\mathbf{w}) \leq d(\mathbf{u},\mathbf{v}) - d(\mathbf{v},\mathbf{w}) \leq d(\mathbf{u},\mathbf{w})$. 5. Note, de modo similar ao Exercício (4), que $|\|\mathbf{u}_n\| - \|\mathbf{u}\|| \leq \|\mathbf{u}_n - \mathbf{u}\|$, para cada $n \in \mathbb{N}$. 6. $\theta = \arccos(\frac{1}{\sqrt{5}})$. 7. Note que $\mathbf{w} = \mathbf{u} - \mathbf{v}$ é o terceiro lado e $\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\langle \mathbf{u},\mathbf{v}\rangle$. 8. Basta desenvolver o lado esquerdo. 9. Observe, depois de alguns cálculos, que

$$\begin{array}{rcl}
\frac{1}{2}\|\mathbf{u}-\mathbf{v}\|^2 + 2\|\mathbf{w} - \frac{1}{2}(\mathbf{u}+\mathbf{v})\|^2 & = & \frac{1}{2}(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\langle \mathbf{u},\mathbf{v}\rangle) \\
& + & 2(\|\mathbf{w}\|^2 + \frac{1}{4}\|\mathbf{u}+\mathbf{v}\|^2 - \langle \mathbf{u}+\mathbf{v},\mathbf{w}\rangle) \\
& = & \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\|\mathbf{w}\|^2 - 2\langle \mathbf{u},\mathbf{w}\rangle) \\
& - & 2\langle \mathbf{v},\mathbf{w}\rangle) \\
& = & \|\mathbf{w}-\mathbf{u}\|^2 + \|\mathbf{w}-\mathbf{v}\|^2.
\end{array}$$

10. Pela identidade de polarização, $f(\mathbf{u},\mathbf{v}) = g(\mathbf{u},\mathbf{v})$, para todos $\mathbf{u},\mathbf{v} \in V$. Portanto, $f = g$. 11. (a) Basta notar que se $|x_n - x| \to 0$ e $\|\mathbf{u}_n - \mathbf{u}\| \to 0$, então $\|x_x\mathbf{u}_n - x\mathbf{u}\| \leq |x_n - x|\|\mathbf{u}_n\| + |x|\|\mathbf{u}_n - \mathbf{u}\|$ garante a afirmação. 12. Observe que se $\mathbf{v} \neq \mathbf{0}$, então

$$f(x) = \|\mathbf{u}\|^2 + 2\langle \mathbf{u},\mathbf{v}\rangle x + \|\mathbf{v}\|^2x^2 = \frac{\|\mathbf{u}\|^2\|\mathbf{v}\|^2 - \langle \mathbf{u},\mathbf{v}\rangle^2}{\|\mathbf{v}\|^2} + (\frac{\langle \mathbf{u},\mathbf{v}\rangle}{\|\mathbf{v}\|} + \|\mathbf{v}\|x)^2$$

Assim, o mínimo de $f(x)$ ocorre em $x = -\langle \mathbf{u},\mathbf{v}\rangle\|\mathbf{v}\|^{-2}$. Note que $f(x)$ não possui valor máximo, pois $\lim_{x\to\pm\infty} f(x) = \infty$. 13. Basta notar que $(\|\mathbf{u}\| + \|\mathbf{v}\|)^2 = \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2$ e usar igualdade de Cauchy-Schwarz. 14. Seja $S = \{\mathbf{x} \in V : \|\mathbf{x}\|_g = 1\}$. Então a função $h : S \to \mathbb{R}$ definida como $h(\mathbf{x}) = \|\mathbf{x}\|_f^{-1}\|\mathbf{x}\|_g$ é positiva e contínua. Assim, possui um valor máximo $c = \max\{h(\mathbf{x}) : \mathbf{x} \in S\}$. Logo, $\|\mathbf{x}\|_g \leq c\|\mathbf{x}\|_f$, para todo $\mathbf{x} \in S$, continue! 15. Se $\mathbf{0} \in \{\mathbf{u},\mathbf{v},\mathbf{w}\}$, nada há para ser provado. Caso contrário, pondo $\mathbf{u}_1 = \|\mathbf{u}\|^{-2}\mathbf{u}, \mathbf{v}_1 = \|\mathbf{v}\|^{-2}\mathbf{v}$ e $\mathbf{w}_1 = \|\mathbf{w}\|^{-2}\mathbf{w}$, obtemos $\|\mathbf{u}_1 - \mathbf{v}_1\| = \|\mathbf{u} - \mathbf{v}\|^2\|\mathbf{u}\|^{-2}\|\mathbf{v}\|^{-2}$. Assim, a desigualdade triangular, $\|\mathbf{u}_1 - \mathbf{v}_1\| \leq \|\mathbf{u}_1 - \mathbf{w}_1\| + \|\mathbf{w}_1 - \mathbf{v}_1\|$ implica o resultado. 16. Note que $\|\mathbf{u}\|^2 = \mathbf{X}^t\mathbf{A}\mathbf{X}$, com $\mathbf{A}$ uma matriz simétrica definida positiva, e confira a Proposição 7.10.

Seção 7.3

1. Basta notar que $\mathbf{A}^*\mathbf{A} = (\mathbf{C}_i^*\mathbf{C}_j) = (d_{ij}\delta_{ij}) = \mathbf{D}$. 2. Imite o Exemplo 7.17. 3.

Cálculo direto. 4. Note, depois de alguns cálculos, que

$$\|\sum_{i=1}^{n}\mathbf{u}_i\|^2=\sum_{i=1}^{n}\sum_{j=1}^{n}\langle\mathbf{u}_i,\mathbf{u}_j\rangle=\sum_{i=1}^{n}\langle\mathbf{u}_i,\mathbf{u}_i\rangle+\sum_{i\neq j}\langle\mathbf{u}_i,\mathbf{u}_j\rangle=\sum_{i=1}^{n}\|\mathbf{u}_i\|^2.$$

5. Observe que f é contínua e S é "compacto" implicam que f possui pelo menos um ponto de máximo e um ponto de mínimo sobre S. Vamos lembrar que um *máximo* sobre S é um ponto $\mathbf{p}\in S$ tal que $f(\mathbf{p})\geq f(\mathbf{v})$, para todo $\mathbf{v}\in S$. Note que $f(\mathbf{v})=\sqrt{2-2\langle\mathbf{u},\mathbf{v}\rangle}$. Assim, basta minimizar e maximizar $\langle\mathbf{u},\mathbf{v}\rangle$. Por outro lado, pela desigualdade de Cauchy-Schwarz, $-1\leq\langle\mathbf{u},\mathbf{v}\rangle\leq 1$, de modo que -1 é o mínimo e 1 é o máximo de $\langle\mathbf{u},\mathbf{v}\rangle$ sobre S. Portanto, o valor máximo de f é 2 e o valor mínimo de f é 0. (b) $f(\mathbf{v})=\sqrt{2}$ se, e somente se, $\langle\mathbf{u},\mathbf{v}\rangle=0$. 6. É claro que $L_i(x_j)=\delta_{ij}$. Assim, $f(L_i(x),L_j(x))=\sum_{k=0}^{n}L_i(c_k)L_j(c_k)=\sum_{k=0}^{n}\delta_{ik}\delta_{jk}=\delta_{ij}$.

Seção 7.4

1. Sejam $V=F^2$ com o produto interno usual e $\beta=\{\mathbf{u},\mathbf{v}\}$, uma base ortogonal de V, com $\mathbf{u}=(1,1)$ e $\mathbf{v}=(1,-1)$. Escolhendo $\mathbf{w}=(5,-3)\in V$, obtemos $\langle 4\mathbf{u}+(-1)\mathbf{v},\mathbf{w}\rangle=0$, mas $\langle\mathbf{u},\mathbf{w}\rangle=2$ e $\langle\mathbf{v},\mathbf{w}\rangle=8$. 2. $\beta=\{\mathbf{v}_1,\mathbf{v}_2\}$, com $\mathbf{v}_1=(-1,1)$ e $\mathbf{v}_2=(0,1)$. 3. Pela Proposição 7.19, V possui uma base ortonormal $\alpha=\{\mathbf{u}_1,\ldots,\mathbf{u}_n\}$. Assim, $L_\alpha:V\to F^n$ definida como $L_\alpha(\mathbf{v})=(x_1,\ldots,x_n)$, com x_i os coeficientes de Fourier de $\mathbf{v}$, é o isomorfismo desejado. 4. Escolhendo o vetor inicial, digamos $q_1(x)=\|p_1(x)\|^{-1}p_1(x)=\frac{1}{\sqrt{2}}$, de modo que

$$p_2(x)=x\ \text{ e }\ q_2(x)=\sqrt{\frac{2}{3}}x;p_3(x)=x^2-\frac{1}{3}\ \text{ e }\ q_3(x)=\frac{\sqrt{10}}{4}(-1+3x^2),$$

e assim por diante. Os polinômios $q_i(x)$ da base ortonormal $\beta=\{q_i(x):i\in\mathbb{N}\}$ de V chamam-se, a menos de constantes, de *polinômios de Legendre*[3] 5. É fácil verificar que $\alpha=\{(1,1,0),(0,1,1)\}$ é uma base de W e $\beta=\{\mathbf{v}_1,\mathbf{v}_2\}$, com $\mathbf{v}_1=\frac{1}{\sqrt{2}}(1,1,0)$ e $\mathbf{v}_2=\frac{1}{\sqrt{6}}(-1,1,2)$, é uma base ortonormal de W. Assim,

$$\mathbf{w}=\sum_{i=1}^{2}\langle\mathbf{u},\mathbf{v}_i\rangle\mathbf{v}_i=\frac{3}{2}(1,1,0)+\frac{7}{6}(-1,1,2)=\frac{1}{3}(1,8,7)$$

é a melhor aproximação de $\mathbf{u}$ por vetores de W. Portanto, $\mathbf{u}-\mathbf{w}=\frac{1}{3}(2,-2,2)$ e

[3] Adrien-Marie Legendre, 1752-1833, matemático francês.

$d(\mathbf{u}, W) = \|\mathbf{u} - \mathbf{w}\| = \frac{2}{3}\sqrt{3}$.

