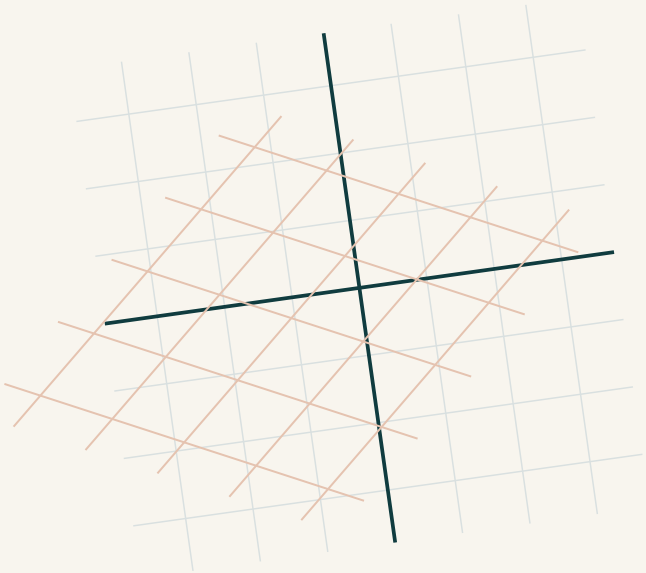
NOTAS DE AULA / MCTB002

ÁLGEBRA LINEAR AVANÇADA

Estruturas, operadores e formas canônicas



Daniel Miranda Universidade Federal do ABC / 22 de Setembro de 2026

PROJETO SOLEIRA

Projeto editorial desenvolvido em colaboração com Soleira – ChatGPT (OpenAI).

Sumário

Prefácio

Estas notas correspondem aos apontamentos para o curso de Álgebra Linear Avançada (MCTB002), ministrado originalmente durante a pandemia de 2020. O texto foi sendo ampliado, revisto e reorganizado desde então, preservando o caráter de notas de aula: uma exposição orientada por exemplos, demonstrações e exercícios.

O livro se destina a estudantes de graduação ou pós-graduação que já tenham cursado ao menos uma disciplina de álgebra linear. Supõe-se familiaridade com espaços vetoriais, matrizes, sistemas lineares e determinantes. Sugestões e correções são bem-vindas.

> *I have not yet any clear view as to the extent to which we are at liberty arbitrarily to create imaginaries and endow them with supernatural properties.*
>
> — *John Graves*

Agradecimentos

Agradecemos a Mariana Rodrigues da Silveira e aos alunos Deborah Fabri e Guilherme Gigeck pelas correções e sugestões.

1

Estruturas Algébricas e Polinômios

Estruturas algébricas são conjuntos munidos de operações que satisfazem certos axiomas. Anéis e corpos são exemplos dessas estruturas, e suas propriedades determinam quais regras de cálculo podemos empregar. Na Álgebra Linear, os elementos de um corpo servem de escalares, enquanto os polinômios de $\mathbb{K}[x]$ permitem estudar operadores lineares.

A teoria de grupos e os grupos de permutações serão tratados no capítulo de determinantes, em conexão com a paridade das permutações e a fórmula de Leibniz.

Nesse capítulo

1.1 Operações e estruturas algébricas

Para definir uma estrutura algébrica, especificamos as operações em um conjunto e as leis que elas devem satisfazer. Esse procedimento se aplica à adição de números, à composição de aplicações, à multiplicação de matrizes e às operações de um espaço vetorial.

Definição 1.1 — Operação binária. Seja A um conjunto não vazio. Uma **operação binária interna** em A é uma função

$$\mu : A \times A \longrightarrow A.$$

Se B é outro conjunto, uma função $B \times A \to A$ será chamada de **operação externa** em A. A multiplicação de vetores por escalares, $\mathbb{K} \times V \to V$, é o exemplo fundamental neste livro.

Em geral, escrevemos $a \cdot b$, $a + b$ ou simplesmente ab em lugar de $\mu(a, b)$. O par (A, μ) indica que o conjunto A está sendo considerado juntamente com a operação μ.

Consequências dos axiomas
Um argumento que depende apenas dos axiomas de uma estrutura vale em todos os exemplos que os satisfazem. Assim, uma mesma demonstração pode ser aplicada a números, matrizes, aplicações ou polinômios.

Exemplos 1.2.

1. A adição, a multiplicação e a subtração são operações binárias em $\mathbb{Z}$. A divisão não é: além de $1/2 \notin \mathbb{Z}$, a expressão $1/0$ nem sequer está definida.
2. A união e a interseção são operações binárias no conjunto das partes $\mathscr{P}(X)$.
3. Se X^X denota o conjunto das aplicações de X em X, a composição é uma operação binária em X^X.
4. A multiplicação usual não é uma operação binária no conjunto das matrizes retangulares $m \times n$ quando $m \neq n$, mas é uma operação em $M_n(\mathbb{K})$.

◁

Quando A é finito, uma operação binária pode ser apresentada por uma **tabela de Cayley**. Adotaremos a convenção de colocar $\mu(a, b)$ na linha de a e na coluna de b.

Exemplo 1.3. No conjunto $A = \{P, I\}$, cujos símbolos representam as paridades par e ímpar, a adição de paridades é descrita pela tabela

$+$	P	I
P	P	I
I	I	P

.

Ela resume, por exemplo, que a soma de dois inteiros ímpares é par.

◁

1.1.1 Leis algébricas

Definição 1.4 — Associatividade e comutatividade. Uma operação em A é **associativa** quando

$$(ab)c = a(bc)$$

para quaisquer $a, b, c \in A$, e é **comutativa** quando $ab = ba$ para quaisquer $a, b \in A$.

A associatividade permite escrever $a_1 a_2 \cdots a_n$ sem especificar a posição dos parênteses. Ela não permite, contudo, alterar a ordem dos fatores: essa possibilidade depende da comutatividade.

Exemplos 1.5.

1. A adição e a multiplicação de números são associativas e comutativas.
2. A composição de aplicações e a multiplicação de matrizes são associativas, mas em geral não são comutativas.
3. A subtração não é associativa, pois normalmente $(a - b) - c \neq a - (b - c)$.

◁

Definição 1.6 — Identidade e inverso. Seja A um conjunto com uma operação binária. Um elemento $e \in A$ é uma **identidade** quando $ea = ae = a$ para todo $a \in A$. Se a operação possui identidade, um elemento $b \in A$ é um **inverso** de a quando $ab = ba = e$.

Proposição 1.7 (Unicidade da identidade e dos inversos).
Uma operação possui no máximo uma identidade. Se, além disso, a operação é associativa, cada elemento possui no máximo um inverso.

Demonstração. Se e e e' são identidades, então $e = ee' = e'$. Agora suponha que b e c sejam inversos de a. Pela associatividade,

$$b = be = b(ac) = (ba)c = ec = c.$$

Logo, tanto a identidade quanto o inverso, quando existem, são únicos. ■

1.1.2 Estruturas e aplicações que as preservam

Definição 1.8 — Estrutura algébrica. Uma **estrutura algébrica** consiste em um conjunto, uma coleção de operações e uma lista de axiomas que essas operações devem satisfazer.

Escolhas diferentes de operações e axiomas dão origem a estruturas distintas. Estudaremos aqui aquelas que intervêm na descrição dos escalares e dos polinômios.

Definição 1.9 — Homomorfismo e isomorfismo. Um **homomorfismo** entre duas estruturas do mesmo tipo é uma aplicação que preserva suas operações. Quando a estrutura inclui constantes distinguidas, a definição deve especificar se elas também são preservadas. Por exemplo, se $f : A \to B$ preserva uma operação binária, então

$$f(x \cdot y) = f(x) \cdot f(y).$$

Um homomorfismo bijetivo é chamado de **isomorfismo**. Estruturas isomorfas possuem as mesmas propriedades algébricas, embora seus elementos possam ter naturezas diferentes.

Preservar a estrutura
Um isomorfismo estabelece uma correspondência entre os elementos de duas estruturas, preservando suas operações. As propriedades algébricas são, portanto, mantidas, ainda que os elementos tenham representações diferentes.

Definição 1.10 — Operação induzida. Seja $(A, \cdot)$ um conjunto com uma operação binária e seja $\phi : A \to B$ uma bijeção. A operação induzida em B é definida por

$$\phi(a_1) * \phi(a_2) = \phi(a_1 \cdot a_2).$$

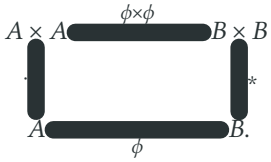
Equivalentemente,

$$b_1 * b_2 = \phi\big(\phi^{-1}(b_1) \cdot \phi^{-1}(b_2)\big).$$

Com essa operação, ϕ torna-se automaticamente um isomorfismo.

A definição anterior é frequentemente chamada de **transporte de estrutura**. Ela é

expressa pelo diagrama comutativo



Comutatividade significa que os dois caminhos de $A \times A$ até B produzem o mesmo resultado.

Exercícios

Ex. 1.1 — Seja $\circ$ uma operação associativa em A e fixe $x \in A$. Defina $a \circ_x b = a \circ (x \circ b)$. Mostre que $\circ_x$ é associativa.

Ex. 1.2 — Construa uma operação em um conjunto de três elementos que possua identidade, mas não seja associativa. Apresente-a por meio de uma tabela de Cayley.

Ex. 1.3 — Seja $(A, \cdot)$ um conjunto com uma operação associativa e com identidade. Mostre que $ab = e$ não implica necessariamente $ba = e$. Dê um exemplo usando composição de aplicações de um conjunto infinito em si mesmo.

Ex. 1.4 — Seja $\phi : A \to B$ uma bijeção e transporte para B uma operação associativa em A. Mostre que a operação induzida em B também é associativa. Verifique o mesmo para comutatividade, identidade e inversos.

1.2 Anéis

Os anéis reúnem duas operações, adição e multiplicação, sem exigir que todo elemento não nulo possua inverso multiplicativo. Essa é a estrutura natural dos números inteiros, das matrizes quadradas e, sobretudo, dos polinômios.

Anéis e polinômios em operadores
Em Álgebra Linear, os escalares pertencem a um corpo, mas os polinômios em um operador formam um anel. Essa diferença explica por que podemos somar e multiplicar polinômios, embora nem todo polinômio não nulo possa ser invertido.

Definição 1.11 — Anel. Um **anel com identidade** é um conjunto R, munido de uma adição e de uma multiplicação, com $0 \neq 1$, que satisfaz as seguintes propriedades:

1. a adição é associativa e comutativa;
2. existe $0 \in R$ tal que $a + 0 = 0 + a = a$ para todo $a \in R$;
3. para cada $a \in R$, existe $-a \in R$ tal que $a + (-a) = 0$;
4. a multiplicação é associativa;
5. $1a = a1 = a$ para todo $a \in R$;
6. a multiplicação é distributiva em relação à adição:

$$a(b + c) = ab + ac, \qquad (a + b)c = ac + bc.$$

O anel é **comutativo** quando $ab = ba$ para quaisquer $a, b \in R$.

Neste livro, a palavra anel sempre significará anel com identidade. Não exigiremos, contudo, que um homomorfismo entre anéis preserve a identidade, a menos que isso seja mencionado explicitamente.

Exemplos 1.12.

1. Os inteiros, com as operações usuais, formam o anel comutativo $\mathbb{Z}$.
2. Se $\mathbb{K}$ é um corpo, então $M_n(\mathbb{K})$ é um anel com a adição e a multiplicação de matrizes. Para $n \geq 2$, esse anel não é comutativo.
3. O conjunto $2\mathbb{Z}$ é fechado por adição e multiplicação, mas não é um subanel de $\mathbb{Z}$ segundo nossa convenção, pois não contém a identidade 1.

◁

Definição 1.13 — Divisor de zero e domínio de integridade. Um elemento não nulo $a \in R$ é um **divisor de zero** se existe $b \neq 0$ tal que $ab = 0$ ou $ba = 0$. Um **domínio de integridade** é um anel comutativo sem divisores de zero.

Assim, $\mathbb{Z}$ é um domínio de integridade. O anel $M_n(\mathbb{K})$, para $n \geq 2$, possui divisores de zero: por exemplo, em $M_2(\mathbb{K})$,

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

Definição 1.14 — Subanel e ideal. Seja R um anel.

1. Um subconjunto $S \subseteq R$ é um **subanel** quando contém 1 e é fechado por subtração e multiplicação.
2. Um subconjunto não vazio $I \subseteq R$ é um **ideal bilateral** quando é fechado por subtração e
$$ra \in I, \qquad ar \in I$$
para quaisquer $a \in I$ e $r \in R$.

Em anéis comutativos, as duas condições de absorção coincidem e diremos simplesmente **ideal**.

Exemplos 1.15.

1. Para cada inteiro m, o conjunto $m\mathbb{Z}$ é um ideal de $\mathbb{Z}$.
2. Os conjuntos $\{0\}$ e R são ideais de qualquer anel R.
3. Se $\mathbb{K}$ é um corpo, seus únicos ideais são $\{0\}$ e $\mathbb{K}$. De fato, se um ideal I contém $a \neq 0$, então contém $a^{-1}a = 1$ e, consequentemente, contém todo elemento de $\mathbb{K}$.

◁

Definição 1.16 — Ideal gerado. Se $A \subseteq R$, o **ideal bilateral gerado por** A, denotado por $\langle A \rangle$, é a interseção de todos os ideais de R que contêm A. Equivalentemente,

$$\langle A \rangle = \left\{ \sum_{i=1}^{m} r_i a_i s_i \;:\; m \geq 0,\ r_i, s_i \in R,\ a_i \in A \right\}.$$

Se R é comutativo, essa expressão se reduz às somas finitas $\sum_i r_i a_i$.

Quando $A = \{a\}$, escrevemos simplesmente $\langle a \rangle$ e dizemos que esse ideal é **principal**.

Definição 1.17 — Domínio de ideais principais. Um **domínio de ideais principais**, ou PID, é um domínio de integridade no qual todo ideal é principal.

Mais adiante provaremos que $\mathbb{K}[x]$ é um domínio de ideais principais. Essa propriedade permitirá interpretar o máximo divisor comum de dois polinômios como o gerador de um ideal.

1.3 Corpos

Os corpos são os sistemas de escalares da Álgebra Linear. Neles podemos realizar as quatro operações aritméticas, com a única restrição habitual de que não se pode dividir por zero.

Definição 1.18 — Corpo. Um **corpo** $\mathbb{K}$ é um anel comutativo no qual todo elemento não nulo possui inverso multiplicativo.

Corpos gerais
Grande parte da Álgebra Linear é válida sobre um corpo arbitrário. Ainda assim, a característica do corpo importa: argumentos que dividem por um inteiro ou usam conjugação exigem hipóteses adicionais.

Exemplos 1.19.

1. Os conjuntos $\mathbb{Q}$, $\mathbb{R}$ e $\mathbb{C}$, com as operações usuais, são corpos e $\mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$.
2. Os inteiros não formam um corpo: por exemplo, 2 não possui inverso multiplicativo em $\mathbb{Z}$.
3. Para $n \geq 2$, o anel $M_n(\mathbb{K})$ não é um corpo. Sua multiplicação não é comutativa e existem matrizes não nulas que não são invertíveis.

◁

Exemplo 1.20. Considere

$$\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \;:\; a, b \in \mathbb{Q}\}.$$

Esse conjunto é fechado por soma, diferença e produto. Se $a+b\sqrt{2} \neq 0$, então $a^2-2b^2 \neq 0$ e

$$(a + b\sqrt{2})^{-1} = \frac{a - b\sqrt{2}}{a^2 - 2b^2} \in \mathbb{Q}(\sqrt{2}).$$

De fato, a igualdade $a^2 = 2b^2$, com $b \neq 0$, implicaria que $a/b = \sqrt{2}$ fosse racional. Portanto, $\mathbb{Q}(\sqrt{2})$ é um corpo.

◁

Exemplo 1.21. Se p é primo, o conjunto das classes de congruência módulo p,

$$\mathbb{Z}/p\mathbb{Z} = \{\overline{0}, \overline{1}, \ldots, \overline{p-1}\},$$

é um corpo com as operações módulo p. Ele será denotado também por $\mathbb{F}_p$ ou $\mathbb{Z}_p$. A primalidade de p é essencial: se $n = ab$ com $1 < a, b < n$, então $\overline{a}\,\overline{b} = \overline{0}$ em $\mathbb{Z}/n\mathbb{Z}$, embora os dois fatores sejam não nulos. Para p primo, por outro lado, multiplicar as classes não nulas por uma classe fixa não nula apenas as permuta; em particular, alguma delas é enviada a $\overline{1}$, fornecendo o inverso multiplicativo.

◁

Proposição 1.22 (Propriedades elementares).
Se $a, b \in \mathbb{K}$, então:

1. $a0 = 0a = 0$;
2. $a(-b) = (-a)b = -(ab)$;
3. $(-a)(-b) = ab$;
4. *se $ab = 0$, então $a = 0$ ou $b = 0$;*
5. *se $a \neq 0$ e $ab = ac$, então $b = c$.*

Demonstração. Pela distributividade,

$$a0 = a(0 + 0) = a0 + a0.$$

Somando o oposto de $a0$ aos dois lados, obtemos $a0 = 0$; analogamente, $0a = 0$. Além disso,

$$ab + a(-b) = a(b - b) = a0 = 0,$$

de modo que $a(-b) = -(ab)$. A igualdade $(-a)b = -(ab)$ é semelhante, e aplicá-la duas vezes fornece $(-a)(-b) = ab$.

Se $ab = 0$ e $a \neq 0$, multiplicar por a^{-1} dá $b = 0$. Finalmente, de $ab = ac$ segue $a(b - c) = 0$; se $a \neq 0$, a propriedade anterior implica $b - c = 0$, isto é, $b = c$. ■

Definição 1.23 — Subcorpo. Um subconjunto $\mathbb{F} \subseteq \mathbb{K}$ é um **subcorpo** quando é um corpo com as operações herdadas de $\mathbb{K}$. Nesse caso, dizemos também que $\mathbb{K}$ é uma extensão de $\mathbb{F}$.

Por exemplo, $\mathbb{Q}$ é subcorpo de $\mathbb{R}$, que é subcorpo de $\mathbb{C}$. As identidades aditiva e multiplicativa de um subcorpo coincidem com as do corpo que o contém.

Definição 1.24 — Característica. A **característica** de um corpo $\mathbb{K}$ é o menor inteiro positivo n para o qual

$$\underbrace{1 + \cdots + 1}_{n \text{ parcelas}} = 0,$$

se tal inteiro existir. Caso contrário, dizemos que $\mathbb{K}$ tem característica zero.

Proposição 1.25.
A característica de um corpo é zero ou um número primo.

Demonstração. Suponha que a característica seja $n > 0$. Se $n = rs$, com $1 < r, s < n$, então

$$(r \cdot 1)(s \cdot 1) = n \cdot 1 = 0.$$

Pela minimalidade de n, os dois fatores são não nulos, contradizendo a ausência de divisores de zero. Logo, n é primo. ■

Os corpos $\mathbb{Q}$, $\mathbb{R}$ e $\mathbb{C}$ têm característica zero, enquanto $\mathbb{F}_p$ tem característica p.

! Em característica p, um inteiro m representa o escalar $m \cdot 1$. Assim, uma divisão por m só é permitida quando p não divide m.

Exercícios

Ex. 1.5 — Considere o conjunto $\mathbb{K} = \{0, 1\}$ com as operações dadas pelas tabelas

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

.

Verifique que $\mathbb{K}$ é um corpo e identifique sua característica.

Ex. 1.6 — Seja $\mathbb{K}$ um corpo. Demonstre diretamente que:

1. $-a = (-1)a$;
2. $(-a)(-b) = ab$;
3. se $a \neq 0$ e $ab = ac$, então $b = c$;
4. $(ab)^{-1} = a^{-1}b^{-1}$ quando $a, b \neq 0$.

Ex. 1.7 — Prove que $\mathbb{Q}(\sqrt{2})$ é um corpo. Em seguida, determine o inverso de $3 + 2\sqrt{2}$.

Ex. 1.8 — Seja d um inteiro positivo que não é quadrado perfeito. Mostre que $\mathbb{Q}(\sqrt{d}) = \{a + b\sqrt{d} \,:\, a, b \in \mathbb{Q}\}$ é um corpo.

Ex. 1.9 — Mostre que todo subcorpo de $\mathbb{C}$ contém uma cópia de $\mathbb{Q}$.

Ex. 1.10 — Mostre que um corpo de característica zero é infinito.

Ex. 1.11 — Seja $\mathbb{K}$ um corpo de característica $p > 0$. Mostre que a aplicação $\mathbb{F}_p \to \mathbb{K}$, definida por $\overline{m} \mapsto m \cdot 1$, é um homomorfismo injetivo. Assim, todo corpo de característica p contém uma cópia de $\mathbb{F}_p$.

Ex. 1.12 — Determine quais dos anéis $\mathbb{Z}/4\mathbb{Z}$, $\mathbb{Z}/5\mathbb{Z}$ e $\mathbb{Z}/6\mathbb{Z}$ são corpos. Nos demais casos, exiba divisores de zero não nulos.

1.4 O anel de polinômios

Seja $\mathbb{K}$ um corpo. Um polinômio com coeficientes em $\mathbb{K}$ é uma soma formal

$$p(x) = \sum_{j=0}^{m} a_j x^j, \qquad a_j \in \mathbb{K},$$

na qual apenas um número finito de coeficientes é não nulo. Dois polinômios são iguais quando possuem os mesmos coeficientes. O conjunto de todos eles é denotado por $\mathbb{K}[x]$.

Polinômios e funções polinomiais
Um polinômio é determinado por sua sequência de coeficientes. Sobre corpos finitos, polinômios diferentes podem definir a mesma função; por isso é importante distinguir o objeto formal de sua função de avaliação.

Definição 1.26 — Grau. Se $p \neq 0$, o **grau** de p é o maior índice m para o qual $a_m \neq 0$ e será denotado por $\operatorname{grau} p$. O termo $a_m x^m$ é o termo principal e a_m é o coeficiente principal. O polinômio é **mônico** quando $a_m = 1$. Adotaremos a convenção $\operatorname{grau} 0 = -\infty$.

A soma e o produto são definidos coeficiente a coeficiente por

$$\left(\sum_{i\geq 0} a_i x^i\right) + \left(\sum_{i\geq 0} b_i x^i\right) = \sum_{i\geq 0} (a_i + b_i) x^i$$

e

$$\left(\sum_{i\geq 0} a_i x^i\right)\left(\sum_{j\geq 0} b_j x^j\right) = \sum_{\ell\geq 0}\left(\sum_{i+j=\ell} a_i b_j\right) x^\ell.$$

As somas são finitas porque os polinômios têm apenas um número finito de coeficientes não nulos.

Exemplo 1.27. Em $\mathbb{Q}[x]$,

$$(2x^2 - x + 1)(x + 3) = 2x^3 + 5x^2 - 2x + 3.$$

O coeficiente de x^2, por exemplo, é $2 \cdot 3 + (-1) \cdot 1 = 5$.

◁

Proposição 1.28.
O conjunto $\mathbb{K}[x]$, com essas operações, é um domínio de integridade. Se $f, g \neq 0$, então

$$\operatorname{grau}(fg) = \operatorname{grau} f + \operatorname{grau} g.$$

Consequentemente, as unidades de $\mathbb{K}[x]$ são exatamente os polinômios constantes não nulos.

Demonstração. As propriedades de anel seguem diretamente das propriedades das operações em $\mathbb{K}$ e das definições dos coeficientes da soma e do produto. Se os termos principais de f e g são $a_m x^m$ e $b_n x^n$, o termo principal de fg é $a_m b_n x^{m+n}$. Como $\mathbb{K}$ não possui divisores de zero, $a_m b_n \neq 0$. Portanto, $fg \neq 0$ e $\operatorname{grau}(fg) = m + n$.

Se $fg = 1$, a igualdade dos graus fornece $\operatorname{grau} f + \operatorname{grau} g = 0$; logo, ambos são constantes não nulas. Reciprocamente, toda constante não nula é invertível porque pertence ao corpo $\mathbb{K}$. ■

Para $m \geq 0$, denotaremos por $\mathbb{K}_{(m)}[x]$ o conjunto dos polinômios de grau no máximo m, incluindo o polinômio zero. No próximo capítulo veremos que esse conjunto é um espaço vetorial de dimensão $m + 1$.

1.5 Divisão euclidiana e máximo divisor comum

Nos inteiros, a divisão com resto sustenta o algoritmo de Euclides. A mesma estrutura existe em $\mathbb{K}[x]$, com o grau ocupando o papel desempenhado pelo valor absoluto.

> **Teorema 1.29** (Divisão Euclidiana).
> *Sejam $f, g \in \mathbb{K}[x]$, com $g \neq 0$. Existem únicos polinômios $q, r \in \mathbb{K}[x]$ tais que*
>
> $$f = gq + r, \qquad r = 0 \quad \textit{ou} \quad \operatorname{grau} r < \operatorname{grau} g.$$

Demonstração. Provaremos a existência por indução forte sobre $\operatorname{grau} f$. Se $f = 0$ ou $\operatorname{grau} f < \operatorname{grau} g$, podemos tomar, respectivamente, $(q, r) = (0, 0)$ ou $(q, r) = (0, f)$.

Suponha agora que f e g tenham termos principais $a_n x^n$ e $b_m x^m$, com $n \geq m$. O polinômio

$$h = f - a_n b_m^{-1} x^{n-m} g$$

é nulo ou possui grau menor que n. Pela hipótese de indução, $h = gq_1 + r$, com $r = 0$ ou $\operatorname{grau} r < m$. Assim,

$$f = g\left(q_1 + a_n b_m^{-1} x^{n-m}\right) + r,$$

o que prova a existência.

Para a unicidade, suponha que $f = gq_1 + r_1 = gq_2 + r_2$, com $r_i = 0$ ou $\operatorname{grau} r_i < \operatorname{grau} g$, para $i = 1, 2$. Então

$$g(q_1 - q_2) = r_2 - r_1.$$

Se $q_1 \neq q_2$, o lado esquerdo tem grau pelo menos $\operatorname{grau} g$, enquanto o lado direito é nulo ou tem grau menor que $\operatorname{grau} g$, uma contradição. Logo, $q_1 = q_2$ e, em seguida, $r_1 = r_2$. ■

Exemplo 1.30 — Divisão de polinômios. Em $\mathbb{Q}[x]$,

$$x^3 - 2x + 4 = (x-1)(x^2 + x - 1) + 3.$$

O quociente é $x^2 + x - 1$ e o resto é 3, cujo grau é menor que o grau de $x - 1$.

◁

Corolário 1.31.
O anel $\mathbb{K}[x]$ *é um domínio de ideais principais.*

Demonstração. Seja $I \neq \{0\}$ um ideal de $\mathbb{K}[x]$ e escolha em I um polinômio não nulo d de grau mínimo. Para cada $f \in I$, escreva $f = dq + r$ pela divisão euclidiana. Como $r = f - dq \in I$ e $r = 0$ ou $\operatorname{grau} r < \operatorname{grau} d$, a minimalidade de d força $r = 0$. Portanto, d divide todo elemento de I e $I = \langle d \rangle$. O ideal nulo é gerado por 0. ■

Definição 1.32 — Máximo divisor comum. Se $f, g \in \mathbb{K}[x]$ não são ambos nulos, o **máximo divisor comum** de f e g é o único polinômio mônico d tal que

1. d divide f e g;
2. todo divisor comum de f e g divide d.

Ele será denotado por $\operatorname{mdc}(f, g)$.

Teorema 1.33 (Identidade de Bézout).
Se $d = \operatorname{mdc}(f, g)$*, existem* $r, s \in \mathbb{K}[x]$ *tais que*

$$d = rf + sg.$$

Demonstração. Como $\mathbb{K}[x]$ é um PID, o ideal $\langle f, g \rangle$ possui um gerador d_0. Multiplicando-o por uma constante não nula, podemos supor que d_0 é mônico. Como $f, g \in \langle d_0 \rangle$, o polinômio d_0 divide ambos. Se c é outro divisor comum, então divide toda combinação $rf + sg$ e, portanto, divide d_0. Logo, $d_0 = \operatorname{mdc}(f, g)$.

Essa caracterização também prova a unicidade: dois polinômios mônicos com as propriedades do máximo divisor comum dividem um ao outro e, portanto, são iguais.

Finalmente, $d_0 \in \langle f, g \rangle$, de modo que existem r, s com $d_0 = rf + sg$. ■

Proposição 1.34.
Se $f = gq + r$*, então*

$$\operatorname{mdc}(f, g) = \operatorname{mdc}(g, r).$$

Demonstração. Um polinômio divide simultaneamente f e g se, e somente se, divide simultaneamente g e $r = f - gq$. Os dois pares têm, portanto, os mesmos divisores comuns e o mesmo máximo divisor comum mônico. ■

Aplicando repetidamente a divisão euclidiana, obtemos uma sequência de restos cujos graus decrescem estritamente. O processo termina, e o último resto não nulo, depois de dividido por seu coeficiente principal, é o máximo divisor comum. Esse procedimento é o **algoritmo de Euclides**.

Normalização do mdc
O último resto não nulo do algoritmo de Euclides é um múltiplo escalar do mdc. Como adotamos a convenção de que o mdc é mônico, é necessário dividir o resto por seu coeficiente principal.

Exemplo 1.35 — Algoritmo de Euclides. Para $f = x^3 - 1$ e $g = x^2 - 1$,

$$f = xg + (x - 1), \qquad g = (x + 1)(x - 1).$$

Logo, $\operatorname{mdc}(f, g) = x - 1$. A primeira igualdade fornece também a identidade de Bézout

$$x - 1 = f - xg.$$

◁

1.6 Raízes e multiplicidades

Definição 1.36 — Raiz. Se $p \in \mathbb{K}[x]$, um elemento $\alpha \in \mathbb{K}$ é uma **raiz** de p quando $p(\alpha) = 0$.

Proposição 1.37 (Teorema do Fator).
Um elemento $\alpha \in \mathbb{K}$ é raiz de $p \in \mathbb{K}[x]$ se, e somente se, $x - \alpha$ divide p.

Demonstração. Pela divisão euclidiana, existem $q \in \mathbb{K}[x]$ e uma constante r tais que

$$p = (x - \alpha)q + r.$$

Substituindo $x = \alpha$, obtemos $r = p(\alpha)$. Portanto, o resto é nulo se, e somente se, α é raiz de p. ■

Teorema 1.38.
Um polinômio não nulo de grau n possui no máximo n raízes distintas em $\mathbb{K}$.

Demonstração. Procedemos por indução sobre n. Um polinômio constante não nulo não possui raízes. Se $n \geq 1$ e p não possui raízes, nada há a provar. Caso contrário, seja α uma raiz. Pelo Teorema do Fator, $p = (x - \alpha)q$, com $\operatorname{grau} q = n - 1$. Se $\beta \neq \alpha$ é outra raiz de p, então

$$0 = p(\beta) = (\beta - \alpha)q(\beta).$$

Como $\beta - \alpha \neq 0$ e um corpo não possui divisores de zero, $q(\beta) = 0$. Assim, toda raiz de p diferente de α é raiz de q, que possui no máximo $n - 1$ raízes pela hipótese de indução. Assim, p possui no máximo n raízes. ■

Definição 1.39 — Multiplicidade. Se $p \neq 0$ e α é raiz de p, a **multiplicidade** de α é o único inteiro $m \geq 1$ para o qual

$$p(x) = (x - \alpha)^m s(x), \qquad s(\alpha) \neq 0.$$

Quando $m = 1$, a raiz é simples; quando $m \geq 2$, é múltipla.

A multiplicidade está bem definida: o Teorema do Fator garante ao menos um fator $x - \alpha$, e o grau de p impede que as potências divisoras $(x - \alpha)^k$ tenham expoentes arbitrariamente grandes.

A derivada formal de $p(x) = \sum_{j=0}^{n} a_j x^j$ é

$$p'(x) = \sum_{j=1}^{n} j a_j x^{j-1}.$$

Essa definição é puramente algébrica e continua válida em qualquer característica.

Proposição 1.40 (Multiplicidade e derivadas).
Se $\mathbb{K}$ tem característica zero, $p \neq 0$ e $m \geq 1$, então α é raiz de multiplicidade m se, e somente se,

$$p^{(j)}(\alpha) = 0 \quad (0 \leq j < m), \qquad p^{(m)}(\alpha) \neq 0.$$

Demonstração. Escreva $p = (x - \alpha)^r s$, com $s(\alpha) \neq 0$, onde r é a multiplicidade de α. Pela regra de Leibniz, todas as derivadas de ordem menor que r ainda contêm um fator $x - \alpha$ e se anulam em α. Na derivada de ordem r, o único termo que não contém esse fator fornece

$$p^{(r)}(\alpha) = r!\, s(\alpha) \neq 0.$$

A última desigualdade usa a hipótese de característica zero. Portanto, r é exatamente a ordem da primeira derivada que não se anula em α, o que prova as duas implicações. ∎

1.7 Irredutibilidade e fatoração

Definição 1.41 — Polinômio irredutível. Um polinômio não constante $p \in \mathbb{K}[x]$ é **irredutível** sobre $\mathbb{K}$ quando, em toda fatoração $p = fg$, um dos fatores é constante. Caso contrário, p é **redutível**.

O corpo faz diferença
Irredutibilidade depende dos escalares permitidos. O polinômio $x^2 + 1$ é irredutível em $\mathbb{R}[x]$, mas se fatora como $(x - i)(x + i)$ em $\mathbb{C}[x]$.

Lema 1.42 (Lema de Euclides).
Se $p \in \mathbb{K}[x]$ é irredutível e p divide ab, então p divide a ou p divide b.

Demonstração. Suponha que p não divida a. Como p é irredutível, seus únicos divisores mônicos são 1 e p; logo, $\operatorname{mdc}(p, a) = 1$. Pela identidade de Bézout, existem $r, s \in \mathbb{K}[x]$ tais que $rp + sa = 1$. Multiplicando por b,

$$b = rpb + sab.$$

As duas parcelas do lado direito são divisíveis por p, pois p divide ab. Portanto, p divide b. ■

Teorema 1.43 (Fatoração única).
Todo polinômio não constante $f \in \mathbb{K}[x]$ admite uma fatoração

$$f = c\, p_1 p_2 \cdots p_r,$$

na qual $c \in \mathbb{K} \setminus \{0\}$ e os p_i são irredutíveis mônicos. Essa fatoração é única a menos da ordem dos fatores.

Demonstração. A existência é demonstrada por indução sobre grau f. Se f é irredutível, basta dividir f por seu coeficiente principal. Se é redutível, escrevemos $f = gh$, com $0 < \text{grau}\, g, \text{grau}\, h < \text{grau}\, f$, e aplicamos a hipótese de indução aos dois fatores.

Para a unicidade, suponha

$$c\, p_1 \cdots p_r = d\, q_1 \cdots q_s,$$

com todos os fatores irredutíveis e mônicos. O Lema de Euclides implica que p_1 divide algum q_j. Como ambos são irredutíveis e mônicos, $p_1 = q_j$. Cancelando esse fator e repetindo o argumento, obtemos $r = s$ e, após uma reordenação, $p_i = q_i$ para todo i. A comparação dos coeficientes principais fornece $c = d$. ■

Proposição 1.44.
Um polinômio de grau 2 ou 3 é irredutível sobre $\mathbb{K}$ se, e somente se, não possui raízes em $\mathbb{K}$.

Demonstração. Se f possui uma raiz α, o Teorema do Fator produz a fatoração não trivial $f = (x-\alpha)q$. Reciprocamente, se f é redutível e tem grau 2 ou 3, numa fatoração não trivial ao menos um dos fatores tem grau 1. Todo polinômio linear possui uma raiz em $\mathbb{K}$, que também é raiz de f. ■

O resultado não se estende ao grau 4: o polinômio $x^4 + 3x^2 + 2 = (x^2+1)(x^2+2)$ não possui raízes reais, mas é redutível em $\mathbb{R}[x]$.

1.8 Polinômios sobre os reais e os complexos

Definição 1.45 — Corpo algebricamente fechado. Um corpo $\mathbb{K}$ é **algebricamente fechado** quando todo polinômio não constante de $\mathbb{K}[x]$ possui uma raiz em $\mathbb{K}$.

Teorema 1.46.
Se $\mathbb{K}$ é algebricamente fechado, todo polinômio $f \in \mathbb{K}[x]$ de grau $n \geq 1$ se fatora

como

$$f(x) = c\prod_{j=1}^{n}(x-\alpha_j),$$

onde $c \neq 0$ e as raízes são contadas com multiplicidade.

Demonstração. Por indução sobre n. Como $\mathbb{K}$ é algebricamente fechado, f possui uma raiz α_1. Pelo Teorema do Fator, $f = (x-\alpha_1)g$, com $\operatorname{grau} g = n-1$. Aplicamos a hipótese de indução a g. O caso inicial $n = 1$ é imediato. ■

Teorema 1.47 (Teorema Fundamental da Álgebra).
Todo polinômio não constante de $\mathbb{C}[x]$ possui uma raiz em $\mathbb{C}$. Equivalentemente, $\mathbb{C}$ é algebricamente fechado.

A demonstração do Teorema Fundamental da Álgebra utiliza ferramentas que não fazem parte deste livro. O resultado será usado aqui como teorema clássico; a fatoração em fatores lineares segue da proposição anterior.

Lema 1.48.
Se $f \in \mathbb{R}[x]$ e $\alpha \in \mathbb{C}$ é raiz de f, então $\overline{\alpha}$ também é raiz de f.

Demonstração. Como os coeficientes de f são reais,

$$f(\overline{\alpha}) = \overline{f(\alpha)} = 0.$$

■

Teorema 1.49 (Fatoração real).
Todo polinômio não constante de $\mathbb{R}[x]$ é produto de polinômios lineares e quadráticos irredutíveis sobre $\mathbb{R}$.

Demonstração. Pelo Teorema Fundamental da Álgebra, f se fatora em fatores lineares sobre $\mathbb{C}$. Como f tem coeficientes reais, $\overline{f(z)} = f(\overline{z})$; portanto, as raízes não reais aparecem em pares conjugados com a mesma multiplicidade. Se $\alpha = a + bi$, com $b \neq 0$, então

$$(x-\alpha)(x-\overline{\alpha}) = x^2 - 2ax + (a^2+b^2),$$

um polinômio real de discriminante $-4b^2 < 0$ e, portanto, irredutível em $\mathbb{R}[x]$. Agrupando cada par conjugado e mantendo os fatores associados às raízes reais, obtemos a fatoração desejada. ■

Teorema 1.50.
Os polinômios irredutíveis em $\mathbb{R}[x]$ são exatamente os polinômios lineares e os quadráticos com discriminante negativo.

Demonstração. Todo polinômio linear é irredutível. Um quadrático é irredutível se, e somente se, não possui raiz real, o que equivale a ter discriminante negativo. Por outro lado, o teorema anterior mostra que todo polinômio real de grau maior que 2 se decompõe em fatores de graus 1 e 2 e, portanto, é redutível. ■

Em particular, todo polinômio real de grau ímpar maior que 1 possui uma raiz real e é redutível. A ressalva sobre o grau é necessária: polinômios lineares são irredutíveis.

Exercícios

Ex. 1.13 — Seja R um anel. Prove que, para quaisquer $a, b \in R$,

1. $0a = a0 = 0$;
2. $a(-b) = (-a)b = -(ab)$;
3. $(-a)(-b) = ab$;
4. $(-1)a = -a$ e $(-1)(-1) = 1$.

Ex. 1.14 — Seja R um anel comutativo. Mostre que

$$\langle a \rangle = \{ra : r \in R\}$$

é um ideal de R para todo $a \in R$.

Ex. 1.15 — Verifique diretamente, a partir das definições das operações, que $\mathbb{K}[x]$ é um anel comutativo com identidade.

Ex. 1.16 — Sejam $p, q \in \mathbb{K}[x]$ não nulos.

1. Se $p \neq -q$, mostre que $\operatorname{grau}(p+q) \leq \max\{\operatorname{grau} p, \operatorname{grau} q\}$ e determine quando ocorre igualdade.
2. Prove que $\operatorname{grau}(pq) = \operatorname{grau} p + \operatorname{grau} q$.

Ex. 1.17 — Seja $S \subseteq \mathbb{K}[x]$ não vazio. Mostre que

$$M(S) = \left\{ \sum_{p \in F} a_p p \ : \ F \subseteq S \text{ é finito e } a_p \in \mathbb{K}[x] \right\}$$

é o ideal gerado por S.

Ex. 1.18 — Efetue a divisão euclidiana de f por g nos casos seguintes.

1. $f = x^4 - 1$ e $g = x - 1$;
2. $f = 2x^4 + x^3 - 3x + 1$ e $g = x^2 + 1$;
3. $f = x^5 + 1$ e $g = x^2 + x + 1$.

Ex. 1.19 — Use o algoritmo de Euclides para determinar o máximo divisor comum dos pares abaixo. Em cada caso, faça a retro-substituição e obtenha uma identidade de Bézout.

1. $2x^5 - x^3 - 3x^2 - 6x + 4$ e $x^4 + x^3 - x^2 - 2x - 2$;
2. $3x^4 + 8x^2 - 3$ e $x^3 + 2x^2 + 3x + 6$;
3. $x^4 - 2x^3 - 2x^2 - 2x - 3$ e $x^3 + 6x^2 + 7x + 1$.

Ex. 1.20 — Sejam $p_1, \dots, p_n \in \mathbb{K}[x]$ coprimos dois a dois. Se cada p_i divide q, prove que $p_1 \cdots p_n$ divide q.

Ex. 1.21 — Mostre que $x - \alpha$ e $x - \beta$ são coprimos se, e somente se, $\alpha \neq \beta$.

Ex. 1.22 — Sejam $p, q \in \mathbb{K}[x]$ e sejam $a_1, \dots, a_m$ elementos distintos de $\mathbb{K}$. Suponha que $p(a_i) = q(a_i)$ para todo i e que

$$m > \max\{\text{grau } p, \text{grau } q\}.$$

Prove que $p = q$.

Ex. 1.23 — Determine quais dos subconjuntos de $\mathbb{Q}[x]$ abaixo são ideais. Quando houver um ideal, determine seu gerador mônico.

1. Os polinômios de grau ímpar;
2. os polinômios de grau pelo menos 5;
3. os polinômios p tais que $p(0) = 0$;
4. os polinômios p tais que $p(2) = p(4) = 0$;

5. os polinômios da forma

$$\sum_{i=0}^{n} \frac{c_i}{i+1} x^{i+1}, \qquad c_0, \dots, c_n \in \mathbb{Q}.$$

Ex. 1.24 — Em um corpo de característica zero, considere $p(x) = (x-1)^3(x+2)^2$. Determine as multiplicidades de suas raízes e verifique o resultado calculando derivadas formais sucessivas nos pontos 1 e -2.

Ex. 1.25 — Seja $\mathbb{K}$ um corpo de característica zero e seja $D : \mathbb{K}[x] \to \mathbb{K}[x]$ a derivada formal. Se grau $p \leq n$, prove a fórmula de Taylor

$$p(x) = \sum_{k=0}^{n} \frac{D^k p(a)}{k!} (x-a)^k.$$

Dica: faça a mudança de variável $y = x - a$ e compare os coeficientes de $p(a+y)$.

Ex. 1.26 — Seja $p \in \mathbb{C}[x]$. Prove que p possui apenas raízes simples se, e somente se, $\operatorname{mdc}(p, p') = 1$.

Ex. 1.27 — Mostre que, se p é irredutível e não divide q, então $\operatorname{mdc}(p, q) = 1$.

Ex. 1.28 — Dizemos que um polinômio não constante p é *primo* quando $p \mid ab$ implica $p \mid a$ ou $p \mid b$. Prove que, em $\mathbb{K}[x]$, um polinômio é primo se, e somente se, é irredutível.

Ex. 1.29 — Seja $p \in \mathbb{K}[x]$ não constante.

1. Mostre que todo polinômio de grau 1 é irredutível.
2. Mostre que, nos graus 2 e 3, irredutibilidade equivale à ausência de raízes em $\mathbb{K}$.
3. Dê um polinômio de grau 4 sem raízes em $\mathbb{K}$ que seja redutível.

Ex. 1.30 — Use o Teorema Fundamental da Álgebra para provar que:

1. todo polinômio complexo não constante se fatora em fatores lineares;
2. se $p \in \mathbb{R}[x]$ e $\alpha \in \mathbb{C}$ é raiz de p, então $\overline{\alpha}$ também é raiz;
3. $(x-\alpha)(x-\overline{\alpha})$ é um quadrático real irredutível quando $\alpha \notin \mathbb{R}$;
4. todo polinômio real é produto de fatores lineares e quadráticos irredutíveis;

5. todo polinômio real de grau ímpar possui ao menos uma raiz real.

Ex. 1.31 — Admita o seguinte caso do critério de Eisenstein: se um primo q divide todos os coeficientes não líderes de um polinômio inteiro, não divide o coeficiente líder e q^2 não divide o termo constante, então o polinômio é irredutível em $\mathbb{Q}[x]$. Use-o para provar que $x^n - q$ é irredutível para todo $n \geq 1$. Conclua que existem polinômios irredutíveis de qualquer grau em $\mathbb{Q}[x]$.

Ex. 1.32 — Assumindo o Teorema Fundamental da Álgebra, prove que dois polinômios complexos são coprimos se, e somente se, não possuem raiz comum.

Para retomada posterior. Os problemas seguintes utilizam conceitos de matrizes e determinantes desenvolvidos em outros capítulos.

Ex. 1.33 — Suponha que $\operatorname{char} \mathbb{K} \neq 2$.

1. Mostre que, para cada $n \geq 2$, existe $p \in \mathbb{K}[x]$, com $\operatorname{grau} p \leq n-1$, tal que x^n divide $1 + x - p^2$.

2. Deduza que, se $N \in M_n(\mathbb{K})$ é nilpotente, então $I + N$ possui uma raiz quadrada.

3. Dê um exemplo explícito para $n = 3$.

Ex. 1.34 — Sejam $A, B \in M_n(\mathbb{Z})$ matrizes não singulares cujos determinantes são coprimos. Prove que existem $U, V \in M_n(\mathbb{Z})$ tais que

$$AU + BV = I_n.$$

Dica: use a identidade de Bézout para $\det A$ e $\det B$ e as matrizes adjugadas de A e B.

2
Espaços Vetoriais

> *When did mathematicians first generally accept the definition of a vector space as a set of elements subject to suitably axiomatized operations of addition and multiplication by scalars, and not just as a set of n-tuples of scalars closed under these two operations? This abstract description did not come into early use, but has the evident advantages of conceptual simplicity and geometric invariance.*
>
> — Mac Lane

A álgebra linear tem raízes históricas em diversos ramos da matemática. Na álgebra, elas incluem equações lineares, formas bilineares, matrizes, quatérnios e outros sistemas hipercomplexos; na geometria, incluem segmentos orientados — os vetores geométricos — e as geometrias afim e projetiva; na análise, aparecem em equações diferenciais lineares e em espaços de dimensão infinita. A física de Maxwell, Gibbs e Heaviside também contribuiu decisivamente para seu desenvolvimento.

A noção geral de espaço vetorial, comum a esses desenvolvimentos, foi isolada e adotada relativamente tarde.

O tratamento moderno e abstrato foi formulado por Giuseppe Peano, em 1888, no contexto da geometria. A formulação teve pouca influência imediata e tampouco se difundiu quando Weyl a retomou em 1918. Por volta de 1920, ideias semelhantes reapareceram nos trabalhos de Banach, Hahn, Wiener e Noether; a partir daí, a noção de espaço vetorial se desenvolveu rapidamente [27].

Nesse capítulo

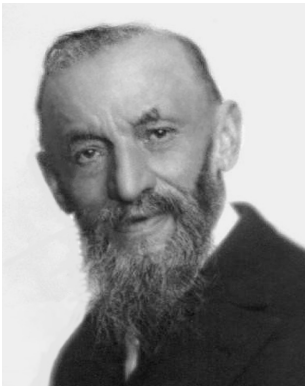


Figura 2.1 Giuseppe Peano

2.1 Definições e Exemplos

No restante do texto, salvo indicação em contrário, assumiremos que $\mathbb{K}$ é um corpo arbitrário.

Operações e axiomas
Os axiomas especificam as propriedades da adição e da multiplicação por escalares, sem exigir uma natureza particular para os vetores. Por isso, a mesma teoria se aplica a vetores geométricos, polinômios, matrizes, sequências e funções.

Definição 2.1 — Espaço Vetorial. Um espaço vetorial V sobre $\mathbb{K}$ é um conjunto não vazio munido de duas aplicações: a adição $(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} + \mathbf{y}$, de $V \times V$ em V, e a multiplicação por escalares $(\lambda, \mathbf{x}) \mapsto \lambda\mathbf{x}$, de $\mathbb{K} \times V$ em V. Essas operações devem satisfazer os seguintes axiomas:

- **V1** $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ para todos os $\mathbf{x}, \mathbf{y} \in V$.
- **V2** $\mathbf{x} + (\mathbf{y} + \vec{z}) = (\mathbf{x} + \mathbf{y}) + \vec{z}$ para todos os $\mathbf{x}, \mathbf{y}, \vec{z} \in V$.
- **V3** Existe um elemento $\vec{0} \in V$ tal que $\vec{0} + \mathbf{x} = \mathbf{x}$ para todo $\mathbf{x} \in V$.
- **V4** Para cada $\mathbf{x} \in V$, existe um $\mathbf{y} \in V$ tal que $\mathbf{x} + \mathbf{y} = \vec{0}$.
- **V5** $(\lambda_1\lambda_2)\mathbf{x} = \lambda_1(\lambda_2\mathbf{x})$ para todos os $\lambda_1, \lambda_2 \in \mathbb{K}$ e $\mathbf{x} \in V$.
- **V6** $\lambda_1(\mathbf{x} + \mathbf{y}) = \lambda_1\mathbf{x} + \lambda_1\mathbf{y}$ para todos os $\lambda_1 \in \mathbb{K}$ e $\mathbf{x}, \mathbf{y} \in V$.
- **V7** $(\lambda_1 + \lambda_2)\mathbf{x} = \lambda_1\mathbf{x} + \lambda_2\mathbf{x}$ para todos os $\lambda_1, \lambda_2 \in \mathbb{K}$ e $\mathbf{x} \in V$.
- **V8** $1\mathbf{x} = \mathbf{x}$ para todos os $\mathbf{x} \in V$.

Assim como ocorre com os corpos, um espaço vetorial inclui não apenas o conjunto V, mas também as duas operações especificadas acima. Em outras palavras, trata-se da tripla formada por V, pela adição e pela multiplicação por escalares, sujeita aos axiomas V1–V8.

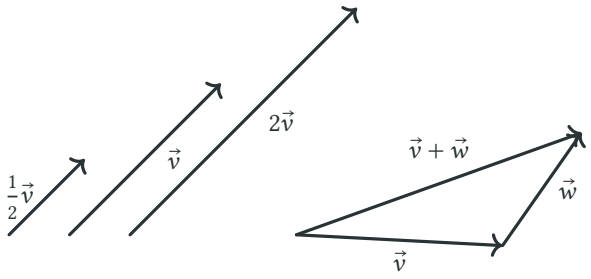


Figura 2.2 Representação geométrica da multiplicação de um vetor por um escalar e da soma de vetores. Essa intuição será útil mesmo quando os vetores forem polinômios, matrizes ou funções.

Em geral, um mesmo conjunto pode receber diferentes estruturas de espaço vetorial sobre $\mathbb{K}$. Quando não houver risco de confusão, deixaremos as operações implícitas e usaremos simplesmente V para denotar o espaço.

Se V for um espaço vetorial sobre $\mathbb{K}$, os elementos de V serão chamados **vetores**, e os elementos de $\mathbb{K}$, **escalares**.

Apresentamos algumas consequências simples dessa definição.

Proposição 2.2. [a] *O elemento $\vec{0} \in V$ é único. Ou seja, existe um único $\vec{0}$ tal que $\vec{0} + \mathbf{x} = \mathbf{x}$ para todo $\mathbf{x} \in V$.*

[b] *O elemento oposto é único. Ou seja, para todo $\mathbf{x} \in V$, existe um único $\mathbf{y} \in V$ tal que $\mathbf{x} + \mathbf{y} = \vec{0}$. Esse vetor será denotado por $-\mathbf{x}$.*

[c] $0\mathbf{x} = \lambda\vec{0} = \vec{0}$ *para todos os* $\lambda \in \mathbb{K}$ *e* $\mathbf{x} \in V$.

[d] $(-1)\mathbf{x} = -\mathbf{x}$.

[e] *Se* $\lambda\mathbf{x} = \vec{0}$, *então* $\lambda = 0$ *ou* $\mathbf{x} = \vec{0}$.

[f] *Se* $\mathbf{x} + \mathbf{y} = \mathbf{z} + \mathbf{y}$ *então* $\mathbf{x} = \mathbf{z}$.

Demonstração. [a] Se $\vec{0}$ e $\vec{0}'$ satisfazem a propriedade do vetor nulo, então $\vec{0} = \vec{0} + \vec{0}' = \vec{0}'$. Logo o vetor nulo é único.

[b] Se $\mathbf{y}$ e $\mathbf{z}$ são opostos de $\mathbf{x}$, então

$$\mathbf{y} = \mathbf{y} + \vec{0} = \mathbf{y} + (\mathbf{x} + \mathbf{z}) = (\mathbf{y} + \mathbf{x}) + \mathbf{z} = \vec{0} + \mathbf{z} = \mathbf{z}.$$

Portanto, o oposto de cada vetor é único.

[c] Como $0\mathbf{x} + 0\mathbf{x} = (0 + 0)\mathbf{x} = 0\mathbf{x}$, o cancelamento aditivo fornece $0\mathbf{x} = \vec{0}$. Analogamente, $\lambda\vec{0} + \lambda\vec{0} = \lambda(\vec{0} + \vec{0}) = \lambda\vec{0}$, de modo que $\lambda\vec{0} = \vec{0}$.

[d] Temos $\mathbf{x} + (-1)\mathbf{x} = 1\mathbf{x} + (-1)\mathbf{x} = (1 - 1)\mathbf{x} = 0\mathbf{x} = \vec{0}$. Pela unicidade do oposto, $(-1)\mathbf{x} = -\mathbf{x}$.

[e] Se $\lambda\mathbf{x} = \vec{0}$ e $\lambda \neq 0$, multiplicar a igualdade por λ^{-1} dá $\mathbf{x} = (\lambda^{-1}\lambda)\mathbf{x} = \lambda^{-1}\vec{0} = \vec{0}$.

[f] Somando $-\mathbf{y}$ aos dois lados de $\mathbf{x} + \mathbf{y} = \mathbf{z} + \mathbf{y}$, obtemos $\mathbf{x} = \mathbf{z}$. ■

Observação 2.3. A expressão $\lambda_1\mathbf{x}_1 + \ldots + \lambda_n\mathbf{x}_n = \sum \lambda_i\mathbf{x}_i$ é definida unicamente para qualquer $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ e $\mathbf{x}_1, \ldots, \mathbf{x}_n \in V$: devido à associatividade da adição, não é necessário inserir parênteses indicando ordem para o cálculo de somas múltiplas. Analogamente, a expressão $\lambda_1\lambda_2 \ldots \lambda_n\mathbf{x}$ está bem definida.

Uma expressão da forma $\sum_{i=1}^{n} a_i\mathbf{x}_i$ é uma **combinação linear** dos vetores $\mathbf{x}_1, \ldots, \mathbf{x}_n$. Os escalares a_i são os **coeficientes** da combinação.

Atenção
Uma combinação linear é sempre uma soma **finita**. Séries de vetores exigem uma noção adicional de convergência.

Exemplo 2.4 — Espaço de dimensão zero. O conjunto $V = \{\vec{0}\}$ torna-se um espaço vetorial quando definimos $\lambda\vec{0} = \vec{0}$ para todo $\lambda \in \mathbb{K}$ e $\vec{0} + \vec{0} = \vec{0}$.

◁

Espaços de dimensão zero sobre corpos distintos são estruturas distintas; por isso o corpo de escalares faz parte da especificação do espaço vetorial.

Exemplo 2.5 — Corpo. O conjunto $V = \mathbb{K}$, com a adição e a multiplicação do próprio corpo, é um espaço vetorial sobre $\mathbb{K}$.

De maneira mais geral, dados um corpo $\mathbb{K}$ e um subcorpo $\mathbb{F}$ deste, o corpo $\mathbb{K}$ pode ser visto como um espaço vetorial sobre $\mathbb{F}$. Por exemplo, o corpo de números complexos $\mathbb{C}$ é um espaço vetorial sobre o corpo dos números reais $\mathbb{R}$, e este é um espaço vetorial sobre o corpo dos números racionais $\mathbb{Q}$.

◁

Exemplo 2.6 — Espaço de Coordenadas, $\mathbb{K}^n$. Para cada $n \in \mathbb{N}$, denotamos por $\mathbb{K}^n = \mathscr{M}_{n,1}(\mathbb{K})$ o espaço dos vetores coluna com n entradas em $\mathbb{K}$. Assim,

$$\mathbb{K}^n = \left\{ \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \mid x_i \in \mathbb{K} \right\}.$$

A adição e a multiplicação por escalares são definidas componente a componente:

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} := \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix},$$

$$\lambda \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} := \begin{bmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{bmatrix}.$$

Em particular, quando $n = 1$, temos que $\mathbb{K}$ é um espaço vetorial sobre $\mathbb{K}$.

◁

Se A e B são dois conjuntos, vamos denotar o conjunto de funções de A para B por B^A. Assim, $B^A = \{f : A \to B \mid f \text{ é uma função }\}$.

No Exemplo 2.6, $\mathbb{K}^n$ pode ser identificado com o conjunto das funções de $\{1, \dots, n\}$ em $\mathbb{K}$: ao vetor coluna de entradas $x_1, \dots, x_n$ associamos a função $i \mapsto x_i$. Essa observação sugere a seguinte generalização.

Exemplo 2.7 — Espaços de Funções. Sejam S um conjunto arbitrário e $\mathbb{K}(S) = \mathbb{K}^S$ o conjunto de funções em S com valores no corpo $\mathbb{K}$. Se $f : S \to \mathbb{K}$ é uma função, então $f(s)$ indica o valor de f no elemento $s \in S$.

A adição e multiplicação de funções por um escalar são definidas pontualmente:

- $(f + g)(s) := f(s) + g(s)$ para todos os $s \in S$,
- $(af)(s) := af(s)$ para todos os $a \in \mathbb{K}$ e $s \in S$.

Como já observamos, se $S = \{1, \dots, n\}$, então $\mathbb{K}(S)$ pode ser identificado com $\mathbb{K}^n$: a função f corresponde ao vetor coluna cujas entradas são $f(1), \dots, f(n)$. As operações são compatíveis com essa identificação.

A todo elemento $s \in S$ podemos associar uma função delta δ_s centrada em $\{s\}$ definida como

$$\delta_s(s) = 1 \quad \text{e} \quad \delta_s(t) = 0 \text{ se } t \neq s.$$

Quando $S = \{1, \dots, n\}$, usaremos também a notação do delta de Kronecker, $\delta_i(k) = \delta_{ik}$.

Se o conjunto S for finito, uma função $f \in \mathbb{K}(S)$ poderá ser representada unicamente por uma combinação linear de funções delta:

$$f = \sum_{s \in S} f(s)\delta_s.$$

A igualdade pode ser verificada ponto a ponto. Reciprocamente, se $f = \sum_{s\in S} a_s\delta_s$, avaliar ambos os lados em s fornece $f(s) = a_s$.

Se S for infinito, essa representação vale apenas para as funções de suporte finito, pois combinações lineares contêm somente um número finito de parcelas. Algumas somas infinitas podem ser definidas quando o espaço possui uma topologia ou uma norma; voltaremos a esse tema na Seção 6.7.

◁

Exemplo 2.8 — Funções Vetoriais. Seja V um espaço vetorial sobre $\mathbb{K}$ e A um conjunto arbitrário. Então, o conjunto V^A de todas as funções de A em V é um espaço vetorial sobre $\mathbb{K}$ quando definimos as operações pontualmente. Se $f, g \in V^A$ e $\lambda \in \mathbb{K}$, então

$$(f + g)(a) = f(a) + g(a), \qquad (\lambda f)(a) = \lambda f(a), \quad a \in A.$$

◁

Se A for finito, com n elementos, identificaremos V^A com V^n depois de escolher uma ordem em A. Em particular, para $V = \mathbb{K}$, recuperamos o espaço de coordenadas $\mathbb{K}^n$ do Exemplo 2.6.

Exemplo 2.9 — Matrizes. Denotaremos por $\mathscr{M}_{m,n}(\mathbb{K})$ o conjunto das matrizes $m \times n$ com entradas em $\mathbb{K}$. A adição usual, $(a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$, e a multiplicação por escalares, $\lambda(a_{ij}) = (\lambda a_{ij})$, fazem de $\mathscr{M}_{m,n}(\mathbb{K})$ um espaço vetorial sobre $\mathbb{K}$.

◁

Exemplo 2.10 — Vetores linha e coluna. Pela convenção adotada neste livro,

$$\mathbb{K}^n = \mathscr{M}_{n,1}(\mathbb{K})$$

é o espaço dos vetores coluna. Os vetores linha pertencem a $\mathscr{M}_{1,n}(\mathbb{K})$. A transposta de um vetor coluna é um vetor linha, e vice-versa; os dois espaços são naturalmente isomorfos, mas não serão identificados.

◁

Exemplo 2.11 — Polinômios. Seja $\mathbb{K}[x]$ o conjunto dos polinômios em uma variável x com coeficientes em $\mathbb{K}$. Um elemento típico é uma soma finita $a_n x^n + \cdots + a_0$, com $n \in \mathbb{N} \cup \{0\}$ e $a_0, \dots, a_n \in \mathbb{K}$. A adição de polinômios e a multiplicação por escalares fazem de $\mathbb{K}[x]$ um espaço vetorial sobre $\mathbb{K}$.

◁

A análise fornece outros exemplos de espaços vetoriais.

Exemplo 2.12 — Sequências. Uma sequência com valores em $\mathbb{K}$ é uma função $a : \mathbb{N} \to \mathbb{K}$; escrevemos usualmente $a_n := a(n)$. O conjunto $\mathrm{Seq}(\mathbb{K})$ de todas as sequências infinitas com valores num corpo $\mathbb{K}$ é um espaço vetorial com as operações definidas termo a termo:

$$(s_n) + (t_n) := (s_n + t_n) \quad \text{e} \tag{2.1}$$

$$a(s_n) := (as_n) \tag{2.2}$$

Também denotaremos o espaço $\mathrm{Seq}(\mathbb{K})$ como $\mathbb{K}^\infty$ e veremos uma sequência como uma lista ordenada. Assim, veremos (a_n) como a lista ordenada $(a_1, a_2, \ldots)$ e escreveremos

$$\mathbb{K}^\infty = \mathrm{Seq}(\mathbb{K}) = \{(a_1, a_2, \ldots) \mid a_i \in \mathbb{K} \text{ para todo } i \in \mathbb{N}\},$$

com as operações definidas coordenada a coordenada.

◁

Exemplo 2.13 — Funções Contínuas e Diferenciáveis. Seja I um intervalo fechado. Vamos denotar por $C(I)$ as funções reais contínuas em I. Denotaremos por $C^k(I)$ as funções de $C(I)$ que são k-vezes diferenciáveis no interior de I. Então $C(I) \supseteq C^1(I) \supseteq C^2(I) \supseteq \cdots$. Esses conjuntos são espaços vetoriais sobre $\mathbb{R}$ quando dotados da adição usual pontual $(f+g)(x) = f(x)+g(x), x \in I$ e multiplicação escalar $(\lambda f)(x) = \lambda(f(x))$.

◁

Exemplo 2.14 — Funções integráveis. Seja $B = [a_1, b_1] \times \cdots \times [a_n, b_n] \subseteq \mathbb{R}^n$ um retângulo fechado. O conjunto $\mathscr{R}(B)$ das funções reais Riemann-integráveis em B é um espaço vetorial sobre $\mathbb{R}$, com as operações pontuais definidas como no Exemplo 2.13.

◁

Observação 2.15. Geralmente, é muito conveniente, mas não totalmente consistente, denotar o elemento zero e a adição em $\mathbb{K}$ e V pelos mesmos símbolos.

Apresentaremos a seguir dois exemplos nos quais uma notação diferente é preferível.

Exemplo 2.16. Seja $L = \{x \in \mathbb{R} \mid x > 0\}$. Consideramos a multiplicação usual como a "adição" de vetores e definimos a multiplicação de $x \in L$ pelo escalar $a \in \mathbb{R}$ como x^a. As leis das potências verificam os axiomas da Definição 2.1, que nessa notação assumem uma forma pouco habitual:

- ☐ o vetor zero em L é o elemento 1, ou seja, $\vec{0} = 1$.
- ☐ a identidade $1\mathbf{x} = \mathbf{x}$ adquire a forma $x^1 = x$
- ☐ a identidade $a(b\mathbf{x}) = (ab)\mathbf{x}$ adquire a forma $(x^b)^a = x^{ba}$
- ☐ a identidade $(a+b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$ torna-se $x^{a+b} = x^a x^b$.

◁

Exemplo 2.17. Seja V um espaço vetorial sobre o corpo dos números complexos $\mathbb{C}$. Defina um novo espaço vetorial $\overline{V}$ com o mesmo grupo aditivo V, mas com uma multiplicação por escalar diferente:

$$(a, \mathbf{x}) \mapsto \overline{a}\mathbf{x},$$

onde $\overline{a}$ denota o conjugado complexo de a. Das fórmulas $\overline{a+b} = \overline{a} + \overline{b}$ e $\overline{ab} = \overline{a}\overline{b}$ segue sem maiores dificuldades que $\overline{V}$ é um espaço vetorial.

◁

Exercícios

Ex. 2.1 — Seja V um espaço vetorial sobre um corpo $\mathbb{K}$. Mostre que:

1. O vetor nulo $\vec{0}$ é único.
2. O vetor oposto $-\mathbf{v}$ a cada vetor $\mathbf{v} \in V$ é único.
3. Prove que $-1 \cdot \mathbf{v} = -\mathbf{v}$ e conclua que $-(-\mathbf{v}) = \mathbf{v}$ para todo $\mathbf{v} \in V$.
4. Prove que $\alpha\mathbf{v} = \vec{0}$ se, e somente se, $\alpha = 0$ ou $\mathbf{v} = \vec{0}$.
5. Prove que $\alpha\mathbf{v} = \mathbf{v}$ se, e somente se, $\alpha = 1$ ou $\mathbf{v} = \vec{0}$.
6. Mostre que se $\mathbf{x} + \mathbf{y} = \mathbf{z} + \mathbf{y}$ então $\mathbf{x} = \mathbf{z}$.
7. Mostre que o axioma de comutatividade da soma pode ser deduzido dos demais axiomas da definição de espaço vetorial.

Ex. 2.2 — Determine se os seguintes conjuntos são espaços vetoriais sobre o corpo $\mathbb{K}$ especificado em cada caso.

1. $\mathbb{R}^3$ sobre $\mathbb{K} = \mathbb{R}$ com as operações

$$\begin{aligned}(x_1, x_2, x_3) + (y_1, y_2, y_3) &= (x_1 + y_1, x_2 + y_2, x_3 + y_3)\\ \alpha \cdot (x_1, x_2, x_3) &= (\alpha \cdot x_1, x_2, x_3)\end{aligned}$$

2. O conjunto $V = \{(a, b) \in \mathbb{R}^2 \,:\, a, b > 0\}$ sobre $\mathbb{K} = \mathbb{R}$ com as operações

$$\begin{aligned}(a, b) + (c, d) &= (ac, bd)\\ \alpha \cdot (a, b) &= (a^\alpha, b^\alpha),\ \alpha \in \mathbb{R}\end{aligned}$$

3. O conjunto $\mathbb{Q}(\sqrt{2})$ dos elementos $a + b\sqrt{2}$ com $a, b \in \mathbb{Q}$ sobre $\mathbb{K} = \mathbb{Q}$ com as operações:

$$\begin{aligned}(a + b\sqrt{2}) + (c + d\sqrt{2}) &= (a + c) + (b + d)\sqrt{2}\\ \alpha \cdot (a + b\sqrt{2}) &= \alpha \cdot a + \alpha \cdot b\sqrt{2},\ \alpha \in \mathbb{Q}\end{aligned}$$

Ex. 2.3 — Mostre que $\mathbb{N}$ pode receber uma estrutura de espaço vetorial sobre $\mathbb{Q}$. (Dica: transporte as operações de $\mathbb{Q}$ por uma bijeção $\mathbb{N} \to \mathbb{Q}$.)

Ex. 2.4 — Seja V um espaço vetorial. Use os axiomas de espaço vetorial para provar que se $\mathbf{v} \in V$ e $n \in \mathbb{N}$ então

$$n\mathbf{v} = \underbrace{\mathbf{v} + \ldots + \mathbf{v}}_{n \text{ parcelas}}.$$

Ex. 2.5 — Sejam $\mathbf{u}, \mathbf{v}$ vetores não nulos de V. Prove que $\mathbf{v}$ é múltiplo de $\mathbf{u}$ se, e somente se, $\mathbf{u}$ é múltiplo de $\mathbf{v}$.

Ex. 2.6 — Prove detalhadamente que o conjunto de funções $\mathbb{K}(S) = \mathbb{K}^S$ é um espaço vetorial quando a adição e a multiplicação por escalares são definidas pontualmente:

- $(f + g)(s) \coloneqq f(s) + g(s)$ para todos os $s \in S$,
- $(af)(s) \coloneqq af(s)$ para todos os $a \in \mathbb{K}$ e $s \in S$.

Ex. 2.7 — Se S é um conjunto finito com n elementos, quantos elementos possui o espaço vetorial $\mathbb{Z}_2^S$?

Ex. 2.8 — Suponha que $\mathrm{char}(\mathbb{K}) \neq 2$. Defina a média de dois vetores por $\mathbf{u} \star \mathbf{v} = \frac{1}{2}\mathbf{u} + \frac{1}{2}\mathbf{v}$. Prove que $(\mathbf{u} \star \mathbf{v}) \star \mathbf{w} = \mathbf{u} \star (\mathbf{v} \star \mathbf{w})$ se, e somente se, $\mathbf{u} = \mathbf{w}$.

Ex. 2.9 — Dados os espaços vetoriais V_1, V_2, considere o conjunto $V = V_1 \times V_2$ (produto cartesiano de V_1 por V_2), cujos elementos são os pares ordenados $v = (v_1, v_2)$, com $v_1 \in V_1$ e $v_2 \in V_2$. Defina operações que tornem V um espaço vetorial. Verifique a validade de cada um dos axiomas.

Ex. 2.10 — Em $\mathbb{R}^2$ mantenhamos a definição de produto αv de um número por um vetor mas modifiquemos, de três maneiras diferentes, a definição de soma $u + v$ de vetores $u = (x, y)$ e $v = (x', y')$. Em cada tentativa, dizer quais axiomas de espaço vetorial continuam válidos e quais são violados:

1. $u + v = (x + y', x' + y)$,
2. $u + v = (xx', yy')$,
3. $u + v = (3x + 3x', 5x + 5x')$.

Ex. 2.11 — Seja $\mathbb{K}$ um corpo finito com q elementos. Quantos elementos possui o espaço vetorial $\mathbb{K}^n$?

O teste de subespaço

Para reconhecer um subespaço, não é preciso verificar novamente todos os axiomas de espaço vetorial. Basta confirmar que o conjunto contém o vetor zero e é fechado por soma e por multiplicação por escalares. É por isso que conjuntos de soluções de sistemas homogêneos aparecem naturalmente como subespaços.

2.2 Subespaços

Em muitas estruturas algébricas podemos reconhecer subconjuntos que herdam as operações da estrutura original. Os espaços vetoriais não são exceção. Fixemos um espaço vetorial V sobre $\mathbb{K}$.

Definição 2.18 — Subespaço Vetorial. Um subconjunto não vazio W de V é um **subespaço vetorial** de V se W for um espaço vetorial com as mesmas operações de adição de vetores e multiplicação escalar que V.

A proposição seguinte fornece um critério para reconhecer subespaços.

Proposição 2.19.
Um subconjunto não vazio W de V é um subespaço vetorial se, e somente se, for fechado em relação à adição de vetores e à multiplicação por escalares de V.

Demonstração. Se W é um subespaço, o fechamento decorre dos axiomas de espaço vetorial. Reciprocamente, suponha que W seja não vazio e fechado por soma e multiplicação por escalares. Escolha $\mathbf{w} \in W$. Como $\vec{0} = 0\mathbf{w} \in W$ e $-\mathbf{w} = (-1)\mathbf{w} \in W$, o vetor nulo e os opostos pertencem a W. Todos os demais axiomas são herdados de V; portanto, W é um espaço vetorial com as operações de V. ■

Exemplo 2.20. O conjunto dos polinômios de grau menor ou igual a n, juntamente com o polinômio nulo, é um subespaço de $\mathbb{K}[x]$. Esse espaço será denotado por $\mathbb{K}_n[x]$.

◁

Exemplos 2.21. Os conjuntos $C([a,b])$, $C^k([a,b])$ e $\mathscr{R}([a,b])$ são subespaços de $\mathbb{R}^{[a,b]}$.

◁

Exemplo 2.22. Vamos apresentar alguns exemplos de subespaços do espaço de sequências.

O conjunto c_0 das sequências complexas que convergem para 0 é um espaço vetorial, assim como o conjunto ℓ^∞ das sequências complexas limitadas. Além disso, se $p \geq 1$ é um número real, o conjunto ℓ^p das sequências complexas (s_n) para as quais

$$\sum_{n=1}^{\infty} |s_n|^p < \infty$$

é um espaço vetorial com operações definidas componente a componente. O fechamento em relação à adição decorre da desigualdade de Minkowski:

$$\left(\sum_{n=1}^{\infty} |s_n + t_n|^p\right)^{1/p} \leq \left(\sum_{n=1}^{\infty} |s_n|^p\right)^{1/p} + \left(\sum_{n=1}^{\infty} |t_n|^p\right)^{1/p}$$

◁

Exemplo 2.23. O conjunto

$$(\mathbb{K}^\infty)_0 = \{(a_i)_{i\geq 1} \in \mathbb{K}^\infty \mid a_i \neq 0 \text{ para apenas finitos índices } i\}$$

é um subespaço de $\mathbb{K}^\infty$.

◁

Exemplo 2.24 — Sistema Linear. Seja $A \in \mathscr{M}_{m,n}(\mathbb{K})$. O conjunto das soluções $X \in \mathbb{K}^n$ do sistema homogêneo $AX = \vec{0}$ é um subespaço de $\mathbb{K}^n$. De fato, se $AX = AY = \vec{0}$ e $\lambda \in \mathbb{K}$, então $A(\lambda X + Y) = \lambda AX + AY = \vec{0}$.

Essa distributividade pode ser verificada diretamente nas entradas. Se $B, C \in \mathscr{M}_{n,p}(\mathbb{K})$, então

$$A(\lambda B + C) = \lambda(AB) + AC$$

para todo $\lambda \in \mathbb{K}$; com efeito,

$$\begin{aligned}
[A(\lambda B + C)]_{ij} &= \sum_k A_{ik}(\lambda B + C)_{kj} \\
&= \sum_k (\lambda A_{ik} B_{kj} + A_{ik} C_{kj}) \\
&= \lambda \sum_k A_{ik} B_{kj} + \sum_k A_{ik} C_{kj} \\
&= \lambda (AB)_{ij} + (AC)_{ij} \\
&= [\lambda(AB) + AC]_{ij}.
\end{aligned}$$

◁

Exemplo 2.25. Considere o seguinte sistema de equações diferenciais lineares:

$$x_1' = a_{11}x_1 + \cdots + a_{1n}x_n \tag{2.3}$$
$$\vdots \tag{2.4}$$
$$x_n' = a_{n1}x_1 + \cdots + a_{nn}x_n \tag{2.5}$$

com $x_1, \ldots, x_n \in C^1(I)$, onde I é um intervalo aberto em $\mathbb{R}$, x_i' denota a derivada de x_i e $a_{ij} \in \mathbb{R}$. A matriz $A = (a_{ij}) \in \mathscr{M}_{n,n}(\mathbb{R})$ é a **matriz do sistema**. Defina $\mathbf{x} = (x_1, \ldots, x_n)^{\mathrm{t}}$. Com essa notação, o sistema assume a forma

$$\mathbf{x}' = A\mathbf{x} \text{ com } \mathbf{x} \in C^1(I)^n.$$

O conjunto de soluções é $V = \{\mathbf{x} \in C^1(I)^n \mid \mathbf{x}' = A\mathbf{x}\}$. A linearidade da derivação e da multiplicação por A mostra que V é um subespaço de $C^1(I)^n$.

◁

Uma coleção $\mathcal{S} = \{W_i \mid i \in I\}$ de subespaços de V permite construir novos subespaços de duas maneiras naturais.

A interseção $\bigcap_{i\in I} W_i$ é um subespaço de V, pois todas as condições do teste de subespaço são preservadas por interseções.

Se, para cada $i, j \in I$, existir $k \in I$ tal que $W_i \cup W_j \subseteq W_k$, então $\bigcup_{i\in I} W_i$ também é um subespaço: dois de seus vetores pertencem simultaneamente a algum W_k, onde sua soma e seus múltiplos escalares permanecem.

Em geral, porém, a união de dois subespaços não é um subespaço. De fato, $W_1 \cup W_2$ é um subespaço se, e somente se, $W_1 \subseteq W_2$ ou $W_2 \subseteq W_1$. A verificação é proposta como exercício.

> **Teorema 2.26.**
> *Seja V um espaço vetorial sobre um corpo infinito $\mathbb{K}$. Então V não pode ser a união de um número finito de subespaços próprios.*

Demonstração. Suponha, por contradição, que $V = W_1 \cup \cdots \cup W_n$, com todos os W_i próprios, e escolha tal cobertura com n mínimo. Ela é irredundante; em particular, existe

$$\mathbf{x} \in W_1 \setminus \bigcup_{i=2}^{n} W_i.$$

Como $W_1 \neq V$, escolha também $\mathbf{y} \in V \setminus W_1$. A reta afim $L = \{\mathbf{x} + \lambda\mathbf{y} \mid \lambda \in \mathbb{K}\}$ é infinita: se duas de suas expressões fossem iguais, teríamos $(\lambda - \mu)\mathbf{y} = \vec{0}$ e, portanto, $\lambda = \mu$.

Como L está coberta pelos n subespaços, algum $L \cap W_j$ é infinito. Escolha dois de seus pontos distintos, $\mathbf{x} + \lambda\mathbf{y}$ e $\mathbf{x} + \mu\mathbf{y}$. A diferença mostra que $(\lambda - \mu)\mathbf{y} \in W_j$ e, como $\lambda \neq \mu$, temos $\mathbf{y} \in W_j$. Subtraindo $\lambda\mathbf{y}$ do primeiro ponto, obtemos $\mathbf{x} \in W_j$. Se $j = 1$, isso contradiz $\mathbf{y} \notin W_1$; se $j \neq 1$, contradiz a escolha de $\mathbf{x}$. Logo tal cobertura não existe.

■

Exemplo 2.27. Se $\mathbb{K}$ é finito, então o teorema anterior é falso em geral. Por exemplo, seja $V = (\mathbb{F}_2)^2$. Então $V = W_1 \cup W_2 \cup W_3$, em que $W_1 = \{(0,0),(1,0)\}$, $W_2 = \{(0,0),(0,1)\}$ e $W_3 = \{(0,0),(1,1)\}$.

◁

> **Definição 2.28** — Subespaço Gerado. Dado $S \subseteq V$, o **subespaço gerado** por S é o menor subespaço de V que contém S. Explicitamente,
>
> $$\langle S \rangle := \bigcap \{W \mid W \text{ é subespaço de } V \text{ e } S \subseteq W\}.$$

Denotaremos por $\mathscr{P}(V)$ o conjunto dos subconjuntos de V e por $\mathscr{S}(V)$ o conjunto de seus subespaços. A aplicação $\langle \cdot \rangle : \mathscr{P}(V) \to \mathscr{S}(V)$ envia S em $\langle S \rangle$. Ela é sobrejetiva, pois sua restrição a $\mathscr{S}(V)$ é a identidade.

> **Definição 2.29** — Combinação Linear. Seja $B = \{\mathbf{v}_i \mid i \in I\} \subseteq V$. Um vetor $\mathbf{v} \in V$ é uma combinação linear dos vetores de B se existe uma família de escalares $(c_i)_{i \in I}$, com apenas um número finito de termos não nulos, tal que
>
> $$\vec{v} = \sum_{i \in I} c_i \vec{v}_i.$$

Observação 2.30. Se escolhermos todos $c_i = 0$, obteremos

$$\vec{0} = \sum c_i \mathbf{v}_i.$$

Esta é a combinação linear trivial dos vetores em B. Qualquer outra combinação linear é não trivial.

Se $B = \emptyset$, a única combinação linear é a soma vazia, cujo valor é, por convenção, o vetor nulo de V.

Teorema 2.31.
A função $\langle \cdot \rangle : \mathscr{P}(V) \to \mathscr{S}(V)$ possui as seguintes propriedades

[a] *Para $S \in \mathscr{P}(V), \langle S \rangle$ é o subespaço de V de todas as combinações lineares finitas de vetores de S. Assim,*

$$\langle S \rangle = \left\{ \sum_{i=1}^{n} x_i \mathbf{x}_i \mid x_i \in \mathbb{K}, \mathbf{x}_i \in S, n \geq 0 \right\}$$

[b] *Se $S_1 \subseteq S_2$, $\langle S_1 \rangle \subseteq \langle S_2 \rangle$.*

[c] *Se $\mathbf{x} \in \langle S \rangle$, existe um subconjunto finito $S' \subseteq S$ de modo que $\mathbf{x} \in \langle S' \rangle$.*

[d] *$S \subseteq \langle S \rangle$ para todos os $S \in \mathscr{P}(V)$.*

[e] *Para cada $S \in \mathscr{P}(V), \langle \langle S \rangle \rangle = \langle S \rangle$.*

[f] *Se $\mathbf{y} \in \langle S \cup \{\mathbf{x}\} \rangle$ e $\mathbf{y} \notin \langle S \rangle$, então $\mathbf{x} \in \langle S \cup \{\mathbf{y}\} \rangle$. Aqui $\mathbf{x}, \mathbf{y} \in V$ e $S \in \mathscr{P}(V)$.*

Demonstração. [a] Seja L o conjunto das combinações lineares finitas de vetores de S. Como $\langle S \rangle$ é um subespaço que contém S, temos $L \subseteq \langle S \rangle$. Por outro lado, L é um subespaço que contém S; pela minimalidade de $\langle S \rangle$, segue que $\langle S \rangle \subseteq L$.

[b] Se $S_1 \subseteq S_2$, então $\langle S_2 \rangle$ é um subespaço que contém S_1. Pela minimalidade de $\langle S_1 \rangle$, $\langle S_1 \rangle \subseteq \langle S_2 \rangle$.

[c] Pelo item [a], cada vetor de $\langle S \rangle$ é uma combinação linear de um número finito de vetores de S; basta reunir esses vetores em S'.

[d] Cada $\mathbf{x} \in S$ é a combinação linear $1\mathbf{x}$, logo $S \subseteq \langle S \rangle$.

[e] O conjunto $\langle S \rangle$ já é subespaço; portanto, tomar seu gerado não o altera.

[f] Se $\mathbf{y} \in \langle S \cup \{\mathbf{x}\} \rangle \setminus \langle S \rangle$, podemos escrever

$$\mathbf{y} = a\mathbf{x} + a_1\mathbf{x}_1 + \cdots + a_n\mathbf{x}_n, \qquad \mathbf{x}_1, \ldots, \mathbf{x}_n \in S.$$

O coeficiente a não pode ser nulo, pois isso colocaria $\mathbf{y}$ em $\langle S \rangle$. Assim,

$$\mathbf{x} = a^{-1}\mathbf{y} - a^{-1}a_1\mathbf{x}_1 - \cdots - a^{-1}a_n\mathbf{x}_n \in \langle S \cup \{\mathbf{y}\} \rangle.$$

■

Exercícios

Ex. 2.12 — Seja W um subconjunto não vazio de V. Prove que, se W for fechado em relação à adição e à multiplicação por escalares, então W é um subespaço de V.

Ex. 2.13 — Para os conjuntos seguintes, determine se os conjuntos dados são espaços vetoriais reais, se a adição e a multiplicação forem as usuais. Para aqueles que não forem diga quais axiomas de espaços vetoriais não são satisfeitos.

1. O conjunto dos polinômios de grau menor ou igual a n
2. O conjunto de todas as funções reais tais que $f(0) = f(1)$
3. O conjunto das funções tais que $f(0) = 1 + f(1)$
4. O conjunto das funções reais crescentes.
5. O conjunto das funções reais pares.
6. O conjunto das funções reais ímpares.
7. O conjunto das funções contínuas em $[0, 1]$ tais que $\int_0^1 f(x)dx = 0$
8. O conjunto das funções contínuas em $[0, 1]$ tais que $\int_0^1 f(x)dx \geq 0$
9. O conjunto dos vetores (x, y, z) em $\mathbb{R}^3$ tais que $x = 0$ ou $y = 0$
10. O conjunto das matrizes 2×2 cujo traço é zero
11. O conjunto das matrizes 2×2 cujo determinante é zero
12. O conjunto das matrizes 2×2 simétricas, isto é, tais que $A = A^{\mathrm{t}}$.
13. O conjunto dos vetores (x, y, z) que satisfaz a equação linear $a_1x+a_2y+a_3z = 0$

Ex. 2.14 — Considere o subconjunto S de $C^n(\mathbb{R})$ formado pelas funções que são soluções da equação diferencial linear homogênea de ordem n com coeficientes constantes

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \cdots + a_1y'(t) + a_0y(t) = 0$$

onde $a_0, a_1, \ldots, a_{n-1} \in \mathbb{R}$. Mostre que S é um subespaço vetorial de $C^n(\mathbb{R})$.

Ex. 2.15 — Seja $\mathscr{F}(X, \mathbb{K})$ o espaço das funções, onde $\mathbb{K}$ é um corpo. No caso particular em que $X = \mathbb{N}$ o espaço é denominado **espaço de sequências** e vamos denotá-lo por $\mathbb{K}^{\mathbb{N}}$. Um elemento $f \in \mathbb{K}^{\mathbb{N}}$ é uma função $f : \mathbb{N} \to \mathbb{K}$ dada por $n \mapsto x_n \in \mathbb{K}$ e será denotado por $f = (x_n)_{n\in\mathbb{N}}$. Quais dos seguintes conjuntos são subespaços de $\mathbb{K}^{\mathbb{N}}$?

1. O conjunto das sequências com apenas um número finito de coordenadas diferentes de zero.
2. Nenhuma coordenada igual a 1.

Nos próximos itens $\mathbb{K} = \mathbb{R}$

3. O conjunto das sequências de Cauchy, isto é, das sequências tais que, dado $\epsilon > 0$ existe $N > 0$ tal que $|x_n - x_m| < \epsilon$ para $n, m > N$.

4. As sequências tais que $\sum_{n=1}^{\infty} |x_n| < \infty$.

5. As sequências limitadas, isto é, as sequências $(x_n)_{n\in\mathbb{N}}$ para as quais existe $M > 0$ tal que $|x_n| \leq M$ para todo $n \in \mathbb{N}$.

Ex. 2.16 — Seja S um subespaço de V e seja $v \in V$. O conjunto $v + S = \{v + s \mid s \in S\}$ é um subespaço afim de V.

1. Quando um subespaço afim de V é subespaço de V?

2. Mostre que dois subespaços afins $x + S$ e $y + S$ ou são iguais ou são disjuntos.

Ex. 2.17 — Prove que

1. Se $S_1 \subseteq S_2$, $\langle S_1 \rangle \subseteq \langle S_2 \rangle$.

2. Se $\mathbf{x} \in \langle S \rangle$, existe um subconjunto finito $S' \subseteq S$ de modo que $\mathbf{x} \in \langle S' \rangle$.

3. $S \subseteq \langle S \rangle$ para todos os $S \in \mathscr{P}(V)$.

4. Para cada $S \in \mathscr{P}(V)$, $\langle \langle S \rangle \rangle = \langle S \rangle$.

Ex. 2.18 — Seja $X \subseteq V$ um subconjunto de um espaço vetorial sobre $\mathbb{K}$ V. Mostre que $\langle X \rangle = \{\sum_{v\in F} \alpha_v \cdot v : F \subseteq X \text{ é um subconjunto finito e } \alpha_v \in \mathbb{K},\ \forall v \in F\}$.

Ex. 2.19 — Dê um contraexemplo que mostre que a união de subespaços não é necessariamente um subespaço.

Ex. 2.20 — Mostre que, se W_1 e W_2 são subespaços de V, então $W_1 \cup W_2$ é um subespaço se, e somente se, $W_1 \subseteq W_2$ ou $W_2 \subseteq W_1$.

2.3 Bases e Dimensão

Definição 2.32 — Dependência e Independência Linear. Suponha que V seja um espaço vetorial arbitrário sobre um corpo $\mathbb{K}$ e seja S um subconjunto de V.

- S é dito **linearmente dependente** sobre $\mathbb{K}$ se existirem vetores distintos $\mathbf{x}_1, \ldots, \mathbf{x}_n \in S$ e escalares $a_1, \ldots, a_n \in \mathbb{K}$, não todos nulos, tais que $a_1\mathbf{x}_1 + \cdots + a_n\mathbf{x}_n = \vec{0}$.
- S é dito **linearmente independente** (sobre $\mathbb{K}$) se S não for linearmente

Gerar sem redundância
Uma base cumpre simultaneamente duas tarefas: seus vetores geram todo o espaço e nenhum deles é dispensável. A condição de geração garante existência das coordenadas; a independência linear garante sua unicidade.

dependente.

Portanto, se S for linearmente independente, toda relação $\sum_{i=1}^{n} a_i \mathbf{x}_i = \vec{0}$, com $\mathbf{x}_1, \dots, \mathbf{x}_n$ distintos em S, terá todos os coeficientes nulos.

Pela definição, o conjunto vazio é linearmente independente por vacuidade. Omitiremos a expressão "sobre $\mathbb{K}$" quando o corpo estiver claro no contexto. Se houver mais de um corpo envolvido, porém, a dependência pode mudar, como mostra o exemplo seguinte.

Exemplo 2.33. Considere $V = \mathbb{R}$ como espaço vetorial tanto sobre $\mathbb{Q}$ quanto sobre $\mathbb{R}$. O conjunto $S = \{1, \sqrt{2}\}$ é linearmente independente sobre $\mathbb{Q}$, pois $a + b\sqrt{2} = 0$, com $a, b \in \mathbb{Q}$, implica $a = b = 0$. Sobre $\mathbb{R}$, porém, ele é linearmente dependente, pois $\sqrt{2} \cdot 1 - 1 \cdot \sqrt{2} = 0$.

◁

Definição 2.34 — Base. Um subconjunto S de V é dito **base** de V se S for linearmente independente sobre $\mathbb{K}$ e $\langle S \rangle = V$.

Atenção
Uma base pode ser infinita, mas cada vetor usa apenas um número **finito** de seus elementos.

Se S é uma base de V, todo vetor não nulo $\mathbf{x} \in V$ possui uma expressão única $\mathbf{x} = x_1\mathbf{x}_1 + \cdots + x_n\mathbf{x}_n$, na qual $\mathbf{x}_1, \dots, \mathbf{x}_n$ são elementos distintos de S e os coeficientes $x_1, \dots, x_n$ são não nulos.

Notação 2.35. As bases serão denotadas por $\underline{\mathrm{B}}, \underline{\mathrm{x}}, \dots$.

Vejamos alguns exemplos de bases.

Exemplo 2.36. O conjunto vazio $\emptyset$ é uma base para o subespaço nulo (0) de qualquer espaço vetorial V. Se considerarmos um corpo $\mathbb{K}$ como um espaço vetorial sobre si mesmo, qualquer elemento diferente de zero $\mathbf{x}$ de $\mathbb{K}$ é uma base de $\mathbb{K}$.

◁

Exemplo 2.37. Seja $V = \mathbb{K}^n$. Para $i = 1, \dots, n$, denote por $\mathbf{e}_i$ o vetor coluna cuja i-ésima entrada é 1 e cujas demais entradas são nulas. Como

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n,$$

o conjunto $\underline{\mathrm{B}} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ é uma base de $\mathbb{K}^n$, denominada **base canônica**.

◁

Exemplo 2.38. Seja $V = (\mathbb{K}^\infty)_0$. Para cada $i \in \mathbb{N}$, denote por $\mathbf{e}_i$ a sequência cuja i-ésima entrada é 1 e cujas demais entradas são nulas. Todo $\mathbf{x} \in (\mathbb{K}^\infty)_0$ tem a forma

$$\mathbf{x} = (x_1, \dots, x_n, 0, 0, \dots) = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$$

para algum n. Logo $\underline{\mathrm{B}} = \{\mathbf{e}_1, \mathbf{e}_2, \dots\}$ é a base canônica de $(\mathbb{K}^\infty)_0$.

◁

Exemplo 2.39. Seja $V = \mathscr{M}_{m,n}(\mathbb{K})$. Para $1 \leq i \leq m$ e $1 \leq j \leq n$, denote por E_{ij} a matriz cuja entrada (i, j) é 1 e cujas demais entradas são nulas. Como $(a_{ij}) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} E_{ij}$, o conjunto $\underline{\mathrm{B}} = \{E_{ij} \mid 1 \leq i \leq m,\ 1 \leq j \leq n\}$ é uma base de V. Seus elementos são as **unidades matriciais**.

◁

Exemplo 2.40. Seja $V = \mathbb{K}[x]$. Seja $\underline{\mathrm{B}}$ o conjunto dos monômios mônicos em x, assim, $\underline{\mathrm{B}} = \{1, x, x^2, \ldots\}$ e é uma base de $\mathbb{K}[x]$.

◁

Exemplo 2.41. O conjunto dos números reais é um espaço vetorial sobre o corpo dos números racionais.

Como seria uma base desse espaço? Não há uma descrição elementar. Como $\mathbb{Q}$ é enumerável, o subespaço de $\mathbb{R}$ gerado por qualquer subconjunto enumerável também é enumerável. Como $\mathbb{R}$ não é enumerável, nenhuma base de $\mathbb{R}$ sobre $\mathbb{Q}$ pode ser enumerável. O teorema seguinte garante a existência de tal base, mas não a constrói explicitamente.

◁

Leitura opcional: bases em dimensão infinita. Os resultados a seguir usam o Lema de Zorn e aritmética de cardinais. Em uma primeira leitura voltada a espaços de dimensão finita, esta passagem pode ser omitida.

Todo espaço vetorial possui uma base, como afirma o teorema seguinte:

> **Teorema 2.42** (Base).
> *Todo espaço vetorial possui uma base.*

De fato, temos um resultado ligeiramente mais forte: qualquer subconjunto linearmente independente S de V pode ser expandido para uma base. É esse fato que provaremos.

> **Teorema 2.43** (Extensão).
> *Seja V um espaço vetorial sobre $\mathbb{K}$ e suponha que S seja um subconjunto linearmente independente de V. Então, existe uma base $\underline{B}$ de V tal que $\underline{B} \supseteq S$.*

Demonstração. Denotaremos por $\mathscr{I}$ o conjunto de todos os subconjuntos linearmente independentes de V que contêm S:

$$\mathscr{I} = \{A \in \mathscr{P}(V) \mid A \supseteq S \text{ e } A \text{ é linearmente independente sobre } \mathbb{K}\}.$$

Como $S \in \mathscr{I}$, temos $\mathscr{I} \neq \emptyset$. Ordenamos $\mathscr{I}$ por inclusão. Assim, para $A_1, A_2 \in \mathscr{I}$,

$$A_1 \leq A_2 \text{ se, e somente se, } A_1 \subseteq A_2.$$

Dessa forma $(\mathscr{I}, \subseteq)$ é um conjunto parcialmente ordenado.

Seja $\mathcal{T} = \{A_i \mid i \in I\}$ uma cadeia em $\mathcal{I}$ e defina $A = \bigcup_{i \in I} A_i$. Temos $S \subseteq A$ e $A_i \subseteq A$ para todo i. Se A fosse linearmente dependente, uma relação não trivial envolveria apenas um subconjunto finito de A. Como $\mathcal{T}$ é totalmente ordenada, esse subconjunto finito estaria contido em algum A_{i_0}, contradizendo a independência linear de A_{i_0}. Portanto, $A \in \mathcal{I}$ e A é um limitante superior de $\mathcal{T}$.

O Lema de Zorn fornece um elemento maximal $\underline{\mathrm{B}} \in \mathcal{I}$. Esse conjunto é linearmente independente e contém S. Se $\langle \underline{\mathrm{B}} \rangle \neq V$, existiria $\mathbf{x} \in V \setminus \langle \underline{\mathrm{B}} \rangle$; então $\underline{\mathrm{B}} \cup \{\mathbf{x}\}$ ainda seria linearmente independente, contrariando a maximalidade de $\underline{\mathrm{B}}$. Logo $\langle \underline{\mathrm{B}} \rangle = V$, e $\underline{\mathrm{B}}$ é a base desejada. ■

do Teorema 2.42. Suponha $V \neq \{\vec{0}\}$ e seja $\mathbf{v} \neq \vec{0}$. Então $\{\mathbf{v}\}$ é um conjunto linearmente independente e pelo Teorema 2.43 pode ser completado a uma base de V. ■

A demonstração tem uma limitação importante: o Lema de Zorn garante a existência de uma base, mas em geral não fornece um procedimento para construí-la. Em muitos espaços de dimensão infinita, essa base de Hamel não admite uma descrição explícita simples.

Exemplo 2.44. Uma base para o espaço vetorial $C^k(I)$ é não enumerável e não possui descrições simples. No entanto, como $\mathbb{R}[x] \subseteq C^k(I)$, o Teorema 2.43 garante que existe uma base de $C^k(I)$ que contém os monômios $1, x, x^2, \ldots$.

◁

O Teorema 2.43 afirma que qualquer subconjunto linearmente independente de V pode ser estendido a uma base. O resultado complementar diz que todo conjunto gerador contém uma base.

> **Teorema 2.45** (Redução).
> *Seja V um espaço vetorial sobre $\mathbb{K}$ e suponha que $V = \langle S \rangle$. Então S contém uma base de V.*

Demonstração. Se $S = \emptyset$ ou $\{\vec{0}\}$, então $V = (0)$. Nesse caso, $\emptyset$ é uma base de V contida em S. Portanto, podemos supor que S contém um vetor diferente de zero $\mathbf{x}$.

Seja $\mathcal{I} = \{A \subseteq S \mid A \text{ linearmente independente sobre } \mathbb{K}\}$. Claramente, $\{\mathbf{x}\} \in \mathcal{I}$. Ordenaremos parcialmente $\mathcal{I}$ por inclusão.

Se $\mathcal{T} = \{A_i \mid i \in I\}$ é um subconjunto totalmente ordenado de $\mathcal{I}$, então $\bigcup_{i \in I} A_i$ é um limite superior para $\mathcal{T}$ em $\mathcal{I}$.

Portanto, $(\mathcal{I}, \subseteq)$ é indutivo. Aplicando o Lema de Zorn temos que $\mathcal{I}$ possui um elemento maximal $\underline{\mathrm{B}}$.

Afirmamos que $\underline{\mathrm{B}}$ é uma base de V. Já sabemos que $\underline{\mathrm{B}} \subseteq S$ e que $\underline{\mathrm{B}}$ é linearmente independente. Se $\langle \underline{\mathrm{B}} \rangle \neq V = \langle S \rangle$, então $S \not\subseteq \langle \underline{\mathrm{B}} \rangle$; logo existe $\mathbf{y} \in S \setminus \langle \underline{\mathrm{B}} \rangle$. Nesse caso, $\underline{\mathrm{B}} \cup \{\mathbf{y}\} \in \mathcal{I}$ seria estritamente maior que $\underline{\mathrm{B}}$, em contradição com sua maximalidade. Portanto, $\langle \underline{\mathrm{B}} \rangle = V$. ■

Um espaço vetorial costuma possuir muitas bases. Por exemplo, se $\lambda \in \mathbb{K}$ e $\lambda \neq -1$, então

$$\{\mathbf{e}_1, \dots, \mathbf{e}_{n-1}, \mathbf{e}_n + \lambda \mathbf{e}_1\}$$

é uma base de $\mathbb{K}^n$. O que todas as bases de um mesmo espaço têm em comum é sua cardinalidade, como mostraremos a seguir.

Lema 2.46 (Troca de Steinitz).
Sejam $\mathbf{v}_1, \dots, \mathbf{v}_m$ vetores linearmente independentes em V e sejam $\mathbf{w}_1, \dots, \mathbf{w}_n$ vetores que geram V. Então $m \leq n$ e, possivelmente após reordenar os vetores $\mathbf{w}_i$, o conjunto $\{\mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{w}_{m+1}, \dots, \mathbf{w}_n\}$ gera V.

Demonstração. Desejamos mostrar que $m \leq n$ e que, após rearranjar os $\mathbf{w}_j$ se necessário, o conjunto $\{\mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{w}_{m+1}, \dots, \mathbf{w}_n\}$ gera V. Vamos proceder por indução em m.

Para o caso base, suponha que m seja zero. Nesse caso, a afirmação é verdadeira porque não há vetores $\mathbf{v}_i$, e o conjunto $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ gera V por hipótese.

Para o passo indutivo, assuma que a proposição é verdadeira para $m-1$. Pela hipótese de indução, podemos reorganizar os $\mathbf{w}_i$ de modo que $\{\mathbf{v}_1, \dots, \mathbf{v}_{m-1}, \mathbf{w}_m, \dots, \mathbf{w}_n\}$ gere V. Como $\mathbf{v}_m \in V$, existem coeficientes $a_1, \dots, a_n$ tais que

$$\mathbf{v}_m = \sum_{j=1}^{m-1} a_j \mathbf{v}_j + \sum_{j=m}^{n} a_j \mathbf{w}_j.$$

Pelo menos um dos $a_m, \dots, a_n$ deve ser diferente de zero, pois, caso contrário, essa igualdade contradiria a independência linear de $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$; portanto, $m \leq n$. Reordenando os vetores $\mathbf{w}_m, \dots, \mathbf{w}_n$ se necessário, podemos assumir que a_m é não nulo. Portanto,

$$\mathbf{w}_m = \frac{1}{a_m}\left(\mathbf{v}_m - \sum_{j=1}^{m-1} a_j \mathbf{v}_j - \sum_{j=m+1}^{n} a_j \mathbf{w}_j\right).$$

Em outras palavras, $\mathbf{w}_m$ está no espaço gerado por $\{\mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{w}_{m+1}, \dots, \mathbf{w}_n\}$. Como esse espaço contém o conjunto que, pela hipótese de indução, gera V, ele coincide com V. ∎

O lema contém duas informações. Primeiro, um conjunto finito linearmente independente não pode ter mais elementos que um conjunto gerador finito. Segundo, parte do conjunto gerador pode substituir, uma a uma, as direções introduzidas pelo conjunto independente, sem perder a propriedade de gerar V.