Seção 7.5

1. Observe que $W = F[(1,1,0)]$ e imite o Exemplo 7.22. 2. Basta observar que $W^{\perp} = F[(1,-1,1)]$ e $V = W \oplus W^{\perp}$ implica, por exemplo, que $T(\mathbf{w}_i) = \mathbf{w}_i$ e $T(1,-1,1) = (0,0,0)$. Podemos também imitar o Exemplo 7.30. 3. Confira Exercício (1). 4. Confira a Proposição 7.25 ou imite o Exemplo 7.30. 5., 6. e 7. Cálculo direto. 8. Imite o Exemplo 7.30. 9. Pondo $n = k + m$, é fácl verificar que

$$\langle x^k, x^m \rangle = \int_{-1}^{1} t^n dt = \begin{cases} \frac{2}{n+1}, & \text{se } \; n \text{ é par} \\ 0, & \text{se } \; n \text{ é ímpar.} \end{cases}$$

Depois de alguns cálculos, obtemos $\beta = \{f_i\} = \{1, x, 3x^2 - 1, 5x^3 - 3x\}$ uma base ortogonal de W e

$$\begin{array}{rcl} P(e^x) & = & \sum_{i=0}^{3} \langle e^x, f_i \rangle \|f_i\|^{-2} f_i \\ & = & \frac{1}{2}(e - e^{-1}) f_0 + 3e^{-1} f_1 + \frac{5}{4}(e - 7e^{-1}) f_2 - \frac{7}{4}(5e - 37e^{-1}) f_3 \end{array}$$

é a melhor aproximação. 10. Observe que $W = F[\mathbf{E}_{11}, \mathbf{E}_{12} + \mathbf{E}_{21}, \mathbf{E}_{22}]$. Assim, dado $\mathbf{A} \in V$, obtemos $\operatorname{tr}(\mathbf{B}^*\mathbf{A}) = \operatorname{tr}(\mathbf{BA}) = 0$, para todo $\mathbf{B} \in W$, se e somente se, $\operatorname{tr}(\mathbf{F}_{ij}\mathbf{A}) = 0$, se e somente se, $\mathbf{A} = \operatorname{diag}(d, -d)$, para todo $d \in F$. Portanto, $W^{\perp} = F[\operatorname{diag}(1,-1)]$. 11. Note que se $h \in V$ for uma função ímpar, então

$$\int_{-1}^{1} h(t)dt = -\int_{-1}^{1} h(t)dt \Rightarrow \int_{-1}^{1} h(t)dt = 0.$$

Portanto, pelo Teorema 7.24, $W^{\perp}$ é o subespaço das funções pares. 12. (a) Se $\mathbf{v} \in U$, então $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, para todo $\mathbf{u} \in U^{\perp}$, de modo que $\mathbf{v} \in U^{\perp\perp}$. Portanto, $U \subseteq U^{\perp\perp}$. Observe que não necessitamos da dimensão na prova. Por outro lado, use que $U \oplus U^{\perp} = V = U^{\perp} \oplus U^{\perp\perp}$ e a dimensão. Note que se $\mathbf{v} \in W^{\perp}$, então $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, para todo $\mathbf{u} \in W$. Em particular, $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, para todo $\mathbf{u} \in U$. Portanto, $W^{\perp} \subseteq U^{\perp}$. (b) Na afirmação $(U + W)^{\perp} = U^{\perp} \cap W^{\perp}$ não necessitamos da dimensão na prova. Como $U \cap W \subseteq U, W$ temos, pelo item (a), que $U^{\perp}, W^{\perp} \subseteq (U \cap W)^{\perp}$, de modo que $U^{\perp} + W^{\perp} \subseteq (U \cap W)^{\perp}$. Por outro lado,

$$\begin{array}{rcl} \dim(U^{\perp} + W^{\perp}) & = & \dim U^{\perp} + \dim W^{\perp} - \dim(U^{\perp} \cap W^{\perp}) \\ & = & \dim V - \dim U - \dim W + \dim V - \dim(U + W)^{\perp} \\ & = & \dim V - \dim U - \dim W + \dim(U + W) \\ & = & \dim V - \dim(U \cap W) = \dim(U \cap W)^{\perp}; \end{array}$$

Portanto, $(U \cap W)^\perp = U^\perp + W^\perp$. 13. Como $\|\mathbf{u} \pm \mathbf{v}\|^2 = 2 \pm 2\,\mathrm{Re}\langle \mathbf{u}, \mathbf{v}\rangle \geq 0$ temos que $|\langle \mathbf{u}, \mathbf{v}\rangle| \leq 1$. Se $\mathbf{0} \in \{\mathbf{u}, \mathbf{v}\}$, nada há para ser provado. Caso contrário

$$|\langle \frac{1}{\|\mathbf{u}\|}\mathbf{u}, \frac{1}{\|\mathbf{v}\|}\mathbf{v}\rangle| \leq 1 \Leftrightarrow |\langle \mathbf{u}, \mathbf{v}\rangle| \leq \|\mathbf{u}\|\|\mathbf{v}\|.$$

14. Note que a condição é equivalente a: $2\,\mathrm{Re}\langle \mathbf{u}, \mathbf{w}\rangle \leq \|\mathbf{w}\|^2$, para todo $\mathbf{w} \in W$, de modo que $\|\mathbf{u} - \mathbf{w}\|^2 = \|\mathbf{u}\|^2 - 2\,\mathrm{Re}\langle \mathbf{u}, \mathbf{w}\rangle + \|\mathbf{w}\|^2 \geq \|\mathbf{u}\|^2$ e o resultado segue do Exemplo 7.18. Uma prova direta. Pondo $\mathbf{w} = \langle \mathbf{u}, \mathbf{v}\rangle\|\mathbf{v}\|^{-2}\mathbf{v}$, para todo $\mathbf{v} \in W$, temos que $\langle \mathbf{u}, \mathbf{w}\rangle = |\langle \mathbf{u}, \mathbf{v}\rangle|^2\|\mathbf{v}\|^{-2}$ é um número real, de modo que

$$\|\mathbf{w}\|^2 - 2\,\mathrm{Re}\langle \mathbf{u}, \mathbf{w}\rangle \geq 0 \Rightarrow -|\langle \mathbf{u}, \mathbf{v}\rangle|^2\|\mathbf{v}\|^{-2} \geq 0$$

Portanto, $\langle \mathbf{u}, \mathbf{v}\rangle = 0$. 15. (a) Pelo processo de Gram-Schmidt, V possui uma base ortonormal $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ tal que $\{\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}\}$ seja uma base ortonormal de H. Assim, o vetor $\mathbf{n} = \mathbf{u}_n$ ou a normalização da melhor aproximação de qualquer $\mathbf{u} \in V$ possui as propriedades desejadas. 16. Note que $\langle \mathbf{u}, \mathbf{w}\rangle = 0$ implica, para cada $\mathbf{z} \in S \cap r$, que $\|\mathbf{u}\|^2 = \|\mathbf{z}\|^2 = \|\mathbf{u} + t\mathbf{w}\|^2 = \|\mathbf{u}\|^2 + |t|^2\|\mathbf{w}\|^2$, de modo que $t = 0$ e $\mathbf{z} = \mathbf{u}$. 17. Basta notar que $\|\mathbf{v} - \mathbf{u}\|^2 = \|\mathbf{x} - \mathbf{u}\|^2 + \|\mathbf{x} - \mathbf{v}\|^2$. 18. Observe que $\langle T(\mathbf{X}), \mathbf{X}\rangle = 0$, para todo $\mathbf{X} \in V$, implica que $\langle T_{\mathbf{A}}(\mathbf{X}), \mathbf{Y}\rangle = -\langle T_{\mathbf{A}}(\mathbf{Y}), \mathbf{X}\rangle$, para todos $\mathbf{X}, \mathbf{Y} \in V$. Como $a_{ij} = \langle T_{\mathbf{A}}(\mathbf{E}_{ij}), \mathbf{E}_{ij}\rangle$ temos que $a_{11} = a_{22} = 0$ e $a_{12} = \langle T_{\mathbf{A}}(\mathbf{E}_{12}), \mathbf{E}_{21}\rangle = -\langle T_{\mathbf{A}}(\mathbf{E}_{21}), \mathbf{E}_{12}\rangle = -a_{21}$. Escolha $\lambda = a_{12} \in F$. 19. Como $\mathbf{v} = \sum_{i=1}^n \langle \mathbf{v}, \mathbf{u}_i\rangle\mathbf{u}_i$, para todo $\mathbf{v} \in V$, temos, em particular, que $T(\mathbf{v}_j) = \sum_{i=1}^n \langle T(\mathbf{v}_j), \mathbf{u}_i\rangle\mathbf{u}_i$. Portanto, $[T]_\alpha^\alpha = (\langle T(\mathbf{u}_j), \mathbf{u}_i\rangle)$. 20. Seja $\mathbf{A}$ a matriz cujas colunas são os vetores de α. Então $\mathbf{G}_\alpha = \mathbf{A}^*\mathbf{A}$, $\mathbf{G}_\alpha^* = \mathbf{G}_\alpha$ e $|\det \mathbf{G}_\alpha| \geq 0$. Dado $\mathbf{X} = (x_1, \ldots, x_n)^t \in F^{n\times 1}$, depois de alguns cálculos, obtemos

$$\mathbf{X}^*\mathbf{G}_\alpha\mathbf{X} = (\mathbf{A}\mathbf{X})^*(\mathbf{A}\mathbf{X}) = \|x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n\|^2 \geq 0.$$