O resultado seguinte justifica a definição da dimensão de um espaço vetorial.

Teorema 2.47 (da Invariância da Dimensão).
Seja V um espaço vetorial sobre $\mathbb{K}$ e suponha que $\underline{B}_1$ e $\underline{B}_2$ sejam duas bases de V. Então $\left|\underline{B}_1\right| = \left|\underline{B}_2\right|$.

Demonstração. Dividimos essa prova em dois casos.

Caso 1: suponha que V tenha uma base $\underline{\mathrm{B}}_1$ finita.

Suponha que $\underline{\mathrm{B}}_1 = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$. Se $\underline{\mathrm{B}}_2$ for finita, o Lema de Steinitz, aplicado nos dois sentidos, fornece $\left|\underline{\mathrm{B}}_1\right| = \left|\underline{\mathrm{B}}_2\right|$. Se $\underline{\mathrm{B}}_2$ fosse infinita, conteria um subconjunto linearmente independente com $n+1$ elementos, o que também contradiria o Lema de Steinitz, pois $\underline{\mathrm{B}}_1$ gera V.

Caso 2: suponha que nenhuma base de V seja finita.

Nesse caso, $\underline{\mathrm{B}}_1$ e $\underline{\mathrm{B}}_2$ são conjuntos infinitos. Para cada $\mathbf{y} \in \underline{\mathrm{B}}_2$, seja $I_{\mathbf{y}} \subseteq \underline{\mathrm{B}}_1$ o suporte finito da expressão de $\mathbf{y}$ na base $\underline{\mathrm{B}}_1$. Como $\underline{\mathrm{B}}_2$ gera V, a união desses suportes gera V; sendo um subconjunto de $\underline{\mathrm{B}}_1$, ela deve ser igual a $\underline{\mathrm{B}}_1$. Portanto,

$$\left|\underline{\mathrm{B}}_1\right| \leq \left|\underline{\mathrm{B}}_2\right| \cdot \aleph_0 = \left|\underline{\mathrm{B}}_2\right|.$$

Invertendo os papéis de $\underline{\mathrm{B}}_1$ e $\underline{\mathrm{B}}_2$, obtemos a desigualdade contrária. Pelo teorema de Cantor–Schröder–Bernstein, $\left|\underline{\mathrm{B}}_1\right| = \left|\underline{\mathrm{B}}_2\right|$. ∎

Definição 2.48 — Dimensão. A cardinalidade comum das bases de V é denominada **dimensão** de V e denotada por $\dim V$. Para explicitar o corpo de escalares, escreveremos $\dim_{\mathbb{K}} V$. Assim, $\dim V = \left|\underline{\mathrm{B}}\right|$ para qualquer base $\underline{\mathrm{B}}$ de V.

Retomemos alguns exemplos. No Exemplo 2.33, $\dim_{\mathbb{R}} \mathbb{R} = 1$ e $\dim_{\mathbb{Q}} \mathbb{R} = |\mathbb{R}|$. No Exemplo 2.36, $\dim_{\mathbb{K}} \{\vec{0}\} = 0$. Pelos Exemplos 2.37, 2.39 e 2.40, temos, respectivamente, $\dim \mathbb{K}^n = n$, $\dim \mathscr{M}_{m,n}(\mathbb{K}) = mn$ e $\dim \mathbb{K}[x] = |\mathbb{N}|$.

Se a dimensão de um espaço vetorial V for infinita, como nos Exemplos 2.33 e 2.40, em geral não faremos nenhuma tentativa de distinguir qual a cardinalidade de $\dim V$. Em vez disso, escreveremos simplesmente $\dim V = \infty$. Se V tiver uma base finita $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$, diremos que V é um espaço vetorial de dimensão finita e escreveremos $\dim V < \infty$, ou, mais precisamente, $\dim V = n < \infty$. Por exemplo, $\dim_{\mathbb{R}} C^k(I) = \infty$, enquanto $\dim_{\mathbb{R}} \mathbb{R}^n = n < \infty$.

Teorema 2.49.
Seja V um espaço vetorial sobre $\mathbb{K}$.

- [a] *Se W for um subespaço de V, então* $\dim W \leq \dim V$.
- [b] *Se V for de dimensão finita e W é um subespaço de V tal que* $\dim W = \dim V$, $W = V$.

Demonstração. Se $\underline{\mathrm{B}}_W$ é uma base de W, o Teorema 2.43 permite estendê-la a uma base $\underline{\mathrm{B}}_V$ de V. Logo $\dim W = \left|\underline{\mathrm{B}}_W\right| \leq \left|\underline{\mathrm{B}}_V\right| = \dim V$, provando [a]. Se V tem dimensão finita e essas dimensões são iguais, a extensão não pode acrescentar vetores; portanto, $\underline{\mathrm{B}}_W = \underline{\mathrm{B}}_V$ e $W = V$, o que prova [b]. ∎

Se V não é de dimensão finita, então [b] é falso em geral. Um exemplo simples ilustra esse ponto.

Exemplo 2.50. Seja $V = \mathbb{K}[x]$ e seja W o subespaço dos polinômios pares. Então $\{1, x^2, x^4, \ldots\}$ é uma base de W, enquanto $\{1, x, x^2, \ldots\}$ é uma base de V. Ambas são enumeráveis, de modo que $\dim V = \dim W$, embora $W \neq V$.

◁

Exercícios

Ex. 2.21 — Suponha que S seja finito. Exiba uma base para o espaço de funções $\mathbb{K}(S) = \mathbb{K}^S$.

Ex. 2.22 — Dado um subespaço $L \subseteq V$, prove que, se $\dim L = \dim V < \infty$, então $L = V$.

Ex. 2.23 —

1. Prove que os únicos subespaços de $\mathbb{R}$ são o próprio $\mathbb{R}$ e o subespaço nulo.
2. Prove que todos os subespaços de $\mathbb{R}^2$ são o próprio $\mathbb{R}^2$, o subespaço nulo ou o subespaço constituído pelos múltiplos de um vetor não nulo fixo de $\mathbb{R}^2$.
3. Quais são todos os subespaços de $\mathbb{R}^3$?

Ex. 2.24 — Seja L um espaço n-dimensional sobre um corpo finito com q elementos.

1. Calcule o número de subespaços k-dimensionais de L, para $1 \leq k \leq n$.
2. Calcule o número de pares de subespaços L_1 e L_2 com $\dim L_1$, $\dim L_2$ e $\dim(L_1 \cap L_2)$ fixos. Verifique que, quando $q \to \infty$, a razão entre o número de pares em posição geral e o número total de pares com $\dim L_1$ e $\dim L_2$ fixadas tende a 1.

Ex. 2.25 — Seja X um subconjunto de um espaço vetorial V sobre $\mathbb{K}$. Mostre que

$$\langle X \rangle = \left\{ \sum_{v \in F} \alpha_v v \,:\, F \subseteq X \text{ é finito e } \alpha_v \in \mathbb{K} \text{ para todo } v \in F \right\}.$$

Ex. 2.26 — Seja $\underline{\mathrm{B}}$ um subconjunto de um espaço vetorial V. Mostre que $\underline{\mathrm{B}}$ é linearmente dependente se, e somente se, existir $v \in \underline{\mathrm{B}}$ que pode ser escrito como combinação linear dos elementos de $\underline{\mathrm{B}} \setminus \{v\}$.

Ex. 2.27 — Se $\underline{\mathrm{E}} = \{\mathbf{e}_i\}_{i \in I}$ é uma base de V e $\underline{\mathrm{F}} = \{\vec{f}_j\}_{j \in J}$ é uma base de W, mostre que $\{(\mathbf{e}_i, 0)\}_{i \in I} \cup \{(0, \vec{f}_j)\}_{j \in J}$ é uma base de $V \boxplus W$.

Ex. 2.28 — Seja V um espaço vetorial sobre $\mathbb{K}$ de dimensão não necessariamente finita e seja $\underline{B}$ um conjunto linearmente independente em V. Mostre que se existir um elemento $v \in V$ que não seja combinação linear de elementos de $\underline{B}$, então o conjunto $\underline{B} \cup \{v\}$ é linearmente independente.

Ex. 2.29 — Seja V um espaço vetorial sobre $\mathbb{K}$. Mostre que:

1. Dados vetores não nulos v_1, v_2, prove que eles são linearmente independentes se, e somente se, $\langle v_1 \rangle \cap \langle v_2 \rangle = \{0\}$.

2. Dados vetores v_1, v_2, v_3 não nulos. Prove que $\langle v_1 \rangle \cap \langle v_2 \rangle \cap \langle v_3 \rangle = \{0\}$ não implica que os vetores v_1, v_2, v_3 sejam linearmente independentes.

3. Dado $S = \{s_1, \dots, s_n\} \subset V$. Prove que S é linearmente independente se, e somente se, $\langle S \setminus \{s_i\} \rangle \neq \langle S \rangle$ para todo $s_i \in S$.

4. Prove que, se $A, B \subseteq V$, então $\langle A \rangle + \langle B \rangle = \langle A \cup B \rangle$.

Ex. 2.30 — Mostre que se os coeficientes $a_1, \dots, a_n$ não são todos iguais a zero, o hiperplano

$$H = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid a_1 x_1 + \cdots + a_n x_n = 0\}$$

é um subespaço vetorial de dimensão $n - 1$ em $\mathbb{R}^n$.

Ex. 2.31 — Prove que em qualquer conjunto de vetores S existe um subconjunto S' linearmente independente tal que $\langle S \rangle = \langle S' \rangle$.

Ex. 2.32 — Se $\mathbb{K} = \mathbb{Z}_2$, o subconjunto $\{(\overline{1}, \overline{1}, \overline{0}), (\overline{1}, \overline{0}, \overline{1}), (\overline{0}, \overline{1}, \overline{1})\}$ de $\mathbb{K}^3$ é linearmente dependente? E se $\mathbb{K} = \mathbb{Z}_{13}$?

Ex. 2.33 — Seja $V = \mathscr{F}(\mathbb{R}, \mathbb{C})$ o $\mathbb{C}$-espaço vetorial de todas as funções de $\mathbb{R}$ em $\mathbb{C}$. Prove que $\{f_1, f_2, f_3\}$ é linearmente independente, onde $f_1(x) = 1$, $f_2(x) = e^{ix} = \cos(x) + i\,\mathrm{sen}(x)$ e $f_3(x) = e^{-ix}$ para todo $x \in \mathbb{R}$.

Ex. 2.34 — Considere o espaço das funções $\mathscr{F}(\mathbb{R}, \mathbb{R})$. Mostre que os seguintes subconjuntos são linearmente independentes:

1. $(f_n)_{n \in \mathbb{N}}$ onde $f_n : x \mapsto e^{nx}$

2. $(f_a)_{a \in \mathbb{R}}$ onde $f_a : x \mapsto |x - a|$

Ex. 2.35 — Seja V um espaço vetorial sobre $\mathbb{R}$ e considere no conjunto $V_{\mathbb{C}} = \{(u, v) : u, v \in V\}$ as seguintes operações de adição e multiplicação por um número complexo:

$$(u_1, v_1) + (u_2, v_2) = (u_1 + u_2, v_1 + v_2)$$
$$(\alpha + i\beta) \cdot (u, v) = (\alpha u - \beta v, \beta u + \alpha v)$$

1. Mostre que $V_{\mathbb{C}}$ é um espaço vetorial sobre $\mathbb{C}$.

2. Seja $\{v_1, v_2, \dots, v_n\} \subseteq V$ um subconjunto linearmente independente. Mostre que $\{(v_1, 0), (v_2, 0), \dots, (v_n, 0)\}$ e $\{(0, v_1), (0, v_2), \dots, (0, v_n)\}$ são subconjuntos linearmente independentes de $V_{\mathbb{C}}$.

Ex. 2.36 — Para um $\mathbb{C}$-espaço vetorial V, denotaremos por $V_{\mathbb{R}}$ o conjunto V visto como $\mathbb{R}$-espaço vetorial. Mostre que se $\{v_1, v_2, \dots, v_n\}$ for um subconjunto linearmente independente de V, então $\{v_1, v_2, \dots, v_n\}$ e $\{v_1, v_2, \dots, v_n\} \cup \{iv_1, iv_2, \dots, iv_n\}$ são subconjuntos linearmente independentes de $V_{\mathbb{R}}$.

Ex. 2.37 — Seja V um espaço vetorial e seja $\underline{\mathrm{B}} \subseteq V$. Mostre que são equivalentes:

1. $\underline{\mathrm{B}}$ é uma base de V;

2. $\underline{\mathrm{B}}$ é um conjunto linearmente independente maximal, isto é, a adição de qualquer vetor de $V \setminus \underline{\mathrm{B}}$ produz um conjunto linearmente dependente;

3. $\underline{\mathrm{B}}$ é um conjunto gerador minimal, isto é, a remoção de qualquer vetor de $\underline{\mathrm{B}}$ produz um conjunto que não gera V.

Ex. 2.38 — Seja $V = \langle X \rangle$. Mostre que existe uma base $\underline{\mathrm{B}}$ de V tal que $\underline{\mathrm{B}} \subseteq X$. Dica: Defina um conjunto parcialmente ordenado e use o lema de Zorn.

Ex. 2.39 — Ache uma base de $\mathscr{M}_{n,n}(\mathbb{C})$ como espaço vetorial sobre $\mathbb{R}$. Qual é a $\dim_{\mathbb{R}}(\mathscr{M}_{n,n}(\mathbb{C}))$?

Ex. 2.40 — Seja S o $\mathbb{R}$-espaço vetorial do Exercício 2.14. Mostre que $\dim_{\mathbb{R}} S = n$.

Dica. Use o Teorema de Existência e Unicidade de soluções. Considere a equação

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \cdots + a_1 y'(t) + a_0 y(t) = 0 \tag{2.6}$$

onde $a_0, a_1, \dots, a_{n-1} \in \mathbb{R}$. Dados $A_0, A_1, \dots, A_{n-1} \in \mathbb{R}$, existe uma única solução $y : \mathbb{R} \to \mathbb{R}$ da equação (2.6) verificando $y(0) = A_0$, $y'(0) = A_1, \dots, y^{(n-1)}(0) = A_{n-1}$ (condições iniciais da equação (2.6)).

Construa n soluções que formarão uma base do espaço das soluções de (2.6), da seguinte forma: considerando as condições iniciais $A_0 = 1$ e $A_1 = \cdots = A_{n-1} = 0$ o Teorema de Existência e Unicidade garante que existe uma única solução $y_1 : \mathbb{R} \to \mathbb{R}$ de (2.6) que verifica as condições $y_1(0) = 1$ e $y_1'(0) = \cdots = y_1^{(n-1)}(0) = 0$. Repita o procedimento considerando as condições iniciais $A_i = 1$ e $A_j = 0$ para todo $j \neq i$ com i variando de 1 a $n - 1$.

Ex. 2.41 — Determine se os espaços abaixo têm dimensão finita. Em caso afirmativo, determine a dimensão e uma base para o espaço:

1. O conjunto de todas as sequências reais.

2. O conjunto das sequências reais que satisfazem $a_k = a_{k-1} + a_{k-2}$ para $k \geqslant 3$.

3. $\mathbb{C}^n$ visto como um espaço vetorial sobre $\mathbb{C}$ e visto como um espaço vetorial

sobre $\mathbb{R}$.

4. O conjunto das sequências com apenas um número finito de termos não nulos.

5. O espaço das funções em $\mathscr{F}(X, \mathbb{R})$, $|X| < \infty$ que se anulam em todos os pontos de um subconjunto $X_0 \subset X$.

6. O espaço das funções $C([0,1], \mathbb{R})$.

Ex. 2.42 — Seja $\mathbb{F}$ um corpo. Um subcorpo $\mathbb{K}$ é um subconjunto de $\mathbb{F}$ que é corpo quando restringimos as operações de $\mathbb{F}$ a $\mathbb{K}$.

1. Mostre que $\mathbb{F}$ é espaço vetorial sobre $\mathbb{K}$.

2. Suponha que V seja um espaço vetorial m-dimensional sobre $\mathbb{F}$ e que $\mathbb{F}$ seja um espaço vetorial n-dimensional sobre $\mathbb{K}$. Qual a dimensão de V sobre $\mathbb{K}$?

Ex. 2.43 — Seja W um subespaço de V com $\dim(V) < \infty$. Prove que:

1. $\dim(W) \leq \dim(V)$

2. $\dim(W) = \dim(V)$ se, e somente se, $W = V$.

Ex. 2.44 — Mostre que $\mathbb{R}$ é um espaço vetorial de dimensão infinita sobre $\mathbb{Q}$.

2.4 Soma de Subespaços

Soma e união
A união de dois subespaços não é, em geral, um subespaço: a soma de um vetor de cada conjunto pode não pertencer à união. A soma $U + W$ é o menor subespaço que contém ambos.

O conjunto $\mathcal{S}(V)$ de todos os subespaços de V é parcialmente ordenado pela inclusão. Nessa ordem, $\{\vec{0}\}$ é o menor elemento e V é o maior. Essa estrutura de ordem motiva as construções seguintes.

Se $S, T \in \mathcal{S}(V)$, então $S \cap T$ é o maior subespaço contido simultaneamente em S e T. Mais geralmente, a interseção de uma família $\{S_i \mid i \in I\}$ é seu maior limite inferior:

$$\inf\{S_i \mid i \in I\} = \bigcap_{i \in I} S_i$$

Para determinar o menor subespaço de V contendo os subespaços S e T, fazemos a seguinte definição.

Definição 2.51 — Soma. Se S e T são subespaços de V, sua **soma** é o conjunto

$$S + T \coloneqq \{\mathbf{x} + \mathbf{y} \mid \mathbf{x} \in S,\ \mathbf{y} \in T\}$$

Para uma família $\{S_i \mid i \in I\}$, definimos sua soma como o conjunto de todas as somas finitas com uma parcela em cada subespaço escolhido:

$$\sum_{i \in I} S_i \coloneqq \{\mathbf{x}_{i_1} + \cdots + \mathbf{x}_{i_r} \mid r \geq 0,\ i_1, \ldots, i_r \in I,\ \mathbf{x}_{i_j} \in S_{i_j}\}.$$

Proposição 2.52.
Dada uma coleção $\{S_i \mid i \in I\}$ de subespaços de V, as condições seguintes são equivalentes:

1. *W é o menor subespaço que contém S_i para todo $i \in I$;*
2. $W = \langle \bigcup \{S_i \mid i \in I\} \rangle$
3. $W = \sum_{i \in I} S_i$

A equivalência decorre diretamente da caracterização de $\langle S \rangle$ como o conjunto das combinações lineares finitas de elementos de S.

2.4.1 Soma Direta

Uma soma direta descreve um espaço por componentes independentes: cada vetor se decompõe de maneira única em parcelas pertencentes aos subespaços dados.

Somas Diretas Internas

Definição 2.53 — Soma Direta. Seja V um espaço vetorial e seja $\mathcal{S} = \{S_i \mid i \in I\}$ uma família de subespaços. Dizemos que V é a **soma direta interna** dessa família se todo vetor $\mathbf{v} \in V$ possui uma única representação como soma finita

$$\mathbf{v} = \mathbf{v}_{i_1} + \cdots + \mathbf{v}_{i_r}, \qquad \mathbf{v}_{i_j} \in S_{i_j},$$

depois de desconsiderar parcelas nulas e a ordem das parcelas.
Se V é a soma direta de $\mathcal{S}$, escrevemos

$$V = \bigoplus_{i \in I} S_i.$$

Se $\mathcal{S} = \{S_1, \ldots, S_n\}$ é uma família finita, podemos denotar a soma direta como

$$V = S_1 \oplus \cdots \oplus S_n$$

Uma soma é direta se, e somente se, toda igualdade $\mathbf{u}_{i_1} + \cdots + \mathbf{u}_{i_r} = \vec{0}$, com índices distintos e $\mathbf{u}_{i_j} \in S_{i_j}$, implica $\mathbf{u}_{i_j} = \vec{0}$ para todo j. Em outras palavras, basta verificar a unicidade da representação do vetor nulo. Essa afirmação motiva a seguinte definição e a próxima proposição.

Definição 2.54 — Espaços Independentes. Uma família finita $\{S_1, \ldots, S_k\}$ de subespaços de V é **independente** se $\mathbf{s}_1 + \cdots + \mathbf{s}_k = \vec{0}$, com $\mathbf{s}_i \in S_i$, implica $\mathbf{s}_i = \vec{0}$ para todo i.

Assim, para subespaços $S_1, \ldots, S_k$ de V, a igualdade $V = S_1 \oplus \cdots \oplus S_k$ vale se, e somente se,

- **a** $V = S_1 + \cdots + S_k$ e
- **b** a família $\{S_1, \ldots, S_k\}$ é independente.

A seguinte caracterização de somas diretas é bastante útil.

Teorema 2.55.
Um espaço vetorial V é a soma direta de uma família $\mathcal{S} = \{S_i \mid i \in I\}$ de subespaços se, e somente se,

a *V é a soma dos S_i, $V = \sum_{i\in I} S_i$*

b *Para cada $i \in I$, $S_i \cap (\sum_{j\neq i} S_j) = \{\vec{0}\}$*

Demonstração. Suponha primeiro que V seja a soma direta de $\mathcal{S}$. Então a é verdadeiro e se

$$\mathbf{v} \in S_i \cap (\sum_{j\neq i} S_j)$$

então $\mathbf{v} = \mathbf{s}_i$ para algum $\mathbf{s}_i \in S_i$ e

$$\mathbf{v} = \mathbf{s}_{j_1} + \cdots + \mathbf{s}_{j_n}$$

onde $\mathbf{s}_{j_k} \in S_{j_k}$ e $j_k \neq i$ para todos os $k = 1, \ldots, n$. Portanto, pela unicidade das representações de soma direta, $\mathbf{s}_i = 0$ e, portanto, $\mathbf{v} = 0$. Logo demonstramos b.

Para a recíproca, suponha que a e b sejam válidos. Precisamos apenas verificar a condição de unicidade

$$\mathbf{v} = \mathbf{s}_{j_1} + \cdots + \mathbf{s}_{j_n} \quad \text{e} \tag{2.7}$$
$$\mathbf{v} = \mathbf{t}_{k_1} + \cdots + \mathbf{t}_{k_m} \tag{2.8}$$

onde $\mathbf{s}_{j_i} \in S_{j_i}$ e $\mathbf{t}_{k_i} \in S_{k_i}$. Acrescentando parcelas nulas, podemos supor que os conjuntos de índices das duas representações coincidem, digamos, com $\{i_1, \ldots, i_p\}$; assim,

$$\mathbf{v} = \mathbf{s}_{i_1} + \cdots + \mathbf{s}_{i_p} \quad \text{e} \tag{2.9}$$
$$\mathbf{v} = \mathbf{t}_{i_1} + \cdots + \mathbf{t}_{i_p} \tag{2.10}$$

A diferença das duas igualdades fornece $(\mathbf{s}_{i_1} - \mathbf{t}_{i_1}) + \cdots + (\mathbf{s}_{i_p} - \mathbf{t}_{i_p}) = \vec{0}$. Para cada u, temos

$$\mathbf{s}_{i_u} - \mathbf{t}_{i_u} = -\sum_{v\neq u} (\mathbf{s}_{i_v} - \mathbf{t}_{i_v}) \in S_{i_u} \cap \sum_{j\neq i_u} S_j = \{\vec{0}\}.$$

Logo $\mathbf{s}_{i_u} = \mathbf{t}_{i_u}$ para todo u, o que prova a unicidade. ■

Se tivermos apenas dois subespaços $\{W_1,\ W_2\}$, essa condição simplesmente indicará $W_1 \cap W_2 = \{\vec{0}\}$. Se tivermos mais de dois subespaços, essa condição é mais forte que a condição $W_i \cap W_j = \{\vec{0}\}$ para $i \neq j$.

Exemplo 2.56. Suponha que $\mathrm{char}(\mathbb{K})$ não divida n. Seja $\mathscr{M}_{n,n}(\mathbb{K})$ o espaço vetorial das matrizes $n \times n$ com entradas em $\mathbb{K}$, seja $\mathfrak{sl}_n(\mathbb{K})$ o subespaço das matrizes de traço zero e seja $D = \langle \mathrm{I}_n \rangle$.

Como $\mathrm{tr}(\lambda\, \mathrm{I}_n) = n\lambda$, temos $\mathfrak{sl}_n(\mathbb{K}) \cap D = \{\vec{0}\}$. Para qualquer $A \in \mathscr{M}_{n,n}(\mathbb{K})$, defina $\alpha = \mathrm{tr}(A)/n$ e $B = A - \alpha\, \mathrm{I}_n$. Então $\mathrm{tr}(B) = 0$ e $A = B + \alpha\, \mathrm{I}_n$. Portanto, $\mathscr{M}_{n,n}(\mathbb{K}) = \mathfrak{sl}_n(\mathbb{K}) \oplus D$.

◁

Exemplo 2.57. Suponha que $\mathrm{char}(\mathbb{K}) \neq 2$. Toda matriz $A \in \mathscr{M}_{n,n}(\mathbb{K})$ pode ser decomposta como

$$A = \frac{1}{2}(A + A^{\mathrm{t}}) + \frac{1}{2}(A - A^{\mathrm{t}}) = B + C \tag{2.11}$$

onde A^{t} é a transposta de A. A matriz B é simétrica e C é antissimétrica; assim, A é a soma de uma matriz simétrica e uma matriz antissimétrica.

Como os conjuntos $\mathrm{Sim}_{n,n}$ e $\mathrm{ASim}_{n,n}$ de todas as matrizes simétricas e assimétricas em $\mathscr{M}_{n,n}(\mathbb{K})$ são subespaços de $\mathscr{M}_{n,n}(\mathbb{K})$, temos que

$$\mathscr{M}_{n,n}(\mathbb{K}) = \mathrm{Sim}_{n,n} + \mathrm{ASim}_{n,n}$$

Além disso, se $S + T = S' + T'$, onde S e S' são simétricas e T e T' são antissimétricas, então a matriz

$$U = S - S' = T' - T$$

é simétrica e antissimétrica. Como $\mathrm{char}(\mathbb{K}) \neq 2$, devemos ter $U = 0$ e, portanto, $S = S'$ e $T = T'$. Assim, a soma é direta:

$$\mathscr{M}_{n,n}(\mathbb{K}) = \mathrm{Sim}_{n,n} \oplus \mathrm{ASim}_{n,n}$$

◁

2.4.2 Decomposição em Somas Diretas

Proposição 2.58.
Sejam $W_1, \ldots, W_k$ subespaços de V, seja $\underline{B}_i$ uma base de W_i e ponha $\underline{B} = \bigcup_{i=1}^{k} \underline{B}_i$. Então

a $\underline{B}$ *gera V se, e somente se, $V = W_1 + \cdots + W_k$.*

b $\underline{B}$ *é linearmente independente se, e somente se, $\{W_1, \ldots, W_k\}$ for independente.*

c $\underline{B}$ *é uma base para V se, e somente se, $V = W_1 \oplus \cdots \oplus W_k$.*

A demonstração da proposição é proposta como exercício.

Corolário 2.59.
Seja V um espaço vetorial de dimensão finita. Se $V = W_1 \oplus \cdots \oplus W_k$, então $\dim V = \dim W_1 + \cdots + \dim W_k$.

A reunião de bases das parcelas é uma base de V; contar seus elementos prova o corolário.

Definição 2.60 — Complemento. Se W_1 é um subespaço de V, dizemos que W_2 é um **complemento** de W_1 quando $V = W_1 \oplus W_2$.

Teorema 2.61.
Seja V um espaço vetorial e seja W um subespaço de V. Então W possui um complemento W'. Além disso, $\dim V = \dim W + \dim W'$.

Demonstração. Seja $\underline{B}$ uma base de W. Pelo Teorema 2.43, existe uma base $\underline{B}'$ de V tal que $\underline{B} \subseteq \underline{B}'$. Defina $W' = \langle \underline{B}' \setminus \underline{B} \rangle$. A reunião das duas bases gera V e é linearmente independente; portanto, $V = W \oplus W'$. Como $\underline{B}' = \underline{B} \dot{\cup} (\underline{B}' \setminus \underline{B})$, a igualdade de cardinais $\dim V = \dim W + \dim W'$ também segue, inclusive em dimensão infinita.

■

Observação 2.62. O subespaço W' construído no Teorema 2.61 é chamado um **complemento** de W em V. Em geral, o complemento não é único, embora todos os complementos sejam isomorfos. Nos casos extremos, se $W = \{\vec{0}\}$, o único complemento é V; se $W = V$, o único complemento é $\{\vec{0}\}$.

Teorema 2.63.
Se V tem dimensão finita e W_1, W_2 são subespaços de V, então

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

Demonstração. Seja $\underline{B}_0 = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ uma base de $W_1 \cap W_2$. No caso particular em que $W_1 \cap W_2 = (0)$, faremos $\underline{B}_0 = \emptyset$. Pelo Teorema da Extensão 2.43, podemos completar $\underline{B}_0$ para uma base $\underline{B}_1 = \{\mathbf{x}_1, \dots, \mathbf{x}_n, \mathbf{y}_1, \dots, \mathbf{y}_m\}$ de W_1. De modo análogo, podemos completar $\underline{B}_0$ para uma base $\underline{B}_2 = \{\mathbf{x}_1, \dots, \mathbf{x}_n, \vec{z}_1, \dots, \vec{z}_p\}$ de W_2. Assim, $\dim W_1 \cap W_2 = n$, $\dim W_1 = n + m$, e $\dim W_2 = n + p$.

Afirmamos que $\underline{B} = \{\mathbf{x}_1, \dots, \mathbf{x}_n, \mathbf{y}_1, \dots, \mathbf{y}_m, \vec{z}_1, \dots, \vec{z}_p\}$ é uma base de $W_1 + W_2$. Claramente $\langle \underline{B} \rangle = W_1 + W_2$.

Resta provar que $\underline{B}$ é linearmente independente. Suponha que

$$\sum_{i=1}^{n} x_i \mathbf{x}_i + \sum_{i=1}^{m} y_i \mathbf{y}_i + \sum_{i=1}^{p} z_i \mathbf{z}_i = \vec{0}.$$

Então

$$\sum_{i=1}^{p} z_i \mathbf{z}_i = -\sum_{i=1}^{n} x_i \mathbf{x}_i - \sum_{i=1}^{m} y_i \mathbf{y}_i.$$

Comparando ambos os lados da equação temos que $\displaystyle\sum_{i=1}^{p} z_i \vec{z}_i \in W_1 \cap W_2 = \langle \{\mathbf{x}_1, \dots, \mathbf{x}_n\} \rangle$.

Logo podemos escrever $\displaystyle\sum_{i=1}^{p} z_i \vec{z}_i$ como:

$$\sum_{i=1}^{p} z_i \vec{z}_i = \sum_{i=1}^{n} w_i \vec{x}_i$$

Como $\underline{B}_2$ é uma base de W_2, comparar essa expressão com $\sum_{i=1}^{n} w_i \mathbf{x}_i$ fornece $z_1 = \cdots = z_p = 0$. A relação original se reduz a $\sum_{i=1}^{n} x_i \mathbf{x}_i + \sum_{i=1}^{m} y_i \mathbf{y}_i = \vec{0}$; como $\underline{B}_1$ é uma base de W_1, todos os coeficientes x_i e y_i também são nulos.

Logo $\underline{B}$ é uma base de $W_1 + W_2$, de modo que $\dim(W_1 + W_2) = n + m + p$. Substituir as quatro dimensões obtidas conclui a prova. ∎

O Teorema 2.63 continua válido em dimensão infinita quando as dimensões são entendidas como cardinais. Nesse caso, entretanto, a aritmética cardinal pode ocultar parte da informação geométrica: a soma de dois cardinais infinitos é igual ao maior deles.

O mesmo argumento fornece uma estimativa para a interseção de uma família finita de subespaços.

Corolário 2.64.
Seja V um espaço vetorial de dimensão n e sejam $W_1, \dots, W_k$ subespaços de V. Para $i = 1, \dots, k$, defina $c_i = n - \dim W_i$. Então

- $\dim(W_1 \cap \cdots \cap W_k) = n - \sum_{i=1}^{k} c_i + \sum_{j=1}^{k-1} \left[n - \dim\left((W_1 \cap \cdots \cap W_j) + W_{j+1}\right) \right]$;
- $\dim(W_1 \cap \cdots \cap W_k) \geq n - \sum_{i=1}^{k} c_i$;
- $\dim(W_1 \cap \cdots \cap W_k) = n - \sum_{i=1}^{k} c_i$ *se, e somente se, para todo* $i = 1, \dots, k$,

$$W_i + \left(\bigcap_{j \neq i} W_j\right) = V$$

Demonstração. Ponha $U_j = W_1 \cap \cdots \cap W_j$. O Teorema 2.63, aplicado a U_j e W_{j+1}, fornece

$$\dim U_{j+1} = \dim U_j + \dim W_{j+1} - \dim(U_j + W_{j+1}).$$

Iterando essa identidade, obtemos o primeiro item. Cada parcela entre colchetes é não negativa, o que prova o segundo.

A igualdade no segundo item equivale a $\operatorname{codim}(W_1 \cap \cdots \cap W_k) = \sum_i \operatorname{codim} W_i$. Ela ocorre exatamente quando cada W_i é transversal à interseção dos demais, isto é, quando $W_i + \bigcap_{j \neq i} W_j = V$. Essa equivalência também resulta da fórmula de dimensões do Teorema 2.63, aplicada a cada um desses pares. ∎

2.5 Soma Direta Externa

Se V e W forem espaços vetoriais sobre um corpo $\mathbb{K}$, o produto cartesiano

$$V \times W = \{(\mathbf{v},\ \mathbf{w}) \,:\, \mathbf{v} \in V,\ \mathbf{w} \in W\}$$

pode ser munido de uma estrutura de espaço vetorial definindo as operações

$$(\mathbf{v}_1,\ \mathbf{w}_1) + (\mathbf{v}_2,\ \mathbf{w}_2) \coloneqq (\mathbf{v}_1 + \mathbf{v}_2,\ \mathbf{w}_1 + \mathbf{w}_2) \tag{2.12}$$

$$\lambda(\mathbf{v},\ \mathbf{w}) \coloneqq (\lambda\mathbf{v},\ \lambda\mathbf{w})\,. \tag{2.13}$$

O produto cartesiano $V \times W$, munido dessas operações, é denominado **soma direta externa** de V e W.

A construção se estende a qualquer família finita.

Interna e externa
Na soma direta interna, os subespaços pertencem a um mesmo espaço. Na soma direta externa, construímos um novo espaço a partir dos espaços dados. Uma decomposição interna determina um isomorfismo com a soma direta externa correspondente, mas as duas construções não são o mesmo objeto.

Definição 2.65 — Soma Direta Externa. Sejam $V_1, \dots, V_n$ espaços vetoriais sobre $\mathbb{K}$. Sua soma direta externa, denotada por $V_1 \boxplus \cdots \boxplus V_n$, é o produto cartesiano $V_1 \times \cdots \times V_n$ munido das operações coordenada a coordenada. Seus elementos são as ênuplas ordenadas

$$V_1 \boxplus \cdots \boxplus V_n = \{(\mathbf{x}_1, \ \dots, \ \mathbf{x}_n) \mid \mathbf{x}_i \in V_i, \ i = 1, \ \dots, \ n\}$$

e as operações são

$$(\mathbf{y}_1, \ \dots, \ \mathbf{y}_n) + (\mathbf{x}_1, \ \dots, \ \mathbf{x}_n) \coloneqq (\mathbf{y}_1 + \mathbf{x}_1, \ \dots, \ \mathbf{y}_n + \mathbf{x}_n) \tag{2.14}$$

$$\lambda(\mathbf{x}_1, \ \dots, \ \mathbf{x}_n) \coloneqq (\lambda\mathbf{x}_1, \ \dots, \ \lambda\mathbf{x}_n) \tag{2.15}$$

Exemplo 2.66. O espaço vetorial $\mathbb{K}^n$ é naturalmente isomorfo à soma direta externa de n cópias de $\mathbb{K}$:

$$\mathbb{K}^n = \mathbb{K} \boxplus \cdots \boxplus \mathbb{K}$$

◁

Essa construção pode ser generalizada a uma família arbitrária de espaços vetoriais. Representaremos seus elementos por famílias $(v_i)_{i\in I}$, com $v_i \in V_i$, sem atribuir uma estrutura vetorial à união dos V_i.

Definição 2.67 — Produto Direto. Seja $\mathscr{F} = \{V_i \mid i \in I\}$ uma família de espaços vetoriais sobre $\mathbb{K}$. O **produto direto** da família é

$$\prod_{i\in I} V_i \coloneqq \{(v_i)_{i\in I} \mid v_i \in V_i \text{ para todo } i \in I\},$$

com adição e multiplicação por escalares definidas coordenada a coordenada.

Será mais útil restringir o conjunto de funções àquelas com suporte finito.

Definição 2.68 — Suporte. Seja $\mathscr{F} = \{V_i \mid i \in I\}$ uma família de espaços vetoriais sobre $\mathbb{K}$. O suporte de $v = (v_i)_{i\in I} \in \prod_{i\in I} V_i$ é o conjunto

$$\text{suporte}(v) \coloneqq \{i \in I \mid v_i \neq 0\}.$$

Assim, uma família tem suporte finito quando apenas um número finito de suas coordenadas é não nulo.

Definição 2.69 — Soma Direta Externa. A **soma direta externa** da família $\mathscr{F}$ é o subespaço do produto direto dado por

$$\boxed{+}_{i\in I} V_i \coloneqq \{(v_i)_{i\in I} \in \prod_{i\in I} V_i \mid \text{suporte}(v) \text{ é finito}\}.$$

Um caso especial importante ocorre quando $V_i = V$ para todo $i \in I$. Se V^I denota o conjunto das funções de I em V e $(V^I)_0$ o subconjunto das funções de suporte finito, então

$$\prod_{i\in I} V = V^I \qquad \text{e} \qquad \boxed{+}_{i\in I} V = (V^I)_0$$

Observe que o produto direto e a soma direta externa são os mesmos para uma família de espaços vetoriais finita.

Exemplos 2.70. O espaço vetorial $\mathbb{K}^\infty$ é produto direto de $\mathbb{K}$ e $(\mathbb{K}^\infty)_0$ é soma direta externa:

$$\mathbb{K}^\infty = \prod_{i\in\mathbb{N}} \mathbb{K} \quad \text{e} \quad (\mathbb{K}^\infty)_0 = \boxed{+}_{i\in\mathbb{N}} \mathbb{K}$$

◁

Exercícios

Ex. 2.45 — Mostre que, se $U \subseteq S$, então

$$S \cap (T + U) = (S \cap T) + U$$

Essa propriedade é denominada lei modular do reticulado $\mathcal{S}(V)$.

Ex. 2.46 — Para quais espaços vetoriais a lei distributiva de subespaços

$$S \cap (T + U) = (S \cap T) + (S \cap U)$$

é verdadeira? Determine, portanto, quando o reticulado de subespaços é distributivo.

Ex. 2.47 — Seja S um subespaço de V e seja $v \in V$. O conjunto $v + S = \{v + s : s \in S\}$ é chamado subespaço afim de V.

1. Quando um subespaço afim de V é subespaço de V?

2. Mostre que dois subespaços afins $x + S$ e $y + S$ ou são iguais ou são disjuntos.

Ex. 2.48 — Seja X um subconjunto de um espaço vetorial V sobre $\mathbb{K}$. Mostre que

$$\langle X\rangle = \left\{ \sum_{v\in F} \alpha_v v \,:\, F \subseteq X \text{ é finito e } \alpha_v \in \mathbb{K} \text{ para todo } v \in F \right\}.$$

Ex. 2.49 — Prove que

1. Se $S_1 \subseteq S_2$, $\langle S_1\rangle \subseteq \langle S_2\rangle$.

2. Se $\mathbf{x} \in \langle S\rangle$, existe um subconjunto finito $S' \subseteq S$ de modo que $\mathbf{x} \in \langle S'\rangle$.

3. $S \subseteq \langle S\rangle$ para todos os $S \in \mathcal{P}(V)$.

4. Para cada $S \in \mathcal{P}(V)$, $\langle\langle S\rangle\rangle = \langle S\rangle$.

5. Se $\mathbf{y} \in \langle S \cup \{\mathbf{x}\}\rangle$ e $\mathbf{y} \notin \langle S\rangle$, então $\mathbf{x} \in \langle S \cup \{\mathbf{y}\}\rangle$. Aqui $\mathbf{x}, \mathbf{y} \in V$ e $S \in \mathcal{P}(V)$.

Ex. 2.50 — Dê um contraexemplo que mostre que a união de subespaços não é necessariamente um subespaço.

Ex. 2.51 — Mostre que, se W_1 e W_2 são subespaços de V, então $W_1 \cup W_2$ é um subespaço vetorial se, e somente se, $W_1 \subseteq W_2$ ou $W_2 \subseteq W_1$.

Exercícios

Ex. 2.52 — Seja V um espaço vetorial e sejam $W_1, \dots, W_k$ subespaços de V. Para cada i, seja $\underline{\mathrm{B}}_i$ uma base de W_i e ponha $\underline{\mathrm{B}} = \bigcup_{i=1}^{k} \underline{\mathrm{B}}_i$. Mostre que:

1. $\underline{\mathrm{B}}$ gera V se, e somente se, $V = W_1 + \cdots + W_k$.
2. $\underline{\mathrm{B}}$ é linearmente independente se, e somente se, $\{W_1, \ \dots \ , \ W_k\}$ for independente.
3. $\underline{\mathrm{B}}$ é uma base para V se, e somente se, $V = W_1 \oplus \cdots \oplus W_k$.

Ex. 2.53 — Seja V um espaço vetorial de dimensão n e sejam $W_1, \dots, W_k$ subespaços, com $n_i = \dim W_i$.

1. Se $n_1 + \cdots + n_k > n$, então $\{W_1, \ \dots \ , \ W_k\}$ não é independente.
2. Se $n_1 + \cdots + n_k < n$, então $V \neq W_1 + \cdots + W_k$.
3. Se $n_1 + \cdots + n_k = n$, prove que são equivalentes:

   a)) $V = W_1 \oplus \cdots \oplus W_k$.

   b)) $V = W_1 + \cdots + W_k$

   c)) $\{W_1, \ \dots \ , \ W_k\}$ é independente.

Ex. 2.54 — Seja V um espaço vetorial de dimensão finita e sejam $W_1, \dots, W_k$ subespaços tais que $V = W_1 \oplus \cdots \oplus W_k$. Prove que $\dim V = \dim W_1 + \cdots + \dim W_k$.

Ex. 2.55 — Encontre um subespaço complementar a $L_3 = \{p(x) \in \mathbb{K}_n[x] \ : \ p(1) = 0\}$ em $\mathbb{K}_n[x]$.

Ex. 2.56 — Seja V um espaço vetorial de dimensão finita n e sejam $W_1, \dots, W_k$ subespaços de V. Para $1 \leq i \leq k$, defina $c_i = n - \dim W_i$. Mostre que:

1. $$\dim\left(W_1 \cap \cdots \cap W_k\right) = n - \sum_{i=1}^{k} c_i + \sum_{j=1}^{k-1} \left[n - \dim\left(\left(W_1 \cap \cdots \cap W_j\right) + W_{j+1}\right)\right].$$
2. $\dim\left(W_1 \cap \cdots \cap W_k\right) \geq n - \sum_{i=1}^{k} c_i$.
3. A igualdade $\dim\left(W_1 \cap \cdots \cap W_k\right) = n - \sum_{i=1}^{k} c_i$ vale se, e somente se, para todo

$i = 1, \dots, k$,

$$W_i + (\bigcap_{j \neq i} W_j) = V$$

2.6 Coordenadas

Vetores e coordenadas
A coluna de coordenadas de um vetor depende da base escolhida. Uma mudança de base altera essa coluna, mas não o vetor. Essa distinção também se aplica à representação matricial das transformações lineares.

Suponha que V seja um espaço vetorial de dimensão finita n sobre $\mathbb{K}$.

Definição 2.71 — Base Ordenada. Uma **base ordenada** para V é uma n-tupla ordenada $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ de vetores para os quais o conjunto $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ é uma base para V.

Se $\underline{\mathrm{x}} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ for uma base ordenada para V, então para todo $\mathbf{y} \in V$ haverá uma única n-tupla $(y_1, \dots, y_n)$ de escalares para os quais

$$\mathbf{y} = y_1\mathbf{x}_1 + \cdots + y_n\mathbf{x}_n$$

Assim, se $\underline{\mathrm{x}} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ é uma base ordenada de V, temos uma aplicação natural $[\,\cdot\,]_{\underline{\mathrm{x}}} : V \to \mathbb{K}^n$, denominada aplicação de coordenadas.

Definição 2.72 — Coordenadas de um vetor. Se $\underline{\mathrm{x}} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ é uma base de V, então definimos as **coordenadas** do vetor $\mathbf{y} = \sum_{i=1}^{n} y_i\mathbf{x}_i$ na base $\underline{\mathrm{x}}$ como

$$[\mathbf{y}]_{\underline{\mathrm{x}}} \coloneqq (y_1, \dots, y_n)^{\mathrm{t}} \in \mathbb{K}^n$$

Como $\underline{\mathrm{x}}$ é uma base de V, a representação de um determinado vetor $\mathbf{y}$ como uma combinação linear de $\mathbf{x}_1, \dots, \mathbf{x}_n$ é única e assim temos que a função está bem definida.

A aplicação $[\,\cdot\,]_{\underline{\mathrm{x}}} : V \to \mathbb{K}^n$ é bijetiva e preserva somas e multiplicações por escalares. Em particular,

$$[\lambda_1\mathbf{y} + \lambda_2\mathbf{z}]_{\underline{\mathrm{x}}} = \lambda_1[\mathbf{y}]_{\underline{\mathrm{x}}} + \lambda_2[\mathbf{z}]_{\underline{\mathrm{x}}}.$$

Como veremos em breve, trata-se de uma transformação linear.

Definição 2.73 — Componentes. As entradas do vetor coluna $[\mathbf{y}]_{\underline{\mathrm{x}}}$ são denominadas **componentes** de $\mathbf{y}$ na base $\underline{\mathrm{x}}$.

Exemplo 2.74. Suponha que $\mathrm{char}(\mathbb{K}) \neq 2$. Seja $V = \mathscr{M}_{2,2}(\mathbb{K})$ e considere os vetores

$$\vec{a}_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \vec{a}_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \vec{a}_3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \vec{a}_4 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

As quatro matrizes são linearmente independentes e, como $\dim \mathscr{M}_{2,2}(\mathbb{K}) = 4$, formam uma base ordenada $\underline{\mathrm{A}} = (\vec{a}_1, \vec{a}_2, \vec{a}_3, \vec{a}_4)$. Em particular, toda matriz $A = (a_{ij})$ pode ser escrita como

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \frac{a_{12} + a_{21}}{2}\vec{a}_1 + \frac{a_{12} - a_{21}}{2}\vec{a}_2 + \frac{a_{11} + a_{22}}{2}\vec{a}_3 + \frac{a_{11} - a_{22}}{2}\vec{a}_4$$

Logo,

$$\left[\!\left[\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \right]\!\right]_{\underline{A}} = \begin{bmatrix} \frac{a_{12}+a_{21}}{2} \\ \frac{a_{12}-a_{21}}{2} \\ \frac{a_{11}+a_{22}}{2} \\ \frac{a_{11}-a_{22}}{2} \end{bmatrix}$$

◁

2.6.1 Mudança de Base

Duas bases de um mesmo espaço atribuem, em geral, coordenadas diferentes a um mesmo vetor. A mudança de base consiste em passar de uma dessas colunas de coordenadas à outra. O vetor permanece o mesmo; o que muda são os coeficientes usados para escrevê-lo como combinação linear dos vetores da base.

Exemplo 2.75 — Duas bases de $\mathbb{R}^2$. Considere a base canônica $\underline{E}$ de $\mathbb{R}^2$ e a base $\underline{B} = (\vec{u}_1, \vec{u}_2)$, onde

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \vec{u}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Os dois vetores são linearmente independentes, pois o sistema $a + b = 0$, $a + 2b = 0$ tem apenas a solução $a = b = 0$. Se $\vec{z} = a\vec{u}_1 + b\vec{u}_2$, então

$$[\vec{z}]_{\underline{B}} = \begin{bmatrix} a \\ b \end{bmatrix}, \qquad [\vec{z}]_{\underline{E}} = \begin{bmatrix} a+b \\ a+2b \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}.$$

As colunas dessa matriz são precisamente as coordenadas de $\vec{u}_1$ e $\vec{u}_2$ na base canônica. Por exemplo,

$$\vec{z} = \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \vec{u}_1 + 2\vec{u}_2$$

tem coordenadas $(3, 5)^{\mathrm{t}}$ na base $\underline{E}$ e $(1, 2)^{\mathrm{t}}$ na base $\underline{B}$. São duas representações do mesmo vetor.

◁

Para fazer essa passagem em um espaço qualquer, basta conhecer as coordenadas dos vetores de uma base em relação à outra. Sejam $\underline{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ e $\underline{y} = (\mathbf{y}_1, \dots, \mathbf{y}_n)$ bases ordenadas de V.

Definição 2.76 — Matriz de Mudança de Base. A **matriz de mudança de base** de $\underline{y}$ para $\underline{x}$, denotada por $\mathrm{M}_{\underline{y}\to\underline{x}}$, é a matriz $n \times n$ cujas colunas são as coordenadas dos vetores de $\underline{y}$ na base $\underline{x}$:

$$\mathrm{M}_{\underline{y}\to\underline{x}} \coloneqq \left([\mathbf{y}_1]_{\underline{x}} \;\middle|\; \cdots \;\middle|\; [\mathbf{y}_n]_{\underline{x}} \right). \tag{2.16}$$

O sentido indicado na notação é o da passagem entre coordenadas: $\mathrm{M}_{\underline{y}\to\underline{x}}$ recebe coordenadas na base $\underline{y}$ e fornece coordenadas na base $\underline{x}$. Por isso suas colunas são formadas pelos vetores da base de partida, escritos na base de chegada. No exemplo anterior, a matriz obtida é $\mathrm{M}_{\underline{B}\to\underline{E}}$, e não $\mathrm{M}_{\underline{E}\to\underline{B}}$.

Teorema 2.77 (Mudança de Base).
Para todo $\vec{z} \in V$,

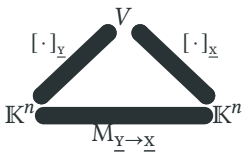
$$\mathrm{M}_{\underline{\mathrm{Y}}\to\underline{\mathrm{X}}}[\vec{z}]_{\underline{\mathrm{Y}}} = [\vec{z}]_{\underline{\mathrm{X}}}.$$

Demonstração. Sejam $c_1, \dots, c_n$ as coordenadas de $\vec{z}$ na base $\underline{\mathrm{Y}}$. Então $\vec{z} = \sum_{i=1}^{n} c_i \mathrm{y}_i$. Como a aplicação de coordenadas preserva combinações lineares,

$$\begin{aligned}
[\vec{z}]_{\underline{\mathrm{X}}} &= \sum_{i=1}^{n} c_i [\mathrm{y}_i]_{\underline{\mathrm{X}}} \\
&= \begin{pmatrix} [\mathrm{y}_1]_{\underline{\mathrm{X}}} & \cdots & [\mathrm{y}_n]_{\underline{\mathrm{X}}} \end{pmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \\
&= \mathrm{M}_{\underline{\mathrm{Y}}\to\underline{\mathrm{X}}}[\vec{z}]_{\underline{\mathrm{Y}}}.
\end{aligned}$$

A segunda igualdade é a interpretação do produto de uma matriz por uma coluna como combinação linear das colunas da matriz. ∎

Podemos resumir essa relação no diagrama



(2.17)

Um diagrama representa espaços e aplicações por objetos e setas. Ele é *comutativo* quando quaisquer dois caminhos com a mesma origem e o mesmo destino definem a mesma aplicação. Aqui, obter diretamente as coordenadas de $\vec{z}$ na base $\underline{\mathrm{X}}$ dá o mesmo resultado que obter primeiro suas coordenadas na base $\underline{\mathrm{Y}}$ e depois multiplicar por $\mathrm{M}_{\underline{\mathrm{Y}}\to\underline{\mathrm{X}}}$.

A passagem no sentido contrário é feita por $\mathrm{M}_{\underline{\mathrm{X}}\to\underline{\mathrm{Y}}}$. Efetuar as duas mudanças sucessivamente devolve a coluna inicial. Como toda coluna de $\mathbb{K}^n$ é a coluna de coordenadas de algum vetor de V, segue que

$$\mathrm{M}_{\underline{\mathrm{X}}\to\underline{\mathrm{Y}}}\,\mathrm{M}_{\underline{\mathrm{Y}}\to\underline{\mathrm{X}}} = I_n, \qquad \mathrm{M}_{\underline{\mathrm{Y}}\to\underline{\mathrm{X}}}\,\mathrm{M}_{\underline{\mathrm{X}}\to\underline{\mathrm{Y}}} = I_n.$$

Portanto, toda matriz de mudança de base é invertível e

$$\mathrm{M}_{\underline{\mathrm{X}}\to\underline{\mathrm{Y}}} = \mathrm{M}_{\underline{\mathrm{Y}}\to\underline{\mathrm{X}}}^{-1}. \tag{2.18}$$

No Exemplo 2.75, por exemplo,

$$\mathrm{M}_{\underline{\mathrm{B}}\to\underline{\mathrm{E}}} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \qquad \mathrm{M}_{\underline{\mathrm{E}}\to\underline{\mathrm{B}}} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}.$$

A segunda matriz permite recuperar as coordenadas na base $\underline{\mathrm{B}}$ a partir das coordenadas usuais: ela leva $(3,5)^{\mathrm{t}}$ a $(1,2)^{\mathrm{t}}$.

Se $\underline{\mathrm{W}}$ é uma terceira base, o mesmo argumento fornece

$$\mathrm{M}_{\underline{\mathrm{Y}}\to\underline{\mathrm{W}}}\,\mathrm{M}_{\underline{\mathrm{X}}\to\underline{\mathrm{Y}}} = \mathrm{M}_{\underline{\mathrm{X}}\to\underline{\mathrm{W}}}. \tag{2.19}$$

O fator da direita faz a primeira mudança, de $\underline{\mathrm{X}}$ para $\underline{\mathrm{Y}}$; o da esquerda faz a segunda, de $\underline{\mathrm{Y}}$ para $\underline{\mathrm{W}}$. O produto realiza a passagem direta de $\underline{\mathrm{X}}$ para $\underline{\mathrm{W}}$.

Exemplo 2.78 — Coordenadas de um polinômio. Seja $V = \mathbb{K}_3[x]$ o espaço dos polinômios de grau menor ou igual a 3. Considere as bases ordenadas

$$\underline{\mathrm{B}} = (p_1, p_2, p_3, p_4) = (1, x, x^2, x^3)$$

e

$$\underline{\mathrm{C}} = (q_1, q_2, q_3, q_4) = (1, x-1, (x-1)^2, (x-1)^3).$$

Para passar de coordenadas na base $\underline{\mathrm{C}}$ para coordenadas na base $\underline{\mathrm{B}}$, escrevemos cada q_i como combinação dos p_j:

$$\begin{aligned} q_1 &= p_1, \\ q_2 &= -p_1 + p_2, \\ q_3 &= p_1 - 2p_2 + p_3, \\ q_4 &= -p_1 + 3p_2 - 3p_3 + p_4. \end{aligned}$$

Os coeficientes de cada linha formam uma *coluna* da matriz:

$$\mathrm{M}_{\underline{\mathrm{C}}\to\underline{\mathrm{B}}} = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Para a passagem inversa, podemos inverter essa matriz ou escrever $x = (x-1) + 1$ e desenvolver as potências de x. Assim,

$$\mathrm{M}_{\underline{\mathrm{B}}\to\underline{\mathrm{C}}} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Por exemplo, para $r(x) = x^2 + 2x + 3$, temos

$$[r]_{\underline{\mathrm{B}}} = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \qquad [r]_{\underline{\mathrm{C}}} = \mathrm{M}_{\underline{\mathrm{B}}\to\underline{\mathrm{C}}}[r]_{\underline{\mathrm{B}}} = \begin{bmatrix} 6 \\ 4 \\ 1 \\ 0 \end{bmatrix}.$$

A segunda coluna de coordenadas expressa a identidade

$$x^2 + 2x + 3 = 6 + 4(x-1) + (x-1)^2.$$

Novamente, o polinômio não mudou; apenas sua expressão na base escolhida.

◁

2.6.2 Espaços Linha e Coluna

Seja $A \in \mathscr{M}_{m,n}(\mathbb{K})$. Suas linhas geram um subespaço de $\mathscr{M}_{1,n}(\mathbb{K})$, denominado **espaço de linhas** de A e denotado por $\mathrm{EspLin}(A)$. Suas colunas geram um subespaço de $\mathbb{K}^m$, denominado **espaço de colunas** de A e denotado por $\mathrm{EspCol}(A)$.

As dimensões desses espaços são denominadas posto por linha e posto por coluna, respectivamente. Denotamos o posto por linha por $\mathrm{posto}_l(A)$ e o posto por coluna por $\mathrm{posto}_c(A)$.

Demonstraremos que o posto por linha é igual ao posto por coluna, embora, quando $m \neq n$, os dois espaços pertençam a ambientes diferentes.

Nossa demonstração desse fato depende da seguinte proposição sobre matrizes.

> **Proposição 2.79.**
> *Seja $A \in \mathscr{M}_{m,n}(\mathbb{K})$. Operações elementares por linhas ou por colunas não alteram nem o posto por linha nem o posto por coluna de A.*

Demonstração. Uma operação elementar por linhas equivale a substituir A por EA, onde E é invertível. As linhas de EA são combinações invertíveis das linhas de A, portanto $\text{EspLin}(EA) = \text{EspLin}(A)$. Além disso, as colunas de EA são as imagens, pelo automorfismo $\mathbf{x} \mapsto E\mathbf{x}$ de $\mathbb{K}^m$, das colunas de A; logo seus espaços têm a mesma dimensão.

De modo análogo, uma operação elementar por colunas substitui A por AF, com F invertível. Ela faz combinações invertíveis das colunas e aplica o automorfismo $\mathbf{x} \mapsto \mathbf{x}F$ de $\mathscr{M}_{1,n}(\mathbb{K})$ às linhas. Assim, preserva os dois postos. ■

> **Teorema 2.80.**
> *Se $A \in \mathscr{M}_{m,n}(\mathbb{K})$, então $\text{posto}_l(A) = \text{posto}_c(A)$. Esse número é o **posto** de A e será denotado por $\text{posto}(A)$.*

Demonstração. Por operações elementares por linhas e colunas, podemos reduzir A a uma matriz em blocos da forma $M = \left[\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right]$. Pela proposição anterior, nenhum dos dois postos se altera durante o processo. As primeiras r linhas e as primeiras r colunas de M são linearmente independentes, e todas as demais são nulas. Portanto,

$$\text{posto}_l(A) = \text{posto}_l(M) = \text{posto}_c(M) = \text{posto}_c(A)$$

como desejado. ■

Exercícios

Ex. 2.57 — Seja $\mathbb{K}_n[x]$ o espaço dos polinômios de grau no máximo n com coeficientes em $\mathbb{K}$. Mostre que:

1. $(1, x, \dots, x^n)$ é uma base de $\mathbb{K}_n[x]$. As coordenadas de um polinômio nessa base são seus coeficientes.

2. $(1, x-a, (x-a)^2, \dots, (x-a)^n)$ é uma base de $\mathbb{K}_n[x]$. Se $\text{char}(\mathbb{K}) = 0$ ou $\text{char}(\mathbb{K}) = p > n$, mostre que

$$[f]_{(1,x-a,\dots,(x-a)^n)} = \left(f(a), f'(a), \frac{f''(a)}{2!}, \dots, \frac{f^{(n)}(a)}{n!}\right)^{\mathrm{t}}.$$

Ex. 2.58 —

1. Verifique que $\underline{B} = (1 + x, 1 + x^2, 1 + 2x - 2x^2)$ é uma base de $\mathbb{K}_2[x]$.
2. Calcule os vetores de coordenadas $[1]_{\underline{B}}$, $[x]_{\underline{B}}$ e $[x^2]_{\underline{B}}$.

Ex. 2.59 — Sejam $\vec{f}_1(x) = -\frac{1}{6}(x-1)(x-2)(x-3)$, $\vec{f}_2(x) = \frac{1}{2}x(x-2)(x-3)$, $\vec{f}_3(x) = -\frac{1}{2}x(x-1)(x-3)$, $\vec{f}_4(x) = \frac{1}{6}x(x-1)(x-2)$.

1. Prove que $\underline{B} = (\vec{f}_1, \vec{f}_2, \vec{f}_3, \vec{f}_4)$ é uma base de $\mathbb{R}_3[x]$.
2. Se $g(x) \in \mathbb{R}_3[x]$, prove que $[g]_{\underline{B}} = \begin{bmatrix} g(0) \\ g(1) \\ g(2) \\ g(3) \end{bmatrix}$

Ex. 2.60 — Seja $\underline{B} = (\vec{f}_1, \vec{f}_2, \vec{f}_3, \vec{f}_4)$ a base de $\mathbb{R}_3[x]$ do exercício anterior. Calcule, em relação a $\underline{B}$, os vetores de coordenadas dos elementos da base canônica $(1, x, x^2, x^3)$.

Ex. 2.61 — Seja $\underline{B}$ uma base de um espaço vetorial V de dimensão finita n sobre $\mathbb{K}$, e sejam $\mathbf{u}_1, \dots, \mathbf{u}_k \in V$. Prove que $\langle \mathbf{u}_1, \dots, \mathbf{u}_k \rangle = V$ se, e somente se, $\langle [\mathbf{u}_1]_{\underline{B}}, \dots, [\mathbf{u}_k]_{\underline{B}} \rangle = \mathbb{K}^n$.

Ex. 2.62 — Seja $\underline{B}$ uma base de um espaço vetorial V de dimensão n sobre $\mathbb{K}$, e sejam $\mathbf{u}_1, \dots, \mathbf{u}_n \in V$. Prove que $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ é uma base de V se, e somente se, $([\mathbf{u}_1]_{\underline{B}}, \dots, [\mathbf{u}_n]_{\underline{B}})$ é uma base de $\mathbb{K}^n$.

2.7 Bandeiras

Subespaços e bases adaptadas
Uma bandeira é uma sequência crescente de subespaços. Uma base adaptada descreve cada um deles por um segmento inicial da base e permite acompanhar essa sequência em coordenadas.

Definição 2.81 — Bandeira. Uma **bandeira** em V é uma sequência ascendente de subespaços $V_0 \subset V_1 \subset \cdots \subset V_n$.

O número n é o **comprimento** da bandeira.

Definição 2.82 — Bandeira Maximal Finita. Uma bandeira finita

$$\{0\} = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_n = V$$

é dita **maximal** se nenhum subespaço puder ser inserido estritamente entre V_i e V_{i+1}. Isto é, se $V_i \subseteq W \subseteq V_{i+1}$, então $W = V_i$ ou $W = V_{i+1}$.

Exemplo 2.83. No espaço $\mathbb{K}_n[x]$ dos polinômios de grau no máximo n, temos a bandeira maximal

$$\{\vec{0}\} \subsetneq \mathbb{K}_0[x] \subsetneq \mathbb{K}_1[x] \subsetneq \cdots \subsetneq \mathbb{K}_n[x].$$

◁

Uma bandeira de comprimento n pode ser construída a partir de uma base ordenada $\underline{\mathrm{E}} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$: basta definir $V_0 = \{\vec{0}\}$ e $V_i = \langle \mathbf{e}_1, \ldots, \mathbf{e}_i \rangle$ para $i = 1, \ldots, n$. Essa é a **bandeira canônica associada** à base $\underline{\mathrm{E}}$.

Essa bandeira é maximal. Reciprocamente, em dimensão finita, toda bandeira maximal é obtida desse modo.

> **Teorema 2.84.**
> *Se V tem dimensão finita, então sua dimensão é igual ao comprimento de qualquer bandeira maximal finita.*

Demonstração. Seja $\{0\} = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_n = V$ uma bandeira maximal. Para cada $i = 1, \ldots, n$, selecione $\vec{e}_i \in V_i \setminus V_{i-1}$. Mostraremos que $\{\vec{e}_1, \ldots, \vec{e}_i\}$ é uma base de V_i.

Como $\langle \mathbf{e}_1, \ldots, \mathbf{e}_{i-1} \rangle \subseteq V_{i-1}$ e $\mathbf{e}_i \notin V_{i-1}$, uma indução mostra que $\{\mathbf{e}_1, \ldots, \mathbf{e}_i\}$ é linearmente independente para todo i.

Mostraremos também, por indução, que esses vetores geram V_i. Suponha que $V_{i-1} = \langle \mathbf{e}_1, \ldots, \mathbf{e}_{i-1} \rangle$ e ponha $V' = \langle \mathbf{e}_1, \ldots, \mathbf{e}_i \rangle$. Então $V_{i-1} \subsetneq V' \subseteq V_i$. Pela maximalidade da bandeira, não há subespaço estritamente intermediário, logo $V' = V_i$.

Para $i = n$, obtemos uma base $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ de V. Portanto, $n = \dim V$. ■

Toda bandeira em um espaço de dimensão finita pode ser estendida a uma bandeira maximal; em particular, seu comprimento não excede $\dim V$. De fato, cada inclusão estrita permite escolher um novo vetor linearmente independente dos anteriores, e esse processo não pode produzir mais que $\dim V$ vetores.

Dizemos que um espaço vetorial satisfaz a **condição da cadeia ascendente** (acc) se toda sequência ascendente de subespaços

$$W_1 \subseteq W_2 \subseteq W_3 \subseteq \cdots$$

se estabiliza, isto é, se existe n tal que $W_n = W_{n+1} = W_{n+2} = \cdots$. De modo análogo, V satisfaz a **condição da cadeia descendente** (dcc) se toda sequência

$$W_1 \supseteq W_2 \supseteq W_3 \supseteq \cdots$$

se estabiliza.

> **Teorema 2.85.**
> *Seja V um espaço vetorial sobre $\mathbb{K}$. Então, são equivalentes:*
>
> a *V é de dimensão finita;*
>
> b *V tem uma bandeira maximal finita;*

c *V satisfaz a condição da cadeia ascendente (acc);*

d *V satisfaz a condição da cadeia descendente (dcc).*

A demonstração das equivalências é proposta nos exercícios a seguir.

Exercícios

Ex. 2.63 — Seja V um espaço vetorial sobre $\mathbb{K}$. Mostre que são equivalentes:

1. V é de dimensão finita;
2. V tem uma bandeira maximal finita;
3. o reticulado de subespaços de V satisfaz a condição da cadeia ascendente;
4. o reticulado de subespaços de V satisfaz a condição da cadeia descendente.

Ex. 2.64 — Seja $0 = V_0 \subsetneq V_1 \subsetneq \ldots \subsetneq V_n = W_1$ uma bandeira maximal para W_1 e $0 = L_0 \subsetneq L_1 \subsetneq \ldots \subsetneq L_m = W_2$ uma bandeira maximal para W_2. Mostre que

$$0 = V_0 \subsetneq V_1 \subsetneq \ldots \subsetneq V_n \subsetneq V_n \oplus L_1 \subsetneq V_n \oplus L_2 \subsetneq \ldots \subsetneq V_n \oplus L_m = W_1 \oplus W_2 = V$$

é uma bandeira maximal para $V = W_1 \oplus W_2$. Conclua, mais uma vez, que a soma direta de espaços vetoriais de dimensão finita tem dimensão igual à soma das dimensões.

3
Transformações Lineares

As transformações lineares são as aplicações compatíveis com a estrutura dos espaços vetoriais: preservam somas e produtos por escalares. Em dimensão finita, escolhidas bases no domínio e no contradomínio, cada transformação linear é representada por uma matriz.

Essa representação permite calcular a ação e a composição das transformações. Para estudar suas propriedades sem depender das coordenadas, consideraremos também o núcleo, a imagem e as decomposições em subespaços invariantes.

Nesse capítulo

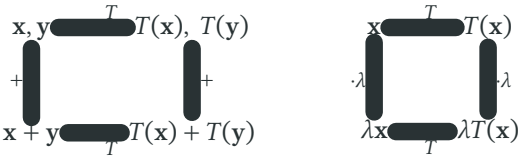
3.1 Definição e Exemplos

Definição 3.1 — Transformação Linear. Sejam V e W espaços vetoriais sobre um corpo $\mathbb{K}$. Uma função $T : V \to W$ é denominada **transformação linear** ou **homomorfismo** se

$$T(\lambda_1 \mathbf{x} + \lambda_2 \mathbf{y}) = \lambda_1 T(\mathbf{x}) + \lambda_2 T(\mathbf{y})$$

para todos os $\lambda_1, \lambda_2 \in \mathbb{K}$ e $\mathbf{x}, \mathbf{y} \in V$.

Determinada por uma base
Uma transformação linear preserva combinações lineares. Consequentemente, depois de escolher uma base do domínio, basta conhecer as imagens dos vetores dessa base: todos os demais valores ficam determinados por linearidade.



Uma transformação linear de V em V é dita **operador linear** em V. Uma transformação linear de V em $\mathbb{K}$ é denominada **funcional linear** em V.

Os exemplos mais elementares são os seguintes.

Exemplos 3.2.

- □ A aplicação $0 : V \to W$ que envia todo vetor de V a $\vec{0} \in W$ é linear e recebe o nome de **transformação nula**.
- □ A transformação linear **identidade** $\mathrm{I}_V : V \to V$ é definida por $\mathrm{I}_V(\mathbf{v}) = \mathbf{v}$ para todo $\mathbf{v} \in V$. Quando não houver risco de confusão, escreveremos simplesmente I.

◁

Em geral, uma aplicação que preserva uma estrutura algébrica é chamada de homomorfismo; daí a expressão *homomorfismo linear*. Adotaremos ainda as designações seguintes.

Definição 3.3.
- □ **endomorfismo**, para um operador linear;
- □ **monomorfismo**, para uma transformação linear injetiva;
- □ **epimorfismo**, para uma transformação linear sobrejetiva;
- □ **isomorfismo**, para uma transformação linear bijetiva;
- □ **automorfismo**, para um operador linear bijetivo.

Exemplo 3.4. Se V tem dimensão finita e base ordenada $\underline{\mathrm{x}} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$, então a aplicação de coordenadas $[\,\cdot\,]_{\underline{\mathrm{x}}} : V \to \mathbb{K}^n$ é linear e bijetiva; portanto, é um isomorfismo.

◁

Exemplo 3.5. A transposição $A \mapsto A^{\mathrm{t}}$ é uma aplicação linear de $\mathscr{M}_{m,n}(\mathbb{K})$ em $\mathscr{M}_{n,m}(\mathbb{K})$.

◁

Exemplo 3.6. Se $V = \mathscr{M}_{m,n}(\mathbb{K})$ e $A \in \mathscr{M}_{m,m}(\mathbb{K})$, a multiplicação à esquerda por A define a transformação linear $\mathrm{L}_A : V \to V$ dada por $\mathrm{L}_A(B) = AB$.

◁

Os Exemplos 3.4 e 3.6 mostram que as aplicações do diagrama comutativo 2.17 são lineares.

Exemplo 3.7. Seja $V = C^\infty(\mathbb{R})$ o espaço vetorial das funções $f : \mathbb{R} \to \mathbb{R}$ infinitamente diferenciáveis.

1. Para $a \in \mathbb{R}$, a avaliação em a é o funcional $\mathrm{E}_a : V \to \mathbb{R}$ definido por $\mathrm{E}_a(f) = f(a)$. Também é linear o operador $\overline{\mathrm{E}}_a : V \to V$ que associa a f a função constante de valor $f(a)$.
2. A diferenciação $\mathrm{D} : V \to V$, definida por $\mathrm{D}(f) = f'$, é uma transformação linear.

3 Para um número real a, seja $\mathrm{I}_a : V \to V$ a integração com ponto inicial $t = a$, isto é,

$$\mathrm{I}_a(f)(x) = \int_a^x f(t)\,dt.$$

Então I_a é uma transformação linear.

4 Também temos o funcional linear $I_a^b = \mathrm{E}_b \circ \mathrm{I}_a$. Dessa forma,

$$I_a^b(f) = (\mathrm{E}_b \circ \mathrm{I}_a)(f) = \int_a^b f(t)\,dt.$$

◁

Teorema 3.8 (Fundamental do Cálculo). 1 $\mathrm{D} \circ \mathrm{I}_a = \mathrm{I}$;

2 $\mathrm{I}_a \circ \mathrm{D} = \mathrm{I} - \overline{\mathrm{E}}_a$.

Exemplo 3.9. Em $V = \mathbb{K}[x]$, definimos a derivada formal por

$$D\left(\sum_{i=0}^{n} a_i x^i\right) = \sum_{i=1}^{n} i a_i x^{i-1}.$$

Essa aplicação, denominada **derivação formal**, é linear.

◁

Exemplo 3.10. Seja $V = \mathscr{R}(B)$, como no Exemplo 2.14. A aplicação $T : V \to \mathbb{R}$ definida por $T(f) = \int_B f\,dA$ é linear.

◁

Exemplo 3.11. No espaço vetorial das funções integráveis $f : \mathbb{R} \to \mathbb{R}$ para as quais $xf(x)$ também é integrável, a aplicação

$$M_1(f) = \int_{\mathbb{R}} xf(x)\,dx$$

é um funcional linear. Quando f é uma densidade de probabilidade de uma variável aleatória X, seu valor é $M_1(f) = \mathbb{E}[X]$. Observe que o conjunto das densidades, isoladamente, não é um espaço vetorial.

◁

Exemplo 3.12. Seja $V = \mathbb{K}^\infty$ o espaço vetorial definido no Exemplo 2.6. Definimos os deslocamentos à esquerda e à direita, respectivamente, por

$$\mathrm{L}([a_1,\ a_2,\ a_3,\ \ldots]) = [a_2,\ a_3,\ a_4,\ \ldots],$$

$$\mathrm{R}([a_1,\ a_2,\ a_3,\ \ldots]) = [0,\ a_1,\ a_2,\ \ldots].$$

As aplicações $\mathrm{L}, \mathrm{R} : V \to V$ são lineares.

Podemos restringir L e R ao subespaço $W = (\mathbb{K}^\infty)_0$ definido no Exemplo 2.38. Assim, $\mathrm{L} : W \to W$ e $\mathrm{R} : W \to W$ também são transformações lineares.

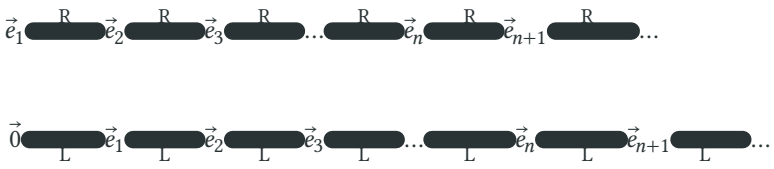
Nesse mesmo exemplo, vimos que $\underline{B} = (\vec{e}_i \mid i \in \mathbb{N})$ é uma base de W, onde

$$\vec{e}_i = \underbrace{(0, \dots, 0, 1, 0, \dots)}_{1 \text{ na } i\text{-ésima posição}}.$$

Os deslocamentos agem nos vetores da base da seguinte maneira:

$$\mathrm{R}(\vec{e}_i) = \vec{e}_{i+1} \tag{3.1}$$

$$\mathrm{L}(\vec{e}_i) = \begin{cases} \vec{e}_{i-1} & \text{se } i > 1 \\ \vec{0} & \text{se } i = 1 \end{cases} \tag{3.2}$$



◁

Proposição 3.13.
Sejam $T, S : V \to W$ e $L : W \to Z$ transformações lineares e sejam $\lambda_1, \lambda_2 \in \mathbb{K}$. Então:

a $T(\vec{0}) = \vec{0}$.

b *A composta $L \circ T : V \to Z$ é uma transformação linear.*

c $\lambda_1 T + \lambda_2 S$ *é uma transformação linear de V em W.*

d *Se T for invertível, então $T^{-1} : W \to V$ é uma transformação linear.*

Demonstração. a Para qualquer $\mathbf{v} \in V$, temos $T(\vec{0}) = T(0\mathbf{v}) = 0T(\mathbf{v}) = \vec{0}$.

b Se $a, b \in \mathbb{K}$ e $\mathbf{x}, \mathbf{y} \in V$, então

$$(L \circ T)(a\mathbf{x} + b\mathbf{y}) = L\big(aT(\mathbf{x}) + bT(\mathbf{y})\big) = a(L \circ T)(\mathbf{x}) + b(L \circ T)(\mathbf{y}).$$

c Se $a, b \in \mathbb{K}$ e $\mathbf{x}, \mathbf{y} \in V$, então

$$\begin{aligned}
(\lambda_1 T + \lambda_2 S)(a\mathbf{x} + b\mathbf{y}) &= \lambda_1 T(a\mathbf{x} + b\mathbf{y}) + \lambda_2 S(a\mathbf{x} + b\mathbf{y}) && (3.3)\\
&= \lambda_1 aT(\mathbf{x}) + \lambda_1 bT(\mathbf{y}) + \lambda_2 aS(\mathbf{x}) + \lambda_2 bS(\mathbf{y}) && (3.4)\\
&= a(\lambda_1 T(\mathbf{x}) + \lambda_2 S(\mathbf{x})) + b(\lambda_1 T(\mathbf{y}) + \lambda_2 S(\mathbf{y})) && (3.5)\\
&= a(\lambda_1 T + \lambda_2 S)(\mathbf{x}) + b(\lambda_1 T + \lambda_2 S)(\mathbf{y}). && (3.6)
\end{aligned}$$

Portanto, $\lambda_1 T + \lambda_2 S$ é linear.

d Como T é bijetiva, quaisquer $\mathbf{y}_1, \mathbf{y}_2 \in W$ podem ser escritos como $\mathbf{y}_1 = T(\mathbf{x}_1)$ e $\mathbf{y}_2 = T(\mathbf{x}_2)$. Assim,

$$\begin{aligned}
T^{-1}(\lambda_1 \mathbf{y}_1 + \lambda_2 \mathbf{y}_2) &= T^{-1}(\lambda_1 T(\mathbf{x}_1) + \lambda_2 T(\mathbf{x}_2)) && (3.7)\\
&= T^{-1}(T(\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2)) && (3.8)\\
&= \lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 && (3.9)\\
&= \lambda_1 T^{-1}(\mathbf{y}_1) + \lambda_2 T^{-1}(\mathbf{y}_2). && (3.10)
\end{aligned}$$

Isso mostra que T^{-1} é linear. ■

O conjunto de todas as transformações lineares de V em W também possui uma estrutura natural de espaço vetorial.

Definição 3.14 — Espaço das Transformações Lineares. Sejam V e W espaços vetoriais sobre $\mathbb{K}$. O espaço vetorial de todas as transformações lineares de V em W será denotado por $\mathrm{Hom}_{\mathbb{K}}(V,W)$ ou $\mathscr{L}_{\mathbb{K}}(V,W)$. O espaço dos operadores lineares em V será denotado por $\mathrm{Hom}_{\mathbb{K}}(V,V)$.

Quando o corpo base estiver claro no contexto, escreveremos simplesmente $\mathrm{Hom}(V,W)$. Esse espaço é o subespaço de W^V, definido no Exemplo 2.8, formado pelas funções lineares de V em W.

Teorema 3.15.
$\mathrm{Hom}(V,W)$ *é um subespaço vetorial de* W^V.

Demonstração. A transformação nula pertence a $\mathrm{Hom}(V,W)$. Se $S,T \in \mathrm{Hom}(V,W)$ e $a,b \in \mathbb{K}$, o item [c] da proposição anterior mostra que $aS+bT$ também é linear. O teste de subespaço conclui a demonstração. ■

Como toda $T \in \mathrm{Hom}(V,W)$ satisfaz $T(\vec{0}) = \vec{0}$, o espaço $\mathrm{Hom}(V,W)$ é um subespaço próprio de W^V sempre que $W \neq \{\vec{0}\}$.

A cada transformação linear estão associados dois subespaços fundamentais.

Definição 3.16 — Núcleo e Imagem. Seja $T \in \mathrm{Hom}(V,\ W)$.

[1] O subespaço
$$\ker T = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \vec{0}\}$$
é denominado **núcleo** de T;

[2] O subespaço
$$\mathrm{im}\, T = \{T(\mathbf{v}) \mid \mathbf{v} \in V\}$$
é denominado **imagem** de T.

[3] A dimensão de $\ker T$ é a **nulidade** de T e é denotada por $\mathrm{nul}\, T$;

[4] A dimensão de $\mathrm{im}\, T$ é o **posto** de T e é denotada por $\mathrm{posto}\, T$.

Segue diretamente da linearidade que $\ker T$ é um subespaço de V e que $\mathrm{im}\, T$ é um subespaço de W.

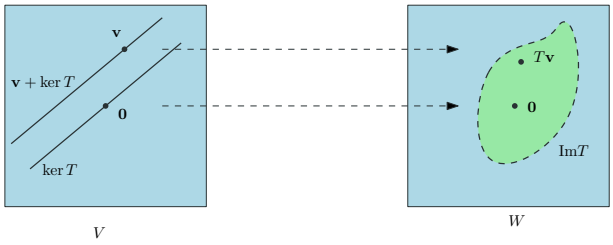


Figura 3.1 Uma aplicação linear $T : V \to W$ envia todos os vetores da classe $\mathbf{v} + \ker T$ ao mesmo ponto $T(\mathbf{v})$. Classes distintas têm imagens distintas, e o conjunto dessas imagens é $\mathrm{im}\, T$. Em particular, a classe $\vec{0} + \ker T$ é enviada à origem de W.

Teorema 3.17.
Seja $T \in \mathrm{Hom}(V, W)$. Então

a *T é injetiva, e somente se, $\ker T = \{\vec{0}\}$;*

b *T é sobrejetiva se, e somente se, $\mathrm{im}\, T = W$.*

Demonstração. Para a primeira afirmação, observe que

$$T(\mathbf{u}) = T(\mathbf{v}) \iff T(\mathbf{u} - \mathbf{v}) = \vec{0} \iff \mathbf{u} - \mathbf{v} \in \ker T.$$

Se $\ker T = \{\vec{0}\}$, essa equivalência mostra que T é injetiva. Reciprocamente, se T é injetiva e $\mathbf{u} \in \ker T$, então $T(\mathbf{u}) = T(\vec{0})$ e, portanto, $\mathbf{u} = \vec{0}$.

A segunda afirmação é direta da definição de sobrejetividade. ■

O teorema seguinte descreve como definir uma transformação linear a partir das imagens dos vetores de uma base.

Teorema 3.18 (da Extensão por Linearidade).
Sejam V e W espaços vetoriais sobre $\mathbb{K}$ e suponha que $\underline{x} = (\mathbf{x}_i \mid i \in I)$ seja uma base de V. Para qualquer família $(\mathbf{y}_i \mid i \in I)$ de vetores de W, existe uma única $T \in \mathrm{Hom}(V, W)$ tal que $T(\mathbf{x}_i) = \mathbf{y}_i$ para todo $i \in I$.

Demonstração. Todo vetor $\mathbf{x} \in V$ possui uma expressão única $\mathbf{x} = \sum_{i \in F} x_i \mathbf{x}_i$, em que $F \subseteq I$ é finito. Defina

$$T(\mathbf{x}) = \sum_{i \in F} x_i \mathbf{y}_i.$$

A unicidade das coordenadas mostra que a definição é bem posta. A distributividade das somas finitas mostra que T é linear e, tomando $\mathbf{x} = \mathbf{x}_i$, obtemos $T(\mathbf{x}_i) = \mathbf{y}_i$. Se S é outra transformação linear com os mesmos valores na base, então

$$S(\mathbf{x}) = \sum_{i \in F} x_i S(\mathbf{x}_i) = \sum_{i \in F} x_i \mathbf{y}_i = T(\mathbf{x}),$$

logo $S = T$. ■

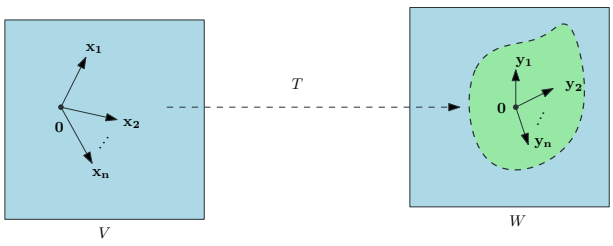


Figura 3.2 Uma aplicação linear $T : V \to W$ fica completamente determinada pelos valores que assume em uma base do domínio.

O teorema anterior admite a seguinte generalização para conjuntos linearmente independentes.

Teorema 3.19 (da Extensão para Conjuntos Linearmente Independentes). *Sejam V e W espaços vetoriais sobre $\mathbb{K}$ e suponha que $(\mathbf{x}_i \mid i \in I)$ seja uma família linearmente independente em V. Para qualquer família $(\mathbf{y}_i \mid i \in I)$ de vetores de W, existe $T \in \mathrm{Hom}(V, W)$ tal que $T(\mathbf{x}_i) = \mathbf{y}_i$ para todo $i \in I$.*

Demonstração. Pelo Teorema da Extensão de Bases, o conjunto $\{\mathbf{x}_i \mid i \in I\}$ pode ser completado a uma base $\underline{\mathrm{B}}$ de V. Associe a $\mathbf{x}_i$ o vetor $\mathbf{y}_i$ e associe o vetor nulo a cada elemento acrescentado à base. O Teorema 3.18 fornece uma transformação linear com os valores prescritos. ■

Essa transformação não é necessariamente única.

Teorema 3.20. *Sejam V e W espaços vetoriais sobre $\mathbb{K}$, com $\dim V = n < \infty$. Cada base ordenada $\underline{\mathrm{x}}$ de V determina um isomorfismo $\tau(\underline{\mathrm{x}}) : \mathrm{Hom}(V, W) \to W^n$.*

Demonstração. Seja $\underline{\mathrm{x}} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ uma base ordenada de V. Defina $\tau(\underline{\mathrm{x}}) : \mathrm{Hom}(V, W) \to W^n$ por

$$\tau(\underline{\mathrm{x}})(T) = (T(\mathbf{x}_1), \dots, T(\mathbf{x}_n)).$$

A linearidade é verificada componente a componente. Defina $\phi : W^n \to \mathrm{Hom}(V, W)$ por

$$\phi(\mathbf{y}_1, \dots, \mathbf{y}_n) = T,$$

onde T é a única transformação linear que satisfaz $T(\mathbf{x}_i) = \mathbf{y}_i$. Pelo Teorema 3.18, as composições $\tau(\underline{\mathrm{x}}) \circ \phi$ e $\phi \circ \tau(\underline{\mathrm{x}})$ são as identidades nos respectivos espaços. Portanto, $\tau(\underline{\mathrm{x}})$ é um isomorfismo. ■

3.1.1 Transformações de $\mathbb{K}^n$ em $\mathbb{K}^m$

Para qualquer matriz $A \in \mathscr{M}_{m,n}(\mathbb{K})$, a aplicação de multiplicação

$$T_A(\mathbf{v}) = A\mathbf{v}$$

é uma transformação linear de $\mathbb{K}^n$ em $\mathbb{K}^m$.

Reciprocamente, toda transformação linear $T \in \mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^m)$ é dada pela multiplicação por uma matriz. Para verificar isso, fixemos as bases canônicas $\underline{\mathrm{E}} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ de $\mathbb{K}^n$ e $\underline{\mathrm{F}} = (\vec{f}_1, \dots, \vec{f}_m)$ de $\mathbb{K}^m$ e definamos

$$[T]_{\underline{\mathrm{E}},\underline{\mathrm{F}}} := \left([T\mathbf{e}_1]_{\underline{\mathrm{F}}} \;\middle|\; \cdots \;\middle|\; [T\mathbf{e}_n]_{\underline{\mathrm{F}}} \right).$$

Assim, para cada vetor da base canônica do domínio,

$$[T]_{\underline{\mathrm{E}},\underline{\mathrm{F}}} [\mathbf{e}_i]_{\underline{\mathrm{E}}} = [T\mathbf{e}_i]_{\underline{\mathrm{F}}}.$$

Consequentemente, $T = T_A$, onde

$$A = [T]_{\underline{\mathrm{E}},\underline{\mathrm{F}}} = \left[T\mathbf{e}_1 \;\middle|\; \cdots \;\middle|\; T\mathbf{e}_n \right].$$

Teorema 3.21. [a] *Se A é uma matriz $m \times n$ sobre $\mathbb{K}$, então $T_A \in \mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^m)$.*

[b] *Se $T \in \mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^m)$, então $T = T_A$, em que*

$$A = \left[T(\mathbf{e}_1) \,\middle|\, \cdots \,\middle|\, T(\mathbf{e}_n) \right].$$

*A matriz A é denominada **matriz de T na base canônica**.*

Exemplos 3.22. Consideremos algumas transformações de $\mathbb{R}^2$ em $\mathbb{R}^2$ e suas interpretações geométricas.

[1] A homotetia $H_\lambda(\mathbf{v}) = \lambda \mathbf{v}$ multiplica todo vetor por λ. Em qualquer base, sua matriz é

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

[2] A dilatação anisotrópica $H_{\lambda,\mu}$ multiplica as direções canônicas por fatores possivelmente diferentes: $H_{\lambda,\mu}(\mathbf{e}_1) = \lambda \mathbf{e}_1$ e $H_{\lambda,\mu}(\mathbf{e}_2) = \mu \mathbf{e}_2$. Sua matriz na base canônica é

$$\begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}.$$

[3] A rotação de ângulo θ, denotada por R_θ, tem matriz

$$[R_\theta] = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

na base canônica.

[4] A reflexão em relação ao eixo x fixa $\mathbf{e}_1$ e envia $\mathbf{e}_2$ em $-\mathbf{e}_2$. Sua matriz na base canônica é

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

[5] Um cisalhamento do plano fixa uma direção $\mathbf{e}_1$ e desloca a outra por um múltiplo da primeira. Em uma base $(\mathbf{e}_1, \mathbf{e}_2)$, ele é determinado por $T(\mathbf{e}_1) = \mathbf{e}_1$ e $T(\mathbf{e}_2) = \mathbf{e}_2 + m\mathbf{e}_1$. Nessa base, sua matriz é

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

◁

O Teorema 3.21 estende-se a $(\mathbb{K}^\infty)_0$, mas as matrizes envolvidas são infinitas. Cada coluna de uma matriz que representa um operador em $(\mathbb{K}^\infty)_0$ deve ter suporte finito, pois ela contém as coordenadas da imagem de um vetor da base. Não se impõe a mesma condição às linhas: como $\mathbf{v}$ tem suporte finito, cada coordenada de $A\mathbf{v}$ é uma soma finita. Além disso, $A\mathbf{v}$ tem suporte finito, pois é combinação linear de um número finito de colunas de suporte finito. O produto de duas matrizes desse tipo também está bem definido e representa a composição dos operadores correspondentes.

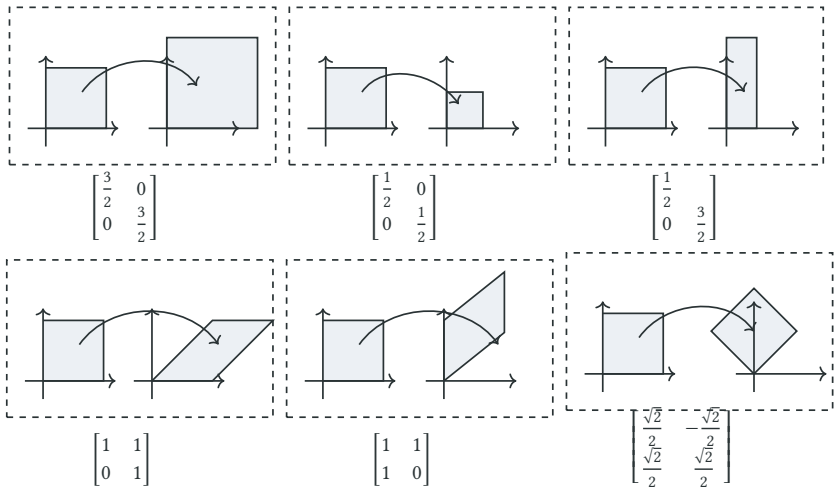


Figura 3.3 Transformações dos Exemplos 3.22.

Exemplo 3.23. O deslocamento para a esquerda admite a representação matricial

$$\mathrm{L} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 0 & \ddots & \\ \vdots & \vdots & & & & \ddots \end{bmatrix}$$

◁

Exercícios

Ex. 3.1 — Seja $T : U \to V$ uma transformação linear. Mostre que:

1. $T(\vec{0}_U) = \vec{0}_V$, onde $\vec{0}_U$ e $\vec{0}_V$ são os vetores nulos de U e V, respectivamente;
2. $T(-\mathbf{u}) = -T(\mathbf{u})$ para todo $\mathbf{u} \in U$;
3. $$T\left(\sum_{i=1}^{m} \alpha_i \mathbf{u}_i\right) = \sum_{i=1}^{m} \alpha_i T(\mathbf{u}_i),$$
   para quaisquer $\alpha_i \in \mathbb{K}$ e $\mathbf{u}_i \in U$.

Ex. 3.2 — Considere o operador de derivação formal $D : \mathbb{R}_n[x] \to \mathbb{R}_n[x]$, definido por $D(p) = p'$.

1. Prove que D é linear.
2. Prove que $D^{n+1} = 0$. Um operador T tal que $T^k = 0$ para algum $k \geq 1$ é denominado **nilpotente**.
3. Calcule $\ker D^k$ para $k = 1, \ldots, n+1$.

Ex. 3.3 — Considere o deslocamento à esquerda $\mathrm{L} : (\mathbb{K}^\infty)_0 \to (\mathbb{K}^\infty)_0$, definido por

$$\mathrm{L}(a_1, a_2, a_3, \ldots) = (a_2, a_3, a_4, \ldots).$$

1. Mostre que L é linear.
2. Calcule $\ker \mathrm{L}^k$ para $k \in \mathbb{N}$.
3. Mostre que, para todo $\mathbf{v} \in (\mathbb{K}^\infty)_0$, existe $k \in \mathbb{N}$ tal que $\mathrm{L}^k(\mathbf{v}) = \vec{0}$, embora L não seja nilpotente.

Ex. 3.4 — Seja $V = L_1 \oplus L_2$. Mostre que os operadores seguintes são lineares.

1. A projeção $\sigma : V \to V$ sobre L_1 ao longo de L_2, definida por
$$\sigma(\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{v}_1, \qquad \mathbf{v}_1 \in L_1, \quad \mathbf{v}_2 \in L_2.$$
2. A reflexão $\tau : V \to V$ em relação a L_1 na direção de L_2, definida por
$$\tau(\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{v}_1 - \mathbf{v}_2.$$

Ex. 3.5 — Sejam V e W espaços vetoriais sobre $\mathbb{K}$, seja $(\mathbf{x}_i \mid i \in \Delta)$ uma família linearmente independente em V e seja $(\mathbf{y}_i \mid i \in \Delta)$ uma família de vetores de W.

1. Mostre que existe $T \in \mathrm{Hom}(V, W)$ tal que $T(\mathbf{x}_i) = \mathbf{y}_i$ para todo $i \in \Delta$.
2. Dê exemplos que mostrem que T não é necessariamente única.

Ex. 3.6 — Demonstre as afirmações seguintes.

1. Se $T, G : U \to V$ são lineares, então $T + G : U \to V$ é linear.
2. Se $T : U \to V$ é linear e $\alpha \in \mathbb{K}$, então $\alpha T : U \to V$ é linear.
3. Se $T : U \to V$ e $G : V \to W$ são lineares, então $G \circ T : U \to W$ é linear.

Ex. 3.7 — No conjunto $V = (0, \infty)$, defina

$$x \oplus y \coloneqq xy \quad \text{e} \quad \alpha \odot x \coloneqq x^\alpha, \qquad x, y \in V, \quad \alpha \in \mathbb{R}.$$

Verifique que $(V, \oplus, \odot)$ é um espaço vetorial sobre $\mathbb{R}$ e que $T : V \to \mathbb{R}$, dada por $T(x) = \ln x$, é linear em relação a essas operações.

Ex. 3.8 — Seja $T : U \to V$ uma transformação linear e sejam $U' \subseteq U$ e $V' \subseteq V$ subespaços. Mostre que $T(U')$ é um subespaço de V e que $T^{-1}(V')$ é um subespaço de U. Conclua que $\operatorname{im} T$ é um subespaço de V e que $\ker T$ é um subespaço de U.

Ex. 3.9 — Sejam V e W espaços vetoriais sobre $\mathbb{K}$ e seja $T : V \to W$ uma função. Mostre que T é linear se, e somente se, seu gráfico

$$G_T = \{(\mathbf{x}, T(\mathbf{x})) \mid \mathbf{x} \in V\}$$

é um subespaço de $V \times W$.

Ex. 3.10 — Seja $\mathbb{R}_n[x]$ o espaço dos polinômios reais de grau menor ou igual a n. Para um número real fixo $h \neq 0$, defina o operador diferença

$$A_h(p)(x) \coloneqq \frac{p(x+h) - p(x)}{h}.$$

Determine o núcleo e a imagem de A_h.

Ex. 3.11 — Seja $T : U \to U$ um operador linear idempotente, isto é, $T^2 = T$. Defina

$$W = \{\mathbf{x} \in U \mid T(\mathbf{x}) = \mathbf{x}\} \quad \text{e} \quad V = \ker T.$$

Prove que:

1. $U = W \oplus V$;
2. $T(U) = W$;
3. $T(V) = \{\vec{0}\}$.

Ex. 3.12 — Mostre que existe uma função $f : \mathbb{R} \to \mathbb{R}$ tal que $f(x+y) = f(x) + f(y)$, $f(x) = 0$ para todo $x \in \mathbb{Q}$ e $f(\sqrt{2}) = 1$.

Ex. 3.13 — Determine todos os funcionais lineares em $\mathbb{Z}_2^3$. Qual é a dimensão do espaço dual $(\mathbb{Z}_2^3)^*$?

Ex. 3.14 — Seja $T \in \mathrm{Hom}(V, V)$ e defina

$$L = \{\mathbf{v} \in V \mid f(T(\mathbf{v})) = 0 \text{ para todo } f \in V^*\}.$$

Prove que $L = \ker T$.

Ex. 3.15 — Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por

$$T(x_1, x_2, x_3) = (x_1 - x_2 + 2x_3,\ 2x_1 + x_2,\ -x_1 - 2x_2 + 2x_3).$$

1. Verifique que T é linear.
2. Determine a imagem de T.
3. Determine o posto de T.

Ex. 3.16 — Seja $\mathscr{M}_{n,n}(\mathbb{K})$ o espaço das matrizes $n\times n$ sobre $\mathbb{K}$ e fixe $B \in \mathscr{M}_{n,n}(\mathbb{K})$. Prove que

$$T : \mathscr{M}_{n,n}(\mathbb{K}) \to \mathscr{M}_{n,n}(\mathbb{K}), \qquad T(A) = AB - BA,$$

é linear. Descreva $\ker T$ e expresse $\operatorname{posto} T$ em termos da dimensão do espaço das matrizes que comutam com B.

Ex. 3.17 — Sejam U, V, W espaços vetoriais de dimensão finita e sejam $T : U \to V$ e $S : V \to W$ transformações lineares. Prove que

$$\dim \ker (S \circ T) \leq \dim \ker S + \dim \ker T.$$

3.2 Isomorfismos

Definição 3.24 — Isomorfismo. Uma transformação linear bijetiva $T : V \to W$ é denominada **isomorfismo** de V em W. Quando existe tal aplicação, dizemos que V e W são isomorfos e escrevemos $V \cong W$.

Dois espaços vetoriais isomorfos são indistinguíveis do ponto de vista de sua estrutura linear: um isomorfismo permite transportar de um para o outro somas, multiplicações por escalares e todas as relações construídas com essas operações.

Exemplo 3.25.
1. A identidade $\mathrm{I}_V : V \to V$ é um isomorfismo; logo $V \cong V$.
2. Se $T : V \to W$ é um isomorfismo, então $T^{-1} : W \to V$ também é; logo $V \cong W$ implica $W \cong V$.
3. Se $T : V \to W$ e $S : W \to Z$ são isomorfismos, então $S \circ T : V \to Z$ é um isomorfismo; logo $V \cong W$ e $W \cong Z$ implicam $V \cong Z$.

Portanto, o isomorfismo define uma relação de equivalência entre espaços vetoriais sobre um mesmo corpo.

◁

Exemplo 3.26. Pela convenção adotada neste texto, $\mathbb{K}^n = \mathscr{M}_{n,1}(\mathbb{K})$ é o espaço dos vetores coluna. A transposição fornece um isomorfismo entre colunas e linhas:

$$\mathbb{K}^n = \mathscr{M}_{n,1}(\mathbb{K}) \longrightarrow \mathscr{M}_{1,n}(\mathbb{K}), \qquad A \longmapsto A^{\mathrm{t}}.$$

◁

Notação 3.27. Os elementos de $\mathbb{K}^n$ serão sempre representados como vetores coluna:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{K}^n.$$

Um vetor linha pertence a $\mathscr{M}_{1,n}(\mathbb{K})$ e não será identificado literalmente com um elemento de $\mathbb{K}^n$.

Exemplo 3.28. O Exemplo 3.4 mostra que todo espaço vetorial V de dimensão n sobre $\mathbb{K}$ é isomorfo a $\mathbb{K}^n$:

$$V \cong \mathbb{K}^n.$$

Consequentemente, todos os espaços vetoriais de dimensão n sobre $\mathbb{K}$ são isomorfos entre si. Em outras palavras, a menos de isomorfismo, existe um único espaço vetorial de cada dimensão finita sobre $\mathbb{K}$.

◁

Recordemos a descrição do espaço V^n introduzido no Exemplo 2.8.