Assim, α é LD se, e somente se, $\mathbf{A}$ for singular se, e somente se, $\det \mathbf{G}_\alpha = 0$. Uma outra prova é dada pelo Lema 7.28. 21. Note que $\mathbf{u} \in \ker T$ implica que $\|\mathbf{u}\| = \|T(\mathbf{u})\| = 0$ e $\mathbf{u} = \mathbf{0}$. 22. Seja $\mathbf{z} = \mathbf{v} - \sum_{i=1}^n \langle \mathbf{v}, \mathbf{u}_i\rangle\mathbf{u}_i$. Então, depois de alguns cálculos, $0 \leq \langle \mathbf{z}, \mathbf{z}\rangle = \langle \mathbf{v}, \mathbf{v}\rangle - \sum_{i=1}^n \langle \mathbf{v}, \mathbf{u}_i\rangle\overline{\langle \mathbf{v}, \mathbf{u}_i\rangle} = \|\mathbf{v}\|^2 - \sum_{i=1}^n |\langle \mathbf{v}, \mathbf{u}_i\rangle|^2$. A equivalência $(a \Leftrightarrow c)$ é clara. $(b \Rightarrow c)$ Isto implica que $\mathbf{z} = \mathbf{0}$. $(c \Rightarrow d)$ Para qualquer $\mathbf{v} \in V$, obtemos $\langle \mathbf{w}, \mathbf{v}\rangle = \langle \sum_{i=1}^n \langle \mathbf{w}, \mathbf{u}_i\rangle\mathbf{u}_i, \mathbf{v}\rangle = \sum_{i=1}^n \langle \mathbf{w}, \mathbf{u}_i\rangle\langle \mathbf{u}_i, \mathbf{v}\rangle$. $(d \Rightarrow b)$ Basta tomar $\mathbf{v} = \mathbf{w}$. 23. É fácil verificar que $\mathbf{v} = \sum_{i=1}^n \langle \mathbf{v}, \mathbf{u}_i\rangle\mathbf{u}_i$, para todo $\mathbf{v} \in V$. Assim, depois de alguns cálculos,

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle \sum_{i=1}^{n} \langle \mathbf{v}, \mathbf{u}_i \rangle \mathbf{u}_i, \sum_{i=1}^{n} \langle \mathbf{w}, \mathbf{u}_j \rangle \mathbf{u}_j \rangle = \sum_{i=1}^{n} \langle \mathbf{v}, \mathbf{u}_i \rangle \langle \mathbf{w}, \mathbf{u}_i \rangle.$$

Observe que isto é uma generalização do Teorema de Pitágoras. 24. Sejam $\mathbf{u}, \mathbf{v} \in V$. Então $\mathbf{u}$ e $\mathbf{v}$ possuem coordenadas polares $(\|\mathbf{u}\|, \theta)$ e $(\|\mathbf{v}\|, \phi)$. Então

$$\cos(\theta - \phi) = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\|\mathbf{v}\|}, \cos\theta = \frac{\langle \mathbf{u}, \mathbf{e}_1 \rangle}{\|\mathbf{u}\|} \text{ e } \operatorname{sen}\theta = \frac{\langle \mathbf{u}, \mathbf{e}_2 \rangle}{\|\mathbf{u}\|}.$$

Assim, pelo Exercício (23),

$$\cos(\theta - \phi) = \frac{\langle \mathbf{u}, \mathbf{e}_1 \rangle \langle \mathbf{v}, \mathbf{e}_1 \rangle + \langle \mathbf{u}, \mathbf{e}_2 \rangle \langle \mathbf{v}, \mathbf{e}_2 \rangle}{\|\mathbf{u}\|\|\mathbf{v}\|} = \cos\theta\cos\phi + \operatorname{sen}\theta\operatorname{sen}\phi.$$

25. Pondo $P(\mathbf{u}) = \sum_{i=1}^{n} c_i \mathbf{v}_i$, obtemos $d(\mathbf{u}, W)^2 + \sum_{i=1}^{n} c_i \langle \mathbf{u}, \mathbf{v}_i \rangle = \|\mathbf{u}\|^2$, pois $d(\mathbf{u}, W)^2 = \|\mathbf{u} - P(\mathbf{u})\|^2$. Como $\langle \mathbf{u}, \mathbf{v}_j \rangle = \langle P(\mathbf{u}), \mathbf{v}_j \rangle = \sum_{i=1}^{n} c_i \langle \mathbf{v}_j, \mathbf{v}_i \rangle$, para $j = 1, \ldots, n$, temos que $d(\mathbf{u}, W)^2$ e $c_1, \ldots, c_n$ são soluções do sistema

$$\begin{cases} x_0 + \langle \mathbf{u}, \mathbf{v}_1 \rangle x_1 + \cdots + \langle \mathbf{u}, \mathbf{v}_n \rangle x_n &= \|\mathbf{u}\|^2 \\ \langle \mathbf{v}_1, \mathbf{v}_1 \rangle x_1 + \langle \mathbf{v}_1, \mathbf{v}_2 \rangle x_2 + \cdots + \langle \mathbf{v}_1, \mathbf{v}_n \rangle x_n &= \langle \mathbf{u}, \mathbf{v}_1 \rangle \\ &\vdots \\ \langle \mathbf{v}_n, \mathbf{v}_1 \rangle x_1 + \langle \mathbf{v}_n, \mathbf{v}_2 \rangle x_2 + \cdots + \langle \mathbf{v}_n, \mathbf{v}_n \rangle x_n &= \langle \mathbf{u}, \mathbf{v}_n \rangle. \end{cases}$$

Portanto, o resultado segue da Regra de Cramer. 26. Confira o Exercício (15) da Seção 7.1 e/ou o Exercício (20). 27. Aplique a desigualdade de Cauchy-Schwarz aos vetores $\mathbf{u} = (\sqrt{a_1^{\lambda}}, \ldots, \sqrt{a_n^{\lambda}})$ e $\mathbf{v} = (\sqrt{a_1^{1-\lambda}}, \ldots, \sqrt{a_n^{1-\lambda}})$. 28. Cálculo direto. 29. Observe, pela desigualdade de Cauchy-Schwarz, que $\langle \mathbf{v}, \mathbf{w} \rangle \langle \mathbf{w}, \mathbf{v} \rangle \leq \langle \mathbf{w}, \mathbf{w} \rangle$. Em particular, para $\mathbf{w} = T(\mathbf{v})$. A igualdade ocorre se, e somente se, $\mathbf{AX}$ e $\mathbf{X}$ são LD. Como $\|\mathbf{X}\| = 1$ temos que $\mathbf{AX} = \lambda\mathbf{X}$, para algum $\lambda \in \mathbb{C}$. Portanto, $\mathbf{X}$ é um autovetor de $\mathbf{A}$.

Capítulo 8

Seção 8.1

1. Vamos provar os itens (1), (2) e (3) do Teorema 8.9: (1) Seja $\mathbf{A} = [T]$. Então

$$\overline{\det \mathbf{A}} = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) \prod_{i=1}^{n} \bar{a}_{i\sigma(i)} = \det \overline{\mathbf{A}} = \det \overline{\mathbf{A}}^t = \det \mathbf{A}^*.$$

(2) A equação $T(\mathbf{u}) = \mathbf{b}$ possui uma solução se, e somente se, $\mathbf{b} \in \operatorname{Im} T = (\ker T^*)^\perp$ se, e somente se, $\mathbf{b} \perp \mathbf{v}$, para todo $\mathbf{v} \in \ker T^*$. (3) Dado $\mathbf{v} \in W^\perp$ e $\mathbf{w} \in W$. Como $T(\mathbf{w}) \in W$ temos que $0 = \langle \mathbf{v}, T(\mathbf{w}) \rangle = \langle T^*(\mathbf{v}), \mathbf{w} \rangle$, de modo que $T^*(\mathbf{v}) \in W^\perp$ ou $T^*(W^\perp) \subseteq W^\perp$ (não necessitamos da dimensão). Para recíproca, use $W^{\perp\perp} = W$. 2. Como

$$\begin{aligned}
\langle (x,y), T^*(z,w) \rangle &= \langle T(x,y), (z,w) \rangle = \langle (x+iy, -2x-y), (z,w) \rangle \\
&= (x+iy)\bar{z} + (-2x-y)\bar{w} = x(\bar{z} - 2\bar{w}) + y(i\bar{z} - \bar{w}) \\
&= x\overline{(z-2w)} + y\overline{(-iz-w)} = \langle (x,y), (z-2w, -iz-w) \rangle
\end{aligned}$$

temos que $T^*(z,w) = (z-2w, -iz-w)$. 3. É fácil verificar que $T^*(a) = a\mathbf{v}$. 4. Observe que

$$\begin{aligned}
\langle (x_1, \ldots, x_n), T^*(y_1, \ldots, y_n) \rangle &= \langle T(x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle \\
&= \langle (0, x_1, \ldots, x_{n-1}), (y_1, \ldots, y_n) \rangle \\
&= x_1\bar{y}_2 + \cdots + x_{n-1}\bar{y}_n \\
&= \langle (x_1, \ldots, x_n), (y_2, \ldots, y_n, 0) \rangle
\end{aligned}$$

implica que $T^*(y_1, \ldots, y_n) = (y_2, \ldots, y_n, 0)$. 5. Como $\mathbf{u} = \sum_{i=1}^n \langle \mathbf{u}, \mathbf{u}_i \rangle \mathbf{u}_i$, para todo $\mathbf{u} \in V$, temos que

$$\begin{aligned}
\langle \mathbf{u}, T^*(\mathbf{v}) \rangle &= \langle T(\mathbf{u}), \mathbf{v} \rangle = \textstyle\sum_{i=1}^n \langle \mathbf{u}, \mathbf{u}_i \rangle \langle T(\mathbf{u}_i), \mathbf{v} \rangle \\
&= \textstyle\sum_{i=1}^n \langle \mathbf{u}, \mathbf{u}_i \rangle \overline{\langle \mathbf{v}, T(\mathbf{u}_i) \rangle} = \sum_{i=1}^n \langle \mathbf{u}, \langle \mathbf{v}, T(\mathbf{u}_i) \rangle \mathbf{u}_i \rangle \\
&= \langle \mathbf{u}, \textstyle\sum_{i=1}^n \langle \mathbf{v}, T(\mathbf{u}_i) \rangle \mathbf{u}_i \rangle.
\end{aligned}$$

Portanto, $T^*(\mathbf{v}) = \sum_{i=1}^n \langle \mathbf{v}, T(\mathbf{u}_i) \rangle \mathbf{u}_i$. 6. É fácil verificar que

$$T^*(x,y,z) = (x+3y+z, x+y+3z, x+y+3z)$$

e $\ker T^* = F[(-4,1,1)]$. Como $\langle (-4,1,1), (3,10,1) \rangle = -1 \neq 0$ temos que $(3,10,1) \notin$

$(\ker T^*)^{\perp}$. 7. Note, por integração por partes, que

$$\langle D(f(x)), g(x)\rangle = \int_0^1 f'(t)\overline{g(t)}dt = f(1)g(1) - f(0)g(0) - \langle f(x), D(g(x))\rangle.$$

Suponhamos, por absurdo, que $D^*(g(x))$ exista, para algum $g(x) \in V$ fixado, tal que $\langle D(f(x)), g(x)\rangle = \langle f(x), D^*(g(x))\rangle$, para todo $f(x) \in V$. Então

$$\langle f(x), D^*(g(x))\rangle = f(1)g(1) - f(0)g(0) - \langle f(x), D(g(x))\rangle,$$

de modo que

$$\langle f(x), D^*(g(x) + D(g(x)))\rangle = f(1)g(1) - f(0)g(0)$$

Como $L(f(x)) = f(1)g(1) - f(0)g(0)$ é um funcional linear temos, pelo Exemplo 8.3, que $L = 0$. Assim, $g(0) = g(1) = 0$. Reciprocamente, se $g(0) = g(1) = 0$, então $D^*(g(x)) = -D(g(x))$ satisfaz $\langle D(f(x)), g(x)\rangle = \langle f(x), D^*(g(x))\rangle$, para todo $f(x) \in V$. Portanto, se $g(0) \neq 0$ ou $g(1) \neq 0$, então D^* não existe. 8. (a) T é injetora se, e somente se, $\ker T = \{\mathbf{0}\}$ se, e somente se, $\operatorname{Im} T^* = (\ker T)^{\perp} = V$ se, e somente se, T^* for sobrejetora. 9. Como $V = W \oplus W^{\perp}$ temos que $E^2(v) = E^2(\mathbf{v}_1 + \mathbf{v}_2) = E(\mathbf{v}_1) = E(\mathbf{v})$ e

$$\begin{aligned} \langle \mathbf{u}, E^*(\mathbf{v})\rangle &= \langle E(\mathbf{u}_1 + \mathbf{u}_2), \mathbf{v}_1 + \mathbf{v}_2\rangle = \langle \mathbf{u}_1, \mathbf{v}_1 + \mathbf{v}_2\rangle = \langle \mathbf{u}_1, \mathbf{v}_1\rangle \\ &= \langle \mathbf{u}, \mathbf{v}_1\rangle = \langle \mathbf{u}, E(\mathbf{v})\rangle. \end{aligned}$$

Portanto, pela unicidade, $E^* = E = E^2$. Reciprocamente, $E^2 = E$ implica que $V = \operatorname{Im} E \oplus \ker E$. Dados $\mathbf{u}, \mathbf{v} \in V$ tais que $\mathbf{u} \in \operatorname{Im} E$ e $\mathbf{v} \in \ker E$, obtemos

$$\langle \mathbf{u}, \mathbf{v}\rangle = \langle E(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, E^*(\mathbf{v})\rangle = \langle \mathbf{u}, E(\mathbf{v})\rangle = 0.$$

Portanto, E é uma projeção ortogonal. 10. $(a \Rightarrow b)$ Suponhamos que T seja uma projeção ortogonal. Então $\mathbf{u} - T(\mathbf{u}) \in \ker T$ e $\mathbf{v} - T(\mathbf{v}) \in \ker T$, de modo que $\langle T(\mathbf{u}), \mathbf{v}\rangle = \langle T(\mathbf{u}), T(\mathbf{v})\rangle = \langle \mathbf{u}, T(\mathbf{v})\rangle$, para todos $\mathbf{u}, \mathbf{v} \in V$. $(b \Rightarrow c)$ Como $\langle \mathbf{u} - T(\mathbf{u}), T(\mathbf{v})\rangle = \langle T(\mathbf{u} - T(\mathbf{u})), \mathbf{v}\rangle = 0$, pois $T^2 = T$, temos que $\|\mathbf{v}\|^2 = \|\mathbf{v} - T(\mathbf{v}) + T(\mathbf{v})\|^2 = \|T(\mathbf{v}) - \mathbf{v}\|^2 + \|T(\mathbf{v})\|^2$. Portanto, $\|T(\mathbf{v})\| \leq \|\mathbf{v}\|$, para todo $\mathbf{v} \in V$. $(c \Rightarrow a)$ Dado $\mathbf{v} \in (\ker T)^{\perp}$, então é claro que $\mathbf{u} = T(\mathbf{v}) - \mathbf{v} \in \ker T$, de modo que $T(\mathbf{v}) = \mathbf{u} + \mathbf{v}$, com $\langle \mathbf{u}, \mathbf{v}\rangle = 0$. Assim, $\|\mathbf{v}\|^2 \geq \|T(\mathbf{v})\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \geq \|\mathbf{v}\|^2$. Logo, $\|\mathbf{u}\|^2 = 0$ ou $\mathbf{u} = \mathbf{0}$, de modo que $T(\mathbf{v}) = \mathbf{v}$ e $(\ker T)^{\perp} \subseteq \operatorname{Im} T$. Por outro lado, se $\mathbf{v} \in \operatorname{Im} T$, então $\mathbf{v} = \mathbf{u} + \mathbf{w}$, onde $\mathbf{u} \in \ker T$ e $\mathbf{w} \in (\ker T)^{\perp}$, implica que

$\mathbf{v} = T(\mathbf{v}) = T(\mathbf{w}) = \mathbf{w}$, pois $(\ker T)^{\perp} \subseteq \operatorname{Im} T$. Logo, $\operatorname{Im} T \subseteq (\ker T)^{\perp}$. Portanto, T é uma projeção ortogonal. 11. Se $T^* = T$, então $\langle T(\mathbf{v}), \mathbf{v}\rangle = \langle \mathbf{v}, T(\mathbf{v})\rangle = \overline{\langle T(\mathbf{v}), \mathbf{v}\rangle}$. Reciprocamente, se $\langle T(\mathbf{v}), \mathbf{v}\rangle$ for real, então $\langle T(\mathbf{v}), \mathbf{v}\rangle = \langle T^*(\mathbf{v}), \mathbf{v}\rangle$, para todo $\mathbf{v} \in V$. Assim,

$$\langle (T - T^*)(\mathbf{v}), \mathbf{v}\rangle = \langle T(\mathbf{v}), \mathbf{v}\rangle - \langle T^*(\mathbf{v}), \mathbf{v}\rangle = 0,$$

para todo $\mathbf{v} \in V$. Portanto, $T^* = T$. Esta implicação é falsa, considere $T \in \mathcal{L}(\mathbb{R}^2)$ definido como $T(x, y) = (y, -x)$. 12. Basta desenvolver $0 = \langle T(\mathbf{u}+\mathbf{v}), \mathbf{u}+\mathbf{v}\rangle$, para todos $\mathbf{u}, \mathbf{v} \in V$. 13. (a) Seja $\lambda \in F$ um autovalor de T. Então existe um $\mathbf{v} \in V$, com $\mathbf{v} \neq \mathbf{0}$, tal que $T(\mathbf{v}) = \lambda\mathbf{v}$. Assim,

$$\lambda\langle \mathbf{v}, \mathbf{v}\rangle = \langle \lambda\mathbf{v}, \mathbf{v}\rangle = \langle T(\mathbf{v}), \mathbf{v}\rangle = \langle \mathbf{v}, T(\mathbf{v})\rangle = \langle \mathbf{v}, \lambda\mathbf{v}\rangle = \bar{\lambda}\langle \mathbf{v}, \mathbf{v}\rangle$$

implica que $\lambda = \bar{\lambda}$, de modo que λ é real. (b) Seja $\mathbf{A}$ a matriz de T em relação à alguma base de V. Então $\mathbf{A}$ pode ser vista como um elemento de $\mathbb{C}^{n\times n}$. Assim, pelo Teorema 5.16, $\mathbf{A}$ possui um autovalor real λ. Logo, existe um $\mathbf{Z} \in \mathbb{C}^{n\times 1}$, com $\mathbf{Z} \neq \mathbf{O}$, tal que $\mathbf{AZ} = \lambda\mathbf{Z}$. Como $\mathbf{Z} = \mathbf{X} + i\mathbf{Y}$ temos que $\mathbf{AX} + i\mathbf{AY} = \lambda\mathbf{X} + i\lambda\mathbf{Y}$, de modo que $\mathbf{AX} = \lambda\mathbf{X}$ ou $\mathbf{AY} = \lambda\mathbf{Y}$. Portanto, $\mathbf{X}$ ou $\mathbf{Y}$ é um autovetor real associado a λ. (c) Cálculo direto. (d) Se $\mathbf{w} \in V_\lambda^{\perp}$, então $\langle \mathbf{v}, \mathbf{w}\rangle = 0$, para todo $\mathbf{v} \in V_\lambda$. Como $T(\mathbf{v}) = \lambda\mathbf{v}$ temos que $\langle \mathbf{v}, T(\mathbf{w})\rangle = \langle T(\mathbf{v}), \mathbf{w}\rangle = \lambda\langle \mathbf{v}, \mathbf{w}\rangle = 0$, para todo $\mathbf{v} \in V_\lambda$. Portanto, $T(\mathbf{w}) \in V_\lambda^{\perp}$ e $V_\lambda^{\perp}$ é invariante sob T. (e) Cálculo direto. (f) É claro que $\ker T \subseteq \ker T^k$. Por outro lado, se $\mathbf{v} \in \ker T^k$ e $k > 1$, então $T^k(\mathbf{v}) = \mathbf{0}$, de modo que $0 = \langle T^k(\mathbf{v}), T^{k-2}(\mathbf{v})\rangle = \langle T^{k-1}(\mathbf{v}), T^{k-1}(\mathbf{v})\rangle$, Assim, $T^{k-1}(\mathbf{v}) = \mathbf{0}$. Portanto, recursivamente, $T(\mathbf{v}) = \mathbf{0}$ e $\mathbf{v} \in \ker T$. (g) Sejam $\lambda \in F$ um autovalor de T e $\mathbf{v} \in V^\lambda$. Então $(T - \lambda I)^k(\mathbf{v}) = \mathbf{0}$, para algum $k \in \mathbb{N}$. Assim, pelo item (e), $(T - \lambda I)(\mathbf{v}) = \mathbf{0}$, pois $T - \lambda I$ é autoadjunto. Portanto, $V_\lambda = V^\lambda$, ou seja, $m_g(\lambda) = m_a(\lambda) = 1$ e o polinômio minimal de T é um produto de fatores lineares distintos. 14. (b) e (c) Basta notar que

$$(T_{\mathbf{u},\mathbf{v}} \circ T_{\mathbf{v},\mathbf{w}})(\mathbf{x}) = \langle \mathbf{x}, \mathbf{w}\rangle T_{\mathbf{u},\mathbf{v}}(\mathbf{v}) = \langle \mathbf{x}, \mathbf{w}\rangle \|\mathbf{v}\|^2\mathbf{u} = \|\mathbf{v}\|^2 T_{\mathbf{u},\mathbf{w}}(\mathbf{x}).$$

e

$$\langle \mathbf{x}, T^*_{\mathbf{u},\mathbf{v}}(\mathbf{y})\rangle = \langle T_{\mathbf{u},\mathbf{v}}(\mathbf{x}), \mathbf{y}\rangle = \langle \langle \mathbf{x}, \mathbf{v}\rangle\mathbf{u}, \mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{v}\rangle\langle \mathbf{u}, \mathbf{y}\rangle = \langle \mathbf{x}, \langle \mathbf{y}, \mathbf{u}\rangle\mathbf{v}\rangle.$$