Exemplo 3.29. Seja V um espaço vetorial sobre $\mathbb{K}$ e seja $n \in \mathbb{N}$. Considere a soma direta

$$\underbrace{V \boxplus \cdots \boxplus V}_{n \text{ vezes}}.$$

Para $A = \{1, \dots, n\}$, o espaço V^A é denotado simplesmente por V^n.

Existe um isomorfismo natural $T : V \boxplus \cdots \boxplus V \to V^n$ que envia $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ à função f definida por $f(i) = \mathbf{x}_i$. A soma e a multiplicação por escalares são calculadas componente a componente, e a inversa envia f para $(f(1), \dots, f(n))$. Portanto,

$$V^n \cong \underbrace{V \boxplus \cdots \boxplus V}_{n \text{ vezes}}$$

Observe que esse isomorfismo não exige a escolha de bases.

◁

Notação 3.30 — Isomorfismo natural. O termo **natural** possui um significado preciso na Teoria das Categorias. Aqui, ele indicará que a construção não depende da escolha de bases. Chamaremos de **acidental** um isomorfismo que dependa de tal escolha.

Usaremos **canônico** para uma escolha ou representação consagrada, como a base canônica de $\mathbb{R}^n$. O uso desse termo varia na literatura e, por vezes, coincide com o de "natural".

A partir deste ponto, usaremos a identificação natural entre a soma direta de n cópias de V e V^n. Se A for apenas um conjunto de cardinalidade n, uma identificação $V^A \cong V^n$ exigirá a escolha de uma enumeração de A.

Exemplo 3.31. O espaço $(\mathbb{K}^\infty)_0$ das sequências sobre $\mathbb{K}$ com suporte finito é isomorfo a $\mathbb{K}[x]$:

$$\mathbb{K}[x] \cong (\mathbb{K}^\infty)_0$$

Um isomorfismo é dado por

$$T(a_0 + a_1 x + \cdots + a_n x^n) = (a_0, a_1, \dots, a_n, 0, \dots).$$

◁

Espaços vetoriais isomorfos compartilham todas as propriedades expressas em termos de sua estrutura linear. Se $T \in \mathrm{Hom}(V, W)$ e $S \subseteq V$, escreveremos

$$T(S) = \{T(\mathbf{s}) \mid \mathbf{s} \in S\}.$$

Teorema 3.32.
Sejam $T \in \mathrm{Hom}(V, W)$ um isomorfismo e $S \subseteq V$. Então:

- a *S gera V se, e somente se, $T(S)$ gera W;*
- b *S é linearmente independente em V se, e somente se, $T(S)$ é linearmente independente em W.*

Demonstração. Como T e T^{-1} preservam combinações lineares, temos $T(\langle S\rangle) = \langle T(S)\rangle$. Isso prova a primeira equivalência. Para a segunda, uma relação linear entre elementos de S é enviada por T a uma relação com os mesmos coeficientes entre suas imagens; aplicando T^{-1}, obtemos a recíproca. ■

Em particular, os isomorfismos podem ser caracterizados por seu efeito sobre as bases.

Teorema 3.33.
Sejam V e W espaços vetoriais sobre $\mathbb{K}$. Então:

- a *uma transformação linear $T \in \mathrm{Hom}(V, W)$ é um isomorfismo se, e somente se, envia alguma base de V a uma base de W. Nesse caso, envia toda base de V a uma base de W;*
- b *$V \cong W$ se, e somente se, $\dim V = \dim W$.*

Demonstração. Se T é um isomorfismo e $\underline{\mathrm{B}}$ é uma base de V, o teorema anterior mostra que $T(\underline{\mathrm{B}})$ é uma base de W. Reciprocamente, se T envia alguma base de V a uma base de W, então o mesmo teorema mostra que T é injetiva e sobrejetiva. Isso prova a.

Para b, um isomorfismo envia uma base de V a uma base de W e, portanto, preserva sua cardinalidade. Na direção inversa, escolha bases de V e W com a mesma cardinalidade, fixe uma bijeção entre elas e use o Teorema 3.18 para estendê-la a um isomorfismo. ■

O resultado pode ser enunciado como um teorema de classificação.

Problema de classificação
Um teorema de classificação descreve os objetos de determinado tipo a menos de uma equivalência. A descrição deve ser completa e não redundante: todo objeto aparece em uma única classe de equivalência.

Teorema 3.34 (de Classificação dos Espaços Vetoriais). a *Para cada $n \in \mathbb{N}$, todo espaço vetorial de dimensão n sobre $\mathbb{K}$ é isomorfo a $\mathbb{K}^n$;*

- b *se κ é um cardinal e B é um conjunto de cardinalidade κ, todo espaço vetorial de dimensão κ sobre $\mathbb{K}$ é isomorfo ao espaço $(\mathbb{K}^B)_0$ das funções de B em $\mathbb{K}$ com suporte finito.*

Demonstração. O primeiro item foi demonstrado no Exemplo 3.28.

Para o segundo, suponha que B tenha cardinalidade κ. Para cada $b \in B$, defina $\delta_b \in (\mathbb{K}^B)_0$ por

$$\delta_b(x) = \begin{cases} 1, & \text{se } x = b, \\ 0, & \text{se } x \neq b. \end{cases}$$

formam uma base de $(\mathbb{K}^B)_0$: toda função de suporte finito se escreve como uma combinação linear finita das δ_b, e uma relação linear entre elas se anula ponto a ponto somente quando todos os coeficientes são nulos. Portanto, $\dim((\mathbb{K}^B)_0) = |B| = \kappa$.

Se V possui dimensão κ, uma bijeção entre uma base de V e a base $(\delta_b)_{b\in B}$ estende-se, pelo Teorema 3.18, a um isomorfismo $V \to (\mathbb{K}^B)_0$. ■

Exercícios

Ex. 3.18 — Mostre que

$$V^n \cong \underbrace{V \boxplus \cdots \boxplus V}_{n \text{ vezes}}.$$

Ex. 3.19 — Sejam A e B subespaços de um espaço vetorial V tais que $V = A + B$, e considere a soma direta externa $E = A \boxplus B$. Defina

$$\tau : E \to V, \qquad \tau(\vec{a}, \vec{b}) = \vec{a} + \vec{b}.$$

1. Prove que τ é linear.

2. Determine $\ker \tau$.

3. Determine quando τ é um isomorfismo.

Ex. 3.20 — Mostre que

$$\mathbb{K}[x] \cong (\mathbb{K}^{\infty})_0.$$

Ex. 3.21 — Seja V um espaço vetorial de dimensão finita e seja $T : V \to V$ linear. Suponha que exista $G : V \to V$ tal que $T \circ G = \mathrm{I}_V$. Prove que T é um isomorfismo e que $G = T^{-1}$. Dê um exemplo que mostre que a afirmação pode falhar quando V tem dimensão infinita.

Ex. 3.22 — Encontre um espaço vetorial V com duas decomposições

$$V = A \oplus B = C \oplus D$$

tais que $A \cong C$, mas $B \ncong D$. Conclua que subespaços isomorfos podem admitir complementos não isomorfos.

Ex. 3.23 — Sejam V, A, B, C espaços vetoriais e sejam $F : A \to C$ e $G : B \to C$ transformações lineares. Defina o produto fibrado

$$A \times_C B \coloneqq \{(\vec{a}, \vec{b}) \in A \times B \mid F(\vec{a}) = G(\vec{b})\}.$$

Mostre que:

1. $\mathrm{Hom}(V, A \oplus B)$ é isomorfo a $\mathrm{Hom}(V, A) \oplus \mathrm{Hom}(V, B)$;
2. $\mathrm{Hom}(V, A \times_C B)$ é isomorfo ao produto fibrado

$$\mathrm{Hom}(V, A) \times_{\mathrm{Hom}(V,C)} \mathrm{Hom}(V, B) \coloneqq \{(S, T) \in \mathrm{Hom}(V, A) \times \mathrm{Hom}(V, B) \mid F \circ S = G \circ T\}.$$

3.3 Teorema do Núcleo-Imagem

Núcleo, imagem e dimensão
Em dimensão finita, a dimensão do núcleo conta as direções independentes anuladas por T, e a dimensão da imagem é o posto de T. O teorema do núcleo-imagem afirma que a soma dessas dimensões é a dimensão do domínio.

Teorema 3.35 (Teorema do Núcleo-Imagem).
Sejam V e W espaços vetoriais sobre $\mathbb{K}$, com V de dimensão finita, e seja $T : V \to W$ uma transformação linear. Então

$$\dim V = \dim(\ker T) + \dim(\operatorname{im} T).$$

Demonstração. Seja $\underline{\mathrm{B}}_1 = (\mathbf{u}_1, \dots, \mathbf{u}_r)$ uma base de $\ker T$. Pelo Teorema da Extensão, complete-a a uma base

$$\underline{\mathrm{B}}_2 = (\mathbf{u}_1, \dots, \mathbf{u}_r, \mathbf{v}_1, \dots, \mathbf{v}_s)$$

de V. Mostraremos que $(T(\mathbf{v}_1), \dots, T(\mathbf{v}_s))$ é uma base de $\operatorname{im} T$.

Se $\mathbf{w} = T(\mathbf{z}) \in \operatorname{im} T$, escreva $\mathbf{z} = \sum_{i=1}^{r} a_i \mathbf{u}_i + \sum_{j=1}^{s} b_j \mathbf{v}_j$. Como $T(\mathbf{u}_i) = \vec{0}$, temos

$$\mathbf{w} = T(\mathbf{z}) = \sum_{j=1}^{s} b_j T(\mathbf{v}_j),$$

logo esses vetores geram $\operatorname{im} T$. Se $\sum_{j=1}^{s} b_j T(\mathbf{v}_j) = \vec{0}$, então $\sum_{j=1}^{s} b_j \mathbf{v}_j \in \ker T$ e, para certos $a_i \in \mathbb{K}$,

$$\sum_{j=1}^{s} b_j \mathbf{v}_j = \sum_{i=1}^{r} a_i \mathbf{u}_i.$$

A independência linear de $\underline{\mathrm{B}}_2$ implica que todos os b_j são nulos. Portanto, $(T(\mathbf{v}_1), \dots, T(\mathbf{v}_s))$ é uma base de $\operatorname{im} T$. Como $\dim V = r + s$, concluímos que

$$\dim V = \dim(\ker T) + \dim(\operatorname{im} T).$$

■

Segunda demonstração. Apresentaremos uma formulação mais abstrata do Teorema do Núcleo-Imagem. Seja $T \in \mathrm{Hom}(V, W)$. Escolha um complemento K de $\ker T$ em V, de modo que

$$V = \ker T \oplus K.$$

A restrição $T\restriction_K : K \to W$ é injetiva, pois

$$\ker(T\restriction_K) = \ker T \cap K = \{\vec{0}\}.$$

Além disso, sua imagem é $\operatorname{im} T$. De fato, uma inclusão é imediata. Para a outra, escreva $\mathbf{v} = \mathbf{u} + \mathbf{w}$, com $\mathbf{u} \in \ker T$ e $\mathbf{w} \in K$; então

$$T(\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{w}) = T(\mathbf{w}) \in \operatorname{im}(T\restriction_K).$$

Logo $T\restriction_K$ define um isomorfismo

$$K \cong \operatorname{im} T.$$

Provamos o seguinte teorema.

Teorema 3.36 (Teorema do Núcleo-Imagem).
Seja $T \in \operatorname{Hom}(V, W)$.

a *Todo complemento de $\ker T$ em V é isomorfo a $\operatorname{im} T$;*

b $\dim(\ker T) + \dim(\operatorname{im} T) = \dim V$.

Se $A \in \mathscr{M}_{m,n}(\mathbb{K})$, a imagem de T_A é o espaço das colunas de A. Portanto,

$$\dim(\ker T_A) + \operatorname{posto}(A) = n.$$

Isso fornece o seguinte resultado útil.

Teorema 3.37.
Seja $A \in \mathscr{M}_{m,n}(\mathbb{K})$. Então:

1 *$T_A : \mathbb{K}^n \to \mathbb{K}^m$ é injetiva se, e somente se, $\operatorname{posto}(A) = n$;*

2 *$T_A : \mathbb{K}^n \to \mathbb{K}^m$ é sobrejetiva se, e somente se, $\operatorname{posto}(A) = m$.*

Teorema 3.38.
Sejam V e W espaços vetoriais sobre $\mathbb{K}$ e seja $T \in \operatorname{Hom}(V, W)$. Então:

a *Se T é sobrejetiva, então $\dim V \geq \dim W$;*

b *Se $\dim V = \dim W < \infty$, então T é um isomorfismo se, e somente se, é injetiva ou sobrejetiva.*

Demonstração. a Se T é sobrejetiva, então $W = \operatorname{im} T$; pelo Teorema do Núcleo-Imagem, $\dim W = \dim(\operatorname{im} T) \leq \dim V$.

b Se T é um isomorfismo, então é injetiva e sobrejetiva. Se T é injetiva, $\ker T = \{\vec{0}\}$ e o Teorema do Núcleo-Imagem dá $\dim \operatorname{im} T = \dim V = \dim W$; logo $\operatorname{im} T = W$. Se T é sobrejetiva, $\operatorname{im} T = W$ e o mesmo teorema dá $\dim \ker T = 0$; logo T é injetiva.

■

Observe que o resultado de b não é verdadeiro, em geral, para espaços de dimensão infinita.

Proposição 3.39.
Sejam U, V, W espaços vetoriais de dimensão finita e sejam $T : U \to V$ e $S : V \to W$ transformações lineares. Então

$$\dim \ker (S \circ T) \leq \dim \ker S + \dim \ker T.$$

Demonstração. Como $\ker (S \circ T) = T^{-1}(\ker S)$, considere a restrição

$$T' : \ker (S \circ T) \longrightarrow \ker S, \qquad T'(\mathbf{u}) = T(\mathbf{u}).$$

Seu núcleo é $\ker T$. Pelo Teorema do Núcleo-Imagem,

$$\dim \ker (S \circ T) = \dim \ker T + \dim \operatorname{im} T' \leq \dim \ker T + \dim \ker S.$$

■

Leitura opcional: complexos e sequências exatas. Terminamos esta seção com uma formulação que será útil em contextos mais avançados.

Definição 3.40 — Complexo de Cadeias. Um **complexo de cadeias** de espaços vetoriais sobre $\mathbb{K}$ é uma família $C = ((V_i, d_i) \mid i \in \mathbb{Z})$, em que $d_i : V_i \to V_{i-1}$ é linear e $d_i \circ d_{i+1} = 0$ para todo $i \in \mathbb{Z}$.

A condição $d_i \circ d_{i+1} = 0$ nos diz que $\operatorname{im} d_{i+1} \subseteq \ker d_i$.

Representamos um complexo de cadeias pelo diagrama

$$C : \cdots \longrightarrow V_{i+1} \xrightarrow{d_{i+1}} V_i \xrightarrow{d_i} V_{i-1} \longrightarrow \cdots \tag{3.11}$$

Se apenas um número finito de espaços for não nulo, o diagrama assume a forma

$$C : 0 \longrightarrow V_n \xrightarrow{d_n} V_{n-1} \xrightarrow{d_{n-1}} \cdots \longrightarrow V_1 \xrightarrow{d_1} V_0 \longrightarrow 0 \tag{3.12}$$

Os espaços e as aplicações que não aparecem na Equação 3.12 são nulos.

Definição 3.41 — Exato. Um complexo de cadeias

$$C : \cdots \longrightarrow V_{i+1} \xrightarrow{d_{i+1}} V_i \xrightarrow{d_i} V_{i-1} \longrightarrow \cdots \tag{3.13}$$

é dito **exato** se $\operatorname{im} d_{i+1} = \ker d_i$ para todo $i \in \mathbb{Z}$.

Proposição 3.42.
Seja $T : X \to Y$ uma transformação linear. Então:

1. *A sequência $X \xrightarrow{T} Y \to 0$ é exata se, e somente se, T é sobrejetiva.*
2. *A sequência $0 \to X \xrightarrow{T} Y$ é exata se, e somente se, T é injetiva.*
3. *A sequência $0 \to X \xrightarrow{T} Y \to 0$ é exata se, e somente se, T é bijetiva.*

Exemplo 3.43. Sejam V e W espaços vetoriais sobre $\mathbb{K}$ e seja $T \in \mathrm{Hom}(V, W)$. Então

$$C : 0 \longrightarrow \ker T \xrightarrow{i} V \xrightarrow{T} \operatorname{im} T \longrightarrow 0 \tag{3.14}$$

é um complexo de cadeias exato, em que i denota a inclusão de $\ker T$ em V.

◁

O exemplo motiva a definição seguinte.

Definição 3.44 — Sequência Exata Curta. Uma **sequência exata curta** é um complexo de cadeias exato da forma

$$C : 0 \longrightarrow V_2 \xrightarrow{d_2} V_1 \xrightarrow{d_1} V_0 \longrightarrow 0 \tag{3.15}$$

No Exemplo 3.43, temos uma sequência exata curta com $V_2 = \ker T$, $d_2 = i$, $V_1 = V$ e $d_1 = T$. Pela definição de exatidão, o complexo C da Equação 3.15 é uma sequência exata curta se, e somente se, d_2 é injetiva, d_1 é sobrejetiva e $\operatorname{im} d_2 = \ker d_1$.

Se os espaços têm dimensão finita, o Teorema 3.36 b implica

$$\dim V_2 - \dim V_1 + \dim V_0 = 0.$$

Essa é outra forma do Teorema do Núcleo-Imagem.

Teorema 3.45 (Teorema do Núcleo-Imagem).
Seja

$$C : 0 \longrightarrow V_2 \xrightarrow{d_2} V_1 \xrightarrow{d_1} V_0 \longrightarrow 0 \tag{3.16}$$

uma sequência exata curta de espaços vetoriais de dimensão finita. Então

$$\dim V_2 - \dim V_1 + \dim V_0 = 0.$$

Podemos agora provar a generalização seguinte.

Teorema 3.46.
Suponha que

$$C : 0 \longrightarrow V_n \xrightarrow{d_n} V_{n-1} \xrightarrow{d_{n-1}} \cdots \longrightarrow V_1 \xrightarrow{d_1} V_0 \longrightarrow 0 \tag{3.17}$$

seja um complexo de cadeias exato de espaços de dimensão finita. Então

$$\sum_{i=0}^{n} (-1)^i \dim V_i = 0.$$

Demonstração. O complexo de cadeias C se decompõe nas seguintes sequências exatas

curtas:

$$C_1 \;:\; 0 \longrightarrow \ker d_1 \longrightarrow V_1 \xrightarrow{d_1} V_0 \longrightarrow 0 \tag{3.18}$$

$$C_2 \;:\; 0 \longrightarrow \ker d_2 \longrightarrow V_2 \xrightarrow{d_2} \ker d_1 \longrightarrow 0$$

$$\vdots$$

$$C_n \;:\; 0 \longrightarrow \ker d_n \longrightarrow V_n \xrightarrow{d_n} \ker d_{n-1} \longrightarrow 0$$

Aplicando o Teorema 3.36 b a cada C_i e somando as igualdades com sinais alternados, os termos $\dim \ker d_i$ se cancelam aos pares. Resta $\sum_{i=0}^{n} (-1)^i \dim V_i = 0$. ■

Exercícios

Ex. 3.24 — Seja $T : \mathbb{R}_n[x] \to \mathbb{R}_n[x]$ o operador definido por

$$T(p) = 3p - 6p' + p''.$$

1. Mostre que $\ker T = \{\vec{0}\}$.

2. Conclua que, para todo $q \in \mathbb{R}_n[x]$, existe $p \in \mathbb{R}_n[x]$ tal que $3p - 6p' + p'' = q$.

Ex. 3.25 — Seja $T : V \to V$ um operador em um espaço vetorial de dimensão finita. Mostre que

$$V = \ker T \oplus \operatorname{im} T$$

se, e somente se, $\ker T = \ker T^2$.

Ex. 3.26 — Sejam V e W espaços vetoriais de dimensão finita e seja $T : V \to W$ linear. Prove que:

1. se T é sobrejetiva, então $\dim V \geq \dim W$;

2. se T é injetiva, então $\dim V \leq \dim W$.

Ex. 3.27 — Sejam V e W espaços vetoriais de dimensão finita.

1. Se $T : V \to W$ é sobrejetiva, prove que existe uma transformação linear $S : W \to V$ tal que $T \circ S = \mathrm{I}_W$.

2. Se $T : V \to W$ é injetiva, prove que existe uma transformação linear $S : W \to V$ tal que $S \circ T = \mathrm{I}_V$.

Ex. 3.28 — Seja $A : V \to W$ uma transformação linear. Mostre que existe um espaço vetorial U, uma transformação sobrejetiva $S : V \to U$ e uma transformação injetiva $T : U \to W$ tais que $A = T \circ S$. Verifique também que é possível escolher um espaço U', uma transformação injetiva $T' : V \to U'$ e uma transformação sobrejetiva $S' : U' \to W$ tais que $A = S' \circ T'$.

Ex. 3.29 — Seja T um operador em um espaço vetorial V de dimensão n. Prove que

$$\ker T^n = \ker T^{n+1} \quad \text{e} \quad \operatorname{im} T^n = \operatorname{im} T^{n+1}.$$

Ex. 3.30 — Sob as hipóteses do exercício anterior, prove que

$$V = \operatorname{im} T^n \oplus \ker T^n.$$

Ex. 3.31 — Uma *sequência* de espaços vetoriais de dimensão finita

$$V_0 \xrightarrow{T_0} V_1 \xrightarrow{T_1} \cdots \xrightarrow{T_{n-1}} V_n$$

é *exata* se $\ker T_i = \operatorname{im} T_{i-1}$ para $i = 1, \dots, n-1$.

1. Mostre que a sequência $0 \to V_1 \xrightarrow{T_1} V_2$ é exata se, e somente se, T_1 é injetiva.

2. Mostre que a sequência $V_0 \xrightarrow{T_0} V_1 \to 0$ é exata se, e somente se, T_0 é sobrejetiva.

3. Considere a sequência exata

   $$0 \to V_1 \xrightarrow{T_1} V_2 \xrightarrow{T_2} V_3 \xrightarrow{T_3} V_4 \to 0.$$

   Se $\dim_{\mathbb{K}} V_i = d_i$, mostre que $d_1 + d_3 = d_2 + d_4$.

Ex. 3.32 — Um *complexo de cocadeias* é um par (C, d), em que $C = (C_i)_{i \in \mathbb{Z}}$ é uma família de espaços vetoriais e $d = (d_i)_{i \in \mathbb{Z}}$ é uma família de transformações lineares $d_i : C_i \to C_{i+1}$ tais que $d_{i+1} \circ d_i = 0$. Representamos o complexo pelo diagrama

$$C : \quad \cdots \xrightarrow{d_{i-2}} C_{i-1} \xrightarrow{d_{i-1}} C_i \xrightarrow{d_i} C_{i+1} \xrightarrow{d_{i+1}} \cdots.$$

1. O *i-ésimo espaço de cohomologia* de (C, d) é

   $$H^i(C, d) \coloneqq \ker d_i / \operatorname{im} d_{i-1}.$$

   Mostre que esse quociente está bem definido, isto é, que $\operatorname{im} d_{i-1} \subseteq \ker d_i$ para todo $i \in \mathbb{Z}$.

2. Suponha que cada C_i tenha dimensão finita e que o complexo seja limitado: existem $m, n \in \mathbb{Z}$ tais que $C_i = \{\vec{0}\}$ quando $i \notin \{m, \dots, n\}$. Defina sua *característica de Euler* por

   $$\chi(C, d) \coloneqq \sum_{k=m}^{n} (-1)^k \dim C_k.$$

   Mostre que

   $$\chi(C, d) = \sum_{k=m}^{n} (-1)^k \dim H^k(C, d).$$

3.4 Representação Matricial dos Homomorfismos

Exercício avançado: complexos de cocadeias.

Dependência das bases
A transformação linear existe independentemente de coordenadas. Sua matriz depende de uma base no domínio e de outra no contradomínio. Por isso, a mesma aplicação pode ter matrizes diferentes sem que seu comportamento intrínseco tenha mudado.

Teorema 3.47.
Suponha que V e W sejam espaços vetoriais tais que $\dim V = n$ e $\dim W = m$. Se $\underline{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ é uma base de V e $\underline{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_m)$ é uma base de W, então o par $(\underline{X}, \underline{Y})$ determina o isomorfismo $[\,\cdot\,]_{\underline{X},\underline{Y}} : \mathrm{Hom}(V, W) \to \mathscr{M}_{m,n}(\mathbb{K})$ definido por

$$T \longmapsto [T]_{\underline{X},\underline{Y}}, \tag{3.19}$$

onde $[T]_{\underline{X},\underline{Y}} = \Big([T(\mathbf{x}_1)]_{\underline{Y}} \,\Big|\, \cdots \,\Big|\, [T(\mathbf{x}_n)]_{\underline{Y}}\Big)$. Em particular, sua i-ésima coluna é $[T(\mathbf{x}_i)]_{\underline{Y}}$. Além disso, o diagrama comuta:

$$\begin{array}{ccc} V & \xrightarrow{\;T\;} & W \\ {\scriptstyle [\,\cdot\,]_{\underline{X}}}\downarrow & & \downarrow{\scriptstyle [\,\cdot\,]_{\underline{Y}}} \\ \mathbb{K}^n & \xrightarrow[\;[T]_{\underline{X},\underline{Y}}\;]{} & \mathbb{K}^m \end{array} \tag{3.20}$$

Demonstração. **Linearidade.** Se $T_1, T_2 \in \mathrm{Hom}(V, W)$ e $\lambda_1, \lambda_2 \in \mathbb{K}$, então

$$[\lambda_1 T_1 + \lambda_2 T_2]_{\underline{X},\underline{Y}} = \Big([(\lambda_1 T_1 + \lambda_2 T_2)(\mathbf{x}_1)]_{\underline{Y}} \,\Big|\, \cdots \,\Big|\, [(\lambda_1 T_1 + \lambda_2 T_2)(\mathbf{x}_n)]_{\underline{Y}}\Big) \tag{3.21}$$

$$= \Big(\lambda_1 [T_1(\mathbf{x}_1)]_{\underline{Y}} + \lambda_2 [T_2(\mathbf{x}_1)]_{\underline{Y}} \,\Big|\, \cdots \,\Big|\, \lambda_1 [T_1(\mathbf{x}_n)]_{\underline{Y}} + \lambda_2 [T_2(\mathbf{x}_n)]_{\underline{Y}}\Big) \tag{3.22}$$

$$= \lambda_1 \Big([T_1(\mathbf{x}_1)]_{\underline{Y}} \,\Big|\, \cdots \,\Big|\, [T_1(\mathbf{x}_n)]_{\underline{Y}}\Big) + \lambda_2 \Big([T_2(\mathbf{x}_1)]_{\underline{Y}} \,\Big|\, \cdots \,\Big|\, [T_2(\mathbf{x}_n)]_{\underline{Y}}\Big) \tag{3.23}$$

$$= \lambda_1 [T_1]_{\underline{X},\underline{Y}} + \lambda_2 [T_2]_{\underline{X},\underline{Y}}. \tag{3.24}$$

Portanto, $[\,\cdot\,]_{\underline{X},\underline{Y}}$ é linear.

Bijetividade.

Suponha que $T \in \ker [\,\cdot\,]_{\underline{X},\underline{Y}}$. Então $[T(\mathbf{x}_i)]_{\underline{Y}} = 0$ para $i = 1, \dots, n$. Pelo Teorema 5.19, temos $T(\mathbf{x}_i) = \vec{0}$ para todo i. Como uma transformação linear é determinada pelos valores que assume em uma base, $T = 0$. Logo $[\,\cdot\,]_{\underline{X},\underline{Y}}$ é injetiva.

Para provar a sobrejetividade, seja $A = (a_{ji}) \in \mathscr{M}_{m,n}(\mathbb{K})$ e, para $i = 1, \dots, n$, defina

$$\mathbf{z}_i = \sum_{j=1}^{m} a_{ji} \mathbf{y}_j.$$

Então $[\mathbf{z}_i]_{\underline{Y}}$ é a i-ésima coluna de A. Pelo Teorema 3.18, existe uma única $T \in \mathrm{Hom}(V, W)$ tal que $T(\mathbf{x}_i) = \mathbf{z}_i$ para todo i. Consequentemente, $[T]_{\underline{X},\underline{Y}} = A$, o que prova a sobrejetividade.

Comutatividade do diagrama.

No diagrama, a seta superior atua sobre vetores; a inferior multiplica colunas de coordenadas. Partindo de $\mathbf{v} \in V$, podemos aplicar T e depois tomar coordenadas na base $\underline{\mathrm{Y}}$, obtendo $[T(\mathbf{v})]_{\underline{\mathrm{Y}}}$. Podemos também tomar primeiro $[\mathbf{v}]_{\underline{\mathrm{X}}}$ e multiplicar por $[T]_{\underline{\mathrm{X}},\underline{\mathrm{Y}}}$. Precisamos mostrar que esses dois percursos produzem a mesma coluna. Isso não identifica literalmente um vetor de W com suas coordenadas: compara duas maneiras de calcular essas coordenadas.

Para cada vetor $\mathbf{x}_i$ da base $\underline{\mathrm{X}}$, temos

$$\begin{aligned}([\,\cdot\,]_{\underline{\mathrm{Y}}}T)(\mathbf{x}_i) &= [T(\mathbf{x}_i)]_{\underline{\mathrm{Y}}} \\ &= \mathrm{Col}_i([T]_{\underline{\mathrm{X}},\underline{\mathrm{Y}}}) \\ &= [T]_{\underline{\mathrm{X}},\underline{\mathrm{Y}}}\,(0,\dots,1,\dots,0)^{\mathrm{t}} \\ &= [T]_{\underline{\mathrm{X}},\underline{\mathrm{Y}}}\,[\mathbf{x}_i]_{\underline{\mathrm{X}}}.\end{aligned}$$

Os dois caminhos do diagrama são transformações lineares de V em $\mathbb{K}^m$ e coincidem em todos os vetores da base $\underline{\mathrm{X}}$; portanto, coincidem em todo V. ■

Definição 3.48 — Matriz da Transformação Linear. A matriz $[T]_{\underline{\mathrm{X}},\underline{\mathrm{Y}}}$ é denominada **matriz da transformação linear** T em relação às bases $\underline{\mathrm{X}}$ e $\underline{\mathrm{Y}}$.

Como as aplicações de coordenadas e $[\,\cdot\,]_{\underline{\mathrm{X}},\underline{\mathrm{Y}}}$ são isomorfismos, é comum identificar V, W e $\mathrm{Hom}(V,W)$, depois de escolhidas as bases, com $\mathbb{K}^n$, $\mathbb{K}^m$ e $\mathscr{M}_{m,n}(\mathbb{K})$, respectivamente. Ainda assim, convém distinguir a transformação linear de sua matriz.

Exemplo 3.49. Seja $V = \mathbb{K}_3[x]$ o espaço dos polinômios de grau menor ou igual a 3. O operador de diferenciação formal D do Exemplo 3.9 leva V em si mesmo. Considere a base ordenada $\underline{\mathrm{B}} = (p_1, p_2, p_3, p_4) = (1, x, x^2, x^3)$. Então

$$\begin{aligned} Dp_1 &= 0, & Dp_1 &= 0p_1 + 0p_2 + 0p_3 + 0p_4 \\ Dp_2 &= 1, & Dp_2 &= 1p_1 + 0p_2 + 0p_3 + 0p_4 \\ Dp_3 &= 2x, & Dp_3 &= 0p_1 + 2p_2 + 0p_3 + 0p_4 \\ Dp_4 &= 3x^2, & Dp_4 &= 0p_1 + 0p_2 + 3p_3 + 0p_4 \end{aligned}$$

Logo a matriz de D na base $\underline{\mathrm{B}}$ é

$$[D]_{\underline{\mathrm{B}}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

◁

Corolário 3.50.
Sejam $\underline{v}$ e $\underline{w}$ bases de V e W, respectivamente, sejam $T, S \in \mathrm{Hom}(V,W)$ e seja $c \in \mathbb{K}$. Então

1. $[cT]_{\underline{v},\underline{w}} = c[T]_{\underline{v},\underline{w}}$;
2. $[T+S]_{\underline{v},\,\underline{w}} = [T]_{\underline{v},\,\underline{w}} + [S]_{\underline{v},\,\underline{w}}$.

Seja A uma matriz $m \times n$ com colunas $\vec{a}_1, \ldots, \vec{a}_n$ e seja $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{K}^n$. Então

$$A\mathbf{x} = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n.$$

Essa identidade fornece imediatamente o resultado seguinte.

Teorema 3.51.
Sejam V e W espaços vetoriais de dimensão finita, com bases $\underline{v}$ e $\underline{w}$, respectivamente, e seja $T : V \to W$ linear. Então, para todo $\mathbf{v} \in V$,

$$[T(\mathbf{v})]_{\underline{W}} = [T]_{\underline{v},\underline{W}}[\mathbf{v}]_{\underline{v}}.$$

Se A tem colunas $\vec{a}_1, \ldots, \vec{a}_n$ e B é uma matriz $p \times m$, então a j-ésima coluna de BA é $B\vec{a}_j$. Portanto,

$$BA = \left(B\vec{a}_1 \,\middle|\, B\vec{a}_2 \,\middle|\, \cdots \,\middle|\, B\vec{a}_n\right).$$

Essa observação traduz a composição de aplicações em multiplicação de matrizes.

Teorema 3.52.
Sejam V, W, X espaços vetoriais de dimensão finita, com bases $\underline{v}$, $\underline{w}$, $\underline{x}$, respectivamente, e sejam $T : V \to W$ e $S : W \to X$ transformações lineares. Então

$$[S \circ T]_{\underline{v},\underline{x}} = [S]_{\underline{w},\underline{x}}[T]_{\underline{v},\underline{w}}.$$

Demonstração. Ponha $A = [T]_{\underline{v},\underline{w}}$ e $B = [S]_{\underline{w},\underline{x}}$. Se $\vec{a}_j$ é a j-ésima coluna de A, então, pelo Teorema 3.51,

$$[(S \circ T)(\mathbf{v}_j)]_{\underline{x}} = [S(T(\mathbf{v}_j))]_{\underline{x}} = B[T(\mathbf{v}_j)]_{\underline{w}} = B\vec{a}_j.$$

Esse vetor é tanto a j-ésima coluna de $[S \circ T]_{\underline{v},\underline{x}}$ quanto a j-ésima coluna de BA. Logo as duas matrizes são iguais. ∎

Essa compatibilidade é uma das motivações para a definição do produto de matrizes.

3.4.1 Mudança de Bases

Na mudança de coordenadas, um vetor permanece fixo enquanto sua coluna de coordenadas se altera. Agora consideramos uma transformação linear $T : V \to W$ e perguntamos como sua matriz se altera quando mudamos as bases. A transformação continua sendo a mesma, mas sua representação depende de duas escolhas: a base usada para escrever os vetores de entrada e a base usada para escrever suas imagens. Para um operador, a notação $[T]_{\underline{B}}$ indica que a mesma base $\underline{B}$ é usada no domínio e no contradomínio.

Exemplo 3.53 — Uma transformação em duas bases. Considere $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por

$$T\begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} s \\ 2t \end{pmatrix}.$$

Na base canônica $\underline{\mathrm{E}}$, sua matriz é

$$[T]_{\underline{\mathrm{E}}} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

Retomemos a base $\underline{\mathrm{B}} = (\vec{u}_1, \vec{u}_2)$ do Exemplo 2.75, com $\vec{u}_1 = (1,1)^{\mathrm{t}}$ e $\vec{u}_2 = (1,2)^{\mathrm{t}}$. Como

$$T(\vec{u}_1) = \vec{u}_2, \qquad T(\vec{u}_2) = -2\vec{u}_1 + 3\vec{u}_2,$$

as coordenadas dessas imagens fornecem as colunas de

$$[T]_{\underline{\mathrm{B}}} = \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix}.$$

As duas matrizes são diferentes, mas representam a mesma aplicação. A primeira atua sobre coordenadas na base $\underline{\mathrm{E}}$; a segunda atua sobre coordenadas na base $\underline{\mathrm{B}}$.

◁

Para relacionar essas matrizes sem recalcular todas as imagens, precisamos acompanhar o sentido de cada mudança de coordenadas. Sejam $\underline{\mathrm{x}}$ e $\underline{\mathrm{y}}$ as bases inicialmente escolhidas em V e W, e sejam $\underline{\mathrm{x}}$' e $\underline{\mathrm{y}}$' as novas bases. Escrevamos

$$A = [T]_{\underline{\mathrm{x}},\underline{\mathrm{y}}}, \qquad A' = [T]_{\underline{\mathrm{x}}',\underline{\mathrm{y}}'}.$$

Conhecemos A e queremos obter A'. Para usar A em um vetor cuja coluna foi dada na base $\underline{\mathrm{x}}$', primeiro devemos converter essa coluna para a base $\underline{\mathrm{x}}$. Depois de aplicar A, devemos converter a coluna obtida na base $\underline{\mathrm{y}}$ para a base $\underline{\mathrm{y}}$'. Definimos, portanto,

$$P = \mathrm{M}_{\underline{\mathrm{x}}' \to \underline{\mathrm{x}}}, \qquad Q = \mathrm{M}_{\underline{\mathrm{y}} \to \underline{\mathrm{y}}'}.$$

O percurso completo é

$$[\mathbf{v}]_{\underline{\mathrm{x}}'} \xrightarrow{P} [\mathbf{v}]_{\underline{\mathrm{x}}} \xrightarrow{A} [T(\mathbf{v})]_{\underline{\mathrm{y}}} \xrightarrow{Q} [T(\mathbf{v})]_{\underline{\mathrm{y}}'}.$$

Assim, a matriz procurada é $A' = QAP$: o produto é aplicado da direita para a esquerda, na ordem indicada pelo percurso. As matrizes P e Q são quadradas e invertíveis, mas A pode ser retangular.

Teorema 3.54 (Mudança de Base).
Sejam V e W espaços vetoriais de dimensões n e m sobre $\mathbb{K}$, respectivamente. Sejam $\underline{\mathrm{x}}$ e $\underline{\mathrm{x}}$' bases de V, e sejam $\underline{\mathrm{y}}$ e $\underline{\mathrm{y}}$' bases de W. Para todo $T \in \mathrm{Hom}(V,W)$,

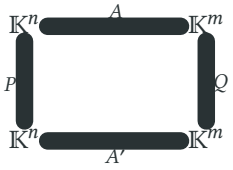
$$\begin{aligned} [T]_{\underline{\mathrm{x}}',\underline{\mathrm{y}}'} &= \mathrm{M}_{\underline{\mathrm{y}} \to \underline{\mathrm{y}}'} [T]_{\underline{\mathrm{x}},\underline{\mathrm{y}}} \mathrm{M}_{\underline{\mathrm{x}}' \to \underline{\mathrm{x}}} \\ &= \mathrm{M}_{\underline{\mathrm{y}} \to \underline{\mathrm{y}}'} [T]_{\underline{\mathrm{x}},\underline{\mathrm{y}}} \mathrm{M}^{-1}_{\underline{\mathrm{x}} \to \underline{\mathrm{x}}'}. \end{aligned} \tag{3.25}$$

Demonstração. Com a notação acima, para todo $\mathbf{v} \in V$ temos

$$\begin{aligned} A'[\mathbf{v}]_{\underline{\mathrm{x}}'} &= [T(\mathbf{v})]_{\underline{\mathrm{y}}'} \\ &= Q[T(\mathbf{v})]_{\underline{\mathrm{y}}} \\ &= QA[\mathbf{v}]_{\underline{\mathrm{x}}} \\ &= QAP[\mathbf{v}]_{\underline{\mathrm{x}}'}. \end{aligned}$$

As igualdades usam a fórmula de representação matricial do Teorema 3.51 e a mudança de coordenadas do Teorema 2.77. Como $[\mathbf{v}]_{\underline{\mathrm{x}}'}$ percorre todo $\mathbb{K}^n$, as matrizes A' e QAP coincidem. Por fim, $\mathrm{M}_{\underline{\mathrm{x}}'\to\underline{\mathrm{x}}} = \mathrm{M}^{-1}_{\underline{\mathrm{x}}\to\underline{\mathrm{x}}'}$ pela Equação 2.18, o que fornece a segunda expressão. ∎

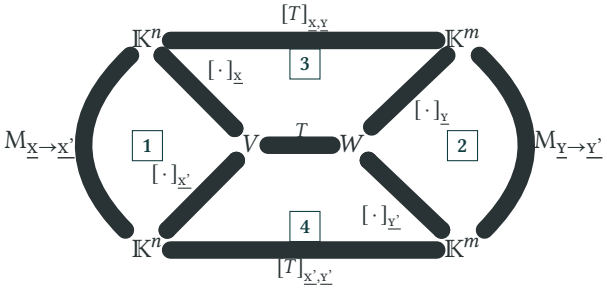
O diagrama seguinte resume o cálculo. Na linha superior usamos as bases $\underline{\mathrm{x}}$ e $\underline{\mathrm{y}}$; na inferior, as bases $\underline{\mathrm{x}}'$ e $\underline{\mathrm{y}}'$. As duas cópias de $\mathbb{K}^n$ recebem coordenadas do mesmo espaço V, mas em bases diferentes; o mesmo vale para as duas cópias de $\mathbb{K}^m$ e W. A seta P sobe porque converte as novas coordenadas de entrada nas antigas; a seta Q desce porque converte as antigas coordenadas de saída nas novas.



(3.26)

Partimos da posição inferior esquerda, com a coluna $[\mathbf{v}]_{\underline{\mathrm{x}}'}$. A seta inferior aplica A' diretamente. O outro caminho sobe por P, segue por A e desce por Q. Ambos terminam em $[T(\mathbf{v})]_{\underline{\mathrm{y}}'}$: a igualdade $A' = QAP$ expressa precisamente essa concordância.

Para relacionar esse cálculo com a transformação $T : V \to W$, reunimos agora o triângulo de mudança de coordenadas 2.17 e o quadrado de representação matricial 3.20 no diagrama completo abaixo. No centro estão os espaços V e W e a transformação T, que não dependem das bases. As setas diagonais associam a cada vetor suas coordenadas; as setas curvas convertem uma coluna de coordenadas na outra, sem alterar o vetor que ela representa.



(3.27)

Podemos ler o diagrama por partes. Na região 1, obter diretamente $[\mathbf{v}]_{\underline{\mathrm{x}}'}$ equivale a obter primeiro $[\mathbf{v}]_{\underline{\mathrm{x}}}$ e aplicar $\mathrm{M}_{\underline{\mathrm{x}}\to\underline{\mathrm{x}}'} = P^{-1}$. A região 2 expressa a mesma mudança de descrição em W, agora por $\mathrm{M}_{\underline{\mathrm{y}}\to\underline{\mathrm{y}}'} = Q$. Esses dois triângulos comutam pelo Teorema 2.77.

As regiões 3 e 4 representam a mesma transformação T nas bases antigas e nas novas, respectivamente. Em cada uma delas, aplicar T e depois tomar coordenadas dá o mesmo resultado que tomar coordenadas primeiro e multiplicar pela matriz correspondente, A ou A'. Sua comutatividade é a do Teorema 3.47.

Consideremos agora apenas os dois percursos exteriores, partindo da coluna $[\mathbf{v}]_{\underline{\mathrm{X}}}$ no canto superior esquerdo:

$$[\mathbf{v}]_{\underline{\mathrm{X}}} \xrightarrow{A} [T(\mathbf{v})]_{\underline{\mathrm{Y}}} \xrightarrow{Q} [T(\mathbf{v})]_{\underline{\mathrm{Y}}'},$$
$$[\mathbf{v}]_{\underline{\mathrm{X}}} \xrightarrow{P^{-1}} [\mathbf{v}]_{\underline{\mathrm{X}}'} \xrightarrow{A'} [T(\mathbf{v})]_{\underline{\mathrm{Y}}'}.$$

Os dois chegam às coordenadas de $T(\mathbf{v})$ na base $\underline{\mathrm{Y}}'$. Como isso vale para todo $\mathbf{v} \in V$, obtemos $QA = A'P^{-1}$; multiplicando à direita por P, recuperamos $A' = QAP$. A inversa na lateral esquerda aparece porque seguimos das coordenadas antigas para as novas, no sentido contrário ao percorrido pela seta P no diagrama 3.26.

Domínio e contradomínio
Se apenas a base do domínio muda, a nova matriz é AP. Se apenas a base do contradomínio muda, ela é QA. São escolhas independentes; por isso, na fórmula geral, os fatores dos dois lados não precisam ser inversos um do outro.

Quando $T : V \to V$ é um operador e usamos a mesma base no domínio e no contradomínio, escrevemos $[T]_{\underline{\mathrm{X}}}$ em lugar de $[T]_{\underline{\mathrm{X}},\underline{\mathrm{X}}}$. Ao substituir $\underline{\mathrm{X}}$ por $\underline{\mathrm{X}}'$ nos dois lados, a mudança de entrada é $P = \mathrm{M}_{\underline{\mathrm{X}}'\to\underline{\mathrm{X}}}$ e a mudança de saída é $\mathrm{M}_{\underline{\mathrm{X}}\to\underline{\mathrm{X}}'} = P^{-1}$. Nesse caso, a fórmula geral se reduz a

$$[T]_{\underline{\mathrm{X}}'} = P^{-1}[T]_{\underline{\mathrm{X}}}P. \tag{3.28}$$

A presença da inversa tem, portanto, uma razão precisa: as coordenadas de entrada são convertidas das novas para as antigas, e as de saída fazem o percurso contrário. Essa forma da fórmula corresponde ao uso de uma única base para representar tanto a entrada quanto a saída. Quando as bases são escolhidas independentemente, mesmo que $V = W$, devemos usar a fórmula geral.

No Exemplo 3.53, temos $P = \mathrm{M}_{\underline{\mathrm{B}}\to\underline{\mathrm{E}}}$ e

$$[T]_{\underline{\mathrm{B}}} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix}.$$

Podemos conferir a relação em um vetor. Para $\vec{z} = (3, 5)^{\mathrm{t}}$, temos $[\vec{z}]_{\underline{\mathrm{B}}} = (1, 2)^{\mathrm{t}}$ e

$$[T(\vec{z})]_{\underline{\mathrm{B}}} = \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -4 \\ 7 \end{bmatrix}.$$

De fato, $-4\vec{u}_1 + 7\vec{u}_2 = (3, 10)^{\mathrm{t}} = T(\vec{z})$. A coluna final não é a coluna de coordenadas canônicas de $T(\vec{z})$: ela ainda deve ser interpretada na base $\underline{\mathrm{B}}$.

Uma mudança de base nem sempre altera as entradas da matriz. O exemplo seguinte mostra uma situação em que a nova base conserva as mesmas relações entre os vetores e suas imagens.

Exemplo 3.55 — Diferenciação em duas bases de polinômios. Seja $V = \mathbb{K}_3[x]$ e seja D o operador de diferenciação formal do Exemplo 3.49. Considere as bases

$$\underline{\mathrm{B}} = (p_1, p_2, p_3, p_4) = (1, x, x^2, x^3)$$

e

$$\underline{\mathrm{C}} = (q_1, q_2, q_3, q_4) = (1, x-1, (x-1)^2, (x-1)^3).$$

Já calculamos as mudanças entre essas bases no Exemplo 2.78. Pondo $P = \mathrm{M}_{\underline{\mathrm{C}}\to\underline{\mathrm{B}}}$, temos $P^{-1} = \mathrm{M}_{\underline{\mathrm{B}}\to\underline{\mathrm{C}}}$. Pela fórmula de mudança de base,

$$[D]_{\underline{\mathrm{C}}} = P^{-1}[D]_{\underline{\mathrm{B}}}P.$$

Neste caso, o produto fornece

$$[D]_{\underline{\mathrm{C}}} = [D]_{\underline{\mathrm{B}}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

A igualdade também se explica diretamente pelas imagens dos vetores da nova base:

$$Dq_1 = 0, \qquad Dq_2 = q_1, \qquad Dq_3 = 2q_2, \qquad Dq_4 = 3q_3.$$

São as mesmas relações que a derivação satisfaz na base $\underline{\mathrm{B}}$. Por isso as colunas das duas matrizes coincidem, embora as coordenadas de um polinômio nas duas bases possam ser diferentes.

◁

Exercícios

Ex. 3.33 — Seja $T : \mathscr{M}_{2,2}(\mathbb{R}) \to \mathscr{M}_{2,2}(\mathbb{R})$ definida por

$$T\begin{bmatrix} x & y \\ z & w \end{bmatrix} = \begin{bmatrix} 0 & x \\ z - w & 0 \end{bmatrix}.$$

1. Determine a matriz de T em relação à base canônica.

2. Determine a matriz de T em relação à base

$$\underline{\mathrm{B}} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \right).$$

3. Exiba a matriz de mudança de base M tal que

$$[T]_{\underline{\mathrm{B}}} = M^{-1}[T]_{\mathrm{Can}} M.$$

Ex. 3.34 — Seja $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

1. Prove que não existem matrizes quadradas A e B sobre $\mathbb{K}$ tais que $AB - BA = I$.

2. Conclua que não existem operadores T e G em um espaço vetorial de dimensão finita tais que $T \circ G - G \circ T = \mathrm{I}$.

3. Mostre que a conclusão pode falhar em dimensão infinita. Sugestão: considere os operadores de multiplicação por x e de derivação em $\mathbb{K}[x]$.

Ex. 3.35 — Considere as bases

$$\underline{\mathrm{B}} = (1, x + x^2, 1 + x^2) \quad \text{e} \quad \underline{\mathrm{C}} = (1, x, x^2)$$

de $\mathbb{R}_2[x]$. Seja $T : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$ definida, em coordenadas, pela regra

$$[p]_{\underline{\mathrm{B}}} = (a, b, c) \quad \Longrightarrow \quad [T(p)]_{\underline{\mathrm{C}}} = (2c - 2b, a + c, a + b + c).$$

1. Verifique que T é linear.

2. Existe um polinômio não nulo p tal que $T(p) = p$? Justifique.

Ex. 3.36 — Seja W um subespaço de V, com $m = \dim W < \dim V = n$, e defina

$$Z = \{T \in \mathrm{Hom}(V,V) \mid T(\mathbf{w}) = \vec{0} \text{ para todo } \mathbf{w} \in W\}.$$

Mostre que Z é um subespaço de $\mathrm{Hom}(V,V)$ e calcule sua dimensão.

Ex. 3.37 — Sejam V e W espaços vetoriais de dimensão finita sobre $\mathbb{K}$, com bases $\underline{\mathrm{v}}$ e $\underline{\mathrm{w}}$, respectivamente. Seja $T : V \to W$ uma transformação linear e ponha $A = [T]_{\underline{\mathrm{v}},\underline{\mathrm{w}}}$. Prove que

$$\mathbf{v} \in \ker T \quad \Longleftrightarrow \quad [\mathbf{v}]_{\underline{\mathrm{v}}} \in \{\mathbf{x} \in \mathbb{K}^n \mid A\mathbf{x} = \vec{0}\}.$$

O subespaço à direita é denominado **espaço nulo** da matriz A.

3.5 Somas Diretas e Projeções

Decompor um espaço em soma direta permite estudar um operador parcela por parcela. As projeções fazem o caminho inverso: extraem de cada vetor sua componente em uma das parcelas.

3.5.1 Soma Direta de Transformações Lineares

Definição 3.56 — Soma Direta de Transformações Lineares. Suponha que $U = V \oplus W$ e sejam $T_1 : V \to V$ e $T_2 : W \to W$ transformações lineares. A **soma direta** de T_1 e T_2 é a transformação linear $T_1 \oplus T_2 : U \to U$ definida por

$$(T_1 \oplus T_2)(\mathbf{v} + \mathbf{w}) = T_1(\mathbf{v}) + T_2(\mathbf{w}),$$

para $\mathbf{v} \in V$ e $\mathbf{w} \in W$.

Os subespaços V e W são invariantes por $T_1 \oplus T_2$, pois

$$(T_1 \oplus T_2)(V) \subseteq V \quad \text{e} \quad (T_1 \oplus T_2)(W) \subseteq W.$$

Concatenação
Dadas duas listas ordenadas $\underline{\mathrm{L}}_1 = (a_1, \dots, a_n)$ e $\underline{\mathrm{L}}_2 = (b_1, \dots, b_m)$, a **concatenação** de $\underline{\mathrm{L}}_1$ e $\underline{\mathrm{L}}_2$, denotada $\underline{\mathrm{L}}_1 + \underline{\mathrm{L}}_2$, é a lista ordenada

$$\underline{\mathrm{L}}_1 + \underline{\mathrm{L}}_2 := (a_1, \dots, a_n, b_1, \dots, b_m).$$

Teorema 3.57.
Para $1 \le i \le j$, seja V_i um espaço vetorial de dimensão k_i, com base $\underline{B}_i = (\mathbf{v}_{i1}, \dots, \mathbf{v}_{i,k_i})$, e seja $T_i : V_i \to V_i$ um operador linear. Considere

$$T = T_1 \oplus \cdots \oplus T_j \quad \textit{em} \quad V = V_1 \oplus \cdots \oplus V_j.$$

Na base ordenada $\underline{B} = \underline{B}_1 + \underline{B}_2 + \cdots + \underline{B}_j$, obtida por concatenação, a matriz de T é

$$[T]_{\underline{B}} = \begin{bmatrix} [T_1]_{\underline{B}_1} & 0 & \cdots & 0 \\ 0 & [T_2]_{\underline{B}_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & [T_j]_{\underline{B}_j} \end{bmatrix},$$

em que $[T_i]_{\underline{B}_i}$ é um bloco de tamanho $k_i \times k_i$.

Demonstração. A lista $\underline{\mathrm{B}}$ é uma base de V. Se $[T_i]_{\underline{\mathrm{B}}_i} = (c_{rs}^{(i)})$, então, como $T(V_i) \subseteq V_i$,

$$T(\mathbf{v}_{i,m}) = \sum_{r=1}^{k_i} c_{rm}^{(i)} \mathbf{v}_{i,r}.$$

Assim, a coluna correspondente a $\mathbf{v}_{i,m}$ possui entradas apenas no bloco de coordenadas de V_i. Isso vale para todos os vetores de $\underline{\mathrm{B}}_i$ e para todo i, de modo que a matriz de T é diagonal em blocos, com o i-ésimo bloco igual a $[T_i]_{\underline{\mathrm{B}}_i}$. ■

Reciprocamente, suponha que $U = V \oplus W$ e que o operador $T : U \to U$ deixe V e W invariantes. As restrições $T_1 = T{\restriction}_V$ e $T_2 = T{\restriction}_W$ são operadores em V e W, respectivamente, e $T = T_1 \oplus T_2$. Assim, o estudo de T reduz-se ao estudo de suas restrições aos dois subespaços.

Definição 3.58 — Subespaço Invariante. Dado $T \in \mathrm{Hom}(V, V)$, dizemos que um subespaço $W \subseteq V$ é **invariante** por T, ou **T-invariante**, se $T(W) \subseteq W$.

Se $T \in \mathrm{Hom}(V, V)$ e $W \subseteq V$ é T-invariante, a restrição de T a W define um operador $T{\restriction}_W \in \mathrm{Hom}(W, W)$.

Teorema 3.59.
Seja $T : V \to V$ um operador e suponha que V se decomponha em subespaços T-invariantes

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_j,$$

com $\dim V_i = n_i$, para $1 \leq i \leq j$. Escolha uma base $\underline{B}_i$ de cada V_i e ponha $\underline{B} = \underline{B}_1 \mathbin{+\!\!+} \cdots \mathbin{+\!\!+} \underline{B}_j$. Se $T_i = T{\restriction}_{V_i}$, então a representação de T na base $\underline{B}$ é

$$[T]_{\underline{B}} = \begin{bmatrix} [T_1]_{\underline{B}_1} & 0 & \cdots & 0 \\ 0 & [T_2]_{\underline{B}_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & [T_j]_{\underline{B}_j} \end{bmatrix},$$

em que $[T_i]_{\underline{B}_i}$ é um bloco de tamanho $n_i \times n_i$.

Demonstração. Como cada V_i é T-invariante, a imagem de um vetor de $\underline{\mathrm{B}}_i$ é combinação linear apenas dos vetores dessa mesma base. Portanto, as colunas correspondentes a $\underline{\mathrm{B}}_i$ têm entradas nulas fora do i-ésimo bloco, que é precisamente a matriz de T_i em $\underline{\mathrm{B}}_i$. ■

Portanto, uma decomposição de V em soma direta de subespaços invariantes produz uma matriz diagonal em blocos para T. Reciprocamente, toda matriz diagonal em blocos determina, na base correspondente, uma decomposição de V em subespaços invariantes.

Projeções e complementos
Uma projeção sobre V ao longo de W pressupõe uma decomposição $U = V \oplus W$. Ela conserva a componente em V e anula a componente em W; assim, sua imagem é V e seu núcleo é W.

3.5.2 Projeções

Definição 3.60 — Projeção. Seja $V = V_1 \oplus \cdots \oplus V_n$. Para cada i, a **projeção** de V sobre V_i é o operador $P_i : V \to V$ definido por

$$P_i\left(\sum_{j=1}^{n} \mathbf{v}_j\right) = \mathbf{v}_i, \qquad \mathbf{v}_j \in V_j.$$

Dizemos também que P_i é a projeção sobre V_i ao longo do subespaço complementar $\bigoplus_{j\neq i} V_j$.

As propriedades fundamentais desses operadores seguem diretamente da unicidade da decomposição em soma direta.

Proposição 3.61.
As projeções P_i associadas à decomposição $V = V_1 \oplus \cdots \oplus V_n$ têm as seguintes propriedades:

a *cada $P_i : V \to V$ é um operador linear;*

b $P_i^2 = P_i$ *para todo i;*

c $P_i P_j = 0$ *se $i \neq j$;*

d $\operatorname{im} P_i = V_i$ *e* $\ker P_i = \bigoplus_{j\neq i} V_j$;

e $P_1 + \cdots + P_n = \mathrm{I}_V$.

Identificando V com $V_1 \times \cdots \times V_n$, a i-ésima projeção assume a forma

$$P_i(\mathbf{v}_1, \ldots, \mathbf{v}_n) = (\vec{0}, \ldots, \vec{0}, \mathbf{v}_i, \vec{0}, \ldots, \vec{0}).$$

Essas projeções são definidas por uma decomposição em soma direta e não devem ser confundidas com projeções ortogonais, que exigem a escolha adicional de um produto interno.

Proposição 3.62.
Seja $V = V_1 \oplus \cdots \oplus V_n$. Então a projeção $P_i : V \to V$ pode ser escrita como a soma direta

$$P_i = 0_{V_1} \oplus \cdots \oplus 0_{V_{i-1}} \oplus \mathrm{I}_{V_i} \oplus 0_{V_{i+1}} \oplus \cdots \oplus 0_{V_n}.$$

Se $\underline{B}_i$ é uma base de V_i, $r_i = \dim V_i$ e $\underline{B}$ é a base ordenada obtida pela concatenação $\underline{B}_1 \mathbin{+\!\!+} \underline{B}_2 \mathbin{+\!\!+} \cdots \mathbin{+\!\!+} \underline{B}_n$, então

$$[P_i]_{\underline{B},\underline{B}} = \begin{bmatrix} 0_{r_1} & & & & \\ & \ddots & & & \\ & & I_{r_i} & & \\ & & & \ddots & \\ & & & & 0_{r_n} \end{bmatrix}.$$

A propriedade de ser idempotente, $P^2 = P$, caracteriza as projeções associadas a decomposições em soma direta. Já vimos que, se P e $Q = \mathrm{I} - P$ são as projeções associadas a $V = V_1 \oplus V_2$, então

1. $P^2 = P$ e $Q^2 = Q$;
2. $PQ = QP = 0$;
3. $P + Q = \mathrm{I}$.

A recíproca também é verdadeira.

Proposição 3.63.
Se $P : V \to V$ é um operador linear idempotente, isto é, $P^2 = P$, então $V = \operatorname{im} P \oplus \ker P$ e P é a projeção de V em $\operatorname{im} P$ ao longo de $\ker P$. O operador $Q = \mathrm{I} - P$ também é idempotente, com

$$\operatorname{im} Q = \operatorname{im}(\mathrm{I} - P) = \ker P \quad e \quad \ker Q = \ker(\mathrm{I} - P) = \operatorname{im} P,$$

e projeta V sobre $\ker P$ ao longo de $\operatorname{im} P$.

Demonstração. Temos $Q^2 = (\mathrm{I} - P)^2 = \mathrm{I} - 2P + P^2 = \mathrm{I} - P = Q$. Além disso,

$$\begin{aligned} \mathbf{v} \in \ker Q &\iff (\mathrm{I} - P)(\mathbf{v}) = \vec{0} \\ &\iff P(\mathbf{v}) = \mathbf{v} \iff \mathbf{v} \in \operatorname{im} P, \end{aligned}$$

onde, na última equivalência, usamos $P^2 = P$. De modo análogo, $\operatorname{im} Q = \ker P$.

Para todo $\mathbf{v} \in V$,

$$\mathbf{v} = P(\mathbf{v}) + (\mathrm{I} - P)(\mathbf{v}),$$

com $P(\mathbf{v}) \in \operatorname{im} P$ e $(\mathrm{I} - P)(\mathbf{v}) \in \ker P$. Logo $V = \operatorname{im} P + \ker P$. Se $\mathbf{v} \in \operatorname{im} P \cap \ker P$, então $\mathbf{v} = P(\mathbf{w})$ para algum $\mathbf{w}$ e $\mathbf{v} = P(\mathbf{w}) = P^2(\mathbf{w}) = P(\mathbf{v}) = \vec{0}$. Portanto, a soma é direta. Finalmente, para $\mathbf{u} \in \operatorname{im} P$ e $\mathbf{w} \in \ker P$, temos $P(\mathbf{u} + \mathbf{w}) = \mathbf{u}$; assim, P é precisamente a projeção afirmada. ■

3.6 Mudança de Corpo

Leitura opcional. Esta seção trata da restrição de escalares, da descomplexificação e da complexificação. Ela pode ser omitida em uma primeira leitura sem interromper a sequência principal do curso.

O corpo faz parte da estrutura
O corpo de escalares faz parte da estrutura vetorial. A descomplexificação restringe os escalares complexos aos reais; a complexificação constrói um espaço complexo a partir de um espaço real. As dimensões e as propriedades dos operadores devem ser examinadas em relação ao corpo considerado.

O corpo de escalares faz parte da estrutura de um espaço vetorial. Se V é complexo, podemos conservar a adição e restringir a multiplicação por escalares de $\mathbb{C}$ a $\mathbb{R}$; obtemos assim um espaço vetorial real.

Definição 3.64 — Descomplexificação. Dado um espaço vetorial V sobre $\mathbb{C}$, obtemos um espaço vetorial sobre $\mathbb{R}$ restringindo a multiplicação por escalares aos números reais. Esse espaço é a **descomplexificação** de V e será denotado por $V_{\mathbb{R}}$.

Exemplo 3.65. O espaço $\mathbb{C}$ é unidimensional sobre si mesmo. Sua descomplexificação $\mathbb{C}_{\mathbb{R}}$ é um espaço vetorial real bidimensional, com base $\{1, i\}$.

◁

A mesma construção aplica-se às transformações lineares.

Definição 3.66 — Descomplexificação de uma Transformação. Sejam V e W espaços vetoriais sobre $\mathbb{C}$ e seja $T : V \to W$ uma transformação linear. Considerada apenas em relação aos escalares reais, T define uma transformação linear $T^{\mathbb{R}} : V_{\mathbb{R}} \to W_{\mathbb{R}}$, denominada **descomplexificação** de T.

Teorema 3.67. [a] *Se $\underline{E} = (\vec{e}_1, \dots, \vec{e}_n)$ é uma base de um espaço V sobre $\mathbb{C}$, então*

$$\underline{E}_{\mathbb{R}} := (\vec{e}_1, \dots, \vec{e}_n, i\vec{e}_1, \dots, i\vec{e}_n)$$

é uma base de $V_{\mathbb{R}}$ sobre $\mathbb{R}$. Em particular, $\dim_{\mathbb{R}} V_{\mathbb{R}} = 2\dim_{\mathbb{C}} V$.

[b] *Se $A = B + iC$ é a matriz da transformação linear $T : V \to W$ nas bases $\underline{E} = (\vec{e}_1, \dots, \vec{e}_n)$ e $\underline{F} = (\vec{f}_1, \dots, \vec{f}_m)$, onde B e C são matrizes reais, então a matriz de $T^{\mathbb{R}} : V_{\mathbb{R}} \to W_{\mathbb{R}}$ nas bases $\underline{E}_{\mathbb{R}}$ e $\underline{F}_{\mathbb{R}}$ é*

$$[T^{\mathbb{R}}]_{\underline{E}_{\mathbb{R}}, \underline{F}_{\mathbb{R}}} = \begin{bmatrix} B & -C \\ C & B \end{bmatrix}.$$

Demonstração. [a] Para qualquer $\mathbf{x} \in V$, existem $a_k = b_k + ic_k \in \mathbb{C}$, com $b_k, c_k \in \mathbb{R}$, tais que

$$\mathbf{x} = \sum_{k=1}^{n} a_k \vec{e}_k = \sum_{k=1}^{n} b_k \vec{e}_k + \sum_{k=1}^{n} c_k (i\vec{e}_k).$$

Logo $\underline{E}_{\mathbb{R}}$ gera $V_{\mathbb{R}}$. Por outro lado, se

$$\sum_{k=1}^{n} b_k \vec{e}_k + \sum_{k=1}^{n} c_k (i\vec{e}_k) = \vec{0},$$

então $\sum_{k=1}^{n} (b_k + ic_k)\vec{e}_k = \vec{0}$. A independência linear de $\underline{E}$ sobre $\mathbb{C}$ implica $b_k + ic_k = 0$ e, portanto, $b_k = c_k = 0$ para todo k. Assim, $\underline{E}_{\mathbb{R}}$ é uma base real.

[b] Escreva $B = (b_{kj})$ e $C = (c_{kj})$. Para $j = 1, \dots, n$, temos

$$T(\vec{e}_j) = \sum_{k=1}^{m} (b_{kj} + ic_{kj})\vec{f}_k = \sum_{k=1}^{m} b_{kj}\vec{f}_k + \sum_{k=1}^{m} c_{kj}(i\vec{f}_k).$$

Como T é linear sobre $\mathbb{C}$,

$$T(i\vec{e}_j) = iT(\vec{e}_j) = -\sum_{k=1}^{m} c_{kj}\vec{f}_k + \sum_{k=1}^{m} b_{kj}(i\vec{f}_k).$$

Portanto, a matriz de $T^{\mathbb{R}}$ é

$$\begin{bmatrix} B & -C \\ C & B \end{bmatrix},$$

o que completa a prova. ■

Corolário 3.68.
Seja $T : V \to V$ um operador linear sobre um espaço vetorial complexo de dimensão finita. Então

$$\det T^{\mathbb{R}} = |\det T|^2 .$$

Demonstração. Se $A = B + iC$ representa T em uma base complexa, o teorema anterior mostra que $T^{\mathbb{R}}$ é representado por

$$M = \begin{bmatrix} B & -C \\ C & B \end{bmatrix}.$$

Sobre $\mathbb{C}$, a matriz M é semelhante à matriz em blocos $\operatorname{diag}(B + iC, B - iC)$. Logo

$$\det T^{\mathbb{R}} = \det(B + iC)\det(B - iC) = \det(A)\overline{\det(A)} = |\det T|^2 .$$

■

A construção anterior é um caso particular da restrição de escalares. Seja $\mathbb{K}$ um corpo, seja $\mathbb{F}$ um subcorpo de $\mathbb{K}$ e seja V um espaço vetorial sobre $\mathbb{K}$. Ao restringir a multiplicação por escalares a elementos de $\mathbb{F}$, obtemos um espaço vetorial $V_{\mathbb{F}}$ sobre $\mathbb{F}$. Analogamente, uma transformação $\mathbb{K}$-linear $T : V \to W$ induz uma transformação $\mathbb{F}$-linear $T^{\mathbb{F}} : V_{\mathbb{F}} \to W_{\mathbb{F}}$. Essas operações constituem a **restrição do corpo de escalares** de $\mathbb{K}$ para $\mathbb{F}$.

O próprio corpo $\mathbb{K}$ pode ser considerado um espaço vetorial sobre $\mathbb{F}$. Se $\mathbb{K}$ tem dimensão finita sobre $\mathbb{F}$ e V tem dimensão finita sobre $\mathbb{K}$, então

$$\dim_{\mathbb{F}} V_{\mathbb{F}} = \dim_{\mathbb{F}} \mathbb{K} \cdot \dim_{\mathbb{K}} V.$$

De fato, se $(\vec{e}_1, \dots, \vec{e}_n)$ é uma base de V sobre $\mathbb{K}$ e $(k_1, \dots, k_m)$ é uma base de $\mathbb{K}$ sobre $\mathbb{F}$, então

$$(k_1\vec{e}_1, \dots, k_1\vec{e}_n, \dots, k_m\vec{e}_1, \dots, k_m\vec{e}_n)$$

é uma base de $V_{\mathbb{F}}$ sobre $\mathbb{F}$.

Complexificação. Dado um espaço vetorial real V, considere a soma direta externa $V \oplus V$. Nela, definimos a **estrutura complexa** J por

$$\mathrm{J}(\mathbf{x}, \mathbf{y}) = (-\mathbf{y}, \mathbf{x}).$$

Observe que $\mathrm{J}^2 = -\mathrm{I}$.

Definição 3.69. No espaço vetorial real $V \oplus V$, definimos a multiplicação por números complexos por

$$(a + bi)(\mathbf{x}, \mathbf{y}) \coloneqq a(\mathbf{x}, \mathbf{y}) + b\,\mathrm{J}(\mathbf{x}, \mathbf{y}).$$

Teorema 3.70.
O espaço $V \oplus V$, munido da adição usual e da multiplicação por números complexos da Definição 3.69, é um espaço vetorial complexo, denotado por $V^{\mathbb{C}}$. Sua descomplexificação é $V^{\mathbb{C}}_{\mathbb{R}} = V \oplus V$.

Demonstração. Como J é linear sobre $\mathbb{R}$, as duas leis distributivas decorrem das definições. Para verificar a associatividade da multiplicação por escalares, sejam $a, b, c, d \in \mathbb{R}$ e $\mathbf{z} \in V \oplus V$. Usando $\mathrm{J}^2 = -\mathrm{I}$, obtemos

$$(a+bi)\big((c+di)\mathbf{z}\big) = a\big(c\mathbf{z} + d\,\mathrm{J}(\mathbf{z})\big) + b\,\mathrm{J}\big(c\mathbf{z} + d\,\mathrm{J}(\mathbf{z})\big) \tag{3.29}$$

$$= (ac - bd)\mathbf{z} + (ad + bc)\,\mathrm{J}(\mathbf{z}) \tag{3.30}$$

$$= \big((a+bi)(c+di)\big)\mathbf{z}. \tag{3.31}$$

O escalar 1 atua como a identidade, e os demais axiomas aditivos são os de $V \oplus V$. Portanto, a estrutura definida é a de um espaço vetorial complexo. ■

Definição 3.71 — Complexificação. O espaço $V^{\mathbb{C}}$ é denominado **complexificação** do espaço real V.

A igualdade $V^{\mathbb{C}} = V \oplus V$ descreve uma decomposição de espaços vetoriais reais, não de espaços vetoriais complexos.

Identifiquemos V com o subespaço real $V \oplus \{\vec{0}\}$ de $V^{\mathbb{C}}$. Como $i(\mathbf{v}, \vec{0}) = \mathrm{J}(\mathbf{v}, \vec{0}) = (\vec{0}, \mathbf{v})$, todo vetor de $V^{\mathbb{C}}$ pode ser escrito na forma

$$(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \vec{0}) + (\vec{0}, \mathbf{y}) = (\mathbf{x}, \vec{0}) + i(\mathbf{y}, \vec{0}) = \mathbf{x} + i\mathbf{y}.$$

Consequentemente, qualquer base de V sobre $\mathbb{R}$ também é uma base de $V^{\mathbb{C}}$ sobre $\mathbb{C}$. Em particular, $\dim_{\mathbb{R}} V = \dim_{\mathbb{C}} V^{\mathbb{C}}$.

Definição 3.72. Seja $T : V \to W$ uma transformação linear entre espaços vetoriais reais. A **complexificação** de T é a aplicação $T^{\mathbb{C}} : V^{\mathbb{C}} \to W^{\mathbb{C}}$ definida por

$$T^{\mathbb{C}}(\mathbf{x}, \mathbf{y}) = (T(\mathbf{x}), T(\mathbf{y})).$$

A aplicação $T^{\mathbb{C}}$ é linear sobre $\mathbb{R}$ e comuta com J, pois

$$T^{\mathbb{C}}\,\mathrm{J}(\mathbf{x}, \mathbf{y}) = T^{\mathbb{C}}(-\mathbf{y}, \mathbf{x}) = (-T(\mathbf{y}), T(\mathbf{x})) = \mathrm{J}\,T^{\mathbb{C}}(\mathbf{x}, \mathbf{y}).$$

Portanto, $T^{\mathbb{C}}$ é linear sobre $\mathbb{C}$.

Proposição 3.73.
Sejam $\underline{v}$ e $\underline{w}$ bases dos espaços vetoriais reais V e W, respectivamente. Pela identificação $\mathbf{v} \mapsto (\mathbf{v}, \vec{0})$, podemos considerá-las também bases de $V^{\mathbb{C}}$ e $W^{\mathbb{C}}$. Nessas bases, a matriz de $T^{\mathbb{C}}$ é a mesma matriz real que representa T:

$$[T]_{\underline{v},\underline{w}} = [T^{\mathbb{C}}]_{\underline{v},\underline{w}}.$$

Demonstração. Para cada vetor $\mathbf{v}_j$ da base $\underline{v}$, a expansão de $T^{\mathbb{C}}(\mathbf{v}_j) = T(\mathbf{v}_j)$ na base $\underline{w}$ tem os mesmos coeficientes reais. Logo as matrizes têm as mesmas colunas. ■

Definição 3.74. Uma **conjugação** em um espaço vetorial complexo V é uma função $c : V \to V$ tal que

a $c(\mathbf{u}+\mathbf{v}) = c(\mathbf{u}) + c(\mathbf{v})$ para todos os $\mathbf{u}, \mathbf{v} \in V$;

b $c(z\mathbf{v}) = \overline{z}\, c(\mathbf{v})$ para todos os $z \in \mathbb{C}$ e $\mathbf{v} \in V$;

c $c(c(\mathbf{v})) = \mathbf{v}$ para todo $\mathbf{v} \in V$.

Uma conjugação é linear sobre $\mathbb{R}$, mas, em geral, não é linear sobre $\mathbb{C}$, pois

$$c(i\mathbf{v}) = -ic(\mathbf{v}).$$

Exemplo 3.75. Em $\mathbb{C}^n$, a aplicação $c(z_1, \dots, z_n) = (\overline{z_1}, \dots, \overline{z_n})$, que conjuga cada coordenada, é uma conjugação.

◁

Exemplo 3.76. Em $V^{\mathbb{C}} = V \oplus V$, a aplicação $(\mathbf{u}, \mathbf{v}) \mapsto (\mathbf{u}, -\mathbf{v})$ é uma conjugação.

◁

Exercícios

Ex. 3.38 — Mostre que as complexificações de $\mathbb{R}^n$ e de $\mathscr{M}_{n,n}(\mathbb{R})$ são naturalmente isomorfas a $\mathbb{C}^n$ e a $\mathscr{M}_{n,n}(\mathbb{C})$, respectivamente.

Ex. 3.39 — Mostre que $\mathbb{R}[x]^{\mathbb{C}}$ é naturalmente isomorfo a $\mathbb{C}[x]$.

Ex. 3.40 — Mostre que, se $W = \{\vec{0}\}$, então $W^{\mathbb{C}} = \{\vec{0}\}$. Se $W \neq \{\vec{0}\}$ e $(\mathbf{e}_j)$ é uma base de W sobre $\mathbb{R}$, mostre que $((\mathbf{e}_j, \vec{0}))$ é uma base de $W^{\mathbb{C}}$ sobre $\mathbb{C}$. Em particular,

$$\dim_{\mathbb{C}} W^{\mathbb{C}} = \dim_{\mathbb{R}} W$$

para todo espaço vetorial real W de dimensão finita.

Ex. 3.41 — Seja $\phi : W \to W'$ uma transformação linear entre espaços vetoriais reais. Mostre que sua complexificação $\phi^{\mathbb{C}} : W^{\mathbb{C}} \to (W')^{\mathbb{C}}$ satisfaz

$$\ker(\phi^{\mathbb{C}}) = (\ker \phi)^{\mathbb{C}} \quad \text{e} \quad \operatorname{im}(\phi^{\mathbb{C}}) = (\operatorname{im} \phi)^{\mathbb{C}}.$$

Ex. 3.42 — Se V é um espaço vetorial complexo, uma $\mathbb{R}$-forma de V é um subespaço vetorial real $W \subseteq V$ cuja complexificação é isomorfa a V. Mostre que dois exemplos de $\mathbb{R}$-formas de $\mathscr{M}_{2,2}(\mathbb{C})$ são

$$\mathscr{M}_{2,2}(\mathbb{R}) \quad \text{e} \quad \left\{ \begin{bmatrix} a & b+ci \\ b-ci & d \end{bmatrix} \mid a, b, c, d \in \mathbb{R} \right\}.$$

4

Espaços Quocientes

Fixado um subespaço S de um espaço vetorial V, podemos considerar equivalentes dois vetores cuja diferença pertence a S. O espaço quociente V/S tem essas classes de equivalência como elementos. Geometricamente, cada translação de S corresponde a um único elemento do novo espaço.

As translações de subespaços, ou subespaços afins, permitem descrever essas classes. A estrutura linear do quociente é compatível com a de V: toda transformação linear que anula S induz uma transformação em V/S. Os teoremas de isomorfismo descrevem as relações daí decorrentes entre núcleos, imagens e quocientes.

Nesse capítulo

4.1 Espaços Afins

Começamos com uma generalização da noção de subespaço vetorial. Um subespaço afim é obtido pela translação de um subespaço vetorial e, por isso, preserva suas direções, mas não precisa conter o vetor zero.

Quando duas translações coincidem
As classes $\mathbf{x} + U$ e $\mathbf{y} + U$ são iguais exatamente quando $\mathbf{x} - \mathbf{y} \in U$. Esse critério permite descrever os elementos do quociente V/U como subespaços afins.

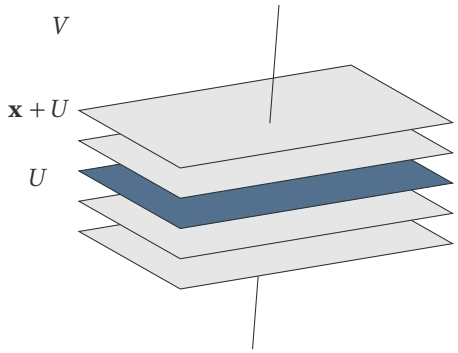
Definição 4.1 — Subespaço Afim. Um subconjunto não vazio X de V é um **subespaço afim** de V paralelo ao subespaço $U \subseteq V$ se existe $\mathbf{x}_0 \in X$ tal que

$$X = \mathbf{x}_0 + U = \{\mathbf{x}_0 + \mathbf{u} \mid \mathbf{u} \in U\}.$$

O subespaço U é chamado de **espaço diretor** de X.

Se $X = \mathbf{x}_0 + U$ e $\mathbf{x}_1 \in X$, então $\mathbf{x}_1 - \mathbf{x}_0 \in U$ e, portanto, $X = \mathbf{x}_1 + U$. Assim, qualquer ponto de X pode servir de origem afim. Além disso, o espaço diretor é único: se $X = \mathbf{x}_0 + U = \mathbf{x}_0 + W$, então $U = W$. Podemos, portanto, definir

$$\dim X := \dim U.$$



Proposição 4.2.
Seja X um subespaço afim paralelo a U. Se $\mathbf{x}_0 \in X$ e $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ é uma base de U, então cada $\mathbf{x} \in X$ admite uma única expressão da forma

$$\mathbf{x} = \mathbf{x}_0 + \sum_{i=1}^{n} c_i \mathbf{u}_i, \qquad c_i \in \mathbb{K}.$$

Demonstração. Se $\mathbf{x} \in X = \mathbf{x}_0 + U$, então $\mathbf{x} - \mathbf{x}_0 \in U$. As coordenadas desse vetor na base $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ existem e são únicas, o que fornece a expressão desejada. ■

Exemplo 4.3. Se for não vazio, o conjunto solução do sistema linear $A\mathbf{x} = \vec{b}$, com $A \in \mathscr{M}_{m,n}(\mathbb{K})$ e $\vec{b} \in \mathbb{K}^m$, é um subespaço afim de $\mathbb{K}^n$ paralelo a $\ker A$. De fato, escolhida uma solução $\mathbf{x}_0$, temos

$$\{\mathbf{x} \in \mathbb{K}^n \mid A\mathbf{x} = \vec{b}\} = \mathbf{x}_0 + \ker A.$$

◁

Definição 4.4. Denotaremos por $\mathrm{Afim}[V]$ o conjunto dos subespaços afins de V e por $\mathrm{Afim}[V, U]$ o conjunto daqueles cujo espaço diretor é U.

Assim, $A \in \mathrm{Afim}[V]$ se, e somente se, $A = \mathbf{x} + W$ para algum subespaço $W \subseteq V$ e algum $\mathbf{x} \in V$.

Teorema 4.5.
Sejam V e V' espaços vetoriais sobre $\mathbb{K}$.

- **a** *A interseção de uma família de subespaços afins de V é vazia ou é um subespaço afim de V.*
- **b** *Se $A, B \in \mathrm{Afim}[V]$, então $A + B \in \mathrm{Afim}[V]$.*
- **c** *Se $A \in \mathrm{Afim}[V]$ e $\lambda \in \mathbb{K}$, então $\lambda A \in \mathrm{Afim}[V]$.*
- **d** *Se $A \in \mathrm{Afim}[V]$ e $\mathbf{x} \in V$, então $\mathbf{x} + A \in \mathrm{Afim}[V]$.*
- **e** *Se $A \in \mathrm{Afim}[V]$ e $T \in \mathrm{Hom}(V, V')$, então $T(A) \in \mathrm{Afim}[V']$.*

Demonstração. Para [a], escreva $A_i = \mathbf{x}_i + W_i$ e suponha que $\bigcap_{i\in\Delta} A_i$ seja não vazia. Escolha $\mathbf{y} \in \bigcap_{i\in\Delta} A_i$. Como $\mathbf{y} \in A_i$, temos $A_i = \mathbf{y} + W_i$ para todo i e, portanto,

$$\bigcap_{i\in\Delta} A_i = \bigcap_{i\in\Delta} (\mathbf{y} + W_i) = \mathbf{y} + \bigcap_{i\in\Delta} W_i.$$

A última igualdade decorre diretamente de $\mathbf{x} \in \mathbf{y} + W_i \iff \mathbf{x} - \mathbf{y} \in W_i$.

Os demais itens seguem das identidades

$$\begin{aligned}
(\mathbf{x} + U) + (\mathbf{y} + W) &= (\mathbf{x} + \mathbf{y}) + (U + W),\\
\lambda(\mathbf{x} + U) &= \begin{cases} \lambda\mathbf{x} + U, & \lambda \neq 0,\\ \{\vec{0}\}, & \lambda = 0, \end{cases}\\
\mathbf{y} + (\mathbf{x} + U) &= (\mathbf{y} + \mathbf{x}) + U,\\
T(\mathbf{x} + U) &= T(\mathbf{x}) + T(U).
\end{aligned}$$

■

Em particular, se fixarmos o espaço diretor, obtemos o seguinte resultado.

Corolário 4.6.
Seja U um subespaço de V.

[a] *Se $A, B \in \mathrm{Afim}[V,U]$, então $A + B \in \mathrm{Afim}[V,U]$.*

[b] *Se $A \in \mathrm{Afim}[V,U]$ e $\lambda \in \mathbb{K}$ é não nulo, então $\lambda A \in \mathrm{Afim}[V,U]$.*

[c] *Se $A \in \mathrm{Afim}[V,U]$ e $\mathbf{x} \in V$, então $\mathbf{x} + A \in \mathrm{Afim}[V,U]$.*

Os conjuntos solução de sistemas lineares são casos particulares do próximo resultado.

Teorema 4.7 (Imagem Inversa).
Sejam $T : V \to W$ uma transformação linear e $\mathbf{w}_0 \in W$. Se $T^{-1}(\mathbf{w}_0)$ for não vazio, então esse conjunto é um subespaço afim de V paralelo a $\ker T$.

Demonstração. Escolha $\mathbf{v}_0 \in V$ tal que $T(\mathbf{v}_0) = \mathbf{w}_0$. Para todo $\mathbf{v} \in V$,

$$T(\mathbf{v}) = \mathbf{w}_0 \iff T(\mathbf{v} - \mathbf{v}_0) = \vec{0} \iff \mathbf{v} - \mathbf{v}_0 \in \ker T.$$

Logo $T^{-1}(\mathbf{w}_0) = \mathbf{v}_0 + \ker T$. ■

Podemos generalizar o Teorema 4.7 por meio de transformações afins. Para $\mathbf{x} \in V$, a translação por $\mathbf{x}$ é a aplicação $S_\mathbf{x} : V \to V$ dada por $S_\mathbf{x}(\mathbf{y}) = \mathbf{x} + \mathbf{y}$. Assim, $\mathbf{x} + W = S_\mathbf{x}(W)$. Se $\mathbf{x} \neq \vec{0}$, essa translação não é linear.

Definição 4.8. Sejam V e V' espaços vetoriais sobre $\mathbb{K}$. Uma aplicação $f : V \to V'$ é uma **transformação afim** se existem $T \in \mathrm{Hom}_\mathbb{K}(V,V')$ e $\mathbf{x} \in V'$ tais que

$$f(\mathbf{v}) = T(\mathbf{v}) + \mathbf{x}$$

para todo $\mathbf{v} \in V$. Denotaremos o conjunto dessas transformações por $\mathrm{Afim}_\mathbb{K}(V,V')$.

Temos $\mathrm{Hom}_{\mathbb{K}}(V,V') \subseteq \mathrm{Afim}_{\mathbb{K}}(V,V') \subseteq (V')^V$. O Teorema 4.5 e admite então a seguinte extensão.

> **Teorema 4.9.**
> *Se $A \in \mathrm{Afim}[V]$ e $f \in \mathrm{Afim}_{\mathbb{K}}(V,V')$, então $f(A) \in \mathrm{Afim}[V']$.*

Demonstração. Se $A = \mathbf{v} + U$ e $f(\mathbf{z}) = T(\mathbf{z}) + \mathbf{x}$, então

$$f(A) = \big(T(\mathbf{v}) + \mathbf{x}\big) + T(U),$$

que é um subespaço afim de V'. ■

Exercícios

Ex. 4.1 — Seja U um subespaço de V e seja $X \subseteq V$ não vazio. Mostre que, se $X = \mathbf{x}_0 + U$ para algum $\mathbf{x}_0 \in X$, então $X = \mathbf{x} + U$ para todo $\mathbf{x} \in X$.

Ex. 4.2 — Seja X um subespaço afim paralelo a U. Fixe $\mathbf{x}_0 \in X$ e uma base $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ de U. Mostre que a aplicação

$$\mathbb{K}^n \longrightarrow X, \qquad (c_1, \dots, c_n) \longmapsto \mathbf{x}_0 + \sum_{i=1}^{n} c_i \mathbf{u}_i$$

é bijetiva. Explique por que ela não precisa ser uma transformação linear.

Ex. 4.3 — Prove as seguintes afirmações.

1. Se $A \in \mathrm{Afim}[V]$ e $\mathbf{x} \in V$, então $\mathbf{x} + A \in \mathrm{Afim}[V]$.
2. Se $A \in \mathrm{Afim}[V]$ e $T \in \mathrm{Hom}(V,V')$, então $T(A) \in \mathrm{Afim}[V']$.
3. Se $A' \in \mathrm{Afim}[V']$ e $T \in \mathrm{Hom}(V,V')$, então $T^{-1}(A')$ é vazio ou é um subespaço afim de V. Determine seu espaço diretor quando A' é paralelo a $U' \subseteq V'$.

Ex. 4.4 — Sejam $\mathbf{v}_1, \dots, \mathbf{v}_r \in V$. Mostre que o conjunto das combinações afins desses vetores,

$$\left\{ \sum_{i=1}^{r} \lambda_i \mathbf{v}_i \;\middle|\; \lambda_i \in \mathbb{K} \text{ e } \sum_{i=1}^{r} \lambda_i = 1 \right\},$$

é o menor subespaço afim de V que contém $\mathbf{v}_1, \dots, \mathbf{v}_r$.

4.2 Espaço Quociente

Classes módulo um subespaço
Passar de V para V/S significa identificar vetores que diferem por um elemento de S. Cada translação de S constitui uma classe, e os vetores do próprio subespaço representam a classe nula.

Seja S um subespaço de V. Dois vetores serão considerados equivalentes quando diferirem por um elemento de S. Mais precisamente, definimos

$$\mathbf{u} \equiv \mathbf{v} \pmod{S} \iff \mathbf{u} - \mathbf{v} \in S.$$

Essa é uma relação de equivalência em V. Quando o subespaço estiver claro, escreveremos apenas $\mathbf{u} \equiv \mathbf{v}$.

As classes de equivalência são precisamente as translações de S, pois

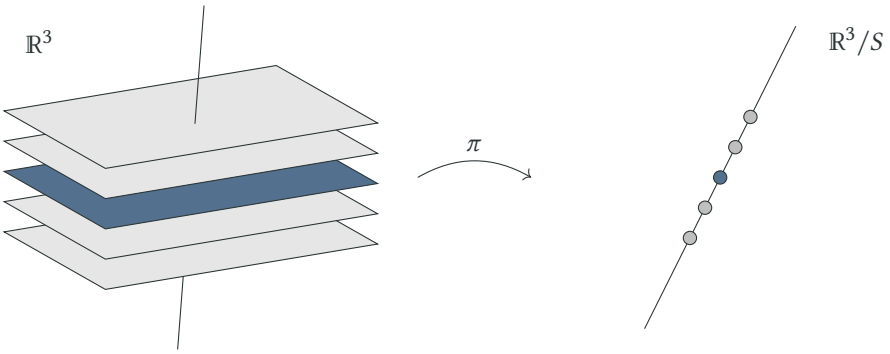
$$[\mathbf{v}] = \{\mathbf{u} \in V \mid \mathbf{u} - \mathbf{v} \in S\} = \{\mathbf{v} + \mathbf{s} \mid \mathbf{s} \in S\} = \mathbf{v} + S.$$

O conjunto $\mathbf{v} + S$ é a **classe lateral** de $\mathbf{v}$ módulo S, e $\mathbf{v}$ é um de seus **representantes**. Dois vetores representam a mesma classe se, e somente se, sua diferença pertence a S.

Definição 4.10 — Espaço Quociente. O conjunto das classes laterais de S em V,

$$V/S = \{\mathbf{v} + S \mid \mathbf{v} \in V\},$$

é denominado **espaço quociente** de V por S.



As operações naturais entre classes são

$$(\mathbf{u} + S) + (\mathbf{v} + S) \coloneqq (\mathbf{u} + \mathbf{v}) + S, \qquad \lambda(\mathbf{u} + S) \coloneqq \lambda\mathbf{u} + S.$$

Como uma classe possui muitos representantes, é necessário verificar que o resultado não depende dos representantes escolhidos.

Teorema 4.11.
As operações acima estão bem definidas e tornam V/S um espaço vetorial sobre $\mathbb{K}$. Seu vetor zero é a classe $\vec{0} + S = S$ e o oposto de $\mathbf{v} + S$ é $-\mathbf{v} + S$.

Demonstração. Suponha que $\mathbf{u}_1 + S = \mathbf{u}_2 + S$ e $\mathbf{v}_1 + S = \mathbf{v}_2 + S$. Então $\mathbf{u}_1 - \mathbf{u}_2, \mathbf{v}_1 - \mathbf{v}_2 \in S$. Como S é um subespaço, para quaisquer $\lambda_1, \lambda_2 \in \mathbb{K}$,

$$(\lambda_1\mathbf{u}_1 + \lambda_2\mathbf{v}_1) - (\lambda_1\mathbf{u}_2 + \lambda_2\mathbf{v}_2) \in S.$$

Portanto,

$$\lambda_1\mathbf{u}_1 + \lambda_2\mathbf{v}_1 + S = \lambda_1\mathbf{u}_2 + \lambda_2\mathbf{v}_2 + S,$$

o que prova, simultaneamente, que a adição e a multiplicação por escalares estão bem definidas. Os axiomas de espaço vetorial são herdados das operações em V; por exemplo,

$$(\mathbf{u} + S) + (\mathbf{v} + S) = (\mathbf{u} + \mathbf{v}) + S = (\mathbf{v} + \mathbf{u}) + S.$$

Os demais axiomas são verificados da mesma maneira. ∎

Exemplo 4.12 — Quociente do plano por uma reta. Considere $V = \mathbb{R}^2$ e $S = \langle (1,0) \rangle$. Como

$$(x, y) + S = (0, y) + S,$$

toda classe é determinada apenas pela segunda coordenada. A aplicação

$$\Phi : V/S \longrightarrow \mathbb{R}, \qquad \Phi\big((x, y) + S\big) = y,$$

está bem definida, é linear e bijetiva. Portanto, $\mathbb{R}^2/S \cong \mathbb{R}$. Geometricamente, cada reta horizontal é comprimida a um único ponto, e o quociente conserva apenas a altura.

◁

Exemplo 4.13 — Polinômios módulo os termos de grau menor. Seja $V = P_2(\mathbb{R})$ e $S = P_1(\mathbb{R})$. Para $p(x) = a + bx + cx^2$, temos

$$p + S = cx^2 + S,$$

pois $a + bx \in S$. Assim, a classe de p registra somente o coeficiente quadrático. A transformação

$$\Psi : V/S \longrightarrow \mathbb{R}, \qquad \Psi(p + S) = c,$$

é um isomorfismo. Em particular, $\dim(V/S) = 1$.

◁

Definição 4.14 — Projeção Canônica. Seja S um subespaço de V. A aplicação

$$\pi_S : V \longrightarrow V/S, \qquad \pi_S(\mathbf{v}) = \mathbf{v} + S,$$

é denominada **projeção canônica**, ou **projeção natural**, de V sobre V/S.

Quando não houver risco de confusão, escreveremos simplesmente π.

Projeção e fatoração
A projeção canônica $\pi : V \to V/S$ envia cada vetor à sua classe lateral e tem núcleo S. Toda aplicação linear que anula S se fatora de maneira única por π. Essa é a propriedade universal do quociente.

Proposição 4.15.
A projeção canônica $\pi_S : V \to V/S$ é linear, sobrejetiva e satisfaz $\ker \pi_S = S$.

Demonstração. Para $\lambda_1, \lambda_2 \in \mathbb{K}$ e $\mathbf{x}, \mathbf{y} \in V$,

$$\pi_S(\lambda_1 \mathbf{x} + \lambda_2 \mathbf{y}) = \lambda_1(\mathbf{x} + S) + \lambda_2(\mathbf{y} + S),$$

logo π_S é linear. Toda classe $\mathbf{v} + S$ é a imagem de $\mathbf{v}$, de modo que π_S é sobrejetiva. Por fim,

$$\mathbf{v} \in \ker \pi_S \iff \mathbf{v} + S = S \iff \mathbf{v} \in S.$$

■

Teorema 4.16 (Teorema da Correspondência).
Seja S um subespaço de V. A atribuição

$$T \longmapsto T/S$$

é uma bijeção, que preserva inclusões, entre os subespaços T de V que contêm S e os subespaços de V/S.

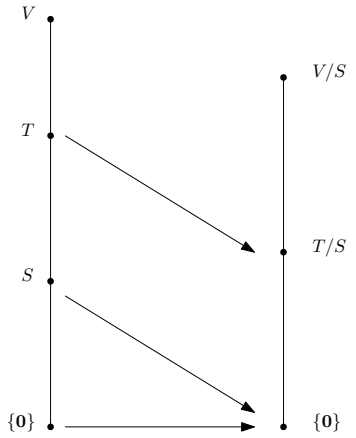


Figura 4.1 Correspondência entre os subespaços de V que contêm S e os subespaços de V/S.

Demonstração. Se $S \subseteq T \subseteq V$, então $T/S = \pi_S(T)$ é um subespaço de V/S. Reciprocamente, se X é um subespaço de V/S, então $\pi_S^{-1}(X)$ é um subespaço de V que contém S. Além disso,

$$\pi_S^{-1}(T/S) = T \quad \text{e} \quad \pi_S\big(\pi_S^{-1}(X)\big) = X,$$

pois π_S é sobrejetiva. As duas construções são, portanto, inversas e ambas preservam inclusões. ■

Exercícios

Ex. 4.5 — Seja $V = \langle\{(1,-1)\}\rangle \subseteq \mathbb{R}^2$.

1. Represente geometricamente V em $\mathbb{R}^2$.
2. Represente as classes
$$C_1 = (1,1) + V, \qquad C_2 = (2,1) + V.$$
3. Descreva o espaço quociente $\mathbb{R}^2/V$.
4. Represente a classe $C_3 = (-2)C_1 + C_2$.
5. Determine se $C_3 = (-1,0) + V$.

Ex. 4.6 — Seja
$$V = \left\{ p \in \mathbb{R}_2[x] \mid \int_{-1}^{1} p(t)\,\mathrm{d}t = 0 \right\}.$$

1. Determine uma base de V.
2. Determine uma base de $\mathbb{R}_2[x]/V$.

Ex. 4.7 — Considere o espaço $\mathbb{K}^\infty$ das sequências de escalares e o subespaço

$$U = \{(0, a_2, a_3, \ldots) \mid a_i \in \mathbb{K}\}.$$

Descreva $\mathbb{K}^\infty/U$ e exiba um isomorfismo desse quociente com um espaço conhecido.

Ex. 4.8 — Suponha que V e seu subespaço S tenham dimensão infinita. A dimensão de V/S precisa ser finita? Dê um exemplo em que seja finita e outro em que seja infinita.

Ex. 4.9 — Demonstre o Teorema da Correspondência construindo explicitamente as duas aplicações inversas entre os conjuntos de subespaços.

Ex. 4.10 — Seja S um subespaço de V. Partindo de uma base de S, descreva como obter uma base de V/S. Justifique por que as classes encontradas são linearmente independentes e geram o quociente.

Ex. 4.11 — Sejam $V \subseteq W$ subespaços de U. Para $\mathbf{u} \in U$, denote por $[\mathbf{u}]_V$ e $[\mathbf{u}]_W$ as classes $\mathbf{u} + V$ e $\mathbf{u} + W$, respectivamente. Mostre que, se

$$[\mathbf{u}_1]_V, \ldots, [\mathbf{u}_k]_V$$

são linearmente dependentes em U/V, então $[\mathbf{u}_1]_W, \ldots, [\mathbf{u}_k]_W$ são linearmente dependentes em U/W. Conclua que, se U tem dimensão finita, então

$$\dim(U/V) \geq \dim(U/W).$$

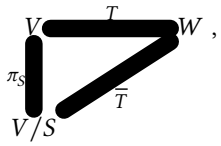
Ex. 4.12 — Seja W um subespaço de dimensão finita de V. Se V/W também tem dimensão finita, mostre que V tem dimensão finita.

4.3 Teoremas de Isomorfismo

Os teoremas de isomorfismo descrevem como núcleos, imagens, somas e quocientes interagem. Sua forma básica é a propriedade universal a seguir: toda transformação linear que anula S pode ser definida diretamente sobre V/S.

Núcleo e aplicação induzida
A aplicação induzida por uma transformação linear no quociente pelo seu núcleo é injetiva e tem a mesma imagem. O Primeiro Teorema de Isomorfismo expressa essa relação: $V/\ker T \cong \operatorname{im} T$.

Teorema 4.17 (Propriedade Universal do Quociente).
Sejam S um subespaço de V e $T \in \operatorname{Hom}(V, W)$ tais que $S \subseteq \ker T$. Existe uma única transformação linear $\overline{T} : V/S \to W$ que torna comutativo o diagrama



isto é, $\overline{T} \circ \pi_S = T$. Essa transformação satisfaz

$$\ker \overline{T} = (\ker T)/S \qquad e \qquad \operatorname{im} \overline{T} = \operatorname{im} T.$$

Demonstração. Para que $\overline{T} \circ \pi_S = T$, é necessário definir

$$\overline{T}(\mathbf{v} + S) = T(\mathbf{v}).$$

Essa definição não depende do representante escolhido, pois

$$\begin{aligned} \mathbf{v} + S = \mathbf{u} + S &\iff \mathbf{v} - \mathbf{u} \in S \\ &\implies \mathbf{v} - \mathbf{u} \in \ker T \\ &\iff T(\mathbf{v}) = T(\mathbf{u}). \end{aligned}$$

Além disso, para $\lambda_1, \lambda_2 \in \mathbb{K}$,

$$\overline{T}\big(\lambda_1(\mathbf{u} + S) + \lambda_2(\mathbf{v} + S)\big) = \lambda_1 T(\mathbf{u}) + \lambda_2 T(\mathbf{v}),$$

de modo que $\overline{T}$ é linear. Pela definição,

$$\begin{aligned} \operatorname{im} \overline{T} &= \{T(\mathbf{v}) \mid \mathbf{v} \in V\} = \operatorname{im} T, \\ \ker \overline{T} &= \{\mathbf{v} + S \mid T(\mathbf{v}) = \vec{0}\} = (\ker T)/S. \end{aligned}$$

Por fim, se $R : V/S \to W$ também satisfaz $R \circ \pi_S = T$, então

$$R(\mathbf{v} + S) = R(\pi_S(\mathbf{v})) = T(\mathbf{v}) = \overline{T}(\mathbf{v} + S)$$

para toda classe $\mathbf{v} + S$. Logo $R = \overline{T}$. ■

Tomando $S = \ker T$, obtemos a forma mais conhecida do resultado.