(d) Suponhamos que $S = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\mathbf{u}_i \otimes \mathbf{u}_j = O$. Então, depois de alguns cálculos,

$$0 = \langle \mathbf{u}_p, S(\mathbf{u}_q)\rangle = \langle \mathbf{u}_p, \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\langle \mathbf{u}_q, \mathbf{u}_j\rangle \mathbf{u}_i\rangle = \bar{a}_{pq}.$$

Assim, β é LI e $\dim \mathcal{L}(V) = n^2$ implica que β é uma base. (e) Dados $\mathbf{u}, \mathbf{v} \in V$, com $\mathbf{u} \neq \mathbf{0}$ e $\mathbf{v} \neq \mathbf{0}$, $T_{\mathbf{u},\mathbf{v}}$ é autoadjunto se, e somente se, $T_{\mathbf{u},\mathbf{v}} = T_{\mathbf{v},\mathbf{u}}$ se, e somente se, $\langle \mathbf{x}, \mathbf{v}\rangle \mathbf{u} = \langle \mathbf{x}, \mathbf{u}\rangle \mathbf{v}$, para todo $\mathbf{x} \in V$. Em particular, pondo $\mathbf{x} = \mathbf{v}$, obtemos $\mathbf{u} = \lambda \mathbf{v}$, com $\lambda = \|\mathbf{v}\|^{-2}\langle \mathbf{v}, \mathbf{u}\rangle$. (f) Já sabemos que $a_{ij} = \langle T_{\mathbf{u},\mathbf{v}}(\mathbf{e}_j), \mathbf{e}_i\rangle$. Como $T_{\mathbf{u},\mathbf{v}}(\mathbf{e}_j) = \langle \mathbf{e}_j, \mathbf{v}\rangle \mathbf{u} = \bar{z}_j \mathbf{u}$ temos que $a_{ij} = y_i \bar{z}_j$. Note que $\mathbf{C}_j = \mathbf{u}^t \bar{z}_j$ implica que $\rho(\mathbf{A}) = \rho([\mathbf{u}^t]) = 1$. 15. (a) A função $f : V \to \mathbb{R}$ definida como $f(\mathbf{x}) = \det(\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}, \mathbf{x})$ é, pelo Teorema 1.24, um funcional linear. Assim, pelo Teorema 8.1, existe um único $\mathbf{w} \in V$ tal que $f(\mathbf{x}) = \langle \mathbf{x}, \mathbf{w}\rangle$, para todo $\mathbf{x} \in V$. (b) $\langle \mathbf{v}_i, \mathbf{w}\rangle = 0$, pois o determinante possui duas colunas iguais. (c) $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ são LD se, e somente se, $\langle \mathbf{x}, \mathbf{w}\rangle = 0$, para todo $\mathbf{x} \in V$ se, e somente se, $\mathbf{w} = \mathbf{0}$. (d) Note que $\mathbf{w} = \sum_{i=1}^{n} x_i \mathbf{e}_i$ implica que $\langle \mathbf{e}_i, \mathbf{w}\rangle = x_i$. Por outro lado, pela fórmula de Laplace em relação n-ésima coluna, obtemos

$$x_i = \det(\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}, \mathbf{e}_i) = (-1)^{i+n}\det(\mathbf{A}_i),$$

em que $\mathbf{A}_i$ é a submatriz de $\mathbf{A}$ obtida eliminando-se a i-ésima linha. Portanto, $\mathbf{w} = \sum_{i=1}^{n}(-1)^{n+i}\Delta_i \cdot \mathbf{e}_i$. (e) Seja $\mathbf{B}$ a matriz cujas colunas são os vetores $\mathbf{f}_j$, de modo que $\mathbf{B}\mathbf{B}^t = \mathbf{I}$. Portanto, $\det(\mathbf{f}_1, \ldots, \mathbf{f}_n) \in \{-1, 1\}$. (f) Segue dos itens (b) e (e). 16. (a) Como $\mathbf{x} - E(\mathbf{x}) \in H^{\perp} = \mathbb{R}[\mathbf{w}]$, para todo $\mathbf{x} \in V$, temos que $\mathbf{x} - E(\mathbf{x}) = a\mathbf{w}$, para algum $a \in \mathbb{R}$. Assim, $\langle E(\mathbf{x}), \mathbf{w}\rangle = 0$ implica que $a = \frac{\langle \mathbf{w}, \mathbf{x}\rangle}{\|\mathbf{w}\|^2}$. (b) Como $\langle \mathbf{v}_i, \mathbf{w}\rangle = 0, i = 1, \ldots, n-1$, temos que $H \subseteq \{\mathbf{x} \in V : \langle \mathbf{x}, \mathbf{w}\rangle = 0\}$. A reciproca é análogo.

Seção 8.2

1. Primeiro observe que a matriz deve possuir entradas complexas, por exemplo,

$$\mathbf{A} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \text{ e } \mathbf{B} = \frac{1}{\sqrt{3}}\begin{pmatrix} i & -2 \\ 2 & i \end{pmatrix}$$

2. (a) Defina $\phi : V \to \mathbb{R}^2$ como $\phi(a + bi) = (a, b)$. Então

$$\langle \phi(a+bi), \phi(c+di)\rangle = ac + bd = \operatorname{Re}(a+bi)(\overline{c+di}) = \langle a+bi, c+di\rangle.$$

(b) Como

$$\langle x, T_z^*(y)\rangle = \langle T_z(x), y\rangle = \operatorname{Re}(T_z(x)\bar{y}) = \operatorname{Re}(zx\bar{y}) = \operatorname{Re}(x\overline{\bar{z}y}) = \langle x, \bar{z}y\rangle$$

temos que $T_z^* = T_{\bar{z}}$. (c) Observe que $T_z = T_z^*$ se, e somente se, $z = \bar{z}$. (d) Note que $T_z T_z^* = I$ se, e somente se, $|z| = 1$. (e) T_z é positivo se $\langle T_z(x), x\rangle > 0$, para todo $x \in V$, com $x \neq 0$, se, e somente se, $\operatorname{Re}(z) > 0$. (f) Note que $z^2 - 2\operatorname{Re}(z)z + |z|^2 = 0$, para todo $z \in V$, de moto que $T_z^2 - 2\operatorname{Re}(z)T_z + |z|^2 I = 0$. Portanto, $\det T_z = |z|^2$ e $\operatorname{tr}(T_z) = 2\operatorname{Re}(z)$. (g) Se $z = a + bi$, então

$$[T_z]_\alpha^\alpha = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

(h) T deve ser linear sobre $\mathbb{C}$, ou seja, se, e somente se, $T(i) = iT(1)$, confira Exercício (8) da Seção 5.2. (i) $T(x) = \bar{x}$. 3. Seja $T \in \mathcal{L}(\mathbb{R}^{3\times 1})$ definida como $T(\mathbf{X}) = \mathbf{QX}$. Então é fácil verificar que $\mathbf{QQ}^t = \mathbf{I}, \mathbf{Q}^t = \mathbf{Q}, \mathbf{Q}^2 = \mathbf{I}$ e $\det \mathbf{Q} = -1$. Assim, pelo Exemplo 8.18, T é uma reflexão sobre um plano, $\ker(\mathbf{I} - \mathbf{A}) = \mathbb{R}[(1,1,0),(1,0,1)]$, seguida de uma rotação em torno de uma reta normal ao plano, $\ker(\mathbf{I} + \mathbf{A}) = \mathbb{R}[(-1,1,1)]$. 4. Seja $T \in \mathcal{L}(\mathbb{R}^{3\times 1})$ definida como $T(\mathbf{X}) = \mathbf{AX}$. Então é fácil verificar que $\mathbf{A}^t\mathbf{A} = \mathbf{I}, \mathbf{A}^t \neq \mathbf{A}$ e $\det \mathbf{A} = 1$. Assim, pelo Exemplo 8.18, T é uma rotação em torno de uma reta, $\ker(\mathbf{I} - \mathbf{A}) = \mathbb{R}[\frac{1}{\sqrt{2}}(1,0,1)]$. Seja θ o ângulo de rotação. Então $1 + 2\cos\theta = \operatorname{tr}(\mathbf{A}) = \frac{5}{3}$ implica que $\cos\theta = \frac{1}{3}$. Por outro lado, para encontrar o sinal de $\operatorname{sen}\theta$ devemos determinar um vetor ortogonal a $\mathbf{X}_1 = \frac{1}{\sqrt{2}}(1,0,1)^t$, digamos $\mathbf{E}_2$. Logo, $\operatorname{sen}\theta = \det \mathbf{B} = -\frac{4}{3\sqrt{2}}$. Portanto, $\theta = -\arccos(\frac{1}{3})$. 5. Se T for uma isometria e $T^2 = I$, então $E^2 = E$ e $E^* = E$. A recíproca é análoga. 6. (a) Cálculo direto. (b) Como