Teorema 4.18 (Primeiro Teorema de Isomorfismo).
Se $T \in \operatorname{Hom}(V, W)$, então

$$V/\ker T \cong \operatorname{im} T.$$

Demonstração. Considere T com contradomínio restrito a $\operatorname{im} T$. Pelo Teorema 4.17, a aplicação induzida

$$\overline{T} : V/\ker T \longrightarrow \operatorname{im} T, \qquad \overline{T}(\mathbf{v} + \ker T) = T(\mathbf{v}),$$

é sobrejetiva e tem núcleo $(\ker T)/(\ker T) = \{\vec{0}\}$. Portanto, $\overline{T}$ é um isomorfismo. ■

Proposição 4.19.
Se

$$0 \longrightarrow X \xrightarrow{A} Y \xrightarrow{B} Z \longrightarrow 0$$

é uma sequência exata, então a aplicação induzida por B fornece um isomorfismo $Y/\operatorname{im} A \cong Z$.

Demonstração. Pela exatidão, B é sobrejetiva e $\ker B = \operatorname{im} A$. O Primeiro Teorema de Isomorfismo fornece

$$Y/\operatorname{im} A = Y/\ker B \cong \operatorname{im} B = Z.$$

■

O Segundo Teorema de Isomorfismo relaciona soma, interseção e quociente de dois subespaços.

Teorema 4.20 (Segundo Teorema de Isomorfismo).
Se W e W' são subespaços de V, então

$$\frac{W+W'}{W} \cong \frac{W'}{W\cap W'}.$$

Demonstração. Considere a transformação linear

$$T : W' \longrightarrow (W+W')/W, \qquad T(\mathbf{y}) = \mathbf{y} + W.$$

Seu núcleo é $W\cap W'$. Ela é sobrejetiva: se $\mathbf{z}+W \in (W+W')/W$, escreva $\mathbf{z}=\mathbf{x}+\mathbf{y}$, com $\mathbf{x}\in W$ e $\mathbf{y}\in W'$. Então $\mathbf{z}+W = \mathbf{y}+W = T(\mathbf{y})$. O Primeiro Teorema de Isomorfismo fornece

$$\frac{W'}{W\cap W'} = \frac{W'}{\ker T} \cong \operatorname{im} T = \frac{W+W'}{W}.$$

■

O resultado seguinte mostra que tomar quocientes em duas etapas equivale a tomar um único quociente pelo maior dos subespaços.

Teorema 4.21 (Terceiro Teorema de Isomorfismo).
Se $W \subseteq W'$ são subespaços de V, então

$$\frac{V/W}{W'/W} \cong \frac{V}{W'}.$$

Demonstração. Sejam $\pi_W : V \to V/W$ e $\pi_{W'/W} : V/W \to (V/W)/(W'/W)$ as projeções canônicas. A composição

$$T = \pi_{W'/W} \circ \pi_W : V \longrightarrow \frac{V/W}{W'/W}$$

é sobrejetiva. Além disso,

$$\begin{aligned} \mathbf{v} \in \ker T &\iff \mathbf{v}+W \in W'/W \\ &\iff \mathbf{v} \in W'. \end{aligned}$$

Logo $\ker T = W'$, e o Primeiro Teorema de Isomorfismo dá

$$\frac{V}{W'} = \frac{V}{\ker T} \cong \operatorname{im} T = \frac{V/W}{W'/W}.$$

■

Proposição 4.22.
Se $V = V_1 \oplus \cdots \oplus V_n$, então, para cada i,

$$\frac{V}{V_i} \cong V_1 \oplus \cdots \oplus \widehat{V_i} \oplus \cdots \oplus V_n,$$

em que o acento circunflexo indica a omissão do termo correspondente.

Demonstração. Ponha $U_i = \bigoplus_{j \neq i} V_j$. Então $V = V_i \oplus U_i$ e $V_i \cap U_i = \{\vec{0}\}$. Pelo Segundo Teorema de Isomorfismo,

$$V/V_i = (V_i + U_i)/V_i \cong U_i/(V_i \cap U_i) = U_i.$$

■

Definição 4.23 — Codimensão. Se W é um subespaço de V, a **codimensão** de W em V é

$$\operatorname{codim}_V W \coloneqq \dim(V/W).$$

Proposição 4.24.
Se W_2 é um complemento de W_1 em V, então $\operatorname{codim}_V W_1 = \dim W_2$.

Demonstração. Como $V = W_1 \oplus W_2$, a proposição anterior fornece $V/W_1 \cong W_2$. Portanto, esses espaços têm a mesma dimensão. ■

Corolário 4.25.
Se V tem dimensão n e $W \subseteq V$ tem dimensão k, então

$$\dim(V/W) = \operatorname{codim}_V W = n - k.$$

Demonstração. Estenda uma base de W a uma base de V. Os $n-k$ vetores acrescentados geram um complemento de W, ao qual se aplica a proposição anterior. ■

Outra construção natural de espaços quocientes é o conúcleo de uma transformação linear.

Definição 4.26 — Conúcleo. Seja $T : V \to W$ uma transformação linear. O **conúcleo** de T é

$$\operatorname{coker} T \coloneqq W/\operatorname{im} T.$$

Conúcleo e sobrejetividade
O núcleo reúne os vetores que T envia a zero; o conúcleo $W/\operatorname{im} T$ descreve o contradomínio módulo a imagem. Em dimensão finita, sua dimensão é a deficiência de posto da transformação no contradomínio.

Corolário 4.27.
Se V tem dimensão finita e $T : V \to V$ é linear, então

$$\dim(\ker T) = \dim(\operatorname{coker} T).$$

Demonstração. Pelo Teorema do Posto,

$$\dim(\ker T) = \dim V - \dim(\operatorname{im} T) = \dim(V/\operatorname{im} T) = \dim(\operatorname{coker} T).$$

■

Cisão e soma direta
Uma sequência exata curta cinde quando o termo central pode ser identificado com a soma direta dos termos das extremidades, de modo compatível com as aplicações da sequência. A cisão exige uma escolha linear de representantes das classes, e não apenas uma igualdade de dimensões.

4.3.1 Somas diretas e sequências exatas cindidas

Considere $Z = X \oplus Y$. Todo $\mathbf{z} \in Z$ possui uma decomposição única $\mathbf{z} = \mathbf{x} + \mathbf{y}$, com $\mathbf{x} \in X$ e $\mathbf{y} \in Y$. Se $i : X \to Z$ é a inclusão e $\pi : Z \to Y$ é a projeção dada por $\pi(\mathbf{x} + \mathbf{y}) = \mathbf{y}$, então

$$0 \longrightarrow X \xrightarrow{\ i\ } Z \xrightarrow{\ \pi\ } Y \longrightarrow 0$$

é uma sequência exata. A inclusão $j : Y \to Z$ satisfaz $\pi \circ j = \mathrm{I}_Y$ e escolhe, de maneira linear, um representante em Z para cada vetor de Y.

Motivados por esse exemplo, dizemos que uma sequência exata curta

$$0 \longrightarrow X \xrightarrow{\ i\ } Z \xrightarrow{\ \pi\ } Y \longrightarrow 0 \tag{4.1}$$

é **cindida** se π possui uma inversa à direita linear, isto é, se existe $j : Y \to Z$ tal que $\pi \circ j = \mathrm{I}_Y$. A aplicação j é chamada de **seção** de π; ela é necessariamente injetiva.

> **Proposição 4.28.**
> *A sequência* (4.1) *é cindida se, e somente se, existe um isomorfismo* $\Phi : X \oplus Y \to Z$ *tal que*
>
> $$\Phi(\mathbf{x}, \vec{0}) = i(\mathbf{x}) \qquad e \qquad \pi\big(\Phi(\mathbf{x}, \mathbf{y})\big) = \mathbf{y}$$
>
> *para todos* $\mathbf{x} \in X$ *e* $\mathbf{y} \in Y$.

Demonstração. Suponha que a sequência seja cindida e escolha uma seção $j : Y \to Z$. Defina

$$\Phi : X \oplus Y \longrightarrow Z, \qquad \Phi(\mathbf{x}, \mathbf{y}) = i(\mathbf{x}) + j(\mathbf{y}).$$

Para todo $\mathbf{z} \in Z$,

$$\mathbf{z} - j(\pi(\mathbf{z})) \in \ker \pi = \operatorname{im} i,$$

logo Φ é sobrejetiva. Se $\Phi(\mathbf{x}, \mathbf{y}) = \vec{0}$, a aplicação de π fornece $\mathbf{y} = \vec{0}$; então $i(\mathbf{x}) = \vec{0}$ e, como i é injetiva, $\mathbf{x} = \vec{0}$. Portanto, Φ é também injetiva. Por construção, $\Phi(\mathbf{x}, \vec{0}) = i(\mathbf{x})$ e $\pi(\Phi(\mathbf{x}, \mathbf{y})) = \mathbf{y}$.

Reciprocamente, dado um isomorfismo Φ com essas propriedades, defina $j : Y \to Z$ por $j(\mathbf{y}) = \Phi(\vec{0}, \mathbf{y})$. Então $\pi \circ j = \mathrm{I}_Y$, e a sequência é cindida. ∎

Exercícios

Ex. 4.13 — Use o Primeiro Teorema de Isomorfismo para demonstrar o Teorema do Posto:

$$\dim(\ker T) + \operatorname{posto} T = \dim V,$$

para $T \in \operatorname{Hom}(V, W)$ e $\dim V < \infty$.

Ex. 4.14 — Sejam $T \in \operatorname{Hom}(V, V)$ e S um subespaço de V. Considere a fórmula

$$T'(\mathbf{v} + S) = T(\mathbf{v}) + S.$$

1. Determine a condição necessária e suficiente para que essa fórmula defina uma aplicação $T' : V/S \to V/S$.

2. Quando bem definida, prove que T' é linear e determine $\ker T'$ e $\operatorname{im} T'$ em termos de T e S.

Ex. 4.15 — Sejam M e N subespaços de L. Prove diretamente que a aplicação

$$\frac{M+N}{N} \longrightarrow \frac{M}{M\cap N}, \qquad m+n+N \longmapsto m+(M\cap N),$$

é um isomorfismo linear. Em particular, verifique que ela está bem definida.

Ex. 4.16 — Seja L um subespaço unidimensional de $\mathbb{R}^n$. Prove que $\mathbb{R}^n/L \cong \mathbb{R}^{n-1}$.

Ex. 4.17 — Seja $V = C([a,b],\mathbb{R})$ e seja $W \subseteq V$ o subespaço das funções constantes.

1. Prove que $W \cong \mathbb{R}$.

2. Considere $T : V \to V$ dada por $T(f) = f - f(a)$. Determine seu núcleo e sua imagem e conclua que V/W é isomorfo ao espaço das funções contínuas em $[a,b]$ que se anulam em a.

Ex. 4.18 — Se $L = M \oplus N$, prove que a aplicação canônica

$$M \longrightarrow L/N, \qquad m \longmapsto m+N,$$

é um isomorfismo.

Ex. 4.19 — Mostre que a lei de cancelamento para somas diretas pode falhar em dimensão infinita. Mais precisamente, construa espaços isomorfos

$$V = S \oplus B \qquad \text{e} \qquad W = S \oplus D$$

para os quais $B \not\cong D$. Conclua que $V \cong W$ não implica, em geral, $V/S \cong W/S$.

Ex. 4.20 — Suponha que

$$V = S_1 \oplus T_1 = S_2 \oplus T_2$$

e que S_1 e S_2 tenham codimensão finita em V. Prove que $S_1 \cap S_2$ também tem codimensão finita e que

$$\operatorname{codim}_V(S_1 \cap S_2) \le \dim T_1 + \dim T_2.$$

Ex. 4.21 — Sob as hipóteses do exercício anterior, construa uma decomposição $V = W \oplus X$ tal que

1. W tenha codimensão finita;

2. $W \subseteq S_1 \cap S_2$;

3. $X \supseteq T_1 + T_2$.

Ex. 4.22 — Seja $f \in \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$ não nulo. Mostre que $V/\ker f \cong \mathbb{K}$.

Ex. 4.23 — Seja $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in \mathbb{R}[x]$ e considere

$$W = f\mathbb{R}[x] = \{p(x)f(x) \mid p(x) \in \mathbb{R}[x]\}.$$

1. Mostre que W é um subespaço de $\mathbb{R}[x]$.

2. Use o algoritmo da divisão para mostrar que $\dim(\mathbb{R}[x]/W) = n$.

5
Espaços Duais

Um funcional linear transforma vetores em escalares. O espaço dual é formado por todos esses funcionais. Seu estudo relaciona as coordenadas dos vetores, as equações que definem subespaços e as transformações lineares entre espaços. Essas relações são descritas pelas bases duais, pelos aniquiladores e pela transformação transposta.

Nesse capítulo

5.1 Espaços Duais

Definição 5.1 — Funcional Linear. Seja V um espaço vetorial sobre $\mathbb{K}$. Uma transformação linear $f : V \to \mathbb{K}$ é denominada **funcional linear** (ou simplesmente **funcional**) em V.

Interpretação dos funcionais
Um funcional pode ser interpretado como uma medição linear: a cada vetor associa um escalar. O espaço dual reúne todas essas aplicações e permite somá-las e multiplicá-las por escalares.

Os funcionais lineares também são chamados de **covetores**, sobretudo na literatura de Física.

Definição 5.2 — Espaço Dual. O espaço vetorial de todos os funcionais lineares em V é denominado **espaço dual** de V e denotado por

$$V^* = \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K}).$$

Exemplo 5.3. Cada vetor $a = (a_1, \dots, a_n) \in \mathbb{K}^n$ determina o funcional

$$f_a(x_1, \dots, x_n) = \sum_{i=1}^{n} a_i x_i.$$

Reciprocamente, veremos que todo funcional em $\mathbb{K}^n$ tem essa forma. Sobre $\mathbb{R}$, por exemplo, a escolha $a = (1/n, \dots, 1/n)$ fornece o funcional média:

$$\text{média}(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

◁

Exemplo 5.4. A aplicação $E_0 : \mathbb{K}[x] \to \mathbb{K}$, definida por $E_0(p) = p(0)$, é um funcional linear, chamado de **avaliação em** 0. Mais geralmente, para cada $c \in \mathbb{K}$, a expressão $E_c(p) = p(c)$ define o funcional de avaliação em c.

◁

Exemplo 5.5. No espaço $C[a,b]$ das funções reais contínuas em $[a,b]$, a aplicação

$$I : C[a,b] \longrightarrow \mathbb{R}, \qquad I(\alpha) = \int_a^b \alpha(x)\,dx,$$

é um funcional linear.

◁

Exemplo 5.6. No espaço $\mathscr{M}_{n,n}(\mathbb{K})$ das matrizes quadradas $n \times n$, a aplicação traço, $A \mapsto \operatorname{tr}(A)$, é um funcional linear.

◁

Como $\operatorname{im} f$ é um subespaço de $\mathbb{K}$, há apenas duas possibilidades:

- $\operatorname{im} f = \{0\}$, caso em que f é o funcional nulo;
- $\operatorname{im} f = \mathbb{K}$, caso em que f é sobrejetivo.

Portanto, todo funcional linear não nulo é sobrejetivo. Se, além disso, V tem dimensão finita, o teorema do núcleo e da imagem fornece

$$\dim(\ker f) = \dim(V) - 1.$$

Teorema 5.7. [a] *Para cada vetor não nulo $\mathbf{v} \in V$, existe $f \in V^*$ tal que $f(\mathbf{v}) \neq 0$.*

[b] *Um vetor $\mathbf{v} \in V$ é nulo se, e somente se, $f(\mathbf{v}) = 0$ para todo $f \in V^*$.*

[c] *Se $f \in V^*$ e $f(\mathbf{x}) \neq 0$, então*

$$V = \langle \mathbf{x} \rangle \oplus \ker f.$$

[d] *Dois funcionais não nulos $f, g \in V^*$ têm o mesmo núcleo se, e somente se, existe $\lambda \in \mathbb{K}$, com $\lambda \neq 0$, tal que $f = \lambda g$.*

Demonstração. [a] Estenda o conjunto linearmente independente $\{\mathbf{v}\}$ a uma base $\underline{\mathrm{B}}$ de V. Existe um único funcional f que assume o valor 1 em $\mathbf{v}$ e o valor 0 nos demais vetores de $\underline{\mathrm{B}}$. Em particular, $f(\mathbf{v}) = 1 \neq 0$.

[b] Se $\mathbf{v} = \vec{0}$, então $f(\mathbf{v}) = 0$ para todo $f \in V^*$. A recíproca é a contrapositiva do item [a].

[c] Para qualquer $\mathbf{v} \in V$,

$$\mathbf{v} = \frac{f(\mathbf{v})}{f(\mathbf{x})}\mathbf{x} + \left(\mathbf{v} - \frac{f(\mathbf{v})}{f(\mathbf{x})}\mathbf{x}\right),$$

e o termo entre parênteses pertence a $\ker f$. Além disso, se $a\mathbf{x} \in \ker f$, então $af(\mathbf{x}) = 0$ e, portanto, $a = 0$. Logo, a soma é direta.

d Se $f = \lambda g$, com $\lambda \neq 0$, então $\ker f = \ker g$. Reciprocamente, suponha que $K = \ker f = \ker g$ e escolha $\mathbf{x} \notin K$. Pelo item c, $V = \langle \mathbf{x} \rangle \oplus K$. Tomando $\lambda = f(\mathbf{x})/g(\mathbf{x})$, os funcionais f e λg coincidem em $\mathbf{x}$ e se anulam em K; logo, coincidem em todo V. ■

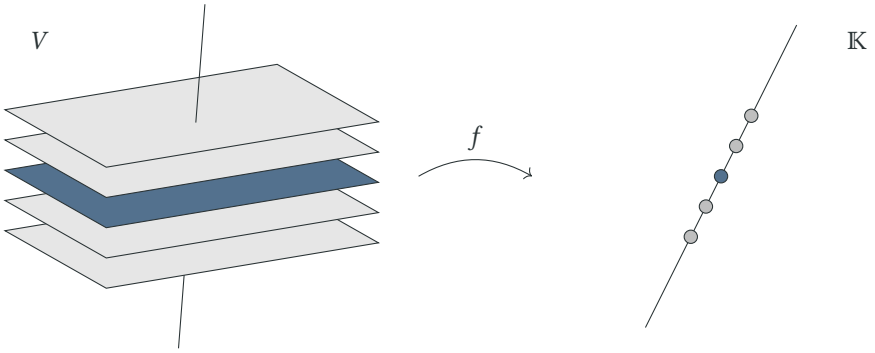


Figura 5.1 Se V tem dimensão finita n e $f \neq 0$, as fibras de f são subespaços afins paralelos a $\ker f$, de dimensão $n - 1$. Cada fibra é levada a um único ponto de $\mathbb{K}$.

5.1.1 Bases duais

Funcionais coordenada
Se $(v_1, \ldots, v_n)$ é uma base, o funcional dual v_i^* devolve a i-ésima coordenada do vetor nessa base. Esses funcionais são determinados pelos valores que assumem nos vetores da base original.

Seja V um espaço vetorial com base $\underline{\mathrm{B}} = \{\mathbf{x}_i \mid i \in \Delta\}$. Para cada $i \in \Delta$, definimos o funcional coordenada $\mathbf{x}_i^* \in V^*$ por

$$\mathbf{x}_i^*(\mathbf{x}_j) = \delta_{ij},$$

onde δ_{ij} é o delta de Kronecker: vale 1 se $i = j$ e 0 caso contrário. Como todo vetor de V é combinação linear de um número finito de elementos de $\underline{\mathrm{B}}$, esses valores determinam $\mathbf{x}_i^*$ de maneira única por linearidade.

Teorema 5.8.
Seja V um espaço vetorial com base $\underline{B} = \{\mathbf{x}_i \mid i \in \Delta\}$.

a *O conjunto $\underline{B}^* = \{\mathbf{x}_i^* \mid i \in \Delta\}$ é linearmente independente.*

b *Se V tem dimensão finita, então $\underline{B}^*$ é uma base de V^*, denominada **base dual** de $\underline{B}$.*

Demonstração. a Considere uma combinação linear finita

$$a_1 \mathbf{x}_{i_1}^* + \cdots + a_r \mathbf{x}_{i_r}^* = 0.$$

Aplicando ambos os lados a $\mathbf{x}_{i_k}$, obtemos $a_k = 0$. Portanto, $\underline{\mathrm{B}}^*$ é linearmente independente.

b Escreva $\underline{\mathrm{B}} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$. Para todo $f \in V^*$, temos

$$f = \sum_{j=1}^{n} f(\mathbf{x}_j)\mathbf{x}_j^*,$$

pois os dois lados assumem o mesmo valor em cada vetor da base $\underline{B}$. Assim, $\underline{B}^*$ também gera V^*. ■

Segue-se do teorema anterior que, se V tem dimensão finita, então

$$\dim(V^*) = \dim(V).$$

Em dimensão infinita, os funcionais coordenada continuam linearmente independentes, mas em geral não geram todo o espaço dual.

Leitura opcional: o dual em dimensão infinita. O restante desta subseção compara cardinalidades em dimensão infinita e pode ser omitido numa primeira leitura.

Exemplo 5.9. Seja V um espaço vetorial de dimensão infinita sobre $\mathbb{Z}_2 = \{0, 1\}$, com base $\underline{B}$. Cada vetor de V corresponde a um subconjunto finito de $\underline{B}$: o conjunto dos vetores da base que aparecem com coeficiente 1 em sua expansão. Portanto,

$$|V| = |\mathscr{P}_0(\underline{B})| = |\underline{B}|,$$

onde $\mathscr{P}_0(\underline{B})$ denota o conjunto das partes finitas de $\underline{B}$. Por outro lado, um funcional em V é determinado livremente por seus valores em $\underline{B}$. Logo, cada funcional corresponde ao subconjunto dos elementos de $\underline{B}$ nos quais assume o valor 1, e

$$|V^*| = |\mathscr{P}(\underline{B})| > |\underline{B}| = |V|.$$

Em particular, V e V^* não podem ser isomorfos. Como o corpo é finito, a cardinalidade de um espaço vetorial de dimensão infinita coincide com a cardinalidade de uma de suas bases; assim, neste exemplo, $\dim(V^*) > \dim(V)$.

◁

Teorema 5.10.
Seja V um espaço vetorial. Então

$$\dim(V) \leq \dim(V^*).$$

Se V tem dimensão finita, ocorre a igualdade.

Demonstração. Os funcionais coordenada associados a uma base de V formam um conjunto linearmente independente em V^*; portanto, $\dim(V) \leq \dim(V^*)$. A igualdade no caso de dimensão finita foi provada no Teorema 5.8. ■

Em dimensão infinita, a desigualdade é sempre estrita. Esse fato é uma consequência do teorema de Erdős–Kaplansky sobre a dimensão do espaço de todas as funções de uma base em $\mathbb{K}$. Como sua demonstração exige resultados de aritmética cardinal que não serão usados adiante, registramos aqui o resultado, mas não o incluímos no enunciado demonstrado.

Exemplo 5.11. Todo funcional em $\mathbb{K}^n$ tem a forma

$$f_a(x_1, \dots, x_n) = a_1 x_1 + \cdots + a_n x_n$$

para um único vetor de coeficientes $a = (a_1, \dots, a_n)$. Assim, $(\mathbb{K}^n)^*$ é canonicamente isomorfo ao espaço dos vetores-linha de comprimento n.

◁

Definição 5.12 — Hiperplano. Seja V um espaço vetorial de dimensão finita n. Um subespaço $H \subseteq V$ é denominado **hiperplano** quando $\dim(H) = n - 1$.

Teorema 5.13.
Seja V um espaço vetorial de dimensão finita.

1. *Se $0 \neq f \in V^*$, então $\ker f$ é um hiperplano.*
2. *Todo hiperplano de V é o núcleo de algum funcional não nulo em V^*.*

Demonstração. O primeiro item decorre do teorema do núcleo e da imagem, pois $\dim(\operatorname{im} f) = 1$. Para o segundo, seja $(\mathbf{v}_1, \dots, \mathbf{v}_{n-1})$ uma base de um hiperplano H e complete-a com $\mathbf{v}_n$ para obter uma base de V. O funcional definido por $f(\mathbf{v}_i) = 0$ para $i < n$ e $f(\mathbf{v}_n) = 1$ satisfaz $\ker f = H$. ■

5.1.2 Bidual

Avaliação e bidual
Cada vetor $v \in V$ define naturalmente um funcional sobre V^* por avaliação: $f \mapsto f(v)$. Em dimensão finita, isso identifica V com V^{**} sem escolher base. Em dimensão infinita, a aplicação continua injetiva, mas pode não ser sobrejetiva.

Aplicando novamente a construção do dual, obtemos o espaço dos funcionais lineares definidos em V^*.

Definição 5.14 — Espaço Bidual. O **espaço bidual** de V é

$$V^{**} := \operatorname{Hom}(V^*, \mathbb{K}).$$

Cada vetor $\mathbf{v} \in V$ determina um elemento de V^{**} por avaliação. Isso define a **aplicação natural**

$$\tau_V : V \longrightarrow V^{**}, \qquad \tau_V(\mathbf{v})(f) = f(\mathbf{v}).$$

A notação τ_V ressalta que a aplicação depende apenas do espaço V, e não da escolha de uma base.

Teorema 5.15.
A aplicação natural $\tau_V : V \to V^{**}$ é linear e injetiva. Se V tem dimensão finita, então τ_V é um isomorfismo.

Demonstração. Para $a, b \in \mathbb{K}$, $\mathbf{u}, \mathbf{v} \in V$ e $f \in V^*$, temos

$$\tau_V(a\mathbf{u} + b\mathbf{v})(f) = f(a\mathbf{u} + b\mathbf{v}) = \big(a\tau_V(\mathbf{u}) + b\tau_V(\mathbf{v})\big)(f).$$

Logo, τ_V é linear. Se $\tau_V(\mathbf{v}) = 0$, então $f(\mathbf{v}) = 0$ para todo $f \in V^*$; pelo Teorema 5.7, $\mathbf{v} = \vec{0}$. Portanto, τ_V é injetiva. Quando V tem dimensão finita,

$$\dim(V^{**}) = \dim(V^*) = \dim(V),$$

de modo que a injetividade implica a sobrejetividade. ■

Um espaço V é denominado **algebricamente reflexivo** quando a aplicação natural τ_V é sobrejetiva. Em dimensão finita, podemos então identificar canonicamente V e V^{**}. O teorema anterior, junto com a desigualdade estrita de Erdős–Kaplansky registrada acima, mostra que um espaço vetorial é algebricamente reflexivo se, e somente se, tem dimensão finita.

Exemplo 5.16. Considere o espaço

$$V = \bigoplus_{k \geq 1} \mathbb{Z}_2 \vec{e}_k,$$

formado pelas sequências binárias com apenas um número finito de entradas iguais a 1. Os funcionais coordenada $\vec{e}_k^*$ são linearmente independentes em V^*. Pelo teorema de extensão de funcionais, existe $F \in V^{**}$ tal que

$$F(\vec{e}_k^*) = 1 \qquad \text{para todo } k \geq 1.$$

Por outro lado, para cada $\mathbf{v} \in V$, temos $\tau_V(\mathbf{v})(\vec{e}_k^*) = \vec{e}_k^*(\mathbf{v}) = 0$ para todo k suficientemente grande. Logo, F não pertence à imagem de τ_V, e V não é algebricamente reflexivo.

◁

Exercícios

Ex. 5.1 — Seja $V = \mathbb{R}_2[x]$ o espaço dos polinômios reais de grau menor ou igual a dois.

1. Mostre que os funcionais $f_1(p) = \displaystyle\int_0^1 p(x)\,dx$, $f_2(p) = \int_0^2 p(x)\,dx$ e $f_3(p) = \int_0^3 p(x)\,dx$ formam uma base de V^*.

2. Encontre a base de V dual à base (f_1, f_2, f_3) de V^*.

Ex. 5.2 — Seja V um espaço vetorial e sejam $\mathbf{v}_1 \neq \mathbf{v}_2$ em V. Prove que existe $f \in V^*$ tal que $f(\mathbf{v}_1) \neq f(\mathbf{v}_2)$.

Ex. 5.3 — Sejam $f, g \in V^*$ não nulos. Suponha que $f(\mathbf{v}) = 0$ se, e somente se, $g(\mathbf{v}) = 0$. Prove que $f = \lambda g$ para algum $\lambda \in \mathbb{K}$ não nulo.

Ex. 5.4 — Sejam $\mathbf{y} \in V \setminus \{\vec{0}\}$ e $f \in V^* \setminus \{0\}$. Defina $T : V \to V$ por $T(\mathbf{x}) = f(\mathbf{x})\mathbf{y}$. Uma transformação definida dessa maneira é denominada **díade**.

1. Mostre que $\dim(\operatorname{im} T) = 1$.
2. Se $S \in \operatorname{Hom}(V, V)$ satisfaz $\dim(\operatorname{im} S) = 1$, mostre que S é uma díade.

Ex. 5.5 — Seja $V = C([0, 1])$ o espaço real das funções contínuas em $[0, 1]$. Defina

$$\phi : V \longrightarrow V^*, \qquad \phi(f)(g) = \int_0^1 f(x)g(x)\,dx.$$

Mostre que ϕ é linear e injetiva.

Ex. 5.6 — Seja V um espaço vetorial de dimensão finita e seja $W = V \oplus V^*$. Usando a aplicação natural $\tau_V : V \to V^{**}$ e a identificação $W^* \cong V^* \oplus V^{**}$, mostre que

$$\Phi : W \longrightarrow W^*, \qquad \Phi(\mathbf{x}, f) = (f, \tau_V(\mathbf{x})),$$

é um isomorfismo.

Ex. 5.7 — Prove que $(V_1 \oplus \cdots \oplus V_n)^* \cong V_1^* \oplus \cdots \oplus V_n^*$.

Ex. 5.8 — Seja $\mathbf{v} \in V$ fixo. Prove que a aplicação de avaliação

$$\tau_V(\mathbf{v}) : V^* \longrightarrow \mathbb{K}, \qquad \tau_V(\mathbf{v})(f) = f(\mathbf{v}),$$

é um elemento de V^{**}.

Ex. 5.9 — Seja $V = \mathbb{R}[x]$. Quais das aplicações a seguir são elementos de V^*?

1. $T(p) = \displaystyle\int_0^1 p(x)dx$.
2. $T(p) = \displaystyle\int_0^1 p(x)^2dx$.
3. $T(p) = \displaystyle\int_0^1 x^2p(x)dx$.
4. $T(p) = p'$.
5. $T(p) = p'(0)$.

Ex. 5.10 — Suponha que $\mathbb{K}$ seja um corpo finito e que $\dim(V) = n$. Para cada $m \leq n$, mostre que o número de subespaços de V de dimensão m é igual ao número de subespaços de dimensão $n - m$.

Ex. 5.11 — Defina $\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii}$ para $A = (a_{ij}) \in \mathscr{M}_{n,n}(\mathbb{K})$. Mostre que

$$\operatorname{tr} \in \left(\mathscr{M}_{n,n}(\mathbb{K})\right)^*.$$

Ex. 5.12 — Mostre que $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ para todo $A, B \in \mathscr{M}_{n,n}(\mathbb{K})$.

Ex. 5.13 — Sejam $m, n \in \mathbb{N}$ e $f_1, \dots, f_m \in (\mathbb{K}^n)^*$. Defina $T : \mathbb{K}^n \to \mathbb{K}^m$ por

$$T(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x})).$$

1. Mostre que $T \in \operatorname{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m)$.
2. Mostre que todo $T \in \operatorname{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m)$ é dado desta maneira para alguns $f_1, \dots, f_m$.

Ex. 5.14 — Seja V um espaço vetorial de dimensão finita sobre $\mathbb{C}$. Dados vetores não nulos $\mathbf{x}_1, \dots, \mathbf{x}_n \in V$, mostre que existe $f \in V^*$ tal que $f(\mathbf{x}_k) \neq 0$ para todo $k = 1, \dots, n$.

5.2 Aniquiladores

Definição 5.17 — Aniquilador. Sejam V um espaço vetorial sobre $\mathbb{K}$ e $M \subseteq V$. O **aniquilador** de M em V^* é o subespaço

$$M^{\perp} = \operatorname{Aniq}_{V^*}(M) \coloneqq \{f \in V^* \mid f(\mathbf{u}) = 0 \text{ para todo } \mathbf{u} \in M\}.$$

Ortogonalidade sem produto interno
O aniquilador de U reúne os funcionais que valem zero em todo vetor de U. Ele desempenha no dual um papel semelhante ao complemento ortogonal, mas sua definição usa apenas a estrutura linear, sem exigir produto interno.

Assim, $M^{\perp}$ reúne os funcionais que anulam todos os vetores de M. Ele é um subespaço de V^* mesmo quando M não é um subespaço de V.

Teorema 5.18. [a] $M^{\perp} = (\langle M \rangle)^{\perp}$.

[b] *Se $M \subseteq N \subseteq V$, então $N^{\perp} \subseteq M^{\perp}$.*

[c] *Se V tem dimensão finita, então*

$$\tau_V(\langle M \rangle) = M^{\perp\perp}.$$

Em particular, sob a identificação natural entre V e V^{**}, temos $S^{\perp\perp} = S$ para todo subespaço $S \subseteq V$.

[d] *Se $S, T \subseteq V$ são subespaços, então*

$$(S \cap T)^{\perp} = S^{\perp} + T^{\perp} \quad e \quad (S + T)^{\perp} = S^{\perp} \cap T^{\perp}.$$

Demonstração. [a] Um funcional se anula em M se, e somente se, pela linearidade, se anula em todas as combinações lineares de elementos de M.

[b] Se f se anula em N, então se anula no subconjunto M.

[c] Ponha $S = \langle M \rangle$. Se $\mathbf{s} \in S$ e $f \in S^\perp$, então $\tau_V(\mathbf{s})(f) = f(\mathbf{s}) = 0$; logo, $\tau_V(S) \subseteq S^{\perp\perp}$. Como V tem dimensão finita, todo elemento de V^{**} é da forma $\tau_V(\mathbf{v})$. Se $\mathbf{v} \notin S$, estenda uma base de S acrescentando $\mathbf{v}$ e defina $g \in V^*$ de modo que g se anule em S e $g(\mathbf{v}) = 1$. Então $g \in S^\perp$, mas $\tau_V(\mathbf{v})(g) \neq 0$. Portanto, $\tau_V(\mathbf{v}) \notin S^{\perp\perp}$, o que prova a inclusão inversa.

[d] Um funcional se anula em $S + T$ se, e somente se, anula S e T; daí

$$(S + T)^\perp = S^\perp \cap T^\perp.$$

Além disso, $S^\perp + T^\perp \subseteq (S \cap T)^\perp$. Para provar a inclusão inversa, escolha decomposições

$$S = S' \oplus (S \cap T), \qquad T = (S \cap T) \oplus T'$$

e um complemento U tais que

$$V = S' \oplus (S \cap T) \oplus T' \oplus U.$$

Se $f \in (S \cap T)^\perp$, defina g como f em $S' \oplus U$ e como zero em $(S \cap T) \oplus T'$. Defina h como f em T' e como zero nos demais termos. Então $g \in T^\perp$, $h \in S^\perp$ e $f = g + h$. ■

5.2.1 Aniquiladores e Somas Diretas

Teorema 5.19.
Seja $V = S \oplus T$.

[a] *A extensão por zero em S define um isomorfismo $T^* \cong S^\perp$.*

[b] *Se V tem dimensão finita, então*

$$\dim(S^\perp) = \operatorname{codim}_V(S) = \dim(V) - \dim(S).$$

Demonstração. Dado $f \in T^*$, defina $\widetilde{f} \in V^*$ por

$$\widetilde{f}(\mathbf{s} + \mathbf{t}) = f(\mathbf{t}), \qquad \mathbf{s} \in S, \quad \mathbf{t} \in T.$$

A aplicação $f \mapsto \widetilde{f}$ é linear, tem imagem em $S^\perp$ e sua inversa é a restrição a T. Isso prova o primeiro item. Se V tem dimensão finita, então

$$\dim(S^\perp) = \dim(T^*) = \dim(T) = \dim(V) - \dim(S).$$

■

Exemplo 5.20. A hipótese de dimensão finita no segundo item é essencial. Seja V o espaço sobre $\mathbb{Z}_2$ com base enumerável $(\vec{e}_1, \vec{e}_2, \ldots)$, seja $S = \langle \vec{e}_1 \rangle$ e seja $T = \langle \vec{e}_2, \vec{e}_3, \ldots \rangle$. Então $V = S \oplus T$, e T é isomorfo a V. Consequentemente, $S^\perp \cong T^* \cong V^*$, embora S tenha codimensão 1.

◁

O aniquilador fornece uma maneira de descrever o espaço dual de uma soma direta.

Teorema 5.21.
Se $V = S \oplus T$, então

$$V^* = S^\perp \oplus T^\perp.$$

Demonstração. Como $S \cap T = \{\vec{0}\}$, o Teorema 5.18 dá

$$V^* = (S \cap T)^\perp = S^\perp + T^\perp.$$

Além disso, um funcional que pertence a $S^\perp \cap T^\perp$ anula $S + T = V$ e, portanto, é nulo. A soma é, assim, direta. ■

Exercícios

Ex. 5.15 — Seja $A = \{f_1, \dots, f_r\} \subseteq V^*$. Defina o aniquilador de A em V por

$$A^\perp = \{\mathbf{v} \in V \mid f(\mathbf{v}) = 0 \text{ para todo } f \in A\}.$$

Mostre que $A^\perp = \bigcap_{i=1}^{r} \ker f_i$.

Ex. 5.16 — Seja V de dimensão finita, seja $A = \{f_1, \dots, f_r\} \subseteq V^*$ e suponha que $g \in V^*$ se anule em $A^\perp$. Mostre que $g \in \langle A \rangle$.

Ex. 5.17 — Seja V um espaço vetorial de dimensão finita. Para subespaços $W_1, W_2 \subseteq V$, mostre que:

1. $\text{Aniq}(W_1 \cap W_2) = \text{Aniq}(W_1) + \text{Aniq}(W_2)$.
2. $\text{Aniq}(W_1 + W_2) = \text{Aniq}(W_1) \cap \text{Aniq}(W_2)$.

Ex. 5.18 — Seja $U \subseteq V$ um subespaço. Mostre diretamente que $U^\perp$ é um subespaço de V^*.

Ex. 5.19 — Determine os aniquiladores de $\{\vec{0}\}$ e de V.

Ex. 5.20 — Seja V um espaço vetorial de dimensão finita e seja $U \subseteq V$ um subespaço não nulo. Mostre que $U^\perp \neq V^*$.

Ex. 5.21 — Seja V de dimensão finita e seja $U \subseteq V$ um subespaço. Mostre que $\dim V = \dim U + \dim U^\perp$.

Ex. 5.22 — Sejam $U, W \subseteq V$ subespaços. Mostre que $U \subseteq W$ implica $W^\perp \subseteq U^\perp$.

Ex. 5.23 — Suponha que $m < n$ e que $f_1, \dots, f_m$ sejam funcionais em um espaço n-dimensional V. Prove que existe $\mathbf{x} \neq \vec{0}$ tal que $f_j(\mathbf{x}) = 0$ para todo $j = 1, \dots, m$. O que esse resultado afirma sobre sistemas lineares homogêneos?

Ex. 5.24 — Suponha que $m < n$ e que $f_1, \dots, f_m \in V^*$, onde $\dim(V) = n$. Determine as condições sobre $\alpha_1, \dots, \alpha_m \in \mathbb{K}$ para que exista $\mathbf{x} \in V$ tal que $f_j(\mathbf{x}) = \alpha_j$ para todo j. Interprete o resultado em termos de sistemas lineares.

Ex. 5.25 — Seja $V_0 \subset V_1 \subset \cdots \subset V_n$ uma bandeira maximal em V. Mostre que

$$V_n^\perp \subset V_{n-1}^\perp \subset \cdots \subset V_0^\perp$$

é uma bandeira maximal em V^*.

Ex. 5.26 — Seja V um espaço vetorial de dimensão n e seja $W \subsetneq V$ um subespaço de dimensão k.

1. Mostre que existem funcionais $f_1, \dots, f_{n-k} \in V^*$ tais que
$$W = \bigcap_{i=1}^{n-k} \ker f_i.$$
2. Conclua, usando o Teorema 5.13, que W é a interseção de $n-k$ hiperplanos.

5.3 Transposta de uma Transformação Linear

Definição 5.22 — Transposta de uma Transformação. Seja $T : V \to W$ uma transformação linear. A **transposta algébrica** de T é a transformação

$$T^* : W^* \longrightarrow V^*, \qquad T^*(f) = f \circ T.$$

Assim, para $f \in W^*$ e $\mathbf{v} \in V$,

$$\big(T^*(f)\big)(\mathbf{v}) = f\big(T(\mathbf{v})\big).$$

Também usaremos a notação T^{t} para essa transformação.

As setas se invertem
Se $T : V \to W$, a transposta algébrica leva funcionais de W a funcionais de V por composição: $f \mapsto f \circ T$. O sentido da seta se inverte porque uma medição feita depois de T pode ser transportada de volta ao domínio.

Teorema 5.23 (Propriedades da Transposta de uma Transformação).

a *Para $T, S \in \mathrm{Hom}(V, W)$ e $a, b \in \mathbb{K}$,*
$$(aT + bS)^* = aT^* + bS^*.$$

b *Para $S \in \mathrm{Hom}(V, W)$ e $T \in \mathrm{Hom}(W, U)$,*
$$(T \circ S)^* = S^* \circ T^*.$$

c *Se $T : V \to W$ é invertível, então*

$$(T^{-1})^* = (T^*)^{-1}.$$

Demonstração. As igualdades seguem diretamente da definição. Para $f \in W^*$,

$$(aT + bS)^*(f) = f \circ (aT + bS) = a(f \circ T) + b(f \circ S).$$

Para $f \in U^*$,

$$(T \circ S)^*(f) = f \circ T \circ S = S^*\big(T^*(f)\big).$$

Por fim,

$$T^* \circ (T^{-1})^* = (T^{-1} \circ T)^* = \mathrm{I}_{V^*},$$

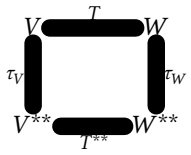
e, analogamente, $(T^{-1})^* \circ T^* = \mathrm{I}_{W^*}$. ■

Aplicando duas vezes a construção anterior, obtemos $T^{**} : V^{**} \to W^{**}$. Embora T e T^{**} tenham domínios e contradomínios diferentes, eles são compatíveis com as aplicações naturais.

Teorema 5.24.
Seja $T : V \to W$ linear. Então

$$T^{**} \circ \tau_V = \tau_W \circ T.$$

Em outras palavras, o diagrama a seguir é comutativo:



Se V e W têm dimensão finita, as aplicações naturais identificam T^{**} com T.

Demonstração. Para $\mathbf{v} \in V$ e $f \in W^*$,

$$\big(T^{**}(\tau_V(\mathbf{v}))\big)(f) = \tau_V(\mathbf{v})(T^* f) = (f \circ T)(\mathbf{v}) = \tau_W(T\mathbf{v})(f).$$

Isso prova a igualdade das transformações. ■

O próximo resultado descreve o núcleo e a imagem da transposta de uma transformação linear.

Teorema 5.25.
Seja $T : V \to W$ linear. Então

a $\ker(T^*) = (\operatorname{im} T)^\perp$;

b $\operatorname{im}(T^*) = (\ker T)^\perp$.

Demonstração. [a] Um funcional $g \in W^*$ pertence a $\ker(T^*)$ se, e somente se, $g \circ T = 0$, isto é, se, e somente se, g se anula em $\operatorname{im} T$.

[b] Se $f = T^*(g) = g \circ T$, então f se anula em $\ker T$. Logo, $\operatorname{im}(T^*) \subseteq (\ker T)^\perp$.

Reciprocamente, seja $f \in (\ker T)^\perp$ e escolha um complemento S de $\ker T$ em V. A restrição $T\restriction_S : S \to \operatorname{im} T$ é um isomorfismo. Defina $g_0 : \operatorname{im} T \to \mathbb{K}$ por

$$g_0(T\mathbf{s}) = f(\mathbf{s}).$$

Essa definição é válida, e g_0 pode ser estendido a um funcional $g \in W^*$. Se $\mathbf{v} = \mathbf{u} + \mathbf{s}$, com $\mathbf{u} \in \ker T$ e $\mathbf{s} \in S$, então

$$T^*(g)(\mathbf{v}) = g(T\mathbf{s}) = f(\mathbf{s}) = f(\mathbf{v}).$$

Portanto, $f = T^*(g)$. ∎

Corolário 5.26.
Seja $T : V \to W$ linear, com V e W de dimensões finitas. Então

$$\operatorname{posto}(T) = \operatorname{posto}(T^*).$$

Demonstração. Pelo teorema anterior e pelo Teorema 5.19,

$$\dim(\operatorname{im} T^*) = \dim((\ker T)^\perp) = \dim(V) - \dim(\ker T) = \operatorname{posto}(T).$$

∎

Teorema 5.27.
Seja $T : V \to W$ linear, com V e W de dimensões finitas. Se $\underline{B} = (\vec{b}_1, \ldots, \vec{b}_n)$ e $\underline{C} = (\vec{c}_1, \ldots, \vec{c}_m)$ são bases de V e W, respectivamente, então

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

Demonstração. Escreva $T(\vec{b}_j) = \sum_{i=1}^{m} a_{ij}\vec{c}_i$. Para todo i e j,

$$\left(T^*(\vec{c}_i^{\,*})\right)(\vec{b}_j) = \vec{c}_i^{\,*}(T(\vec{b}_j)) = a_{ij}.$$

Portanto, as coordenadas de $T^*(\vec{c}_i^{\,*})$ na base $\underline{B}^*$ formam a i-ésima linha de $[T]_{\underline{B},\underline{C}}$, o que prova a fórmula. ∎

Exercícios

Ex. 5.27 — Se $V = S \oplus T$, prove que $V^* = S^\perp \oplus T^\perp$.

Ex. 5.28 — Mostre que a transposta da transformação nula é nula e que $(\mathrm{I}_V)^* = \mathrm{I}_{V^*}$.

Ex. 5.29 — Seja $S \subseteq V$ um subespaço. Prove que $(V/S)^* \cong S^\perp$.

Ex. 5.30 — Verifique as seguintes identidades:

1. $(T+S)^* = T^* + S^*$, para $T, S \in \mathrm{Hom}(V,W)$.
2. $(\lambda T)^* = \lambda T^*$, para $\lambda \in \mathbb{K}$ e $T \in \mathrm{Hom}(V,W)$.

Ex. 5.31 — Sejam V e W de dimensões finitas e $T \in \mathrm{Hom}(V,W)$. Prove que $\mathrm{posto}(T) = \mathrm{posto}(T^*)$.

Ex. 5.32 — Se $T : U \to V$ e $S : V \to W$ são lineares, prove que

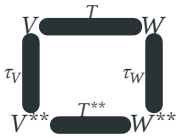
$$(S \circ T)^* = T^* \circ S^*.$$

Observe a inversão da ordem na transposta de uma composição.

Ex. 5.33 — Se $T : V \to W$ é invertível, prove que $T^* : W^* \to V^*$ também é invertível e que $(T^{-1})^* = (T^*)^{-1}$.

Ex. 5.34 — Se T é uma díade em V, mostre que T^* é uma díade em V^*.

Ex. 5.35 — Sejam $\tau_V : V \to V^{**}$ e $\tau_W : W \to W^{**}$ as aplicações naturais. Mostre que, para todo $T \in \mathrm{Hom}(V,W)$, o diagrama a seguir é comutativo:



Ex. 5.36 — Se

é uma sequência exata curta de espaços vetoriais, mostre que

é exata.

Ex. 5.37 — Sejam V, W espaços vetoriais de dimensão finita e $T \in \mathrm{Hom}(V,W)$. Mostre que:

1. T é injetiva se, e somente se, T^* é sobrejetiva.
2. T é sobrejetiva se, e somente se, T^* é injetiva.
3. T é bijetiva se, e somente se, T^* é bijetiva.

5.4 Sistemas de Equações

Equações são funcionais
Cada linha de uma matriz define um funcional linear. Resolver $Ax = 0$ significa encontrar os vetores anulados simultaneamente por todos esses funcionais. Assim, o espaço solução é o aniquilador do espaço gerado pelas linhas de A.

Um sistema de m equações lineares em n variáveis pode ser escrito como

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

onde A é a matriz dos coeficientes. Nesta seção, usaremos funcionais lineares para interpretar o sistema homogêneo $A\mathbf{x} = \vec{0}$.

Seja $\underline{\mathrm{E}} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ a base padrão de $\mathbb{K}^n$, e seja $\underline{\mathrm{E}}^*$ sua base dual. Todo funcional $f \in (\mathbb{K}^n)^*$ pode ser escrito como

$$f = a_1\mathbf{e}_1^* + \cdots + a_n\mathbf{e}_n^*.$$

Para $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$, temos

$$f(\mathbf{x}) = a_1x_1 + \cdots + a_nx_n.$$

Assim, uma equação linear homogênea é precisamente a condição de pertencer ao núcleo de um funcional.

Teorema 5.28.
Seja $A = (a_{ij}) \in \mathscr{M}_{m,n}(\mathbb{K})$ e, para $i = 1, \dots, m$, defina

$$f_i = a_{i1}\mathbf{e}_1^* + \cdots + a_{in}\mathbf{e}_n^*.$$

Então

$$\{\mathbf{x} \in \mathbb{K}^n \mid A\mathbf{x} = \vec{0}\} = \bigcap_{i=1}^{m} \ker f_i.$$

Demonstração. A i-ésima coordenada de $A\mathbf{x}$ é $f_i(\mathbf{x})$. Portanto, $A\mathbf{x} = \vec{0}$ se, e somente se, $f_i(\mathbf{x}) = 0$ para todo i. ∎

Logo, resolver um sistema homogêneo equivale a determinar a interseção dos núcleos dos funcionais associados às linhas de sua matriz.

As operações elementares sobre as linhas admitem a seguinte interpretação na lista $(f_1, \dots, f_m)$:

- trocar duas linhas corresponde a permutar dois funcionais;
- multiplicar uma linha por $\lambda \neq 0$ corresponde a substituir f_i por λf_i;
- somar λ vezes a linha i à linha j corresponde a substituir f_j por $f_j + \lambda f_i$.

Cada operação preserva o subespaço gerado pelos funcionais. Portanto, se U é obtida de A por operações elementares, os funcionais associados às linhas de A e de U geram o mesmo subespaço de $(\mathbb{K}^n)^*$. Como as linhas não nulas de uma matriz escalonada reduzida são linearmente independentes, elas formam uma base desse subespaço.

Já demonstramos no Teorema 2.80, por meio de operações elementares, que os postos por linhas e por colunas coincidem. A dualidade fornece agora uma segunda demonstração.

Teorema 5.29.
Seja A uma matriz sobre $\mathbb{K}$. Então seu posto por linhas é igual ao seu posto por colunas:

$$\text{posto}_l\, A = \text{posto}_c\, A.$$

Demonstração. Se T_A é a transformação representada por A, então $\text{posto}_c\, A = \text{posto}(T_A)$. A matriz da transposta T_A^* nas bases duais é A^{t}; portanto, $\text{posto}_l\, A = \text{posto}_c(A^{\mathrm{t}}) = \text{posto}(T_A^*)$. Pelo corolário anterior, $\text{posto}(T_A^*) = \text{posto}(T_A)$, e o resultado segue. ■

Proposição 5.30.
Seja $A \in \mathscr{M}_{m,n}(\mathbb{K})$ e seja U sua forma escalonada reduzida. Então

$$U\mathbf{v} = \vec{0} \iff A\mathbf{v} = \vec{0}.$$

Demonstração. Sejam $f_1, \dots, f_m$ os funcionais associados às linhas de A e $g_1, \dots, g_r$ os associados às linhas não nulas de U. As operações elementares preservam o espaço gerado pelas linhas; logo,

$$\langle f_1, \dots, f_m \rangle = \langle g_1, \dots, g_r \rangle.$$

Os dois espaços gerados têm o mesmo aniquilador em $\mathbb{K}^n$. Pelo teorema anterior, esse aniquilador é o conjunto de soluções tanto de $A\mathbf{v} = \vec{0}$ quanto de $U\mathbf{v} = \vec{0}$. ■

6
Produto Interno

Os axiomas de espaço vetorial descrevem a adição e a multiplicação por escalares, mas não definem comprimento, ângulo ou perpendicularidade. O produto interno acrescenta essas noções à estrutura linear.

O produto escalar de $\mathbb{R}^n$ fornece o modelo para a definição abstrata. A partir de suas propriedades, definimos a norma, a ortogonalidade e as projeções; a adjunta relaciona o produto interno aos operadores. A seção opcional sobre espaços de Hilbert trata dessas construções quando a dimensão é infinita e a completeza precisa ser considerada.

Nesse capítulo

Definição 6.1 — Produto Escalar. Sejam $\mathbf{u} = (u_1, \dots, u_n)$ e $\mathbf{v} = (v_1, \dots, v_n)$ vetores de $\mathbb{R}^n$. O **produto escalar** de $\mathbf{u}$ e $\mathbf{v}$ é

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{n} u_i v_i = u_1 v_1 + \cdots + u_n v_n.$$

O comprimento de $\mathbf{u}$ e o ângulo entre dois vetores não nulos podem ser recuperados desse único número:

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}}, \qquad \cos\theta(\mathbf{u}, \mathbf{v}) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}.$$

Em particular, $\|\mathbf{u}\|$ é a distância do ponto $\mathbf{u}$ à origem.

As propriedades algébricas que tornam possível essa interpretação são as seguintes.

Teorema 6.2.
Sejam $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$. Então:

1. $\mathbf{u} \cdot \mathbf{u} \geq 0$ *e* $\mathbf{u} \cdot \mathbf{u} = 0$ *se, e somente se,* $\mathbf{u} = \vec{0}$. *Dizemos que o produto escalar é positivo definido.*
2. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$. *Dizemos que o produto escalar é simétrico.*
3. $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$. *Dizemos que o produto escalar é aditivo no primeiro argumento.*
4. $(\lambda \mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (\lambda \mathbf{v}) = \lambda(\mathbf{u} \cdot \mathbf{v})$. *Dizemos que o produto escalar é homogêneo em relação aos escalares.*

Demonstração. Escreva $\mathbf{u} = (u_1, \dots, u_n)$, $\mathbf{v} = (v_1, \dots, v_n)$ e $\mathbf{w} = (w_1, \dots, w_n)$. Então

$$\mathbf{u} \cdot \mathbf{u} = \sum_{i=1}^{n} u_i^2 \geq 0,$$

e essa soma é zero exatamente quando $u_i = 0$ para todo i, isto é, quando $\mathbf{u} = \vec{0}$. Além disso,

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{n} u_i v_i = \sum_{i=1}^{n} v_i u_i = \mathbf{v} \cdot \mathbf{u},$$

$$(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \sum_{i=1}^{n} (u_i + v_i) w_i = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w},$$

e

$$(\lambda \mathbf{u}) \cdot \mathbf{v} = \sum_{i=1}^{n} \lambda u_i v_i = \lambda(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (\lambda \mathbf{v}).$$

■

Tomaremos essas propriedades como axiomas e as transportaremos para espaços vetoriais abstratos. No caso complexo, a simetria precisa envolver conjugação: é ela que garante que $\langle \mathbf{v}, \mathbf{v} \rangle$ seja real e não negativo.

6.1 Produto Interno

Geometria codificada em escalares
O produto interno transforma pares de vetores em números que registram comprimento, ângulo e ortogonalidade. No caso complexo, a conjugação em uma das entradas é indispensável para garantir que $\langle v, v \rangle$ seja real e não negativo.

Neste capítulo, $\mathbb{K}$ denotará $\mathbb{R}$ ou $\mathbb{C}$, e $\overline{\lambda}$ indicará o conjugado de $\lambda \in \mathbb{C}$.

Definição 6.3 — Produto Interno. Seja V um espaço vetorial sobre $\mathbb{K}$. Um **produto interno** em V é uma função $\langle\ ,\ \rangle : V \times V \to \mathbb{K}$ com as seguintes propriedades:

a **Linear na primeira coordenada**: para todos $\mathbf{u}, \mathbf{v} \in V$ e $\lambda_1, \lambda_2 \in \mathbb{K}$

$$\langle \lambda_1 \mathbf{u} + \lambda_2 \mathbf{v},\ \mathbf{w} \rangle = \lambda_1 \langle \mathbf{u},\ \mathbf{w} \rangle + \lambda_2 \langle \mathbf{v},\ \mathbf{w} \rangle$$

b **Simétrico**:

$$\langle \mathbf{u},\ \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle} \quad \text{(simetria Hermitiana)}$$

c **Positivo definido**: Para todo $\mathbf{v} \in V$,

$$\langle \mathbf{v}, \mathbf{v} \rangle \geq 0 \text{ e } \langle \mathbf{v}, \mathbf{v} \rangle = 0 \text{ se, e somente se, } \mathbf{v} = \vec{0}.$$

Quando $\mathbb{K} = \mathbb{R}$ a condição de simetria se reduz a

$$\langle \mathbf{u},\ \mathbf{v} \rangle = \langle \mathbf{v},\ \mathbf{u} \rangle \quad \text{(simetria)}\ .$$

Um espaço vetorial V, munido de um produto interno, é denominado espaço com produto interno.

Observe que um subespaço vetorial S de um espaço com produto interno V também é um espaço com produto interno com a restrição do produto interno de V a S.

Proposição 6.4. [a] *Se $\mathbb{K} = \mathbb{R}$, então o produto interno é linear em ambas as coordenadas, ou seja, o produto interno é bilinear.*

[b] *No entanto, se $\mathbb{K} = \mathbb{C}$, então o produto interno é linear na primeira coordenada e conjugado-linear na segunda, isto é,*

$$\langle \mathbf{w}, \lambda_1 \mathbf{u} + \lambda_2 \mathbf{v} \rangle = \overline{\lambda_1} \langle \mathbf{w}, \mathbf{u} \rangle + \overline{\lambda_2} \langle \mathbf{w}, \mathbf{v} \rangle.$$

[c] $\langle \vec{0}, \mathbf{v} \rangle = \langle \mathbf{v}, \vec{0} \rangle = 0$.

Demonstração. Se $\mathbb{K} = \mathbb{R}$, a linearidade na segunda coordenada segue da simetria e da linearidade na primeira. De fato,

$$\begin{aligned} \langle \mathbf{w}, \lambda_1 \mathbf{u} + \lambda_2 \mathbf{v} \rangle &= \langle \lambda_1 \mathbf{u} + \lambda_2 \mathbf{v}, \mathbf{w} \rangle \\ &= \lambda_1 \langle \mathbf{u}, \mathbf{w} \rangle + \lambda_2 \langle \mathbf{v}, \mathbf{w} \rangle \\ &= \lambda_1 \langle \mathbf{w}, \mathbf{u} \rangle + \lambda_2 \langle \mathbf{w}, \mathbf{v} \rangle. \end{aligned}$$

Se $\mathbb{K} = \mathbb{C}$, então

$$\langle \mathbf{w}, \lambda_1 \mathbf{u} + \lambda_2 \mathbf{v} \rangle = \overline{\langle \lambda_1 \mathbf{u} + \lambda_2 \mathbf{v}, \mathbf{w} \rangle} = \overline{\lambda_1} \overline{\langle \mathbf{u}, \mathbf{w} \rangle} + \overline{\lambda_2} \overline{\langle \mathbf{v}, \mathbf{w} \rangle} = \overline{\lambda_1} \langle \mathbf{w}, \mathbf{u} \rangle + \overline{\lambda_2} \langle \mathbf{w}, \mathbf{v} \rangle$$

Finalmente, pela linearidade na primeira coordenada, $\langle \vec{0}, \mathbf{v} \rangle = \langle \vec{0} + \vec{0}, \mathbf{v} \rangle = 2\langle \vec{0}, \mathbf{v} \rangle$, logo $\langle \vec{0}, \mathbf{v} \rangle = 0$. Pela simetria hermitiana, também $\langle \mathbf{v}, \vec{0} \rangle = 0$. ■

Assim, um produto interno complexo é linear na primeira coordenada e conjugado-linear na segunda. Dizemos, por isso, que ele é **sesquilinear**: o prefixo latino *sesqui* significa "uma vez e meia".

Exemplo 6.5. O espaço vetorial $\mathbb{R}^n$ é um espaço com produto interno com o produto escalar, definido por

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle = x_1 y_1 + \cdots + x_n y_n$$

O espaço com produto interno $\mathbb{R}^n$ é geralmente denominado espaço euclidiano n-dimensional.

◁

Exemplo 6.6. O espaço vetorial $\mathbb{C}^n$ é um espaço com produto interno com o produto escalar definido por

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle = x_1 \overline{y}_1 + \cdots + x_n \overline{y}_n$$

Esse espaço com produto interno costuma ser denominado espaço hermitiano n-dimensional.

◁

Exemplo 6.7. Se o espaço vetorial V possui uma base $\underline{\mathrm{B}} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, então

$$\langle \mathbf{x}, \mathbf{y} \rangle = [\mathbf{x}]_{\underline{\mathrm{B}}}^T [\overline{\mathbf{y}}]_{\underline{\mathrm{B}}}$$

define um produto interno em V.

◁

Exemplo 6.8. Para $A, B \in \mathscr{M}_{n,n}(\mathbb{K})$ definimos

$$\langle A, B \rangle = \mathrm{tr}(A^{\mathrm{t}}\overline{B}) .$$

Então $\langle \cdot, \cdot \rangle$ é um produto interno em $\mathscr{M}_{n,n}(\mathbb{K})$ conhecido como produto interno da Frobenius.

◁

Exemplo 6.9. O espaço vetorial $C[a, b]$ de todas as funções contínuas reais no intervalo fechado $[a, b]$ é um espaço de produto interno com

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

Vamos provar apenas que é positivo definido.

Ou seja, vamos provar que se

- f é contínuo no intervalo fechado $[a, b]$.
- $f \geq 0$ em (a, b).

Então, $f = 0$ em $[a, b]$ ou $\int_a^b f(t)dt > 0$.

Sejam m e M os valores mínimo e máximo de f no intervalo $[a, b]$. Portanto, temos $0 \leq m \leq f(t) \leq M$ para todos os t em $[a, b]$.

Se $M = 0$, então f é identicamente nula.

Então agora assuma que $M > 0$. Se $m = M$, então $f(t) = M$ para todos os t e $\int_a^b f(t)dt = M(b - a) > 0$ e terminamos a demonstração.

Assim podemos assumir que $m < M$. Então o número $A = \frac{1}{2}(M + m)$ é positivo e estritamente menor que M. Seja $f(c) = M$, com c em $[a, b]$. Como f é contínua em $[a, b]$, temos que f é maior que A se estivermos perto o suficiente de $t = c$. Logo existe uma vizinhança $[p, q]$ contendo c, de modo que $f(t) \geq A$ nesse intervalo, onde temos: $a \leq p < q \leq b$ e $p \leq c \leq q$. Então, como $f(t) \geq 0$ em $[a, b]$, temos $\int_a^p f(t)dt \geq 0$ e $\int_q^b f(t)dt \geq 0$, então obtemos:

$$\begin{aligned} \int_a^b f(t)dt &= \int_a^p f(t)dt + \int_p^q f(t)dt + \int_q^b f(t)dt \\ &\geq \int_p^q f(t)dt \geq \int_p^q A dt = A(q - p) > 0. \end{aligned}$$

Então $\int_a^b f(t)dt > 0$ e obtemos o resultado desejado. Aplicando esse resultado à função não negativa f^2, concluímos que $\langle f, f\rangle = 0$ implica $f = 0$, como exigido pela positividade definida.

◁

Exemplo 6.10. O espaço vetorial $C^1[a,\ b]$ de todas as funções com a primeira derivada contínua de valor complexo no intervalo fechado $[a,\ b]$ é um espaço complexo de produto interno com

$$\langle f,\ g\rangle = \int_a^b f(x)\overline{g(x)} + f'(x)\overline{g'(x)}dx$$

◁

Exemplo 6.11. No espaço vetorial real $\mathscr{M}_{n,n}(\mathbb{R})$ o produto interno usual é definido da seguinte forma:

$$\langle A, B\rangle = \operatorname{tr}(B^t A)$$

◁

Exemplo 6.12 — ℓ^2. Um dos exemplos de espaços de produto interno mais fundamentais é o espaço vetorial ℓ^2 de todas as sequências reais (ou complexas) (s_n) com a propriedade que

$$\sum |s_n|^2 < \infty$$

munido do produto interno

$$\langle (s_n),\ (t_n)\rangle = \sum_{n=0}^{\infty} s_n\overline{t_n} \tag{6.1}$$

Tais sequências são denominadas quadrado somáveis. Para que o produto interno esteja bem definido a soma à direita na Equação 6.1 deve convergir. Para provar isso, observamos que, se $(s_n), (t_n) \in \ell^2$, então

$$0 \leq (|s_n| - |t_n|)^2 = |s_n|^2 - 2|s_n||t_n| + |t_n|^2$$

e consequentemente

$$2|s_n t_n| \leq |s_n|^2 + |t_n|^2$$

o que implica que a série $\sum_n |s_n t_n|$ converge. Portanto, a série da Equação 6.1 converge absolutamente. Deixamos ao leitor verificar os demais axiomas do produto interno.

◁

Exemplo 6.13 — Matriz de Gram. Seja V um espaço vetorial de dimensão finita e base $\underline{\mathrm{B}} = \{\mathbf{v}_1,\ \ldots,\ \mathbf{v}_n\}$,. Sejam $\mathbf{v} = \sum_{i=1}^n a_i\mathbf{v}_i$ e $\mathbf{u} = \sum_{j=1}^n b_j\mathbf{v}_j$ vetores em V. Então

$$\langle \mathbf{v}, \mathbf{u}\rangle = \sum_{i,j=1}^{n} a_i\overline{b_j}\langle \mathbf{v}_i, \mathbf{v}_j\rangle = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} G \begin{bmatrix} \overline{b_1} \\ \vdots \\ \overline{b_n} \end{bmatrix}.$$

onde $G = [g_{ij}]$ é a matriz definida por $g_{ij} = \langle \mathbf{v}_i, \mathbf{v}_j\rangle$ para $1 \leq i, j \leq n$. A matriz G é conhecida como matriz de Gram na base $\underline{\mathrm{B}}$.

◁

Definição 6.14 — Ângulo. Se V é um espaço com produto interno real e $\mathbf{u}, \mathbf{v} \in V$ são não nulos, o **ângulo** entre eles é o número $\theta(\mathbf{u}, \mathbf{v}) \in [0, \pi]$ definido por

$$\cos\theta(\mathbf{u}, \mathbf{v}) = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\|\mathbf{v}\|}.$$

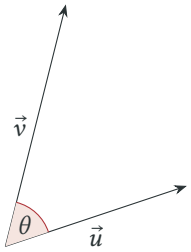


Figura 6.1 Ângulo entre vetores

Definição 6.15 — Vetores Ortogonais. Sejam $\mathbf{u}, \mathbf{v}$ vetores no espaço com produto interno V. Esses vetores são ditos **ortogonais** ou **perpendiculares** se $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. Nesse caso escrevemos $\mathbf{u} \perp \mathbf{v}$.

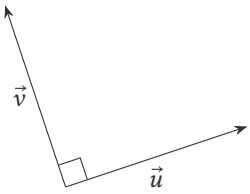


Figura 6.2 Vetores ortogonais

O seguinte resultado simples é bastante útil.

Proposição 6.16.
Se V é um espaço com produto interno $\langle \mathbf{u}, \mathbf{x} \rangle = \langle \mathbf{v}, \mathbf{x} \rangle$ para todo $\mathbf{x} \in V$, então $\mathbf{u} = \mathbf{v}$.

Demonstração. Tomando $\mathbf{x} = \mathbf{u} - \mathbf{v}$, obtemos

$$0 = \langle \mathbf{u}, \mathbf{x} \rangle - \langle \mathbf{v}, \mathbf{x} \rangle = \langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle = \|\mathbf{u} - \mathbf{v}\|^2.$$

Como o produto interno é positivo definido, $\mathbf{u} - \mathbf{v} = \vec{0}$ e, portanto, $\mathbf{u} = \mathbf{v}$. ■

6.2 Normas

Todo produto interno determina uma noção de comprimento. Para $\mathbf{v} \in V$, definimos a **norma** induzida por

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

Normas e produtos internos
Todo produto interno induz uma norma por $\|v\| = \sqrt{\langle v, v \rangle}$. A recíproca é falsa: uma norma determina um produto interno exatamente quando satisfaz a regra do paralelogramo.

Um vetor de norma 1 é denominado **unitário**.

Teorema 6.17. [a] $\|\mathbf{v}\| \geq 0$ *e* $\|\mathbf{v}\| = 0$ *se, e somente se,* $\mathbf{v} = \vec{0}$.

[b] *Para todos* $\lambda \in \mathbb{K}$ *e* $\mathbf{v} \in V$,

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

[c] *(Pitágoras) Se* $\mathbf{v} \perp \mathbf{u}$*, temos*

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

[d] *(Desigualdade de Cauchy-Schwarz-Bunyakovsky) Para todos* $\mathbf{u}, \mathbf{v} \in V$,

$$|\langle \mathbf{u},\ \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|$$

com igualdade se, e somente se, $\mathbf{u}$ *e* $\mathbf{v}$ *forem múltiplos escalares um do outro.*

[e] *(Desigualdade triangular) Para todos* $\mathbf{u}, \mathbf{v} \in V$,

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

com igualdade se, e somente se, um dos vetores for zero ou $\mathbf{u} = \lambda \mathbf{v}$ *para algum escalar real* $\lambda \geq 0$.

[f] *Para todos* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$,

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

[g] *Para todos* $\mathbf{u}, \mathbf{v} \in V$,

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

[h] *(Regra do Paralelogramo) Para todos* $\mathbf{u}, \mathbf{v} \in V$,

$$\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$$

Demonstração. As afirmações [a] e [b] seguem diretamente da definição:

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} \geq 0, \qquad \|\lambda \mathbf{v}\|^2 = \langle \lambda \mathbf{v}, \lambda \mathbf{v} \rangle = |\lambda|^2 \|\mathbf{v}\|^2.$$

Além disso, $\|\mathbf{v}\| = 0$ se, e somente se, $\langle \mathbf{v}, \mathbf{v} \rangle = 0$, o que equivale a $\mathbf{v} = \vec{0}$.

Para demonstrar Pitágoras observamos que

$$\|\mathbf{v} + \mathbf{u}\|^2 = \langle \mathbf{v} + \mathbf{u},\ \mathbf{v} + \mathbf{u} \rangle = \langle \mathbf{v},\ \mathbf{v} \rangle + \langle \mathbf{v},\ \mathbf{u} \rangle + \langle \mathbf{u},\ \mathbf{v} \rangle + \langle \mathbf{u},\ \mathbf{u} \rangle = \|\mathbf{v}\|^2 + \|\mathbf{u}\|^2,$$

pois $\mathbf{v}$ e $\mathbf{u}$ são ortogonais.

Para demonstrar Cauchy–Schwarz, o caso em que um dos vetores é zero é imediato. Se $\mathbf{v} \neq 0$, tome $\lambda = \langle \mathbf{u}, \mathbf{v} \rangle / \langle \mathbf{v}, \mathbf{v} \rangle$. Então

$$0 \leq \|\mathbf{u} - \lambda \mathbf{v}\|^2 = \|\mathbf{u}\|^2 - \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|^2}{\|\mathbf{v}\|^2}.$$

o que equivale à desigualdade de Cauchy–Schwarz–Bunyakovsky. Além disso, a igualdade vale se, e somente se, $\|\mathbf{u} - \lambda \mathbf{v}\|^2 = 0$, isto é, se, e somente se, $\mathbf{u} - \lambda \mathbf{v} = \vec{0}$, o que equivale a $\mathbf{u}$ e $\mathbf{v}$ serem múltiplos escalares um do outro.

Para provar a desigualdade triangular, utilizaremos a desigualdade de Cauchy-Schwarz--Bunyakovsky

$$\begin{aligned}\|\mathbf{u}+\mathbf{v}\|^2 &= \langle \mathbf{u}+\mathbf{v},\ \mathbf{u}+\mathbf{v}\rangle \\ &= \|\mathbf{u}\|^2 + 2\,\mathrm{Re}\langle \mathbf{u},\mathbf{v}\rangle + \|\mathbf{v}\|^2 \\ &\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2 \\ &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2\end{aligned}$$

Na desigualdade acima usamos primeiro $\mathrm{Re}\, z \leq |z|$ e depois Cauchy–Schwarz–Bunyakovsky. Assim, temos a desigualdade triangular. Se nenhum dos vetores é nulo, a igualdade ocorre exatamente quando

$$\mathrm{Re}\langle \mathbf{u},\mathbf{v}\rangle = |\langle \mathbf{u},\mathbf{v}\rangle| = \|\mathbf{u}\|\|\mathbf{v}\|.$$

Pelo caso de igualdade de Cauchy–Schwarz, isso significa que $\mathbf{u} = \lambda\mathbf{v}$; a primeira igualdade força λ a ser real e não negativo. A recíproca é imediata, e o caso em que um vetor é nulo também é imediato.

A afirmação **f** é a desigualdade triangular aplicada a

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

Para **g**, a desigualdade triangular fornece $\|\mathbf{u}\| \leq \|\mathbf{u}-\mathbf{v}\| + \|\mathbf{v}\|$ e, trocando $\mathbf{u}$ e $\mathbf{v}$, também $\|\mathbf{v}\| \leq \|\mathbf{u}-\mathbf{v}\| + \|\mathbf{u}\|$. As duas desigualdades juntas dão o resultado.

Finalmente, expandindo os dois termos do lado esquerdo, obtemos

$$\begin{aligned}\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 &= \langle \mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v}\rangle + \langle \mathbf{u}-\mathbf{v}, \mathbf{u}-\mathbf{v}\rangle \\ &= 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2,\end{aligned}$$

que prova **h**. ■

Podemos generalizar o conceito de norma de modo a torná-lo independente do produto interno.

Definição 6.18 — Norma. Seja V um espaço vetorial sobre o corpo $\mathbb{K}$. Uma **norma** em V é uma aplicação $\|\ \| : V \to [0, \infty)$ satisfazendo às seguintes propriedades:

- **a** $\|\mathbf{v}\| > 0$ se $\mathbf{v} \neq 0$;
- **b** $\|\lambda\mathbf{v}\| = |\lambda|\|\mathbf{v}\|$, para $\lambda \in \mathbb{K}$;
- **c** $\|\mathbf{v}+\mathbf{u}\| \leq \|\mathbf{v}\| + \|\mathbf{u}\|$.

Se V possui uma norma, dizemos que V é um espaço normado.

Exemplo 6.19. Seja V um espaço com produto interno. Então $\|\mathbf{v}\| := \langle \mathbf{v},\mathbf{v}\rangle^{1/2}$ define uma norma em V.

◁

Exemplos 6.20.

1. No espaço vetorial real $\mathbb{R}^n$ temos as seguintes normas
   a) Norma do Máximo: $\|x\|_\infty = \max\{|x_i| \ ; \ 1 \leq i \leq n\}$
   b) Norma do Táxi: $\|x\|_1 = \sum_{i=1}^{n} |x_i|$
2. No espaço vetorial real $\mathscr{M}_{n,n}(\mathbb{K})(\mathbb{R})$. temos as seguintes normas
   a) $\|A\|_\infty = \max\{\sum_{j=1}^{n} |a_{ij}| \ ; \ 1 \leq i \leq n\}$
   b) $\|A\|_1 = \max\{\sum_{i=1}^{n} |a_{ij}| \ ; \ 1 \leq j \leq n\}$

◁

Observamos que nem todas as normas provêm de um produto interno.

O produto interno em V pode ser recuperado da norma associada. Assim, conhecer o comprimento de todos os vetores em V equivale a conhecer todos os produtos internos dos vetores em V.

Teorema 6.21 (Identidades de Polarização). [a] *Se V é um espaço real com produto interno, então*

$$\langle \mathbf{u},\ \mathbf{v}\rangle = \frac{1}{4}(\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2)$$

[b] *Se V é um espaço complexo com produto interno, então*

$$\langle \mathbf{u},\ \mathbf{v}\rangle = \frac{1}{4}(\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2) + \frac{1}{4}i(\|\mathbf{u}+i\mathbf{v}\|^2 - \|\mathbf{u}-i\mathbf{v}\|^2)$$

Demonstração. Expandindo os quadrados pela sesquilinearidade, obtemos

$$\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2 = 4\,\mathrm{Re}\langle \mathbf{u},\mathbf{v}\rangle$$

e, no caso complexo,

$$i\big(\|\mathbf{u}+i\mathbf{v}\|^2 - \|\mathbf{u}-i\mathbf{v}\|^2\big) = 4i\,\mathrm{Im}\langle \mathbf{u},\mathbf{v}\rangle.$$

Somando as duas identidades, obtemos a fórmula complexa; no caso real, a primeira identidade já fornece o resultado. ■

Teorema 6.22 (Jordan–von Neumann).
Seja V um espaço vetorial real ou complexo. Uma norma em V é induzida por um produto interno se, e somente se, satisfaz a regra do paralelogramo

$$\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2.$$

Nesse caso, o produto interno é único e é dado pela identidade de polarização correspondente.

Esboço da demonstração. Se a norma é induzida por um produto interno, a regra do paralelogramo segue expandindo os dois lados. Reciprocamente, no caso real, defina

$$B(\mathbf{x},\mathbf{y}) = \frac{1}{4}\big(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2\big).$$

A regra do paralelogramo implica as identidades $B(\mathbf{x}+\mathbf{z},\mathbf{y}) = B(\mathbf{x},\mathbf{y}) + B(\mathbf{z},\mathbf{y})$ e $B(-\mathbf{x},\mathbf{y}) = -B(\mathbf{x},\mathbf{y})$. Daí segue a homogeneidade para escalares inteiros e racionais. Pela desigualdade triangular invertida, $|\|\mathbf{x}\|-\|\mathbf{y}\|| \le \|\mathbf{x}-\mathbf{y}\|$; portanto, a norma e a função B são contínuas. Aproximando um escalar real por racionais, estendemos a homogeneidade a todos os escalares reais. A simetria é imediata e $B(\mathbf{x},\mathbf{x}) = \|\mathbf{x}\|^2$, logo B é um produto interno que induz a norma dada.

No caso complexo, usa-se a identidade de polarização complexa para definir o produto interno; o mesmo argumento fornece a aditividade, e os termos com i fornecem a homogeneidade complexa. A unicidade, em ambos os casos, decorre diretamente das identidades de polarização. ■

A norma pode ser usada para definir a distância entre dois vetores em um espaço com produto interno.

Definição 6.23 — Distância. Se V é um espaço com produto interno. Definimos a **distância** $d(\mathbf{u},\ \mathbf{v})$ entre os vetores $\mathbf{u}$ e $\mathbf{v}$ em V como

$$d(\mathbf{u},\ \mathbf{v}) = \|\mathbf{u}-\mathbf{v}\|$$

Proposição 6.24.
Se V é um espaço munido da distância $d(\mathbf{u},\ \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$ então:

a $d(\mathbf{u},\ \mathbf{v}) \geq 0$ *e* $d(\mathbf{u},\ \mathbf{v}) = 0$ *se, e somente se,* $\mathbf{u} = \mathbf{v}$.

b $d(\mathbf{u},\ \mathbf{v}) = d(\mathbf{v},\ \mathbf{u})$. *Nesse caso dizemos que d é simétrica.*

c $d(\mathbf{u},\ \mathbf{v}) \leq d(\mathbf{u},\ \mathbf{w}) + d(\mathbf{w},\ \mathbf{v})$. *Nesse caso dizemos que d satisfaz a desigualdade triangular.*

Demonstração. A afirmação a segue da positividade da norma, pois $d(\mathbf{u}, \mathbf{v}) = 0$ se, e somente se, $\mathbf{u} - \mathbf{v} = \vec{0}$. Para b,

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \|-(\mathbf{v} - \mathbf{u})\| = d(\mathbf{v}, \mathbf{u}).$$

Finalmente,

$$d(\mathbf{u}, \mathbf{v}) = \|(\mathbf{u} - \mathbf{w}) + (\mathbf{w} - \mathbf{v})\| \leq d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v}),$$

pela desigualdade triangular da norma. ■

Qualquer conjunto não vazio V, juntamente com uma função $d : V \times V \to \mathbb{R}$ que satisfaça as propriedades da Proposição 6.24, é denominado **espaço métrico** e a função d é dita **métrica** em V. Assim, qualquer espaço com produto interno é um espaço métrico com a métrica induzida pelo produto interno.

A presença de um produto interno e, portanto, de uma métrica, permite a definição de diversos conceitos analíticos em V como convergência de sequências e séries infinitas, funções contínuas e conjuntos abertos, fechados e compactos.

6.3 Bases Ortonormais

Coordenadas ortonormais
Cada coordenada de um vetor em uma base ortonormal é obtida por um produto interno. A ortogonalidade dos vetores da base permite calcular o quadrado da norma pela soma dos quadrados dos módulos das coordenadas.

Definição 6.25 — Conjuntos Ortogonais e Ortonormais. Seja V um espaço com produto interno. Um conjunto não vazio $O = \{\mathbf{u}_i \mid i \in \Delta\} \subset V$ é **ortogonal** se $\mathbf{u}_i \perp \mathbf{u}_j$ sempre que $i \neq j$. Se, além disso, todos os seus vetores são unitários, então O é **ortonormal**. Ou seja, O é ortonormal se

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \delta_{i,j}$$

para todos $i, j \in \Delta$.

Definição 6.26 — Subespaço Ortogonal. Seja V um espaço com produto interno e $A \subset V$. Então definimos o **subespaço ortogonal** a A como

$$A^{\perp} = \{\mathbf{v} \in V \text{ tais que } \mathbf{v} \perp \vec{a},\ \forall \vec{a} \in A\}$$

Lema 6.27.
O conjunto $S^{\perp}$ é um subespaço de V, mesmo que S não o seja. Se S é um subespaço de V então $S \cap S^{\perp} = \{\vec{0}\}$.

Demonstração. O vetor nulo pertence a $S^{\perp}$. Se $\mathbf{u}, \mathbf{v} \in S^{\perp}$ e $\lambda, \mu \in \mathbb{K}$, então, para todo $\mathbf{s} \in S$,

$$\langle \lambda\mathbf{u} + \mu\mathbf{v}, \mathbf{s} \rangle = \lambda\langle \mathbf{u}, \mathbf{s} \rangle + \mu\langle \mathbf{v}, \mathbf{s} \rangle = 0.$$

Logo $\lambda\mathbf{u} + \mu\mathbf{v} \in S^{\perp}$, e $S^{\perp}$ é um subespaço. Se S também é um subespaço e $\mathbf{v} \in S \cap S^{\perp}$, então $\langle \mathbf{v}, \mathbf{v} \rangle = 0$; portanto, $\mathbf{v} = \vec{0}$. ■

Exemplo 6.28. Seja $C[-1,\ 1]$ o espaço vetorial de todas as funções contínuas reais no intervalo fechado $[-1,\ 1]$ munido do produto interno

$$\langle f,\ g \rangle = \int_{-1}^{1} f(x)g(x)dx$$

Então o conjunto $\{\text{sen}(n\pi x)\}_{n=1}^{\infty}$ é ortonormal.

$$\int_{-1}^{1} \text{sen}(n\pi x)\,\text{sen}(m\pi x)dx = \begin{cases} 0 & \text{se } n \neq m \\ 1 & \text{se } m = n \end{cases}$$

$\mathbf{m = n}$:

$$\int_{-1}^{1} \frac{1}{2}\left(\cos(n\pi x - n\pi x) - \cos(n\pi x + n\pi x)\right) \tag{6.2}$$
$$= \frac{1}{2}\int_{-1}^{1} \cos(0) - \frac{1}{2}\int_{-1}^{1} \cos(2n\pi x)dx \tag{6.3}$$
$$= \frac{1}{2}[1-(-1)] - \frac{1}{2}[0-0] \tag{6.4}$$
$$= 1 \tag{6.5}$$

$\mathbf{m \neq n}$:

$$\int_{-1}^{1} \frac{1}{2}\left(\cos(m\pi x - n\pi x) - \cos(m\pi x + n\pi x)\right) dx \tag{6.6}$$
$$= \frac{1}{2}\int_{-1}^{1} \cos((m-n)\pi x)dx - \frac{1}{2}\int_{-1}^{1} \cos((m+n)\pi x)dx \tag{6.7}$$
$$= 0 \tag{6.8}$$

◁

A ortogonalidade é mais forte que a independência linear.

Lema 6.29.
Todo conjunto ortogonal formado por vetores não nulos é linearmente independente.

Demonstração. Sejam $\mathbf{x}_1, \ldots, \mathbf{x}_m \in X$ tais que

$$\alpha_1\mathbf{x}_1 + \ldots + \alpha_m\mathbf{x}_m = \vec{0}.$$

Então

$$0 = \langle \vec{0},\ \mathbf{x}_i\rangle = \langle \alpha_1\mathbf{x}_1 + \ldots + \alpha_m\mathbf{x}_m,\ \mathbf{x}_i\rangle = \alpha_1\langle \mathbf{x}_1,\ \mathbf{x}_i\rangle + \ldots + \alpha_m\langle \mathbf{x}_m,\ \mathbf{x}_i\rangle = \alpha_i\langle \mathbf{x}_i,\ \mathbf{x}_i\rangle.$$

Como $\langle \mathbf{x}_i, \mathbf{x}_i\rangle = \|\mathbf{x}_i\|^2 \neq 0$, temos que $\alpha_i = 0$. ■

Seja O a coleção de todos os conjuntos ortonormais em um espaço com produto interno V. Observe que qualquer conjunto unitário $\{\mathbf{v}\}$ com $\|\mathbf{v}\| = 1$ é um conjunto ortonormal em V. Como uma subfamília do conjunto das partes $\mathscr{P}(V)$ a família O é parcialmente ordenada pela ordem dada pela inclusão de conjuntos.

Definição 6.30 — Sistema Ortonormal Maximal. Um conjunto ortonormal maximal A em um espaço com produto interno V é chamado de **sistema ortonormal maximal**.

Em dimensão infinita, um sistema ortonormal maximal não precisa ser uma base do espaço vetorial. Para evitar ambiguidade, denominaremos uma base no sentido puramente algébrico de **base de Hamel**.

Teorema 6.31.
Seja V um espaço com produto interno e A seja um conjunto ortonormal em V. As seguintes asserções são equivalentes

a *A é um conjunto ortonormal maximal em V.*

b *Não há vetor unitário $\mathbf{v}$ para o qual $A \bigcup \{\mathbf{v}\}$ seja um conjunto ortonormal.*

c *Se $\mathbf{v} \perp A$, $\mathbf{v} = \vec{0}$ (isto é, $A^{\perp} = \{\vec{0}\}$).*

Demonstração. Seja A um conjunto ortonormal em um espaço com produto interno V.

a $\Rightarrow$ b Se existe um vetor unitário $\mathbf{v}$ em V para o qual $A \bigcup \{\mathbf{v}\}$ é um conjunto ortonormal, $A \bigcup \{\mathbf{v}\}$ é um conjunto ortonormal que inclui propriamente A (pois $\mathbf{v} \perp A$). Portanto, A não é um conjunto ortonormal maximal em V.

b $\Rightarrow$ c Se houver um vetor diferente de zero $\mathbf{v}$ em V, de modo que $\mathbf{v} \perp A$, existe um vetor unitário $\mathbf{v}' = \frac{\mathbf{v}}{\|\mathbf{v}\|}$ em V de modo que $A \bigcup \{\mathbf{v}'\}$ é um conjunto ortonormal.

c $\Rightarrow$ a Se a não for verdadeiro, então existe um conjunto ortonormal A' em V que contém propriamente A de modo que $A' \backslash A \neq \emptyset$. Seja $\mathbf{v}$ em $A' \backslash A$, que é um vetor diferente de zero (na verdade, $\mathbf{v}$ é um vetor unitário) ortogonal a A (porque $\mathbf{v} \in A'$, $A \subset A'$ e A' é um conjunto ortonormal). Assim c não é verdadeiro.

■

O Lema de Zorn pode ser usado para mostrar que qualquer espaço com produto interno possui um sistema ortonormal maximal.

Teorema 6.32.
Se A é um conjunto ortonormal em um espaço com produto interno V então existe um sistema ortonormal maximal B em V tal que $A \subseteq B$.

Demonstração. Seja A um conjunto ortonormal em um espaço com produto interno V. Definimos

$$O_A = \{S \in \mathscr{P}(V) \,:\, S \text{ é um conjunto ortonormal em } V \text{ e } A \subseteq S\}$$

a família de todos os conjuntos ortonormais em V que incluem A.

Como O_A é uma subfamília não vazia (pois $A \in O_A$) do conjunto das partes $\mathscr{P}(V)$, ela é parcialmente ordenada pela inclusão. Considere uma cadeia arbitrária $C = \{C_i \,:\, i \in \Delta\}$ em O_A e a união $U = \bigcup_{i \in \Delta} C_i$. Se $\mathbf{v}$ e $\mathbf{u}$ são vetores distintos de U, existem $C_i, C_j \in C$ tais que $\mathbf{v} \in C_i$ e $\mathbf{u} \in C_j$. Como C é uma cadeia, $C_i \subseteq C_j$ ou $C_j \subseteq C_i$. Sem perda de generalidade, suponha $C_i \subseteq C_j$; então $\mathbf{v}, \mathbf{u} \in C_j$ e, portanto, são ortogonais. Além disso, todo vetor de U é unitário e $A \subseteq U$. Logo U é ortonormal e pertence a O_A.

Como U é um limite superior para C, toda cadeia em O_A tem um limite superior em O_A. Pelo Lema de Zorn, O_A possui um elemento maximal B. Esse conjunto é ortonormal e contém A. Se existisse um vetor unitário $\mathbf{v}$ tal que $B \cup \{\mathbf{v}\}$ fosse ortonormal, esse conjunto pertenceria a O_A e conteria propriamente B, em contradição com a maximalidade. Portanto, B é um sistema ortonormal maximal em V que contém A.

■

O próximo resultado fornece um algoritmo que transforma uma sequência linearmente independente em uma sequência ortogonal.

Teorema 6.33 (Processo de Ortogonalização de Gram-Schmidt).
Seja $\underline{B} = (\mathbf{v}_1,\ \mathbf{v}_2,\ \ldots)$ *uma sequência de vetores linearmente independentes em um espaço com produto interno* V*, então a sequência* $O = (\mathbf{u}_1,\ \mathbf{u}_2,\ \ldots)$ *definida por*

$$\mathbf{u}_k = \mathbf{v}_k - \sum_{i=1}^{k-1} \frac{\langle \mathbf{v}_k, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$$

é uma sequência ortogonal em V *com a propriedade que*

$$\langle \mathbf{u}_1, \ldots, \mathbf{u}_k \rangle = \langle \mathbf{v}_1, \ldots, \mathbf{v}_k \rangle$$

para todos os $k > 0$.
A partir da sequência ortogonal $(\mathbf{u}_i)$*, obtemos a sequência ortonormal* $(\mathbf{w}_i)$*, onde* $\mathbf{w}_i = \mathbf{u}_i / \|\mathbf{u}_i\|$.

Subespaços preservados
O processo de Gram–Schmidt altera os vetores, mas preserva a sequência dos subespaços gerados. Em cada etapa, os primeiros vetores obtidos geram o mesmo subespaço que os primeiros vetores da sequência original.

Demonstração. A demonstração é por indução. Para $k = 1$, temos $\mathbf{u}_1 = \mathbf{v}_1 \neq \vec{0}$, e a igualdade dos espaços gerados é imediata.

Suponha agora que $\mathbf{u}_1, \ldots, \mathbf{u}_k$ sejam não nulos e ortogonais e que

$$\langle \mathbf{u}_1, \ldots, \mathbf{u}_k \rangle = \langle \mathbf{v}_1, \ldots, \mathbf{v}_k \rangle.$$

Defina

$$\mathbf{u}_{k+1} = \mathbf{v}_{k+1} - \sum_{j=1}^{k} \frac{\langle \mathbf{v}_{k+1}, \mathbf{u}_j \rangle}{\langle \mathbf{u}_j, \mathbf{u}_j \rangle} \mathbf{u}_j.$$

Para $1 \leq i \leq k$, a ortogonalidade dos vetores já construídos dá

$$\begin{aligned} \langle \mathbf{u}_{k+1}, \mathbf{u}_i \rangle &= \langle \mathbf{v}_{k+1}, \mathbf{u}_i \rangle - \sum_{j=1}^{k} \frac{\langle \mathbf{v}_{k+1}, \mathbf{u}_j \rangle}{\langle \mathbf{u}_j, \mathbf{u}_j \rangle} \langle \mathbf{u}_j, \mathbf{u}_i \rangle \\ &= \langle \mathbf{v}_{k+1}, \mathbf{u}_i \rangle - \langle \mathbf{v}_{k+1}, \mathbf{u}_i \rangle = 0. \end{aligned}$$

Logo $\mathbf{u}_{k+1}$ é ortogonal aos vetores anteriores. Ele não pode ser nulo: se fosse, a fórmula mostraria que $\mathbf{v}_{k+1} \in \langle \mathbf{u}_1, \ldots, \mathbf{u}_k \rangle = \langle \mathbf{v}_1, \ldots, \mathbf{v}_k \rangle$, contrariando a independência linear da sequência $(\mathbf{v}_i)$.

A própria fórmula mostra que $\mathbf{u}_{k+1} \in \langle \mathbf{v}_1, \ldots, \mathbf{v}_{k+1} \rangle$, enquanto a identidade obtida ao isolarmos $\mathbf{v}_{k+1}$ mostra a inclusão inversa entre os espaços gerados. Portanto,

$$\langle \mathbf{u}_1, \ldots, \mathbf{u}_{k+1} \rangle = \langle \mathbf{v}_1, \ldots, \mathbf{v}_{k+1} \rangle.$$

Isso completa a indução. Como nenhum $\mathbf{u}_i$ é nulo, sua normalização produz a sequência ortonormal anunciada. ■

Exemplo 6.34. Considere o espaço vetorial $P_3(\mathbb{R})$ munido do produto interno

$$\langle p,\ q\rangle = \int_{-1}^{1} p(x)q(x)dx.$$

Podemos a partir da base $\underline{\mathrm{B}} = \{1,\ x,\ x^2,\ x^3\}$ obter uma base ortogonal utilizando o processo de ortogonalização de Gram–Schmidt:

$$P_0(x) = 1, \quad P_1(x) = x, \quad P_2(x) = x^2 - \frac{1}{3} \quad \text{e} \quad P_3(x) = x^3 - \frac{3}{5}x.$$

Os polinômios P_0, P_1, P_2, P_3 são denominados polinômios ortogonais de Legendre.

◁

Exemplo 6.35. Considere o espaço vetorial $P_3(\mathbb{R})$ munido do produto interno

$$\langle p,\ q\rangle = \int_{0}^{\infty} \exp(-x)p(x)q(x)dx.$$

Utilizando o processo de ortogonalização de Gram-Schmidt na base $\underline{\mathrm{B}} = \{1,\ x,\ x^2,\ x^3\}$, obtemos os polinômios

$$\begin{aligned} L_0(x) &= 1 \\ L_1(x) &= x - 1 \\ L_2(x) &= x^2 - 4x + 2 \\ L_3(x) &= x^3 - 9x^2 + 18x - 6 \end{aligned}$$

que são denominados polinômios ortogonais de Laguerre.

Na construção dos polinômios de Laguerre utilizamos que $\int_0^{\infty} \exp(-x)x^n dx = n!$ para $n \in \mathbb{N}$

◁

Teorema 6.36.
Seja V um espaço com produto interno.

- **a** *Se $\dim V = n$ for finita, então V terá uma base ortonormal de tamanho n e todos os sistemas ortonormais maximais terão tamanho n.*
- **b** *Se V possuir um sistema ortonormal maximal finito com tamanho n, então $\dim V = n$.*

Demonstração. Para a parte **a**, a aplicação do processo de Gram–Schmidt a uma base de Hamel fornece uma base ortonormal de tamanho n. Um sistema ortonormal tem no máximo n elementos; se tiver menos, pode ser ampliado aplicando Gram–Schmidt após acrescentar um vetor fora de seu espaço gerado. Logo todo sistema ortonormal maximal tem tamanho n.

Para a parte **b**, se $\dim V > n$, podemos acrescentar a um sistema ortonormal maximal de tamanho n um vetor fora de seu espaço gerado e aplicar Gram–Schmidt, obtendo um sistema ortonormal maior, contradição.

■

6.4 Melhor Aproximação

O erro deve ser ortogonal
Se w é a melhor aproximação de v dentro de um subespaço W, então o erro $v - w$ é ortogonal a todas as direções disponíveis em W. Caso contrário, ainda seria possível deslocar w dentro de W e diminuir a distância.

Bases ortonormais têm uma grande vantagem sobre bases arbitrárias. Do ponto de vista computacional, se $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ é uma base para V, então cada $\mathbf{v} \in V$ pode ser escrito na forma

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_n \mathbf{v}_n$$

Em geral, no entanto, determinar as coordenadas λ_i requer a resolução de um sistema de equações lineares de tamanho $n \times n$.

Por outro lado, se $O = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ é uma base ortonormal para V e

$$\mathbf{v} = \lambda_1 \mathbf{u}_1 + \cdots + \lambda_n \mathbf{u}_n$$

então os coeficientes são facilmente calculados:

$$\langle \mathbf{v}, \mathbf{u}_i \rangle = \langle \lambda_1 \mathbf{u}_1 + \cdots + \lambda_n \mathbf{u}_n, \mathbf{u}_i \rangle = \lambda_i \langle \mathbf{u}_i, \mathbf{u}_i \rangle = \lambda_i$$

Mesmo quando $O = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ não é uma base (mas apenas um conjunto ortonormal), ainda podemos considerar a expansão

$$\hat{\mathbf{v}} = \langle \mathbf{v}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \cdots + \langle \mathbf{v}, \mathbf{u}_n \rangle \mathbf{u}_n$$

Definição 6.37 — Coeficiente de Fourier. Seja V um espaço com produto interno. Seja $O = (\mathbf{u}_k)_{k \geq 1}$ uma sequência ortonormal de vetores de V. Para $\mathbf{v} \in V$, o k-ésimo **coeficiente de Fourier** de $\mathbf{v}$ em relação a O é definido como

$$f_k := \langle \mathbf{v}, \mathbf{u}_k \rangle$$

Exemplo 6.38. Seja $C[-1, 1]$ o espaço vetorial de todas as funções contínuas reais no intervalo fechado $[-1, 1]$ munido do produto interno

$$\langle f, g \rangle = \int_{-1}^{1} f(x) g(x) dx$$

Já vimos que conjunto $\underline{\mathrm{B}} = \{\mathrm{sen}(n\pi x)\}_{n=1}^{\infty}$ é ortogonal.

Os coeficientes de Fourier da função $f(x) = x$, para $x \in [-1, 1]$, com relação ao conjunto ortogonal $\underline{\mathrm{B}}$, são dados por:

$$f_k = \begin{cases} \frac{2}{k\pi} & \text{se } k \text{ ímpar} \\ -\frac{2}{k\pi} & \text{se } k \text{ par} \end{cases}$$

◁

Teorema 6.39 (Expansão de Fourier).
Seja V um espaço com produto interno. Seja $O = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ um conjunto ortonormal de vetores de V. Para qualquer $\mathbf{v} \in V$, a expansão de Fourier de $\mathbf{v}$ em relação a O é

$$\hat{\mathbf{v}} := \langle \mathbf{v}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \cdots + \langle \mathbf{v}, \mathbf{u}_k \rangle \mathbf{u}_k$$

Nesse caso, a desigualdade de Bessel vale para todos os $\mathbf{v} \in V$*, isto é,*

$$\|\hat{\mathbf{v}}\| \leq \|\mathbf{v}\|$$

Além disso, são equivalentes:

a *O conjunto* O *é uma base ortonormal para* V*.*

b *Todo vetor é igual à sua expansão de Fourier, ou seja, para todos os* $\mathbf{v} \in V$

$$\hat{\mathbf{v}} = \mathbf{v}$$

c *A identidade de Bessel vale para todos os* $\mathbf{v} \in V$*, isto é,*

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

d *A identidade de Parseval vale para todos os* $\mathbf{v}, \mathbf{w} \in V$*, ou seja*

$$\langle \mathbf{v},\ \mathbf{w}\rangle = \langle \mathbf{v},\ \mathbf{u}_1\rangle\overline{\langle \mathbf{w}, \mathbf{u}_1\rangle} + \cdots + \langle \mathbf{v},\ \mathbf{u}_k\rangle\overline{\langle \mathbf{w}, \mathbf{u}_k\rangle}$$

Demonstração. Ponha $\vec{r} = \mathbf{v} - \hat{\mathbf{v}}$. Para $1 \leq j \leq k$,

$$\langle \vec{r}, \mathbf{u}_j\rangle = \langle \mathbf{v}, \mathbf{u}_j\rangle - \sum_{i=1}^{k} \langle \mathbf{v}, \mathbf{u}_i\rangle\langle \mathbf{u}_i, \mathbf{u}_j\rangle = 0.$$

Logo $\vec{r}$ é ortogonal a $\langle O\rangle$ e, em particular, a $\hat{\mathbf{v}}$. Por Pitágoras,

$$\|\mathbf{v}\|^2 = \|\hat{\mathbf{v}}\|^2 + \|\vec{r}\|^2, \qquad \|\hat{\mathbf{v}}\|^2 = \sum_{i=1}^{k} |\langle \mathbf{v}, \mathbf{u}_i\rangle|^2.$$

Da primeira igualdade segue a desigualdade de Bessel.

Se O é uma base ortonormal, a fórmula das coordenadas em uma base ortonormal dá $\mathbf{v} = \hat{\mathbf{v}}$ para todo $\mathbf{v}$, provando a ⇒ b. Reciprocamente, se b vale, todo vetor pertence a $\langle O\rangle$; portanto, O é uma base, e temos b ⇒ a.

A implicação b ⇒ c é imediata. Se c vale, a identidade de Pitágoras acima fornece $\|\vec{r}\| = 0$, logo $\mathbf{v} = \hat{\mathbf{v}}$; assim, c ⇒ b.

Sob b, expandimos os dois vetores na base O e obtemos

$$\langle \mathbf{v}, \mathbf{w}\rangle = \sum_{i=1}^{k} \langle \mathbf{v}, \mathbf{u}_i\rangle\overline{\langle \mathbf{w}, \mathbf{u}_i\rangle},$$

que é d. Finalmente, tomando $\mathbf{w} = \mathbf{v}$ em d, obtemos $\|\mathbf{v}\|^2 = \sum_i |\langle \mathbf{v}, \mathbf{u}_i\rangle|^2 = \|\hat{\mathbf{v}}\|^2$, isto é, c. ∎

Vimos que, se S é um subespaço de um espaço com produto interno V, então $S \cap S^\perp = \{\vec{0}\}$. Isso levanta a questão de saber se o complemento ortogonal $S^\perp$ é um complemento de espaço vetorial de S, ou seja, se $V = S \oplus S^\perp$.

Se S é um subespaço de dimensão finita de V, a resposta é sim, mas para subespaços de dimensão infinita em geral, $V \neq S \oplus S^\perp$.

Exemplo 6.40. Seja $V = \ell^2$ e seja S o subespaço de todas as sequências de suporte finito, isto é, o subespaço gerado pelos vetores

$$\vec{e_i} = (0, \ \ldots, \ 0, 1, 0, \ \ldots)$$

onde $\vec{e_i}$ tem um 1 na i-ésima coordenada e zeros nos outros lugares. Se $\mathbf{x} = (x_n) \in S^\perp$, então

$x_i = \langle \mathbf{x}, \vec{e_i} \rangle = 0$ para todos os i e, portanto, $\mathbf{x} = \vec{0}$. Portanto, $S^\perp = \{\vec{0}\}$. No entanto,

$$S \oplus S^\perp = S \neq \ell^2.$$

◁

6.5 Projeção

Duas noções de projeção
Toda decomposição em soma direta produz uma projeção linear. A projeção ortogonal é especial: além de ser linear e idempotente, escolhe o ponto do subespaço mais próximo do vetor original.

Como mostra o próximo teorema, no caso de dimensão finita, os complementos ortogonais também são complementos de espaço vetorial.

> **Teorema 6.41** (Projeção).
> *Se S for um subespaço finito-dimensional de um espaço com produto interno V (que não precisa ser finito-dimensional), então*
>
> $$V = S \oplus S^\perp$$

Demonstração. Seja $O = \{\mathbf{u}_1, \ \ldots, \ \mathbf{u}_k\}$ uma base ortonormal de S. Para cada $\mathbf{v} \in V$, considere a expansão de Fourier

$$\hat{\mathbf{v}} = \langle \mathbf{v}, \ \mathbf{u}_1 \rangle \mathbf{u}_1 + \cdots + \langle \mathbf{v}, \ \mathbf{u}_k \rangle \mathbf{u}_k$$

em relação a O. Podemos escrever

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

onde $\hat{\mathbf{v}} \in S$. Além disso, $\mathbf{v} - \hat{\mathbf{v}} \in S^\perp$, já que

$$\langle \mathbf{v} - \hat{\mathbf{v}}, \ \mathbf{u}_i \rangle = \langle \mathbf{v}, \ \mathbf{u}_i \rangle - \langle \hat{\mathbf{v}}, \ \mathbf{u}_i \rangle = 0.$$

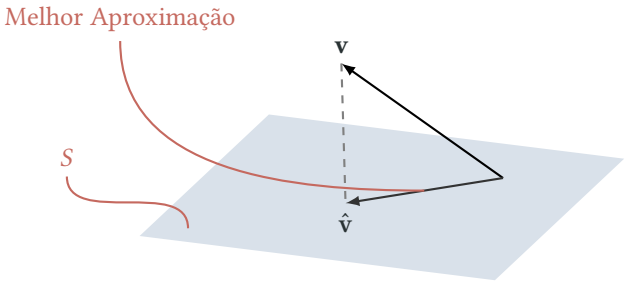
Como os $\mathbf{u}_i$ geram S, a linearidade mostra que $\mathbf{v} - \hat{\mathbf{v}}$ é ortogonal a todo vetor de S. Portanto, $V = S + S^\perp$. Já observamos que $S \cap S^\perp = \{\vec{0}\}$ e, portanto, $V = S \oplus S^\perp$. ■

De acordo com a prova do teorema da projeção, o componente de $\mathbf{v}$ que está em S é a expansão de Fourier de $\mathbf{v}$ em relação a qualquer base ortonormal O de S.

O teorema da projeção implica que, se $\mathbf{v} = \hat{\mathbf{v}} + \vec{s}^\perp$ em que $\hat{\mathbf{v}} \in S$ e $\vec{s}^\perp \in S^\perp$ então $\hat{\mathbf{v}}$ é o elemento de S mais próximo de $\mathbf{v}$, ou seja, $\hat{\mathbf{v}}$ é a melhor aproximação para $\mathbf{v}$ em S. Pois se $\vec{t} \in S$, como $\mathbf{v} - \hat{\mathbf{v}} \in S^\perp$, temos $(\mathbf{v} - \hat{\mathbf{v}}) \perp (\hat{\mathbf{v}} - \vec{t})$ e assim

$$\|\mathbf{v} - \vec{t}\|^2 = \|\vec{v} - \hat{\mathbf{v}} + \hat{\mathbf{v}} - \vec{t}\|^2 = \|\mathbf{v} - \hat{\mathbf{v}}\|^2 + \|\hat{\mathbf{v}} - \vec{t}\|^2.$$

Desse fato resulta que $\|\vec{v} - \vec{t}\|$ é mínimo, e apenas é mínimo, quando $\vec{t} = \hat{\mathbf{v}}$. Além disso, observamos que $\hat{\mathbf{v}}$ é o único vetor em S para o qual $\mathbf{v} - \hat{\mathbf{v}} \perp S$. Assim, podemos dizer que a melhor aproximação de $\mathbf{v}$ a partir de S é o único vetor $\vec{s} \in S$ para o qual $(\mathbf{v} - \vec{s}) \perp S$ e que esse vetor é a expansão de Fourier $\hat{\mathbf{v}}$ de $\mathbf{v}$.



Definição 6.42 — Soma Direta Ortogonal. Seja V um espaço com produto interno e sejam $S_1, \dots, S_n$ subespaços de V. Então V é dito **soma direta ortogonal** de $S_1, \dots, S_n$, escrita

$$V = S_1 \odot \cdots \odot S_n$$

se

a $V = S_1 \oplus \cdots \oplus S_n$

b $S_i \perp S_j$ por $i \neq j$

O teorema 6.41 afirma que $V = S \odot S^\perp$, para qualquer subespaço de dimensão finita S de um espaço vetorial V. O seguinte resultado simples é muito útil.

Teorema 6.43.
Seja V um espaço com produto interno. Então são equivalentes

a $V = S \odot T$

b $V = S \oplus T$ *e* $T = S^\perp$

c $V = S \oplus T$ *e* $T \subseteq S^\perp$

Demonstração. Suponha que a seja válido. Então $V = S \oplus T$ e $S \perp T$, o que implica que $T \subseteq S^\perp$. Mas se $\vec{w} \in S^\perp$ então $\vec{w} = \vec{s} + \vec{t}$ para $\vec{s} \in S, \vec{t} \in T$ e assim

$$0 = \langle \vec{s},\ \vec{w} \rangle = \langle \vec{s},\ \vec{s} \rangle + \langle \vec{s},\ \vec{t} \rangle = \langle \vec{s},\ \vec{s} \rangle.$$

Portanto, $\vec{s} = \vec{0}$ e $\vec{w} \in T$, o que implica que $S^\perp \subseteq T$. Portanto, $S^\perp = T$, o que fornece b. Obviamente, b implica c. Finalmente, se c vale, então $T \subseteq S^\perp$, o que implica que $S \perp T$ e, portanto, a vale. ■

Teorema 6.44.
Seja V um espaço com produto interno.

a *Se* $\dim V < \infty$ *e S for um subespaço de V,*

$$\dim(S^\perp) = \dim V - \dim(S)$$

b *Se S é um subespaço de dimensão finita de V, então*

$$S^{\perp\perp} = S$$

c *Se X for um subconjunto de V e* $\dim(\langle X \rangle) < \infty$, *então*

$$X^{\perp\perp} = \langle X \rangle$$

Demonstração. Como $V = S \oplus S^\perp$, temos $\dim V = \dim(S) + \dim(S^\perp)$, o que prova a parte a. Quanto à parte b, está claro que $S \subseteq S^{\perp\perp}$. Por outro lado, se $\mathbf{v} \in S^{\perp\perp}$ então pelo teorema da projeção

$$\mathbf{v} = \vec{s} + \vec{s}'$$

onde $\vec{s} \in S$ e $\vec{s}' \in S^\perp$. Mas $\mathbf{v} \in S^{\perp\perp}$ implica que $0 = \langle \mathbf{v}, \vec{s}' \rangle = \langle \vec{s}', \vec{s}' \rangle$ e assim $\vec{s}' = \vec{0}$, mostrando que $\mathbf{v} \in S$. Portanto, $S^{\perp\perp} \subseteq S$ e $S^{\perp\perp} = S$.

Para c, observe que $X^\perp = \langle X \rangle^\perp$: ser ortogonal a cada elemento de X equivale, por linearidade, a ser ortogonal a toda combinação linear desses elementos. Aplicando b ao subespaço finito-dimensional $\langle X \rangle$, obtemos

$$X^{\perp\perp} = (\langle X \rangle^\perp)^\perp = \langle X \rangle.$$

■

6.6 Adjuntas

Transportar um operador
A adjunta T^* é caracterizada por $\langle Tv, w \rangle = \langle v, T^* w \rangle$. Ela transporta a ação de T de uma entrada do produto interno para a outra; em espaços de funções, integração por partes é o modelo básico dessa operação.

Nesta seção, o símbolo T^* denota a *adjunta* determinada pelos produtos internos. No capítulo anterior, o mesmo símbolo foi usado para a transposta algébrica entre espaços duais; em bases ortonormais, as duas construções correspondem pela representação de Riesz.

Em nosso primeiro resultado, mostramos que em dimensão finita existe uma correspondência canônica entre os vetores do espaço dual V^* e os vetores de V.

> **Teorema 6.45** (de Representação de Riesz, dimensão finita).
> *Todo funcional linear $f : V \to \mathbb{K}$ num espaço com produto interno de dimensão finita pode ser escrito como um produto interno. Mais precisamente, existe um único $\mathbf{y} \in V$ tal que*
>
> $$f(\mathbf{x}) = \langle \mathbf{x},\ \mathbf{y} \rangle \qquad \text{para todo } \mathbf{x} \in V.$$

Demonstração. Considere uma base ortonormal $\underline{\mathrm{x}} = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ de V. Como $\mathbf{x} = \langle \mathbf{x}, \mathbf{x}_1 \rangle \mathbf{x}_1 + \ldots + \langle \mathbf{x}, \mathbf{x}_n \rangle \mathbf{x}_n$, temos

$$\begin{aligned} f(\mathbf{x}) &= \langle \mathbf{x},\ \mathbf{x}_1 \rangle f(\mathbf{x}_1) + \ldots + \langle \mathbf{x},\ \mathbf{x}_n \rangle f(\mathbf{x}_n) \\ &= \langle \mathbf{x},\ \overline{f(\mathbf{x}_1)} \mathbf{x}_1 \rangle + \ldots + \langle \mathbf{x}, \overline{f(\mathbf{x}_n)} \mathbf{x}_n \rangle \\ &= \langle \mathbf{x},\ \overline{f(\mathbf{x}_1)} \mathbf{x}_1 + \ldots + \overline{f(\mathbf{x}_n)} \mathbf{x}_n \rangle. \end{aligned}$$

Defina $\mathbf{y} = \overline{f(\mathbf{x}_1)} \mathbf{x}_1 + \ldots + \overline{f(\mathbf{x}_n)} \mathbf{x}_n$. Se também $f(\mathbf{x}) = \langle \mathbf{x}, \mathbf{z} \rangle$ para todo $\mathbf{x} \in V$, então $\langle \mathbf{x}, \mathbf{y} - \mathbf{z} \rangle = 0$ para todo $\mathbf{x} \in V$; tomando $\mathbf{x} = \mathbf{y} - \mathbf{z}$, obtemos $\mathbf{y} = \mathbf{z}$.

■

Para $f \in V^*$, seja f^* o vetor $\mathbf{y}$ em V de modo que $f(\mathbf{x}) = \langle \mathbf{x}, \mathbf{y} \rangle$. A função $f \to f^*$ de V^* a V é aditiva. Se o corpo for os reais, o mapa $f \to f^*$ é linear. No caso, em que o corpo é os complexos o mapa não será linear, mas linear conjugada pois $(\lambda f)^* = \overline{\lambda} f^*$.

Corolário 6.46.
Se V é um espaço com produto interno real, a aplicação $f \mapsto f^*$ é um isomorfismo entre V^* e V.

Teorema 6.47.
Sejam V, W espaços com produto interno de dimensão finita. Dada uma aplicação linear $T : V \to W$, existe uma única aplicação linear $T^* : W \to V$, denominada adjunta de T, satisfazendo

$$\langle T\mathbf{v},\ \mathbf{w}\rangle = \langle \mathbf{v},\ T^*\mathbf{w}\rangle \ \textit{para todo}\ \mathbf{v} \in V,\ \mathbf{w} \in W.$$

Demonstração. Para $\mathbf{w} \in W$ fixo, a aplicação $\mathbf{v} \mapsto \langle T\mathbf{v}, \mathbf{w}\rangle$ pertence ao dual de V. O Teorema de Representação de Riesz garante então que existe um único $\mathbf{y} \in V$ tal que

$$\langle T\mathbf{v},\ \mathbf{w}\rangle = \langle \mathbf{v},\ \mathbf{y}\rangle.$$

Definimos $T^*\mathbf{w} = \mathbf{y}$. Temos assim uma aplicação $T^* : W \to V$.

Para demonstrar a linearidade de T^* sejam $\mathbf{u}, \mathbf{w} \in W$ e $\lambda \in \mathbb{K}$. Então

$$\langle \mathbf{v},\ T^*(\mathbf{u} + \lambda\mathbf{w})\rangle = \langle T\mathbf{v},\ \mathbf{u} + \lambda\mathbf{w}\rangle = \langle T\mathbf{v},\ \mathbf{u}\rangle + \overline{\lambda}\langle T\mathbf{v},\ \mathbf{w}\rangle = \langle \mathbf{v},\ T^*\mathbf{u} + \lambda T^*\mathbf{w}\rangle.$$

Se $S^* : W \to V$ fosse outra aplicação linear tal que $\langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, S^*\mathbf{w}\rangle$, então $\langle \mathbf{v}, T^*\mathbf{w} - S^*\mathbf{w}\rangle = 0$ para todo $\mathbf{v} \in V$. Escolhendo $\mathbf{v} = T^*\mathbf{w} - S^*\mathbf{w}$, concluímos que $T^* = S^*$ ■

Definição 6.48 — Adjunta. Sejam V, W espaços com produto interno. A aplicação linear $T^* : W \to V$ que satisfaz

$$\langle T\mathbf{v},\ \mathbf{w}\rangle = \langle \mathbf{v},\ T^*\mathbf{w}\rangle \text{ para todos } \mathbf{v} \in V,\ \mathbf{w} \in W$$

é denominada **adjunta** de T.

O resultado a seguir enumera algumas propriedades do mapa $T \mapsto T^*$ de $\mathrm{Hom}(V, W)$ em $\mathrm{Hom}(W, V)$.

Teorema 6.49.
Sejam V, W e Z espaços de produto interno de dimensão finita. Então:

1. *Se $S, T \in \mathrm{Hom}(V,\ W)$ então $(S + T)^* = S^* + T^*$;*
2. *Se $T \in \mathrm{Hom}(V,\ W)$ e $\lambda \in \mathbb{K}$, então $(\lambda T)^* = \overline{\lambda}T^*$.;*
3. *Se $S \in \mathrm{Hom}(V,\ W)$ e $T \in \mathrm{Hom}(W,\ Z)$, então $(TS)^* = S^*T^*$;*
4. *Se $T \in \mathrm{Hom}(V,\ W)$ então $(T^*)^* = T$; e*
5. *$I_V^* = I_V$.*

Demonstração. Munimos V, W e Z dos produtos internos indicados no enunciado.

1 Seja $\mathbf{v} \in V, \mathbf{w} \in W$. Então

$$\begin{aligned}\langle \mathbf{v},\ (S+T)^*(\mathbf{w})\rangle_V &= \langle (S+T)(\mathbf{v}), \mathbf{w}\rangle_W \\ &= \langle S(\mathbf{v}) + T(\mathbf{v}), \mathbf{w}\rangle_W \\ &= \langle S(\mathbf{v}), \mathbf{w}\rangle_W + \langle T(\mathbf{v}), \mathbf{w}\rangle_W \\ &= \langle \mathbf{v}, S^*(\mathbf{w})\rangle_V + \langle \mathbf{v}, T^*(\mathbf{w})\rangle_V \\ &= \langle \mathbf{v}, S^*(\mathbf{w}) + T^*(\mathbf{w})\rangle_V \\ &= \langle \mathbf{v},\ (S^* + T^*)(\mathbf{w})\rangle_V.\end{aligned}$$

Consequentemente, $(S+T)^*(\mathbf{w}) = S^*(\mathbf{w}) + T^*(\mathbf{w})$ para todos os $\mathbf{w} \in W$ e, portanto, $(S+T)^* = S^* + T^*$.

2 Sejam $\mathbf{v} \in V, \mathbf{w} \in W$ e λ um escalar. Então

$$\begin{aligned}\langle \mathbf{v},\ (\lambda T)^*(\mathbf{w})\rangle_V &= \langle (\lambda T)(\mathbf{v}), \mathbf{w}\rangle_W \\ &= \langle \lambda T(\mathbf{v}), \mathbf{w}\rangle_W \\ &= \lambda \langle T(\mathbf{v}), \mathbf{w}\rangle_W \\ &= \lambda \langle \mathbf{v}, T^*(\mathbf{w})\rangle_V \\ &= \langle \mathbf{v}, \overline{\lambda} T^*(\mathbf{w})\rangle_V.\end{aligned}$$

Podemos então concluir que $(\lambda T)^* = \overline{\lambda} T^*$.

3 Se $S : V \to W, T : W \to Z, \mathbf{v} \in V$ e $\mathbf{z} \in Z$, então

$$\langle (T \circ S)\mathbf{v}, \mathbf{z}\rangle_Z = \langle S\mathbf{v}, T^*\mathbf{z}\rangle_W = \langle \mathbf{v}, S^*T^*\mathbf{z}\rangle_V.$$

Pela unicidade da adjunta, $(T \circ S)^* = S^* \circ T^*$.

4 Para $\mathbf{v} \in V$ e $\mathbf{w} \in W$, temos

$$\begin{aligned}\langle (T^*)^*\mathbf{v}, \mathbf{w}\rangle_W &= \overline{\langle \mathbf{w}, (T^*)^*\mathbf{v}\rangle_W} \\ &= \overline{\langle T^*\mathbf{w}, \mathbf{v}\rangle_V} \\ &= \langle \mathbf{v}, T^*\mathbf{w}\rangle_V = \langle T\mathbf{v}, \mathbf{w}\rangle_W.\end{aligned}$$

Como isso vale para todo $\mathbf{w}$, segue que $(T^*)^*\mathbf{v} = T\mathbf{v}$ para todo $\mathbf{v}$.

5 Por fim, $\langle I_V\mathbf{v}, \mathbf{w}\rangle_V = \langle \mathbf{v}, \mathbf{w}\rangle_V = \langle \mathbf{v}, I_V\mathbf{w}\rangle_V$; pela unicidade da adjunta, $I_V^* = I_V$. ■

Em seguida, apresentaremos algumas relações entre a imagem e o núcleo de $T \in \mathrm{Hom}(V,\ W)$ e de $T^* \in \mathrm{Hom}(W,\ V)$.

Teorema 6.50.
Sejam V, W espaço com produto interno com dimensões finitas e $T \in \mathrm{Hom}(V,\ W)$. Então

1 $\ker(T^*) = \mathrm{im}(T)^\perp$;

2 $\mathrm{im}(T^*) = \ker(T)^\perp$;

3 $\ker(T) = \mathrm{im}(T^*)^\perp$; *e*

4 $\mathrm{im}(T) = \ker(T^*)^\perp$.

Demonstração. Sejam $(V,\ \langle\cdot,\cdot\rangle_V)$, $(W,\ \langle\cdot,\cdot\rangle_W)$.

1 Suponha $\mathbf{w} \in \ker(T^*)$. Então $\langle \mathbf{v}, T^*(\mathbf{w})\rangle_V = \langle \mathbf{v}, \vec{0}\rangle_V = 0$ para todos os $\mathbf{v} \in V$. Pela definição de $T^*, \langle \mathbf{v}, T^*(\mathbf{w})\rangle_V = \langle T(\mathbf{v}), \mathbf{w}\rangle_W$. Isso implica que $\mathbf{w} \perp T(\mathbf{v})$ para todos os $\mathbf{v} \in V$ e, portanto, $\mathbf{w} \in \mathrm{im}(T)^\perp$. Portanto, $\ker(T^*) \subseteq \mathrm{im}(T)^\perp$.

Seja $\mathbf{w} \in \mathrm{im}(T)^\perp$. Então, para todos os $\mathbf{v} \in V, \langle T(\mathbf{v}), \mathbf{w}\rangle_W = 0$. Mas então, pela definição de $T^*, \langle \mathbf{v}, T^*(\mathbf{w})\rangle_V = 0$. Em particular, $\langle T^*(\mathbf{w}), T^*(\mathbf{w})\rangle_V = 0$, e logo , $T^*(\mathbf{w}) = \vec{0}$ e $\mathbf{w} \in \ker(T^*)$.

Como $(T^*)^* = T$, temos que 3 é consequência de 1.

De 1 também deduzimos que $\ker(T^*)^\perp = [\mathrm{im}(T)^\perp]^\perp = \mathrm{im}(T)$ e, consequentemente, 4 também é verdadeiro.

Finalmente, como $\ker(T) = \mathrm{im}(T^*)^\perp$, temos que $\ker(T)^\perp = [\mathrm{im}(T^*)^\perp]^\perp = \mathrm{im}(T^*)$, e assim 2 é válido. ■

O teorema seguinte relaciona as matrizes de T e T^* calculadas em bases ortonormais de V e W.

Teorema 6.51.
Sejam V e W espaços com produto interno com bases ortonormais $\underline{v} = (\mathbf{v}_1,\ \ldots,\ \mathbf{v}_n)$ e $\underline{w} = (\mathbf{w}_1,\ \ldots,\ \mathbf{w}_m)$ de V e W, respectivamente. Seja $A = [T]_{\underline{v},\underline{w}}$ e $B = [T^*]_{\underline{w},\underline{v}}$. Então $B = \overline{A}^{\mathrm{t}}$

Demonstração. Munimos V e W dos produtos internos indicados no enunciado. Então

$$T(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + \cdots + a_{mj}\mathbf{w}_m. \tag{6.9}$$

$$T^*(\mathbf{w}_i) = b_{1i}\mathbf{v}_1 + \cdots + b_{ni}\mathbf{v}_n. \tag{6.10}$$

Precisamos provar que $b_{ji} = \overline{a_{ij}}$ ou equivalente, que $a_{ij} = \overline{b_{ji}}$. Fazemos isso calculando $\langle T(\mathbf{v}_j), \mathbf{w}_i\rangle_W = \langle \mathbf{v}_j, T^*(\mathbf{w}_i)\rangle_V$ usando as Equações 6.9 e 6.10.

Por um lado

$$\langle T(\mathbf{v}_j),\ \mathbf{w}_i\rangle_W = \sum_{k=1}^{m} a_{kj}\langle \mathbf{w}_k,\ \mathbf{w}_i\rangle_W = a_{ij},$$

na última igualdade, utilizamos que $\underline{w}$ é uma base ortonormal de W. Por outro lado,

$$\langle \mathbf{v}_j,\ T^*(\mathbf{w}_i)\rangle_V = \sum_{l=1}^{n} \overline{b_{li}}\langle \mathbf{v}_j,\ \mathbf{v}_l\rangle_V = \overline{b_{ji}}.$$

Assim, $a_{ij} = \overline{b_{ji}}$. ■

6.7 Espaços de Hilbert

Leitura opcional. Esta seção pode ser omitida numa primeira leitura voltada a espaços de dimensão finita. Em dimensão infinita, a completeza requer uma hipótese adicional: uma sequência de Cauchy pode não ter limite no espaço considerado.

Completude e aproximação
Em um espaço completo, toda sequência cujos termos ficam arbitrariamente próximos entre si, a partir de certo índice, converge para um elemento do próprio espaço. A condição é formulada em termos das distâncias entre os termos, sem pressupor a existência de um limite.

Seja V um espaço com produto interno. A norma induz a métrica $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$. Dizemos que uma sequência $(\mathbf{v}^n)$ **converge** para $\mathbf{v} \in V$ quando

$$\lim_{n\to\infty} \|\mathbf{v}^n - \mathbf{v}\| = 0.$$

Definição 6.52 — Sequência de Cauchy. Seja (M, d) um espaço métrico. Uma sequência (x^n) em M é uma **sequência de Cauchy** se, para todo $\epsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que

$$n, m \geq n_0 \implies d(x^n, x^m) < \epsilon.$$

Definição 6.53 — Espaço Completo. Um espaço métrico (M, d) é **completo** se toda sequência de Cauchy em M converge para um limite que ainda pertence a M.

Exemplo 6.54. O corpo $\mathbb{R}$ é completo em sua métrica usual. O mesmo vale para $\mathbb{C}$: se (z_n) é de Cauchy, então $(\operatorname{Re} z_n)$ e $(\operatorname{Im} z_n)$ são sequências de Cauchy em $\mathbb{R}$. Se seus limites são x e y, respectivamente, então $z_n \to x + \mathrm{i}y \in \mathbb{C}$.

◁

Definição 6.55 — Espaço de Hilbert. Um **espaço de Hilbert** é um espaço com produto interno que é completo para a métrica induzida por sua norma.

Todo espaço com produto interno de dimensão finita é completo; seus subespaços são fechados e suas transformações lineares são contínuas. Em dimensão infinita, nenhuma dessas conclusões é automática. O exemplo central é o espaço das sequências quadrado somáveis.

Teorema 6.56.
Os espaços $\mathbb{R}^n$, $\mathbb{C}^n$ e ℓ^2 são espaços de Hilbert.

Demonstração. A completeza de $\mathbb{R}^n$ e $\mathbb{C}^n$ segue da completeza do corpo coordenada a coordenada. Provaremos o caso de ℓ^2.

Seja $(\mathbf{x}^k)$ uma sequência de Cauchy em ℓ^2, com

$$\mathbf{x}^k = (x_n^k)_{n\geq 1}.$$

Para cada índice fixo n,

$$|x_n^k - x_n^l| \leq \|\mathbf{x}^k - \mathbf{x}^l\|,$$

logo $(x_n^k)_k$ é de Cauchy em $\mathbb{K}$. Denote seu limite por a_n e ponha $\vec{a} = (a_n)_{n\geq 1}$; resta provar que $\vec{a} \in \ell^2$ e que $\mathbf{x}^k \to \vec{a}$.

Dado $\epsilon > 0$, escolha k_0 tal que $\|\mathbf{x}^k - \mathbf{x}^l\| < \epsilon$ sempre que $k, l \geq k_0$. Para todo N e todo $k, l \geq k_0$,

$$\sum_{n=1}^{N} |x_n^k - x_n^l|^2 \leq \|\mathbf{x}^k - \mathbf{x}^l\|^2 < \epsilon^2.$$

Fazendo $l \to \infty$ nessa soma finita, obtemos

$$\sum_{n=1}^{N} |x_n^k - a_n|^2 \leq \epsilon^2.$$

Como isso vale para todo N, segue que

$$\sum_{n=1}^{\infty} |x_n^k - a_n|^2 \le \epsilon^2 \qquad (k \ge k_0).$$

Em particular, $\mathbf{x}^{k_0} - \vec{a} \in \ell^2$ e, portanto, $\vec{a} = \mathbf{x}^{k_0} - (\mathbf{x}^{k_0} - \vec{a}) \in \ell^2$. A mesma desigualdade mostra que $\|\mathbf{x}^k - \vec{a}\| \le \epsilon$ para $k \ge k_0$. Assim, $\mathbf{x}^k \to \vec{a}$, e ℓ^2 é completo. ■

O teorema de projeção finito-dimensional admite agora sua forma completa. A hipótese de que o subespaço seja fechado substitui a hipótese de dimensão finita.

Teorema 6.57 (Projeção sobre subespaços fechados).
Sejam V um espaço de Hilbert e $S \subseteq V$ um subespaço fechado. Então

$$V = S \oplus S^{\perp}.$$

Equivalentemente, para cada $\mathbf{x} \in V$ existe um único $\mathbf{s} \in S$ que minimiza a distância $\|\mathbf{x} - \mathbf{s}\|$, e $\mathbf{x} - \mathbf{s} \in S^{\perp}$.

Demonstração. Fixe $\mathbf{x} \in V$ e ponha

$$d = \inf_{\mathbf{s} \in S} \|\mathbf{x} - \mathbf{s}\|.$$

Escolha $(\mathbf{s}_n)$ em S de modo que $\|\mathbf{x} - \mathbf{s}_n\|^2 < d^2 + 1/n$. Pela regra do paralelogramo,

$$\begin{aligned} \|\mathbf{s}_n - \mathbf{s}_m\|^2 &= 2\|\mathbf{x} - \mathbf{s}_n\|^2 + 2\|\mathbf{x} - \mathbf{s}_m\|^2 \\ &\quad - 4\|\mathbf{x} - \tfrac{1}{2}(\mathbf{s}_n + \mathbf{s}_m)\|^2 \\ &\le \frac{2}{n} + \frac{2}{m}, \end{aligned}$$

pois $(\mathbf{s}_n + \mathbf{s}_m)/2 \in S$. Logo $(\mathbf{s}_n)$ é de Cauchy. A completeza de V e o fato de S ser fechado fornecem $\mathbf{s}_n \to \mathbf{s} \in S$, e então $\|\mathbf{x} - \mathbf{s}\| = d$.

Escreva $\vec{r} = \mathbf{x} - \mathbf{s}$. Se existisse $\mathbf{u} \in S$ com $\langle \vec{r}, \mathbf{u} \rangle \neq 0$, tomando $\lambda = \langle \vec{r}, \mathbf{u} \rangle / \|\mathbf{u}\|^2$ teríamos

$$\|\vec{r} - \lambda \mathbf{u}\|^2 = \|\vec{r}\|^2 - \frac{|\langle \vec{r}, \mathbf{u} \rangle|^2}{\|\mathbf{u}\|^2} < \|\vec{r}\|^2,$$

contradizendo a minimalidade de $\mathbf{s}$. Portanto, $\vec{r} \in S^{\perp}$. A decomposição é única porque $S \cap S^{\perp} = \{\vec{0}\}$. ■

Podemos agora caracterizar os sistemas ortonormais maximais: em um espaço de Hilbert, são exatamente aqueles cujo espaço gerado é denso.

Proposição 6.58.
Seja A um conjunto ortonormal em um espaço com produto interno V.

a *Se $\overline{\langle A \rangle} = V$, então A é ortonormal maximal.*

b *Se V é um espaço de Hilbert e A é ortonormal maximal, então $\overline{\langle A \rangle} = V$.*

Demonstração. Temos $A^\perp = (\langle A\rangle)^\perp = (\overline{\langle A\rangle})^\perp$; a segunda igualdade segue da continuidade do produto interno. Se $\overline{\langle A\rangle} = V$, então $A^\perp = \{\vec{0}\}$, e nenhum vetor unitário pode ser acrescentado a A preservando a ortonormalidade. Logo A é maximal.

Reciprocamente, suponha que V seja de Hilbert e que A seja maximal. O Teorema 6.57, aplicado ao subespaço fechado $\overline{\langle A\rangle}$, fornece

$$V = \overline{\langle A\rangle} \oplus (\overline{\langle A\rangle})^\perp.$$

A maximalidade implica $A^\perp = \{\vec{0}\}$; portanto, o segundo somando é nulo e $\overline{\langle A\rangle} = V$. ■

Definição 6.59 — Base de Hilbert. Uma **base de Hilbert** de V é um conjunto ortonormal B tal que

$$\overline{\langle B\rangle} = V.$$

Uma base de Hilbert não deve ser confundida com uma base vetorial, também chamada de **base de Hamel**. A primeira permite recuperar vetores por limites de somas; a segunda permite apenas combinações lineares finitas.

Exemplo 6.60. Em ℓ^2, seja $M = \{\vec{e}_1, \vec{e}_2, \ldots\}$, onde $\vec{e}_i$ tem valor 1 na i-ésima coordenada e zero nas demais. O conjunto M é ortonormal. Se $\mathbf{v} = (x_n)$ é ortogonal a todos os $\vec{e}_i$, então

$$x_i = \langle \mathbf{v}, \vec{e}_i\rangle = 0$$

para todo i; logo $\mathbf{v} = \vec{0}$. Assim, M é maximal e, pela proposição anterior, é uma base de Hilbert de ℓ^2.

Entretanto, $\langle M\rangle$ contém apenas as sequências de suporte finito e é um subespaço próprio de ℓ^2. Portanto, M não é uma base de Hamel.

◁

6.7.1 Espaços de Hilbert Separáveis e Séries de Fourier

Definição 6.61 — Separável. Um espaço de Hilbert é **separável** se possui um subconjunto denso enumerável. Equivalentemente, ele admite uma base de Hilbert finita ou enumerável.

Proposição 6.62.
Seja V um espaço de Hilbert separável. Se V tem dimensão infinita, então V é isometricamente isomorfo a ℓ^2.

A demonstração será dada como consequência do teorema a seguir.

Para formular a expansão de Fourier como uma expansão em base, precisamos de uma noção que admita séries convergentes.

Definição 6.63 — Base de Schauder. Uma sequência $(\mathbf{v}_n)$ em um espaço normado V é uma **base de Schauder** se, para cada $\mathbf{v} \in V$, existe uma única sequência de escalares (a_n) tal que

$$\mathbf{v} = \sum_{n=1}^{\infty} a_n \mathbf{v}_n,$$

isto é,

$$\lim_{N\to\infty} \Big\| \mathbf{v} - \sum_{n=1}^{N} a_n \mathbf{v}_n \Big\| = 0.$$

Teorema 6.64 (Bases de Hilbert e séries de Fourier).
Seja V um espaço de Hilbert separável e seja $O = \{e_n\}_{n\geq 1}$ um sistema ortonormal. As seguintes afirmações são equivalentes:

1. *O é completo, isto é, $\overline{\langle O\rangle} = V$.*
2. *Se $x \perp e_n$ para todo $n \geq 1$, então $x = 0$.*
3. *Todo $x \in V$ possui a expansão de Fourier*

$$x = \sum_{n=1}^{\infty} \langle x, e_n\rangle e_n;$$

em particular, (e_n) é uma base de Schauder de V.

4. *Vale a identidade de Parseval*

$$\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, e_n\rangle|^2$$

para todo $x \in V$.

Demonstração. A equivalência entre 1 e 2 segue de $(\langle O\rangle)^{\perp} = \{0\}$ se, e somente se, $\overline{\langle O\rangle} = V$.

Suponha 2 e fixe $x \in V$. Escreva $a_n = \langle x, e_n\rangle$ e $y_N = \sum_{n=1}^{N} a_n e_n$. Pela desigualdade de Bessel, as somas $\sum_{n=1}^{N} |a_n|^2$ são crescentes e limitadas por $\|x\|^2$; portanto, a série $\sum_n |a_n|^2$ converge. Se $M > N$, a ortonormalidade fornece

$$\|y_M - y_N\|^2 = \sum_{n=N+1}^{M} |a_n|^2.$$

Assim, (y_N) é de Cauchy e converge, pela completeza de V, para algum $y \in V$. Para cada j, a desigualdade de Cauchy–Schwarz permite passar ao limite e obter $\langle y, e_j\rangle = a_j = \langle x, e_j\rangle$. Logo $x - y$ é ortogonal a todos os e_j; por 2, $x = y$. Os coeficientes são únicos: se $x = \sum_n b_n e_n$, tomar o produto interno com e_j fornece $b_j = \langle x, e_j\rangle$. Isso prova 3.

Se 3 vale, então

$$\|x\|^2 = \lim_{N\to\infty} \|y_N\|^2 = \lim_{N\to\infty} \sum_{n=1}^{N} |\langle x, e_n\rangle|^2,$$

que é 4. Finalmente, se 4 vale e $x \perp e_n$ para todo n, então $\|x\|^2 = 0$; logo $x = 0$ e vale 2. ∎

Demonstração da Proposição 6.62. Como V é separável e tem dimensão infinita, ele possui uma base de Hilbert enumerável $\{e_n\}_{n\geq 1}$. Defina

$$U : V \longrightarrow \ell^2, \qquad Ux = (\langle x, e_n\rangle)_{n\geq 1}.$$

A identidade de Parseval mostra que U está bem definida, é linear e preserva a norma. Para $(a_n) \in \ell^2$, as somas parciais $\sum_{n=1}^{N} a_n e_n$ formam uma sequência de Cauchy em V e, pela completeza, convergem para algum $x \in V$. A continuidade do produto interno dá $\langle x, e_n\rangle = a_n$ para todo n, de modo que U é sobrejetiva. Portanto, U é um isomorfismo isométrico. ∎

Definição 6.65 — $L^2([a,b])$. O espaço $L^2([a,b])$ é formado pelas classes de equivalência de funções mensuráveis $f : [a,b] \to \mathbb{K}$ tais que $\int_a^b |f(x)|^2\,dx < \infty$, identificando funções iguais quase em toda parte. Seu produto interno é

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}\,dx.$$

Equivalentemente, $L^2([a,b])$ é o completamento de $C([a,b])$ para essa norma.

$L^2([a,b])$ é um espaço de Hilbert separável. Encerramos esta leitura opcional registrando, sem demonstração, alguns sistemas ortogonais completos clássicos.

Teorema 6.66 (Sistemas ortogonais clássicos).
Os seguintes sistemas são completos nos espaços indicados:

1. *Série de Fourier em $L^2[-\pi,\pi]$:*
$$\left\{\frac{1}{\sqrt{2\pi}}, \frac{\cos(nx)}{\sqrt{\pi}}, \frac{\operatorname{sen}(nx)}{\sqrt{\pi}} \;\middle|\; n = 1, 2, \ldots\right\}.$$

2. *Polinômios de Legendre em $L^2[-1,1]$. Se*
$$P_k(x) = \frac{1}{2^k k!}\frac{d^k}{dx^k}(x^2-1)^k, \qquad k = 0, 1, \ldots,$$
então $\left\{\sqrt{(2k+1)/2}\,P_k \mid k \geq 0\right\}$ é ortonormal e completo.

3. *Funções de Bessel em $L^2[0,1]$. Para $\nu > -1$, sejam $j_{\nu,k}$ os zeros positivos de J_ν. Então*
$$\left\{\frac{\sqrt{2x}}{|J_{\nu+1}(j_{\nu,k})|} J_\nu(j_{\nu,k}x) \;\middle|\; k = 1, 2, \ldots\right\}$$
é um sistema ortonormal completo.

4. *Polinômios de Hermite em $L^2(\mathbb{R}, e^{-x^2}dx)$,*
$$H_k(x) = (-1)^k \exp\left(x^2\right)\frac{d^k}{dx^k}\left(\exp\left(-x^2\right)\right), k = 0, 1, \ldots$$
normalizados pelos fatores $(2^k k!\sqrt{\pi})^{-1/2}$.

5. *Polinômios de Chebyshev do primeiro tipo em $L^2((-1,1), (1-x^2)^{-1/2}dx)$:*
$$T_k(x) = \cos(k \arccos x), \qquad k = 0, 1, \ldots.$$
O sistema ortonormal correspondente é $\{T_0/\sqrt{\pi}\} \cup \{\sqrt{2/\pi}\,T_k \mid k \geq 1\}$.

Exercícios

Ex. 6.1 — Seja $T : W \to V$ uma transformação linear *injetiva* entre $\mathbb{K}$-espaços vetoriais. Suponha que $\langle , \rangle$, é um produto interno em V. Mostre que $\langle w_1, w_2 \rangle = \langle T(w_1), T(w_2) \rangle$ para todos $w_1, w_2 \in W$ define um produto interno em W. [Em particular, se V é um espaço vetorial com produto interno $\langle , \rangle$, W um subespaço de V e $T : W \to V$ é a inclusão natural, teremos como consequência que $\langle , \rangle$ restrito aos elementos de W é um produto interno em W.]

Ex. 6.2 — Dizemos que uma matriz $A \in M_n(\mathbb{K})$ hermitiana (i.e. $A^* = A$) é *positiva definida* se para toda matriz coluna não nula $\mathrm{X} \in M_{n\times 1}(\mathbb{K})$ tem-se que $\mathrm{X}^* A \mathrm{X} > 0$.

1. Mostre que se V é um espaço vetorial sobre $\mathbb{K}$ com produto interno $\langle , \rangle$ e se $\underline{\mathrm{B}} = \{v_1, v_2, \dots, v_n\}$ é uma base de V então $A = (a_{ij}) \in M_n(\mathbb{K})$, onde $a_{ij} = \langle v_i, v_j \rangle$ para todo $1 \leq i, j \leq n$ é uma matriz positiva definida.

2. Mostre que se $A = (a_{ij}) \in M_n(\mathbb{K})$ é uma matriz positiva definida, $\underline{\mathrm{B}}$ é uma base de um espaço vetorial sobre $\mathbb{K}$ V então

$$\langle u, v \rangle := [u]_{\underline{\mathrm{B}}}^{\mathrm{t}} A \overline{[v]_{\underline{\mathrm{B}}}} \quad \text{para todo } u, v \in V$$

   define um produto interno em V.

Ex. 6.3 — Seja V um espaço com produto interno $\langle , \rangle$ de dimensão finita n. Mostre que um conjunto de vetores $\{v_1, v_2, \dots, v_n\}$ é linearmente independente se, e só se, a matriz $A = \left(\langle v_i, v_j \rangle\right) \in M_n(\mathbb{K})$ é não singular.

Ex. 6.4 — Seja V um $\mathbb{K}$-e.v. com produto interno.

1. Se $\mathbb{K} = \mathbb{R}$, mostre que $\langle u, v \rangle = 0 \Leftrightarrow \|u + v\|^2 = \|u\|^2 + \|v\|^2$

2. Mostre que isso não é verdade se $\mathbb{K} = \mathbb{C}$.

3. Se $\mathbb{K} = \mathbb{C}$, mostre que para $u, v \in V$ $\langle u, v \rangle = 0 \Leftrightarrow \|\alpha u + \beta v\|^2 = \|\alpha u\|^2 + \|\beta v\|^2$, para todos $\alpha, \beta \in \mathbb{C}$.

Ex. 6.5 — Seja V um $\mathbb{K}$-e.v. com produto interno. Sejam $u, v \in V$, prove:

1. **Lei do Paralelogramo**: $\|u + v\|^2 + \|u - v\|^2 = 2\left(\|u\|^2 + \|v\|^2\right)$.

2. **Desigualdade de Bessel**: seja $\{v_1, \dots, v_n\}$ um conjunto ortonormal de V, então $\sum_{k=1}^{n} |\langle v, v_k \rangle|^2 \leq \|v\|^2$.

3. **Identidade de Parseval**: seja $\{v_1, \dots, v_n\}$ uma base ortonormal de V. Mostre que $\langle v, u \rangle = \sum_{k=1}^{n} \langle v, v_k \rangle \langle v_k, u \rangle$.

Ex. 6.6 — Seja V um $\mathbb{C}$-e.v. com produto interno

1. Fixe vetores $w, u \in V$. Mostre que o operador $T : V \to V$ definido por $T(v) = \langle v, w\rangle\, u$, admite adjunta e a descreva explicitamente.

2. Suponha $\dim V < \infty$ e que $\langle T(v), v\rangle = 0$, para todo $v \in V$. Mostre que $T = 0$. Mostre que o resultado não é necessariamente verdadeiro se V é um $\mathbb{R}$-e.v.

Ex. 6.7 — Seja V um $\mathbb{R}$-espaço vetorial de dimensão finita com um produto interno. Seja $W \subseteq V$ um subespaço, mostre que cada classe de equivalência do quociente V/W possui exatamente um vetor ortogonal a W.

Ex. 6.8 — Determine o mínimo do conjunto : $\left\{\int_0^1 (t^2 - at - b)^2 \mathrm{d}t \,:\, a, b \in \mathbb{R}\right\}$.

Ex. 6.9 — Seja $V = C([0,1], \mathbb{R})$ o espaço das funções contínuas no intervalo $[0,1]$ com o produto interno $\langle f, g\rangle = \int_0^1 f(t)g(t)dt$.

1. Exiba uma base ortonormal do subespaço de V gerado pelos polinômios 1, t e t^2.

2. Encontre o polinômio de grau menor ou igual a 2 que melhor aproxima $f(t) = \cos t$ no intervalo $[0,1]$.

Ex. 6.10 — Seja V um $\mathbb{K}$-e.v. com produto interno. Suponha que U e W sejam subespaços de V. Mostre que:

1. $W \subseteq (W^\perp)^\perp$ e se $V = W \oplus W^\perp$ então $W = (W^\perp)^\perp$.

2. $(U + W)^\perp = U^\perp \cap W^\perp$.

3. $U^\perp + W^\perp \subseteq (U \cap W)^\perp$.

4. Se $V = U \oplus U^\perp = W \oplus W^\perp$, então $U^\perp + W^\perp = (U \cap W)^\perp$.

Ex. 6.11 — Considere $\mathbb{R}[x]$ com o produto interno $\langle p, q\rangle = \int_0^1 p(x)q(x)dx$ e seja W o subespaço constituído por todos os polinômios com termo constante igual a zero. Mostre que $W^\perp = \{0\}$ e então $W^{\perp\perp} = \mathbb{R}[x] \neq W$. Também $\mathbb{R}[x] \neq W \oplus W^\perp$.

Ex. 6.12 — Seja W um subespaço de dimensão finita de V. Existem, em geral, vários operadores projeções cuja imagem é W. Prove que se $E \in \mathrm{Hom}(V, V)$ é uma projeção cuja imagem é W e $\|E(v)\| \leq \|v\|$, para todo $v \in V$, então E é a projeção ortogonal em W, i.e. $E = E_W$.

Ex. 6.13 — Seja W um subespaço vetorial de dimensão finita de um espaço V com produto interno. Seja E_W a projeção ortogonal de V em W. Mostre que $\mathrm{I} - E_W$ é a projeção ortogonal de V em $W^\perp$, é um operador projeção e $\ker(\mathrm{I} - E_W) = W$. (Obs: $W^\perp$ não necessariamente terá dimensão finita).

Ex. 6.14 — Seja V um $\mathbb{K}$-e.v. com produto interno $\langle , \rangle$, $\alpha \in \mathbb{K}$ e sejam $T, S \in \mathrm{Hom}(V, V)$ que admitem adjuntos. Mostre que os seguintes operadores admitem adjuntos:

1. $T + S$ e $(T + S)^* = T^* + S^*$.
2. αT e $(\alpha T)^* = \overline{\alpha} T^*$.
3. $T \circ S$ e $(T \circ S)^* = S^* \circ T^*$.
4. T^* e $(T^*)^* = T$.

   Se $T, S \in \mathrm{Hom}(V, V)$ são operadores normais então é verdade que αT, $T + S$ e $T \circ S$ são normais?

Ex. 6.15 — Seja T um operador linear em V que possui um adjunto T^*. Mostre que:

1. $\ker T^* = (\mathrm{im}\, T)^\perp$ e que $\mathrm{im}\, T^* \subset (\ker T)^\perp$.
2. Se $\dim V < \infty$, então $\mathrm{im}\, T^* = (\ker T)^\perp$.
3. Se $\dim V < \infty$, T é injetiva se, e somente se, T^*é sobrejetiva e vice-versa.
4. Seja $V = C([0, 1], \mathbb{R})$, com o produto interno $\langle f, g \rangle = \displaystyle\int_0^1 f(t)g(t)dt$, e considere o operador

$$\begin{aligned} T : V &\to V \\ f &\mapsto T(f) : [0, 1] \mapsto \mathbb{R} \\ &\qquad t \mapsto t f(t) \end{aligned}$$

   Determine T^* e mostre que $\mathrm{im}\, T^* \neq (\ker T)^\perp$.

Ex. 6.16 — Seja $T \in L(V)$. Prove que se T é invertível, então T^* é invertível e $(T^*)^{-1} = (T^{-1})^*$.

Exercícios avançados: adjuntas e teoria espectral. Os exercícios seguintes pressupõem as definições de operador normal, autoadjunto e unitário e, quando indicado, resultados espectrais posteriores.

Ex. 6.17 — Suponha que V é de dimensão finita. Sejam $T : V \to V$ um operador linear e $\lambda \in \mathbb{K}$. Mostre que λ é um autovalor de T se, e somente se, $\overline{\lambda}$ é um autovalor de T^*.

Ex. 6.18 — Seja V um espaço com produto interno complexo de dimensão finita. Seja $T : V \to V$ um operador linear. Definimos

$$T_1 = \frac{1}{2}(T + T^*),\ T_2 = \frac{1}{2i}(T - T^*).$$

Mostre que

1. T_1 e T_2 são autoadjuntos.
2. $T = T_1 + iT_2$.
3. O que pode dizer sobre a unicidade dessa decomposição?
4. Escreva T^* em função de T_1 e T_2.
5. Encontre condições em T_1 e T_2 para que o operador T seja: i- autoadjunto; ii-unitário; iii- normal.

Ex. 6.19 — Seja E um operador linear em V tal que $E^2 = E$ e tal que E possui um adjunto E^*. Prove que E é autoadjunto se, e somente se é normal. Prove também que, neste caso, E é a projeção ortogonal em $W = \operatorname{im} E$.

Ex. 6.20 — Sejam V um espaço com produto interno sobre $\mathbb{C}$, de dimensão finita, e $T \in \operatorname{Hom}(V, V)$. Prove que T é autoadjunto se, e somente se $\langle T(v), v\rangle \in \mathbb{R}$, para todo $v \in V$. Conclua que os autovalores de um operador autoadjunto são reais.

Ex. 6.21 — Sejam $S, T \in \operatorname{Hom}(V, V)$. Suponha que S e T admitem adjunto. Mostre que:

1. Se S e T são autoadjuntos, então ST é autoadjunto se, e somente se, $ST = TS$.
2. T^*T é autoadjunto.
3. Se T é autoadjunto, então S^*TS é autoadjunto.

Ex. 6.22 — Suponha que $\dim V$ é finita e seja $T : V \to V$ um operador normal. Mostre que:

1. Se T é nilpotente então $T = 0$.
2. Se T é um operador projeção então $T^* = T$.
3. Se $T^3 = T^2$ então T é um operador projeção.

Ex. 6.23 — Sejam V um $\mathbb{C}$-e.v. com produto interno de dimensão finita e $T \in \operatorname{Hom}(V, V)$ um operador normal. Prove que:

1. T é autoadjunto se, e somente se, todo autovalor de T é um número real.
2. T é unitário se, e somente se, todo autovalor de T é um número complexo de

módulo igual a 1.

3. T é o operador nulo se, e somente se todos os autovalores de T são nulos.

Ex. 6.24 — Sejam T e S dois operadores autoadjuntos em um espaço vetorial V de dimensão finita. Prove que se T e S comutam, então existe uma base ortonormal de V que diagonaliza estes dois operadores simultaneamente.

Ex. 6.25 — Um operador linear $T \in \mathrm{Hom}(V, V)$ é *não negativo* se T é autoadjunto e $\langle T(v), v\rangle \geq 0$ para todo $v \in V$. Mostre que se V é um $\mathbb{C}$-e.v. de dimensão finita:

1. T é não negativo se, e somente se, T é normal e os autovalores de T são números reais não negativos.

2. Se T é não negativo então existe um único operador linear não-negativo $R \in \mathrm{Hom}(V, V)$ que é raiz quadrada de T, isto é, que satisfaz $R^2 = T$.

Ex. 6.26 — Seja $\mathbb{R}_2[x]$ com o produto interno dado por $\langle f, g\rangle = \int_0^1 f(x)g(x)\,\mathrm{dx}$. Considere o operador $T : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$ dado por $T(ax^2 + bx + c) = bx$.

1. Mostre que T não é autoadjunto.

2. A matriz de T na base canônica é

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

que é simétrica. Explique por que isso não contradiz o item 1..

Ex. 6.27 — Sejam V um $\mathbb{C}$-e.v. com produto interno de dimensão finita e $T \in \mathrm{Hom}(V, V)$ um operador normal. Mostre que:

1. $T = \lambda_1 P_1 + \ldots + \lambda_k P_k$ onde $\mathrm{I} = P_1 + \ldots + P_k$, $P_i P_j = 0$ se $i \neq j$ e $P_i^2 = P_i = P_i^*$.

2. Existe $g \in \mathbb{C}[x]$ tal que $T^* = g(T)$.

3. Todo subespaço de V T-invariante é também T^*-invariante.

Ex. 6.28 — Seja V um $\mathbb{C}$-espaço vetorial com produto de interno complexo de dimensão finita. Considere $T \in \mathrm{Hom}(V, V)$ um operador qualquer e $\lambda_1, \ldots, \lambda_n$ os autovalores de T (aparecendo tantas vezes quanto sua multiplicidade algébrica). Mostre que:

1. $\sum_{i=1}^{n} |\lambda_i|^2 \leq \mathrm{tr}(T^* \circ T)$

2. A igualdade vale se, e somente se, T é normal.

Ex. 6.29 — Seja

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix} \in M_3(\mathbb{R}).$$

Ache uma matriz ortogonal $P \in M_3(\mathbb{R})$ tal que $P^T AP$ é diagonal.

Ex. 6.30 — Considere um $\mathbb{C}$-espaço vetorial V com o produto interno e dimensão finita. Dizemos que um operador $T : V \to V$ é *anti-hermitiano* ou *anti-autoadjunto* se para quaisquer vetores $u, v \in V$ tem-se que $\langle T(u), v\rangle = -\langle u, T(v)\rangle$. Assuma que T é anti-hermitiano e mostre que:

1. Os autovalores de T são imaginários puros.

2. Se $v_1, v_2 \in V$ são autovetores associados a autovalores distintos $\lambda_1, \lambda_2 \in \mathbb{C}$ então eles são ortogonais.

3. Existe uma base ortonormal de V formada por autovetores de T.

Ex. 6.31 — Determine a forma polar da matriz

$$A = \begin{bmatrix} 1+i & -1 & 1 \\ 0 & i & i \\ 1 & 0 & 1+i \end{bmatrix} \in M_3(\mathbb{C}).$$

7
Álgebras

Uma álgebra é um espaço vetorial no qual também podemos multiplicar vetores. Essa estrutura reúne exemplos já estudados: polinômios, matrizes e operadores lineares. A multiplicação nem sempre é comutativa, como ocorre com matrizes e com a composição de operadores.

Os quatérnios, descobertos por William Rowan Hamilton em 1843, constituem outro exemplo. Eles formam um sistema numérico real de dimensão quatro cuja multiplicação é associativa, mas não comutativa. O conceito de álgebra permite estudar as propriedades comuns a esses sistemas, relacionando a multiplicação à estrutura de espaço vetorial.

Nesse capítulo

7.1 Definições e exemplos

Definição 7.1 — Álgebra associativa. Uma **álgebra associativa** sobre um corpo $\mathbb{K}$ é um espaço vetorial $\mathfrak{A}$ sobre $\mathbb{K}$, munido de uma multiplicação

$$\mathfrak{A} \times \mathfrak{A} \longrightarrow \mathfrak{A}, \qquad (a, b) \longmapsto ab,$$

que é bilinear e associativa. Assim, para todos $a, b, c \in \mathfrak{A}$ e $\alpha \in \mathbb{K}$,

$$a(b+c) = ab + ac, \qquad (a+b)c = ac + bc,$$
$$(\alpha a)b = \alpha(ab) = a(\alpha b), \qquad (ab)c = a(bc).$$

Multiplicação e bilinearidade
A bilinearidade garante que uma tabela de produtos entre vetores de uma base determine toda a multiplicação. Em dimensão finita, essa tabela descreve a multiplicação por meio de uma quantidade finita de coeficientes.

Neste capítulo, a palavra *álgebra* significará sempre álgebra associativa. Não exigiremos comutatividade nem a existência de identidade multiplicativa, a menos que isso seja dito explicitamente. Uma álgebra é **comutativa** quando $ab = ba$ para todos $a, b \in \mathfrak{A}$.

Definição 7.2 — Álgebra unital. Uma álgebra $\mathfrak{A}$ é **unital** se existe um elemento $1_{\mathfrak{A}} \in \mathfrak{A}$ tal que

$$1_{\mathfrak{A}} a = a 1_{\mathfrak{A}} = a$$

para todo $a \in \mathfrak{A}$.

Quando não houver risco de confusão, escreveremos apenas 1. A identidade é única:

se e e e' são identidades de $\mathfrak{A}$, então $e = ee' = e'$.

Definição 7.3 — Elemento invertível. Se $\mathfrak{A}$ é unital, um elemento $a \in \mathfrak{A}$ é **invertível** se existe $b \in \mathfrak{A}$ tal que $ab = ba = 1$. Nesse caso, escrevemos $b = a^{-1}$.

O inverso também é único. Com efeito, se $ab = ba = 1$ e $ac = ca = 1$, então

$$b = b(ac) = (ba)c = c.$$

Exemplo 7.4 — Álgebra de polinômios. O espaço $\mathbb{K}[x]$, com as operações usuais, é uma álgebra comutativa unital. Se

$$f(x) = \sum_{i=0}^{r} \alpha_i x^i, \qquad g(x) = \sum_{j=0}^{s} \beta_j x^j,$$

então

$$(fg)(x) = \sum_{k=0}^{r+s} \left(\sum_{i+j=k} \alpha_i \beta_j \right) x^k.$$

A família $\{1, x, x^2, \ldots\}$ é uma base de $\mathbb{K}[x]$; portanto, essa álgebra tem dimensão infinita. Seus elementos invertíveis são precisamente os polinômios constantes não nulos.

◁

Exemplo 7.5 — Álgebra de matrizes. O espaço $\mathscr{M}_{n,n}(\mathbb{K})$, com a multiplicação de matrizes, é uma álgebra unital. Sua identidade é a matriz I_n. Para $n > 1$, ela não é comutativa. As unidades matriciais E_{ij}, com $1 \le i, j \le n$, formam uma base e satisfazem

$$E_{ij} E_{k\ell} = \delta_{jk} E_{i\ell}.$$

◁

Exemplo 7.6 — Álgebra de operadores. Para qualquer espaço vetorial V, o espaço $\operatorname{Hom}(V, V)$ é uma álgebra com o produto dado pela composição:

$$(ST)(v) = S(T(v)).$$

A identidade é o operador identidade I_V. Se $\dim V > 1$, em geral $ST \neq TS$. Os elementos invertíveis dessa álgebra são exatamente os isomorfismos de V em V.

Quando V tem dimensão finita e uma base é escolhida, o isomorfismo $\operatorname{Hom}(V, V) \cong \mathscr{M}_{n,n}(\mathbb{K})$ transforma composição de operadores em multiplicação de matrizes. Assim, a álgebra de matrizes é uma representação concreta da álgebra de operadores.

◁

Exemplo 7.7. Todo corpo L que contém $\mathbb{K}$ é uma álgebra comutativa unital sobre $\mathbb{K}$. Em particular, $\mathbb{C}$ é uma álgebra real de dimensão dois, com base $\{1, i\}$.

◁

Subálgebras e ideais
Uma subálgebra é fechada pela multiplicação de seus próprios elementos. Um ideal é, além disso, estável pela multiplicação, à esquerda e à direita, por qualquer elemento da álgebra. Essa propriedade permite definir a multiplicação na álgebra quociente independentemente dos representantes.

7.2 Subálgebras, ideais e centro

Definição 7.8 — Subálgebra. Uma **subálgebra** de $\mathfrak{A}$ é um subespaço vetorial $\mathfrak{B} \subseteq \mathfrak{A}$ fechado para a multiplicação. Se $\mathfrak{A}$ é unital, dizemos que $\mathfrak{B}$ é uma **subálgebra unital** quando $1_{\mathfrak{A}} \in \mathfrak{B}$.

Uma subálgebra pode possuir identidade própria sem ser uma subálgebra unital. Por exemplo,

$$\mathfrak{B} = \left\{ \begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 0 \end{bmatrix} : a, b, c, d \in \mathbb{K} \right\} \subseteq \mathscr{M}_{3,3}(\mathbb{K})$$

tem identidade $\operatorname{diag}(1, 1, 0)$, diferente de I_3.

Se $S \subseteq \mathfrak{A}$, a **subálgebra gerada por** S, denotada por $\operatorname{Alg}(S)$, é a interseção de todas as subálgebras que contêm S. Equivalentemente,

$$\operatorname{Alg}(S) = \operatorname{span}\{s_1 s_2 \cdots s_m \ : \ m \geq 1,\ s_i \in S\}.$$

Se $\mathfrak{A}$ é unital, a subálgebra unital gerada por S é $\operatorname{Alg}(S \cup \{1\})$. Em particular,

$$\operatorname{Alg}(1, a) = \{f(a) \ : \ f \in \mathbb{K}[x]\}.$$

Definição 7.9 — Ideal bilateral. Um **ideal bilateral** de uma álgebra $\mathfrak{A}$ é um subespaço vetorial $J \subseteq \mathfrak{A}$ tal que

$$ax \in J \quad \text{e} \quad xa \in J$$

para todos $a \in \mathfrak{A}$ e $x \in J$.

Neste capítulo, *ideal* significará sempre ideal bilateral. Se $\mathfrak{A}$ é unital, o ideal gerado por um subconjunto S é

$$\langle S \rangle = \left\{ \sum_{j=1}^{m} a_j s_j b_j \ : \ m \geq 1,\ a_j, b_j \in \mathfrak{A},\ s_j \in S \right\}.$$

De fato, o conjunto à direita é um ideal que contém S — basta tomar $a_j = b_j = 1$ — e está contido em todo ideal que contém S.

Definição 7.10 — Centro. O **centro** de $\mathfrak{A}$ é a subálgebra

$$Z(\mathfrak{A}) = \{a \in \mathfrak{A} \ : \ ab = ba \text{ para todo } b \in \mathfrak{A}\}.$$

Se $\mathfrak{A}$ é unital, então $\mathbb{K}1 \subseteq Z(\mathfrak{A})$. A álgebra é denominada **central** quando $Z(\mathfrak{A}) = \mathbb{K}1$.

Definição 7.11 — Álgebra simples. Uma álgebra $\mathfrak{A}$ é **simples** se $\mathfrak{A}^2 \neq \{0\}$ e seus únicos ideais são $\{0\}$ e $\mathfrak{A}$. Uma álgebra unital que é simples e central é chamada **álgebra simples central**.

Aqui, $\mathfrak{A}^2$ denota o subespaço gerado pelos produtos ab, com $a, b \in \mathfrak{A}$. A condição $\mathfrak{A}^2 \neq \{0\}$ exclui a álgebra não nula cuja multiplicação é identicamente zero.

Exemplo 7.12. A álgebra $\mathscr{M}_{n,n}(\mathbb{K})$ é simples. Para provar isso, seja $J \neq \{0\}$ um ideal e escolha $A = [a_{rs}] \in J$ com $a_{ij} \neq 0$. Para quaisquer k, ℓ,

$$E_{ki} A E_{j\ell} = a_{ij} E_{k\ell} \in J.$$

Logo, todas as unidades matriciais pertencem a J e, portanto, $J = \mathscr{M}_{n,n}(\mathbb{K})$.

◁

Se $\mathfrak{A}_1, \dots, \mathfrak{A}_r$ são álgebras sobre $\mathbb{K}$, o produto cartesiano $\mathfrak{A}_1 \times \cdots \times \mathfrak{A}_r$ é uma álgebra com operações componente a componente. Cada fator se identifica com um ideal do produto. Quando os fatores são unitais, a identidade é $(1_{\mathfrak{A}_1}, \dots, 1_{\mathfrak{A}_r})$.

7.3 Homomorfismos e álgebras quocientes

Preservar as duas estruturas
Um homomorfismo de álgebras deve preservar a estrutura linear e a multiplicação. Por isso, seu núcleo não é apenas um subespaço: é um ideal. A imagem, por sua vez, é uma subálgebra.

Definição 7.13 — Homomorfismo de álgebras. Se $\mathfrak{A}$ e $\mathfrak{B}$ são álgebras sobre $\mathbb{K}$, um mapa $\varphi : \mathfrak{A} \to \mathfrak{B}$ é um **homomorfismo de álgebras** se é linear e

$$\varphi(ab) = \varphi(a)\varphi(b)$$

para todos $a, b \in \mathfrak{A}$. Se as duas álgebras são unitais, o homomorfismo é **unital** quando $\varphi(1_{\mathfrak{A}}) = 1_{\mathfrak{B}}$.

Um homomorfismo bijetivo é um **isomorfismo**. Quando $\mathfrak{A} = \mathfrak{B}$, um isomorfismo é chamado **automorfismo**.

Proposição 7.14.
Se $\varphi : \mathfrak{A} \to \mathfrak{B}$ é um homomorfismo de álgebras, então $\ker \varphi$ é um ideal de $\mathfrak{A}$ e $\operatorname{im} \varphi$ é uma subálgebra de $\mathfrak{B}$.

Demonstração. Como φ é linear, seu núcleo é um subespaço de $\mathfrak{A}$ e sua imagem é um subespaço de $\mathfrak{B}$. Se $x \in \ker \varphi$ e $a \in \mathfrak{A}$, então

$$\varphi(ax) = \varphi(a)\varphi(x) = 0, \qquad \varphi(xa) = \varphi(x)\varphi(a) = 0.$$

Logo, $ax, xa \in \ker \varphi$. Se $y = \varphi(a)$ e $z = \varphi(b)$ pertencem à imagem, então $yz = \varphi(ab)$ também pertence à imagem. Isso prova as duas afirmações. ■

Exemplo 7.15 — Cálculo funcional polinomial. Se $\mathfrak{A}$ é unital e $a \in \mathfrak{A}$, a avaliação

$$\operatorname{ev}_a : \mathbb{K}[x] \longrightarrow \mathfrak{A}, \qquad \sum_{j=0}^{m} \alpha_j x^j \longmapsto \sum_{j=0}^{m} \alpha_j a^j$$

é um homomorfismo unital. A linearidade é imediata. Para monômios, $\operatorname{ev}_a(x^r x^s) = a^{r+s} = a^r a^s$; pela distributividade, a igualdade se estende a quaisquer polinômios. Sua imagem é precisamente $\operatorname{Alg}(1, a)$.

◁

Se J é um ideal de $\mathfrak{A}$, definimos em $\mathfrak{A}$ a relação

$$x \sim y \iff x - y \in J.$$

Ela é uma relação de equivalência, e a classe de x é $[x] = x + J$. O conjunto das classes será denotado por $\mathfrak{A}/J$.

Proposição 7.16 (Álgebra quociente).
As operações

$$[x] + [y] = [x + y], \qquad \alpha[x] = [\alpha x], \qquad [x][y] = [xy]$$

estão bem definidas e tornam $\mathfrak{A}/J$ uma álgebra sobre $\mathbb{K}$. Se $\mathfrak{A}$ é unital e $J \neq \mathfrak{A}$, então $\mathfrak{A}/J$ é unital, com identidade $[1]$.

Demonstração. As duas primeiras operações são as operações usuais do espaço vetorial quociente. Para verificar a multiplicação, suponha que $[x] = [x']$ e $[y] = [y']$. Então $x - x', y - y' \in J$ e

$$xy - x'y' = x(y - y') + (x - x')y' \in J,$$

pois J é um ideal bilateral. Logo, $[xy] = [x'y']$. A bilinearidade e a associatividade são herdadas de $\mathfrak{A}$. No caso unital, $[1][x] = [x][1] = [x]$. ∎

Teorema 7.17 (Primeiro Teorema do Isomorfismo).
Se $\varphi : \mathfrak{A} \to \mathfrak{B}$ é um homomorfismo de álgebras, então

$$\mathfrak{A}/\ker\varphi \cong \operatorname{im}\varphi.$$

Demonstração. Defina

$$\tilde{\varphi} : \mathfrak{A}/\ker\varphi \longrightarrow \operatorname{im}\varphi, \qquad \tilde{\varphi}([a]) = \varphi(a).$$

Se $[a] = [a']$, então $a - a' \in \ker\varphi$ e, portanto, $\varphi(a) = \varphi(a')$; logo, o mapa está bem definido. Além disso,

$$\tilde{\varphi}(\alpha[a] + \beta[b]) = \alpha\tilde{\varphi}([a]) + \beta\tilde{\varphi}([b]),$$
$$\tilde{\varphi}([a][b]) = \varphi(ab) = \varphi(a)\varphi(b),$$

de modo que $\tilde{\varphi}$ é um homomorfismo de álgebras. Ele é sobrejetivo pela definição de $\operatorname{im}\varphi$. Por fim,

$$\tilde{\varphi}([a]) = 0 \iff a \in \ker\varphi \iff [a] = [0],$$

o que prova a injetividade. ∎

Aplicando o teorema ao homomorfismo do Exemplo 7.15, obtemos

$$\mathbb{K}[x]/\ker(\mathrm{ev}_a) \cong \mathrm{Alg}(1, a).$$

Em dimensão finita, o núcleo é um ideal não nulo de $\mathbb{K}[x]$; seu gerador mônico é o polinômio minimal de a.

Uma tabela de multiplicação
Fixada uma base, as constantes de estrutura registram como multiplicar todos os pares de vetores básicos. Pela bilinearidade, esses coeficientes determinam toda a multiplicação da álgebra, assim como uma matriz determina uma transformação linear.

7.4 Constantes de estrutura

Se $\mathfrak{A}$ tem dimensão finita e $\mathscr{B} = \{z_1, \dots, z_n\}$ é uma base, existem escalares únicos $\gamma_{ij}^k \in \mathbb{K}$ tais que

$$z_i z_j = \sum_{k=1}^{n} \gamma_{ij}^k z_k.$$

Os n^3 escalares γ_{ij}^k são chamados **constantes de estrutura** de $\mathfrak{A}$ na base $\mathscr{B}$. Eles determinam toda a multiplicação, pois, se $a = \sum_i \alpha_i z_i$ e $b = \sum_j \beta_j z_j$, então

$$ab = \sum_{i,j,k=1}^{n} \alpha_i \beta_j \gamma_{ij}^k z_k.$$

Reciprocamente, uma escolha de constantes de estrutura define, por bilinearidade, uma multiplicação no espaço vetorial. A condição de associatividade pode ser verificada apenas nos elementos da base.

Proposição 7.18.
A multiplicação determinada por γ_{ij}^k é associativa se, e somente se,

$$\sum_{r=1}^{n} \gamma_{ij}^r \gamma_{r\ell}^k = \sum_{r=1}^{n} \gamma_{j\ell}^r \gamma_{ir}^k$$

para todos i, j, ℓ, k. Ela é comutativa se, e somente se, $\gamma_{ij}^k = \gamma_{ji}^k$ para todos i, j, k.

Demonstração. Pela definição da multiplicação,

$$(z_i z_j) z_\ell = \sum_{r=1}^{n} \gamma_{ij}^r z_r z_\ell = \sum_{k=1}^{n} \left(\sum_{r=1}^{n} \gamma_{ij}^r \gamma_{r\ell}^k \right) z_k,$$
$$z_i (z_j z_\ell) = \sum_{r=1}^{n} \gamma_{j\ell}^r z_i z_r = \sum_{k=1}^{n} \left(\sum_{r=1}^{n} \gamma_{j\ell}^r \gamma_{ir}^k \right) z_k.$$

Como $\mathscr{B}$ é uma base, essas expressões são iguais precisamente quando os coeficientes indicados coincidem. Pela trilinearidade, a associatividade nos elementos da base equivale à associatividade em toda a álgebra. De modo análogo, $z_i z_j = z_j z_i$ equivale às igualdades entre as constantes, e a bilinearidade estende a comutatividade a todos os elementos. ∎

Exemplo 7.19 — Quatérnios. Se $\mathbb{K}$ é um subcorpo de $\mathbb{R}$, o espaço vetorial

$$\mathbb{H}_{\mathbb{K}} = \mathbb{K}1 \oplus \mathbb{K}i \oplus \mathbb{K}j \oplus \mathbb{K}k$$

torna-se uma álgebra quando a multiplicação é determinada por

$$i^2 = j^2 = k^2 = -1, \qquad ij = k, \quad jk = i, \quad ki = j,$$

e pela troca de sinal ao inverter a ordem de dois entre i, j, k. Essa multiplicação é associativa e não comutativa.

Para $q = a_0 + a_1 i + a_2 j + a_3 k$, defina $\overline{q} = a_0 - a_1 i - a_2 j - a_3 k$. Um cálculo direto fornece

$$q\overline{q} = \overline{q}q = (a_0^2 + a_1^2 + a_2^2 + a_3^2)1.$$

Como $\mathbb{K} \subseteq \mathbb{R}$, essa soma é não nula quando $q \neq 0$. Portanto, todo quatérnio não nulo é invertível e

$$q^{-1} = \frac{\overline{q}}{a_0^2 + a_1^2 + a_2^2 + a_3^2}.$$

Para $\mathbb{K} = \mathbb{R}$, escrevemos simplesmente $\mathbb{H}$.

◁

7.5 Álgebras de grupo

Linearizar simetrias
A álgebra de grupo transforma elementos de um grupo em vetores de uma base e prolonga a multiplicação do grupo por linearidade. Com isso, problemas sobre simetrias podem ser estudados por métodos de álgebra linear.

Se G é um grupo finito, a **álgebra de grupo** $\mathbb{K}G$ é o espaço vetorial que possui G como base:

$$\mathbb{K}G = \left\{ \sum_{g \in G} \alpha_g g \,:\, \alpha_g \in \mathbb{K} \right\}.$$

A multiplicação é a extensão bilinear da multiplicação de G. Assim,

$$\left(\sum_{g \in G} \alpha_g g \right) \left(\sum_{h \in G} \beta_h h \right) = \sum_{u \in G} \left(\sum_{h \in G} \alpha_{uh^{-1}} \beta_h \right) u.$$

A associatividade vem da associatividade em G, e o elemento neutro de G é a identidade da álgebra. Além disso, $\mathbb{K}G$ é comutativa se, e somente se, G é abeliano.

Exemplo 7.20. Considere $C_4 = \{1, \omega, \omega^2, \omega^3\}$, com $\omega^4 = 1$, e sejam, em $\mathbb{Q}C_4$,

$$x = \omega + \omega^2, \qquad y = \omega + \frac{1}{2}\omega^2 + \frac{1}{3}\omega^3.$$

Então

$$\begin{aligned} xy &= \omega^2 + \frac{1}{2}\omega^3 + \frac{1}{3} + \omega^3 + \frac{1}{2} + \frac{1}{3}\omega \\ &= \frac{5}{6} + \frac{1}{3}\omega + \omega^2 + \frac{3}{2}\omega^3. \end{aligned}$$

◁

O comportamento de uma álgebra de grupo pode depender do corpo de escalares, como mostra o exemplo seguinte.

Exemplo 7.21. Seja $C_3 = \{1, \omega, \omega^2\}$ e defina

$$e = 1 + \omega + \omega^2, \qquad J = \mathbb{K}e \subseteq \mathbb{K}C_3.$$

Para $y = \alpha + \beta\omega + \gamma\omega^2$, temos

$$ey = ye = (\alpha + \beta + \gamma)e.$$

Logo, J é um ideal unidimensional. Como $e^2 = 3e$, se $\operatorname{char}\mathbb{K} \neq 3$, o elemento $e/3$ é a identidade de J. Em particular, sobre $\mathbb{Q}$, temos $J \cong \mathbb{Q}$ como álgebras.

Se $\operatorname{char}\mathbb{K} = 3$, então $e^2 = 0$ e o produto de quaisquer dois elementos de J é zero. O mesmo subespaço, portanto, possui comportamentos algébricos radicalmente diferentes conforme o corpo base.

◁

As álgebras de grupo relacionam a teoria de representações à álgebra linear: uma representação de G em um espaço vetorial V estende-se, por linearidade, a um homomorfismo de álgebras $\mathbb{K}G \to \mathrm{Hom}(V, V)$. Desse modo, uma ação de grupo pode ser estudada por meio de operadores lineares.

8
Determinantes

O determinante associa um escalar a cada matriz quadrada. Sobre os reais, seu módulo mede a variação de volume produzida pela transformação linear, e seu sinal registra a orientação. Essa interpretação motiva a definição algébrica. A construção pela fórmula de Leibniz utiliza permutações e seus sinais, que estudaremos antes de definir o determinante.

Nesse capítulo

8.1 Grupos, permutações e sinal

A fórmula explícita do determinante envolve todas as maneiras de escolher uma entrada em cada linha e em cada coluna de uma matriz. Essas escolhas são descritas por permutações. Antes de construir o determinante, desenvolveremos a linguagem mínima necessária para organizar essas permutações e determinar se cada uma contribui com sinal positivo ou negativo.

Permutações e determinantes
O grupo simétrico descreve as escolhas de entradas que aparecem na fórmula de Leibniz. Cada termo corresponde a uma permutação, cujo sinal registra a alteração de orientação.

8.1.1 Grupos

Definição 8.1 — Grupo. Um **grupo** é um conjunto não vazio G, munido de uma operação binária $(g, h) \mapsto gh$, que satisfaz:

1. $(gh)k = g(hk)$ para quaisquer $g, h, k \in G$;
2. existe $e \in G$ tal que $eg = ge = g$ para todo $g \in G$;
3. para cada $g \in G$, existe $h \in G$ tal que $gh = hg = e$.

O grupo é **abeliano** quando $gh = hg$ para quaisquer $g, h \in G$.

A identidade e o inverso são únicos, como vimos no capítulo inicial. Por isso, denotamos a identidade por e e o inverso de g por g^{-1}. Quando a operação é escrita aditivamente, usamos 0 para a identidade e $-g$ para o inverso.

Proposição 8.2 (Propriedades elementares).
Se G é um grupo, então:

1. *valem as leis de cancelamento;*
2. $(gh)^{-1} = h^{-1}g^{-1}$ *e* $(g^{-1})^{-1} = g$*;*
3. *as equações $gx = k$ e $yg = k$ possuem, respectivamente, as únicas soluções $x = g^{-1}k$ e $y = kg^{-1}$.*

Demonstração. Se $gh = gk$, multiplicar à esquerda por g^{-1} fornece $h = k$; o cancelamento à direita é análogo. Além disso,

$$(gh)(h^{-1}g^{-1}) = g(hh^{-1})g^{-1} = e$$

e, da mesma forma, $(h^{-1}g^{-1})(gh) = e$. Pela unicidade do inverso, $(gh)^{-1} = h^{-1}g^{-1}$. Como g é inverso de g^{-1}, também $(g^{-1})^{-1} = g$.

Por fim, $gx = k$ implica $x = g^{-1}k$ pelo cancelamento, e esse elemento de fato satisfaz a equação. O argumento para $yg = k$ é semelhante. ■

Exemplos 8.3.

1. $(\mathbb{Z}, +)$ e $(\mathbb{R}, +)$ são grupos abelianos.
2. $\mathbb{C} \setminus \{0\}$, com a multiplicação, é um grupo abeliano.
3. O conjunto $\mathrm{GL}_n(\mathbb{K})$ das matrizes invertíveis $n \times n$ é um grupo sob multiplicação. Para $n \geq 2$, ele em geral não é abeliano.
4. Se X é um conjunto, as bijeções $X \to X$ formam um grupo sob composição.

◁

O último exemplo representa um princípio importante: transformações invertíveis que preservam uma estrutura costumam formar um grupo. No caso do determinante, interessam as bijeções de um conjunto finito.

Exercícios

Ex. 8.1 — Seja G um grupo. Demonstre diretamente que

$$(g_1 g_2 \cdots g_r)^{-1} = g_r^{-1} \cdots g_2^{-1} g_1^{-1}.$$

Ex. 8.2 — Mostre que, se $xy = e$, então $y = x^{-1}$ e, portanto, $yx = e$.

Ex. 8.3 — Suponha que G tenha exatamente dois elementos. Mostre que $g^2 = e$ para todo $g \in G$.

Ex. 8.4 — Suponha que G tenha exatamente três elementos. Mostre, sem usar o Teorema de Lagrange, que $g^3 = e$ para todo $g \in G$.

Ex. 8.5 — Verifique que as matrizes invertíveis $n \times n$ formam um grupo sob multiplicação. Dê duas matrizes em $\mathrm{GL}_2(\mathbb{R})$ que não comutam.

Ex. 8.6 — Se $f : G \to H$ é um homomorfismo de grupos, prove que $f(e_G) = e_H$ e $f(g^{-1}) = f(g)^{-1}$ para todo $g \in G$.

8.1.2 Permutações e ciclos

Escreveremos $[n] = \{1, 2, \dots, n\}$.

Definição 8.4 — Grupo simétrico. O **grupo simétrico** S_n é o grupo das bijeções $\sigma : [n] \to [n]$, com a composição. Adotaremos a convenção

$$(\sigma\tau)(k) = \sigma\big(\tau(k)\big),$$

isto é, a permutação mais à direita atua primeiro.

Uma permutação pode ser apresentada por duas linhas:

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

Para compreender sua estrutura, porém, a notação por ciclos é mais informativa.

Definição 8.5 — Ciclo. Dados elementos distintos $i_1, \dots, i_k \in [n]$, o **k-ciclo** $(i_1\ i_2\ \cdots\ i_k)$ é a permutação que envia

$$i_1 \mapsto i_2 \mapsto \cdots \mapsto i_k \mapsto i_1$$

e fixa todos os demais elementos. O **suporte** de uma permutação σ é

$$\operatorname{supp}(\sigma) = \{i \in [n] \,:\, \sigma(i) \neq i\}.$$

Assim, ciclos de comprimento 1 representam pontos fixos e normalmente são omitidos. Ciclos cujos suportes são disjuntos comutam.

Exemplo 8.6. Se $\sigma = (1\ 2)$ e $\tau = (1\ 2\ 3)$, então

$$\sigma\tau = (2\ 3), \qquad \tau\sigma = (1\ 3).$$

O exemplo mostra simultaneamente que a composição é lida da direita para a esquerda e que S_n não é abeliano para $n \geq 3$.

◁

Teorema 8.7 (Decomposição em ciclos).
Toda permutação é produto de ciclos com suportes dois a dois disjuntos. Se os pontos fixos forem omitidos, essa decomposição é única a menos da ordem dos ciclos e de rotações na escrita de cada ciclo.

Demonstração. Fixe $i \in [n]$. Como $[n]$ é finito, a sequência

$$i, \sigma(i), \sigma^2(i), \dots$$

eventualmente repete um termo. A injetividade de σ implica que o primeiro termo repetido é i. Obtemos, assim, um ciclo formado pela órbita de i sob as potências de σ.

Escolhendo em seguida um elemento que ainda não pertence a nenhum ciclo e repetindo o procedimento, particionamos $[n]$ em órbitas disjuntas. Em cada órbita, a ação de σ determina um ciclo. As órbitas são determinadas unicamente por σ, o que prova a unicidade declarada. ∎

8.1.3 Transposições e sinal

Um ciclo de comprimento 2 é chamado de **transposição**. Todo ciclo pode ser escrito como produto de transposições:

$$(i_1\ i_2\ \cdots\ i_k) = (i_1\ i_k)(i_1\ i_{k-1}) \cdots (i_1\ i_2).$$

Pelo teorema anterior, toda permutação é, portanto, produto de transposições. Essa fatoração não é única, mas a paridade do número de fatores é invariável.

Teorema 8.8 (Paridade).
Se uma permutação $\sigma \in S_n$ admite duas decomposições em transposições,

$$\sigma = \tau_1 \cdots \tau_r = \rho_1 \cdots \rho_s,$$

então r e s possuem a mesma paridade.

Demonstração. Considere o polinômio de Vandermonde

$$\Delta(x_1, \dots, x_n) = \prod_{1 \le i < j \le n} (x_j - x_i).$$

Permutar as variáveis apenas reorganiza os fatores e possivelmente troca a ordem em alguns deles. Portanto, para cada $\sigma \in S_n$ existe $\varepsilon(\sigma) \in \{1, -1\}$ tal que

$$\Delta(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \varepsilon(\sigma)\Delta(x_1, \dots, x_n).$$

A substituição sucessiva das variáveis mostra que $\varepsilon(\sigma\rho) = \varepsilon(\sigma)\varepsilon(\rho)$. Uma transposição troca exatamente duas variáveis; os fatores que envolvem apenas uma delas são permutados aos pares, enquanto o fator que envolve ambas muda de sinal. Logo, $\varepsilon(\tau) = -1$ para toda transposição τ.

Aplicando essa identidade às duas fatorações de σ, obtemos

$$\varepsilon(\sigma) = (-1)^r = (-1)^s.$$

Consequentemente, r e s têm a mesma paridade. ∎

Definição 8.9 — Sinal de uma permutação. Se σ é produto de r transposições, definimos

$$\operatorname{sgn}(\sigma) = (-1)^r.$$

A permutação é **par** quando seu sinal é 1 e **ímpar** quando seu sinal é -1.

Proposição 8.10.
Para quaisquer $\sigma, \rho \in S_n$,

$$\operatorname{sgn}(e) = 1, \qquad \operatorname{sgn}(\sigma\rho) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\rho), \qquad \operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma).$$

Além disso, um k-ciclo tem sinal $(-1)^{k-1}$.

Demonstração. A identidade é produto de zero transposições. Concatenar decomposições de σ e ρ prova a fórmula do produto. Aplicando-a a $\sigma\sigma^{-1} = e$, obtemos $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)^{-1} = \operatorname{sgn}(\sigma)$. Finalmente, a decomposição exibida de um k-ciclo possui $k-1$ transposições. ∎

Um par (i, j), com $i < j$, é uma **inversão** de σ quando $\sigma(i) > \sigma(j)$. Se $N(\sigma)$ é o número de inversões, então

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

De fato, trocar duas entradas adjacentes altera o número de inversões em uma unidade. Ordenar a lista $\sigma(1), \dots, \sigma(n)$ por trocas adjacentes requer exatamente $N(\sigma)$ trocas; essas trocas são transposições e reduzem σ à identidade.

O sinal na fórmula de Leibniz
Uma permutação escolhe uma entrada de cada coluna sem repetir linhas. Seu sinal fornece o coeficiente $+1$ ou -1 do termo correspondente na fórmula de Leibniz.

Na seção dedicada à fórmula de Leibniz, essa linguagem será usada para obter

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, a_{\sigma(1),1} \cdots a_{\sigma(n),n}.$$

Exercícios

Ex. 8.7 — Mostre que ciclos com suportes disjuntos comutam.

Ex. 8.8 — Escreva

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 6 & 5 & 1 & 3 \end{pmatrix}$$

como produto de ciclos disjuntos.

Ex. 8.9 — Calcule os produtos abaixo, lembrando que o fator à direita atua primeiro, e escreva o resultado como produto de ciclos disjuntos.

1. $(1\ 2)(1\ 3)$ em S_3;
2. $(1\ 2)(1\ 3)$ em S_6;
3. $(1\ 2)(1\ 2\ 3\ 4\ 5)$;
4. $(1\ 2\ 3\ 4\ 5)(1\ 2)$;

5. $(1\ 2\ 3)^2$;

6. $(1\ 2\ 3\ 4)^2$.

Ex. 8.10 — Determine as inversas das permutações seguintes.

1. $(1\ 2)$;

2. $(1\ 2\ 3)$;

3. $(i_1\ i_2\ \cdots\ i_k)$.

Ex. 8.11 — Verifique as identidades

$$(1\ 2\ \cdots\ k) = (1\ k)(1\ k-1)\cdots(1\ 2)$$

e

$$(1\ k)(1\ 2\ \cdots\ k-1) = (1\ 2\ \cdots\ k).$$

Ex. 8.12 — A ordem de $\sigma \in S_n$ é o menor inteiro positivo m tal que $\sigma^m = e$.

1. Mostre que um k-ciclo tem ordem k.

2. Mostre que a ordem de um produto de ciclos disjuntos é o mínimo múltiplo comum de seus comprimentos.

Ex. 8.13 — Calcule o sinal de cada permutação do terceiro exercício de duas maneiras: por uma decomposição em transposições e pelo número de inversões.

Ex. 8.14 — Prove que $\operatorname{sgn}(\sigma) = 1$ se, e somente se, σ pode ser escrita como produto de um número par de transposições.

8.2 Geometria dos Volumes

Escala orientada
O determinante mede o fator pelo qual uma transformação linear altera volumes orientados. O módulo fornece a mudança de volume; o sinal registra se a orientação foi preservada ou invertida. Determinante zero significa colapso em dimensão menor.

Sobre $\mathbb{R}$, o determinante de uma matriz A tem uma interpretação geométrica simples: ele é o volume orientado da imagem do cubo unitário pela transformação linear T_A.

Começaremos descrevendo as propriedades esperadas desse volume orientado e, em seguida, as adotaremos como ponto de partida para a definição algébrica.

Doravante, por concisão, diremos simplesmente volume no lugar de volume orientado.

Para motivar essas propriedades, trabalhemos em $\mathbb{R}^2$, onde o volume se reduz à área. Seja $A = [\mathbf{v}_1 \mid \mathbf{v}_2]$. Como $T_A(\mathbf{e}_1) = \mathbf{v}_1$ e $T_A(\mathbf{e}_2) = \mathbf{v}_2$, a imagem do quadrado unitário é o paralelogramo P determinado pelas colunas de A.

A área orientada de um paralelogramo satisfaz as duas propriedades seguintes:

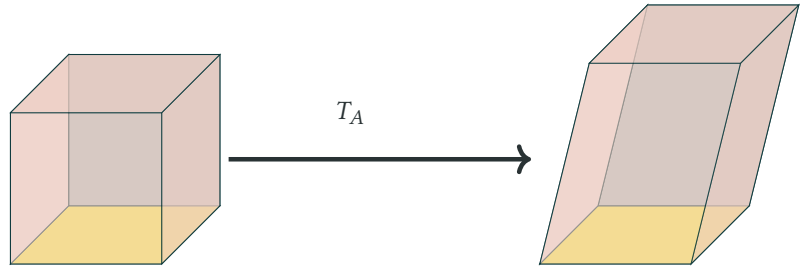
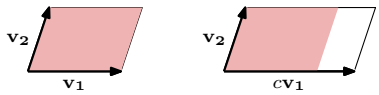
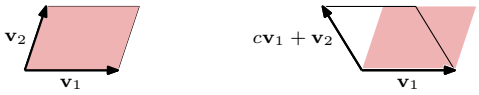


Figura 8.1 O determinante representa um fator de volume orientado.

A1 Se substituirmos $\mathbf{v}_1$ por $c\mathbf{v}_1$, a área orientada será multiplicada por c.



A2 Se substituirmos $\mathbf{v}_2$ por $\mathbf{v}_2 + c\mathbf{v}_1$, a área orientada não mudará. Geometricamente, essa operação altera a forma do paralelogramo, mas preserva sua base e sua altura.



A propriedade A1, em particular, permanece válida para $c = 0$, quando um dos lados se torna o vetor zero; nesse caso, o paralelogramo degenera em um segmento de reta (ou em um ponto, se os dois lados forem nulos), e sua área é 0.

Passemos agora a um corpo arbitrário $\mathbb{K}$ e a matrizes $n \times n$. As propriedades A1 e A2 motivam uma noção algébrica de volume orientado, formulada diretamente em termos das colunas.

Definição 8.11 — Função Volume. Uma **função de volume** é uma aplicação $\mathrm{vol} : \mathscr{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ que satisfaz:

V1 Para qualquer escalar c e qualquer i,

$$\mathrm{vol}\left(\left[\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_{i-1} \mid c\mathbf{v}_i \mid \mathbf{v}_{i+1} \mid \cdots \mid \mathbf{v}_n\right]\right) \tag{8.1}$$
$$= c\,\mathrm{vol}\left(\left[\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_{i-1} \mid \mathbf{v}_i \mid \mathbf{v}_{i+1} \mid \cdots \mid \mathbf{v}_n\right]\right) \tag{8.2}$$

V2 Para qualquer escalar c e qualquer $j \neq i$,

$$\mathrm{vol}\left(\left[\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_{i-1} \mid \mathbf{v}_i + c\mathbf{v}_j \mid \mathbf{v}_{i+1} \mid \cdots \mid \mathbf{v}_n\right]\right) \tag{8.3}$$
$$= \mathrm{vol}\left(\left[\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_{i-1} \mid \mathbf{v}_i \mid \mathbf{v}_{i+1} \mid \cdots \mid \mathbf{v}_n\right]\right) \tag{8.4}$$

Ainda não mostramos que uma função de volume existe. Por enquanto, deduziremos as propriedades que qualquer função desse tipo deve satisfazer; depois, construiremos explicitamente uma delas.

Observamos também que a função vol não é única, pois podemos multiplicá-la por um fator arbitrário e obter outra função volume. Depois de fixarmos uma escala, obteremos

uma função única, indicada por det: o determinante. É conveniente trabalhar primeiro com funções volume arbitrárias e normalizar o resultado apenas ao final, exigindo que o volume orientado do cubo unitário n-dimensional, com as colunas na ordem padrão, seja 1.

Proposição 8.12.
Seja $\mathrm{vol} : \mathscr{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ *uma função de volume. Então:*

1. *Se alguma coluna de* A *é nula, então* $\mathrm{vol}(A) = 0$.
2. *Se as colunas de* A *são linearmente dependentes, então* $\mathrm{vol}(A) = 0$. *Em particular, isso ocorre quando duas colunas são iguais.*
3. *Se* $A^{(ij)}$ *é obtida de* A *pela troca das colunas* i *e* j*, então*
$$\mathrm{vol}\big(A^{(ij)}\big) = -\mathrm{vol}(A).$$
4. *Se* $A[j \leftarrow \mathbf{u}]$ *denota a matriz obtida substituindo a coluna* j *de* A *por* $\mathbf{u}$*, então*
$$\mathrm{vol}\big(A[j \leftarrow a\mathbf{u} + b\mathbf{w}]\big) = a\,\mathrm{vol}\big(A[j \leftarrow \mathbf{u}]\big) + b\,\mathrm{vol}\big(A[j \leftarrow \mathbf{w}]\big).$$

Demonstração.

1 Seja $\mathbf{v}_i = 0$. Então $\mathbf{v}_i = 0\mathbf{v}_i$, então pela propriedade V1

$$\mathrm{vol}\left(\left[\mathbf{v}_1 \;\middle|\; \cdots \;\middle|\; \mathbf{v}_i \;\middle|\; \cdots \;\middle|\; \mathbf{v}_n\right]\right) = 0 \cdot \mathrm{vol}\left(\left[\mathbf{v}_1 \;\middle|\; \cdots \;\middle|\; \mathbf{v}_i \;\middle|\; \cdots \;\middle|\; \mathbf{v}_n\right]\right) = 0.$$

2 Se as colunas de A são linearmente dependentes, existe uma relação

$$a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \vec{0}$$

com algum coeficiente $a_i \neq 0$. Portanto,

$$\mathbf{v}_i = -\sum_{j \neq i} \frac{a_j}{a_i}\mathbf{v}_j.$$

Aplicando repetidamente V2 à coluna i, removemos cada parcela dessa soma sem alterar o valor de vol. Ao final, a coluna i torna-se nula; pelo item anterior, $\mathrm{vol}(A) = 0$.

3

$$\begin{aligned}
\mathrm{vol}\left(\left[\mathbf{v}_1 \;\middle|\; \cdots \;\middle|\; \mathbf{v}_j \;\middle|\; \cdots \;\middle|\; \mathbf{v}_i \;\middle|\; \cdots \;\middle|\; \mathbf{v}_n\right]\right) &= \mathrm{vol}\left(\left[\mathbf{v}_1 \;\middle|\; \cdots \;\middle|\; \mathbf{v}_j \;\middle|\; \cdots \;\middle|\; \mathbf{v}_j + \mathbf{v}_i \;\middle|\; \cdots \;\middle|\; \mathbf{v}_n\right]\right) \\
&= \mathrm{vol}\left(\left[\mathbf{v}_1 \;\middle|\; \cdots \;\middle|\; -\mathbf{v}_i \;\middle|\; \cdots \;\middle|\; \mathbf{v}_j + \mathbf{v}_i \;\middle|\; \cdots \;\middle|\; \mathbf{v}_n\right]\right) \\
&= \mathrm{vol}\left(\left[\mathbf{v}_1 \;\middle|\; \cdots \;\middle|\; -\mathbf{v}_i \;\middle|\; \cdots \;\middle|\; \mathbf{v}_j \;\middle|\; \cdots \;\middle|\; \mathbf{v}_n\right]\right) \\
&= -\mathrm{vol}\left(\left[\mathbf{v}_1 \;\middle|\; \cdots \;\middle|\; \mathbf{v}_i \;\middle|\; \cdots \;\middle|\; \mathbf{v}_j \;\middle|\; \cdots \;\middle|\; \mathbf{v}_n\right]\right)
\end{aligned}$$

4 Se $\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n\}$ não for linearmente independente. Então, a afirmação é direta.

Suponha que $\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n\}$ seja linearmente independente. Podemos estender esse conjunto para uma base $\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n, \mathbf{z}\}$ de $\mathbb{K}^n$. Então

podemos escrever

$$\mathbf{u} = c_1\mathbf{v}_1 + \cdots + c_{i-1}\mathbf{v}_{i-1} + c_{i+1}\mathbf{v}_{i+1} + \cdots + c_n\mathbf{v}_n + c'\mathbf{z},$$
$$\mathbf{w} = d_1\mathbf{v}_1 + \cdots + d_{i-1}\mathbf{v}_{i-1} + d_{i+1}\mathbf{v}_{i+1} + \cdots + d_n\mathbf{v}_n + d'\mathbf{z}.$$

Seja $\mathbf{v} = a\mathbf{u} + b\mathbf{w}$. Então

$$\mathbf{v} = e_1\mathbf{v}_1 + \cdots + e_{i-1}\mathbf{v}_{i-1} + e_{i+1}\mathbf{v}_{i+1} + \cdots + e_n\mathbf{v}_n + e'\mathbf{z}$$

onde $e' = ac' + bd'$.

Aplicando a propriedade V2 repetidamente e a propriedade V1, vemos que

$$\begin{aligned}
\operatorname{vol}\left([\mathbf{v}_1 \mid \cdots \mid \mathbf{v} \mid \cdots \mid \mathbf{v}_n]\right) &= e' \operatorname{vol}\left([\mathbf{v}_1 \mid \cdots \mid \mathbf{z} \mid \cdots \mid \mathbf{v}_n]\right) \\
\operatorname{vol}\left([\mathbf{v}_1 \mid \cdots \mid \mathbf{u} \mid \cdots \mid \mathbf{v}_n]\right) &= c' \operatorname{vol}\left([\mathbf{v}_1 \mid \cdots \mid \mathbf{z} \mid \cdots \mid \mathbf{v}_n]\right) \\
\operatorname{vol}\left([\mathbf{v}_1 \mid \cdots \mid \mathbf{w} \mid \cdots \mid \mathbf{v}_n]\right) &= d' \operatorname{vol}\left([\mathbf{v}_1 \mid \cdots \mid \mathbf{z} \mid \cdots \mid \mathbf{v}_n]\right)
\end{aligned}$$

o que conclui a demonstração desse item. ■

Teorema 8.13.
Uma aplicação $f : \mathscr{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ é uma função de volume se, e somente se, satisfaz:

1 *Multilinearidade: se $A = [\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n]$ e $\mathbf{v}_i = a\mathbf{u} + b\mathbf{w}$ para algum i, então*

$$\begin{aligned}
f\left([\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_i \mid \cdots \mid \mathbf{v}_n]\right) = af\left([\mathbf{v}_1 \mid \cdots \mid \mathbf{u} \mid \cdots \mid \mathbf{v}_n]\right) \\
+bf\left([\mathbf{v}_1 \mid \cdots \mid \mathbf{w} \mid \cdots \mid \mathbf{v}_n]\right)
\end{aligned}$$

2 *Alternatividade: se $A = [\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n]$ e $\mathbf{v}_i = \mathbf{v}_j$ para $i \neq j$, então*

$$f\left([\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n]\right) = 0.$$

Demonstração. Já provamos que qualquer função de volume satisfaz os itens 3 e 4 da Proposição 8.12, que fornecem alternatividade e multilinearidade. Reciprocamente, a multilinearidade implica V1, e a alternatividade, combinada com a multilinearidade, implica V2. ■

As condições do Teorema 8.13 são geralmente consideradas como a definição de uma função de volume.

Observação 8.14. Em característica 2, antissimetria não implica alternatividade. Por exemplo, a aplicação bilinear

$$f\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ab$$

satisfaz $f(\mathbf{v}_2, \mathbf{v}_1) = -f(\mathbf{v}_1, \mathbf{v}_2)$, pois $-1 = 1$ no corpo, mas em geral $f(\mathbf{v}, \mathbf{v}) \neq 0$.

Definição 8.15 — Determinante. Uma aplicação $\det : \mathscr{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ é denominada **determinante** se satisfaz:

D1 Multilinearidade: se $A = [\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n]$ e $\mathbf{v}_i = a\mathbf{u} + b\mathbf{w}$ para algum i, então

$$\det\left([\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_i \mid \cdots \mid \mathbf{v}_n]\right) = a\det\left([\mathbf{v}_1 \mid \cdots \mid \mathbf{u} \mid \cdots \mid \mathbf{v}_n]\right) + b\det\left([\mathbf{v}_1 \mid \cdots \mid \mathbf{w} \mid \cdots \mid \mathbf{v}_n]\right)$$

D2 Alternatividade: se $A = [\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n]$ e $\mathbf{v}_i = \mathbf{v}_j$ para $i \neq j$, então

$$\det\left([\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n]\right) = 0.$$

D3 Normalização:

$$\det\left([\mathbf{e}_1 \mid \cdots \mid \mathbf{e}_n]\right) = 1.$$

Caracterização do determinante
Multilinearidade, alternância e a normalização na matriz identidade caracterizam unicamente o determinante. A fórmula de Leibniz permite construir uma função com essas propriedades.

Primeiramente provaremos a unicidade da função determinante.

Teorema 8.16 (Unicidade do Determinante).
Suponha que exista uma função de volume não nula $\mathrm{vol} : \mathscr{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$. *Então existe uma única função determinante. Além disso, toda função de volume é da forma* $a\det$ *para algum* $a \in \mathbb{K}$.

Demonstração. Como vol não é nula, existe A tal que $\mathrm{vol}(A) \neq 0$. Pela Proposição 8.12, A é invertível. Portanto, uma sequência de operações elementares sobre as colunas transforma A em I. Cada operação multiplica o valor de vol por um escalar não nulo; logo, $\mathrm{vol}(\mathrm{I}) \neq 0$.