$$\langle \mathbf{x}, T^t(\mathbf{y})\rangle = \langle T(\mathbf{x}), \mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle + k\langle \mathbf{x}, \mathbf{v}\rangle\langle \mathbf{v}, \mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y} + k\langle \mathbf{y}, \mathbf{v}\rangle\mathbf{v}\rangle$$

temos que $T^t = T$, para todo $k \in \mathbb{R}$. (c) Observe que $\|\mathbf{x}\|^2 = \|T(x)\|^2$ se, e somente se, $0 = 2k|\langle \mathbf{x}, \mathbf{v}\rangle|^2 + k^2|\langle \mathbf{x}, \mathbf{v}\rangle|^2$ se, e somente se, $k = -2$ ou $k = 0$. 7. Seja $\mathbf{v} \in \ker T$. Então $\|\mathbf{v}\| = \|T(\mathbf{v})\| = 0$ implica que $\mathbf{v} = \mathbf{0}$, de modo que T é injetora e T^{-1} existe. Dados $\mathbf{u}, \mathbf{v} \in V$, existem únicos $\mathbf{z}, \mathbf{w} \in V$ tais que $\mathbf{u} = T^{-1}(\mathbf{z})$ e $\mathbf{v} = T^{-1}(\mathbf{w})$. Assim, $\langle \mathbf{z}, \mathbf{w}\rangle = \langle T^{-1}(\mathbf{z}), T^{-1}(\mathbf{w})\rangle$, para todos $\mathbf{z}, \mathbf{w} \in V$, de modo que $\langle T(\mathbf{u}), \mathbf{v}\rangle = \langle T^{-1}(T(\mathbf{u})), T^{-1}(\mathbf{v})\rangle = \langle \mathbf{u}, T^{-1}(\mathbf{v})\rangle$, para todos $\mathbf{u}, \mathbf{v} \in V$. Finalmente, dados $\mathbf{u}, \mathbf{v} \in V$ e $a \in F$, e depois de alguns cálculos, $\langle T(\mathbf{u} + a\mathbf{v}), \mathbf{w}\rangle = \langle \mathbf{u} + a\mathbf{v}, T^{-1}(\mathbf{w})\rangle = \langle T(\mathbf{u}) + aT(\mathbf{v}), \mathbf{w}\rangle$, para todo $\mathbf{w} \in V$. Portanto, T é linear. 8. (a) Tome $J \in \mathcal{L}(\mathbb{C}^2)$ definido como $J(x, y) = (-x, y)$. 9. Pelo Teorema de Schur, $\mathbf{P}^*\mathbf{AP} = \mathbf{U}$, de modo que $\mathbf{U}^* = \mathbf{P}^*\mathbf{A}^*\mathbf{P} = \mathbf{P}^*\mathbf{AP} = \mathbf{U}$. Assim, $u_{ii} = \bar{u}_{ii}$ e $u_{ij} = 0$, se $i \neq j$, pois $\mathbf{U}$ é triangular superior e $\mathbf{U}^*$ é triangular inferior. 10. Note que $S = T|_W$ é injetora, de modo que S é sobrejetora. Portanto, $W = S(W) = T(W)$. Assim, dado $\mathbf{w} \in$

W, existe um único $\mathbf{u} \in W$ tal que $\mathbf{w} = T(\mathbf{u})$. Logo, dado $\mathbf{v} \in W^{\perp}$, obtemos $\langle T(\mathbf{v}), \mathbf{w}\rangle = \langle T(\mathbf{v}), T(\mathbf{u})\rangle = \langle \mathbf{v}, \mathbf{u}\rangle = 0$. Portanto, $T(W^{\perp}) \subseteq W^{\perp}$. 11. Observe que $\|T(\mathbf{u})\| = \|\mathbf{u}\|$, para todo $\mathbf{u} \in V$, de modo que

$$2\langle T(\mathbf{u}), T(\mathbf{v})\rangle = \|T(\mathbf{u})\|^2 + \|T(\mathbf{v})\|^2 - \|T(\mathbf{u}) - T(\mathbf{v})\|^2 = 2\langle \mathbf{u}, \mathbf{v}\rangle.$$

Assim, dados $\mathbf{u}, \mathbf{v} \in V$ e $a \in F$,

$$\begin{aligned}
\|T(\mathbf{u}+a\mathbf{v}) - (T(\mathbf{u}) + aT(\mathbf{v}))\|^2 &= \|T(\mathbf{u}+a\mathbf{v})\|^2 + \|T(\mathbf{u})\|^2 + a^2\|T(\mathbf{v})\|^2 \\
&- 2\langle T(\mathbf{u}+a\mathbf{v}), T(\mathbf{u})\rangle \\
&- 2a\langle T(\mathbf{u}+a\mathbf{v}), T(\mathbf{v})\rangle + 2a\langle T(\mathbf{u}), T(\mathbf{v})\rangle \\
&= \|\mathbf{u}+a\mathbf{v}\|^2 + \|\mathbf{u}\|^2 + a^2\|\mathbf{v}\|^2 \\
&- 2\langle \mathbf{u}+a\mathbf{v}, \mathbf{u}\rangle - 2a\langle \mathbf{u}+a\mathbf{v}, \mathbf{v}\rangle \\
&+ 2a\langle \mathbf{u}, \mathbf{v}\rangle = \|\mathbf{u}+a\mathbf{v} - (\mathbf{u}+a\mathbf{v})\|^2 = 0.
\end{aligned}$$

Portanto, $T(\mathbf{u}+a\mathbf{v}) = T(\mathbf{u}) + aT(\mathbf{v})$, ou seja, T é linear. 12. Aplique o Exercício (11) a $S(\mathbf{x}) = T(\mathbf{x}) - T(\mathbf{0})$. 13. Lembre-se que $\langle \mathbf{u}+\mathbf{v}, \mathbf{u}-\mathbf{v}\rangle = \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2$. Assim, as implicações $(a \Rightarrow b \Rightarrow c \Rightarrow d)$ são diretas. $(d \Rightarrow a)$ Note, para todos $\mathbf{u}, \mathbf{v} \in V - \{\mathbf{0}\}$, que $\|\frac{1}{\|\mathbf{u}\|}\mathbf{u}\| = \|\frac{1}{\|\mathbf{v}\|}\mathbf{v}\|$ implica que $\|T(\mathbf{u})\| = k\|\mathbf{u}\|$, em que $k = \frac{\|T(\mathbf{v})\|}{\|\mathbf{v}\|}$. Portanto, $T = k(k^{-1}T) = kU$, com U ortogonal. 14. (a) Dado $\mathbf{u} \in V$,

$$\|\mathbf{u} + iT(\mathbf{u})\|^2 = \|\mathbf{u}\|^2 + \|T(\mathbf{u})\|^2 + \langle iT(\mathbf{u}), \mathbf{u}\rangle + \langle \mathbf{u}, iT(\mathbf{u})\rangle = \|\mathbf{u}\|^2 + \|T(\mathbf{u})\|^2.$$

(b) Direto de (a). (c) Dado $\mathbf{w} \in V$, existe um único $\mathbf{u} \in V$ tal que $\mathbf{w} = (I + iT)(\mathbf{u})$ ou $\mathbf{u} = (I + iT)^{-1}(\mathbf{w})$. Assim,

$$\|U(\mathbf{u})\| = \|(I - iT)(I + iT)^{-1}(\mathbf{u})\| = \|(I - iT)(\mathbf{w})\| = \|(I + iT)(\mathbf{w})\| = \|\mathbf{u}\|.$$

Portanto, U é unitário. Se $U(\mathbf{u}) = \mathbf{u}$, então

$$(I - iT)(I + iT)^{-1}(\mathbf{u}) = \mathbf{u} = (I + iT)(I + iT)^{-1}(\mathbf{u}),$$

de modo que $2i(I + iT)^{-1}(\mathbf{u}) = \mathbf{0}$. Assim, $\mathbf{u} = \mathbf{0}$. Portanto, 1 não é autovalor de U. (d) Se 1 não for autovalor de U, então $I - U$ é não singular e $U^* = U^{-1}$ implicam que

$$\begin{aligned}
T^* &= -i(I + U^*)(I - U^*)^{-1} = -i(I + U^{-1})UU^{-1}(I - U^{-1})^{-1} \\
&= -i(I + U)U^{-1}(I - U^{-1})^{-1} = -i(I + U)((I - U^{-1})U)^{-1} \\
&= i(I + U)(I - U)^{-1} = T.
\end{aligned}$$

Portanto, T é autoadjunto.

Seção 8.3

2. Vamos usar indução sobre n. Se $n = 1$, nada há para ser provado. Suponhamos que o resultado seja válido para todo k, com $1 \leq k < n$ e $n > 1$. Observe que

$$\mathbf{A} = \begin{pmatrix} \mathbf{B} & \mathbf{b} \\ \mathbf{b}^* & a_{nn} \end{pmatrix}$$

implica que existe uma matriz $\mathbf{J}_1 = \operatorname{diag}(\pm 1, \ldots, \pm 1)$ tal que $\det(\mathbf{B}+\mathbf{J}_1) \neq 0$. Assim, pelo Exercício (6) da Seção 1.3,

$$\det(\mathbf{A}+\mathbf{J}) = \det \begin{pmatrix} \mathbf{B}+\mathbf{J}_1 & \mathbf{b} \\ \mathbf{b}^* & a_{nn} \pm 1 \end{pmatrix} = k \det(\mathbf{B}+\mathbf{J}_1),$$

onde $k = a_{nn} \pm 1 - \mathbf{b}^*(\mathbf{B}+\mathbf{J}_1)^{-1}\mathbf{b} \in F$. É fácil verificar que

$$\det \begin{pmatrix} \mathbf{B}+\mathbf{J}_1 & \mathbf{b} \\ \mathbf{b}^* & a_{nn} + 1 \end{pmatrix} - \det \begin{pmatrix} \mathbf{B}+\mathbf{J}_1 & \mathbf{b} \\ \mathbf{b}^* & a_{nn} - 1 \end{pmatrix} = 2\det(\mathbf{B}+\mathbf{J}_1) \neq 0.$$

Logo, pelo menos um dos determinantes, é não nulo. Portanto, $\det(\mathbf{A}+\mathbf{J}) \neq 0$. 3. Seja $f(x) = \det(\mathbf{P} + ix\mathbf{Q}) \in \mathbb{C}[x]$. Então $f(1) \neq 0$. Assim, pelo Exercício (11) da Seção 1.3, existe um $x_0 = -ia^{-1}b \in \mathbb{C}$ (puramente imaginário) tal que $f(x_0) \neq 0$. Portanto, existem $a, b \in \mathbb{R}$ tais que $a\mathbf{P}+b\mathbf{Q}$ é não singular. 4. É claro que $\mathbf{B} = \mathbf{A}+\mathbf{A}^t$ é simétrica. Então existe um $\mathbf{X} \in V$ tal que $\mathbf{B}\mathbf{X} = \lambda\mathbf{X}$, onde $\lambda \in \mathbb{R}$. Consideremos $W = \mathbb{R}[\mathbf{X}, \mathbf{A}\mathbf{X}]$. Se $\mathbf{X}$ e $\mathbf{A}\mathbf{X}$ são LD, então existe um $c \in \mathbb{R}$ tal que $\mathbf{A}\mathbf{X} = c\mathbf{X}$. Assim, $W = \mathbb{R}[\mathbf{X}]$ é invariante sob $\mathbf{A}$, com $\dim W = 1$. Se $\mathbf{X}$ e $\mathbf{A}\mathbf{X}$ são LI, então $\lambda\mathbf{X} = \mathbf{A}\mathbf{X}+\mathbf{A}^t\mathbf{X} = \mathbf{A}\mathbf{X}+\mathbf{A}^{-1}\mathbf{X}$ implica que $\mathbf{A}^2\mathbf{X} = \lambda\mathbf{A}\mathbf{X}-\mathbf{X} \in W$. Assim, dado $\mathbf{Y} \in W$, existem $a, b \in \mathbb{R}$ tais que $\mathbf{Y} = a\mathbf{X} + b\mathbf{A}\mathbf{X}$. Logo, $\mathbf{A}\mathbf{Y} = a\mathbf{A}\mathbf{X} + b\mathbf{A}^2\mathbf{X} = -b\mathbf{X} + (a+b\lambda)\mathbf{A}\mathbf{X} \in W$. Portanto, W é invariante sob $\mathbf{A}$, com $\dim W = 2$. 5. Seja $T = T_{\mathbf{A}} \in \mathcal{L}(V)$. Vamos usar indução sobre n. Se $n = 1$, nada há para ser provado. Suponhamos que o resultado seja válido para todo k, com $1 \leq k < n$ e $n > 1$. Então, pelo Exercício (4), V possui um subespaço invariante sob T, com $\dim W = 1$ e/ou 2. Assim, $\dim W^\perp < n$ e $S = T|_{W^\perp}$ é ortogonal. Logo, $W^\perp$ pode ser escrito como uma soma direta de subespaço invariante sob T, com dimensão 1 e/ou 2. Portanto, segue de $V = W \oplus W^\perp$ e do Exemplo 8.14. 6. (a) Note que

$$T = \frac{1}{2}(T+T^*) + i\frac{1}{2i}(T-T^*) = R + iS, \text{ com } R^* = R \text{ e } S^* = S.$$

Por outro lado, se $T = T_1 + iT_2$, então $T^* = T_1 - iT_2$, de modo que $T_1 = R$ e $T_2 = S$. (b) Consequência direta de (a). 7. Seja $S = TT^*$. Então $S^* = S$. Assim, pelo item (f) do Exercício (13) da Seção 8.1, $\ker S^m = \ker S$. Logo, se $\mathbf{v} \in \ker T^m$, então $S(\mathbf{v}) = \mathbf{0}$, pois $TT^* = T^*T$. Portanto, $0 = \langle S(\mathbf{v}), \mathbf{v}\rangle = \langle T(\mathbf{v}), T(\mathbf{v})\rangle = \|T(\mathbf{v})\|^2$, de modo que $T(\mathbf{v}) = \mathbf{0}$ ou $\mathbf{v} \in \ker T$. 8. Segue do item (3) do Teorema 8.22 e/ou pelo Teorema de Schur. 9. (a) Basta notar que $TT^* = T^*T$ implica que $(TT^*)^m = T^m(T^*)^m$, para todo $m \in \mathbb{N}$. (b) Como $p(T)^* = \bar{p}(T^*)$ temos, pelo Teorema 8.19, que $\|p(T)(\mathbf{u})\| = \|\bar{p}(T^*)(\mathbf{u})\|$. Portanto, $p(T)(\mathbf{u}) = \mathbf{0}$ se, e somente se, $\bar{p}(T^*)(\mathbf{u}) = \mathbf{0}$. 10. Suponhamos que T seja normal. Então existe uma base ortonormal $\alpha = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ de V tal que $T(\mathbf{v}_i) = \lambda_i\mathbf{v}_i$ e $T(\mathbf{v}_i) = \bar{\lambda}_i\mathbf{v}_i$. Defina $U \in \mathcal{L}(V)$ como $U(\mathbf{v}_i) = \frac{\bar{\lambda}_i}{\lambda_i}\mathbf{v}_i$, se $\lambda_i \neq 0$, e $U(\mathbf{v}_i) = \mathbf{v}_i$, se $\lambda_i = 0$. Então $U^*U = I$ e $T^* = TU$. Reciprocamente, se $T^* = TU$, então $T = U^*T^*$, de modo que $UT = T^*$. Assim, $UT = TU$ e $T^*T = TUT = TTU = TT^*$. Portanto, T é normal. 11. (a) Note que $\mathbf{X}_i^t\mathbf{X} = x_i$, $\mathbf{AX} = \sum_{i=1}^n \lambda_i x_i \mathbf{X}_i$ e $\mathbf{X}^t\mathbf{AX} = \sum_{i=1}^n \lambda_i x_i^2$. Portanto,

$$F(\mathbf{X}) = \frac{\mathbf{X}^t\mathbf{AX}}{\mathbf{X}^t\mathbf{X}} = \frac{x_1^2\lambda_1 + \cdots + x_n^2\lambda_n}{\|\mathbf{X}\|^2}.$$

(b) Basta observar, pelo item (a), que

$$\lambda_1 = \frac{\lambda_1(x_1^2 + \cdots + x_n^2)}{\|\mathbf{X}\|^2} \leq F(\mathbf{X}) \leq \frac{\lambda_n(x_1^2 + \cdots + x_n^2)}{\|\mathbf{X}\|^2} = \lambda_n.$$

Portanto, $\lambda_1 = \min\{\mathbf{X}^t\mathbf{AX} : \|\mathbf{X}\|^2 = 1\}$. 12. Seja $\mathbf{A} = [T]_\alpha^\alpha$, para alguma base α de V. Então, pelo Teorema de Schur, $\mathbf{Q}^*\mathbf{AQ} = \mathbf{U}$ ou $\mathbf{A} = \mathbf{QUQ}^*$, em que $\mathbf{Q}$ é uma matriz unitária e $\mathbf{U}$ é uma matriz triangular superior. Assim,

$$\operatorname{tr}(\mathbf{A}^*\mathbf{A}) = \operatorname{tr}(\mathbf{QU}^*\mathbf{UQ}^*) = \operatorname{tr}(\mathbf{U}^*\mathbf{U}) \geq \sum_{i=1}^n |\lambda_i|^2.$$

Lembre-se que $\mathbf{A}$ é normal se, e somente se, $\mathbf{U}$ for diagonal. 13. Já sabemos que

$$\mathbf{C} = \left(\begin{array}{c|cc} \mathbf{O} & \multicolumn{2}{c}{\mathbf{I}_2} \\ \hline 1 & b+3 & b \end{array}\right)$$

é a matriz companheira de $f(x)$, o qual é seu polinômio característico. Neste caso,

$\mathrm{tr}(\mathbf{C}) = b, \mathrm{tr}(\mathrm{adj}(\mathbf{C})) = -(b+3)$ e $\det(\mathbf{C}) = 1$. Observe que

$$\mathbf{P}^{-1}\mathbf{CP} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & b+1 & 1 \\ 0 & 1 & -1 \end{pmatrix} = \mathbf{B}, \text{ com } \mathbf{P} = \begin{pmatrix} -1 & 1 & -b-1 \\ 1 & -1 & b+2 \\ -1 & 2 & -b-3 \end{pmatrix}$$

Como $\mathbf{B}^t = \mathbf{B}$ temos que todas as raízes de $f(x)$ são reais.

Seção 8.4

1. (a) Posto $k = 2$ e assinatura $(1,1)$, de modo que ela representa uma hipérbole. (b) Posto $k = 2$ e assinatura $(2,0)$, de modo que ela representa uma elipse. (c) Hipérbole. (d) Hipérbole. (d) Elipse. 2. (a) Posto $k = 3$ e assinatura $(3,0)$, de modo que ela representa um elipsoide. (b) Posto $k = 3$ e assinatura $(2,1)$, de modo que ela representa um hiperboloide de uma folha. (c) Posto $k = 3$ e assinatura $(1,2)$, de modo que ela representa um hiperboloide de duas folhas. (d) Hiperboloide de duas folhas. (e) Posto $k = 2$ e assinatura $(1,1)$, de modo que ela representa um cilindro hiperbólico. (f) Posto $k = 2$ e assinatura $(1,1)$, de modo que ela representa um cone $2x^2 - y^2 = 0$. 3. Como $(3,0)$ é a assinatura temos que ela representa um elipsoide e q definida positiva implica que $q(x,y,z) = 4(xy - yz) > 0$. 4. Observe que se $a \neq 0$, então $q(x,y) = ax^2 + 2bxy + cy^2 = a^{-1}((ax+by)^2 + Dy^2)$, com $D = ac - b^2$. Note que $q(1,0) = a$ e $q(b,-a) = aD$. Pondo $u = ax + by$ e $v = y$ (operador linear $(u,v) = T(x,y)$), obtemos $q(u,v) = a^{-1}u^2 + a^{-1}Dv^2$. Assim, se $D \neq 0$, então as principais propriedades de q são determinadas pelos sinais de a^{-1} e $a^{-1}D$. Por exemplo, $q(a,0) = a > 0$ e $q(0,-a) = aD > 0$ 5. Seja $q(\mathbf{X}) = \mathbf{X}^t\mathbf{AX}$. (a) $a_{ii} = q(\mathbf{E}_i) > 0$. (b) Segue do Teorema 8.29. (c) Considere $W = \mathbb{R}[\mathbf{E}_i, \mathbf{E}_j]$, com $i \neq j$. Então $p = q|_W$ é uma forma quadrática sobre W tal que

$$p(x_i\mathbf{E}_i + x_j\mathbf{E}_j) = a_{ii}x_i^2 + 2a_{ij}x_ix_j + a_{jj}x_j^2.$$