Todo múltiplo escalar de uma função de volume ainda é uma função de volume. Portanto, podemos normalizar vol definindo $\det(A) = \mathrm{vol}(A)/\mathrm{vol}(\mathrm{I})$; assim, $\det(\mathrm{I}) = 1$. Para $a \in \mathbb{K}$, ponha

$$\mathrm{vol}_a(A) = a\det(A).$$

Agora, seja f uma função de volume qualquer e ponha $a = f(\mathrm{I})$. Se A é singular, então $f(A) = 0 = \mathrm{vol}_a(A)$. Se A é invertível, escolha uma sequência de operações sobre as colunas que transforme I em A. O efeito de cada operação sobre o valor de uma função de volume independe da função escolhida. Se b é o produto dos fatores correspondentes, então

$$f(A) = bf(\mathrm{I}) = ba = b\,\mathrm{vol}_a(\mathrm{I}) = \mathrm{vol}_a(A)\,,$$

Logo, $f = \mathrm{vol}_a$. Em particular, toda função de volume que vale 1 em I coincide com det. ∎

O teorema prova apenas a unicidade, sob a hipótese de existência de uma função de volume não nula. A existência será estabelecida na próxima seção por meio de uma fórmula explícita.

De onde vem o sinal
Na expansão por permutações, cada termo escolhe exatamente uma entrada de cada linha e de cada coluna. O sinal da permutação registra quantas inversões de orientação são necessárias para reorganizar essa escolha.

8.3 A fórmula de Leibniz

Seja $A = \left[\mathbf{v}_1 \mid \dots \mid \mathbf{v}_n\right]$. Escrevendo cada um desses vetores em termos da base canônica de $\mathbb{K}^n$, obtemos

$$\begin{aligned}
\mathbf{v}_1 &= v_{11}\mathbf{e}_1 + \dots + v_{n1}\mathbf{e}_n, \\
\mathbf{v}_2 &= v_{12}\mathbf{e}_1 + \dots + v_{n2}\mathbf{e}_n, \\
&\vdots \qquad \vdots \\
\mathbf{v}_n &= v_{1n}\mathbf{e}_1 + \dots + v_{nn}\mathbf{e}_n
\end{aligned}$$

Dessa forma

$$\begin{aligned}
\det(\mathbf{v}_1, \dots, \mathbf{v}_n) &= \det(v_{11}\mathbf{e}_1 + \dots + v_{n1}\mathbf{e}_n, \dots, v_{1n}\mathbf{e}_1 + \dots + v_{nn}\mathbf{e}_n) && (8.5) \\
&= \sum_{i_1,\dots,i_n=1}^{n} v_{i_1 1} v_{i_2 2} \cdots v_{i_n n} \det(\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_n}) && (8.6)
\end{aligned}$$

No somatório 8.6, anulam-se todos os termos em que algum dos índices $i_1, \dots, i_n$ se repete. Nesse caso, $\mathbf{e}_{i_k} = \mathbf{e}_{i_j}$ para algum $j \neq k$ e, pela propriedade D2, o determinante correspondente é zero.

Restam apenas as escolhas em que $i_1, \dots, i_n$ são distintos, isto é, as escolhas descritas pelas permutações de $[n]$. Portanto,

$$\det(\mathbf{v}_1, \dots, \mathbf{v}_n) = \sum_{\sigma \in S_n} v_{\sigma(1)1} v_{\sigma(2)2} \cdots v_{\sigma(n)n} \det(\mathbf{e}_{\sigma(1)}, \dots, \mathbf{e}_{\sigma(n)}),$$

em que σ percorre o grupo simétrico S_n. Como toda permutação é um produto de transposições e cada transposição troca duas colunas, temos

$$\det(\mathbf{e}_{\sigma(1)}, \dots, \mathbf{e}_{\sigma(n)}) = \operatorname{sgn}(\sigma) \det(\mathbf{e}_1, \dots, \mathbf{e}_n) = \operatorname{sgn}(\sigma).$$

E logo

$$\det(\mathbf{v}_1, \dots, \mathbf{v}_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) v_{\sigma(1)1} v_{\sigma(2)2} \cdots v_{\sigma(n)n},$$

que é a expressão do determinante em termos de permutações.

Podemos agora demonstrar a existência do determinante.

> **Teorema 8.17** (Existência do Determinante).
> *Existe uma única função de volume* $\det : \mathscr{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ *tal que* $\det(\mathrm{I}) = 1$.

Demonstração. Já demonstramos a unicidade. Mostraremos agora que

$$\det(\mathbf{v}_1, \dots, \mathbf{v}_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) v_{\sigma(1)1} v_{\sigma(2)2} \cdots v_{\sigma(n)n}, \tag{8.7}$$

satisfaz as propriedades D1 – D3. No que se segue, σ denotará uma permutação de $\{1, \dots, n\}$.

D2 Suponhamos que as colunas $\mathbf{v}_i$ e $\mathbf{v}_j$ sejam iguais, com $i \neq j$, e seja τ a transposição de i e j. A aplicação $\sigma \mapsto \sigma \circ \tau$ particiona S_n em pares. Compor à direita por τ troca os

fatores provenientes das colunas i e j; como essas colunas são iguais, os produtos das entradas coincidem, enquanto os sinais são opostos. Logo as duas parcelas se cancelam. Esse argumento vale também em característica 2, quando a soma das parcelas iguais é nula. Portanto, a expressão em 8.7 é alternada.

D1 A linearidade é imediata, pois cada parcela de 8.7 contém exatamente uma coordenada do vetor $\mathbf{v}_i + k\mathbf{u}_i$, de modo que

$$\begin{aligned}&\det(\mathbf{v}_1,\ \dots,\ \mathbf{v}_i + k\mathbf{u}_i,\ \dots,\ \mathbf{v}_n)\\ &= \det(\mathbf{v}_1,\ \dots,\ \mathbf{v}_i,\ \dots,\ \mathbf{v}_n) + k\det(\mathbf{v}_1,\ \dots,\ \mathbf{u}_i,\ \dots,\ \mathbf{v}_n).\end{aligned}$$

D3 Vamos mostrar que $\det(\mathbf{e}_1, \dots,\ \mathbf{e}_n) = 1$.

Se $\sigma(i) \neq i$ para algum i, então $v_{\sigma(i)i} = 0$. Assim, apenas a permutação identidade produz um termo não nulo; seu sinal e todos os fatores são iguais a 1. Logo $\det(\mathbf{e}_1, \dots, \mathbf{e}_n) = 1$. ■

Proposição 8.18.
Seja $A \in \mathscr{M}_{n,n}(\mathbb{K})$. Então $\det(A) \neq 0$ se, e somente se, A é invertível.

Demonstração. Se A é singular, suas colunas são linearmente dependentes e a Proposição 8.12 dá $\det(A) = 0$. Se A é invertível, operações elementares sobre as colunas transformam A em I. Cada operação multiplica o determinante por um escalar não nulo; como $\det(\mathrm{I}) = 1$, segue que $\det(A) \neq 0$. ■

Exemplo 8.19 — Quando aparece a singularidade. Para

$$A_t = \begin{bmatrix} 1 & t \\ t & 1 \end{bmatrix},$$

a fórmula de Leibniz fornece

$$\det(A_t) = 1 - t^2 = (1 - t)(1 + t).$$

Logo A_t é invertível exatamente quando $t \neq \pm 1$. Nos dois valores excluídos, as colunas se tornam iguais ou opostas, e o paralelogramo por elas determinado colapsa em um segmento.

◁

Notação 8.20. O determinante de A, denotado por $\det A$ ou $|A|$, também pode ser escrito colocando-se as entradas da matriz entre barras:

$$\det A = \begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{vmatrix}.$$

Compor multiplica escalas
A identidade $\det(AB) = \det(A)\det(B)$ expressa um fato geométrico: aplicar primeiro B e depois A multiplica sucessivamente os fatores de escala orientada. Ela também explica imediatamente por que uma matriz é invertível exatamente quando seu determinante não é zero.

8.4 Propriedades

Deduziremos agora outras propriedades do determinante.

Teorema 8.21.
Sejam $A, B \in \mathscr{M}_{n,n}(\mathbb{K})$. Então

$$\det(AB) = \det(A)\det(B).$$

Demonstração. Defina $f : \mathscr{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ por $f(B) = \det(AB)$. Como a multiplicação por A é linear em cada coluna, f é multilinear; se duas colunas de B coincidem, o mesmo ocorre em AB, logo f é alternada. Portanto, $f(B) = a\det(B)$, com $a = f(\mathrm{I}) = \det(A)$. Assim, $\det(AB) = \det(A)\det(B)$. ■

Proposição 8.22. [1] $\det(A) \neq 0$ *se, e somente se, A for invertível.*

[2] *Se A for invertível,* $\det(A^{-1}) = 1/\det(A)$.

[3] *Se A for invertível, então para qualquer matriz B,* $\det(ABA^{-1}) = \det(B)$.

Demonstração. O primeiro item é a Proposição 8.18. Se A é invertível, então

$$1 = \det(\mathrm{I}) = \det(AA^{-1}) = \det(A)\det(A^{-1}),$$

o que prova o segundo item. Finalmente,

$$\det(ABA^{-1}) = \det(A)\det(B)\det(A^{-1}) = \det(B),$$

e isso prova o terceiro. ■

Proposição 8.23. [1] *Seja A uma matriz diagonal. Então* $\det(A)$ *é o produto de suas entradas diagonais.*

[2] *De maneira mais geral, seja A uma matriz triangular superior ou uma triangular inferior. Então* $\det(A)$ *é o produto de suas entradas diagonais.*

Demonstração.

[1] Se A é diagonal, existe apenas um termo não nulo na Equação 8.7: o correspondente à permutação identidade. Esse termo é o produto das entradas diagonais.

[2] Se σ não é a identidade, existem índices j e k tais que $\sigma(j) < j$ e $\sigma(k) > k$. Assim, tanto para uma matriz triangular superior quanto para uma triangular inferior, a parcela associada a σ contém um fator nulo. Resta novamente apenas o termo diagonal. ■

Exemplo 8.24 — Cálculo por operações de colunas. Considere

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 5 & 1 \\ 0 & 1 & 3 \end{bmatrix}.$$

Somar a uma coluna um múltiplo de outra não altera o determinante. Portanto,

$$\begin{aligned}\det(A) &= \det\begin{bmatrix}1 & 0 & 0\\ 2 & 1 & 1\\ 0 & 1 & 3\end{bmatrix} \quad (C_2 \leftarrow C_2 - 2C_1)\\ &= \det\begin{bmatrix}1 & 0 & 0\\ 2 & 1 & 0\\ 0 & 1 & 2\end{bmatrix} \quad (C_3 \leftarrow C_3 - C_2)\\ &= 1 \cdot 1 \cdot 2 = 2.\end{aligned}$$

A triangularização evita expandir seis parcelas pela fórmula de Leibniz.

◁

Teorema 8.25. 1 *Seja M uma matriz diagonal por bloco,*

$$M = \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}.$$

Então $\det(M) = \det(A)\det(D)$.

2 *Em geral, seja M uma matriz triangular superior por bloco ou uma matriz triangular inferior por bloco,*

$$M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix} \quad \textit{ou} \quad M = \begin{bmatrix} A & 0 \\ C & D \end{bmatrix}$$

Então $\det(M) = \det(A)\det(D)$.

Demonstração. Considere primeiro $M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$. Na fórmula de Leibniz, uma parcela só pode ser não nula se as linhas correspondentes ao bloco inferior forem associadas às colunas do bloco D. As permutações que sobrevivem são, portanto, exatamente os pares formados por uma permutação das linhas e colunas de A e outra das linhas e colunas de D. A soma das parcelas se fatora e fornece

$$\det(M) = \det(A)\det(D).$$

O caso diagonal por blocos é particular. Para $M = \begin{bmatrix} A & 0 \\ C & D \end{bmatrix}$, o mesmo argumento se aplica, agora começando pelas linhas do bloco superior. ■

Proposição 8.26.
Seja A^{t} a matriz transposta de A. Então $\det(A^{\mathrm{t}}) = \det(A)$.

Demonstração. Seja $B = (b_{ij}) = A^{\mathrm{t}}$. Fazendo a mudança de índice $\rho = \sigma^{-1}$ na fórmula

de Leibniz e usando $\operatorname{sgn}(\rho) = \operatorname{sgn}(\rho^{-1})$, obtemos

$$\begin{aligned}\det(A) &= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} a_{\sigma(j),j} \\ &= \sum_{\rho\in S_n} \operatorname{sgn}(\rho) \prod_{i=1}^{n} a_{i,\rho(i)} \\ &= \sum_{\rho\in S_n} \operatorname{sgn}(\rho) \prod_{i=1}^{n} b_{\rho(i),i} = \det(B).\end{aligned}$$

■

Denote por A_{ij} a submatriz complementar à entrada (i, j), obtida pela exclusão da linha i e da coluna j de A. O escalar $\det(A_{ij})$ é o menor associado a essa entrada.

Teorema 8.27 (Expansão de Laplace).
Seja A uma matriz n por n, $A = (a_{ij})$.

1. *Para qualquer i,*
$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) .$$

2. *Para qualquer j,*
$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) .$$

3. *Para qualquer i e para qualquer $k \neq i$,*
$$0 = \sum_{j=1}^{n} (-1)^{i+j} a_{kj} \det(A_{ij}) .$$

4. *Para qualquer j e para qualquer $k \neq j$,*
$$0 = \sum_{i=1}^{n} (-1)^{i+j} a_{ik} \det(A_{ij}) .$$

Ferramenta de cálculo, não definição
A expansão de Laplace é especialmente útil quando uma linha ou coluna contém muitos zeros. Para matrizes grandes e densas, eliminação gaussiana é muito mais eficiente; a importância da fórmula aqui é também teórica.

Demonstração. Demonstraremos 1 e 3 simultaneamente, portanto fixamos k (que pode ou não ser igual a i).

Cada parcela do lado direito é linear em cada coluna; portanto, a soma é multilinear.

Agora mostramos que é alternado. Seja A uma matriz com as colunas p e q iguais, onde $1 \leq p < q \leq n$. Se $j \neq p, q$, então A_{ij} tem duas colunas iguais e $\det(A_{ij}) = 0$. Assim, os únicos dois termos que podem contribuir para a soma são

$$(-1)^{i+p} a_{kp} \det(A_{ip}) + (-1)^{i+q} a_{kq} \det(A_{iq}) .$$

Por hipótese, $a_{kq} = a_{kp}$. Agora

$$A_{ip} = [\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_{p-1} \mid \mathbf{v}_{p+1} \mid \cdots \mid \mathbf{v}_{q-1} \mid \mathbf{v}_q \mid \mathbf{v}_{q+1} \mid \cdots \mid \mathbf{v}_n],$$

$$A_{iq} = [\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_{p-1} \mid \mathbf{v}_p \mid \mathbf{v}_{p+1} \mid \cdots \mid \mathbf{v}_{q-1} \mid \mathbf{v}_{q+1} \mid \cdots \mid \mathbf{v}_n].$$

onde $\mathbf{v}_m$ indica a coluna m depois da exclusão da linha i. Por hipótese, $\mathbf{v}_p = \mathbf{v}_q$; portanto, essas duas matrizes têm as mesmas colunas, em ordens diferentes. Passamos da primeira para a segunda por $q-p-1$ trocas sucessivas de colunas. Logo $\det(A_{iq}) = (-1)^{q-p-1}\det(A_{ip})$. Assim, a contribuição desses dois termos é

$$(-1)^{i+p}a_{kp}\det(A_{ip}) + (-1)^{i+q}a_{kp}(-1)^{q-p-1}\det(A_{ip})$$

e como $(-1)^{i+p}$ e $(-1)^{i+2q-p-1}$ sempre têm sinais opostos, eles cancelam.

Pela unicidade, o lado direito é um múltiplo de $\det(A)$. Avaliando em $A = \mathrm{I}$, esse múltiplo é 1 se $k = i$ e 0 se $k \neq i$, o que prova simultaneamente os itens 1 e 3. Os itens 2 e 4 seguem aplicando esses resultados à matriz transposta. ■

Exemplo 8.28. Para $A \in \mathscr{M}_{3,3}(\mathbb{K})$ temos

$$\det A = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a\begin{vmatrix} \square & \square & \square \\ \square & e & f \\ \square & h & i \end{vmatrix} - b\begin{vmatrix} \square & \square & \square \\ d & \square & f \\ g & \square & i \end{vmatrix} + c\begin{vmatrix} \square & \square & \square \\ d & e & \square \\ g & h & \square \end{vmatrix} \tag{8.8}$$

$$= a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix} \tag{8.9}$$

$$= aei + bfg + cdh - ceg - bdi - afh. \tag{8.10}$$

◁

Definição 8.29 — Cofatores e Matriz Adjugada. Dada $A \in \mathscr{M}_{n,n}(\mathbb{K})$:

- O **cofator** da entrada a_{ij} é o escalar $C_{ij} = (-1)^{i+j}\det(A_{ij})$, em que A_{ij} é a matriz obtida de A pela exclusão da linha i e da coluna j.
- A **matriz dos cofatores** de A é $C = (C_{ij})$.
- A **matriz adjugada** de A, denotada por $\operatorname{adj}(A)$, é a transposta de sua matriz de cofatores.

Proposição 8.30.
Para qualquer matriz A,

1. $\operatorname{adj}(A)A = A\operatorname{adj}(A) = \det(A)\,\mathrm{I}$.
2. *Se A for invertível,*

$$A^{-1} = \frac{1}{\det(A)}\operatorname{adj}(A).$$

Demonstração. A entrada (i,j) de $A\operatorname{adj}(A)$ é

$$\sum_{k=1}^{n} a_{ik}C_{jk}.$$

Se $i = j$, essa soma é a expansão de Laplace de $\det(A)$ pela linha i. Se $i \neq j$, ela é a expansão do determinante da matriz obtida de A substituindo a linha j pela linha i; como essa matriz tem duas linhas iguais, a soma é zero. Portanto, $A\operatorname{adj}(A) = \det(A)\,\mathrm{I}$. Aplicando o mesmo argumento às colunas, obtemos $\operatorname{adj}(A)A = \det(A)\,\mathrm{I}$. Se A é invertível, a segunda afirmação segue dividindo a primeira igualdade por $\det(A) \neq 0$. ■

Observação 8.31 — Determinantes sobre anéis comutativos. A fórmula de Leibniz, a expansão de Laplace e a identidade

$$A\,\mathrm{adj}(A) = \mathrm{adj}(A)A = \det(A)\,\mathrm{I}$$

usam apenas somas e produtos. Por isso, continuam válidas para matrizes com entradas em qualquer anel comutativo com unidade, em particular em $\mathbb{K}[x]$. Nesse contexto, porém, a caracterização correta da invertibilidade é: A é invertível se, e somente se, $\det(A)$ é uma unidade do anel. Não basta exigir $\det(A) \neq 0$.

Teorema 8.32 (Regra de Cramer).
Seja $A \in \mathscr{M}_{n,n}(\mathbb{K})$ invertível e seja $\vec{b} \in \mathbb{K}^n$. Considere o sistema $A\mathbf{x} = \vec{b}$. Então:

1. *O sistema possui uma única solução $\mathbf{x} \in \mathbb{K}^n$.*
2. *Se $\mathbf{x} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^{\mathrm{t}}$, então*

$$x_i = \frac{\det(A_i(\vec{b}))}{\det(A)}, \qquad 1 \leq i \leq n,$$

onde $A_i(\vec{b})$ é obtida de A substituindo sua i-ésima coluna por $\vec{b}$.

Fórmula de Cramer
A regra de Cramer expressa cada coordenada da solução como um quociente de determinantes. Essa expressão pode ser usada em argumentos teóricos; para resolver sistemas por cálculo numérico, em geral se empregam métodos de eliminação.

Demonstração. A invertibilidade de A garante a existência e a unicidade da solução. Denote as colunas de A por $\vec{a}_1, \ldots, \vec{a}_n$. Se $\vec{b} = A\vec{x} = \sum_{j=1}^{n} x_j \vec{a}_j$, a multilinearidade na i-ésima coluna fornece

$$\det A_i(\vec{b}) = \sum_{j=1}^{n} x_j \det A_i(\vec{a}_j) = x_i \det A,$$

pois os termos com $j \neq i$ têm duas colunas iguais, enquanto $A_i(\vec{a}_i) = A$. Como $\det A \neq 0$, segue a fórmula. ∎

A fórmula fornece uma expressão explícita da solução. Para o cálculo de sistemas de maior ordem, contudo, a eliminação evita avaliar separadamente os determinantes que aparecem em cada coordenada.

Definição 8.33 — Posto por Determinante. O **posto por determinante** de A é a maior ordem de uma submatriz quadrada de A com determinante não nulo. Adotamos o valor 0 quando A é a matriz nula. Esse posto será denotado por $\mathrm{posto}_{\det}(A)$.

Teorema 8.34.
Para toda matriz A, os postos por linha, por coluna e por determinante são iguais:

$$\mathrm{posto}(A) = \mathrm{posto}_{\det}(A) = \mathrm{posto}_{l}(A) = \mathrm{posto}_{c}(A).$$

Demonstração. Como os postos por linha e por coluna coincidem, basta comparar o posto por coluna com o posto por determinante.

Se A possui um menor não nulo de ordem k, as k colunas que participam desse menor são linearmente independentes: qualquer relação entre elas continuaria válida depois de restringirmos às k linhas escolhidas, contradizendo a invertibilidade da submatriz. Logo $\text{posto}_c(A) \geq \text{posto}_{\det}(A)$.

Reciprocamente, seja $k = \text{posto}_c(A)$ e escolha k colunas linearmente independentes. A matriz C formada por essas colunas tem posto k. Como o posto por linha de C também é k, existem k linhas de C linearmente independentes. A submatriz quadrada determinada por essas linhas e pelas colunas escolhidas é invertível; portanto, possui determinante não nulo. Assim, $\text{posto}_{\det}(A) \geq k$, e os três postos coincidem.

■

Definimos o determinante para matrizes. Podemos agora definir o determinante de um operador $T : V \to V$ quando V tem dimensão finita.

Definição 8.35 — Determinante de um Operador. Seja $T : V \to V$ um operador linear, com V de dimensão finita. Fixada uma base $\underline{B}$ de V, definimos o **determinante** de T por

$$\det(T) = \det([T]_{\underline{B}}).$$

Essa definição independe da escolha de $\underline{B}$: matrizes do mesmo operador em bases diferentes são semelhantes, e o item 3 da Proposição 8.22 mostra que o determinante é invariante por similaridade.

Definição 8.36 — Grupos Linear e Linear Especial. O **grupo linear geral** e o **grupo linear especial** são, respectivamente,

$$\mathrm{GL}_n(\mathbb{K}) = \{A \in \mathscr{M}_{n,n}(\mathbb{K}) \mid \det A \neq 0\},$$
$$\mathrm{SL}_n(\mathbb{K}) = \{A \in \mathrm{GL}_n(\mathbb{K}) \mid \det A = 1\}.$$

Se V tem dimensão finita, definimos analogamente $\mathrm{GL}(V)$ como o grupo dos automorfismos de V e

$$\mathrm{SL}(V) = \{T \in \mathrm{GL}(V) \mid \det T = 1\}.$$

Proposição 8.37.
A aplicação

$$\det : \mathrm{GL}_n(\mathbb{K}) \longrightarrow \mathbb{K}^\times$$

é um homomorfismo de grupos, e seu núcleo é $\mathrm{SL}_n(\mathbb{K})$. *Consequentemente,* $\mathrm{SL}_n(\mathbb{K})$ *é um subgrupo normal de* $\mathrm{GL}_n(\mathbb{K})$. *O mesmo vale para* $\mathrm{SL}(V)$ *em* $\mathrm{GL}(V)$.

Demonstração. A identidade $\det(AB) = \det(A)\det(B)$ prova que o determinante é um homomorfismo. A descrição do núcleo segue das definições, e todo núcleo de homomorfismo é um subgrupo normal. ■

Exercícios

Ex. 8.15 — Calcule o seguinte determinante, onde $a_i := \sum_{k=1}^{i} k$:

$$\begin{vmatrix} a_1 & a_1 & \cdots & \cdots & a_1 \\ a_1 & a_2 & a_2 & \cdots & a_2 \\ \vdots & a_2 & \ddots & \cdots & \vdots \\ \vdots & \vdots & \vdots & a_{n-1} & a_{n-1} \\ a_1 & a_2 & \cdots & a_{n-1} & a_n \end{vmatrix}$$

Ex. 8.16 — Calcule o seguinte determinante $n \times n$, onde $a, b \in \mathbb{K}$:

$$D_n(a,b) = \begin{vmatrix} a+b & ab & 0 & \cdots & 0 \\ 1 & a+b & ab & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & a+b & ab \\ 0 & \cdots & 0 & 1 & a+b \end{vmatrix}$$

Ex. 8.17 — Seja $V = V_1 \oplus V_2$, com V_1 e V_2 de dimensões finitas, e sejam $f_i \in \mathrm{Hom}(V_i, V_i)$ para $i = 1, 2$. Defina $f \in \mathrm{Hom}(V, V)$ por

$$f(\mathbf{v}_1 + \mathbf{v}_2) = f_1(\mathbf{v}_1) + f_2(\mathbf{v}_2).$$

Prove que $\det(f) = \det(f_1)\det(f_2)$.

Ex. 8.18 — Prove que o determinante da matriz $n \times n$ de entradas $a_{ij} = 1 - \delta_{ij}$ é igual a $(n-1)(-1)^{n-1}$.

Ex. 8.19 — **[Critério de avaliação]** Sejam $f_1, \ldots, f_n : X \to \mathbb{K}$ funções definidas em um conjunto arbitrário X. Mostre que $\{f_1, \ldots, f_n\}$ é linearmente independente em $\mathscr{F}(X, \mathbb{K})$ se, e somente se, existem $x_1, \ldots, x_n \in X$ tais que:

$$\begin{vmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_n(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_n) & f_2(x_n) & \cdots & f_n(x_n) \end{vmatrix} \neq 0.$$

Ex. 8.20 — Seja $A \in M_n(\mathbb{Z})$. Mostre que A possui inversa em $M_n(\mathbb{Z})$ se, e somente se, $\det(A) = \pm 1$ (isto é, se, e somente se, $\det(A)$ é invertível em $\mathbb{Z}$).

Ex. 8.21 — **[Determinante de Vandermonde]** Dado um inteiro $n \geq 2$ e escalares $a_1, \ldots, a_n \in \mathbb{K}$, defina

$$\mathscr{V}_n(a_1, \ldots, a_n) = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{vmatrix}$$

o **determinante de Vandermonde**. Considere também o polinômio $p \in \mathbb{K}[x]$ o polinômio $p(x) = \mathscr{V}_n(a_1, \ldots, a_{n-1}, x)$.

1. Determine grau(p) e suas raízes.

2. Deduza uma expressão de p em função de $\mathscr{V}_{n-1}(a_1, \dots, a_{n-1})$.

3. Deduza uma expressão explícita para $\mathscr{V}_n(a_1, \dots, a_{n-1}, a_n)$.

Ex. 8.22 — [Determinante das multiplicações à esquerda e à direita] Dada uma matriz $A \in M_n(\mathbb{K})$, denote por

$$\begin{aligned} L_A, R_A \,:\, M_n(\mathbb{K}) &\to M_n(\mathbb{K}) \\ L_A(X) &= A \cdot X \\ R_A(X) &= X \cdot A \end{aligned}$$

as multiplicações à esquerda e à direita por A, respectivamente.

1. Verifique que L_A e R_A são lineares.

2. Mostre que $\det(L_A) = \det(R_A) = (\det A)^n$. **Dica**: escolha cuidadosamente uma base de $M_n(\mathbb{K})$.

Ex. 8.23 — [Wronskiano] Sejam $U \subseteq \mathbb{R}$ aberto, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, e $f_1, \dots, f_n : U \to \mathbb{K}$ funções $(n-1)$ vezes diferenciáveis. O **Wronskiano** dessas funções é a aplicação $\mathscr{W}(f_1, \dots, f_n) : U \to \mathbb{K}$ definida por

$$\mathscr{W}(f_1, \dots, f_n)(x) := \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ f_1''(x) & f_2''(x) & \cdots & f_n''(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}.$$

Mostre que, se existe $x \in U$ tal que $\mathscr{W}(f_1, \dots, f_n)(x) \neq 0$, então $\{f_1, \dots, f_n\}$ é linearmente independente. A recíproca é verdadeira?

Ex. 8.24 — Seja V um espaço vetorial de dimensão n sobre $\mathbb{K}$. Se F é uma aplicação n-linear em $(V^*)^n$ e G é uma aplicação n-linear em V^n, escrevemos

$$(F \boxtimes G)(f_1, \dots, f_n, v_1, \dots, v_n) = F(f_1, \dots, f_n)G(v_1, \dots, v_n).$$

1. Sejam Δ e Δ^* funções determinantes em V e V^*, respectivamente. Mostre que a aplicação $(2n)$-linear $\Delta^* \boxtimes \Delta$ assume o mesmo valor em quaisquer dois pares de bases duais, isto é,

   $$\Delta^* \boxtimes \Delta(v_1^*, \dots, v_n^*, v_1, \dots, v_n) = \Delta^* \boxtimes \Delta(w_1^*, \dots, w_n^*, w_1, \dots, w_n).$$

   Duas funções determinantes não nulas Δ e Δ^* em V e V^* são chamadas de **duais** quando

   $$\Delta^*(v_1^*, \dots, v_n^*) \cdot \Delta(v_1, \dots, v_n) = 1 \text{ sempre que } v_i^*(v_j) = \delta_{ij}.$$

2. Seja $\Delta \neq 0$ uma função determinante em V. Defina $\Delta^* : \underbrace{V^* \times \cdots \times V^*}_{n\text{-vezes}} \to \mathbb{K}$ da seguinte maneira. Se os vetores $v_1^*, \dots, v_n^*$ são linearmente dependentes, defina

$\Delta^*(v_1^*, \dots, v_n^*) = 0$. Se forem linearmente independentes, formarão uma base $\underline{C}$ de V^*; nesse caso, defina

$$\Delta^*(v_1^*, \dots, v_n^*) = \frac{1}{\Delta(v_1, \dots, v_n)},$$

onde $(v_1, \dots, v_n)$ é a base de V dual a $\underline{C}$. Prove que Δ^* é uma função determinante em V^* e é dual a Δ. Conclua que toda função determinante não nula em V possui uma única função determinante dual em V^*.

3. Sejam Δ e Δ^* funções determinantes duais. Mostre que

$$\Delta^* \boxtimes \Delta(v_1^*, \dots, v_n^*, u_1, \dots, u_n) = \begin{vmatrix} v_1^*(u_1) & v_1^*(u_2) & \cdots & v_1^*(u_n) \\ v_2^*(u_1) & v_2^*(u_2) & \cdots & v_2^*(u_n) \\ \vdots & \vdots & \ddots & \vdots \\ v_n^*(u_1) & v_n^*(u_2) & \cdots & v_n^*(u_n) \end{vmatrix}.$$

9

Estrutura dos Operadores Lineares

A representação matricial de uma transformação linear depende das bases escolhidas. Quando as bases do domínio e do contradomínio podem variar independentemente, toda transformação de posto r admite uma matriz com um bloco identidade de ordem r e zeros nas demais posições.

Para estudar um operador $T : V \to V$ por similaridade, usamos a mesma base no domínio e no contradomínio. Nesse caso, o posto não basta para a classificação. Introduziremos os autovalores e os polinômios característico e minimal; as formas de Schur, Jordan e racional fornecerão representações adaptadas à estrutura do operador.

Chamaremos de *normal* uma representação especialmente simples e de *canônica* uma representação que, além disso, escolhe um único representante em cada classe de equivalência.

Nesse capítulo

Formas normais e canônicas
Uma forma normal simplifica os representantes de uma classe. Uma forma canônica exige também unicidade: objetos equivalentes devem conduzir ao mesmo representante.

9.1 Equivalência de matrizes

Pela fórmula de mudança de bases,

$$[T]_{\underline{B'},\underline{C'}} = \mathrm{M}_{\underline{C}\to\underline{C'}}\,[T]_{\underline{B},\underline{C}}\,\mathrm{M}^{-1}_{\underline{B}\to\underline{B'}},$$
$$[T]_{\underline{B}',\underline{C}'} = P[T]_{\underline{B},\underline{C}}Q^{-1}$$

para matrizes invertíveis P e Q. Isso motiva a definição seguinte.

Definição 9.1 — Matrizes Equivalentes. Duas matrizes A e B são **equivalentes** se existirem matrizes invertíveis P e Q para as quais

$$B = PAQ^{-1}.$$

Operações elementares de linha correspondem à multiplicação à esquerda por uma matriz invertível; operações elementares de coluna, à multiplicação à direita. Assim, A e B são equivalentes exatamente quando uma pode ser obtida da outra por operações elementares de linhas e colunas.

A fórmula de mudança de bases fornece a interpretação intrínseca dessa relação.

Teorema 9.2.
Sejam V e W espaços vetoriais com $\dim V = n$ e $\dim W = m$. Então as matrizes $A, B \in \mathscr{M}_{m,n}(\mathbb{K})$ são equivalentes se, e somente se, representam a mesma transformação linear $T \in \mathrm{Hom}(V, W)$, possivelmente em bases ordenadas diferentes.

Demonstração. Se A e B representam T em pares de bases diferentes, a fórmula de mudança de bases mostra que $B = PAQ^{-1}$ para matrizes invertíveis P e Q; logo são equivalentes.

Reciprocamente, suponha $B = PAQ^{-1}$ e escolha bases $\underline{\mathrm{B}}$ de V e $\underline{\mathrm{C}}$ de W para as quais $A = [T]_{\underline{\mathrm{B}},\underline{\mathrm{C}}}$. Como toda matriz invertível pode ser realizada como matriz de mudança entre duas bases, podemos escolher novas bases $\underline{\mathrm{B}}'$ e $\underline{\mathrm{C}}'$ cujas matrizes de mudança sejam Q no domínio e P no contradomínio. A fórmula de mudança de bases fornece então $B = [T]_{\underline{\mathrm{B}}',\underline{\mathrm{C}}'}$. ■

O teorema seguinte mostra que o posto classifica as matrizes de um mesmo tamanho quando as bases do domínio e do contradomínio variam independentemente.

$$J_k = \begin{bmatrix} \mathrm{I}_k & 0_{k,n-k} \\ 0_{m-k,k} & 0_{m-k,n-k} \end{bmatrix}$$

Teorema 9.3 (da Forma Normal Zero-Um).
Seja $T : V \to W$ uma transformação linear entre espaços vetoriais de dimensão finita. Então:

1. *Existe uma decomposição direta $V = V_0 \oplus V_1, W = W_1 \oplus W_2$ tal que $\ker T = V_0$ e T induz um isomorfismo de V_1 em W_1.*
2. *Existem bases em V e W tais que a matriz de T nessas bases tem a forma (a_{ij}), com $a_{ii} = 1$ para $1 \leq i \leq r$ e $a_{ij} = 0$ para os valores restantes de i, j.*
3. *Seja $A \in \mathscr{M}_{m,n}(\mathbb{K})$. Então existem matrizes invertíveis $B \in \mathscr{M}_{m,m}(\mathbb{K})$ e $C \in \mathscr{M}_{n,n}(\mathbb{K})$ e um único número $r \leq \min(m, n)$ tais que BAC tenha o formato descrito no item anterior. Esse número r é o posto de A.*

Demonstração.

1. Defina $V_0 = \ker T$ e faça V_1 um complemento direto de V_0. Em seguida, defina $W_1 = \mathrm{im}\, T$ e faça W_2 um complemento direto de W_1. Precisamos apenas verificar se T determina um isomorfismo de V_1 a W_1.

   A aplicação $T|_{V_1} : V_1 \to W_1$ é injetiva, pois $V_0 \cap V_1 = \{\vec{0}\}$. É sobrejetiva porque, se $\vec{v} = \vec{v}_0 + \vec{v}_1$ com $\vec{v}_i \in V_i$, então $T(\vec{v}) = T(\vec{v}_1)$.

2. Ponha $r = \dim V_1 = \dim W_1$. Escolha uma base $(\vec{e}_1, \ldots, \vec{e}_r)$ de V_1 e complete-a com $(\vec{e}_{r+1}, \ldots, \vec{e}_n)$ para obter uma base de V. Como $T|_{V_1} : V_1 \to W_1$ é um isomorfismo, os vetores $\vec{e}'_i = T(\vec{e}_i)$, $1 \leq i \leq r$, formam uma base de W_1. Complete-os com $(\vec{e}'_{r+1}, \ldots, \vec{e}'_m)$ para obter uma base de W. Nessas bases,

$$T(\vec{e}_1,\ \ldots,\ \vec{e}_r,\ \vec{e}_{r+1},\ \ldots,\ \vec{e}_n) = (\vec{e}'_1,\ \ldots, \vec{e}'_r,\ 0,\ \ldots, 0)$$

Dessa forma, nas bases $\underline{B} = (\vec{e}_1, \dots, \vec{e}_n)$ de V e $\underline{B}' = (\vec{e}'_1, \dots, \vec{e}'_m)$ de W,

$$[T]_{\underline{B},\underline{B}'} = \begin{bmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{bmatrix}$$

3 A partir de A, considere $T_A : \mathbb{K}^n \to \mathbb{K}^m$, $T_A(\vec{x}) = A\vec{x}$, e aplique o item 2. Nas novas bases, a matriz de T_A tem a forma zero–um e é obtida de A por multiplicações à esquerda e à direita por matrizes invertíveis. Por fim, $r = \dim \operatorname{im} T_A = \text{posto } A$, o que também prova a unicidade de r.

■

As matrizes J_r formam, portanto, um sistema de representantes canônicos, e o posto é um invariante completo para a equivalência de matrizes.

Invariante completo
Um invariante é completo quando decide a equivalência: dois objetos são equivalentes se, e somente se, possuem o mesmo valor do invariante. Para matrizes sob equivalência, o posto cumpre exatamente esse papel.

9.2 Operadores Similares

Se um operador $T \in \operatorname{Hom}(V,V)$ é representado em duas bases, suas matrizes se relacionam por

$$[T]_{\underline{B}'} = P[T]_{\underline{B}}P^{-1}$$

para alguma matriz invertível P. Aqui, ao contrário da equivalência, a mesma mudança de base aparece dos dois lados.

Definição 9.4 — Operadores Semelhantes. a Dois operadores lineares $T, S \in \operatorname{Hom}(V,V)$ são ditos **similares** ou **semelhantes**, e escrevemos $T \sim S$, se existir um automorfismo P de V para o qual

$$S = PTP^{-1}$$

b Suas matrizes A e B são ditas **similares** ou **semelhantes**, indicadas por $A \sim B$, se existir uma matriz invertível P para a qual

$$B = PAP^{-1}$$

As classes de equivalência associadas a essa relação são chamadas classes de similaridade.

Classes de similaridade
As matrizes escalares $\lambda \mathrm{I}$ formam classes de um único elemento, pois $P(\lambda \mathrm{I})P^{-1} = \lambda \mathrm{I}$. Para uma matriz geral, a classe de similaridade pode ter uma geometria muito mais complicada.

Teorema 9.5.
Seja V um espaço vetorial de dimensão n. Então, dois operadores lineares T e S em V são semelhantes se, e somente se, houver uma matriz $A \in \mathscr{M}_{n,n}(\mathbb{K})$ que represente os dois operadores, mas com relação a bases ordenadas possivelmente diferentes.

Demonstração. Suponha primeiro que $[T]_{\underline{B}} = A = [S]_{\underline{C}}$, com $\underline{B} = (\vec{b}_1, \dots, \vec{b}_n)$ e $\underline{C} = (\vec{c}_1, \dots, \vec{c}_n)$. Seja $P : V \to V$ o automorfismo definido por $P(\vec{b}_i) = \vec{c}_i$. Como T e S têm as mesmas coordenadas nessas bases, vale $SP = PT$; portanto, $S = PTP^{-1}$.

Reciprocamente, suponha $S = PTP^{-1}$. Se $\underline{B} = (\vec{b}_1, \dots, \vec{b}_n)$ é uma base de V, então $P\underline{B} = (P\vec{b}_1, \dots, P\vec{b}_n)$ também é uma base. Da identidade $SP = PT$ segue diretamente

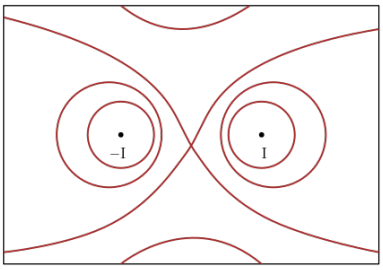


Figura 9.1 As classes de similaridade em $\mathscr{M}_{n,n}(\mathbb{K})$ particionam o espaço das matrizes em subconjuntos disjuntos $[A]$. Algumas classes são pontos, por exemplo $[-\mathrm{I}]$, $[\mathrm{I}]$ e $[0]$. Algumas classes são hiper-superfícies complicadas em $\mathscr{M}_{n,n}(\mathbb{K}) \cong \mathbb{K}^{n^2}$. Uma classe de similaridade pode ter várias componentes conexas.

que

$$[S]_{P\underline{B}} = [T]_{\underline{B}}.$$

■

Teorema 9.6.
Seja V um espaço vetorial de dimensão n. Então, duas matrizes A e $B \in \mathscr{M}_{n,n}(\mathbb{K})$ são semelhantes se, e somente se, representam o mesmo operador linear $T \in \mathrm{Hom}(V)$, mas possivelmente em relação a diferentes bases ordenadas.

Demonstração. É a versão matricial do teorema anterior e também segue diretamente da fórmula de mudança de base.

■

O restante deste capítulo inicia a busca de invariantes e representantes simples para as classes de similaridade.

Exercícios

Ex. 9.1 — Se $A \in \mathscr{M}_{n,n}(\mathbb{K})$ prove que

1. A matriz A comuta com todas as matrizes de $\mathscr{M}_{n,n}(\mathbb{K})$ se, e somente se, $A = \lambda\,\mathrm{I}$, um múltiplo escalar da matriz identidade, para algum $\lambda \in \mathbb{K}$.

2. A matriz A comuta com todas as matrizes invertíveis

   $$\mathrm{GL}(n,\mathbb{K}) = \{M \in \mathscr{M}_{n,n}(\mathbb{K}) \mid \det(M) \neq 0\}$$

   se, e somente se, A comuta com todas as matrizes $\mathscr{M}_{n,n}(\mathbb{K})$.

   Dica: Use as matrizes unitárias U_{ij}, que possuem uma entrada 1 na posição (i,j) e zeros nas demais. Se $i \neq j$, então $\mathrm{I} + U_{ij}$ é invertível.

3. Deduza que uma classe de similaridade $[A]$ em $\mathscr{M}_{n,n}(\mathbb{K})$ consiste em um único ponto se, e somente se, $A = \lambda\,\mathrm{I}$.

Ex. 9.2 — Identifique $\mathscr{M}_{2,2}(\mathbb{K})(\mathbb{R})$ com $\mathbb{R}^4$ pelo isomorfismo linear

$$\Psi(M) = (m_{11}, m_{12}, m_{21}, m_{22}).$$

Seja

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

. Mostre que sua classe de similaridade é

$$[J] = \left\{ \mathrm{I} + \begin{bmatrix} a & b \\ c & -a \end{bmatrix} \mid a^2 + bc = 0,\ (a,b,c) \neq (0,0,0) \right\}.$$

Conclua que, sob a identificação Ψ, essa classe é um cone quadrático bidimensional em um subespaço afim de $\mathbb{R}^4$, com o vértice removido.

Ex. 9.3 — Mostre que a classe de similaridade da matriz

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

não contém nenhuma matriz diagonal. Prove que não existe base $\underline{\mathrm{B}} = \{\vec{f}_1, \vec{f}_2\}$ para a qual $[T_N]^{\underline{\mathrm{B}}}_{\underline{\mathrm{B}}}$ seja diagonal.

9.3 Autovalores e Autovetores

O caso mais simples de subespaço invariante ocorre em dimensão um. Se $W = \langle \vec{w} \rangle$ é T-invariante e $\vec{w} \neq \vec{0}$, então $T\vec{w}$ pertence à mesma reta; portanto, existe $\lambda \in \mathbb{K}$ tal que

$$T(\vec{w}) = \lambda \vec{w} \tag{9.1}$$

Retas invariantes
Cada autovetor gera uma reta preservada por T. O autovalor descreve a ação do operador nessa reta e não depende do representante não nulo escolhido.

Um vetor $\vec{w}$ não nulo que satisfaz a Equação 9.1 é dito autovetor de T. Com efeito, $T(\vec{w}) = \lambda\vec{w}$ equivale a $(T - \lambda\,\mathrm{I})(\vec{w}) = 0$. Assim, os autovetores associados a λ são precisamente os vetores não nulos de $\ker(T - \lambda\,\mathrm{I})$.

Definição 9.7 — Autovalor e Autovetor. Sejam $T : V \to V$ uma transformação linear e $\lambda \in \mathbb{K}$.

- **a** Se $\ker(T - \lambda\,\mathrm{I}) \neq \{\vec{0}\}$, λ é dito **autovalor** de T.
- **b** Nesse caso, qualquer $\mathbf{v}$ não nulo em $\ker(T - \lambda\,\mathrm{I})$ é dito **autovetor** de T.
- **c** O subespaço $\ker(T - \lambda\,\mathrm{I})$ é o **autoespaço** associado a λ.
- **d** Nesse caso, dizemos que λ, $\mathbf{v}$ e o autoespaço $\ker(T - \lambda\,\mathrm{I})$ estão associados.

O conjunto de todos os autovalores de T é denominado **espectro** de T e denotado por $\operatorname{espec} T$. Escreveremos $E_\lambda = \ker(T - \lambda\,\mathrm{I})$ para o autoespaço associado a λ. As multiplicidades serão definidas separadamente mais adiante.

Definição 9.8 — Polinômio Característico. Para $A \in \mathscr{M}_{n,n}(\mathbb{K})$, o **polinômio característico** de A é

$$c_A(x) = \det(x\,\mathrm{I} - A).$$

Se V tem dimensão finita, $T : V \to V$ é linear, $\underline{\mathrm{B}}$ é uma base de V e $A = [T]_{\underline{\mathrm{B}}}$, definimos $c_T(x) = c_A(x)$.

Pela fórmula de Leibniz, $c_A(x)$ é mônico de grau n. Além disso, essa definição para operadores independe da base escolhida. De fato, se $B = PAP^{-1}$, então

$$\det(x\,\mathrm{I} - B) = \det\big(P(x\,\mathrm{I} - A)P^{-1}\big) = \det(x\,\mathrm{I} - A).$$

Exemplos 9.9.

1 Seja $T : V \to V$ que é a multiplicação por um escalar $T\mathbf{v} = \lambda\mathbf{v}$. Neste caso todo vetor não nulo em V é um autovetor associado ao autovalor λ.

2 Reflexão no plano xy. A reflexão no plano xy é a transformação R_{xy} que na base canônica é representada pela matriz:

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

Nesse caso $\operatorname{spec} R_{xy} = \{1, -1\}$. Os autovetores associados ao 1 são $\mathbf{e}_1, \mathbf{e}_2$ e associado ao -1 temos o autovetor $\mathbf{e}_3$.

3 Uma rotação em $\mathbb{R}^2$ pelo ângulo θ, denotada por R_θ, é representada na base canônica pela matriz

$$R_\theta = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

Seu polinômio característico é

$$c_{R_\theta}(x) = x^2 - 2(\cos\theta)x + 1.$$

Portanto, R_θ possui autovalor real se, e somente se, $\operatorname{sen}\theta = 0$, isto é, se $\theta \in \pi\mathbb{Z}$. Se θ é um múltiplo par de π, então $R_\theta = \mathrm{I}$ e o único autovalor é 1; se é um múltiplo ímpar, então $R_\theta = -\mathrm{I}$ e o único autovalor é -1. Nos demais casos não há autovalores reais. Sobre $\mathbb{C}$, os autovalores são $e^{i\theta}$ e $e^{-i\theta}$.

4 Se T numa base $\underline{\mathrm{B}} = \{\mathbf{v}_1, \dots \mathbf{v}_n\}$ possui representação matricial por uma matriz diagonal com os valores na diagonal principal distintos, isto é:

$$[T]_{\underline{\mathrm{B}}} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \text{ com } \lambda_i \neq \lambda_j \text{ se } i \neq j.$$

Então $\lambda_1, \dots \lambda_n$ são autovalores com autovetores associados $\{\mathbf{v}_1, \dots \mathbf{v}_n\}$ respectivamente.

5 Seja $C^\infty[a, b]$ o espaço vetorial das funções suaves em $[a, b]$. Seja $D : C^\infty[a, b] \to C^\infty[a, b]$ o operador derivação. Então a função $f(x)$ é um autovetor se satisfizer

$$Df(x) = \lambda f(x) \to f(x) = Ce^{\lambda x}.$$

No caso particular em que $\lambda = 0$ a função f é constante.

6 Seja $C^\infty[a, b]$ o espaço vetorial das funções suaves em $[a, b]$ e considere o operador de Volterra

$$I(f)(x) = \int_a^x f(u)\, du.$$

Esse operador não possui autovalores. De fato, se $I(f) = \lambda f$ e $\lambda \neq 0$, então $f(a) = 0$ e, derivando, obtemos $f' = \lambda^{-1} f$; a condição inicial força $f = 0$. Se $\lambda = 0$, então $I(f) = 0$ também implica $f = 0$ após derivação.

◁

Definição 9.10 — Autoespaço Generalizado. Seja $T : V \to V$ uma transformação linear e $\lambda \in \mathbb{K}$ seja um autovalor de T.

a Para $k \geq 1$, escrevemos

$$E_\lambda^k = \ker (T - \lambda \mathrm{I})^k.$$

O **autoespaço generalizado** de T associado a λ é

$$E_\lambda^\infty = \bigcup_{k \geq 1} E_\lambda^k = \{\mathbf{v} \in V \mid (T - \lambda \mathrm{I})^k(\mathbf{v}) = 0 \text{ para algum } k \geq 1\}.$$

b Se $\mathbf{v}$ for um vetor diferente de zero nesse autoespaço generalizado, então $\mathbf{v}$ é dito **autovetor generalizado** associado ao autovalor λ.

c Para cada um desses vetores $\mathbf{v}$, o menor número inteiro positivo k para o qual $(T - \lambda \mathrm{I})^k(\mathbf{v}) = 0$ é dito o **índice** de $\mathbf{v}$.

Exemplo 9.11. 1 Sejam $\lambda_1, ..., \lambda_n$ elementos distintos de $\mathbb{K}$ e $A \in \mathscr{M}_{n,n}(\mathbb{K})$ a matriz diagonal

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

Para cada $i = 1, ..., n$, λ_i é um autovalor de A com autoespaço unidimensional E_{λ_i} com base $\{\vec{e}_i\}$.

2 Seja λ um elemento de $\mathbb{K}$ e $A \in \mathscr{M}_{n,n}(\mathbb{K})$ dada por

$$A = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}$$

com entradas λ na diagonal e 1 imediatamente acima da diagonal e 0 em qualquer outra entrada. Para cada $k = 1, ..., n$, $\vec{e}_k$ é um autovetor generalizado do índice k, e o autoespaço generalizado E_λ^k é k-dimensional com base $\{\vec{e}_1, \ldots, \vec{e}_k\}$.

Nessa base a transformação associada a A tem a seguinte descrição:

$$e_k \mapsto \begin{cases} \lambda e_k & \text{se } k = 1 \\ \lambda e_k + e_{k-1} & \text{se } 1 < k \leq n. \end{cases}$$

◁

Em particular, um autovetor generalizado de índice 1 é um autovetor e

$$E_\lambda = E_\lambda^1 \subseteq E_\lambda^2 \subseteq \cdots \subseteq E_\lambda^\infty.$$

Exemplos 9.12. 1 Seja $\mathbb{K}$ um corpo de característica 0 e seja $V = \mathbb{K}[x]$, o espaço de todos os polinômios com coeficientes em $\mathbb{K}$. Seja $\mathrm{D} : V \to V$ o operador diferenciação, $\mathrm{D}(p(x)) = p'(x)$. Então D tem um único autovalor 0 e o autoespaço correspondente E_0 é 1-dimensional consistindo nos polinômios constantes. De maneira mais geral, E_0^k é k-dimensional, consistindo em todos os polinômios de grau no máximo $k - 1$.

2. Seja $V = \mathbb{K}[x]$ o espaço de todos os polinômios com coeficientes em um corpo de característica 0 e seja $T : V \to V$ definida por $T(p(x)) = xp'(x)$. Então, os autovalores de T são os números inteiros não negativos e, para todo número inteiro não negativo m, o autoespaço E_m é unidimensional com base $\{x^m\}$.

3. Seja V o espaço das funções $C^\infty(\mathbb{R})$, e seja $\mathrm{D} : V \to V$ o operador de diferenciação, $\mathrm{D}(f(x)) = f'(x)$. Para qualquer número real λ, E_λ é unidimensional com base $f(x) = e^{\lambda x}$. Além disso, E_λ^k é k-dimensional com base $\{e^{\lambda x},\ xe^{\lambda x},\ \ldots,\ x^{k-1}e^{\lambda x}\}$.

4. Seja $V = (\mathbb{K}^\infty)_0$ e seja $\mathrm{L} : V \to V$ o deslocamento para a esquerda. Então L tem um único autovalor, 0, e $E_0 = \{(a_1, a_2, \ldots) \in V \mid a_i = 0 \text{ para } i > 1\}$. Mais geralmente, $E_0^k = \{(a_1, a_2, \ldots) \in V \mid a_i = 0 \text{ para } i > k\}$, de modo que $\dim E_0^k = k$ e $V = E_0^\infty$. Por outro lado, o deslocamento para a direita $\mathrm{R} : V \to V$ não possui autovalores.

5. Seja $V = \mathbb{K}^\infty$ e seja $\mathrm{L} : V \to V$ o deslocamento para a esquerda. Para todo $\lambda \in \mathbb{K}$, o autoespaço E_λ é unidimensional, com base $\{(1, \lambda, \lambda^2, \ldots)\}$. Além disso, E_λ^k tem dimensão k para todo inteiro positivo k. Por outro lado, o deslocamento para a direita $\mathrm{R} : V \to V$ não possui autovalores.

◁

Teorema 9.13.
Seja V um espaço vetorial de dimensão finita e seja $T : V \to V$ linear. Então λ é autovalor de T se, e somente se, $c_T(\lambda) = 0$.

Demonstração. Seja $\underline{\mathrm{B}}$ uma base de V e faça $A = [T]_{\underline{\mathrm{B}}}$. Então, por definição, λ é um autovalor de T se, e somente se, houver um vetor diferente de zero $\mathbf{v}$ em $\ker(T - \lambda\,\mathrm{I})$, ou seja, se, e somente se, $(A - \lambda\,\mathrm{I})\mathbf{u} = 0$ para algum vetor diferente de zero $\mathbf{u}$ em $\mathbb{K}^n$ (onde $\mathbf{u} = [\mathbf{v}]_{\underline{\mathrm{B}}}$). Isso ocorre se, e somente se, $A - \lambda\,\mathrm{I}$ for singular, ou, equivalentemente, se $\det(A - \lambda\,\mathrm{I}) = 0$. Como $\det(A - \lambda\,\mathrm{I}) = (-1)^n \det(\lambda\,\mathrm{I} - A)$, obtemos $c_T(\lambda) = 0$. ■

A convenção $c_A(x) = \det(x\,\mathrm{I} - A)$ torna o polinômio característico mônico. Para calcular autoespaços, porém, costuma ser mais conveniente usar $A - \lambda\,\mathrm{I}$; os dois núcleos coincidem.

Teorema 9.14.
Sejam $A, B \in \mathscr{M}_{n,n}(\mathbb{K})$ duas matrizes semelhantes, ou seja,

$$A = PBP^{-1}.$$

Então

1. *A e B possuem os mesmos autovalores λ_i;*
2. *Os espaços $\ker(A - \lambda_i\,\mathrm{I})^j$ e $\ker(B - \lambda_i\,\mathrm{I})^j$ possuem a mesma dimensão para todo $j \in \mathbb{N}$ e todo autovalor λ_i.*

Demonstração. Como $A - \lambda \mathrm{I} = P(B - \lambda \mathrm{I})P^{-1}$, temos $(A - \lambda \mathrm{I})^j = P(B - \lambda \mathrm{I})^j P^{-1}$ para todo $j \geq 1$. Assim, P estabelece um isomorfismo entre os respectivos núcleos. A igualdade dos autovalores segue tomando $j = 1$. ■

Para o restante desta seção, assumimos que V é finito dimensional.

Definição 9.15 — Multiplicidade. Sejam $T : V \to V$ e λ um autovalor de T.

- A **multiplicidade algébrica** de λ, algmult(λ), é a multiplicidade de λ como raiz do polinômio característico $c_T(x)$.
- A **multiplicidade geométrica** de λ, geomult(λ), é a dimensão do autoespaço associado $E_\lambda = \ker(T - \lambda \mathrm{I})$.

Exemplo 9.16. Sejam

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \quad \text{e } B = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$$

então 3 é um autovalor de multiplicidade algébrica e geométrica 2 para A. Por outro lado, 3 é um autovalor de multiplicidade algébrica 2 e geométrica 1 para B.

◁

Quando não qualificado, usaremos multiplicidade para significar multiplicidade algébrica, como é o padrão na literatura.

Proposição 9.17.
Sejam $T : V \to V$ e λ seja um autovalor de T. Então $1 \leq \text{geomult}(\lambda) \leq \text{algmult}(\lambda)$.

Demonstração. Por definição, se λ é um autovalor de T, existe um autovetor $\mathbf{v}$ (diferente de zero), portanto, $1 \leq \dim(E_\lambda)$.

Suponha que $\dim(E_\lambda) = d$ e seja $\underline{\mathrm{B}}_1 = \{\mathbf{v}_1, \ldots, \mathbf{v}_d\}$ uma base para E_λ. Estenderemos $\underline{\mathrm{B}}_1$ para uma base, $\underline{\mathrm{B}} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ de V. Então

$$[T]_{\underline{\mathrm{B}}} = \begin{bmatrix} \lambda \mathrm{I} & B \\ 0 & D \end{bmatrix} = A,$$

uma matriz de bloco com o bloco superior esquerdo de tamanho $d \times d$. Então

$$[x\mathrm{I} - T]_{\underline{\mathrm{B}}} = x\mathrm{I} - A = \begin{bmatrix} x\mathrm{I} - \lambda \mathrm{I} & -B \\ 0 & x\mathrm{I} - D \end{bmatrix} = \begin{bmatrix} (x-\lambda)\mathrm{I} & -B \\ 0 & x\mathrm{I} - D \end{bmatrix} \text{ logo}$$

$$c_T(x) = \det(x\mathrm{I} - A) = \det((x-\lambda)\mathrm{I}) \det(x\mathrm{I} - D)$$

$$= (x-\lambda)^d \det(x\mathrm{I} - D)$$

e, portanto, $d \leq \text{algmult}(\lambda)$. ■

Corolário 9.18.
Sejam $T : V \to V$ e λ um autovalor de T com $\text{algmult}(\lambda) = 1$. *Então* $\text{geomult}(\lambda) = 1$.

A existência de autovalores e autovetores depende do corpo $\mathbb{K}$, como mostra o próximo exemplo.

Exemplo 9.19. Para qualquer número racional diferente de zero t, seja A_t a matriz

$$A_t = \begin{bmatrix} 0 & 1 \\ t & 0 \end{bmatrix}$$

dessa forma temos:

$$A_t^2 = \begin{bmatrix} t & 0 \\ 0 & t \end{bmatrix} = t\,\mathrm{I}\,.$$

Seja λ um autovalor de A_t com o autovetor associado v. Então, por um lado,

$$A_t^2(\mathbf{v}) = A_t(A_t(\mathbf{v})) = A_t(\lambda \mathbf{v}) = \lambda A_t(\mathbf{v}) = \lambda^2 v,$$

mas por outro lado,

$$A_t^2(\mathbf{v}) = t\,\mathrm{I}(\mathbf{v}) = t\mathbf{v},$$

então $\lambda^2 = t$.

a Suponha $t = 1$. Então $\lambda^2 = 1, \lambda = \pm 1$, e temos o autovalor $\lambda = 1$ com o autovetor associado $v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ e o autovalor $\lambda = -1$ com o autovetor associado $v = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

b Suponha $t = 2$. Se considerarmos A_t sobre $\mathbb{Q}$, não haverá $\lambda \in \mathbb{Q}$ com $\lambda^2 = 2$, portanto A_t não terá autovalores. Sobre $\mathbb{R}$, $\lambda = \sqrt{2}$ é um autovalor com o autovetor associado $\begin{bmatrix} 1 \\ \sqrt{2} \end{bmatrix}$ e $\lambda = -\sqrt{2}$ é um autovalor com o autovetor associado $\begin{bmatrix} 1 \\ -\sqrt{2} \end{bmatrix}$

◁

Os autovalores descrevem os fatores lineares associados ao operador. Para registrar todas as relações polinomiais satisfeitas por T, introduzimos agora o polinômio minimal.

9.3.1 Ação dos Polinômios em $\mathrm{Hom}(V, V)$

Substituir x por T
Um polinômio $p(x)$ pode agir sobre um operador pela substituição $x \mapsto T$. A identidade $p(T) = 0$ traduz informação algébrica do polinômio em informação estrutural sobre o operador. O polinômio minimal é a relação mônica de menor grau desse tipo.

Se $p(x) = \sum_{i=0}^{d} a_i x^i \in \mathbb{K}[x]$ e $T \in \mathrm{Hom}(V, V)$, definimos

$$p(T) = \sum_{i=0}^{d} a_i T^i,$$

onde $T^0 = \mathrm{I}$ e T^i é a i-ésima potência de T por composição. Essa construção é a avaliação de um polinômio no operador T.

Assim como procuramos escalares nos quais um polinômio se anula, procuraremos agora operadores anulados por polinômios.

Definição 9.20. Dizemos que um polinômio $p(x)$ **anula** uma transformação T se $p(T) = 0$.

Lema 9.21.
Se V é um espaço vetorial de dimensão n e $T \in \mathrm{Hom}(V,V)$ então existe um polinômio $g(x) \in \mathbb{K}[x]$ tal que $g(T) = 0$.

Demonstração. Seja $C = \left\{\mathrm{I}, T, T^2, \dots T^{n^2}\right\}$. Como $\dim \mathrm{Hom}(V,V) = n^2$, temos que C é um conjunto linearmente dependente, logo existem escalares $a_i \in \mathbb{K}$, não todos nulos, tais que

$$a_0 \mathrm{I} + a_1 T + \cdots + a_{n^2} T^{n^2} = 0$$

Logo $g(x) = \sum_{i=0}^{n^2} a_i x^i$ é um polinômio não nulo que anula T. ■

Corolário 9.22.
Se $V \neq \{0\}$ é finito dimensional sobre um corpo algebricamente fechado, qualquer operador linear $T : V \to V$ possui um autovetor.

Demonstração. Pelo lema anterior, existe um polinômio não nulo g tal que $g(T) = 0$. Como o corpo é algebricamente fechado, $g(x) = a\prod_{i=1}^{r}(x - \lambda_i)$. Se todos os operadores $T - \lambda_i \mathrm{I}$ fossem injetivos, sua composição também seria injetiva, o que contradiz $g(T) = 0$ em $V \neq \{0\}$. Logo algum $\ker(T - \lambda_i \mathrm{I})$ é não trivial. ■

O **aniquilador** de T é o conjunto dos polinômios que anulam T:

$$\mathrm{Aniq}(T) = \{p(x) \mid p(T) = 0\}$$

Esse conjunto é um ideal não nulo de $\mathbb{K}[x]$. Como $\mathbb{K}[x]$ é um domínio de ideais principais, existe um único polinômio mônico $m_T(x)$ tal que

$$\mathrm{Aniq}(T) = \langle m_T(x)\rangle = \{a(x)m_T(x) \mid a(x) \in \mathbb{K}[x]\}.$$

Definição 9.23 — Polinômio Minimal. O gerador $m_T(x)$ do ideal $\mathrm{Aniq}(T)$ é dito **polinômio minimal** de T.

O polinômio minimal também pode ser caracterizado como o único polinômio mônico $m_T(x)$ de menor grau com $m_T(T) = 0$. Além disso, $m_T(x)$ divide todo polinômio $p(x)$ tal que $p(T) = 0$.

Lema 9.24.
Matrizes semelhantes têm o mesmo polinômio minimal.

Demonstração. Se $B = PAP^{-1}$, então $p(B) = Pp(A)P^{-1}$ para todo $p \in \mathbb{K}[x]$. Portanto, $p(A) = 0$ se, e somente se, $p(B) = 0$; os dois conjuntos de polinômios anuladores coincidem e, assim, seus geradores mônicos m_A e m_B são iguais. ■

Triangularização e autovalores
Uma matriz triangular tem seus autovalores na diagonal, mesmo quando não é diagonalizável. Assim, a triangularização fornece uma representação na qual os autovalores podem ser identificados sem uma base completa de autovetores.

9.4 Teoremas de Schur e Cayley-Hamilton

A triangularização ocupa uma posição intermediária entre uma matriz arbitrária e uma matriz diagonal. Ela já exibe os autovalores na diagonal e conduz ao Teorema de Cayley–Hamilton, sem exigir uma base completa de autovetores.

Definição 9.25 — Triangularizável. Seja V um espaço vetorial de dimensão finita e seja $T : V \to V$ linear.

1. T é **triangularizável superior** se houver uma base, $\underline{\mathrm{B}}$ de V na qual a matriz $A = [T]_{\underline{\mathrm{B}}}$ é triangular superior, i.e., $A(i,j) = 0$ para todos os $i > j$.
2. T é **triangularizável estritamente superior** se houver uma base, $\underline{\mathrm{B}}$ de V na qual a matriz $A = [T]_{\underline{\mathrm{B}}}$ é triangular estritamente superior, i.e., $A(i,j) = 0$ para todos os $i \geq j$.

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & 0 & & \times & \times \\ & & & & \times \end{bmatrix} \qquad \begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ 0 & & & \times \end{bmatrix} \tag{9.2}$$

Fixada uma base $\underline{\mathrm{B}} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ de V, ponha

$$V_0 = \{\vec{0}\}, \qquad V_i = \langle \mathbf{v}_1, \ldots, \mathbf{v}_i \rangle, \quad 1 \leq i \leq n.$$

A cadeia $V_0 \subset V_1 \subset \cdots \subset V_n = V$ é a **bandeira associada** à base $\underline{\mathrm{B}}$.

Definição 9.26. Dizemos que um mapa linear $T \in \mathrm{Hom}(V,V)$ **estabiliza a bandeira** se $T(V_i) \subseteq V_i$ para $0 \leq i \leq n$. Dizemos que T **estabiliza estritamente a bandeira** se $T(V_i) \subseteq V_{i-1}$ para $1 \leq i \leq n$.

Proposição 9.27.
Suponha $T \in \mathrm{Hom}(V,V)$ e $\underline{B} = \{\vec{v}_1, \ldots \vec{v}_n\}$ base de V. Então são equivalentes:

1. *A matriz de T na base $\underline{B}$ é triangular superior; (estritamente superior)*
2. *$T\vec{v}_k \in \langle \vec{v}_1, \ldots, \vec{v}_k \rangle$ para todo $k = 1, \ldots n$ ($T(\vec{v}_1) = 0$, $T\vec{v}_k \in \langle \vec{v}_1, \ldots, \vec{v}_{k-1} \rangle$ para todo $k = 2, \ldots n$;)*
3. *T estabiliza (estritamente) a bandeira associada a $\underline{B}$.*

Demonstração. A coluna k de $[T]_{\underline{\mathrm{B}}}$ é o vetor de coordenadas de $T\mathbf{v}_k$. Portanto, as entradas abaixo da diagonal nessa coluna são nulas se, e somente se, $T\mathbf{v}_k \in V_k$. Isso prova a equivalência no caso triangular superior. No caso estritamente triangular, também é preciso anular a entrada diagonal, o que equivale a $T\mathbf{v}_k \in V_{k-1}$. Como os vetores $\mathbf{v}_1, \ldots, \mathbf{v}_i$ geram V_i, essas condições são equivalentes, respectivamente, a $T(V_i) \subseteq V_i$ e a $T(V_i) \subseteq V_{i-1}$. ∎

Teorema 9.28 (de Schur).
Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{K}$ e $T : V \to V$ uma transformação linear. Então T é triangularizável se, e somente se, o seu polinômio característico $c_T(x)$ for um produto de fatores lineares.
Em particular, se $\mathbb{K}$ for fechado algebricamente, então todo operador linear $T : V \to V$ é triangularizável.

Demonstração. Se $[T]_{\underline{B}}$ é triangular superior, então seu polinômio característico é o produto dos termos $x - d_i$, onde d_i são as entradas da diagonal.

Reciprocamente, suponha que c_T se fatore em termos lineares. Procedemos por indução em $n = \dim V$. O caso $n = 1$ é imediato. Escolha um autovalor λ e um autovetor $\mathbf{v}_1$, e ponha $V_1 = \langle \mathbf{v}_1 \rangle$. Como V_1 é T-invariante, T induz um operador $\overline{T}$ em V/V_1. Em uma base adaptada a V_1, a matriz de T tem a forma

$$\begin{bmatrix} \lambda & * \\ 0 & B \end{bmatrix},$$

logo $c_T(x) = (x - \lambda)c_{\overline{T}}(x)$. Portanto, $c_{\overline{T}}$ também se fatora em termos lineares. Pela hipótese de indução, existe uma base do quociente na qual $\overline{T}$ é triangular superior. Levantando essa base a V e acrescentando $\mathbf{v}_1$, obtemos uma base na qual T é triangular superior. ■

Proposição 9.29.
Seja $\mathbf{v}$ um autovetor de T com autovalor associado λ e seja $p(x) \in \mathbb{K}[x]$. Então $p(T)(\mathbf{v}) = p(\lambda)\mathbf{v}$. Assim, se $p(\lambda) \neq 0$, então $p(T)(\mathbf{v}) \neq \vec{0}$.

Demonstração. Observamos inicialmente que podemos fatorar qualquer polinômio $p(x) \in \mathbb{K}[x]$ em termos de $x - \lambda$:

$$p(x) = a_n(x - \lambda)^n + a_{n-1}(x - \lambda)^{n-1} + \cdots + a_1(x - \lambda) + a_0.$$

Substituindo $x = \lambda$, vemos que $a_0 = p(\lambda)$.

Se $\mathbf{v}$ for um autovetor de T com o autovalor associado λ,

$$\begin{aligned} p(T)(\mathbf{v}) &= (a_n(T - \lambda\,\mathrm{I})^n + \cdots + a_1(T - \lambda\,\mathrm{I}) + p(\lambda)\,\mathrm{I})(\mathbf{v}) \\ &= p(\lambda)\,\mathrm{I}(\mathbf{v}) = p(\lambda)\mathbf{v} \end{aligned}$$

como todos os termos, exceto o último se anulam. ■

9.4.1 Teorema de Cayley-Hamilton

Teorema 9.30 (de Cayley-Hamilton).
Seja V um espaço vetorial de dimensão finita e seja $T : V \to V$ uma transformação linear. Então

$$c_T(T) = 0.$$

Demonstração. Fixe uma base de V e seja A a matriz de T nessa base. Escreva

$$c_A(x) = a_0 + a_1x + \cdots + a_nx^n \quad \text{e} \quad B(x) = \operatorname{adj}(x\,\mathrm{I} - A) = B_0 + B_1x + \cdots + B_{n-1}x^{n-1}.$$

Da identidade $(x\,\mathrm{I} - A)B(x) = c_A(x)\,\mathrm{I}$, a comparação dos coeficientes fornece

$$-AB_0 = a_0\,\mathrm{I}, \qquad B_{k-1} - AB_k = a_k\,\mathrm{I}\ (1 \le k \le n-1), \qquad B_{n-1} = a_n\,\mathrm{I}.$$

Multiplicando essas identidades, respectivamente, por $\mathrm{I}, A, \ldots, A^n$ e somando, todos os termos que envolvem os B_k se cancelam. Resta

$$a_0\,\mathrm{I} + a_1A + \cdots + a_nA^n = 0,$$

isto é, $c_A(A) = 0$. Portanto, $c_T(T) = 0$. ■

Corolário 9.31.
Sejam V um espaço de dimensão finita e $T \in \operatorname{Hom}(V,V)$. Então o polinômio minimal $m_T(x)$ divide o polinômio característico $c_T(x)$.

Teorema 9.32.
Sejam V um espaço de dimensão n e $T \in \operatorname{Hom}(V,V)$. Então o polinômio característico $c_T(x)$ divide a n potência do polinômio minimal: $(m_T(x))^n$.

Demonstração. Para demonstrar esse fato, fixamos uma base de V e seja M a matriz de T nessa base. Nesse contexto, $c_T = \det(x\,\mathrm{I} - M)$.

Escreva $m_T = \sum_{k=0}^{d} a_kx^k$. Então

$$\begin{aligned}
m_T(x\,\mathrm{I}) &= m_T(x\,\mathrm{I}) - m_T(M) && (9.3)\\
&= \sum_{k=0}^{d} a_k(x^k\,\mathrm{I} - M^k) && (9.4)\\
&= \sum_{k=1}^{d} a_k(x^k\,\mathrm{I} - M^k) && (9.5)\\
&= (x\,\mathrm{I} - M)\sum_{k=1}^{d} a_k \sum_{p=0}^{k-1} x^pM^{k-1-p} && (9.6)\\
&= (x\,\mathrm{I} - M)B && (9.7)
\end{aligned}$$

Onde $B := \sum_{k=1}^{d} a_k \sum_{p=0}^{k-1} x^pM^{k-1-p}$.

Em seguida, calculamos o determinante dessas matrizes em $\mathbb{K}[x]$; a validade das identidades usuais nesse anel foi observada em 8.31: $(m_T(x))^n = (m_T(x))^n \det(\mathrm{I}) = \det(m_T(x)\,\mathrm{I}) = \det((x\,\mathrm{I} - M)B) = \det(x\,\mathrm{I} - M)\det(B) = c_T(x)\det(B)$, e o resultado segue. ■

Como consequência do Corolário 9.31 e do Teorema 9.32, temos:

Corolário 9.33.
Seja $T : V \to V$ uma aplicação linear e $c_T(x) \in \mathbb{K}[x]$ seu polinômio característico. Se

$$c_T(x) = [p_1(x)]^{s_1} \cdots [p_j(x)]^{s_j}$$

é a decomposição de $c_T(x)$ em fatores irredutíveis, com $p_i \neq p_k$ para $i \neq k$, então, o polinômio minimal de T é

$$m_T(x) = [p_1(x)]^{d_1} \cdots [p_j(x)]^{d_j},$$

em que $0 < d_i \leq s_i$ para $i = 1, \ldots, j$. Em outras palavras, o polinômio minimal possui todos os fatores irredutíveis do polinômio característico de T.

Definição 9.34 — Autoespaço Generalizado Associado ao Polinômio. Sejam V um espaço vetorial de dimensão finita e $T : V \to V$ uma transformação linear. Seja $c_T \in \mathbb{K}[t]$ o polinômio característico de T. Se

$$c_T(t) = [p_1(t)]^{s_1} \ldots [p_j(t)]^{s_j}$$

é a decomposição de c_T em fatores irredutíveis, com $p_i \neq p_k$ para $i \neq k$. Definimos, o **autoespaço generalizado associado ao polinômio** p_i como o conjunto de todos os vetores $\mathbf{v} \in V$ para os quais existe um inteiro positivo k tal que

$$(p_i(T))^k \mathbf{v} = \vec{0}.$$

Se V for de dimensão finita a cadeia

$$\vec{0} \subset \ker(p_i(T))^1 \subset \ker(p_i(T))^2 \subset \cdots \ker(p_i(T))^k \cdots$$

estabiliza. O menor inteiro positivo r_i tal que $\ker(p_i(T))^{r_i} = \ker(p_i(T))^{r_i+1}$ é dito o **índice** de $p_i(T)$.

9.5 Teorema da Decomposição Primária

Separar fatores coprimos
Quando o polinômio minimal se decompõe em fatores coprimos, o espaço se separa em componentes invariantes correspondentes. Cada componente concentra um único fator primário; podemos estudar o operador por partes e depois recompor o resultado.

Se $p(x)$ e $q(x)$ são polinômios coprimos tais que $p(T)q(T) = 0$, então $V = \ker p(T) \oplus \ker q(T)$.

Proposição 9.35.
Sejam $p, q \in \mathbb{K}[x]$ coprimos e $T \in \mathrm{Hom}(V, V)$. Sejam N_p, N_q e N_{pq} os núcleos dos operadores $p(T), q(T)$ e $p(T)q(T)$, respectivamente. Então

$$N_{pq} = N_p \oplus N_q \tag{9.8}$$

Demonstração. Pela identidade de Bézout, existem polinômios $a, b \in \mathbb{K}[x]$ tais que $a(x)p(x) + b(x)q(x) = 1$. Assim,

$$a(T)p(T) + b(T)q(T) = \mathrm{I}.$$

Se $\mathbf{v} \in N_{pq}$, então $b(T)q(T)\mathbf{v} \in N_p$. De fato, aplicando $p(T)$ a esse vetor, temos $p(T)b(T)q(T)\mathbf{v} = b(T)p(T)q(T)\mathbf{v} = 0$. Da mesma forma temos $a(T)p(T)\mathbf{v} \in N_q$, se

$\mathbf{v} \in N_{pq}$. Como $b(T)q(T)\mathbf{v} + a(T)p(T)\mathbf{v} = \mathbf{v}$, mostramos que $\mathbf{v} = \mathbf{v}_p + \mathbf{v}_q$, com $\mathbf{v}_p \in N_p$ e $\mathbf{v}_q \in N_q$.

Para mostrar que essa decomposição é única, suponhamos que $\mathbf{v} = \mathbf{v}_p + \mathbf{v}_q = \mathbf{v}'_p + \mathbf{v}'_q$. Mas então $\mathbf{w} := \mathbf{v}_p - \mathbf{v}'_p = \mathbf{v}'_q - \mathbf{v}_q$ pertence, simultaneamente, a N_p e N_q. Aplicando $b(T)q(T) + a(T)p(T) = \mathrm{I}$ em $\mathbf{w}$, temos

$$b(T)q(T)\mathbf{w} + a(T)p(T)\mathbf{w} = \mathbf{w}.$$

Mas $b(T)q(T)\mathbf{w} = 0 = a(T)p(T)\mathbf{w}$, de modo que $\mathbf{w} = 0$, o que implica $\mathbf{v}_p = \mathbf{v}'_p$ e $\mathbf{v}_q = \mathbf{v}'_q$, mostrando a unicidade da decomposição.

■

O argumento de Bézout se estende imediatamente a qualquer família finita de polinômios dois a dois coprimos.

Corolário 9.36.
Seja $T \in \mathrm{Hom}(V,V)$. Se $p_1, p_2, \ldots, p_k \in \mathbb{K}[x]$ são dois a dois coprimos, se N_{p_i} denota o núcleo de $p_i(T)$ e $N_{p_1 \ldots p_k}$ o núcleo de $p_1(T) \ldots p_k(T)$, então

$$N_{p_1 \ldots p_k} = N_{p_1} \oplus \cdots \oplus N_{p_k}.$$

Um caso particularmente interessante ocorre quando aplicamos o corolário anterior a uma fatoração em termos irredutíveis do polinômio minimal.

Teorema 9.37 (Decomposição Primária).
Sejam $T : V \to V$ uma aplicação linear e $m_T(x) \in \mathbb{K}[x]$ seu polinômio minimal. Se

$$m_T(x) = [p_1(x)]^{d_1} \cdots [p_j(x)]^{d_j}$$

é uma decomposição desse polinômio em fatores irredutíveis, com $p_i \neq p_k$ para $i \neq k$, então

$$V = W_1 \oplus \cdots \oplus W_j,$$

onde $W_i = \ker\left(p_i(T)\right)^{d_i}$, são subespaços T-invariantes e o polinômio minimal de $T{\restriction}_{W_i}$ é $p_i(x)^{d_i}$.

Demonstração. Como $m_T(T) = 0$ e como os polinômios na fatoração $m_1(x) = [p_1(x)]^{d_1}, \ldots, m_j(x) = [p_j(x)]^{d_j}$ são coprimos, podemos aplicar o corolário 9.36 e concluir que

$$V = N_{m_1 \cdots m_j} = W_1 \oplus \cdots \oplus W_j. \tag{9.9}$$

Como $W_i = \ker\left(p_i(T)\right)^{d_i}$, temos que o polinômio minimal de $T{\restriction}_{W_i}$ é da forma $p_i(x)^{r_i}$ com $r_i \leq d_i$. Se $r_i < d_i$, então $[p_1(x)]^{d_1} \cdots [p_i(x)]^{r_i} \cdots [p_j(x)]^{d_j}$ seria um polinômio de menor grau que anula T, contradizendo a minimalidade de $m_T(x)$. Em particular, o expoente d_i é o índice de $p_i(T)$: na parcela W_i a potência necessária é exatamente d_i, enquanto nas demais parcelas $p_i(T)$ é invertível pela identidade de Bézout. ■

Quando o polinômio característico se fatora em termos lineares, a decomposição primária separa o espaço segundo os autovalores.

Teorema 9.38 (Decomposição em Autoespaços Generalizados).
Suponha que o polinômio característico de T se fatore como produto de termos lineares sobre $\mathbb{K}$ e sejam $\lambda_1, \dots, \lambda_m$ os autovalores distintos de T. Então

$$V = \bigoplus_{j=1}^{m} E^{\infty}_{\lambda_j}(T).$$

Demonstração. Como o polinômio característico se fatora em termos lineares, o Corolário 9.33 dá

$$m_T(x) = \prod_{j=1}^{m} (x - \lambda_j)^{d_j}.$$

Os fatores são dois a dois coprimos. Aplicando o Teorema da Decomposição Primária, obtemos

$$V = \bigoplus_{j=1}^{m} \ker (T - \lambda_j \mathrm{I})^{d_j}.$$

Em dimensão finita, a cadeia dos núcleos de $(T - \lambda_j \mathrm{I})^k$ estabiliza, e sua união é $E^{\infty}_{\lambda_j}(T)$. Logo cada parcela acima coincide com o autoespaço generalizado correspondente. ■

Exemplo 9.39. Considere o operador T cuja matriz na base canônica $\underline{\mathrm{B}}$ é

$$[T]_{\underline{\mathrm{B}}} = \begin{bmatrix} -4 & 5 & 7 & -4 \\ -3 & 4 & 4 & -2 \\ -3 & 3 & 5 & -2 \\ -1 & 1 & 1 & 1 \end{bmatrix}$$

Seu polinômio característico é $c_T(x) = (x-2)^2(x-1)^2$; portanto, os autovalores são 2 e 1. Para o primeiro deles,

$$\ker (T - 2\,\mathrm{I})^2 = \{[x, y, z, w] \mid w = -x + 2y,\ z = y\},$$

e $\{[1, 0, 0, -1], [1, 1, 1, 1]\}$ é uma base desse autoespaço generalizado. Já os vetores $[2, 0, 2, 1]$ e $[1, 1, 0, 0]$ formam uma base de $\ker (T - \mathrm{I})$. Concluímos ainda que $m_T(x) = (x-2)^2(x-1)$.

Considere a base

$$\underline{\mathrm{E}} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4) = ([1, 0, 0, -1], [1, 1, 1, 1], [2, 0, 2, 1], [1, 1, 0, 0])$$

Calculando a imagem de cada vetor, obtemos

$$\begin{aligned} T\mathbf{e}_1 &= \mathbf{e}_1 - \mathbf{e}_2 \\ T\mathbf{e}_2 &= \mathbf{e}_1 + 3\mathbf{e}_2 \\ T\mathbf{e}_3 &= \mathbf{e}_3 \\ T\mathbf{e}_4 &= \mathbf{e}_4 \end{aligned}$$

Portanto, nessa base,

$$[T]_{\underline{\mathrm{E}}} = \left(\begin{array}{cc|cc} 1 & 1 & 0 & 0 \\ -1 & 3 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

Essa matriz em blocos reflete a decomposição primária. A escolha das bases nos autoespaços generalizados ainda pode simplificar cada bloco; esse será o ponto de partida da Forma de Jordan, no próximo capítulo.

◁

Exemplo 9.40. Seja $\mathscr{F} = C^\infty(\mathbb{R}, \mathbb{C})$ o espaço das funções infinitamente diferenciáveis e considere o operador linear

$$E[f] \coloneqq (D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0)f$$

onde D é a diferenciação. A equação diferencial associada é

$$(D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0)f = 0$$

com coeficientes constantes complexos. Resolver a equação é determinar o núcleo de E, que denotaremos por W.

A teoria das equações diferenciais garante que $\dim W = n$. Considere o polinômio

$$m = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0.$$

Em $\mathbb{C}$, esse polinômio fatora como

$$m = (x - \lambda_1)^{r_1}(x - \lambda_2)^{r_2} \cdots (x - \lambda_k)^{r_k}.$$

O espaço $W = \ker m(D)$ é invariante por D, e o polinômio minimal de $D|_W$ é m. Pelo Teorema 9.38, W se decompõe como soma direta dos espaços de solução W_i das equações diferenciais

$$W_i = \{f \mid (D - \lambda_i \mathrm{I})^{r_i} f = 0\}.$$

Para determinar as soluções de $(D - \lambda_i \mathrm{I})^{r_i} f = 0$, usamos a identidade, obtida por indução,

$$(D - \lambda_i \mathrm{I})^{r_i} f = e^{\lambda_i t} D^{r_i}(e^{-\lambda_i t} f)\,.$$

Portanto, f é uma solução se, e somente se, $D^{r_i}(e^{-\lambda_i t} f) = 0$, isto é, se, e somente se, $e^{-\lambda_i t} f$ é um polinômio de grau no máximo $r_i - 1$. Uma base para o espaço de solução de $(D - \lambda_i \mathrm{I})^{r_i} f = 0$ é então $\{e^{\lambda_i t},\ te^{\lambda_i t},\ \ldots,\ t^{r_i - 1}e^{\lambda_i t}\}$.

◁

Exercícios

Ex. 9.4 — Seja $\{T_i : i \in I\}$ um subconjunto de $\mathrm{Hom}(V, V)$ onde V é um e.v. de dimensão finita sobre um corpo algebricamente fechado $\mathbb{K}$. Suponha que $T_i T_j = T_j T_i$ para todo $i, j \in I$. Mostre que V pode ser escrito como soma direta de autoespaços generalizados comuns a todos os $T_i, i \in I$.

Ex. 9.5 — Considere o operador $T : \mathbb{R}^3 \to \mathbb{R}^3$ cuja matriz na base canônica é

$$\begin{bmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{bmatrix}.$$

Ache a decomposição primária de $\mathbb{R}^3$ e encontre bases para cada um desses subespaços T-invariantes.

9.6 Diagonalizabilidade

A decomposição em autoespaços generalizados permite agora reconhecer quando os autoespaços ordinários já ocupam todo o espaço.

Dimensão dos autoespaços
Diagonalizar significa encontrar uma base formada por autovetores. Ter todos os autovalores no corpo é necessário, mas não suficiente: a soma das dimensões dos autoespaços deve alcançar a dimensão inteira do espaço.

Lema 9.41.
Autoespaços associados a autovalores distintos estão em soma direta.

Demonstração. Sejam $\lambda_1, \dots, \lambda_m$ distintos e suponha que $\mathbf{v}_i \in E_{\lambda_i}$ satisfaça $\mathbf{v}_1 + \cdots + \mathbf{v}_m = \vec{0}$. Fixe i e aplique $\prod_{j \neq i}(T - \lambda_j \mathrm{I})$ à igualdade. Todos os termos, exceto o i-ésimo, se anulam, e obtemos

$$\left(\prod_{j \neq i}(\lambda_i - \lambda_j)\right)\mathbf{v}_i = \vec{0}.$$

O escalar entre parênteses é não nulo; portanto, $\mathbf{v}_i = \vec{0}$. Isso vale para todo i. ■

Definição 9.42 — Diagonalizável.

a Um operador $T : V \to V$ em um espaço de dimensão finita é **diagonalizável** se existe uma base $\underline{\mathrm{B}}$ de V na qual $[T]_{\underline{\mathrm{B}}}$ é diagonal.

b Uma matriz $A \in \mathscr{M}_{n,n}(\mathbb{K})$ é **diagonalizável** se $T_A : \mathbb{K}^n \to \mathbb{K}^n$ é diagonalizável.

Ou seja, a matriz A é diagonalizável se for semelhante a uma matriz diagonal.

Proposição 9.43.
Seja V um espaço vetorial de dimensão finita e $T : V \to V$ seja uma transformação linear. Então T é diagonalizável se, e somente se, V possuir uma base, $\underline{\mathrm{B}}$ consistindo de autovetores de T.

Demonstração. Seja $\underline{\mathrm{B}} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ uma base e seja $D = [T]_{\underline{\mathrm{B}}}$ uma matriz diagonal com entradas diagonais $\mu_1, \dots, \mu_n$. Para cada i,

$$[T(\mathbf{v}_i)]_{\underline{\mathrm{B}}} = [T]_{\underline{\mathrm{B}}}[\mathbf{v}_i]_{\underline{\mathrm{B}}} = D\mathbf{e}_i = \mu_i \mathbf{e}_i = \mu_i [\mathbf{v}_i]_{\underline{\mathrm{B}}},$$

então $T(\mathbf{v}_i) = \mu_i \mathbf{v}_i$ e $\mathbf{v}_i$ é um autovetor.

Por outro lado, se $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ é uma base de autovetores, então $T(\mathbf{v}_i) = \mu_i \mathbf{v}_i$ para todo i, e assim

$$\begin{aligned}
[T]_{\underline{\mathrm{B}}} &= \left[[T(\mathbf{v}_1)]_{\underline{\mathrm{B}}} \mid [T(\mathbf{v}_2)]_{\underline{\mathrm{B}}} \mid \cdots \mid [T(\mathbf{v}_n)]_{\underline{\mathrm{B}}}\right] \\
&= \left[[\mu_1 \mathbf{v}_1]_{\underline{\mathrm{B}}} \mid [\mu_2 \mathbf{v}_2]_{\underline{\mathrm{B}}} \mid \cdots \mid [\mu_n \mathbf{v}_n]_{\underline{\mathrm{B}}}\right] \\
&= \left[\mu_1 \mathbf{e}_1 \mid \mu_2 \mathbf{e}_2 \mid \cdots \mid \mu_n \mathbf{e}_n\right] \quad = D
\end{aligned}$$

é uma matriz diagonal.

■

Teorema 9.44.
Sejam V um espaço vetorial de dimensão finita e $T : V \to V$ uma transformação linear. Se $c_T(x)$ não se decompor como um produto de fatores lineares sobre $\mathbb{K}$, então T não será diagonalizável. Se $c_T(x)$ se decompor como um produto de fatores lineares (o que é sempre o caso se $\mathbb{K}$ for algebricamente fechado), então as seguintes afirmações são equivalentes:

a *T é diagonalizável.*

b *$m_T(x)$ se fatora em um produto de termos lineares distintos.*

c *Se $\lambda_1, \ldots, \lambda_m$ são os autovalores distintos de T, então*

$$V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_m}.$$

d *A soma das multiplicidades geométricas dos autovalores é igual à dimensão de V.*

e *Para cada autovalor λ de T, $\text{geomult}(\lambda) = \text{algmult}(\lambda)$.*

f *Para cada autovalor λ de T, $E_\lambda = E_\lambda^\infty$ (ou seja, todo autovetor generalizado de T é um autovetor de T).*

Demonstração.

a implica **b**

Suponha que T é diagonalizável. Isso significa que tem uma base de autovetores, cujos autovalores distintos são $\lambda_1, \ldots, \lambda_k$. Nesse caso,

$$(x - \lambda_1) \cdots (x - \lambda_k)$$

é um polinômio aniquilador para T. Portanto, o polinômio minimal divide esse produto e também possui apenas fatores lineares distintos.

b implica **a**

Se $m_T(x)$ se fatora em um produto de termos lineares distintos então a composta

$$V \xrightarrow{T-\lambda_k \mathrm{I}} V \xrightarrow{T-\lambda_{k-1} \mathrm{I}} \cdots \xrightarrow{T-\lambda_1 \mathrm{I}} V$$

é 0. Logo

$$\begin{aligned} \dim(V) &= \dim\ker((T - \lambda_1 \mathrm{I}) \circ \cdots \circ (T - \lambda_k \mathrm{I})) && (9.10)\\ &\leq \dim\ker(T - \lambda_1 \mathrm{I}) + \cdots + \dim\ker(T - \lambda_k \mathrm{I}) && (9.11)\\ &= \dim(\ker(T - \lambda_1 \mathrm{I}) \oplus \cdots \oplus \ker(T - \lambda_k \mathrm{I})) && (9.12) \end{aligned}$$

onde usamos $\dim\ker(SR) \leq \dim\ker R + \dim\ker S$ para operadores lineares R e S; a segunda igualdade segue do Lema 9.41. Portanto, a soma dos autoespaços tem a mesma dimensão que V, ou seja, essa soma é V e T é diagonalizável.

b implica **c**

A Equação 9.12 implica que $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_m}$.

c implica **a**

Se $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_m}$, seja $\underline{\mathrm{B}}_i$ uma base para E_{λ_i} e faça $\underline{\mathrm{B}} = \underline{\mathrm{B}}_1 \cup \cdots \cup \underline{\mathrm{B}}_m$. Seja T_i a restrição de T a E_{λ_i}. Então $\underline{\mathrm{B}}$ é uma base para V e

$$[T]_{\underline{\mathrm{B}}} = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_m \end{bmatrix} = A,$$

uma matriz diagonal de bloco com $A_i = [T_i]_{\underline{B}_i}$. Mas, neste caso, A_i é a matriz $\lambda_i \mathrm{I}$ (um múltiplo escalar da matriz identidade).

c se, e somente se, d

Por definição.

d se, e somente se, e

Suponha que $c_T(x) = (x-\mu_1)\cdots(x-\mu_n)$. Os escalares $\mu_1, ..., \mu_n$ podem não ser todos distintos, por isso os agrupamos. Sejam $\lambda_1, ..., \lambda_m$ os autovalores distintos logo $c_T(x) = (x-\lambda_1)^{r_1} \ldots (x-\lambda_m)^{r_m}$ para números inteiros positivos $r_1, ..., r_m$.

Aqui r_i é a multiplicidade algébrica de λ_i e $r_1 + \cdots + r_m = n = \dim V$. Se f_i denota a multiplicidade geométrica, a Proposição 9.17 fornece $1 \leq f_i \leq r_i$. Portanto, $f_1 + \cdots + f_m = n$ se, e somente se, $f_i = r_i$ para todo i. Isso prova a equivalência entre d e e.

c se, e somente se, f

Pelo Teorema 9.38, $V = E^\infty_{\lambda_1} \oplus \cdots \oplus E^\infty_{\lambda_k}$. Portanto, $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$ se, e somente se, $E_{\lambda_i} = E^\infty_{\lambda_i}$ para todo i. Logo c e f são equivalentes.

■

> **Corolário 9.45.**
> *Seja V um espaço vetorial de dimensão finita e $T : V \to V$ uma transformação linear. Suponha que $c_T(x) = (x-\lambda_1)\cdots(x-\lambda_n)$ seja um produto de fatores lineares distintos. Então T é diagonalizável.*

Demonstração. Cada autovalor tem multiplicidade algébrica 1 e, pelo Corolário 9.18, multiplicidade geométrica 1. Assim, a soma das multiplicidades geométricas é $n = \dim V$, e o Teorema 9.44 mostra que T é diagonalizável. ■

Exercícios

Ex. 9.6 — Suponha que W seja um subespaço T-invariante de V. Mostre que T induz um operador linear $\overline{T} : V/W \to V/W$, dado por $\overline{T}(\mathbf{v}+W) = T(\mathbf{v}) + W$, e prove que o polinômio minimal de $\overline{T}$ divide o polinômio minimal de T.

Ex. 9.7 — Seja $A \in M_n(\mathbb{K})$

1. Mostre que c_A é um polinômio mônico de grau n.
2. Prove ainda que $c_A(x) = x^n - \mathrm{tr}(A)x^{n-1} + \cdots + (-1)^n \det(A)$.

Ex. 9.8 — Decida se o operador linear $T : \mathbb{K}^n \to \mathbb{K}^n$ é diagonalizável. Em caso positivo, calcule uma base de autovetores e a sua forma diagonal.

1. $[T]^{\underline{B}}_{\underline{B}} = \begin{bmatrix} 2 & 3 \\ -1 & 1 \end{bmatrix}$ com $\mathbb{K} = \mathbb{C}$

2. $[T]_{\underline{B}}^{\underline{B}} = \begin{bmatrix} -4 & -1 \\ 4 & 0 \end{bmatrix}$ com $\mathbb{K} = \mathbb{R}$

3. $[T]_{\underline{B}}^{\underline{B}} = \begin{bmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{bmatrix}$ com $\mathbb{K} = \mathbb{R}, \mathbb{C}$

Ex. 9.9 — Prove que

1. Suponha que $\dim V = n$. Se $\lambda_1, \dots, \lambda_n$ são os autovalores de T, contados com multiplicidade, então $\lambda_1^k, \dots, \lambda_n^k$ são os autovalores de T^k.

2. Seja V um espaço vetorial de dimensão finita sobre um corpo $\mathbb{K}$ algebricamente fechado e seja $T : V \to V$ um operador linear. Prove que α é um autovalor de $p(T)$ se, e somente se, $\alpha = p(\lambda)$ para algum λ autovalor de T.

Ex. 9.10 — Prove que o polinômio característico da transposta de um operador T^{t} coincide com o polinômio característico de T.

Ex. 9.11 — Seja $T : V \to V$ um operador linear. Mostre que se $\dim \mathrm{Im}(T) = m$, então T tem no máximo $m + 1$ autovalores.

Ex. 9.12 — Sejam $T, S : V \to V$ operadores lineares. Suponha que v é autovetor de T e de S associado aos autovalores λ_1, λ_2 de T e S, respectivamente. Ache um autovetor e um autovalor de:

1. $\alpha S + \beta T$ onde $\alpha, \beta \in \mathbb{R}$

2. $S \circ T$

Ex. 9.13 —

1. Mostre que se $B, M \in M_n(\mathbb{K})$ com M invertível, então $(M^{-1}BM)^n = (M^{-1}B^nM)$ para todo $n \in \mathbb{N}$.

2. Calcule A^n, $n \in \mathbb{N}$ onde $A = \begin{bmatrix} 2 & 4 \\ 3 & 13 \end{bmatrix}$.

3. Seja $A = \begin{bmatrix} 0 & 7 & -6 \\ -1 & 4 & 0 \\ 0 & 2 & -2 \end{bmatrix} \in M_3(\mathbb{C})$. Dado $n \in \mathbb{N}$ determine $B \in M_3(\mathbb{C})$ tal que $B^n = A$

Ex. 9.14 — Seja $A = \begin{bmatrix} 2 & 0 & -2 \\ 1 & 1 & -2 \\ 3 & 0 & -3 \end{bmatrix}$ calcule A^{2020}. Agora seja $B = \begin{bmatrix} 3 & -1 & -1 \\ -1 & 2 & 0 \\ 3 & -2 & 0 \end{bmatrix}$ consegue calcular B^{2020}? **Dica**: Pela divisão euclideana de x^k por P_B temos que $x^k = P_B(x)q_k(x) + r_k(x)$ onde $r_k = 0$ ou $\mathrm{grau}(r_k) < \mathrm{grau}(P_B)$. Encontre r_k.

Ex. 9.15 — Seja $T : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$. A matriz abaixo representa T usando a base $\underline{\mathrm{B}}$ no domínio e a base $\underline{\mathrm{C}}$ no contradomínio:

$$[T]_{\underline{\mathrm{C}}}^{\underline{\mathrm{B}}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

onde $\underline{\mathrm{B}} = \{\frac{1}{2}x^2, -\frac{1}{2}, \frac{1}{2}x\}$ e $\underline{\mathrm{C}} = \{x^2, x, 1\}$. Determine T nos elementos da base canônica, calcule $[T]_{\underline{\mathrm{C}}}^{\underline{\mathrm{C}}}$ e mostre que T é diagonalizável.

Ex. 9.16 — Em $\mathscr{F}(\mathbb{R}, \mathbb{R})$ considere as funções $f_1(x) = e^{2x}\operatorname{sen}(x)$, $f_2(x) = e^{2x}\cos(x)$ e $f_3(x) = e^{2x}$, o subespaço $S = \langle f_1, f_2, f_3\rangle$ e o operador linear $D : S \to S$ definido por $D(f) = f'$. Determine:

1. A matriz de D em relação à base $\underline{\mathrm{B}} = \{f_1, f_2, f_3\}$ de S.

2. Os autovalores de D e as funções de S que são autovetores de D.

Ex. 9.17 — Seja $T : V \to V$ ($\dim V < \infty$) um operador diagonalizável cujos autovalores têm multiplicidade algébrica 1.

1. Prove que qualquer operador $G : V \to V$ tal que $GT = TG$ pode ser representado como um polinômio em T.

2. Prove que a dimensão do espaço vetorial formado por tais operadores (operadores que comutam com T) é igual a dimensão de V.

Ex. 9.18 — Sejam $S, T : V \to V$ operadores diagonalizáveis que comutam. Mostre que eles são simultaneamente diagonalizáveis, isto é, que existe uma base $\underline{\mathrm{B}}$ de V tal que $\underline{\mathrm{B}}$ consiste em autovetores de S e de T.

Ex. 9.19 — Sejam V um $\mathbb{K}$-e.v. de $\dim V < \infty$ e $T : V \to V$ um operador linear. Suponha que T comuta com todo operador diagonalizável. Prove que T é um múltiplo escalar da identidade.

Ex. 9.20 — Seja $m \in \mathbb{R}$ e A a matriz dada por

$$A = \begin{bmatrix} 1+m & 1+m & 1 \\ -m & -m & -1 \\ m & m-1 & 0 \end{bmatrix}.$$

1. Encontre os autovalores de A.

2. Para que valores de m a matriz A é diagonalizável?

3. Determine, de acordo com os diferentes valores de m, o polinômio minimal de A.

Ex. 9.21 — Fixemos um vetor não nulo $a \in \mathbb{R}^3$ e defina a aplicação $T : \mathbb{R}^3 \to \mathbb{R}^3$ por $T(v) = a \times v$ (produto vetorial).