Assim, o resultado segue do Exercício (4). 6. $(a \Rightarrow b)$ Vamos usar indução sobre n. Se $n = 1$, então $\mathbf{A} = (a_{11})$. Assim, basta escolher $\mathbf{B} = (\sqrt{a_{11}})$. Suponhamos que o resultado seja válido para todo k, com $1 \leq k < n$ e $n > 1$. Observe que

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{n-1} & \mathbf{X} \\ \mathbf{X}^* & a_{nn} \end{pmatrix}$$

implica que existe uma matriz triangular superior $\mathbf{B}_{n-1} = (b_{ij})$, com $b_{ii} > 0$, tal que

$\mathbf{A}_{n-1} = \mathbf{B}_{n-1}^t \mathbf{B}_{n-1}$. Como

$$\mathbf{A} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{X}^t \mathbf{A}_{n-1}^{-1} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{A}_{n-1} & \mathbf{X} \\ \mathbf{0} & a_{nn} - \mathbf{X}^t \mathbf{A}_{n-1}^{-1} \mathbf{X} \end{pmatrix}$$

temos que $\det \mathbf{A} = \det \mathbf{A}_{n-1}(a_{nn} - \mathbf{X}^t \mathbf{A}_{n-1}^{-1} \mathbf{X}) > 0$ e $\det \mathbf{A}_{n-1} > 0$ implicam que $a_{nn} > \mathbf{X}^t \mathbf{A}_{n-1}^{-1} \mathbf{X} = \|(\mathbf{B}_{n-1}^{-1})^t \mathbf{X}\|^2$. Assim, existe um $b_{nn} \in \mathbb{R}$, de modo que $b_{nn}^2 = a_{nn} - \|(\mathbf{B}_{n-1}^{-1})^t \mathbf{X}\|^2$. Portanto,

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{n-1} & (\mathbf{B}_{n-1}^{-1})^t \mathbf{X} \\ \mathbf{0} & b_{nn} \end{pmatrix}$$

é tal que $\mathbf{A} = \mathbf{B}^t \mathbf{B}$. $(b \Rightarrow c)$ Para cada $\mathbf{X} \in \mathbb{R}^{n\times 1}$, com $\mathbf{X} \neq \mathbf{O}$, obtemos $\mathbf{X}^t \mathbf{A} \mathbf{X} = \|\mathbf{B}\mathbf{X}\|^2 > 0$. Portanto, $\mathbf{A}$ é definida positiva. $(c \Rightarrow a)$ Seja $\mathbf{X}_k = (x_1, \ldots, x_k)^t \in \mathbb{R}^{k\times 1}$, com $\mathbf{X}_k \neq \mathbf{O}$ e $k = 1, \ldots, n$. Então $\mathbf{X} = (x_1, \ldots, x_k, 0, \ldots, 0)^t \in \mathbb{R}^{n\times 1}$ e $\mathbf{X}_k^t \mathbf{A}_k \mathbf{X}_k = \mathbf{X}^t \mathbf{A} \mathbf{X} > 0$. Portanto, cada $\mathbf{A}_k$ é definida positiva, de modo que $\det(\mathbf{A}_k) > 0$. (d) $\mathbf{T}_{12}(-a^{-1}b)\mathbf{A}\mathbf{T}_{12}(-a^{-1}b) = \mathbf{D} = \operatorname{diag}(a, d - a^{-1}b^2)$ ou $\mathbf{A} = \mathbf{T}_{12}(a^{-1}b)\mathbf{D}\mathbf{T}_{12}(a^{-1}b)$. Ponha $\mathbf{B}^t = \mathbf{T}_{12}(a^{-1}b)\sqrt{\mathbf{D}}$. 7. Note que $h_{\mathbf{kA}} = h_{\mathbf{A}}$, para todo $k \in \mathbb{R}$, de modo que $h_{\mathbf{A}}^{-1} = h_{\mathbf{A}^{-1}}$ é a inversa de $h_{\mathbf{A}}$. (b) $\mathbf{A} \in \ker h$ se, e somente se, $h(\mathbf{A}) = I$ se, e somente se, $h_{\mathbf{A}}(z) = z$, para todo $z \in S^2$, ou seja, $cz^2 + (d-a)z - b = 0$, para todo $z \in S^2$. Assim, pelo Teorema 1.2, $c = b = 0$ e $a = d$. Portanto, $\ker h = \{k\mathbf{I}_2 : k \in \mathbb{R}^\times\}$. (c) Se $z = x + iy \in S^2$ e $w = h_{\mathbf{A}}(z)$, então, depois de alguns cálculos,

$$w - \bar{w} = \frac{ad - bc}{|cz + d|^2}(z - \bar{z}),$$

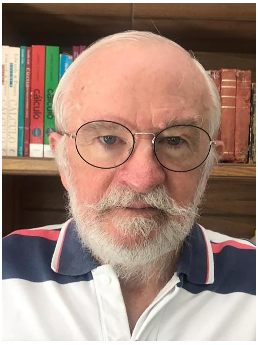
Logo, $\operatorname{Im} h_{\mathbf{A}}(z) \in \mathbb{H}$, para todo $\mathbf{A} \in G^+$ e $z \in \mathbb{H}$. (d) Como $h_1 = h|_{SL_2(\mathbb{R})}$ satisfaz o item (b) e $\ker h_1 = \{\pm\mathbf{I}_2\}$ temos, pelo item (c), que $\mathrm{PSL}_2(\mathbb{R})$ age sobre $\mathbb{H}$. (e) $z \in \mathbb{H}$ é um ponto fixo de $h_{\mathbf{A}}$ se, e somente se, $h_{\mathbf{A}}(z) = z$, ou seja, $cz^2 + (d-a)z - b = 0$. Se $c \neq 0$, então a equação possui duas raízes

$$\frac{a - b \pm \sqrt{\Delta}}{2c}, \text{ com } \Delta = (a-d)^2 + 4bc = \operatorname{tr}(\mathbf{A})^2 - 4.$$

Como $h_{\mathbf{A}}(\infty) \neq \infty$ temos que $h_{\mathbf{A}}$ possui no máximo dois pontos fixos. Se $c = 0$, então $ad = 1$. Assim, $z = (1 - a^2)^{-1}ab$ é o único ponto fixo, quando $a \neq \pm 1$, pois $h_{\mathbf{A}}(\infty) \neq \infty$. Enquanto, se $a = \pm 1$, então $h_{\mathbf{A}}(\infty) = \infty$. Portanto, $h_{\mathbf{A}}$ possui dois pontos fixos se $\operatorname{tr}(\mathbf{A}) \neq \pm 2$ e um único ponto fixo se $\operatorname{tr}(\mathbf{A}) = \pm 2$.

Índice Remissivo

Sobre os autores

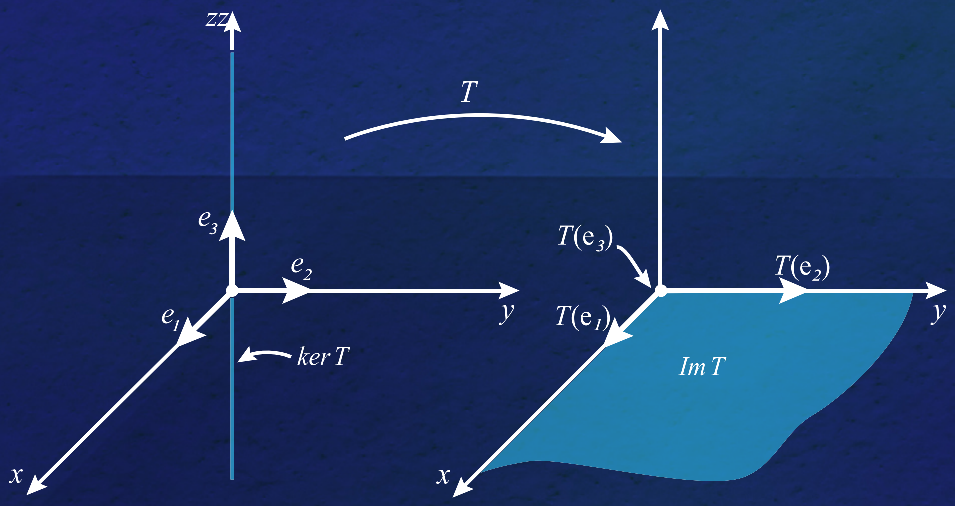


Antônio de Andrade e Silva
Nasceu em Coremas, alto sertão paraibano, em 26 de janeiro de 1954. Veio para João Pessoa em 1974, onde permanece até hoje. Ingressou em 1977 na Universidade Federal da Paraíba (UFPB), onde concluiu o bacharelado em matemática. Obteve o grau de mestre em matemática (1989) na Universidade Federal do Ceará (UFC) e de doutor em engenharia elétrica (1996) na Universidade Estadual de Campinas (UNICAMP). É professor titular aposentado da Universidade Federal da Paraíba.

João Bosco Batista Lacerda
Nasceu em Fortaleza, em 24 de outubro de 1956. Ainda criança seus pais retornaram a terra natal, a cidade de Cajazeiras no alto Sertão Paraibano, onde realizou os seus estudos primário e secundário. Foi para Campina Grande em 1975 onde concluiu os estudos secundários e ingressou na Universidade Regional do Nordeste, concluindo o curso de licenciatura em matemática. Obteve o grau de mestre em matemática (1983) na Universidade de Brasília (UnB) e de doutor em engenharia elétrica (1994) na Universidade Estadual de Campinas (UNICAMP). É professor associado IV da Universidade Federal da Paraíba.

www.PIMENTACULTURAL.com

UM CURSO DE ÁLGEBRA LINEAR



UNIFSM
CENTRO UNIVERSITÁRIO SANTA MARIA

pimenta cultural