1. Prove que T é uma transformação linear.

2. Determine os autovalores e autovetores de T.

Ex. 9.22 — Considere uma matriz real simétrica A de ordem 3 com determinante igual a 6. Suponha que $u = (4, 8, -1)$ e $v = (1, 0, 4)$ sejam autovetores desta matriz associados aos autovalores 1 e 2, respectivamente. Decida se as seguintes afirmações são verdadeiras ou falsas:

1. Os autovalores de A são apenas 1 e 2.

2. O produto vetorial $u \times v$ é autovetor de A.

3. O vetor $(5, 8, 3)$ é autovetor de A.

4. A pode não ser diagonalizável.

Ex. 9.23 — Prove que se $T : V \to V$ é um operador linear, então $\ker T$, $\mathrm{Im}(T)$ são subespaços T-invariantes. Se λ for um autovalor de T, então E_λ é um subespaço T-invariante.

Ex. 9.24 — Prove que a soma e a intersecção de subespaços T-invariantes é T-invariante.

Ex. 9.25 — Prove que se um operador $T \in \mathrm{Hom}(V, V)$ com $\dim V < \infty$ é um isomorfismo então T e T^{-1} possuem os mesmos subespaços invariantes. Vale em dimensão infinita?

Ex. 9.26 — Mostre que se todo subespaço de V for T-invariante então $T = \lambda \mathrm{I}$ para algum $\lambda \in \mathbb{K}$.

Ex. 9.27 — Mostre que $W \subseteq V$ é um subespaço invariante para $T \in \mathrm{Hom}(V, V)$ se, e somente se, seu anulador $\mathrm{Aniq}(W) \subseteq V^*$ é T^{t}-invariante.

Ex. 9.28 — Sejam $T, G : V \to V$ operadores que comutam. Mostre que $\ker T$, $\mathrm{Im}(T)$ e E_λ são G-invariantes.

Ex. 9.29 — Sejam V um $\mathbb{K}$-e.v. de $\dim V = n$, $T : V \to V$ um operador linear, $W \subseteq V$ um subespaço T-invariante. Mostre que $c_{T|_W} \mid c_T$ e $m_{T|_W} \mid m_T$. **Dica**: lembre que se $A \in M_n(\mathbb{K})$ é diagonal por blocos

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix},$$

onde $B \in M_k(\mathbb{K})$ e $D \in M_{n-k}(\mathbb{K})$, então $A^n = \begin{bmatrix} B^n & \widetilde{C} \\ 0 & D^n \end{bmatrix}$.

Ex. 9.30 — Sejam $T : V \to V$ um operador linear, $W \subseteq V$ um subespaço. Mostre que W é $p(T)$-invariante para todo polinômio $p \in \mathbb{K}[x]$ se, e somente se, W for T-invariante.

Ex. 9.31 — Seja $\pi : V \to V$ um operador projeção não trivial (i.e., $\pi \neq 0_{\mathrm{Hom}(V,V)}$ e $\pi \neq \mathrm{I}_V$) mostre que $m_\pi = x^2 - x$.

Ex. 9.32 — Seja V um $\mathbb{C}$-espaço vetorial. Se o polinômio característico de uma transformação linear $T : V \to V$ é $x^2 - x - 1$, então é correto afirmar que:

1. T não é necessariamente invertível.
2. T é invertível e $T^{-1} = T + I$.
3. Não existe T com tal polinômio característico.
4. T é invertível e $T^{-1} = T - I$.
5. T é invertível, mas nenhuma das fórmulas para a inversa de T nos outros itens é válida.

Ex. 9.33 — Seja $T \in \mathrm{Hom}(V, V)$.

1. Mostre que $T : V \to V$ é um operador linear não injetor se e somente se 0 é um autovalor de T.
2. Mostre que T é invertível se, e somente se, o termo independente de seu polinômio minimal é não-nulo.
3. Nestas circunstâncias, T^{-1} é um polinômio em T, i.e., existe $p(x) \in \mathbb{K}[x]$ tal que $T^{-1} = p(T)$.

Ex. 9.34 — Seja A uma matriz complexa tal que $A^k = I$ para algum inteiro k. Prove que A é diagonalizável.

Ex. 9.35 — Prove que uma matriz $n \times n$ sobre $\mathbb{C}$ que satisfaz $A^3 = A$ pode ser diagonalizada.

Ex. 9.36 — Encontre todas as possibilidades para o polinômio minimal de um operador $T : \mathbb{R}^5 \to \mathbb{R}^5$ com polinômio característico c_T dado. É possível concluir que algum deles é necessariamente diagonalizável?

1. $c_T(x) = (x-3)^3(x-2)^2$
2. $c_T(x) = (x-1)(x-2)(x-3)(x-4)(x-5)$

3. $c_T(x) = (x-1)^5$

Ex. 9.37 — Seja V um espaço vetorial de dimensão finita sobre um corpo $\mathbb{K}$ algebricamente fechado e seja $T : V \to V$ um operador. Prove que T é nilpotente se, e somente se, todos os seus autovalores forem iguais a zero.

Ex. 9.38 — Seja V um e.v. sobre um corpo $\mathbb{K}$ algebricamente fechado e $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ os autovalores de T e $q(x) = (x-\lambda_1)\cdots(x-\lambda_r)$. Mostre que $q(T)$ é nilpotente. Qual o índice de nilpotência?

Ex. 9.39 — Dado T, G operadores lineares num espaço n-dimensional sobre um corpo de característica zero. Assuma que $T^n = 0$, $\dim\ker T = 1$ e que $GT - TG = T$. Prove que os autovalores de G são da forma $\alpha, \alpha-1, \alpha-2, \dots, \alpha-(n-1)$ para algum $\alpha \in \mathbb{K}$.

Ex. 9.40 — Seja V um $\mathbb{K}$-e.v. de dimensão finita e $T : V \to V$ um operador linear. Mostre que existem únicos subespaços T-invariantes V_1, V_2 tais que $V = V_1 \oplus V_2$, a restrição $T|_{V_1}$ é nilpotente e $T|_{V_2}$ é um isomorfismo.

Ex. 9.41 — Seja $T : \mathbb{K}^{\mathbb{N}} \to \mathbb{K}^{\mathbb{N}}$ dada por $T(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$. Mostre que T não se escreve como soma direta de um operador nilpotente com um isomorfismo.

Ex. 9.42 — Seja $T : \mathbb{K}^n \to \mathbb{K}^n$ um operador com polinômio característico $c_T(x) = (x-\lambda)^n$. Mostre que o operador $T' = \lambda \mathrm{I} - T$ é nilpotente.

Ex. 9.43 — Seja V um $\mathbb{K}$-e.v. de dimensão finita e $T : V \to V$ um operador linear. Se $\dim \mathrm{im}(T) = 1$ mostre que ou T é diagonalizável ou T é nilpotente.

Ex. 9.44 — Se V é um $\mathbb{C}$-e.v. de dimensão finita e $T : V \to V$ é um operador linear mostre que $T^{n+1} = T$ se, e somente se, $T^{2n} = T^n$ e T diagonalizável.

Ex. 9.45 — Para V espaço vetorial real de dimensão finita, uma involução em V é um operador linear $\varphi : V \to V$ tal que $\varphi^2 = \mathrm{I}$. Dado um operador linear $T : V \to V$, sejam

$$\mathrm{Fix}(T) = \{x \in V : T(x) = x\} \quad \text{e} \quad A(T) = \{x \in V : T(x) = -x\}$$

chamados subespaço de pontos fixos de T e subespaço antipodal de T, respectivamente. Decida se as seguintes afirmações são verdadeiras ou falsas:

1. Se φ é uma involução, então $\det(\varphi) = 1$.
2. Se φ_1, φ_2 são involuções então $\mathrm{Fix}(\varphi_1 \circ \varphi_2) = \ker(\varphi_1 - \varphi_2)$.
3. Se λ é autovalor de uma involução, então $\lambda = \pm 1$.
4. Se φ_1, φ_2 são involuções, então $\varphi_1 \circ \varphi_2$ é também uma involução.
5. Se φ é uma involução, então $A(\varphi) = \mathrm{im}(\mathrm{I} - \varphi)$.

Ex. 9.46 — [Densidade de $\mathrm{GL}_n(\mathbb{C})$ e uma aplicação] Sejam $A \in \mathscr{M}_{n,m}(\mathbb{K})(\mathbb{C})$ e $B \in \mathscr{M}_{m,n}(\mathbb{K})(\mathbb{C})$, com $m, n \geq 1$. O objetivo deste exercício é provar a identidade

$$x^m c_{AB}(x) = x^n c_{BA}(x).$$

Em particular, se $m = n$, então $c_{AB} = c_{BA}$; no caso geral, AB e BA têm os mesmos autovalores não nulos, com as mesmas multiplicidades algébricas.

1. Suponha primeiro que $m = n$ e que A seja invertível. Mostre que AB e BA são semelhantes e conclua que $c_{AB} = c_{BA}$.

2. Mostre que $\mathrm{GL}_n(\mathbb{C})$ é denso em $\mathscr{M}_{n,n}(\mathbb{K})(\mathbb{C})$.
   Dica: para uma matriz M, considere $M + t\,\mathrm{I}$ e observe que $\det(M + t\,\mathrm{I})$ é um polinômio não nulo em t.

3. Usando a continuidade dos coeficientes do polinômio característico como funções das entradas, deduza que $c_{AB} = c_{BA}$ para matrizes quadradas A e B, sem supor A invertível.

4. Aplique o item anterior às matrizes em blocos

   $$\widetilde{A} = \begin{bmatrix} 0 & A \\ 0 & 0 \end{bmatrix}, \qquad \widetilde{B} = \begin{bmatrix} 0 & 0 \\ B & 0 \end{bmatrix}$$

   de ordem $n + m$ e deduza a identidade do enunciado.

5. A afirmação análoga vale para os polinômios minimais de AB e BA? Justifique sua resposta.
   Dica: procure um exemplo com $AB = 0$ e $BA \neq 0$ nilpotente.

10

Forma Normal de Jordan

Uma matriz diagonal descreve um operador por sua ação em retas invariantes. Quando não existe uma base de autovetores, a forma de Jordan fornece, sob a hipótese de fatoração linear, uma representação por blocos. Cada bloco reúne um autovalor e uma parte nilpotente.

Essa forma existe quando o polinômio característico se decompõe em fatores lineares sobre $\mathbb{K}$, como sempre ocorre sobre $\mathbb{C}$. Sobre $\mathbb{R}$, pares de autovalores complexos conjugados dão origem a blocos reais; em corpos nos quais a fatoração linear falha, a forma canônica racional, estudada no próximo capítulo, ocupa seu lugar.

Começamos pelo formato da matriz que queremos obter.

Seja $T : V \to V$ um operador linear em um espaço vetorial de dimensão finita tal que o polinômio característico do operador se decompõe em fatores lineares sobre $\mathbb{K}$ (o que sempre ocorre se $\mathbb{K}$ é algebricamente fechado). Mostraremos que existe uma base $\underline{\mathrm{B}}$ de V, na qual $[T]_{\underline{\mathrm{B}}}$ é uma matriz na forma normal de Jordan:

$$\begin{bmatrix} J_{t_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{t_2}(\lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{t_r}(\lambda_r) \end{bmatrix} \quad \text{com} \quad J_t(\lambda) = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}.$$

As matrizes $J_t(\lambda)$, com $t \in \mathbb{N}^*$ e $\lambda \in \mathbb{K}$, são denominadas **blocos de Jordan**.

Notação 10.1. Ao longo deste livro adotamos a convenção triangular superior: os 1's de $J_t(\lambda)$ ficam imediatamente acima da diagonal. Assim, uma cadeia ordenada $(\mathbf{u}_1, \dots, \mathbf{u}_t)$ satisfaz

$$(T - \lambda\,\mathrm{I})\mathbf{u}_1 = \vec{0}, \qquad (T - \lambda\,\mathrm{I})\mathbf{u}_j = \mathbf{u}_{j-1} \quad (2 \le j \le t).$$

A base na qual o operador assume essa forma é denominada **base de Jordan**. Para construí-la, examinaremos primeiro os operadores nilpotentes, que constituem a parte não diagonal de cada bloco.

Nesse capítulo

Cadeias que terminam em zero
Um operador nilpotente não precisa ser o operador nulo: ele apenas se torna nulo após um número finito de iterações. Sua estrutura é descrita por cadeias $v, Tv, T^2v, \dots$ que eventualmente chegam a zero.

10.1 Operadores Nilpotentes

Definição 10.2 — Nilpotente e unipotente. Um operador linear $T : V \to V$ é dito

1. **nilpotente** se $T^k = 0$ para algum $k \in \mathbb{N}$;
2. **unipotente** se $T = \mathrm{I} + N$, com N nilpotente.

Assim, T é unipotente se, e somente se, $T - \mathrm{I}$ é nilpotente. Matrizes nilpotentes e unipotentes são definidas de maneira análoga.

Exemplo 10.3. Seja $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ em $\mathscr{M}_{2,2}(\mathbb{K})$. Esta é uma matriz nilpotente e, em qualquer corpo, a única raiz de seu polinômio característico $c_A(x) = \det(x\,\mathrm{I} - A) = x^2$ é 0. Existe um autovetor não trivial $\mathbf{e}_1 = (1, 0)$, correspondente ao autovalor 0. Os múltiplos escalares de $\mathbf{e}_1$ são os únicos autovetores de A, e portanto, não existe uma base de autovetores e assim a matriz A não pode ser diagonalizada, independentemente do corpo $\mathbb{K}$.

◁

Um operador nilpotente só pode ser diagonalizável quando é nulo. Precisamos, portanto, de uma descrição que registre a maneira como seus vetores se organizam em cadeias.

Proposição 10.4.
Seja V um espaço vetorial de dimensão finita sobre $\mathbb{K}$. Se $T : V \to V$ for nilpotente então $c_T(x) = \det(x\,\mathrm{I} - T) = x^n$ onde $n = \dim V$ e assim 0 é o único autovalor e o autoespaço associado a 0 é $\ker T$.

Como consequência direta do Teorema de Schur temos:

Proposição 10.5.
Dado um operador nilpotente T em um espaço vetorial de dimensão finita, existe uma base na qual a matriz de T é triangular estritamente superior.

Decomposição de Fitting Para um operador $T : V \to V$ em dimensão finita, ponha $K_i = \ker(T^i)$ e $R_i = \mathrm{im}(T^i)$. Esses espaços formam as cadeias

$$\{0\} \subseteq K_1 \subseteq K_2 \subseteq \cdots \subseteq K_i \subseteq K_{i+1} \subseteq \cdots \tag{10.1}$$

$$V \supseteq R_1 \supseteq R_2 \supseteq \cdots \supseteq R_i \supseteq R_{i+1} \supseteq \cdots, \tag{10.2}$$

Como suas dimensões são monótonas e limitadas, ambas se estabilizam. Além disso, o teorema do núcleo e da imagem fornece $\dim V = \dim K_i + \dim R_i$ em cada etapa; portanto, os núcleos e as imagens se estabilizam no mesmo índice, que denotaremos por r.

As cadeias estabilizadas determinam

- a **imagem estável**, $R_\infty = \bigcap_{i=1}^{\infty} R_i = R_r = R_{r+1} = \cdots$;
- o **núcleo estável**, $K_\infty = \bigcup_{i=1}^{\infty} K_i = K_r = K_{r+1} = \cdots$.

Teorema 10.6 (Decomposição de Fitting).
Sejam V um espaço de dimensão finita e $T : V \to V$. Então V admite uma única decomposição $V = R \oplus K$ tal que

1. *os espaços R, K são T-invariantes;*
2. *$T\restriction_K$ é um operador linear nilpotente em K;*
3. *$T\restriction_R$ é um operador linear bijetivo em R.*

Em termos das restrições do operador, essa decomposição pode ser escrita como

$$T = (T\restriction_R) \oplus (T\restriction_K)$$

com $T\restriction_K$ nilpotente e $T\restriction_R$ bijetivo.

Parte invertível e parte nilpotente
A decomposição de Fitting escreve o espaço como soma direta de dois subespaços invariantes. Em um deles, a restrição de T é invertível; no outro, é nilpotente.

Demonstração. Escolhemos $R = R_\infty$ e $K = K_\infty$.

Afirmamos que T é nilpotente em K_∞ e invertível em R_∞. De fato, $T(R_\infty) = R_\infty$; como esse espaço tem dimensão finita, a sobrejetividade da restrição implica sua invertibilidade. Além disso, $K_\infty = \ker(T^r)$ para algum $r \geq 1$, de modo que a restrição de T a K_∞ é nilpotente.

Afirmamos que $V = K_\infty \oplus R_\infty$. É direto que $K_\infty \cap R_\infty = \{\vec{0}\}$. Portanto, basta mostrar que todo $\mathbf{v} \in V$ pode ser decomposto como $k + r$ com $k \in K_\infty$ e $r \in R_\infty$. Sem perda de generalidade, podemos assumir que $K_\infty = \ker(T^r)$ e $R_\infty = T^r V$. Como $\operatorname{im}(T^{2r}) = \operatorname{im}(T^r)$, temos $T^r(\mathbf{v}) = T^{2r}(\mathbf{w})$ para algum $\mathbf{w} \in V$. Logo $T^r(\mathbf{v} - T^r(\mathbf{w})) = T^r(\mathbf{v}) - T^{2r}(\mathbf{w}) = 0$. Então definimos

$$\mathbf{v}' := \mathbf{v} - T^r(\mathbf{w}) \in K_\infty \quad \text{e} \quad \mathbf{v}'' := T^r(\mathbf{w}) \in R_\infty$$

e temos a decomposição desejada de $\mathbf{v}$.

Finalmente, mostramos que a decomposição $V = K_\infty \oplus R_\infty$ é única. Suponha $V = A \oplus B$ com T nilpotente em A e invertível em B. Então $A \subseteq \ker(T^k)$ para algum inteiro positivo k e $B \subseteq \operatorname{im}(T^k)$ para todo inteiro positivo k. Assim,

$$A \subseteq K_\infty \text{ e } B \subseteq R_\infty.$$

Por considerações de dimensão, devemos ter $A = K_\infty$ e $B = R_\infty$.

■

Definição 10.7 — Índice de Nilpotência. Seja $T : V \to V$ nilpotente, o **índice de nilpotência**, denotado por $\operatorname{ind} T$, é o menor expoente r tal que $T^r = 0$.

Se V é um espaço vetorial de dimensão finita então existe o menor expoente tal que $T^r = 0$ pois a cadeia

$$\{0\} \subseteq K_1 \subseteq K_2 \subseteq \cdots \subseteq K_i \subseteq K_{i+1} \subseteq \cdots$$

se estabiliza em $K_r = V$, sendo $K_i = \ker(T^i)$.

Exemplos 10.8. Sejam

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

então é direto ver que $\operatorname{ind} A = 2$, $\operatorname{ind} B = 3$ e $\operatorname{ind} C = 2$.

◁

Lema 10.9.
Se, para algum vetor $\mathbf{v} \in V$ e algum inteiro m, tivermos:

$$T^{m-1}\mathbf{v} \neq 0 \text{ mas } T^m\mathbf{v} = 0$$

então o conjunto $\{\mathbf{v}, T\mathbf{v}, \dots, T^{m-1}\mathbf{v}\}$ é linearmente independente.

Demonstração. A demonstração deste lema será deixada como exercício ao leitor. ■

Definição 10.10 — Operador cíclico. Se $n = \dim V$, um operador $T \in \operatorname{Hom}(V, V)$ é **cíclico** quando existe $\mathbf{v} \in V$ tal que $\{\mathbf{v}, T\mathbf{v}, \dots, T^{n-1}\mathbf{v}\}$ é uma base de V. Nesse caso, $\mathbf{v}$ é um **vetor cíclico** e a família indicada é uma **base cíclica**.

Proposição 10.11.
Se V é um espaço vetorial finito dimensional e T é um operador nilpotente com índice de nilpotência igual à dimensão de V, então T é cíclico.

Demonstração. Seja $n = \dim V$. Como o índice de nilpotência é n, existe $\mathbf{v}$ tal que $T^{n-1}\mathbf{v} \neq \vec{0}$. Logo, pelo Lema 10.9, o conjunto $\{\mathbf{v}, T\mathbf{v}, \dots, T^{n-1}\mathbf{v}\}$ é linearmente independente e, por dimensão, é uma base de V. ■

Exemplo 10.12. O operador associado à matriz abaixo é nilpotente de índice 4:

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

O vetor $\mathbf{e}_1 = [1, 0, 0, 0] \notin \ker T^3$. Logo o conjunto

$$\underline{\mathrm{B}} = (\mathbf{e}_1, T\mathbf{e}_1, T^2\mathbf{e}_1, T^3\mathbf{e}_1) = ([1, 0, 0, 0], [-2, -2, 0, 1], [-2, -2, 0, 0], [-8, -4, -4, -2])$$

é uma base de V e nessa base

$$[T]_{\underline{\mathrm{B}}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Por outro lado se escolhermos a base na ordem reversa

$$\underline{\mathrm{B}}_2 = (T^3\mathbf{e}_1, T^2\mathbf{e}_1, T\mathbf{e}_1, \mathbf{e}_1) = ([-8, -4, -4, -2], [-2, -2, 0, 0], [-2, -2, 0, 1], [1, 0, 0, 0])$$

temos que

$$[T]_{\underline{B}_2} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Esta é a ordem das cadeias compatível com a convenção triangular superior adotada neste capítulo.

◁

Na próxima seção trataremos o caso geral em que o índice de nilpotência de T é menor que a dimensão de V.

Operadores nilpotentes e vetores cíclicos O resultado seguinte caracteriza bases de Jordan em termos das cadeias de um operador nilpotente.

Proposição 10.13.
Seja V um espaço de dimensão finita e $T \in \mathrm{Hom}(V,V)$ nilpotente então são equivalentes:

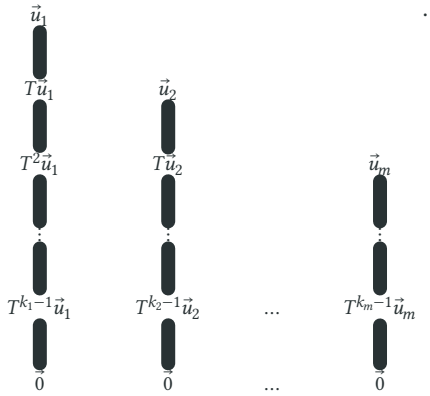
1. $\left(\vec{u}_1^1, \ldots, \vec{u}_{k_1}^1, \vec{u}_1^2, \ldots, \vec{u}_{k_2}^2, \ldots, \vec{u}_1^m, \ldots \vec{u}_{k_m}^m\right)$ *é uma base de Jordan para T;*
2. $\left(\vec{u}_1^1, \ldots, \vec{u}_{k_1}^1, \vec{u}_1^2, \ldots, \vec{u}_{k_2}^2, \ldots, \vec{u}_1^m, \ldots \vec{u}_{k_m}^m\right)$ *é uma base tal que $T\vec{u}_1^l = \vec{0}$ e $T\vec{u}_j^l = \vec{u}_{j-1}^l$ para $2 \leq j \leq k_l$.*
3. *Se $\mathbf{w}_j = \mathbf{u}_{k_j}^j$, então*

$$\{T^{k_1-1}\mathbf{w}_1, \ldots, T\mathbf{w}_1, \mathbf{w}_1, \ldots, T^{k_m-1}\mathbf{w}_m, \ldots, T\mathbf{w}_m, \mathbf{w}_m\}$$

é base de V.

Demonstração. A demonstração será deixada como exercício. ■

Uma representação diagramática da base de Jordan é



Teorema 10.14 (Decomposição Cíclica para Operadores Nilpotentes).
Todo operador T nilpotente agindo num espaço vetorial finito dimensional possui uma base de Jordan $\underline{J}$. Nessa base a matriz de T é uma combinação de blocos da forma

$$[T]_{\underline{J}} = \operatorname{diag}(J_{a_1}(0), \dots, J_{a_k}(0))$$

com $a_1 \geq \cdots \geq a_k \geq 1$ e $a_1 + \cdots + a_k = n$.
Se denotarmos por n_i o número de blocos de Jordan $J_i(0)$ de ordem i que aparecem em $[T]_{\underline{J}}$, então

$$n_i = 2\dim\ker T^i - \dim\ker T^{i-1} - \dim\ker T^{i+1}.$$

Além disso, tal representação de T como uma matriz diagonal de blocos com blocos Jordan é única a menos de permutação dos blocos.

Blocos são cadeias
Cada subespaço cíclico nilpotente fornece uma cadeia de vetores e, numa base adaptada, um bloco de Jordan. O tamanho do bloco é o comprimento da cadeia antes de ela atingir zero.

Demonstração. Faremos a demonstração por indução sobre $\dim V$. Como T é nilpotente, $\dim \operatorname{im} T < \dim V$. Se $\operatorname{im} T = 0$, $T = 0$ e o resultado é trivial, por isso, podemos assumir que $\operatorname{im} T \neq 0$.

Por indução, existem $\vec{u}_1, \dots, \vec{u}_k \in \operatorname{im} T$, de modo que

$$\{\vec{u}_1, T\vec{u}_1, \dots, T^{a_1-1}\vec{u}_1, \dots, \vec{u}_k, T\vec{u}_k, \dots, T^{a_k-1}\vec{u}_k\}$$

é uma base de Jordan para T.

Para $1 \leq i \leq k$ escolha $\vec{v}_i \in V$ tal que $\vec{u}_i = T\vec{v}_i$.

Como o último vetor de cada cadeia pertence ao núcleo,

$$\left\langle T^{a_1-1}\vec{u}_1, \dots, T^{a_k-1}\vec{u}_k \right\rangle \subseteq \ker T.$$

Estendemos esse conjunto linearmente independente a uma base de $\ker T$, adicionando os vetores $\vec{w}_1, \dots, \vec{w}_l$.

Afirmamos que os vetores

$$\{\vec{v}_1, T\vec{v}_1, \dots, T^{a_1}\vec{v}_1, \dots, \vec{v}_k, T\vec{v}_k, \dots, T^{a_k}\vec{v}_k, \vec{w}_1, \dots, \vec{w}_l\}$$

formam uma base para V.

Para verificar a independência linear, suponha que uma combinação desses vetores seja nula:

$$\sum_{i=1}^{k}\sum_{j=0}^{a_i} c_{ij}T^j\vec{v}_i + \sum_{m=1}^{l} b_m\vec{w}_m = 0.$$

Aplicando T, obtemos

$$\sum_{i=1}^{k}\sum_{j=0}^{a_i-1} c_{ij}T^j\vec{u}_i = 0.$$

Como esses vetores pertencem à base de Jordan de $\operatorname{im} T$, segue que $c_{ij} = 0$ para $0 \leq j \leq a_i - 1$. Logo a combinação inicial se reduz a uma combinação dos vetores $T^{a_1}\vec{v}_1, \dots, T^{a_k}\vec{v}_k, \vec{w}_1, \dots, \vec{w}_l$. Mas esses vetores formam uma base de $\ker T$, logo os coeficientes restantes também são nulos. Logo a família é linearmente independente.

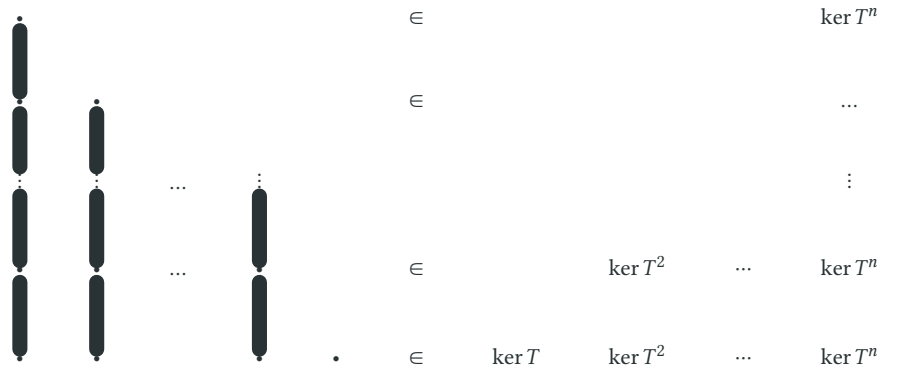
Para mostrar que esses vetores geram V, usamos um argumento dimensional. Sabemos que $\dim\ker T = k+l$ e $\dim \operatorname{im} T = a_1+\ldots+a_k$. Por isso $\dim V = (a_1+1)+\ldots+(a_k+1)+l$, que é o número de vetores acima.

Portanto, construímos cadeias que formam uma base de V. Ordenando cada cadeia do vetor anulado para o seu gerador, obtemos a convenção triangular superior adotada no início do capítulo. Esquematicamente,

$$\begin{aligned} \mathbf{v}_i &\longmapsto T\mathbf{v}_i \longmapsto \cdots \longmapsto T^{a_i}\mathbf{v}_i \longmapsto \vec{0}, \quad 1 \leq i \leq k, \\ \mathbf{w}_m &\longmapsto \vec{0}, \qquad\qquad\qquad\qquad\quad 1 \leq m \leq l. \end{aligned}$$

Os vetores em ocre formam a base de Jordan de $\operatorname{im} T$ usada na hipótese de indução. Pode haver várias cadeias de mesmo comprimento, e o comprimento da maior é o índice de nilpotência de T.

Para calcular a fórmula para n_i, observe a maneira como escrevemos a base $\underline{\mathrm{J}}$ como um diagrama bidimensional. Cada bloco de Jordan em $[T]_{\underline{\mathrm{J}}}$ corresponde a uma coluna de vetores nessa matriz.



Os vetores nas últimas k linhas correspondem a uma base de $\ker T^k$. Assim o número de vetores na k-ésima linha é $\dim \ker T^k - \dim \ker T^{k-1}$. E na $k+1$-ésima linha é $\dim \ker T^{k+1} - \dim \ker T^k$. A diferença desses dois números é exatamente o número de colunas de altura exatamente k.

$$\begin{aligned} n_i &= \dim \ker T^i - \dim \ker T^{i-1} - (\dim \ker T^{i+1} - \dim \ker T^i) && (10.3) \\ &= 2\dim \ker T^i - \dim \ker T^{i-1} - \dim \ker T^{i+1} && (10.4) \end{aligned}$$

Observe que os números n_i não dependem da base específica. Portanto, a representação do bloco Jordan de T é única a menos de uma permutação dos blocos.

■

Exemplo 10.15. O operador linear associado à matriz abaixo é nilpotente de índice 2:

$$\begin{bmatrix} -4 & 6 & 4 & -4 \\ 0 & 2 & -2 & 0 \\ 0 & 2 & -2 & 0 \\ 4 & -1 & -9 & 4 \end{bmatrix}$$

Escalonando a matriz temos a seguinte forma escalonada:

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

As colunas pivôs são a primeira e a segunda; portanto, as colunas correspondentes da matriz original, isto é, $T\mathbf{e}_1$ e $T\mathbf{e}_2$, formam uma base de $\operatorname{im} T$. Escolhemos $\mathbf{e}_1$ e $\mathbf{e}_2$ como preimagens desses vetores. Logo o conjunto

$$\underline{\mathrm{C}} = (\mathbf{e}_1, T\mathbf{e}_1, \mathbf{e}_2, T\mathbf{e}_2) = ([1,0,0,0], [-4,0,0,4], [0,1,0,0], [6,2,2,-1])$$

é uma base de Jordan para V. O diagrama representando esse operador é:

E na base $\underline{C}$ temos que:

$$[T]_{\underline{C}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

As matrizes de mudança de base são

$$M_{\underline{C}\to\underline{B}} = \begin{bmatrix} 1 & -4 & 0 & 6 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 2 \\ 0 & 4 & 0 & -1 \end{bmatrix} \qquad M_{\underline{B}\to\underline{C}} = \begin{bmatrix} 1 & 0 & -\frac{5}{2} & 1 \\ 0 & 0 & \frac{1}{8} & \frac{1}{4} \\ 0 & 1 & -1 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \end{bmatrix}$$

E assim

$$[T]_{\underline{C}} = M_{\underline{B}\to\underline{C}}[T]_{\underline{B}}\, M_{\underline{C}\to\underline{B}}$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -\frac{5}{2} & 1 \\ 0 & 0 & \frac{1}{8} & \frac{1}{4} \\ 0 & 1 & -1 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} -4 & 6 & 4 & -4 \\ 0 & 2 & -2 & 0 \\ 0 & 2 & -2 & 0 \\ 4 & -1 & -9 & 4 \end{bmatrix} \begin{bmatrix} 1 & -4 & 0 & 6 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 2 \\ 0 & 4 & 0 & -1 \end{bmatrix}.$$

◁

10.2 Forma Normal de Jordan

Antes da demonstração, registramos duas descrições equivalentes de uma base de Jordan para $T \in \mathrm{Hom}(V,V)$:

1. $\left(\vec{u}_1^1, \dots, \vec{u}_{k_1}^1, \vec{u}_1^2, \dots, \vec{u}_{k_2}^2, \dots, \vec{u}_1^m, \dots \vec{u}_{k_m}^m\right)$ é uma base de Jordan para T;
2. para cada cadeia, existe $\lambda_l \in \mathbb{K}$ tal que $T\mathbf{u}_1^l = \lambda_l \mathbf{u}_1^l$ e $T\mathbf{u}_j^l = \lambda_l \mathbf{u}_j^l + \mathbf{u}_{j-1}^l$ para $2 \le j \le k_l$.

Estrutura dos blocos de Jordan
Cada bloco tem um autovalor na diagonal e, imediatamente acima dela, entradas iguais a 1. Essas entradas descrevem a ação nilpotente ao longo da cadeia associada. Um operador na forma de Jordan é diagonalizável exatamente quando todos os blocos têm tamanho um.

Teorema 10.16 (Forma Normal de Jordan).
Seja $T : V \to V$ um operador linear em um espaço vetorial de dimensão finita tal que o polinômio característico do operador se decompõe em fatores lineares sobre $\mathbb{K}$. Então:

1. *Existe uma base de Jordan para T, isto é existe uma base de V na qual a matriz de T está na forma normal de Jordan, i.e., existe uma matriz mudança de base M tal que a matriz do operador na base original A pode ser reduzida a forma de Jordan*
$$M^{-1}AM = J$$
2. *Se denotarmos por n_i^λ o número de blocos de Jordan $J_i(\lambda)$ de ordem i que*

aparecem na matriz de Jordan para T, então

$$n_i^{\lambda} = 2\dim\ker(T-\lambda\,\mathrm{I})^{i} - \dim\ker(T-\lambda\,\mathrm{I})^{i-1} - \dim\ker(T-\lambda\,\mathrm{I})^{i+1}.$$

3 *A matriz J é única, a menos de permutação dos blocos de Jordan.*

Demonstração. Pelo Teorema da Decomposição em Autoespaços Generalizados,

$$V = \bigoplus_{i=1}^{m} E^{\infty}_{\lambda_i}(T),$$

onde $\lambda_1, \dots, \lambda_m$ são os autovalores distintos de T. Em cada parcela, o operador

$$N_i = (T-\lambda_i\,\mathrm{I})\restriction_{E^{\infty}_{\lambda_i}(T)}$$

é nilpotente. Pelo Teorema da Decomposição Cíclica para Operadores Nilpotentes, existe uma base $\underline{\mathrm{B}}_i$ dessa parcela na qual

$$[N_i]_{\underline{\mathrm{B}}_i} = \operatorname{diag}\big(J_{t_1^{(i)}}(0), \dots, J_{t_{r_i}^{(i)}}(0)\big).$$

Como $T\restriction_{E^{\infty}_{\lambda_i}(T)} = \lambda_i\,\mathrm{I} + N_i$, na mesma base temos

$$[T\restriction_{E^{\infty}_{\lambda_i}(T)}]_{\underline{\mathrm{B}}_i} = \operatorname{diag}\big(J_{t_1^{(i)}}(\lambda_i), \dots, J_{t_{r_i}^{(i)}}(\lambda_i)\big).$$

Concatenando $\underline{\mathrm{B}}_1, \dots, \underline{\mathrm{B}}_m$, obtemos uma base de V e a forma de Jordan de T. Isso prova a existência.

Unicidade da forma de Jordan

Seja $\underline{\mathrm{B}}$ uma base de Jordan do operador T. Nessa base, as entradas diagonais de $[T]_{\underline{\mathrm{B}}}$ são os autovalores de T. Fixe um autovalor λ e denote por W_λ o subespaço gerado por todas as cadeias dos blocos associados a λ. Em cada uma dessas cadeias,

$$T\mathbf{u}_1 = \lambda\mathbf{u}_1, \qquad T\mathbf{u}_j = \lambda\mathbf{u}_j + \mathbf{u}_{j-1} \quad (j \geq 2).$$

Logo alguma potência de $T-\lambda\,\mathrm{I}$ anula W_λ, e portanto $W_\lambda \subseteq E^{\infty}_{\lambda}$.

Como $V = \bigoplus_i W_{\lambda_i}$ e também $V = \bigoplus_i E^{\infty}_{\lambda_i}(T)$, com $W_{\lambda_i} \subseteq E^{\infty}_{\lambda_i}(T)$, segue que $W_{\lambda_i} = E^{\infty}_{\lambda_i}(T)$ para todo i. Em cada uma dessas parcelas, subtrair $\lambda_i\,\mathrm{I}$ reduz o problema ao caso nilpotente. A fórmula do item 2, já provada nesse caso, determina o número de blocos de cada tamanho. Portanto, a forma de Jordan é única a menos da ordem dos blocos.

■

Corolário 10.17.
Todo operador linear em um espaço vetorial de dimensão finita sobre um corpo $\overline{\mathbb{K}}$ algebricamente fechado possui uma forma de Jordan.

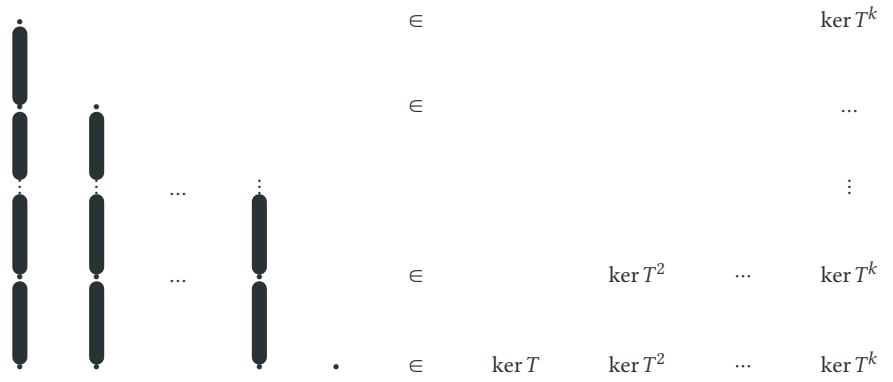
Os núcleos contam blocos
Para um autovalor λ, o crescimento das dimensões de $\ker(T-\lambda I)^k$ revela os tamanhos dos blocos. A diferença entre duas dimensões consecutivas conta quantos blocos têm comprimento pelo menos k.

10.3 Cálculo da Forma de Jordan

O cálculo separa-se em duas etapas:

1. determinar, para cada autovalor, o diagrama que fixa os tamanhos dos blocos;
2. construir uma base de Jordan compatível com esses blocos.

Cálculo dos Diagramas de Jordan Dada uma transformação T cuja matriz numa certa base é A. Considere o diagrama de Jordan correspondente a um autovalor λ de T. Como já observamos, a linha inferior desse diagrama consiste numa base para $\ker(T-\lambda \mathrm{I})$ e assim o número de vetores na linha inferior é $\dim \ker(T-\lambda \mathrm{I})$. De modo análogo temos que as duas linhas inferiores formam uma base para $\ker(T-\lambda \mathrm{I})^2$ e assim sucessivamente.



Logo para determinarmos esse diagrama é suficiente calcularmos:

$$\dim \ker(T-\lambda \mathrm{I}), \dim \ker(T-\lambda \mathrm{I})^2 - \dim \ker(T-\lambda \mathrm{I}), \dim \ker(T-\lambda \mathrm{I})^3 - \dim \ker(T-\lambda \mathrm{I})^2, \ldots.$$

Cada um desses números nos fornece quantos vetores aparecem em cada linha do diagrama.

Cálculo da Matriz Mudança de Base Para calcularmos a matriz mudança de base M é suficiente resolvermos a equação linear:

$$AM = MJ.$$

O sistema linear resultante pode ser indeterminado, mas somente as soluções invertíveis M servem como matrizes de mudança de base.

Exemplo 10.18. Determine a forma de Jordan de

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

O polinômio característico de A é $(x-2)^3$ e, assim, seu único autovalor é 2. Logo

$$A - 2\mathrm{I} = \begin{bmatrix} -1 & -\frac{1}{2} & 0 \\ 2 & 1 & 0 \\ 3 & \frac{3}{2} & 0 \end{bmatrix}.$$

Assim, $\dim \ker(A-2\mathrm{I}) = 2$. Como $(A-2\mathrm{I})^2 = 0$, temos $\dim \ker(A-2\mathrm{I})^2 = 3$. Logo o diagrama associado a essa matriz é:

E assim sua forma de Jordan, na convenção triangular superior, é

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

Para construir uma base de Jordan, tome $\mathbf{v}_1 = \mathbf{e}_1$ e $\mathbf{v}_2 = (A - 2\,\mathrm{I})\mathbf{v}_1 = (-1, 2, 3)^{\mathrm{t}}$. Como $(A - 2\,\mathrm{I})\mathbf{v}_2 = 0$ e $\mathbf{e}_3 \in \ker(A - 2\,\mathrm{I})$, ordenamos a cadeia como $(\mathbf{v}_2, \mathbf{v}_1)$ e obtemos a matriz

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

é invertível e satisfaz

$$M^{-1}AM = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

◁

Exemplo 10.19. Ache a forma de Jordan para

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

Nesse caso o polinômio característico é: $(x - 3)^4$ e assim o único autovalor é 3.

Temos que $\dim\ker(C - 3\,\mathrm{I}) = 1$, $\dim\ker(C - 3\,\mathrm{I})^2 = 2$, $\dim\ker(C - 3\,\mathrm{I})^3 = 3$ e $\dim\ker(C - 3\,\mathrm{I})^4 = 4$.

Logo o diagrama associado a essa matriz é:

E assim sua forma de Jordan é

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

◁

Exemplo 10.20. Ache a forma de Jordan para

$$B = \begin{bmatrix} 7 & 4 & 4 & -4 & 0 \\ -2 & 1 & -2 & 2 & 0 \\ -5 & 0 & -4 & 6 & 1 \\ -2 & 0 & -3 & 6 & 1 \\ 1 & -1 & 2 & -2 & 2 \end{bmatrix}$$

Nesse caso o polinômio característico é: $(x-3)^2(x-2)^3$ e assim seus autovalores são: $3, 2$.

Temos que $\dim \ker (B - 3\,\mathrm{I}) = 2$ e $\dim \ker (B - 3\,\mathrm{I})^2 = 2$ e que

$\dim \ker (B - 2\,\mathrm{I}) = 1$, $\dim \ker (B - 2\,\mathrm{I})^2 = 2$ e $\dim \ker (B - 2\,\mathrm{I})^3 = 3$.

Logo o diagrama de Jordan é

E assim a forma de Jordan é:

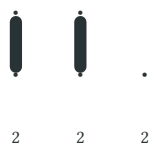
$$\begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}.$$

◁

Exemplo 10.21.

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

Nesse caso $(x-2)^5$, logo 2 é autovalor. Também temos que: $\dim \ker (C - 2\,\mathrm{I}) = 3$ e $\dim \ker (C - 2\,\mathrm{I})^2 = 5$ e logo o diagrama para C é



e consequentemente a forma de Jordan de C é:

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

◁

Exemplo 10.22. Suponha $A \in \mathscr{M}_{6,6}(\mathbb{R})$ com polinômio minimal

$$(x-1)^2(x+1)^2$$

Encontre todas as formas normais de Jordan possíveis e não semelhantes para A.

Nenhum dos blocos de Jordan pode ter tamanho maior que 2, pois isso produziria no polinômio minimal um fator $(x \pm 1)^3$. Como o polinômio minimal contém os fatores $(x \pm 1)^2$, deve haver pelo menos um bloco de tamanho 2 associado a cada um desses autovalores.

Assim, necessariamente aparecem os blocos $J_2(1)$ e $J_2(-1)$. Restam duas dimensões, que podem ser distribuídas entre blocos de tamanho 1 ou usadas para repetir um dos blocos de tamanho 2. Portanto, a menos da ordem dos blocos, há exatamente cinco possibilidades:

$$\begin{aligned}
&J_2(1) \oplus J_2(1) \oplus J_2(-1),\\
&J_2(1) \oplus [1] \oplus [1] \oplus J_2(-1),\\
&J_2(1) \oplus J_2(-1) \oplus J_2(-1),\\
&J_2(1) \oplus J_2(-1) \oplus [-1] \oplus [-1],\\
&J_2(1) \oplus J_2(-1) \oplus [1] \oplus [-1].
\end{aligned}$$

◁

Exercícios

Ex. 10.1 — Seja $A \in M_6(\mathbb{R})$ tal que $A^4 - 8A^2 + 16I_6 = 0$. Quais são as possíveis formas de Jordan não semelhantes para A?

Ex. 10.2 — Seja $T : \mathbb{R}_n[x] \to \mathbb{R}_n[x]$ dado por $T(p(x)) = p(x+1)$.

1. Determine a forma de Jordan de T.
2. Para $n = 4$, encontre uma base $\underline{\mathrm{B}}$ de $\mathbb{R}_4[x]$ tal que $[T]_{\underline{\mathrm{B}}}^{\underline{\mathrm{B}}} = J$.

Ex. 10.3 — Sejam as matrizes

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

$$C = \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \qquad D = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 3 & 0 & 2 & 0 \\ 4 & -5 & 3 & 2 \end{bmatrix},$$

veja que $c_A = c_B = c_C = c_D = (x-2)^4$. Determine a forma de Jordan de cada uma.

Ex. 10.4 — Dados um polinômio $f(x) = (x-\lambda_1)^{d_1}(x-\lambda_2)^{d_2}\dots(x-\lambda_k)^{d_k}$ e um espaço vetorial V com $\dim V = n = d_1 + d_2 + \cdots + d_k$. Quantas formas de Jordan, a menos de permutação dos blocos, podem ocorrer para operadores $T : V \to V$ com $c_T = f$?

Ex. 10.5 — Ache todas as formas de Jordan possíveis para uma transformação linear com polinômio característico $c_T = (x-2)^3(x-1)^2(x-5)$. Ache o polinômio minimal correspondente a cada uma dessas formas de Jordan.

Ex. 10.6 — Considere a transformação linear $T : \mathbb{R}^4 \to \mathbb{R}^4$ dada por

$$T(x,y,z,w) = (3z+w, 2x+2y-4z+2w, z+w, 3w-z)$$

1. Encontre a base e a forma de Jordan de T.

2. Descreva todos os subespaços T-invariantes.

Ex. 10.7 — Seja T um operador linear no $\mathbb{K}$-e.v. V de dimensão finita. Suponha que $m_T = (x-\lambda_1)^{m_1}\dots(x-\lambda_k)^{m_k}$, mostre que existe um operador diagonalizável $D \in \mathrm{Hom}(V,V)$ e um operador nilpotente $N \in \mathrm{Hom}(V,V)$ tal que

1. $T = D + N$,

2. $DN = ND$.

   Prove ainda que os operadores D e N são univocamente determinados por 1. e 2. e cada um deles é um polinômio em T.

Ex. 10.8 — A seguinte afirmação é falsa ou verdadeira: se $A \in M_n(\mathbb{C})$ é uma matriz com entradas complexas de ordem n é tal que $A^4 = I_n$ então

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

pode ser um bloco de Jordan de A.

Ex. 10.9 — Seja $T : \mathbb{R}^6 \to \mathbb{R}^6$ um operador linear com polinômios característico e minimal dados, respectivamente, por $c_T = (x-2)^4(x-1)^2$ e $m_T = (x-2)^2(x-1)^2$. Além disso suponha que $\dim\ker(T - 2\,\mathrm{I}) = 2$. Nestas condições encontre a forma de Jordan de T. Com os polinômios característico e minimal acima é possível supormos que $\dim\ker(T - 2\,\mathrm{I}) = 1$?

Ex. 10.10 —

1. Seja $A \in M_n(\mathbb{C})$ uma matriz complexa invertível. Defina suas partes *real* e *imaginária* $R, S \in M_n(\mathbb{R})$ como sendo

   $$R = \Re(A) \text{ e } S = \Im(A)$$

   de forma que $A = R + iS$. Mostre que existe $\lambda_0 \in \mathbb{R}$ tal que $R + \lambda_0 S$ é invertível.

2. Deduza que se duas matrizes reais $A, B \in M_n(\mathbb{R})$ são semelhantes em $M_n(\mathbb{C})$ então elas são semelhantes em $M_n(\mathbb{R})$.

Ex. 10.11 — Seja $A \in M_n(\mathbb{R})$ uma matriz real satisfazendo a seguinte condição $A^3 = I_n$. Mostre que $\operatorname{tr}(A) \in \mathbb{Z}$.

Ex. 10.12 — Sejam N_1 e N_2 matrizes de ordem 6 nilpotentes. Suponha que elas têm o mesmo polinômio minimal e o mesmo posto. Prove que elas são semelhantes. Mostre que o mesmo resultado não é verdadeiro para matrizes de ordem 7.

Ex. 10.13 — Dê a forma de Jordan de um operador linear $T : \mathbb{R}^7 \to \mathbb{R}^7$ com polinômio característico $c_T(x) = (x-1)^2(x-2)^4(x-3)$ e tal que $\dim \ker(T-2\,\mathrm{I}) = 2$, $\dim \ker(T-\mathrm{I}) = 1$ e $\ker(T-2\,\mathrm{I})^3 \neq \ker(T-2\,\mathrm{I})^2$.

Ex. 10.14 — Seja $N \in M_n(\mathbb{R})$ uma matriz nilpotente tal que $\dim \ker(N) = k$, $0 < k < n$.

1. Mostre que $\dim \ker(N^l) \leq kl$, para todo $l \geq 1$.

2. Prove que $n \leq kr$, onde r é o grau do polinômio minimal de N.

10.4 Forma de Jordan Real

Pares conjugados e blocos reais
Os autovalores não reais de uma matriz real aparecem em pares conjugados. Ao reunir as cadeias correspondentes, obtemos blocos reais construídos com matrizes 2×2. Essas matrizes descrevem a parte de rotação e dilatação, e as demais entradas registram as cadeias de Jordan.

Seja V um espaço vetorial sobre os reais e seja A a matriz do operador $T : V \to V$. Sobre os reais, o polinômio característico $c_A(x)$ de A pode ser fatorado como produto de termos lineares e termos de grau dois que não possuem raízes em $\mathbb{R}$:

$$c_A(x) = (x-c_1)^{m_1} \cdots (x-c_k)^{m_k}\big((x-a_1)^2 + b_1^2\big)^{n_1} \cdots \big((x-a_l)^2 + b_l^2\big)^{n_l}$$

que, por outro lado, pode ser fatorado como

$$c_A(x) = (x-c_1)^{m_1} \cdots (x-c_k)^{m_k}(x-\alpha_1)^{n_1}(x-\overline{\alpha}_1)^{n_1} \cdots (x-\alpha_l)^{n_l}(x-\overline{\alpha}_l)^{n_l}$$

sobre $\mathbb{C}$ com $\alpha_j = a_j + b_j i$ e $\overline{\alpha_j} = a_j - b_j i$.

Partiremos da forma de Jordan da complexificação de T e reuniremos os blocos associados a cada par de raízes conjugadas. A descomplexificação transforma esses pares em blocos reais correspondentes aos fatores quadráticos irredutíveis de $c_T(x)$.

Começamos estendendo a noção de conjugação para matrizes.

> **Definição 10.23.** Sejam $A \in \mathscr{M}_{n,n}(\mathbb{C})$ e $\mathbf{v} \in \mathbb{C}^n$ um vetor qualquer. Definimos $\overline{A} \in \mathscr{M}_{n,n}(\mathbb{C})$ como a matriz obtida ao se tomar o conjugado em cada uma das entradas de A e $\overline{\mathbf{v}} \in \mathbb{C}^n$ como o vetor obtido ao se tomar o conjugado em cada uma das coordenadas de $\mathbf{v}$.

É imediato provar que $\overline{A + \lambda B} = \overline{A} + \overline{\lambda}\overline{B}$, $\overline{AB} = \overline{A}\,\overline{B}$ e $\overline{A\mathbf{v}} = \overline{A}\overline{\mathbf{v}}$ para quaisquer matrizes $A, B \in \mathscr{M}_{n,n}(\mathbb{C})$, $\lambda \in \mathbb{C}$ e $\mathbf{v} \in \mathbb{C}^n$.

As próximas proposições relacionam os autovalores, os autoespaços invariantes, etc. para um operador linear T e sua complexificação $T^{\mathbb{C}}$.

Proposição 10.24.
Sejam V um espaço vetorial real de dimensão finita, $T : V \to V$ um operador linear e $T^{\mathbb{C}}$ sua complexificação. Então:

a *os polinômios característicos de T e $T^{\mathbb{C}}$ são iguais;*

b *se λ é um autovalor de $T^{\mathbb{C}}$, então $\overline{\lambda}$ é também um autovalor de $T^{\mathbb{C}}$;*

c *as multiplicidades algébricas dos autovalores λ e $\overline{\lambda}$ são iguais;*

d *se W' é um subespaço fechado por conjugação, então W' possui uma base formada por vetores reais. Um vetor $\mathbf{v}$ é dito **real** se $\overline{\mathbf{v}} = \mathbf{v}$.*

Demonstração. a As matrizes de T e de sua complexificação $T^{\mathbb{C}}$, na mesma base real considerada sobre $\mathbb{C}$, têm as mesmas entradas.

b Sejam λ um autovalor de $T^{\mathbb{C}}$ e $c_T(z)$ o polinômio característico de $T^{\mathbb{C}}$. Como $c_T(z)$ também é o polinômio característico de T, os coeficientes de $c_T(z)$ são reais.

Tomando o conjugado na equação $c_T(\lambda) = 0$, obtemos $c_T(\overline{\lambda}) = 0$, o que mostra que $\overline{\lambda}$ também é uma raiz do polinômio característico de $T^{\mathbb{C}}$.

c Se λ é raiz de multiplicidade d do polinômio característico, então

$$c_T'(\lambda) = \ldots = c_T^{(d-1)}(\lambda) = 0 \text{ e } c_T^{(d)}(\lambda) \neq 0$$

tomando o conjugado em cada uma dessas equações

$$c_T'(\overline{\lambda}) = \ldots = c_T^{(d-1)}(\overline{\lambda}) = 0 \text{ e } c_T^{(d)}(\overline{\lambda}) \neq 0,$$

mostrando que $\overline{\lambda}$ também tem multiplicidade d.

d Seja $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ uma base de W', com $\mathbf{w}_j = \mathbf{u}_j + i\mathbf{v}_j$ e $\mathbf{u}_j, \mathbf{v}_j \in V$. Como W' é fechado por conjugação,

$$\mathbf{u}_j = \frac{\mathbf{w}_j + \overline{\mathbf{w}_j}}{2} \quad \text{e} \quad \mathbf{v}_j = \frac{\mathbf{w}_j - \overline{\mathbf{w}_j}}{2i}$$

pertencem a W', identificando V com o subespaço dos vetores reais de $V^{\mathbb{C}}$.

Assim, o conjunto $S = \{\mathbf{u}_1, \mathbf{v}_1, \ldots, \mathbf{u}_k, \mathbf{v}_k\}$ é um conjunto de vetores reais que gera W'. Uma base formada de vetores reais é obtida tomando um subconjunto de S com k elementos linearmente independentes em $V^{\mathbb{C}}$.

■

Lema 10.25.
Sejam $T : V \to V$ um operador linear e $T^{\mathbb{C}}$ sua complexificação. Se o subespaço $W' \subset V^{\mathbb{C}}$ possui uma base formada por vetores reais, então ele é a complexificação de um subespaço $W \subset V$.

Demonstração. Todo vetor de $\mathbf{w} \in W'$ pode ser escrito como $\mathbf{w} = \mathbf{u} + i\mathbf{v}$, sendo $\mathbf{u}$ e $\mathbf{v}$ vetores reais. Escrevendo $\mathbf{u}$ e $\mathbf{v}$ em termos dos vetores da base real, segue imediatamente que W' é a complexificação do espaço real W gerado pelos vetores dessa base. ■

Proposição 10.26.
Seja V um espaço vetorial real de dimensão finita, seja $T \in \mathrm{Hom}(V,V)$ e considere sua complexificação $T^{\mathbb{C}} : V^{\mathbb{C}} \to V^{\mathbb{C}}$. Se $\alpha_1, \ldots, \alpha_k$ são os autovalores reais e $\lambda_1, \overline{\lambda_1}, \ldots, \lambda_m, \overline{\lambda_m}$ são os pares de autovalores não reais, então

$$V^{\mathbb{C}} = \bigoplus_{j=1}^{k} W_j \oplus \bigoplus_{j=1}^{m} (Z_j \oplus \overline{Z_j}),$$

onde $W_j = \ker (T^{\mathbb{C}} - \alpha_j \mathrm{I})^{q_j}$ e $Z_j = \ker (T^{\mathbb{C}} - \lambda_j \mathrm{I})^{r_j}$. A conjugação leva Z_j isomorficamente em $\overline{Z_j}$. Se $\{\mathbf{z}_1, \ldots, \mathbf{z}_s\}$ é uma base de Z_j e $\mathbf{z}_h = \mathbf{u}_h + i\mathbf{v}_h$, então

$$\{\mathbf{u}_1, \mathbf{v}_1, \ldots, \mathbf{u}_s, \mathbf{v}_s\}$$

é uma base real do subespaço real cuja complexificação é $Z_j \oplus \overline{Z_j}$.

Demonstração. O Teorema da Decomposição em Autoespaços Generalizados, aplicado a $T^{\mathbb{C}}$, fornece a decomposição indicada. Além disso,

$$\overline{(T^{\mathbb{C}} - \lambda \mathrm{I})^j \mathbf{z}} = (T^{\mathbb{C}} - \overline{\lambda}\, \mathrm{I})^j \overline{\mathbf{z}},$$

de modo que a conjugação identifica Z_j com $\overline{Z_j}$. Por fim, $\mathbf{z}_h$ e $\overline{\mathbf{z}_h}$ geram o mesmo espaço complexo que suas partes real e imaginária, $\mathbf{u}_h$ e $\mathbf{v}_h$. Isso prova a última afirmação.

■

Teorema 10.27 (Forma de Jordan Real).
Se V é um espaço vetorial real e $T : V \to V$ é um operador linear, existe uma base $\underline{B}$ de V em relação à qual a matriz de T tem, ao longo da diagonal, blocos de Jordan (correspondentes aos autovalores reais) e blocos de Jordan aumentados (correspondentes aos autovalores complexos). A soma das ordens dos blocos de Jordan correspondentes a um mesmo autovalor λ é igual à multiplicidade algébrica de λ, se $\lambda \in \mathbb{R}$ e é igual ao dobro da multiplicidade algébrica de λ se $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Os blocos de Jordan aumentados são da forma

$$J_{a,b} = \begin{bmatrix} C_{a,b} & \mathrm{I}_2 & 0 & \cdots & 0 \\ 0 & C_{a,b} & \mathrm{I}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & C_{a,b} & \mathrm{I}_2 \\ 0 & \cdots & \cdots & 0 & C_{a,b} \end{bmatrix}$$

onde $C_{a,b} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$, sendo $a + ib$ um autovalor complexo de $T^{\mathbb{C}}$ e I_2 a matriz identidade 2×2.

Demonstração. Sejam $\alpha_1, \dots, \alpha_k$ os autovalores reais de T e $\lambda_1, \overline{\lambda_1}, \dots, \lambda_m, \overline{\lambda_m}$ os pares de autovalores não reais de $T^{\mathbb{C}}$. Então $V^{\mathbb{C}}$ admite a decomposição

$$V^{\mathbb{C}} = W_1 \oplus \cdots \oplus W_k \oplus Z_1 \oplus \overline{Z_1} \oplus \cdots \oplus Z_m \oplus \overline{Z_m},$$

onde os núcleos são tomados para $T^{\mathbb{C}}$.

Para cada $\alpha_j \in \mathbb{R}$, o operador $T - \alpha_j \mathrm{I}$ restrito ao subespaço real correspondente é nilpotente, e sua base de Jordan pode ser construída como antes.

Suponhamos agora que $T^{\mathbb{C}}$ possua um autovalor $\lambda \in \mathbb{C}\setminus\mathbb{R}$.

Fixe um par $\lambda, \overline{\lambda}$ e escreva $Z_\lambda = \ker(T^{\mathbb{C}} - \lambda \mathrm{I})^r$ e $Z_{\overline{\lambda}} = \ker(T^{\mathbb{C}} - \overline{\lambda}\,\mathrm{I})^r$. Então $\{\mathbf{u}_1,\ \mathbf{v}_1,\ \mathbf{u}_2,\ \mathbf{v}_2,\ \dots,\ \mathbf{u}_k,\ \mathbf{v}_k\}$ é uma base real do subespaço cuja complexificação é $Z_\lambda \oplus Z_{\overline{\lambda}}$.

Finalmente, considere uma cadeia de Jordan $\mathbf{w}_j = \mathbf{u}_j + i\mathbf{v}_j$, $1 \le j \le r$, com $\mathbf{w}_0 = \vec{0}$ e $T^{\mathbb{C}}\mathbf{w}_j = \lambda \mathbf{w}_j + \mathbf{w}_{j-1}$, onde $\lambda = a + ib \in \mathbb{C} \setminus \mathbb{R}$. Então

$$T\mathbf{u}_j + iT\mathbf{v}_j = (a\mathbf{u}_j - b\mathbf{v}_j + \mathbf{u}_{j-1}) + i(b\mathbf{u}_j + a\mathbf{v}_j + \mathbf{v}_{j-1})$$

e logo

$$\begin{aligned} T\mathbf{u}_j &= a\mathbf{u}_j - b\mathbf{v}_j + \mathbf{u}_{j-1}, \\ T\mathbf{v}_j &= b\mathbf{u}_j + a\mathbf{v}_j + \mathbf{v}_{j-1}. \end{aligned}$$

de onde segue que, na base real $\{\mathbf{u}_1,\ \mathbf{v}_1,\ \mathbf{u}_2,\ \mathbf{v}_2,\ \dots,\ \mathbf{u}_r,\ \mathbf{v}_r\}$, o operador T é representado por um bloco aumentado da forma descrita no enunciado. Repetindo o argumento para todas as cadeias e concatenando as bases, obtemos a forma de Jordan real.

■

Exemplo 10.28. Seja

$$A = \begin{bmatrix} 0 & -6 & 10 & -5 \\ -1 & -4 & 10 & -5 \\ -1 & -3 & 8 & -4 \\ -2 & -2 & 9 & -4 \end{bmatrix} \in \mathscr{M}_{4,4}(\mathbb{R})$$

O polinômio característico de A é $c_A(x) = (1 + x^2)^2 = (x - i)^2(x + i)^2$ e coincide com o polinômio minimal de A. Os autovalores sobre os complexos são $i, -i$. Consequentemente temos dois blocos de Jordan, um associado a cada autovalor. Logo a forma normal de Jordan obtida complexificando A é

$$\left[\begin{array}{cc|cc} -i & 1 & 0 & 0 \\ 0 & -i & 0 & 0 \\ \hline 0 & 0 & i & 1 \\ 0 & 0 & 0 & i \end{array}\right]$$

e pelo Teorema 10.27 temos que a forma normal de Jordan sobre $\mathbb{R}$ é:

$$\left[\begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{array}\right]$$

◁

Exercícios

Ex. 10.15 — Encontre uma base e a forma de Jordan real para o operador $T : \mathbb{R}^4 \to \mathbb{R}^4$ tal que

$$[T]_{\underline{B}}^{\underline{B}} = \begin{bmatrix} 2 & 0 & 0 & -2 \\ 5 & 2 & -2 & 2 \\ 3 & 1 & 0 & -3 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Ex. 10.16 — Seja $A \in M_n(\mathbb{R})$, tal que $A^2 + I_n = 0$. Prove que $n = 2k$ e que A é semelhante à matriz

$$B = \begin{bmatrix} 0 & -I_k \\ I_k & 0 \end{bmatrix},$$

onde $I_k \in M_k(\mathbb{R})$ é a matriz identidade.

11
Forma Canônica Racional

Embora o polinômio característico de um operador $T : V \to V$ se decomponha como produto de potências de polinômios irredutíveis em $\mathbb{K}[x]$, esses fatores não precisam ser lineares. Em particular, o polinômio característico pode não se decompor sobre $\mathbb{K}$, e os autovalores de T podem não pertencer ao corpo de escalares. A forma canônica racional contorna essa dificuldade: ela existe sobre o corpo original e substitui os blocos de Jordan por matrizes companheiras.

Seja $T : V \to V$ um operador linear em um espaço vetorial de dimensão finita. Mostraremos que existe uma base $\underline{\mathrm{B}}$ de V, na qual $[T]_{\underline{\mathrm{B}}}$ é uma matriz na forma canônica racional:

$$\begin{bmatrix} C_{p_1} & 0 & \dots & 0 \\ 0 & C_{p_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & C_{p_r} \end{bmatrix} \qquad \text{com} \qquad C_{p_i} = \begin{bmatrix} 0 & 0 & 0 & 0 & -a_{i0} \\ 1 & 0 & 0 & 0 & -a_{i1} \\ 0 & 1 & 0 & 0 & -a_{i2} \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & 1 & -a_{i,d_i-1} \end{bmatrix}$$

onde C_{p_i} é a matriz companheira do polinômio mônico $p_i(x) = x^{d_i} + a_{i,d_i-1}x^{d_i-1} + \cdots + a_{i1}x + a_{i0} \in \mathbb{K}[x]$. Os graus d_i podem variar de um bloco para outro.

Nesse capítulo

11.1 Decomposição Cíclica

Fixe um operador $T : V \to V$. A interseção de subespaços T-invariantes continua sendo T-invariante. Portanto, para cada subconjunto $A \subseteq V$ existe um menor subespaço T-invariante que contém A: a interseção de todos os que contêm esse conjunto. Denotaremos esse espaço por Z_A^T. Quando $A = \{\mathbf{v}\}$, escreveremos $Z_{\mathbf{v}}^T$, ou simplesmente $Z_{\mathbf{v}}$ se o operador estiver fixado.

Um vetor gera um módulo
O subespaço cíclico gerado por v contém tudo o que pode ser obtido aplicando polinômios em T a esse vetor. Assim, estudar T nesse subespaço equivale a estudar um módulo cíclico sobre $\mathbb{K}[x]$.

Definição 11.1 — Subespaço cíclico. O **subespaço T-cíclico gerado por $\mathbf{v}$** é o menor subespaço T-invariante que contém $\mathbf{v}$; denotamo-lo por $Z_{\mathbf{v}}$.

Exemplo 11.2. Seja a transformação linear $T\colon \mathbb{K}^3 \to \mathbb{K}^3$ definida por

$$T(x,\ y,\ z) = (y + 2z,\ -2z,\ x).$$

Considere $\mathbf{v} = (1, 0, 0)$. Temos

$$T\mathbf{v} = (0, 0, 1), \qquad T^2\mathbf{v} = (2, -2, 0).$$

Esses três vetores são linearmente independentes; portanto, $Z_{\mathbf{v}} = \mathbb{K}^3$. Além disso, $T^3\mathbf{v} = -2\mathbf{v} + 2T\mathbf{v}$.

◁

O subespaço $Z_{\mathbf{v}}$ pode ser caracterizado do seguinte modo.

Teorema 11.3.
Sejam V um espaço vetorial de dimensão finita sobre um corpo $\mathbb{K}$ e $T\colon V \to V$ um operador linear. Então, para todo $\mathbf{v} \in V$,

$$Z_{\mathbf{v}} = \{p(T)(\mathbf{v});\ p \in \mathbb{K}[x]\}.$$

A matriz companheira
Quando um único vetor gera todo o espaço, a base $(v, Tv, \dots, T^{n-1}v)$ transforma T numa matriz companheira. Os coeficientes da última coluna registram exatamente a relação polinomial mínima satisfeita por v.

Demonstração. Ponha

$$W = \{p(T)\mathbf{v} \ :\ p \in \mathbb{K}[x]\}.$$

Se $p, q \in \mathbb{K}[x]$ e $a, b \in \mathbb{K}$, então

$$a\,p(T)\mathbf{v} + b\,q(T)\mathbf{v} = (ap + bq)(T)\mathbf{v},$$

de modo que W é um subespaço de V. Além disso, $\mathbf{v} = 1(T)\mathbf{v}$ pertence a W e

$$T\big(p(T)\mathbf{v}\big) = (xp)(T)\mathbf{v} \in W;$$

logo W é T-invariante.

Resta verificar a minimalidade. Se U é um subespaço T-invariante que contém $\mathbf{v}$, então $T^j\mathbf{v} \in U$ para todo $j \geq 0$. Como U é um subespaço, ele contém toda combinação linear desses vetores e, portanto, $p(T)\mathbf{v} \in U$ para todo $p \in \mathbb{K}[x]$. Assim, $W \subseteq U$. Concluímos que W é o menor subespaço T-invariante que contém $\mathbf{v}$, isto é, $W = Z_{\mathbf{v}}$. ■

Para construir uma base de $Z_{\mathbf{v}}$, considere a sequência

$$\mathbf{v},\ T(\mathbf{v}),\ T^2(\mathbf{v}),\ \dots,\ T^r(\mathbf{v}),\ \dots$$

de elementos de $Z_{\mathbf{v}}$. Como V tem dimensão finita, existe um menor inteiro positivo k para o qual $T^k(\mathbf{v})$ é combinação linear dos termos anteriores, digamos

$$T^k(\mathbf{v}) = \lambda_0\mathbf{v} + \lambda_1 T(\mathbf{v}) + \cdots + \lambda_{k-1}T^{k-1}(\mathbf{v})$$

Pela minimalidade de k, os vetores $\mathbf{v}, T\mathbf{v}, \dots, T^{k-1}\mathbf{v}$ são linearmente independentes.

Escrevendo $a_j = -\lambda_j$ para $0 \leq j < k$, obtemos o polinômio

$$m_{\mathbf{v}} = a_0 + a_1 x + \cdots + a_{k-1}x^{k-1} + x^k$$

Ele é o polinômio mônico de menor grau tal que $m_{\mathbf{v}}(T)(\mathbf{v}) = \vec{0}$.

Definição 11.4. Seja $\mathbf{v} \in V$. Então o conjunto

$$\{p(x) \in \mathbb{K}[x] \,:\, p(T)(\mathbf{v}) = 0\}$$

é um ideal de $\mathbb{K}[x]$ e seu gerador mônico $m_{\mathbf{v}}$ é dito T-**aniquilador** de $\mathbf{v}$.

O T-**aniquilador** de $\mathbf{v}$ também será denotado por $\mathrm{Aniq}(\mathbf{v}; T)$.

Exemplo 11.5. Para o operador do Exemplo 11.2, o vetor $\mathbf{v} = (1, 0, 0)$ é cíclico e satisfaz $T^3\mathbf{v} - 2T\mathbf{v} + 2\mathbf{v} = 0$. Logo seu T-aniquilador, que também é o polinômio minimal e o característico de T, é

$$m_{\mathbf{v}}(x) = x^3 - 2x + 2.$$

◁

Com a notação acima, temos o seguinte resultado.

Teorema 11.6.
Sejam V um espaço vetorial de dimensão finita e $T : V \to V$ linear. Se $\mathbf{v} \in V$ tem T-aniquilador

$$m_{\mathbf{v}} = a_0 + a_1 x + \cdots + a_{k-1}x^{k-1} + x^k$$

então o conjunto

$$\underline{B}_{\mathbf{v}} = \{\mathbf{v},\ T(\mathbf{v}),\ \ldots,\ T^{k-1}(\mathbf{v})\}$$

é uma base de $Z_{\mathbf{v}}$ e, portanto, $\dim Z_{\mathbf{v}} = \operatorname{grau} m_{\mathbf{v}}$. Além disso, se $T\restriction_{Z_{\mathbf{v}}} : Z_{\mathbf{v}} \to Z_{\mathbf{v}}$, é a transformação linear induzida no T-subespaço invariante $Z_{\mathbf{v}}$, então a matriz de $T\restriction_{Z_{\mathbf{v}}}$ em relação à base ordenada $\underline{B}_{\mathbf{v}}$ é

$$C_{m_{\mathbf{v}}} = \begin{bmatrix} 0 & 0 & 0 & 0 & -a_0 \\ 1 & 0 & 0 & 0 & -a_1 \\ 0 & 1 & 0 & 0 & -a_2 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & 1 & -a_{k-1} \end{bmatrix}$$

Finalmente, o polinômio minimal de $T\restriction_{Z_{\mathbf{v}}}$ é $m_{\mathbf{v}}$.

Demonstração. Se uma combinação linear de $\mathbf{v}, T\mathbf{v}, \ldots, T^{k-1}\mathbf{v}$ fosse nula, obteríamos um polinômio não nulo de grau menor que k que anula $\mathbf{v}$. Isso contradiria a minimalidade de $m_{\mathbf{v}}$. Portanto, $\underline{\mathrm{B}}_{\mathbf{v}}$ é linearmente independente.

A relação $m_{\mathbf{v}}(T)\mathbf{v} = 0$ fornece

$$T^k\mathbf{v} = -a_0\mathbf{v} - a_1 T\mathbf{v} - \cdots - a_{k-1}T^{k-1}\mathbf{v}.$$

Consequentemente, T leva cada vetor de $\underline{\mathrm{B}}_{\mathbf{v}}$ ao subespaço $\langle \underline{\mathrm{B}}_{\mathbf{v}} \rangle$; esse subespaço é, portanto, T-invariante. Como contém $\mathbf{v}$, ele contém $Z_{\mathbf{v}}$. A inclusão inversa decorre de $T^j\mathbf{v} \in Z_{\mathbf{v}}$ para $0 \le j < k$. Logo $Z_{\mathbf{v}} = \langle \underline{\mathrm{B}}_{\mathbf{v}} \rangle$, e $\underline{\mathrm{B}}_{\mathbf{v}}$ é uma base desse espaço.

Nessa base,

$$\begin{aligned} T\restriction_{Z_\mathbf{v}}(\mathbf{v}) &= T(\mathbf{v}) \\ T\restriction_{Z_\mathbf{v}}[T(\mathbf{v})] &= T^2(\mathbf{v}) \\ &\vdots \\ T\restriction_{Z_\mathbf{v}}[T^{k-2}(\mathbf{v})] &= T^{k-1}(\mathbf{v}) \\ T\restriction_{Z_\mathbf{v}}[T^{k-1}(\mathbf{v})] &= T^k(\mathbf{v}) = -a_0\mathbf{v} - a_1T(\mathbf{v}) - \cdots - a_{k-1}T^{k-1}(\mathbf{v}) \end{aligned}$$

As primeiras $k-1$ colunas da matriz são, assim, as colunas de deslocamento, e a última é $(-a_0, -a_1, \ldots, -a_{k-1})^t$. Essa é precisamente a matriz companheira $C_{m_\mathbf{v}}$.

Finalmente, para todo $q \in \mathbb{K}[x]$,

$$q\big(T\restriction_{Z_\mathbf{v}}\big) = 0 \quad \iff \quad q(T)\mathbf{v} = 0.$$

A implicação da esquerda para a direita é obtida aplicando o operador a $\mathbf{v}$. Reciprocamente, se $q(T)\mathbf{v} = 0$, então, para todo $p(T)\mathbf{v} \in Z_\mathbf{v}$,

$$q(T)p(T)\mathbf{v} = p(T)q(T)\mathbf{v} = 0.$$

Portanto, $T\restriction_{Z_\mathbf{v}}$ e $\mathbf{v}$ têm o mesmo ideal de polinômios anuladores. Seus geradores mônicos coincidem, e o polinômio minimal da restrição é $m_\mathbf{v}$. ■

Exemplo 11.7. Para o operador do Exemplo 11.2, mostramos que $\mathbf{v} = (1, 0, 0)$ é cíclico e que $m_\mathbf{v}(x) = x^3 - 2x + 2$. Logo $m_T = m_\mathbf{v} = c_T$.

◁

Exemplo 11.8. Seja T o operador que na base canônica $\underline{\mathrm{B}}$ é representado pela matriz

$$[T]_{\underline{\mathrm{B}}} = \begin{bmatrix} -2 & 3 & 6 \\ -2 & 2 & 5 \\ -1 & 1 & 3 \end{bmatrix}$$

Vamos calcular o aniquilador de $\mathbf{v}_1 = [1, 1, 0]$. Observamos que $\mathbf{v}_2 = T\mathbf{v}_1 = [1, 0, 0]$, $\mathbf{v}_3 = T^2\mathbf{v}_1 = T\mathbf{v}_2 = [-2, -2, -1]$ e $\mathbf{v}_4 = T^3\mathbf{v}_1 = [-8, -5, -3]$. Os vetores $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ são linearmente independentes e formam uma base de $Z_{\mathbf{v}_1} = V$.

Temos para $\mathbf{v}_4$ a seguinte relação linear $-\mathbf{v}_1 + 3T\mathbf{v}_1 - 3T^2\mathbf{v}_1 + T^3\mathbf{v}_1 = \vec{0}$. Logo $m_\mathbf{v} = -1 + 3x - 3x^2 + x^3$. Nessa base temos Seja

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

a matriz cujas colunas são $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$. Então $P^{-1}[T]_{\underline{\mathrm{B}}}P = C_{m_{\mathbf{v}_1}}$ e, portanto,

$$\begin{bmatrix} -2 & 3 & 6 \\ -2 & 2 & 5 \\ -1 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -2 \\ 1 & 0 & -2 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & -3 \\ 0 & 1 & 3 \end{bmatrix}\begin{bmatrix} 0 & 1 & -2 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

◁

Definição 11.9 — Espaço cíclico. Um subespaço $W \subseteq V$ é T**-cíclico** se é T-invariante e admite uma base da forma $\{\mathbf{v}, T\mathbf{v}, \dots, T^m\mathbf{v}\}$. Essa família é uma **base cíclica** de W, e $\mathbf{v}$ é um **vetor** T**-cíclico** para W.

Pelo Teorema 11.6, o vetor $\mathbf{v}$ é cíclico para $Z_\mathbf{v}$, com base $\underline{\mathrm{B}}_\mathbf{v}$. A matriz $C_{m_\mathbf{v}}$ que representa a restrição de T nessa base é a **matriz companheira** do T-aniquilador $m_\mathbf{v}$.

Exemplo 11.10. Um operador nilpotente N com índice de nilpotência $n = \dim V$ sempre possui um vetor cíclico, conforme a Proposição 10.11.

◁

Reunindo essas ideias, decomporemos V em subespaços cíclicos. A concatenação de suas bases produzirá uma representação diagonal por blocos, cujos blocos são matrizes companheiras.

Teorema 11.11 (Decomposição Cíclica).
Sejam V um espaço vetorial não nulo de dimensão finita sobre $\mathbb{K}$ e $T : V \to V$ um operador linear. Existem vetores não nulos $\mathbf{v}_1, \dots, \mathbf{v}_k$ tais que

$$V = Z_{\mathbf{v}_1} \oplus \cdots \oplus Z_{\mathbf{v}_k}.$$

Se $q_i = m_{\mathbf{v}_i}$, a reunião das bases cíclicas dessas parcelas é uma base de V na qual

$$[T] = C_{q_1} \oplus \cdots \oplus C_{q_k}.$$

Em particular,

$$c_T = q_1 \cdots q_k.$$

Sem precisar de autovalores
A decomposição cíclica funciona sobre qualquer corpo e não exige que o polinômio característico se decomponha em fatores lineares. Essa é a principal vantagem da forma racional sobre a forma de Jordan.

Demonstração. Faremos indução sobre $n = \dim V$. Se $n = 1$, qualquer vetor não nulo gera V, e o resultado é imediato. Suponha $n > 1$ e que o teorema valha em dimensões menores.

Escolha $\mathbf{v}_1 \in V$ de modo que

$$m = \dim Z_{\mathbf{v}_1}$$

seja o maior valor de $\dim Z_\mathbf{v}$ quando $\mathbf{v}$ percorre V. Se $m = n$, então $V = Z_{\mathbf{v}_1}$ e nada mais há a provar. Suponha $m < n$. Construiremos um complemento T-invariante de $Z_{\mathbf{v}_1}$.

A base cíclica

$$\mathbf{v}_1, T\mathbf{v}_1, \dots, T^{m-1}\mathbf{v}_1$$

é linearmente independente. Podemos, portanto, prolongá-la a uma base de V e escolher $f \in V^*$ que satisfaça

$$f(T^j\mathbf{v}_1) = 0 \quad (0 \le j < m-1), \qquad f(T^{m-1}\mathbf{v}_1) = 1.$$

Defina $\tau : V \to \mathbb{K}^m$ por

$$\tau(\mathbf{w}) = \big(f(\mathbf{w}), f(T\mathbf{w}), \dots, f(T^{m-1}\mathbf{w})\big)^t.$$

Mostremos primeiro que $\tau\restriction_{Z_{\mathbf{v}_1}}$ é um isomorfismo. Na base cíclica acima, a entrada da linha $i+1$ e da coluna $j+1$ de sua matriz é $f(T^{i+j}\mathbf{v}_1)$. Essa entrada é zero se $i+j < m-1$

e é igual a 1 se $i + j = m - 1$. Assim, a matriz tem zeros de um lado da antidiagonal e entradas iguais a 1 na antidiagonal. Ao inverter a ordem das linhas, obtemos uma matriz triangular com diagonal formada por uns. Logo a matriz é invertível, como queríamos.

Agora provemos que $\ker \tau$ é T-invariante. Tome $\mathbf{w} \in \ker \tau$. Então

$$f(\mathbf{w}) = f(T\mathbf{w}) = \cdots = f(T^{m-1}\mathbf{w}) = 0.$$

Pela escolha de m, temos $\dim Z_{\mathbf{w}} \leq m$. Se $r = \operatorname{grau} m_{\mathbf{w}}$, então $r \leq m$ e a relação $m_{\mathbf{w}}(T)\mathbf{w} = 0$ escreve $T^r\mathbf{w}$ como combinação linear de $\mathbf{w}, T\mathbf{w}, \ldots, T^{r-1}\mathbf{w}$. Aplicando T^{m-r} a essa relação, concluímos que

$$T^m\mathbf{w} \in \langle \mathbf{w}, T\mathbf{w}, \ldots, T^{m-1}\mathbf{w} \rangle.$$

Portanto, $f(T^m\mathbf{w}) = 0$. Todas as coordenadas de $\tau(T\mathbf{w})$ são nulas, e assim $T\mathbf{w} \in \ker \tau$.

Como $\tau\restriction_{Z_{\mathbf{v}_1}}$ é injetiva, $Z_{\mathbf{v}_1} \cap \ker \tau = \{\vec{0}\}$. Ela também é sobrejetiva, de modo que $\operatorname{im} \tau = \mathbb{K}^m$ e $\dim \ker \tau = n - m$. Consequentemente,

$$\dim Z_{\mathbf{v}_1} + \dim \ker \tau = m + (n - m) = n,$$

e obtemos a soma direta

$$V = Z_{\mathbf{v}_1} \oplus \ker \tau.$$

Se $\ker \tau \neq \{\vec{0}\}$, sua dimensão é menor que n e a hipótese de indução, aplicada a $T\restriction_{\ker \tau}$, decompõe esse complemento em subespaços cíclicos. Juntando essa decomposição a $Z_{\mathbf{v}_1}$, obtemos a decomposição de V. A reunião das bases cíclicas fornece a matriz diagonal por blocos. Por fim, o polinômio característico de uma matriz diagonal por blocos é o produto dos polinômios característicos dos blocos, e o polinômio característico de C_{q_i} é q_i. ■

Exercícios

Ex. 11.1 — Sejam $T \in \operatorname{Hom}(V, V)$ e $\mathbf{u}, \mathbf{v} \in V$. Suponha que os subespaços cíclicos $Z_{\mathbf{u}}$ e $Z_{\mathbf{v}}$ tenham interseção nula.

1. Mostre que
$$m_{\mathbf{u}+\mathbf{v}} = \operatorname{mmc}(m_{\mathbf{u}}, m_{\mathbf{v}}).$$

2. Suponha, além disso, que $m_{\mathbf{u}}$ e $m_{\mathbf{v}}$ sejam coprimos. Prove que
$$Z_{\mathbf{u}+\mathbf{v}} = Z_{\mathbf{u}} \oplus Z_{\mathbf{v}}.$$

Sugestão: use uma identidade de Bézout para recuperar $\mathbf{u}$ e $\mathbf{v}$ a partir de $\mathbf{u} + \mathbf{v}$.

11.2 Forma Canônica Racional

Forma canônica sobre o corpo original
A forma racional classifica operadores por blocos companheiros associados a polinômios. Ela existe sem ampliar o corpo de escalares e é única quando os fatores invariantes são dispostos na ordem de divisibilidade adotada.

O Teorema da Decomposição Cíclica fornece uma soma direta de subespaços cíclicos, mas essa decomposição não é única: uma parcela pode admitir nova decomposição

em parcelas cíclicas menores. Para obter um representante canônico, ordenaremos os polinômios associados de modo que cada um seja divisível pelo seguinte. Esses polinômios serão os **fatores invariantes**; sua lista determinará a classe de similaridade.

Proposição 11.12.
Sejam V um espaço vetorial de dimensão finita sobre um corpo $\mathbb{K}$ e $T : V \to V$ um operador linear com polinômio minimal $m_T = p^t$, onde p é um polinômio irredutível sobre $\mathbb{K}$. Então existem vetores T-cíclicos $\mathbf{v}_1, \dots, \mathbf{v}_k$ e números inteiros positivos $n_1, \dots, n_k$ com cada $n_j \le t$ de modo que

1. $V = \bigoplus_{j=1}^{k} Z_{\mathbf{v}_j}$;
2. *o T-aniquilador de $\mathbf{v}_j$ é p^{n_j};*
3. $\dim V = (n_1 + \cdots + n_k)\,\text{grau}\, p$;
4. $t = n_1 \ge n_2 \ge \cdots \ge n_k \ge 1$;
5. *os inteiros $n_1, \dots, n_k$ e, portanto, os polinômios $p^{n_1}, \dots, p^{n_k}$ são determinados unicamente por T.*

Demonstração. Pelo Teorema 11.11, existem vetores não nulos $\mathbf{v}_1, \dots, \mathbf{v}_k$ tais que

$$V = \bigoplus_{i=1}^{k} Z_{\mathbf{v}_i}.$$

Escreva $q_i = m_{\mathbf{v}_i}$. Como $m_T(T)\mathbf{v}_i = 0$, a definição do aniquilador de $\mathbf{v}_i$ mostra que $q_i \mid m_T$. A hipótese $m_T = p^t$, com p irredutível, implica

$$q_i = p^{n_i}$$

para algum inteiro $1 \le n_i \le t$.

Um polinômio anula T se, e somente se, anula a restrição de T a cada parcela da soma direta. Portanto,

$$m_T = \operatorname{mmc}(q_1, \dots, q_k).$$

Como todos os q_i são potências de p, segue que $t = \max_i n_i$. Reordenando as parcelas, podemos supor

$$t = n_1 \ge n_2 \ge \cdots \ge n_k \ge 1.$$

Além disso, pelo Teorema 11.6,

$$\dim Z_{\mathbf{v}_i} = \text{grau}\, p^{n_i} = n_i\,\text{grau}\, p,$$

e a soma das dimensões fornece a terceira afirmação.

Resta provar que a sequência de expoentes independe da decomposição escolhida. Ponha $d = \text{grau}\, p$. Para $j \ge 0$, a comutatividade entre $p(T)^j$ e T fornece

$$p(T)^j\big(Z_{\mathbf{v}_i}\big) = Z_{p(T)^j \mathbf{v}_i}.$$

Se $j \ge n_i$, esse espaço é nulo. Se $j < n_i$, então, para $q \in \mathbb{K}[x]$,

$$q(T)p(T)^j\mathbf{v}_i = 0 \iff p^{n_i} \mid q p^j \iff p^{n_i - j} \mid q.$$

Logo o aniquilador de $p(T)^j\mathbf{v}_i$ é $p^{n_i - j}$ e

$$\dim p(T)^j\big(Z_{\mathbf{v}_i}\big) = d\max(n_i - j, 0).$$

Como cada $Z_{\mathbf{v}_i}$ é T-invariante, a aplicação $p(T)^j$ preserva a decomposição direta. Assim,

$$\operatorname{im} p(T)^j = \bigoplus_{i=1}^{k} p(T)^j\big(Z_{\mathbf{v}_i}\big)$$

e

$$\dim \operatorname{im} p(T)^j = d \sum_{i=1}^{k} \max(n_i - j, 0).$$

Para todo $j \geq 1$, obtemos

$$\dim \operatorname{im} p(T)^{j-1} - \dim \operatorname{im} p(T)^j = d\,\#\{i : n_i \geq j\}. \tag{11.1}$$

O lado esquerdo depende apenas de T. Ele determina, para cada j, o número de expoentes maiores ou iguais a j. Em particular, determina k pelo caso $j = 1$ e, em seguida, toda a sequência ordenada $n_1, \dots, n_k$. Isso prova a unicidade. ■

Definição 11.13. Quando o polinômio minimal de T é da forma p^t, em que p é irredutível, os polinômios $p^t = p^{n_1}, p^{n_2}, \dots, p^{n_k}$ da cadeia $t = n_1 \geq n_2 \geq \cdots \geq n_k \geq 1$ são denominados **divisores elementares** de T.

Destacamos que o primeiro divisor elementar na sequência é o polinômio minimal de T.

Teorema 11.14 (Forma Canônica Racional por Divisores Elementares). *Sejam V um espaço vetorial de dimensão finita sobre $\mathbb{K}$ e $T : V \to V$ um operador linear. Para cada fator irredutível p_i de m_T, sejam*

$$p_i^{n_{i1}}, \dots, p_i^{n_{ir_i}}, \qquad n_{i1} \geq \cdots \geq n_{ir_i} \geq 1,$$

os divisores elementares da restrição de T ao componente p_i-primário. Então existe uma base de V na qual a matriz de T é diagonal por blocos, e os blocos são precisamente as matrizes companheiras

$$C_{p_i^{n_{ij}}} \qquad (1 \leq i \leq k,\ 1 \leq j \leq r_i).$$

Essa coleção de blocos é determinada por T, a menos de sua ordem.

Demonstração. Escreva $m_T = p_1^{e_1} \cdots p_k^{e_k}$, com os p_i irredutíveis e distintos. Pelo Teorema da Decomposição Primária,

$$V = V_1 \oplus \cdots \oplus V_k, \qquad V_i = \ker p_i(T)^{e_i},$$

e o polinômio minimal de $T_i = T|_{V_i}$ é $p_i^{e_i}$. A Proposição 11.12, aplicada a T_i, fornece uma decomposição

$$V_i = \bigoplus_{j=1}^{r_i} Z_{\mathbf{v}_{ij}}, \qquad m_{\mathbf{v}_{ij}} = p_i^{n_{ij}}.$$

A soma dessas decomposições, para $1 \leq i \leq k$, é direta porque as parcelas V_i já formam uma soma direta. Reunindo as bases cíclicas de todos os $Z_{\mathbf{v}_{ij}}$, obtemos uma base de V na qual a matriz de T tem os blocos $C_{p_i^{n_{ij}}}$.

A fatoração de m_T determina unicamente os componentes primários V_i, e a Proposição 11.12 determina unicamente, em cada componente, os expoentes n_{ij}. Logo a coleção de blocos é determinada por T, restando apenas a liberdade de permutá-los. ■

Pela discussão anterior, esse formato diagonal por bloco, no qual cada bloco A_j é ele próprio uma diagonal por bloco de matrizes companheiras, é único (a menos de permutações de A_j). É denominada **matriz canônica racional por divisores elementares** de T.

Na sequência de divisores elementares pode haver repetições, pois alguns dos n_i podem ser iguais. Por isso, uma mesma matriz companheira pode aparecer mais de uma vez na forma racional.

Exemplo 11.15. Seja

$$A = \begin{bmatrix} -3 & 0 & 12 \\ -2 & 1 & 6 \\ -2 & 0 & 7 \end{bmatrix} \in \mathscr{M}_{3,3}(\mathbb{Q}).$$

Nesse caso $c_T(x) = (x-3)(x-1)^2$ e $m_T(x) = (x-3)(x-1)$. Como a matriz companheira de um polinômio linear $x - a$ é a matriz $[a]$, concluímos que a forma canônica racional de A é $D = \operatorname{diag}(1, 1, 3)$.

◁

Exemplo 11.16. Seja

$$A = \begin{bmatrix} 8 & -4 & -12 & 7 \\ 5 & -5 & -4 & 4 \\ 3 & -3 & -2 & 2 \\ 0 & -3 & 5 & -1 \end{bmatrix} \in \mathscr{M}_{4,4}(\mathbb{Q}).$$

Então $c_T(x) = m_T(x) = (x^2 - 2)(x^2 + 1)$ e, portanto, a forma racional é

$$C_{x^2-2} \oplus C_{x^2+1} = \begin{bmatrix} 0 & 2 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

◁

Exemplo 11.17. Suponha agora que $T\colon \mathbb{Q}^6 \to \mathbb{Q}^6$ possua um polinômio minimal

$$m_T = (x^2 + 1)(x - 2)^2$$

Consequentemente o polinômio característico de T é um dos

$$c_1 = (x^2 + 1)^2(x - 2)^2, \quad c_2 = (x^2 + 1)(x - 2)^4$$

Suponha primeiro que $c_T = c_1$. Nesse caso, temos $\mathbb{Q}^6 = V_1 \oplus V_2$ com $\dim V_1 = 4$ e $\dim V_2 = 2$. A aplicação linear induzida T_1 em V_1 possui o polinômio minimal $m_1 = x^2 + 1$ e a aplicação induzida T_2 em V_2 possui polinômio minimal $m_2 = (x - 2)^2$.

A situação para V_1 é que $C_{x^2+1} \oplus C_{x^2+1}$.

Quanto a V_2, o Teorema 11.14 fornece $2 = n_1 + \cdots + n_k$ e $n_1 = 2$. Como todos os n_i são positivos, não pode haver outro termo: $k = 1$. Portanto, o único divisor elementar de T_2 é $(x-2)^2$. Combinando essas observações, vemos que, neste caso, a matriz canônica racional de T é

$$C_{x^2+1} \oplus C_{x^2+1} \oplus C_{(x-2)^2}.$$

Suponha agora que $c_T = c_2$. Nesse caso, temos $\mathbb{Q}^6 = V_1 \oplus V_2$ com $\dim V_1 = 2$ e $\dim V_2 = 4$. Além disso, a aplicação induzida T_2 em V_2 possui o polinômio minimal $m_2 = (x-2)^2$. Pelo Teorema 11.14, aplicado a T_2, temos $4 = n_1 + \cdots + n_k$ com $n_1 = 2$. Existem, portanto, duas possibilidades, a saber:

- $k = 2$ com $n_1 = n_2 = 2$;
- $k = 3$ com $n_1 = 2, n_2 = n_3 = 1$.

A matriz canônica racional de T neste caso é, portanto, uma das formas

$$C_{x^2+1} \oplus C_{(x-2)^2} \oplus C_{(x-2)^2}$$

$$C_{x^2+1} \oplus C_{(x-2)^2} \oplus C_{x-2} \oplus C_{x-2}.$$

Observe no exemplo acima que o conhecimento dos polinômios característico e minimal geralmente não é suficiente para determinar completamente a forma racional.

◁

Também podemos organizar as parcelas cíclicas de modo que $f_{i+1}(x) \mid f_i(x)$. Essa condição produz a forma canônica racional por fatores invariantes.

Lema 11.18 (Combinação de componentes cíclicas).
Se $q_1, \ldots, q_s \in \mathbb{K}[x]$ são dois a dois coprimos e $\mathbf{w}_i$ é um vetor cíclico com aniquilador q_i, então $\mathbf{u} = \mathbf{w}_1 + \cdots + \mathbf{w}_s$ é cíclico no espaço $Z_{\mathbf{w}_1} \oplus \cdots \oplus Z_{\mathbf{w}_s}$ e seu aniquilador é $q_1 \cdots q_s$.

Demonstração. Escreva

$$W = Z_{\mathbf{w}_1} \oplus \cdots \oplus Z_{\mathbf{w}_s}, \qquad q = q_1 \cdots q_s,$$

e ponha $Q_i = q/q_i$. Como Q_i e q_i são coprimos, existem $a_i, b_i \in \mathbb{K}[x]$ tais que

$$a_i Q_i + b_i q_i = 1.$$

O polinômio $e_i = a_i Q_i$ é congruente a 1 módulo q_i e, para $h \neq i$, é divisível por q_h. Como $q_h(T)\mathbf{w}_h = 0$, temos

$$e_i(T)\mathbf{u} = \mathbf{w}_i.$$

Assim, cada gerador $\mathbf{w}_i$ pertence a $Z_{\mathbf{u}}$ e, como esse subespaço é T-invariante, $Z_{\mathbf{w}_i} \subseteq Z_{\mathbf{u}}$. Portanto, $W \subseteq Z_{\mathbf{u}}$. A inclusão inversa decorre de $\mathbf{u} \in W$ e da T-invariância de W. Logo $Z_{\mathbf{u}} = W$.

Além disso, $q_i(T)\mathbf{w}_i = 0$ para todo i, e cada q_i divide q; portanto, $q(T)\mathbf{u} = 0$. Reciprocamente, se $f(T)\mathbf{u} = 0$, então

$$0 = f(T)\mathbf{u} = \sum_{i=1}^{s} f(T)\mathbf{w}_i.$$

A soma direta implica $f(T)\mathbf{w}_i = 0$ para cada i, de modo que $q_i \mid f$ para todo i. Como os q_i são dois a dois coprimos, $q = q_1 \cdots q_s$ divide f. Logo o ideal dos polinômios que anulam $\mathbf{u}$ é precisamente (q), e q é o aniquilador mônico de $\mathbf{u}$. ■

Teorema 11.19 (Forma Canônica Racional por Fatores Invariantes).
Seja $T \in \mathrm{Hom}(V,V)$, com V de dimensão finita. Existem um inteiro $r > 0$, vetores $\mathbf{u}_1, \dots, \mathbf{u}_r$ e polinômios mônicos $f_1(x), \dots, f_r(x)$ tais que

1. $V = Z_{\mathbf{u}_1} \oplus \cdots \oplus Z_{\mathbf{u}_r}$, *com* $m_{\mathbf{u}_i}(x) = f_i(x)$;
2. $f_1(x) = m_T(x)$ *e* $f_{i+1}(x) \mid f_i(x)$ *para todo* $i = 1, \dots, r-1$.

Além disso, r e os polinômios $f_1, \dots, f_r$ são unicamente determinados por T.

A cadeia de divisibilidade
Adotamos a ordem decrescente: cada fator invariante é divisível pelo seguinte, isto é, $f_{i+1} \mid f_i$. Com essa condição, a lista dos fatores invariantes é determinada unicamente pelo operador.

Demonstração. Para cada fator irredutível p_i do polinômio minimal, escreva os divisores elementares do componente p_i-primário na ordem

$$p_i^{n_{i1}}, \dots, p_i^{n_{ir_i}}, \qquad n_{i1} \geq \cdots \geq n_{ir_i} \geq 1.$$

Seja $r = \max_i r_i$ e, quando $r_i < r$, complete a lista à direita com expoentes nulos. Para $1 \leq j \leq r$, defina

$$f_j(x) = \prod_i p_i(x)^{n_{ij}}, \qquad 1 \leq j \leq r.$$

Para cada i, o maior expoente n_{i1} é o expoente de p_i em m_T. Logo

$$f_1 = \prod_i p_i^{n_{i1}} = m_T.$$

Como $n_{i,j+1} \leq n_{ij}$ para todo i, temos $f_{j+1} \mid f_j$.

Fixe uma coluna j e some os vetores cíclicos não nulos que aparecem nessa coluna, um em cada componente primário. Seus aniquiladores são potências de polinômios irredutíveis distintos e, portanto, são dois a dois coprimos. Pelo Lema 11.18, essa soma é um vetor cíclico $\mathbf{u}_j$ com aniquilador f_j. Como apenas reagrupamos as parcelas da decomposição por divisores elementares, obtemos

$$V = Z_{\mathbf{u}_1} \oplus \cdots \oplus Z_{\mathbf{u}_r}.$$

Por fim, os divisores elementares são determinados unicamente por T. Para cada irredutível p_i, o expoente de p_i em f_j é exatamente n_{ij}. Assim, a fatoração dos f_j recupera todas as listas de divisores elementares, inclusive seus comprimentos. Reciprocamente, essas listas determinam r e cada f_j. Isso prova a unicidade. ■

Definição 11.20. Os polinômios $f_1, \dots, f_r$ são denominados **fatores invariantes** de T.

11.3 Forma Normal de Jordan Generalizada

A forma de Jordan generalizada refina os blocos da forma racional. Em cada parcela cíclica, substitui a matriz companheira de uma potência p^n por uma matriz em blocos construída a partir da companheira de p. Quando todos os fatores irredutíveis são lineares, recuperamos a forma de Jordan usual.

Teorema 11.21.
Seja $\mathbf{v}$ *um vetor cíclico para* $T : V \to V$*, cujo polinômio minimal é* p^n*, onde* $p(x) = x^k + a_{k-1}x^{k-1} + \cdots + a_0$ *é mônico e irredutível. Então existe uma base de* V *na qual a matriz de* T *é a matriz* $kn \times kn$

$$\begin{bmatrix} C_p & N & & & \\ & C_p & N & & \\ & & \ddots & \ddots & \\ & & & C_p & N \\ & & & & C_p \end{bmatrix}$$

onde C_p *é a matriz companheira de* p *e* $N \in \mathscr{M}_{k,k}(\mathbb{K})$ *tem todas as entradas nulas, exceto* $N_{1k} = 1$.

Demonstração. Para $0 \leq j < n$ e $0 \leq \ell < k$, considere os polinômios

$$x^\ell p(x)^j.$$

Eles são mônicos e seus graus $jk + \ell$ percorrem, sem repetição, os inteiros de 0 a $kn - 1$. Portanto, formam uma base do espaço dos polinômios de grau menor que kn.

Considere agora a lista ordenada

$$\underline{\mathrm{B}} = \left(p(T)^j\mathbf{v},\, Tp(T)^j\mathbf{v}, \ldots, T^{k-1}p(T)^j\mathbf{v}\right)_{j=n-1}^{0},$$

em que os blocos são concatenados na ordem decrescente de j. Se houvesse uma relação linear não trivial entre esses vetores, a mesma relação entre os polinômios $x^\ell p^j$ produziria um polinômio não nulo q de grau menor que kn tal que $q(T)\mathbf{v} = 0$. Isso é impossível, pois o aniquilador de $\mathbf{v}$ é p^n, de grau kn. Logo os kn vetores de $\underline{\mathrm{B}}$ são linearmente independentes. Como $\dim V = \operatorname{grau} m_{\mathbf{v}} = kn$, eles formam uma base de V.

Para $0 \leq \ell < k - 1$, o operador T leva $T^\ell p(T)^j\mathbf{v}$ ao vetor seguinte do mesmo bloco. Como

$$p(T) = T^k + a_{k-1}T^{k-1} + \cdots + a_0\,\mathrm{I},$$

a imagem do último vetor do bloco é

$$T^k p(T)^j\mathbf{v} = -a_0 p(T)^j\mathbf{v} - \cdots - a_{k-1}T^{k-1}p(T)^j\mathbf{v} + p(T)^{j+1}\mathbf{v}.$$

Os primeiros k termos dessa expressão fornecem o bloco companheiro C_p na diagonal. Se $j = n - 1$, o termo $p(T)^{j+1}\mathbf{v}$ é nulo porque $p(T)^n\mathbf{v} = 0$. Se $j < n - 1$, ele é o primeiro vetor do bloco imediatamente anterior, pois os expoentes foram ordenados de modo decrescente. Na matriz, sua contribuição é uma entrada 1 na primeira linha e na última coluna do bloco situado logo acima da diagonal, isto é, o bloco N com $N_{1k} = 1$. Todas as demais entradas fora dos blocos descritos são nulas. Obtemos, assim, exatamente a matriz do enunciado. ∎

Definição 11.22. A matriz de blocos descrita no Teorema 11.21 é denominada matriz de Jordan generalizada associada à matriz companheira C_p.

Passando agora para o caso geral, considere uma aplicação linear $T : V \to V$. Com a notação usual, a aplicação induzida T_i em V_i possui polinômio minimal $p_i^{e_i}$. Aplicando o Teorema 11.21 aos subespaços cíclicos da decomposição de T_i, podemos substituir cada bloco companheiro $C_{p_i^{n_j}}$ pela matriz de Jordan generalizada associada a C_{p_i}. A matriz obtida dessa maneira é denominada **forma de Jordan generalizada** de T.

Exemplo 11.23. Seja $T : \mathbb{R}^6 \to \mathbb{R}^6$ linear e tal que

$$c_T = m_T = (x^2 - x + 1)^2(x + 1)^2$$

Pelo Teorema da Decomposição Primária, o polinômio minimal da aplicação induzida T_1 é $(x^2 - x + 1)^2$, enquanto o de T_2 é $(x + 1)^2$. Usando o Teorema 11.14, vemos que o único divisor elementar de T_1 é $(x^2 - x + 1)^2$ e o único divisor elementar de T_2 é $(x + 1)^2$. Consequentemente, a matriz canônica racional de T é

$$C_{(x^2-x+1)^2} \oplus C_{(x+1)^2} = \left[\begin{array}{cccc|cc} 0 & 0 & 0 & -1 & & \\ 1 & 0 & 0 & 2 & & \\ 0 & 1 & 0 & -3 & & \\ 0 & 0 & 1 & 2 & & \\ \hline & & & & 0 & -1 \\ & & & & 1 & -2 \end{array}\right]$$

A matriz de Jordan generalizada é obtida substituindo cada companheira de p_i^2 pelo bloco generalizado associado a C_{p_i}. Assim, obtemos

$$\left[\begin{array}{cccc|cc} 0 & -1 & 0 & 1 & & \\ 1 & 1 & 0 & 0 & & \\ 0 & 0 & 0 & -1 & & \\ 0 & 0 & 1 & 1 & & \\ \hline & & & & -1 & 1 \\ & & & & 0 & -1 \end{array}\right]$$

Quando $k = 1$, temos $p(x) = x - \alpha$. Nesse caso $C_p = [\alpha]$ e o bloco generalizado se reduz ao bloco de Jordan usual associado ao autovalor α. Portanto, quando todos os fatores irredutíveis são lineares, a forma de Jordan generalizada coincide com a forma de Jordan.

◁

Exemplo 11.24. Se $T : \mathbb{R}^7 \to \mathbb{R}^7$ tem polinômio característico $c_T(x) = (x-2)^3(x-3)^4$, o polinômio característico, sozinho, não determina a forma canônica racional. As possibilidades correspondem independentemente às partições de 3 e de 4; portanto, há $3 \cdot 5 = 15$ formas possíveis, a menos de permutação dos blocos.

◁

11.4 Calculando a Forma Canônica Racional

Passamos agora do resultado estrutural ao cálculo dos divisores elementares e dos fatores invariantes.

Fixe um fator irredutível $p(x)$ de grau d no polinômio minimal de T. O diagrama de pontos associado a p será determinado pela Equação 11.1 da Proposição 11.12 e registrará os divisores elementares correspondentes.

Calcule recursivamente, até obter zero,

$$r_1 = \frac{1}{d}\left(\dim(V) - \text{posto}(p(T))\right)$$
$$r_k = \frac{1}{d}\left(\text{posto}(p(T)^{k-1}) - \text{posto}(p(T)^k)\right) \text{ para } k \geq 2$$

A k-ésima linha do diagrama recebe r_k pontos. Se a i-ésima coluna tem s_i pontos, o divisor elementar correspondente é $p(x)^{s_i}$.

Exemplo 11.25. Para $p(x) = x^2 + 1$, suponha que os valores calculados sejam $r_1 = 3$, $r_2 = 3$, $r_3 = 1$, $r_4 = 1$ e $r_5 = 0$.

O diagrama de pontos seria:

$$\begin{array}{ccc} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & & \\ \bullet & & \end{array}$$

As alturas das colunas são 4, 2 e 2; logo os divisores elementares são $(x^2+1)^4$, $(x^2+1)^2$ e $(x^2+1)^2$.

◁

Repetindo o procedimento para cada fator irredutível do polinômio minimal e colocando as companheiras dos divisores elementares na diagonal, obtemos uma primeira forma racional em blocos.

Para recuperar os fatores invariantes, disponha em cada linha os divisores elementares associados a um mesmo fator irredutível, alinhados do maior grau para o menor. Os produtos nas colunas são os fatores invariantes; o maior deles é o polinômio minimal. Por exemplo, se as linhas correspondentes a $x^2 + 1$ e $x - 2$ forem

$$(x^2+1)^4,\ (x^2+1)^2,\ (x^2+1)^2 \qquad \text{e} \qquad (x-2)^3,\ (x-2),$$

então os fatores invariantes são

$$(x^2+1)^4(x-2)^3, \qquad (x^2+1)^2(x-2), \qquad (x^2+1)^2.$$

A forma canônica racional tem, na diagonal, as matrizes companheiras desses fatores invariantes.

Exemplo 11.26. Vamos encontrar a forma racional de

$$A = \begin{bmatrix} -1 & -2 & 6 \\ -1 & 0 & 3 \\ -1 & -1 & 4 \end{bmatrix}$$

Logo o polinômio característico de A é $c_A(x) = (x-1)^3$ e o polinômio minimal é $m_A(x) = (x-1)^2$; o único fator irredutível do polinômio minimal é $(x-1)$. Vamos calcular seu diagrama de pontos.

$$\begin{aligned} r_1 &= (3 - \text{posto}(A - I)) = (3 - 1) = 2 \\ r_2 &= (\text{posto}(A - I) - \text{posto}((A - I)^2)) = (1 - 0) = 1 \\ r_3 &= \text{posto}((A - I)^2) - \text{posto}((A - I)^3) = 0 - 0 = 0 \end{aligned}$$

Portanto, o diagrama de pontos é

$$\begin{array}{cc} \bullet & \bullet \\ \bullet & \end{array}$$

Portanto, os divisores elementares são $(x-1)^2$ e $(x-1)$. Nesse caso, como existe apenas um fator irredutível do polinômio minimal, os fatores invariantes são os mesmos que

os divisores elementares. Assim, a forma canônica racional, obtida por qualquer uma das duas descrições, é:

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

◁

Definição 11.27 — Operador semissimples. Sejam V um espaço vetorial de dimensão finita sobre $\mathbb{K}$ e $T \in \mathrm{Hom}(V,V)$. Dizemos que T é *semissimples* se, para todo subespaço T-invariante $W \subseteq V$, existe um subespaço T-invariante $U \subseteq V$ tal que $V = W \oplus U$.

Semissimplicidade e polinômio minimal
Um operador semissimples é aquele cujos subespaços invariantes admitem complementos invariantes. Quando o polinômio minimal se decompõe em fatores lineares, isso equivale à ausência de fatores repetidos e, portanto, de blocos de Jordan de tamanho maior que um.

Os exercícios a seguir relacionam essa propriedade ao polinômio minimal e, quando o corpo é algebricamente fechado, à diagonalização.

Exercícios

Ex. 11.2 — Mostre que toda matriz $A \in M_n(\mathbb{C})$ é semelhante à sua transposta.

Ex. 11.3 — Seja V um $\mathbb{K}$-e.v. de dimensão finita, $\mathbf{v} \in V$ e $T : V \to V$ um operador linear. Seja

$$\mathscr{C}(\mathbf{v};T) = \{W \subseteq V \mid W \text{ é subespaço } T\text{-invariante e } \mathbf{v} \in W\} \quad \text{e} \quad Z_{\mathbf{v}} = \bigcap_{W \in \mathscr{C}(\mathbf{v};T)} W.$$

Mostre que:

1. $Z_{\mathbf{v}}$ é T-invariante;
2. $Z_{\mathbf{v}}$ é o menor subespaço T-invariante de V que contém $\mathbf{v}$;
3. $Z_{\mathbf{v}} = \{g(T)(\mathbf{v}) \mid g \in \mathbb{K}[x]\}$;
4. $\dim Z_{\mathbf{v}} = 1$ se, e somente se, $\mathbf{v}$ é um autovetor de T.

Ex. 11.4 — Seja $\mathbb{K}$ um corpo e seja $B \in M_n(\mathbb{K})$. Mostre que B é semelhante sobre $\mathbb{K}$ a uma e somente uma matriz que está na forma racional.

Ex. 11.5 — Sejam V um espaço vetorial sobre $\mathbb{K}$ de dimensão finita e $T \in \mathrm{Hom}(V,V)$. Prove que se existe um vetor cíclico para T^2 então existe um vetor cíclico para T. Vale a recíproca?

Ex. 11.6 — Seja V um espaço vetorial sobre $\mathbb{K}$ e $T : V \to V$ um operador linear com polinômio minimal $m_T(x) = p_1(x)^{m_1} \cdots p_k(x)^{m_k}$ onde os $p_i(x)$ são distintos e irredutíveis em $\mathbb{K}[x]$. Seja $V = W_1 \oplus \cdots \oplus W_k$ a decomposição de V dada pelo teorema da decomposição primária. Mostre que para todo subespaço T-invariante W de V temos que

$$W = \bigoplus_{i=1}^{k} (W \cap W_i).$$

Ex. 11.7 — Seja V um espaço vetorial sobre $\mathbb{K}$ e $T : V \to V$ um operador linear. Sejam $p \in \mathbb{K}[x]$ um polinômio irredutível e m um inteiro positivo. Suponha que existe $\mathbf{v} \in V$ tal que $m_{\mathbf{v}} = p^m$. Seja $\mathbf{w} \in Z_{\mathbf{v}}$.

1. Mostre que $\mathbf{w} = q(T)p(T)^l(\mathbf{v})$, onde $q \in \mathbb{K}[x]$ é coprimo com p e $0 \leq l \leq m$.

2. Mostre que $Z_{\mathbf{w}} = Z_{p(T)^l(\mathbf{v})}$. Logo, os únicos subespaços T-cíclicos de $Z_{\mathbf{v}}$ são
$$0 = Z_{p(T)^m(\mathbf{v})} \subset Z_{p(T)^{m-1}(\mathbf{v})} \subset \cdots \subset Z_{p(T)(\mathbf{v})} \subset Z_{\mathbf{v}}.$$

3. Mostre que todo subespaço T-invariante de $Z_{\mathbf{v}}$ é T-cíclico, isto é, tem a forma $Z_{p(T)^l(\mathbf{v})}$ para algum $0 \leq l \leq m$.

Ex. 11.8 — Sejam V um espaço vetorial sobre $\mathbb{K}$ de dimensão finita e $T \in \mathrm{Hom}(V, V)$.

1. Prove que se existe um vetor cíclico para T, então todo subespaço próprio T-invariante de V também tem um vetor cíclico.

2. Vale a recíproca do item 1.?

Ex. 11.9 — Sejam V um espaço vetorial sobre $\mathbb{K}$ de dimensão n e $T \in \mathrm{Hom}(V, V)$ um operador diagonalizável.

1. Mostre que existe um vetor cíclico para T se, e somente se, T tem n autovalores distintos.

2. Mostre que se T tem n autovalores distintos e se $\underline{\mathrm{B}} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ é uma base de autovetores de T, então $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_n$ é um vetor cíclico de T.

Ex. 11.10 — Sejam V um espaço vetorial de dimensão n sobre $\mathbb{K}$ e $T \in \mathrm{Hom}(V, V)$.

1. Prove que $\mathrm{im}\, T$ tem um complementar T-invariante se, e somente se, $\mathrm{im}\, T \cap \ker T = 0$.

2. Se $\mathrm{im}\, T \cap \ker T = 0$, prove que $\ker T$ é o único complementar de $\mathrm{im}\, T$ que é T-invariante.

Ex. 11.11 — Seja $A \in M_3(\mathbb{R})$ a matriz
$$A = \begin{bmatrix} 1 & 3 & 3 \\ 3 & 1 & 3 \\ -3 & -3 & -5 \end{bmatrix}.$$
Encontre uma matriz invertível P tal que $P^{-1}AP$ esteja na forma racional.

Ex. 11.12 — Seja $A \in M_3(\mathbb{R})$ a matriz
$$A = \begin{bmatrix} 3 & -4 & -4 \\ -1 & 3 & 2 \\ 2 & -4 & -3 \end{bmatrix}.$$
Encontre vetores $\mathbf{v}_1, \ldots, \mathbf{v}_r$ que satisfaçam as condições do Teorema da Decomposição Cíclica.

Ex. 11.13 — Sejam V um espaço vetorial de dimensão finita sobre $\mathbb{K}$ e $T \in \mathrm{Hom}(V, V)$. Prove que T tem um vetor cíclico se, e somente se, a seguinte afirmação é verdadeira: "Todo operador linear que comuta com T é um polinômio em T."

Ex. 11.14 — Sejam V um espaço vetorial sobre $\mathbb{K}$ de dimensão finita e $T \in \mathrm{Hom}(V, V)$. Prove que todo vetor $\vec{0} \neq \mathbf{v} \in V$ é um vetor cíclico para T se, e somente se, o polinômio característico de T é irredutível em $\mathbb{K}[x]$.

Ex. 11.15 — Sejam V um espaço vetorial sobre $\mathbb{K}$ de dimensão finita e $T \in \mathrm{Hom}(V, V)$. Suponha que $c_T = m_T = q^m$ é uma potência de um polinômio irredutível. Prove que nenhum subespaço não trivial T-invariante de V tem um complementar que também é T-invariante.

Ex. 11.16 — Seja $T \in \mathrm{Hom}(\mathbb{R}^4, \mathbb{R}^4)$ um operador linear tal que $x^2 + 3$ é um divisor do polinômio minimal de T e 1 é o único autovalor de T. Quais são as possíveis formas racionais de T?

Ex. 11.17 — Seja $T \in \mathrm{Hom}(\mathbb{K}^6, \mathbb{K}^6)$ com polinômio minimal $m_T(x) = (x^2 - 2)(x^2 + 1)$. Ache as possíveis formas racionais de T para $\mathbb{K} = \mathbb{R}$ e $\mathbb{K} = \mathbb{C}$.

Ex. 11.18 — Recorde que um operador é semissimples quando todo subespaço invariante admite um complemento também invariante. Seja V um espaço vetorial de dimensão finita sobre um corpo $\mathbb{K}$ e T um operador linear em V.

1. Mostre que se o polinômio minimal de T for irredutível em $\mathbb{K}$ então T é semissimples.

2. Se $\mathbb{K}$ for algebricamente fechado, prove que T é diagonalizável se, e somente se, é semissimples.

3. Mostre que o operador $T : \mathbb{R}^3 \to \mathbb{R}^3$ tal que

$$[T]_{\underline{B}}^{\underline{B}} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

   não é semissimples.

12
Aplicações

As formas canônicas permitem calcular potências e funções de operadores. Em tempo discreto, as potências de uma matriz descrevem recorrências e sistemas lineares; em tempo contínuo, a exponencial do operador fornece as soluções de equações diferenciais lineares com coeficientes constantes. O estudo dos limites e o teorema de Perron–Frobenius permitem relacionar essas expressões ao comportamento assintótico das soluções.

Nesse capítulo

12.1 Equações de Recorrência

Quando o tempo é discreto, o estado seguinte é calculado a partir de estados anteriores. Essa regra produz uma relação de recorrência; sua versão linear será estudada por meio de potências de matrizes.

Transformar memória em estado
Uma recorrência de ordem k depende dos k valores anteriores. Reunindo esses valores num único vetor de estado, transformamos a recorrência numa dinâmica de primeira ordem governada por uma matriz.

Definição 12.1 — Relação de recorrência. Uma **relação de recorrência** expressa cada termo de uma sequência em função de termos anteriores. Ela é de primeira ordem quando tem a forma

$$u_n = \varphi(n, u_{n-1}), \qquad n > 0,$$

com $\varphi : \mathbb{N} \times \mathbb{K} \to \mathbb{K}$. Fixado u_0, a relação determina sucessivamente todos os demais termos.
Mais geralmente, uma recorrência de ordem k tem a forma

$$u_n = \varphi(n, u_{n-1}, u_{n-2}, \dots, u_{n-k}), \qquad n \geq k,$$

onde $\varphi : \mathbb{N} \times \mathbb{K}^k \to \mathbb{K}$. Nesse caso, são necessários k valores iniciais.

Exemplo 12.2 — Fibonacci. A sequência de Fibonacci é definida recursivamente pela fórmula

$$F_n = F_{n-1} + F_{n-2},$$

e valores iniciais

$$F_1 = 1, \; F_2 = 1.$$

Os primeiros valores são: $1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, \dots$

◁

Exemplo 12.3 — Mapa Logístico.

$$x_{n+1} = r x_n (1 - x_n)$$

com uma constante r dada; fixado o termo inicial x_0, cada termo subsequente é determinado pela relação.

◁

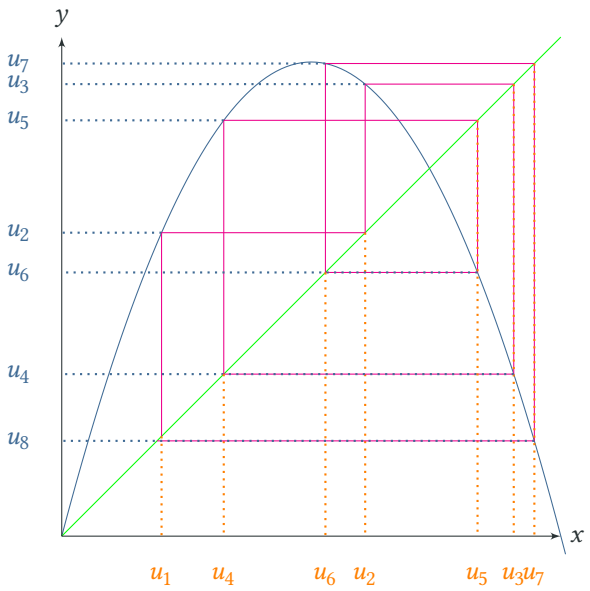


Figura 12.1 $x_n = 3.8 x_{n-1}(1 - x_{n-1})$

Numa equação a diferenças finitas, a incógnita é uma sequência $(u_0, u_1, \dots, u_k, \dots)$ cujos termos devem satisfazer uma relação de recorrência dada.

Uma equação a diferenças finitas é dita **linear** de ordem k se puder ser escrita em termos dos parâmetros $a_1, \dots, a_k$ e b como

$$u_n = a_1 u_{n-1} + \cdots + a_k u_{n-k} + b$$

A equação é dita **homogênea** se $b = 0$ e não homogênea se $b \neq 0$.

Dado o valor inicial u_0, uma recorrência de primeira ordem determina sucessivamente $u_1, u_2, \dots$. De modo análogo, fixados u_0 e u_1, uma recorrência de segunda ordem determina todos os termos seguintes.

> **Teorema 12.4** (Existência e Unicidade).
> *Uma relação de recorrência de ordem k*
>
> $$u_n = \varphi(n, u_{n-1}, u_{n-2}, \dots, u_{n-k}) \quad \textit{para} \quad n \geq k,$$
>
> *juntamente com k valores iniciais $u_0, u_1, \dots u_{k-1}$ possui uma única solução.*

Demonstração. Os valores iniciais determinam os termos de índices $0, \dots, k-1$. Depois disso, u_k é determinado por esses valores, u_{k+1} é determinado pelos k termos anteriores, e assim sucessivamente. Esse procedimento constrói uma solução. Ele também prova a unicidade por indução: duas soluções com os mesmos valores iniciais coincidem nos primeiros k índices e, se coincidem até o índice $n-1$, a recorrência obriga que coincidam também no índice n. ■

Teorema 12.5.
O conjunto das soluções de uma equação linear homogênea de ordem k

$$u_n = a_1 u_{n-1} + \cdots + a_k u_{n-k}$$

é um espaço vetorial de dimensão k.

Demonstração. As operações são feitas termo a termo, por isso a soma de duas soluções e o produto de uma solução por um escalar ainda satisfazem a recorrência homogênea. Considere

$$\Phi : \mathcal{S} \longrightarrow \mathbb{K}^k, \qquad \Phi((u_n)) = (u_0, \dots, u_{k-1}),$$

onde $\mathcal{S}$ é o espaço das soluções. Pelo teorema anterior, cada escolha de valores iniciais determina uma única solução. Logo Φ é um isomorfismo linear e $\dim \mathcal{S} = k$. ■

Exemplo 12.6 — Progressão Aritmética. A solução da equação de primeira ordem $u_{k+1} = u_k + b$, com valor inicial u_0, é a sequência $(u_0, u_0 + b, u_0 + 2b, \dots)$, cujo termo geral é $u_k = u_0 + kb$. Trata-se, portanto, de uma progressão aritmética de razão b.

◁

Exemplo 12.7 — Progressão Geométrica. A equação $u_{k+1} = au_k$ com valor inicial u_0 tem como solução a sequência

$$(u_0,\ au_0,\ a^2u_0,\ \dots,\ a^k u_0,\ \dots)$$

cujo termo geral é $u_k = a^k u_0$, ou seja, uma progressão geométrica de razão a.

◁

Exemplo 12.8. Seja $u_{k+1} = au_k + b$ uma equação linear não homogênea de primeira ordem, com coeficientes constantes. A partir do valor inicial u_0, obtemos sucessivamente:

$$\begin{aligned}
u_1 &= au_0 + b,\\
u_2 &= au_1 + b = a^2u_0 + (1+a)b,\\
u_3 &= au_2 + b = a^3u_0 + (1+a+a^2)b,\\
&\ \ \vdots\\
u_k &= au_{k-1} + b = a^k u_0 + (1 + a + \cdots + a^{k-1})b
\end{aligned}$$

Portanto a solução geral da equação $u_{k+1} = au_k + b$ é

$$u_k = \begin{cases} a^k u_0 + \dfrac{1-a^k}{1-a}\, b, & a \neq 1,\\ u_0 + kb, & a = 1. \end{cases}$$

◁

No caso geral, a e c serão funções de n e temos o seguinte resultado.

Teorema 12.9.
Sejam $a(n)$ e $c(n), n \in \mathbb{N}$, sejam sequências reais. Então a equação da diferença linear de primeira ordem

$$u(n+1) = a(n)u(n) + c(n) \text{ com condição inicial} u(0) = u_0$$

tem solução

$$u(n) = (\prod_{k=0}^{n-1} a(k))u_0 + \sum_{k=0}^{n-1}(\prod_{j=k+1}^{n-1} a(j))c(k)$$

E a solução é única.

Demonstração. Substituindo $n = 0$ na fórmula e adotando a convenção de que o produto vazio vale 1, obtemos $u(0) = u_0$. Para o passo seguinte, isolamos o termo $k = n$ da soma e fatoramos $a(n)$ nos demais termos; isso fornece diretamente

$$u(n+1) = a(n)u(n) + c(n).$$

Logo a fórmula define uma solução. A unicidade segue por indução: duas soluções com o mesmo valor inicial coincidem em $n = 0$ e, se coincidem em n, a recorrência mostra que coincidem em $n + 1$. ■

A equação $u_{k+1} = au_k + b$ também pode ser vista como uma equação linear no espaço das sequências. Defina $L(\mathbf{u})_k = u_{k+1} - au_k$. Então a recorrência equivale a $L(\mathbf{u}) = \vec{b}$, onde $\vec{b} = (b, b, \ldots)$.

Teorema 12.10.
A solução geral da equação $A\vec{x} = \vec{b}$ é a soma de uma solução particular com um elemento qualquer do núcleo de A, isto é, com uma solução da equação homogênea $A\vec{x} = \vec{0}$.

Demonstração. Se $\mathbf{x}_p$ é uma solução particular, então $A\mathbf{x} = \vec{b}$ equivale a $A(\mathbf{x} - \mathbf{x}_p) = \vec{0}$. Portanto, $\mathbf{x} = \mathbf{x}_p + \mathbf{z}$ com $\mathbf{z} \in \ker A$, e todo vetor dessa forma é solução. ■

Se $a \neq 1$, a sequência constante $\hat{c} = (c, c, \ldots)$ é uma solução particular quando $(1-a)c = b$, isto é, $c = b/(1-a)$. O núcleo de L é formado pelas progressões geométricas $(p, ap, a^2p, \ldots)$. Somando a solução particular e impondo o valor inicial, recuperamos

$$u_k = a^k u_0 + \frac{1 - a^k}{1 - a}\, b.$$

12.1.1 Sistemas Lineares

Sistemas de recorrências
Um sistema de recorrências lineares evolui por iteração de um operador. Autovalores indicam taxas de crescimento ou decaimento; blocos não diagonais acrescentam fatores polinomiais ao comportamento exponencial.

Num espaço vetorial V, uma recorrência linear de primeira ordem é determinada por um operador $T : V \to V$ e pela regra

$$\mathbf{v}_{k+1} = T\mathbf{v}_k$$

Esse é um sistema linear homogêneo de equações a diferenças com coeficientes constantes.

Dado $\mathbf{v}_0 \in V$, a única solução é $\mathbf{v}_k = T^k \mathbf{v}_0$.

O problema prático reduz-se, portanto, ao cálculo das potências sucessivas de T. Se T for diagonalizável, escrevemos o vetor inicial como

$$\mathbf{v}_0 = c_1 \mathbf{u}_1 + \cdots + c_n \mathbf{u}_n,$$

onde $T\mathbf{u}_i = \lambda_i \mathbf{u}_i$. Segue-se que

$$T^k \mathbf{v}_0 = c_1 \lambda_1^k \mathbf{u}_1 + \cdots + c_n \lambda_n^k \mathbf{u}_n.$$

Teorema 12.11.
Um sistema linear não homogêneo de k equações de diferenças de primeira ordem pode ser escrito como

$$\begin{aligned} u_1(n+1) &= a_{11}u_1(n) + a_{12}u_2(n) + \cdots + a_{1k}u_k(n) + c_1 \\ u_2(n+1) &= a_{21}u_1(n) + a_{22}u_2(n) + \cdots + a_{2k}u_k(n) + c_2 \\ &\vdots \\ u_k(n+1) &= a_{k1}u_1(n) + a_{k2}u_2(n) + \cdots + a_{kk}u_k(n) + c_k. \end{aligned}$$

Usando a notação matricial

$$\mathbf{u}(n) = \begin{bmatrix} u_1(n) \\ u_2(n) \\ \vdots \\ u_k(n) \end{bmatrix}, \qquad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & & a_{2k} \\ \vdots & \vdots & & \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix}, \qquad \vec{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix}$$

Então a solução do sistema escrito como

$$\mathbf{u}(n+1) = A\mathbf{u}(n) + \vec{c}$$

é

$$\mathbf{u}(n) = A^n \mathbf{u}(0) + \sum_{j=0}^{n-1} A^{n-1-j}\vec{c}, \qquad n \in \mathbb{N}.$$

Demonstração. A forma matricial decorre diretamente da definição do produto de matrizes. Para a fórmula da solução, o caso $n = 0$ é imediato. Se ela vale para n, então

$$\begin{aligned} \mathbf{u}(n+1) &= A\mathbf{u}(n) + \vec{c} \\ &= A^{n+1}\mathbf{u}(0) + \sum_{j=0}^{n-1} A^{n-j}\vec{c} + \vec{c} \\ &= A^{n+1}\mathbf{u}(0) + \sum_{j=0}^{n} A^{n-j}\vec{c}, \end{aligned}$$

que é a fórmula desejada no índice $n+1$. A mesma indução mostra que a solução é única para cada valor inicial. ∎

Quando A não é diagonalizável, a forma de Jordan ainda torna o cálculo de A^n explícito: em cada bloco $J = \lambda \mathrm{I} + N$, usamos o binômio e a nilpotência de N,

$$J^n = \sum_{j=0}^{s-1} \binom{n}{j} \lambda^{n-j} N^j,$$

onde s é o tamanho do bloco.

12.2 Limites de Operadores

Normas e convergência
Para falar em limite de matrizes ou operadores, precisamos medir distâncias. Em dimensão finita, diferentes normas produzem a mesma noção de convergência, embora forneçam estimativas quantitativas distintas.

Para definir limites de operadores em dimensão finita, basta escolher bases e acompanhar a convergência das entradas de suas matrizes. Começamos, portanto, com matrizes sobre $\mathbb{R}$ ou $\mathbb{C}$.

Definição 12.12. Seja $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Uma sequência (A_r) em $\mathscr{M}_{n,n}(\mathbb{K})$ converge para A quando cada entrada de A_r converge para a entrada correspondente de A:

$$|a_{ij}^{(r)} - a_{ij}| \to 0 \text{ quando } r \to \infty$$

para $1 \leq i, j \leq n$, onde $A_r = [a_{ij}^{(r)}]$.

Essa definição equivale à convergência na norma

$$\|A\|_\infty = \max\{|a_{ij}| \mid 1 \leq i, j \leq n\}$$

cuja distância associada é $d(A, B) = \|A - B\|_\infty$. De fato,

$$A_n \to A \text{ quando } n \to \infty \Leftrightarrow \|A_n - A\|_\infty \to 0 \text{ quando } n \to \infty$$

A norma do supremo possui as seguintes propriedades.

Proposição 12.13.
Se $A, B \in \mathscr{M}_{n,n}(\mathbb{K})$ então:

1. *$\|\cdot\|_\infty$ é uma norma.*
2. *Propriedade multiplicativa: $\|AB\|_\infty \leq n \cdot \|A\|_\infty \cdot \|B\|_\infty$.*
3. *Se $A_r \to A$ e $B_r \to B$ na norma do sup e $\lambda_r \to \lambda$ em $\mathbb{C}$, então:*
   - a. *$\lambda_r A_r \to \lambda A$;*
   - b. *$A_r + B_r \to A + B$;*
   - c. *$A_r B \to AB$ e $AB_r \to AB$;*
   - d. *$A_r B_r \to AB$. Assim, a multiplicação de matrizes é uma operação contínua;*
   - e. *se Q for uma matriz invertível, $QA_rQ^{-1} \to QAQ^{-1}$. Portanto, toda transformação de similaridade $A \mapsto QAQ^{-1}$ é uma operação contínua no espaço das matrizes.*

Demonstração. As propriedades de norma seguem das propriedades do módulo em $\mathbb{R}$ ou $\mathbb{C}$. Para cada i, j,

$$|(AB)_{ij}| \leq \sum_{h=1}^{n} |a_{ih}||b_{hj}| \leq n \|A\|_\infty \|B\|_\infty ,$$

o que prova a desigualdade multiplicativa. As afirmações sobre soma e produto por escalar seguem das desigualdades triangular e multiplicativa. Para o produto, note que

uma sequência convergente (A_r) é limitada e escreva

$$A_r B_r - AB = A_r(B_r - B) + (A_r - A)B.$$

Os dois termos tendem a zero. O caso da similaridade é obtido tomando os fatores fixos Q e Q^{-1}. ■

12.3 Funções de Operadores

12.3.1 Exponencial de Matrizes

A exponencial é um fluxo
A identidade $e^{(s+t)A} = e^{sA}e^{tA}$ mostra que a exponencial de matrizes descreve uma evolução compatível com a soma dos tempos. Sua derivada satisfaz $\frac{d}{dt}e^{tA} = Ae^{tA}$.

Definiremos a exponencial de uma matriz por meio da série de Taylor para a exponencial real:

Definição 12.14. Seja A uma matriz $n \times n$. Definimos $e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}$.

Teorema 12.15 (Critério de Convergência de Cauchy).
Uma sequência $\{A_n\}$ em $\mathscr{M}_{n,n}(\mathbb{K})$ converge para algum limite $A_0 = \lim_{n\to\infty} A_n$ na norma $\|\cdot\|_\infty$ se, e somente se, a sequência tiver a propriedade Cauchy: dado $\epsilon > 0$, podemos encontrar um $M > 0$ tal que

$$\|A_m - A_n\|_\infty < \epsilon \text{ para todos os } m, n \geq M$$

Demonstração. Se (A_r) converge, a desigualdade triangular mostra que ela é de Cauchy. Reciprocamente, se (A_r) é de Cauchy nessa norma, então cada sequência de entradas $(a_{ij}^{(r)})$ é de Cauchy em $\mathbb{R}$ ou $\mathbb{C}$ e, portanto, converge para um escalar a_{ij}. Como há apenas um número finito de entradas, para cada $\epsilon > 0$ existe um mesmo índice a partir do qual todas elas estão a menos de ϵ de seus limites. Assim, $A_r \to A = (a_{ij})$ na norma do supremo. ■

Teorema 12.16.
A série e^A converge para toda matriz $A \in \mathscr{M}_{n,n}(\mathbb{K})$.

Demonstração. Se $A \in \mathscr{M}_{n,n}(\mathbb{K})$, a desigualdade multiplicativa implica $\left\|A^k\right\|_\infty \leq n^{k-1}\|A\|_\infty^k$ para $k \geq 1$. Logo a série matricial é absolutamente dominada, entrada a entrada, por uma constante vezes a série escalar $\sum_{k\geq 0}(n\|A\|_\infty)^k/k!$, que converge. Portanto, suas somas parciais formam uma sequência de Cauchy e e^A está bem definida. ■

Exemplo 12.17 — Exponencial da matriz nula. Seja $\vec{0}$ a matriz nula então

$$e^{\vec{0}} = \sum_{n=0}^{\infty} \frac{0^n}{n!} = I$$

◁

Exemplo 12.18 — Exponencial da Matriz Identidade. Como $I^k = I$ para $k \geq 1$, temos $e^I = eI$.

◁

Exemplo 12.19 — Exponencial de uma Matriz Diagonal. De modo análogo ao exemplo anterior se

$$D = \begin{bmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \dots & & \ddots & \\ 0 & 0 & \dots & a_n \end{bmatrix} \text{ então } e^D = \begin{bmatrix} e^{a_1} & 0 & \dots & 0 \\ 0 & e^{a_2} & \dots & 0 \\ \dots & & \ddots & \\ 0 & 0 & \dots & e^{a_n} \end{bmatrix}.$$

Ou seja, para uma matriz diagonal, calcular a sua exponencial é equivalente a exponenciar cada elemento da diagonal.

◁

Exemplo 12.20 — Matrizes Nilpotentes. Observe que se uma matriz é nilpotente, os elementos do somatório a partir de um certo ponto são nulos, isso implica que a série de e^B é finita. Vejamos um exemplo:

Seja $B = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, então $B^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ e $B^3 = 0$, logo:

$$e^B = I + B + \frac{B^2}{2} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \frac{1}{2} & 1 & 1 \end{bmatrix}$$

◁

Teorema 12.21. [1] *Se* $Q = PAP^{-1}$, *então* $e^Q = Pe^AP^{-1}$.

[2] *Se* $AB = BA$, $e^{A+B} = e^Ae^B$.

[3] $e^{-A} = (e^A)^{-1}$.

[4] *se* $A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ *então* $e^A = e^a \begin{bmatrix} \cos b & \operatorname{sen} b \\ -\operatorname{sen} b & \cos b \end{bmatrix}$

[5] $\dfrac{d(e^{tA})}{dt} = Ae^{tA}$.

Demonstração. No item 1, basta observar que $Q^k = PA^kP^{-1}$ e passar ao limite das somas parciais. Se $AB = BA$, o binômio usual vale para $(A+B)^k$; reorganizando o produto das duas séries absolutamente convergentes, obtemos o item 2. Aplicando-o a A e $-A$, segue $e^Ae^{-A} = e^0 = \mathrm{I}$, o que prova 3.

Para 4, escreva $A = a\,\mathrm{I} + bR$, onde $R = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]$ e $R^2 = -\mathrm{I}$. Separando os termos pares e ímpares na série de e^{bR}, obtemos $e^{bR} = \cos(b)\,\mathrm{I} + \operatorname{sen}(b)R$; o item 2 conclui o cálculo. Finalmente, a série de e^{tA} pode ser derivada termo a termo em qualquer intervalo limitado, e

$$\frac{d}{dt}e^{tA} = \sum_{k=1}^{\infty} \frac{kt^{k-1}A^k}{k!} = A\sum_{j=0}^{\infty} \frac{t^jA^j}{j!} = Ae^{tA},$$

o que prova 5. ∎

Cálculo da Exponencial usando a Forma de Jordan

Se $A = MJM^{-1}$, então $e^A = Me^JM^{-1}$. Em cada bloco de Jordan, escrevemos $J = \lambda\,\mathrm{I} + N$, com N nilpotente. O termo escalar $\lambda\,\mathrm{I}$ comuta com N, logo

$$e^J = e^{\lambda}e^N = e^{\lambda}\sum_{k=0}^{s-1} \frac{N^k}{k!},$$

onde s é o tamanho do bloco. Aplicamos essa fórmula bloco a bloco.

12.4 Equações Diferenciais Ordinárias

Sistemas diferenciais lineares
O sistema $x'(t) = Ax(t)$ é resolvido por $x(t) = e^{tA}x(0)$. A forma canônica de A traduz-se diretamente no comportamento das soluções: crescimento, oscilação, rotação e termos polinomiais.

A passagem do tempo discreto ao contínuo substitui potências de A pela exponencial e^{tA}. Seja $I \subseteq \mathbb{R}$ um intervalo aberto que contém a origem e sejam $x_1(t), \dots, x_n(t) \in C^1(I)$. Consideremos o sistema linear

$$\begin{cases} \dfrac{dx_1}{dt} = a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ \dfrac{dx_n}{dt} = a_{n1}x_1 + \cdots + a_{nn}x_n \end{cases} \tag{12.1}$$

Os coeficientes a_{ij} são reais, e procuramos a solução sujeita à condição inicial $x_1(0) = c_1, \dots, x_n(0) = c_n$.

Para escrevê-lo em notação vetorial, definimos $A = (a_{ij}) \in \mathscr{M}_{n,n}(\mathbb{K})(\mathbb{R})$, $\mathbf{x} = (x_1, \dots, x_n)^{\mathrm{t}}$ e $\vec{c} = (c_1, \dots, c_n)^{\mathrm{t}}$. Com essa notação, o Sistema 12.1 torna-se

$$\mathbf{x}' = A\mathbf{x} \text{ com } \mathbf{x}(0) = \vec{c} \tag{12.2}$$

Para resolver o sistema, substituímos A por uma matriz semelhante mais simples. Suponha $J = PAP^{-1}$, com P invertível, e defina $\mathbf{y} = P\mathbf{x}$. Então $\mathbf{y}' = P\mathbf{x}'$ e $\mathbf{y}(0) = P\vec{c}$; além disso, $J\mathbf{y} = PAP^{-1}(P\mathbf{x}) = PA\mathbf{x} = P\mathbf{x}' = \mathbf{y}'$. O problema reduz-se a

$$\mathbf{y}' = J\mathbf{y} \text{ com } \mathbf{y}(0) = P\vec{c} \tag{12.3}$$

Se $\mathbf{y}$ é uma solução para 12.3, então $\mathbf{x} = P^{-1}\mathbf{y}$ é uma solução para 12.2, pois $\mathbf{x}' = P^{-1}\mathbf{y}' = P^{-1}J\mathbf{y} = A\mathbf{x}$. Além disso, $\mathbf{x}(0) = P^{-1}\mathbf{y}(0) = \vec{c}$.

Escolheremos a matriz J como a Forma Normal de Jordan Real de A e nesse caso as equações que obtemos na versão 12.2 serão fáceis de resolver.

Em primeiro lugar, vimos no Teorema 10.27 que J é uma matriz diagonal por blocos:

$$J = \begin{bmatrix} B_1 & & 0 \\ & \vdots & \\ 0 & & B_K \end{bmatrix} \tag{12.4}$$

Cada bloco em 12.4 possui uma das duas formas possíveis:

$$B_j = \begin{bmatrix} c & 1 & 0 \\ & \ddots & 1 \\ 0 & & c \end{bmatrix} \text{ ou } B_j = \begin{bmatrix} D & I_2 & & 0 \\ & D & \ddots & \\ & & \ddots & I_2 \\ 0 & & & D \end{bmatrix} \tag{12.5}$$

Na equação 12.5, D é uma matriz 2×2 da forma

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

Suponha que o tamanho de cada B_j seja $n_j \times n_j$. Então $\mathbf{y}' = J\mathbf{y}$ se decompõe em K sistemas independentes,

$$\mathbf{y}_j' = B_j \mathbf{y}_j, \quad j = 1, \ldots, K.$$

Com $\mathbf{y}_1 = (y_1, \ \ldots, \ y_{n_1}), \mathbf{y}_2 = (y_{n_1+1}, \ \ldots, \ y_{n_1+n_2})$, etc.

Assim, para encontrar uma solução para a equação 12.2, basta saber resolver as equações dos dois tipos a seguir:

$$\begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix} = \begin{bmatrix} c & 1 & 0 \\ & \ddots & 1 \\ 0 & & c \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{Tipo 1}$$

com $x_i(0) = c_i, i = 1, \ldots, n$ e

$$\begin{bmatrix} x_1' \\ \vdots \\ \vdots \\ x_{2n} \end{bmatrix} = \begin{bmatrix} D & I_2 & & 0 \\ & D & \ddots & \\ & & \ddots & I_2 \\ 0 & & & D \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ \vdots \\ x_{2n} \end{bmatrix} \tag{Tipo 2}$$

com $x_i(0) = c_i, i = 1, \ldots, 2n$, e $D = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$

O item 5 do Teorema 12.21 fornece a solução da equação 12.2, a saber $\mathbf{x} = e^{tA}\vec{c}$.

Teorema 12.22 (Existência e Unicidade para EDO's Lineares).
O sistema $\mathbf{x}' = A\mathbf{x}$ *com* $\mathbf{x}(0) = \vec{c}$ *possui uma única solução:*

$$\mathbf{x} = e^{tA}\vec{c}$$

Demonstração. Verificaremos que $\mathbf{x} = e^{tA}\vec{c}$ é solução. Como $\mathbf{x}' = \dfrac{d(e^{tA}\vec{c})}{dt} = Ae^{tA}\vec{c} = A\mathbf{x}$ e $\mathbf{x}(0) = e^{0A}\vec{c} = \vec{c}$.

Para a unicidade, se $\mathbf{y}$ é qualquer solução, então

$$\frac{d}{dt}\big(e^{-tA}\mathbf{y}(t)\big) = -Ae^{-tA}\mathbf{y}(t) + e^{-tA}A\mathbf{y}(t) = 0.$$

Logo $e^{-tA}\mathbf{y}(t) = \mathbf{y}(0) = \vec{c}$ e, portanto, $\mathbf{y}(t) = e^{tA}\vec{c}$. ■

Queremos ver a forma que essa solução assume nos dois casos especiais Tipo 1 e Tipo 2.

Vamos considerar a equação Tipo 1 primeiro. Para isso considere:

$$B = \begin{bmatrix} c & 1 & 0 \\ & \ddots & 1 \\ 0 & & c \end{bmatrix} \quad \text{e} \quad N = \begin{bmatrix} 0 & 1 & 0 \\ & \ddots & 1 \\ 0 & & 0 \end{bmatrix}$$

Então, $B, N \in \mathscr{M}_{n,n}(\mathbb{K})(\mathbb{R})$ e N é nilpotente de índice n. Como $c\,\mathrm{I}_n$ comuta com N, o Teorema 12.21, item 2, implica

$$e^{tB} = e^{tc\,\mathrm{I}_n}e^{tN} = e^{ct}e^{tN}.$$

Portanto, a solução para Tipo 1 é $\mathbf{x} = e^{tB}\vec{c} = e^{ct}e^{tN}\vec{c}$. Usando a definição de e^{tN}, temos

$$e^{tN} = \sum_{k=0}^{n-1}\frac{(tN)^k}{k!} = \begin{bmatrix} 1 & t & \dfrac{t^2}{2!} & \cdots & \dfrac{t^{n-2}}{(n-2)!} & \dfrac{t^{n-1}}{(n-1)!} \\ 0 & 1 & t & \cdots & \dfrac{t^{n-3}}{(n-3)!} & \dfrac{t^{n-2}}{(n-2)!} \\ 0 & 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & t & \dfrac{t^2}{2!} \\ 0 & 0 & \cdots & 0 & 1 & t \\ 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix} \tag{12.6}$$

Substituindo 12.6 em $\mathbf{x} = e^{ct}e^{tN}\vec{c}$, obtemos

$$x_j(t) = e^{ct}\sum_{k=0}^{n-j}\frac{t^k}{k!}c_{j+k} \quad \text{para } j = 1, \ldots, n. \tag{12.7}$$

Agora vamos considerar a equação Tipo 2. Se a matriz nessa equação é apenas D, então o Teorema 12.21, item 4, implica que a solução é

$$x = e^{tD}\vec{c} = e^{ta}\begin{bmatrix} \cos bt & \operatorname{sen} bt \\ -\operatorname{sen} bt & \cos bt \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \tag{12.8}$$

Portanto, podemos assumir $n > 1$ e prosseguir com o caso geral.

Defina $\mu = a - bi$. Para $j = 1, \ldots, n$, seja $z_j = x_{2j-1} + ix_{2j}$ e $w_j = c_{2j-1} + ic_{2j}$. Então a equação Tipo 2 se torna o seguinte sistema:

$$\begin{bmatrix} \mu & 1 & & 0 & 0 \\ 0 & \mu & \ddots & & 0 \\ \vdots & \ddots & \ddots & 1 & \vdots \\ 0 & & 0 & \mu & 1 \\ 0 & & & 0 & \mu \end{bmatrix}\begin{bmatrix} z_1 \\ \\ \vdots \\ \\ z_n \end{bmatrix} = \begin{bmatrix} z_1' \\ \vdots \\ z_n' \end{bmatrix} \tag{12.9}$$

com $z_i(0) = w_i$ para $i = 1, \dots, n$. As equações em 12.9 são resolvidas da mesma maneira que a equação Tipo 1. Portanto,

$$z_j(t) = e^{\mu t} \sum_{k=0}^{n-j} \frac{t^k}{k!} w_{j+k}, \qquad j = 1, \dots, n. \tag{12.10}$$

Como $e^{\mu t} = e^{(a-bi)t} = e^{at}(\cos bt - i \operatorname{sen} bt)$, a Equação 12.10, com $w_j = c_{2j-1} + ic_{2j}$, fornece a solução real:

$$x_{2j-1}(t) = e^{at} \sum_{k=0}^{n-j} \frac{t^k}{k!} \left[c_{2(j+k)-1} \cos(bt) + c_{2(j+k)} \operatorname{sen}(bt) \right] \tag{12.11}$$

$$x_{2j}(t) = e^{at} \sum_{k=0}^{n-j} \frac{t^k}{k!} \left[c_{2(j+k)} \cos(bt) - c_{2(j+k)-1} \operatorname{sen}(bt) \right], \qquad j = 1, \dots, n.$$

Observe que, nos casos Tipo 1 e Tipo 2, as coordenadas da solução são combinações lineares de produtos de polinômios, exponenciais, senos e cossenos. Assim, temos o seguinte teorema:

> **Teorema 12.23.**
> *Sejam $A \in \mathscr{M}_{n,n}(\mathbb{R})$ e $\mathbf{x}(t)$ uma solução de $\mathbf{x}' = A\mathbf{x}$. Então cada coordenada de $\mathbf{x}$ é uma combinação linear de funções da forma $t^k e^{at} \cos(bt)$ e $t^k e^{at} \operatorname{sen}(bt)$, onde $a + bi$ percorre os autovalores complexos de A e k é um inteiro não negativo menor que o tamanho do bloco correspondente.*

Exemplo 12.24 — Um sistema de rotação. Considere

$$\mathbf{x}' = A\mathbf{x}, \qquad A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \qquad \mathbf{x}(0) = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

Como $A^2 = -\mathrm{I}$, a separação dos termos pares e ímpares da série exponencial dá

$$e^{tA} = \cos(t)\,\mathrm{I} + \operatorname{sen}(t) A = \begin{bmatrix} \cos t & -\operatorname{sen} t \\ \operatorname{sen} t & \cos t \end{bmatrix}.$$

Portanto,

$$\mathbf{x}(t) = \begin{bmatrix} c_1 \cos t - c_2 \operatorname{sen} t \\ c_1 \operatorname{sen} t + c_2 \cos t \end{bmatrix}.$$

Além de verificar concretamente a fórmula $\mathbf{x}(t) = e^{tA}\mathbf{x}(0)$, o exemplo mostra que o fluxo preserva a norma euclidiana: as trajetórias são círculos centrados na origem.

◁

12.5 Teorema de Perron–Frobenius

Autovalor dominante e positividade
Uma matriz com entradas estritamente positivas possui um autovalor real positivo maior, em módulo, que os demais, e um autovetor associado com coordenadas positivas. Essas propriedades permitem estudar as iterações da matriz e o comportamento assintótico de modelos definidos por ela.

Usaremos a seguinte forma finito-dimensional do Teorema do Ponto Fixo de Brouwer. Sua demonstração é topológica e foge ao escopo deste texto.

Teorema 12.25 (Ponto Fixo de Brouwer).
Toda aplicação contínua de um simplex compacto em si mesmo possui um ponto fixo.

Definição 12.26. Uma matriz $A = (a_{ij}) \in \mathscr{M}_{n,n}(\mathbb{R})$ é **positiva** se $a_{ij} > 0$ para todos i, j, e é **não negativa** se $a_{ij} \geq 0$. A mesma terminologia será usada para vetores, coordenada a coordenada.

O resultado seguinte caracteriza o autovalor dominante de uma matriz positiva.

Teorema 12.27 (Perron–Frobenius para matrizes positivas).
Se $A \in \mathscr{M}_{n,n}(\mathbb{R})$ é positiva, então existe $r > 0$ tal que:

1. *r é autovalor de A e possui um autovetor positivo $\mathbf{u}$;*
2. *o autoespaço de r é unidimensional e r é uma raiz simples de c_A;*
3. *todo outro autovalor $\lambda \in \mathbb{C}$ satisfaz $|\lambda| < r$;*
4. *todo autovetor não negativo está associado a r.*

Demonstração. Considere o simplex $\Delta = \{\mathbf{x} \geq 0 \mid \sum_i x_i = 1\}$. A aplicação contínua

$$F(\mathbf{x}) = \frac{A\mathbf{x}}{\sum_i (A\mathbf{x})_i}$$

leva Δ em si. Pelo Teorema 12.25, existe $\mathbf{u} \in \Delta$ com $F(\mathbf{u}) = \mathbf{u}$. Assim, $A\mathbf{u} = r\mathbf{u}$ para algum $r > 0$; como A é positiva, $\mathbf{u} > 0$.

Se $A\mathbf{z} = \lambda\mathbf{z}$, escolha $c = \max_i |z_i|/u_i$ e um índice i_0 em que o máximo é atingido. Então

$$|\lambda| c u_{i_0} = |(A\mathbf{z})_{i_0}| \leq \sum_j a_{i_0 j} |z_j| \leq c r u_{i_0},$$

logo $|\lambda| \leq r$. Se houver igualdade, todas as desigualdades triangulares acima são igualdades. Como $a_{ij} > 0$, todas as coordenadas de $\mathbf{z}$ têm a mesma fase e $\mathbf{z}$ é múltiplo complexo de $\mathbf{u}$; em particular, $\lambda = r$. Isso prova a dominância estrita e a unicidade do autoespaço.

Aplicando o argumento a A^{t}, obtemos um autovetor positivo à esquerda $\mathbf{w}$, com $\mathbf{w}^{\mathsf{t}} A = r\mathbf{w}^{\mathsf{t}}$. Se existisse um vetor generalizado $\mathbf{y}$ com $(A - r\,\mathrm{I})\mathbf{y} = \mathbf{u}$, então $0 = \mathbf{w}^{\mathsf{t}}(A - r\,\mathrm{I})\mathbf{y} = \mathbf{w}^{\mathsf{t}}\mathbf{u} > 0$, contradição. Logo r é simples. Finalmente, se $A\mathbf{z} = \lambda\mathbf{z}$ com $\mathbf{z} \geq 0$ e $\mathbf{z} \neq 0$, então

$$\lambda \mathbf{w}^{\mathsf{t}}\mathbf{z} = \mathbf{w}^{\mathsf{t}} A \mathbf{z} = r\mathbf{w}^{\mathsf{t}}\mathbf{z}.$$

Como $\mathbf{w}^{\mathsf{t}}\mathbf{z} > 0$, segue que $\lambda = r$. ∎

12.6 Matrizes Estocásticas

Probabilidade preservada
Uma matriz estocástica envia distribuições de probabilidade em distribuições de probabilidade. Vetores estacionários são autovetores associados ao autovalor 1; a convergência depende de os demais modos espectrais desaparecerem com as iterações.

Definição 12.28. Uma matriz $A = (a_{ij}) \in \mathscr{M}_{n,n}(\mathbb{R})$ é **estocástica por linhas** se $a_{ij} \geq 0$ e $\sum_j a_{ij} = 1$ para todo i. Ela é **regular** ou **primitiva** se A^m é positiva para algum $m \geq 1$.

Proposição 12.29.
Se A é estocástica por linhas, então 1 é autovalor de A e todo autovalor λ satisfaz $|\lambda| \leq 1$.

Demonstração. Para $\mathbf{1} = (1, \dots, 1)^{\mathrm{t}}$, temos $A\mathbf{1} = \mathbf{1}$. Se $A\mathbf{z} = \lambda \mathbf{z}$ e $|z_{i_0}| = \max_i |z_i|$, então

$$|\lambda||z_{i_0}| = \left| \sum_j a_{i_0 j} z_j \right| \leq \sum_j a_{i_0 j} |z_j| \leq |z_{i_0}|.$$

Como $\mathbf{z} \neq 0$, segue que $|\lambda| \leq 1$. ∎

Teorema 12.30 (Convergência de matrizes estocásticas regulares).
Se A é estocástica por linhas e regular, então 1 é um autovalor simples, todos os demais autovalores têm módulo menor que 1 e existe um único vetor positivo $\boldsymbol{\pi}$ tal que

$$A^{\mathrm{t}} \boldsymbol{\pi} = \boldsymbol{\pi}, \qquad \sum_i \pi_i = 1.$$

Além disso,

$$\lim_{k \to \infty} A^k = \mathbf{1}\boldsymbol{\pi}^{\mathrm{t}},$$

uma matriz estocástica positiva cujas linhas são todas iguais a $\boldsymbol{\pi}^{\mathrm{t}}$.

Demonstração. Escolha m com $A^m > 0$. A matriz A^m também é estocástica, logo seu autovalor de Perron é 1. Pelo Teorema 12.27, ele é simples e domina estritamente os demais autovalores de A^m. Se λ é autovalor de A, então λ^m é autovalor de A^m; portanto, $\lambda = 1$ ou $|\lambda| < 1$. A simplicidade de 1 para A^m também exclui blocos de Jordan não triviais para 1 em A.

Aplicando Perron–Frobenius a $(A^{\mathrm{t}})^m$, obtemos o vetor positivo $\boldsymbol{\pi}$, único após a normalização indicada. Na forma de Jordan complexa de A, o bloco associado a 1 é $[1]$ e todos os demais blocos tendem a zero quando elevados à k-ésima potência. Assim, A^k converge para a projeção espectral sobre $\langle \mathbf{1} \rangle$ ao longo de $\ker(\boldsymbol{\pi}^{\mathrm{t}})$, que é precisamente $\mathbf{1}\boldsymbol{\pi}^{\mathrm{t}}$. ∎

13

Teoremas Espectrais e Decomposições

As potências, o núcleo e a imagem de uma matriz diagonal podem ser calculados diretamente a partir de suas entradas. Para obter uma descrição semelhante de um operador, procuramos uma base de autovetores. Quando essa base é ortonormal, a representação em coordenadas também preserva o produto interno do espaço.

O teorema espectral caracteriza os operadores que admitem tal base. Sobre os complexos, são exatamente os operadores normais; sobre os reais, são os operadores autoadjuntos. Desse resultado decorrem a forma canônica dos operadores normais reais, a descrição das isometrias, a raiz quadrada de um operador positivo, a decomposição polar e a decomposição em valores singulares.

Em todo o capítulo, V e W serão espaços com produto interno de dimensão finita sobre $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. Mantemos a convenção do Capítulo 6: o produto interno é linear na primeira variável e conjugado-linear na segunda. Para uma matriz complexa A, escreveremos

$$A^* = \overline{A}^{\mathrm{t}}.$$

No caso real, essa operação se reduz à transposição.

Nesse capítulo

13.1 Operadores Autoadjuntos e Normais

A adjunta relaciona um operador ao produto interno do espaço. As classes introduzidas a seguir são definidas por relações entre o operador e sua adjunta.

Definição 13.1 — Operadores Autoadjuntos e Normais. Seja $T \in \mathrm{Hom}(V, V)$. Dizemos que T é

a **autoadjunto** se $T^* = T$;

b **anti-autoadjunto** se $T^* = -T$;

c **normal** se $T^*T = TT^*$.

Em espaços complexos, os operadores autoadjuntos também são chamados de *hermi-*

Normalidade e diagonalização
Em bases ortonormais, operadores autoadjuntos complexos são representados por matrizes hermitianas. A condição $TT^* = T^*T$ caracteriza os operadores normais. Em dimensão finita sobre $\mathbb{C}$, ela equivale à existência de uma base ortonormal de autovetores.

tianos. Em espaços reais, suas matrizes em bases ortonormais são simétricas. Todo operador autoadjunto ou anti-autoadjunto é normal, mas a classe dos operadores normais é maior.

Exemplo 13.2. Em $\mathbb{C}^2$, os operadores representados na base canônica pelas matrizes

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

são normais. O primeiro é autoadjunto, pois $A^* = A$. O segundo não é autoadjunto, mas é normal porque B e B^* são diagonais e, portanto, comutam.

◁

Teorema 13.3 (Caracterização Matricial).
Seja $\underline{v}$ uma base ortonormal de V e seja $A = [T]_{\underline{v}}$. Então:

a *T é autoadjunto se, e somente se, $A = A^*$;*

b *T é anti-autoadjunto se, e somente se, $A = -A^*$;*

c *T é normal se, e somente se, $A^*A = AA^*$.*

Demonstração. Em uma base ortonormal, a matriz da adjunta é $[T^*]_{\underline{v}} = A^*$. As duas primeiras equivalências seguem imediatamente dessa identidade. Para a terceira, observamos que

$$[T^*T]_{\underline{v}} = A^*A \qquad \text{e} \qquad [TT^*]_{\underline{v}} = AA^*.$$

Como operadores são iguais se, e somente se, têm a mesma matriz em uma base fixada, o resultado segue. ■

Observação 13.4. A condição de normalidade depende do produto interno, pois a adjunta também depende dele. Já a diagonalizabilidade usual depende apenas da estrutura de espaço vetorial. O teorema espectral relacionará essas duas informações ao exigir uma diagonalização por uma base *ortonormal*.

Proposição 13.5.
Para todo $T \in \mathrm{Hom}(V, V)$, os operadores T^*T e TT^* são autoadjuntos. Além disso,

$$\langle T^*T\mathbf{v}, \mathbf{v}\rangle = \|T\mathbf{v}\|^2 \qquad e \qquad \langle TT^*\mathbf{v}, \mathbf{v}\rangle = \|T^*\mathbf{v}\|^2.$$

Demonstração. Das propriedades da adjunta,

$$(T^*T)^* = T^*(T^*)^* = T^*T.$$

A prova para TT^* é idêntica. Finalmente, pela definição de adjunta,

$$\langle T^*T\mathbf{v}, \mathbf{v}\rangle = \langle T\mathbf{v}, T\mathbf{v}\rangle = \|T\mathbf{v}\|^2,$$

e a segunda identidade é obtida trocando T por T^*. ■

Os próximos resultados descrevem as restrições que a adjunta impõe aos autovalores e aos autovetores. Essas relações serão usadas nas demonstrações dos teoremas espectrais.

Teorema 13.6 (Autovalores de Operadores Autoadjuntos).
Todo autovalor de um operador autoadjunto é real. Além disso, autovetores associados a autovalores distintos são ortogonais.

Demonstração. Suponha $T\mathbf{v} = \lambda\mathbf{v}$, com $\mathbf{v} \neq \vec{0}$. Como $T = T^*$,

$$\lambda\|\mathbf{v}\|^2 = \langle T\mathbf{v}, \mathbf{v}\rangle = \langle \mathbf{v}, T\mathbf{v}\rangle = \overline{\lambda}\|\mathbf{v}\|^2.$$

Como $\|\mathbf{v}\| > 0$, temos $\lambda = \overline{\lambda}$; portanto, $\lambda \in \mathbb{R}$.

Agora sejam $T\mathbf{v} = \lambda\mathbf{v}$ e $T\mathbf{w} = \mu\mathbf{w}$, com $\lambda \neq \mu$. Os dois autovalores são reais, e

$$\lambda\langle \mathbf{v}, \mathbf{w}\rangle = \langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, T\mathbf{w}\rangle = \mu\langle \mathbf{v}, \mathbf{w}\rangle.$$

Logo $(\lambda - \mu)\langle \mathbf{v}, \mathbf{w}\rangle = 0$, o que implica $\langle \mathbf{v}, \mathbf{w}\rangle = 0$. ∎

Para operadores normais, a relação entre os autovetores de T e de T^* é dada pelo lema seguinte.

Lema 13.7 (Autovetores de um Operador Normal).
Se T é normal, então, para todo $\lambda \in \mathbb{K}$,

$$\ker(T - \lambda I) = \ker(T^* - \overline{\lambda} I).$$

Em particular, se $T\mathbf{v} = \lambda\mathbf{v}$, então $T^*\mathbf{v} = \overline{\lambda}\mathbf{v}$.

Demonstração. O operador $S = T - \lambda I$ é normal, pois $S^* = T^* - \overline{\lambda} I$ e os termos cruzados nas expressões S^*S e SS^* coincidem. Para todo $\mathbf{v} \in V$, a normalidade fornece

$$\|S\mathbf{v}\|^2 = \langle S^*S\mathbf{v}, \mathbf{v}\rangle = \langle SS^*\mathbf{v}, \mathbf{v}\rangle = \|S^*\mathbf{v}\|^2.$$

Assim, $S\mathbf{v} = \vec{0}$ se, e somente se, $S^*\mathbf{v} = \vec{0}$, o que é exatamente a igualdade dos núcleos anunciada. ∎

Corolário 13.8 (Ortogonalidade dos Autoespaços Normais).
Se T é normal, autovetores associados a autovalores distintos são ortogonais.

Demonstração. Se $T\mathbf{v} = \lambda\mathbf{v}$ e $T\mathbf{w} = \mu\mathbf{w}$, o lema anterior dá $T^*\mathbf{w} = \overline{\mu}\mathbf{w}$. Portanto,

$$\lambda\langle \mathbf{v}, \mathbf{w}\rangle = \langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, T^*\mathbf{w}\rangle = \mu\langle \mathbf{v}, \mathbf{w}\rangle.$$

Se $\lambda \neq \mu$, segue que $\mathbf{v} \perp \mathbf{w}$. ∎

Corolário 13.9.
Se T é normal, então $\ker T = \ker T^*$ *e* $\operatorname{im} T = \operatorname{im} T^*$.

Demonstração. A igualdade dos núcleos é o Lema 13.7 com $\lambda = 0$. Usando as relações entre núcleo, imagem e adjunta demonstradas no Capítulo 6, obtemos

$$\operatorname{im} T = (\ker T^*)^\perp = (\ker T)^\perp = \operatorname{im} T^*.$$

■

13.2 Teoremas Espectrais

Diagonalização em bases ortonormais
Uma diagonalização usa uma base de autovetores. Nos casos cobertos pelo teorema espectral, essa base pode ser escolhida ortonormal. A passagem às coordenadas nessa base preserva o produto interno e, portanto, os comprimentos e os ângulos.

Diagonalizar um operador significa decompor o espaço em retas invariantes. Os teoremas a seguir caracterizam os operadores para os quais essas retas podem ser escolhidas mutuamente ortogonais. Normalizando um vetor em cada reta, obtemos uma base ortonormal de autovetores.

Definição 13.10 — Diagonalização Ortonormal. Um operador $T \in \operatorname{Hom}(V, V)$ é **diagonalizável por uma base ortonormal** se existe uma base ortonormal de V formada por autovetores de T. Equivalentemente, a matriz de T nessa base é diagonal.

Teorema 13.11 (Teorema Espectral Complexo).
Se V é complexo, um operador $T \in \operatorname{Hom}(V, V)$ é normal se, e somente se, V possui uma base ortonormal formada por autovetores de T.

Demonstração. Suponha primeiro que T possua uma base ortonormal de autovetores. Nessa base, T e T^* são representados, respectivamente, por

$$\operatorname{diag}(\lambda_1, \dots, \lambda_n) \quad \text{e} \quad \operatorname{diag}(\overline{\lambda}_1, \dots, \overline{\lambda}_n).$$

As duas matrizes comutam; logo T é normal.

Reciprocamente, suponha que T seja normal. A prova será por indução em $n = \dim V$. O caso $n = 1$ é imediato. Como o corpo é $\mathbb{C}$, o polinômio característico de T possui uma raiz. Portanto, existem $\lambda \in \mathbb{C}$ e um vetor unitário $\mathbf{e}_1$ tais que $T\mathbf{e}_1 = \lambda \mathbf{e}_1$. Pelo Lema 13.7, $T^*\mathbf{e}_1 = \overline{\lambda}\mathbf{e}_1$.

A reta $L = \langle \mathbf{e}_1 \rangle$ e seu complemento ortogonal $L^\perp$ são invariantes por T e por T^*. De fato, se $\mathbf{w} \in L^\perp$, então

$$\langle T\mathbf{w}, \mathbf{e}_1 \rangle = \langle \mathbf{w}, T^*\mathbf{e}_1 \rangle = 0,$$

e o mesmo argumento, trocando T por T^*, mostra a outra invariância. Assim, a restrição $T_1 = T|_{L^\perp}$ é normal; sua adjunta é justamente $T^*|_{L^\perp}$. Pela hipótese de indução, $L^\perp$ possui uma base ortonormal $\mathbf{e}_2, \dots, \mathbf{e}_n$ formada por autovetores de T_1. Juntando $\mathbf{e}_1$ a essa base, obtemos a base ortonormal desejada de V. ■

No caso real, um operador normal pode conter planos invariantes nos quais atua como uma rotação. Por isso, normalidade não basta para garantir uma base real de autovetores. A diagonalização em uma base ortonormal real é caracterizada pela autoadjunção.

Teorema 13.12 (Teorema Espectral Real).
Se V é real, um operador $T \in \mathrm{Hom}(V,V)$ é autoadjunto se, e somente se, V possui uma base ortonormal formada por autovetores de T.

Demonstração. Se existe uma base ortonormal de autovetores, a matriz de T nessa base é diagonal e real. Ela é, portanto, simétrica, e o Teorema 13.3 mostra que T é autoadjunto.

Para a recíproca, começamos justificando a existência de um autovetor real. Escolha uma base ortonormal e seja A a matriz real simétrica de T. Vista como matriz complexa, A é hermitiana e, pelo teorema espectral complexo, possui um autovetor $\mathbf{z} = \mathbf{x} + i\mathbf{y}$ associado a um autovalor real λ. Da igualdade

$$A(\mathbf{x} + i\mathbf{y}) = \lambda(\mathbf{x} + i\mathbf{y})$$

seguem $A\mathbf{x} = \lambda\mathbf{x}$ e $A\mathbf{y} = \lambda\mathbf{y}$. Como $\mathbf{z} \neq \vec{0}$, pelo menos um dos vetores $\mathbf{x}, \mathbf{y}$ é não nulo. Assim, T possui um autovetor real, que podemos normalizar e denotar por $\mathbf{e}_1$.

Agora procedemos por indução em $\dim V$. Como $T = T^*$, o complemento ortogonal de $\langle \mathbf{e}_1 \rangle$ é invariante por T, e a restrição de T a esse complemento continua autoadjunta. A hipótese de indução fornece uma base ortonormal de autovetores no complemento. Acrescentando $\mathbf{e}_1$, obtemos uma base ortonormal de autovetores de V. ∎

Definição 13.13 — Matrizes Unitárias e Ortogonais. Uma matriz complexa Q é **unitária** se $Q^*Q = I$. Uma matriz real Q é **ortogonal** se $Q^{\mathrm{t}}Q = I$.

Corolário 13.14 (Versões Matriciais).
Valem as seguintes afirmações:

a *uma matriz complexa A é normal se, e somente se, existe uma matriz unitária Q tal que Q^*AQ é diagonal;*

b *uma matriz real A é simétrica se, e somente se, existe uma matriz ortogonal Q tal que $Q^{\mathrm{t}}AQ$ é diagonal.*

Demonstração. As colunas de Q são os vetores de uma base ortonormal escritos na base canônica. A expressão $Q^{-1}AQ$ é a matriz do mesmo operador na nova base. Como $Q^{-1} = Q^*$ no caso complexo e $Q^{-1} = Q^{\mathrm{t}}$ no caso real, as afirmações são exatamente as versões matriciais dos dois teoremas espectrais. ∎

Exemplo 13.15 — Diagonalização de uma Matriz Hermitiana. Considere

$$A = \begin{bmatrix} 2 & i \\ -i & 2 \end{bmatrix}.$$

Seus autovalores são 1 e 3. Autovetores unitários correspondentes são

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} -i \\ 1 \end{bmatrix}, \qquad \mathbf{v}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} i \\ 1 \end{bmatrix}.$$

Eles são ortogonais. Se $Q = [\,\mathbf{v}_1\ \mathbf{v}_2\,]$, então Q é unitária e

$$Q^*AQ = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}.$$

O exemplo ilustra a diferença entre diagonalizar e diagonalizar unitariamente: a segunda operação controla também os comprimentos e os ângulos.

◁

13.2.1 Decomposição Espectral e Cálculo Funcional

Cálculo funcional e autovalores
Se $T = \sum_j \lambda_j P_j$, então uma função adequada de T é obtida aplicando a função aos autovalores: $f(T) = \sum_j f(\lambda_j) P_j$. Os projetores preservam os autoespaços; apenas os valores espectrais são transformados.

O teorema espectral pode ser escrito sem escolher vetores dentro de cada autoespaço. Essa formulação é especialmente útil quando um autovalor tem multiplicidade maior que um.

> **Teorema 13.16** (Decomposição Espectral).
> *Se T é normal em um espaço complexo e $\lambda_1, \dots, \lambda_r$ são seus autovalores distintos, então*
>
> $$V = E_{\lambda_1} \odot \cdots \odot E_{\lambda_r},$$
>
> *onde $E_{\lambda_j} = \ker(T - \lambda_j I)$. Se P_j denota a projeção ortogonal sobre E_{λ_j}, então*
>
> $$T = \sum_{j=1}^{r} \lambda_j P_j, \qquad I = \sum_{j=1}^{r} P_j, \qquad P_i P_j = 0 \quad (i \neq j).$$
>
> *As mesmas conclusões valem para operadores autoadjuntos reais.*

Demonstração. Escolha uma base ortonormal de autovetores e agrupe seus elementos segundo o autovalor. O Corolário 13.8 mostra que os autoespaços associados a autovalores distintos são ortogonais. Se $\mathbf{v} = \mathbf{v}_1 + \cdots + \mathbf{v}_r$, com $\mathbf{v}_j = P_j \mathbf{v} \in E_{\lambda_j}$, então

$$T\mathbf{v} = \sum_{j=1}^{r} T\mathbf{v}_j = \sum_{j=1}^{r} \lambda_j P_j \mathbf{v}.$$

Como essa igualdade vale para todo $\mathbf{v}$, obtemos a fórmula para T. As demais identidades seguem das propriedades das projeções ortogonais. ■

Se f é uma função definida no conjunto finito $\{\lambda_1, \dots, \lambda_r\}$, definimos

$$f(T) = \sum_{j=1}^{r} f(\lambda_j) P_j.$$

Quando f é um polinômio, essa definição coincide com a substituição usual de T no polinômio. A vantagem é que também podemos usar funções como $f(t) = \sqrt{t}$ ou $f(t) = 1/t$, desde que estejam definidas no espectro de T.

Exemplo 13.17. Se T é autoadjunto e possui espectro $\{-2, 1, 4\}$, então

$$T^2 = 4P_{-2} + P_1 + 16P_4.$$

Se T for invertível,

$$T^{-1} = -\frac{1}{2}P_{-2} + P_1 + \frac{1}{4}P_4.$$

Essas fórmulas são independentes da escolha de bases dentro dos autoespaços.

◁

Proposição 13.18 (Diagonalização Simultânea).
Se S e T são operadores autoadjuntos e $ST = TS$, então existe uma base ortonormal formada por autovetores simultâneos de S e de T.

Demonstração. Pelo teorema espectral, V é a soma ortogonal dos autoespaços de T. A comutação mostra que cada um desses autoespaços é invariante por S: se $T\mathbf{v} = \lambda\mathbf{v}$, então

$$T(S\mathbf{v}) = S(T\mathbf{v}) = \lambda S\mathbf{v}.$$

A restrição de S a cada autoespaço de T é autoadjunta. Diagonalizando essas restrições e reunindo as bases ortonormais obtidas, construímos uma base de autovetores simultâneos. ■

13.2.2 Triangularização Unitária

Leitura opcional. Esta subseção pode ser omitida em uma primeira leitura. A decomposição de Schur não é necessária para os resultados posteriores do capítulo. Ela estabelece que todo operador complexo admite uma matriz triangular em uma base ortonormal, mesmo quando não admite diagonalização.

O teorema seguinte vale para qualquer operador complexo, sem a hipótese de normalidade. A representação obtida é, em geral, triangular, e não diagonal.

Teorema 13.19 (Decomposição de Schur).
Se $A \in \mathcal{M}_{n,n}(\mathbb{C})$, existe uma matriz unitária Q tal que Q^*AQ é triangular superior.

Demonstração. Apresentamos a versão para operadores. A prova é por indução em n. Escolha um autovetor unitário $\mathbf{e}_1$ de T e escreva $V = \langle \mathbf{e}_1 \rangle \odot \langle \mathbf{e}_1 \rangle^{\perp}$. Seja P a projeção ortogonal sobre $\langle \mathbf{e}_1 \rangle^{\perp}$ e considere o operador comprimido

$$S = P \circ T|_{\langle \mathbf{e}_1 \rangle^{\perp}}.$$

Pela hipótese de indução, existe uma base ortonormal $\mathbf{e}_2, \dots, \mathbf{e}_n$ do complemento na qual S é triangular superior. Como a reta gerada por $\mathbf{e}_1$ é T-invariante, a matriz de T na base $\mathbf{e}_1, \dots, \mathbf{e}_n$ também é triangular superior. Tomando para colunas de Q os vetores dessa base, obtemos Q^*AQ triangular superior. ■

Para uma matriz normal, a forma triangular de Schur é necessariamente diagonal. De fato, se $R = (r_{ij})$ é triangular superior e normal, então $\|R\mathbf{e}_1\| = \|R^*\mathbf{e}_1\|$ implica

$$|r_{11}|^2 = |r_{11}|^2 + |r_{12}|^2 + \cdots + |r_{1n}|^2.$$

Logo a primeira linha não possui entradas fora da diagonal. O argumento se repete no bloco restante. Assim, a decomposição de Schur também oferece uma prova matricial do teorema espectral complexo.

Por que aparecem blocos 2×2
O bloco $\left[\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right]$ representa a multiplicação pelo número complexo $a + ib$ vista no plano real. Seu módulo mede a dilatação e seu argumento, a rotação.

13.3 Operadores Normais Reais

Uma rotação não trivial do plano real é normal, mas não possui autovetores reais. Em vez de uma diagonalização, devemos esperar uma decomposição em retas e planos ortogonais invariantes.

Ideia geométrica.

Ao complexificar o espaço, cada autovalor real ainda fornece uma direção real. Já um par de autovalores não reais conjugados $a \pm ib$ fornece um plano real: nesse plano, o operador é a composição de uma rotação com uma dilatação. A prova consiste em recuperar essas retas e esses planos a partir de uma base ortonormal complexa e verificar que continuam mutuamente ortogonais no espaço real.

Teorema 13.20 (Forma Canônica dos Operadores Normais Reais).
Se V é real, um operador $T \in \mathrm{Hom}(V,V)$ é normal se, e somente se, existe uma base ortonormal na qual sua matriz é diagonal por blocos, sendo cada bloco de um dos tipos

$$[\lambda] \qquad \text{ou} \qquad \begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \quad b > 0.$$

Demonstração. Suponha que T seja normal e estenda os escalares de $\mathbb{R}$ para $\mathbb{C}$. O operador complexificado $T_{\mathbb{C}}$ continua normal e, pelo teorema espectral complexo, possui uma base ortonormal de autovetores.

Se λ é real, o autoespaço complexo correspondente é a complexificação de $\ker(T - \lambda I)$. Escolhendo uma base ortonormal nesse autoespaço real, obtemos blocos unidimensionais $[\lambda]$.

Os autovalores não reais aparecem em pares conjugados. Fixe $\lambda = a + ib$, com $b > 0$, e escolha uma base ortonormal $\mathbf{z}_1, \dots, \mathbf{z}_k$ de seu autoespaço. Escreva $\mathbf{z}_j = \mathbf{x}_j - i\mathbf{y}_j$, com $\mathbf{x}_j, \mathbf{y}_j \in V$. Da igualdade $T_{\mathbb{C}}\mathbf{z}_j = \lambda \mathbf{z}_j$ obtemos

$$T\mathbf{x}_j = a\mathbf{x}_j + b\mathbf{y}_j, \qquad T\mathbf{y}_j = -b\mathbf{x}_j + a\mathbf{y}_j.$$

Como $\overline{\mathbf{z}_1}, \dots, \overline{\mathbf{z}_k}$ são autovetores associados a $\overline{\lambda}$, a ortogonalidade dos autoespaços de um operador normal fornece

$$\langle \mathbf{z}_j, \mathbf{z}_\ell \rangle = \delta_{j\ell}, \qquad \langle \mathbf{z}_j, \overline{\mathbf{z}_\ell} \rangle = 0.$$

Separando partes real e imaginária nessas duas identidades, obtemos

$$\langle \mathbf{x}_j, \mathbf{x}_\ell \rangle = \langle \mathbf{y}_j, \mathbf{y}_\ell \rangle = \frac{1}{2}\delta_{j\ell}, \qquad \langle \mathbf{x}_j, \mathbf{y}_\ell \rangle = 0.$$

Portanto, os vetores $\sqrt{2}\,\mathbf{x}_j, \sqrt{2}\,\mathbf{y}_j$ formam um conjunto ortonormal real. Em cada plano gerado pelo par correspondente, a matriz da restrição de T é

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

Autoespaços correspondentes a diferentes pares conjugados são ortogonais; logo o mesmo cálculo produz planos reais mutuamente ortogonais. Agrupando esses planos

com os autoespaços reais, obtemos uma decomposição ortogonal de V e a matriz anunciada.

Reciprocamente, cada bloco exibido comuta com sua transposta. Uma matriz diagonal por blocos desse tipo também comuta com sua transposta e, pelo Teorema 13.3, representa um operador normal. ■

Exemplo 13.21 — Rotação e Dilatação. A matriz

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

representa, no plano, uma rotação seguida de uma dilatação de fator $\sqrt{13}$. Ela é normal, pois $A^{\mathrm{t}}A = AA^{\mathrm{t}} = 13I$, mas não é simétrica. Seus autovalores complexos são $2 \pm 3i$; por isso, não há uma base real de autovetores.

◁

Pelo Teorema 13.20, um operador normal real é autoadjunto se, e somente se, não aparecem blocos bidimensionais na forma canônica. Nesse caso, a matriz na base ortonormal é diagonal.

13.4 Operadores Unitários e Ortogonais

As isometrias lineares são as transformações que preservam comprimentos. A identidade de polarização mostra que, para uma transformação linear, preservar todos os comprimentos equivale a preservar todo o produto interno.

Definição 13.22 — Isometria Linear. Um operador $U \in \mathrm{Hom}(V, V)$ é uma **isometria linear** se

$$\|U\mathbf{v}\| = \|\mathbf{v}\| \qquad \text{para todo } \mathbf{v} \in V.$$

Em espaços complexos, uma isometria linear é chamada de operador **unitário**; em espaços reais, de operador **ortogonal**.

Teorema 13.23 (Caracterizações das Isometrias).
Para $U \in \mathrm{Hom}(V, V)$, são equivalentes:

a. *U é uma isometria;*
b. *$\langle U\mathbf{v}, U\mathbf{w}\rangle = \langle \mathbf{v}, \mathbf{w}\rangle$ para todos $\mathbf{v}, \mathbf{w} \in V$;*
c. *$U^*U = I$;*
d. *U é invertível e $U^{-1} = U^*$;*
e. *U envia toda base ortonormal de V em uma base ortonormal.*

Demonstração. A implicação de [a] para [b] segue das identidades de polarização: todos os termos do lado direito dessas identidades são normas, e U preserva somas e multiplicações por escalares por ser linear. A implicação inversa é obtida tomando $\mathbf{w} = \mathbf{v}$.

Além disso,

$$\langle U\mathbf{v}, U\mathbf{w}\rangle = \langle \mathbf{v}, U^*U\mathbf{w}\rangle.$$

Logo [b] equivale a $U^*U = I$, o que prova a equivalência com [c]. Essa identidade mostra que U é injetivo; como V tem dimensão finita, U é invertível e $U^* = U^{-1}$. Isso fornece [d], cuja recíproca é imediata.

Finalmente, a preservação do produto interno mostra que a imagem de qualquer base ortonormal é ortonormal. Como U é invertível, essa imagem é uma base. Reciprocamente, se U envia uma base ortonormal em outra, a expansão de um vetor nessa base e a identidade de Parseval mostram que U preserva sua norma. ■

Pela Definição 13.13, as matrizes de um operador unitário ou ortogonal em bases ortonormais são, respectivamente, unitárias ou ortogonais. As colunas dessas matrizes formam uma base ortonormal do espaço coordenado.

Exemplo 13.24 — Rotação e Reflexão. Para todo $\theta \in \mathbb{R}$, a matriz

$$R_\theta = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

é ortogonal e tem determinante 1. Já

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

é ortogonal e tem determinante -1; geometricamente, representa a reflexão em relação ao eixo horizontal.

◁

Corolário 13.25 (Caracterização Espectral das Isometrias Complexas).
Um operador complexo U é unitário se, e somente se, possui uma base ortonormal de autovetores e todos os seus autovalores têm módulo 1.

Demonstração. Um operador unitário é normal, pois $U^*U = UU^* = I$. Pelo teorema espectral complexo, ele possui uma base ortonormal de autovetores. Se $U\mathbf{v} = \lambda\mathbf{v}$, então

$$\|\mathbf{v}\| = \|U\mathbf{v}\| = |\lambda|\|\mathbf{v}\|,$$

de modo que $|\lambda| = 1$. A recíproca segue diretamente: em uma base ortonormal de autovetores, a matriz de U é diagonal e suas entradas têm módulo 1, logo $U^*U = I$. ■

Corolário 13.26 (Forma Canônica das Isometrias Reais).
Todo operador ortogonal real admite uma base ortonormal na qual sua matriz é

diagonal por blocos dos tipos

$$[1], \qquad [-1], \qquad \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}, \quad 0 < \theta < \pi.$$

Demonstração. Um operador ortogonal é normal. Na forma do Teorema 13.20, cada bloco deve ser ortogonal. Nos blocos unidimensionais isso força $\lambda = \pm 1$. Em um bloco bidimensional, $a^2 + b^2 = 1$; portanto, $a = \cos\theta$ e $b = \operatorname{sen}\theta$ para algum $0 < \theta < \pi$. ■

Em particular, o determinante de uma matriz ortogonal é 1 ou -1. Os blocos de rotação têm determinante 1; assim, o sinal do determinante registra a paridade do número de blocos $[-1]$.

13.5 Operadores Positivos e Raízes Quadradas

Positividade e forma quadrática
Para um operador autoadjunto, a positividade é expressa pela desigualdade $\langle Tv, v\rangle \geq 0$ para todo v. Trata-se de uma propriedade da forma quadrática associada, e não de uma condição sobre o sinal de cada entrada da matriz em uma base arbitrária.

Para números reais, a condição $a \geq 0$ pode ser reconhecida pela existência de uma raiz quadrada real. O análogo para operadores exige a autoadjunção e será obtido diretamente do teorema espectral.

Definição 13.27 — Operador Positivo. Um operador $A \in \operatorname{Hom}(V, V)$ é **positivo semidefinido**, ou **semipositivo**, se $A = A^*$ e

$$\langle A\mathbf{v}, \mathbf{v}\rangle \geq 0 \qquad \text{para todo } \mathbf{v} \in V.$$

Ele é **positivo definido** se a desigualdade é estrita para todo $\mathbf{v} \neq \vec{0}$. Escreveremos $A \geq 0$ e $A > 0$, respectivamente.

Exemplo 13.28. Se P é a projeção ortogonal sobre um subespaço S, então

$$\langle P\mathbf{v}, \mathbf{v}\rangle = \langle P\mathbf{v}, P\mathbf{v}\rangle = \|P\mathbf{v}\|^2 \geq 0.$$

Logo $P \geq 0$. Também vimos na Proposição 13.5 que $T^*T \geq 0$ para todo operador T.

◁

Teorema 13.29 (Critérios de Positividade).
Para um operador $A \in \operatorname{Hom}(V, V)$, são equivalentes:

a. $A \geq 0$;

b. *A é autoadjunto e todos os seus autovalores são não negativos;*

c. *existe um operador S tal que $A = S^*S$.*

Além disso, $A > 0$ se, e somente se, A é autoadjunto e todos os seus autovalores são positivos.

Demonstração. Se $A \geq 0$ e $A\mathbf{v} = \lambda\mathbf{v}$, com $\mathbf{v} \neq \vec{0}$, então

$$0 \leq \langle A\mathbf{v}, \mathbf{v}\rangle = \lambda\|\mathbf{v}\|^2.$$

Como A é autoadjunto, λ é real, e a desigualdade implica $\lambda \geq 0$.

Reciprocamente, suponha que A seja autoadjunto. Sejam $\lambda_1, \dots, \lambda_n$ seus autovalores, repetidos de acordo com suas multiplicidades, e suponha que todos sejam não negativos. Escolha uma base ortonormal de autovetores e defina S nessa base por

$$S\mathbf{e}_j = \sqrt{\lambda_j}\,\mathbf{e}_j.$$

Então $S = S^*$ e $A = S^2 = S^*S$.

Por fim, se $A = S^*S$, então A é autoadjunto e

$$\langle A\mathbf{v}, \mathbf{v}\rangle = \langle S\mathbf{v}, S\mathbf{v}\rangle = \|S\mathbf{v}\|^2 \geq 0.$$

Isso fecha o ciclo de equivalências. A afirmação sobre positividade definida segue pelo mesmo cálculo: em uma base ortonormal de autovetores, $\langle A\mathbf{v}, \mathbf{v}\rangle$ é uma soma dos termos $\lambda_j|x_j|^2$. ■

> **Teorema 13.30** (Raiz Quadrada Positiva).
> *Todo operador $A \geq 0$ possui uma única raiz quadrada positiva semidefinida. Ela é denotada por $\sqrt{A}$.*

Demonstração. Se

$$A = \sum_{j=1}^{r} \lambda_j P_j, \qquad \lambda_j \geq 0,$$

é a decomposição espectral de A, defina

$$\sqrt{A} = \sum_{j=1}^{r} \sqrt{\lambda_j}\, P_j.$$

Esse operador é autoadjunto, positivo semidefinido e seu quadrado é A.

Para a unicidade, suponha que $R \geq 0$ e $R^2 = A$. Como R comuta com $A = R^2$, cada autoespaço $E_\lambda(A)$ é invariante por R. A restrição de R a esse autoespaço é positiva semidefinida e satisfaz $R^2 = \lambda I$. Seus autovalores são, portanto, não negativos e têm quadrado igual a λ; todos são iguais a $\sqrt{\lambda}$. Pelo teorema espectral, a restrição é $\sqrt{\lambda}I$. Logo R coincide com o operador construído em cada autoespaço de A. ■

Exemplo 13.31 — Cálculo de uma Raiz Quadrada. A matriz

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

tem autovalores 3 e 1, com autovetores unitários $2^{-1/2}(1,1)$ e $2^{-1/2}(1,-1)$. Aplicando a função raiz quadrada aos autovalores, obtemos

$$\sqrt{A} = \frac{1}{2}\begin{bmatrix} \sqrt{3}+1 & \sqrt{3}-1 \\ \sqrt{3}-1 & \sqrt{3}+1 \end{bmatrix}.$$

A multiplicação direta confirma que $(\sqrt{A})^2 = A$.

◁

13.6 Decomposição Polar

Parte positiva e isometria
A decomposição polar escreve um operador como produto de uma parte isométrica e uma parte positiva. É o análogo matricial de decompor um número complexo em fase e módulo.

Todo número complexo não nulo pode ser escrito como $z = u|z|$, em que $|u| = 1$. A decomposição polar substitui o módulo por um operador positivo e a fase por uma isometria.

Definição 13.32 — Módulo de um Operador. Para $T \in \mathrm{Hom}(V, V)$, definimos

$$|T| = \sqrt{T^*T}.$$

Lema 13.33.
Para todo $\mathbf{v} \in V$,

$$\|T\mathbf{v}\| = \||T|\mathbf{v}\|.$$

Consequentemente, $\ker T = \ker |T|$.

Demonstração. Como $|T|$ é autoadjunto e $|T|^2 = T^*T$,

$$\|T\mathbf{v}\|^2 = \langle T^*T\mathbf{v}, \mathbf{v}\rangle = \langle |T|^2\mathbf{v}, \mathbf{v}\rangle = \langle |T|\mathbf{v}, |T|\mathbf{v}\rangle = \||T|\mathbf{v}\|^2.$$

A igualdade dos núcleos segue porque um vetor tem norma zero se, e somente se, é o vetor nulo. ■

Teorema 13.34 (Decomposição Polar).
Para todo $T \in \mathrm{Hom}(V, V)$, *existe uma isometria* U *tal que*

$$T = U|T|.$$

Se T *é invertível, então* U *é única e* $U = T|T|^{-1}$. *Se* T *não é invertível, a extensão de* U *aos complementos ortogonais, e portanto a decomposição, pode não ser única.*

Demonstração. Defina inicialmente

$$U_0 : \mathrm{im}\,|T| \longrightarrow \mathrm{im}\,T, \qquad U_0(|T|\mathbf{v}) = T\mathbf{v}.$$

O Lema 13.33 mostra simultaneamente que essa aplicação é bem definida e preserva normas. Ela é sobrejetiva por construção. Escolha uma base ortonormal de $\mathrm{im}\,|T|$ e transporte-a por U_0 para uma base ortonormal de $\mathrm{im}\,T$. Como esses subespaços têm a mesma dimensão, podemos completar as duas bases nos respectivos complementos ortogonais e definir uma isometria $U : V \to V$ que estende U_0. Para todo $\mathbf{v} \in V$,

$$U|T|\mathbf{v} = U_0|T|\mathbf{v} = T\mathbf{v}.$$

Se T é invertível, então $|T|$ também é invertível e a igualdade $T = U|T|$ força $U = T|T|^{-1}$. No caso singular, a construção deixa livre a escolha da isometria entre os complementos ortogonais. ■

Exemplo 13.35. Considere a matriz real

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

Temos

$$A^{\mathrm{t}}A = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}, \qquad |A| = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

Portanto,

$$A = \underbrace{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}}_{|A|}.$$

O primeiro fator é uma rotação de 90°; o segundo alonga duas direções ortogonais por fatores 1 e 2.

◁

13.7 Decomposição em Valores Singulares

Duas bases ortonormais
A SVD não exige que a matriz seja quadrada nem diagonalizável: usa uma base ortonormal no domínio e outra no contradomínio. Os valores singulares são obtidos a partir dos autovalores de A^*A e descrevem a variação dos comprimentos nas direções dessas bases.

A diagonalização espectral se aplica a operadores normais e atua no mesmo espaço na partida e na chegada. A decomposição em valores singulares não exige normalidade e vale também para transformações $T : V \to W$ entre espaços de dimensões diferentes. Em vez de usar uma única base de autovetores, ela escolhe uma base ortonormal no domínio e outra no contradomínio.

Teorema 13.36 (Decomposição em Valores Singulares).
Seja $T : V \to W$ uma transformação linear, com $\dim V = n$, $\dim W = m$ e $r = \operatorname{posto} T$. Existem bases ortonormais

$$\underline{v} = (\mathbf{v}_1, \dots, \mathbf{v}_n) \quad e \quad \underline{u} = (\mathbf{u}_1, \dots, \mathbf{u}_m)$$

e números unicamente determinados

$$s_1 \geq s_2 \geq \cdots \geq s_r > 0$$

tais que

$$T\mathbf{v}_j = s_j \mathbf{u}_j \quad (1 \leq j \leq r), \qquad T\mathbf{v}_j = \vec{0} \quad (r < j \leq n).$$

*Os números $s_1, \dots, s_r$ são os **valores singulares** de T.*

Demonstração. O operador T^*T em V é positivo semidefinido. Pelo teorema espectral, existe uma base ortonormal $\mathbf{v}_1, \dots, \mathbf{v}_n$ de autovetores, ordenada de modo que

$$T^*T\mathbf{v}_j = s_j^2 \mathbf{v}_j \quad (1 \leq j \leq r), \qquad T^*T\mathbf{v}_j = \vec{0} \quad (j > r),$$

onde $s_1 \geq \cdots \geq s_r > 0$. O número de autovalores positivos é r, pois $\ker(T^*T) = \ker T$.

Para $1 \leq j \leq r$, defina

$$\mathbf{u}_j = \frac{1}{s_j} T\mathbf{v}_j.$$

Esses vetores são ortonormais. De fato,

$$\begin{aligned}\langle \mathbf{u}_i, \mathbf{u}_j\rangle &= \frac{1}{s_i s_j}\langle T\mathbf{v}_i, T\mathbf{v}_j\rangle \\ &= \frac{1}{s_i s_j}\langle T^*T\mathbf{v}_i, \mathbf{v}_j\rangle = \frac{s_i}{s_j}\langle \mathbf{v}_i, \mathbf{v}_j\rangle.\end{aligned}$$

A última expressão vale 1 se $i = j$ e 0 caso contrário. Complete $\mathbf{u}_1, \dots, \mathbf{u}_r$ a uma base ortonormal de W.

Se $j > r$, então

$$\|T\mathbf{v}_j\|^2 = \langle T^*T\mathbf{v}_j, \mathbf{v}_j\rangle = 0,$$

logo $T\mathbf{v}_j = \vec{0}$. Isso prova a existência da decomposição. Finalmente, os quadrados dos valores singulares positivos são exatamente os autovalores positivos de T^*T, contados com multiplicidade; por isso, a lista ordenada $s_1, \dots, s_r$ é única. ■

Corolário 13.37 (Forma Matricial da SVD).
Para toda matriz $A \in \mathscr{M}_{m,n}(\mathbb{K})$, existem matrizes unitárias — ortogonais, no caso real — $Q \in \mathscr{M}_{m,m}(\mathbb{K})$ e $P \in \mathscr{M}_{n,n}(\mathbb{K})$ tais que

$$A = Q\Sigma P^*,$$

onde Σ é a matriz $m \times n$ cuja diagonal começa com os valores singulares $s_1, \dots, s_r$ e cujas demais entradas são nulas.

Demonstração. Tome para colunas de P os vetores da base ortonormal do domínio e para colunas de Q os vetores da base ortonormal do contradomínio fornecidas pelo Teorema 13.36. Nessa escolha de bases, $Q^*AP = \Sigma$. Multiplicando à esquerda por Q e à direita por P^*, obtemos a fatoração anunciada. ■

Exemplo 13.38. Para

$$A = \begin{bmatrix} 3 & 0 \\ 4 & 0 \end{bmatrix}, \qquad A^{\mathrm{t}}A = \begin{bmatrix} 25 & 0 \\ 0 & 0 \end{bmatrix}.$$

Os valores singulares são 5 e 0. Podemos escolher

$$P = I, \qquad Q = \begin{bmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} 5 & 0 \\ 0 & 0 \end{bmatrix}.$$

Então $A = Q\Sigma P^{\mathrm{t}}$. A primeira direção do domínio é transformada em uma direção do contradomínio com fator de alongamento 5; a segunda pertence ao núcleo.

◁

Exemplo 13.39 — Uma SVD de posto completo. Considere agora

$$A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}.$$

Aqui as direções singulares não são os vetores da base canônica. Primeiro,

$$A^{\mathrm{t}}A = \begin{bmatrix} 25 & 20 \\ 20 & 25 \end{bmatrix}.$$

Os autovalores são 45 e 5, com autovetores unitários

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}, \qquad \mathbf{v}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}.$$

Logo os valores singulares são $s_1 = 3\sqrt{5}$ e $s_2 = \sqrt{5}$. As direções correspondentes no contradomínio são

$$\mathbf{u}_1 = \frac{A\mathbf{v}_1}{s_1} = \frac{1}{\sqrt{10}}\begin{bmatrix}1\\3\end{bmatrix}, \qquad \mathbf{u}_2 = \frac{A\mathbf{v}_2}{s_2} = \frac{1}{\sqrt{10}}\begin{bmatrix}3\\-1\end{bmatrix}.$$

Portanto,

$$A = \underbrace{\frac{1}{\sqrt{10}}\begin{bmatrix}1 & 3\\3 & -1\end{bmatrix}}_{Q}\underbrace{\begin{bmatrix}3\sqrt{5} & 0\\0 & \sqrt{5}\end{bmatrix}}_{\Sigma}\underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}}_{P^{\mathrm{t}}}.$$

Este exemplo separa as três etapas geométricas da SVD. Primeiro mudam-se ortonormalmente as direções do domínio; depois se aplicam dois alongamentos distintos; por fim, mudam-se ortonormalmente as direções do contradomínio.

◁

A construção da decomposição em valores singulares parte do operador positivo T^*T. O teorema espectral fornece uma base ortonormal de autovetores desse operador, e os valores singulares descrevem os fatores de alongamento de T nas direções correspondentes. Assim, a representação de T se reduz a mudanças ortonormais de coordenadas e a uma matriz diagonal, possivelmente retangular.

Exercícios

Ex. 13.1 — Classifique cada matriz abaixo como autoadjunta, normal, unitária ou ortogonal e positiva semidefinida, quando essas expressões fizerem sentido. Uma matriz pode pertencer a mais de uma classe.

$$\begin{bmatrix}2 & 1\\1 & 2\end{bmatrix}, \qquad \begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}, \qquad \begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}, \qquad \begin{bmatrix}1 & 0\\0 & i\end{bmatrix}, \qquad \begin{bmatrix}0 & i\\-i & 0\end{bmatrix}.$$

Nos três primeiros casos, trabalhe sobre $\mathbb{R}$; nos dois últimos, sobre $\mathbb{C}$.

Ex. 13.2 — Seja $T \in \operatorname{Hom}(V, V)$. Prove que T é normal se, e somente se,

$$\|T\mathbf{v}\| = \|T^*\mathbf{v}\| \qquad \text{para todo } \mathbf{v} \in V.$$

Sugestão. Aplique a identidade de polarização ao operador $T^*T - TT^*$.

Ex. 13.3 — Seja T normal.

1. Prove que, se T é nilpotente, então $T = 0$.

2. Prove que, se $T^2 = T$, então T é a projeção ortogonal sobre $\operatorname{im} T$.

Ex. 13.4 — Considere

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

Encontre uma base ortonormal de autovetores, escreva a decomposição espectral de A e obtenha uma fórmula fechada para A^m, com $m \geq 1$.

Ex. 13.5 — Seja $U \in \mathrm{Hom}(V, V)$. Prove que U é simultaneamente autoadjunto e unitário — ortogonal, no caso real — se, e somente se, existe uma projeção ortogonal P tal que

$$U = 2P - I.$$

Interprete geometricamente a ação de U sobre $\operatorname{im} P$ e $\ker P$.

Fundamentos

Ex. 13.6 — Seja T normal complexo, ou autoadjunto real, com autovalores distintos $\lambda_1, \ldots, \lambda_r$. Prove que a projeção espectral sobre E_{λ_j} é dada por

$$P_j = \prod_{k \neq j} \frac{T - \lambda_k I}{\lambda_j - \lambda_k}.$$

Conclua que toda função definida no espectro de T pode ser representada por um polinômio em T.

Ex. 13.7 — Sejam $T_1, \ldots, T_s$ operadores autoadjuntos que comutam dois a dois. Prove, por indução em s, que existe uma base ortonormal formada por autovetores simultâneos de todos eles.

Ex. 13.8 — Se $A \geq 0$, prove que $\sqrt{A}$ comuta com todo operador que comuta com A. Mostre também que

$$\ker \sqrt{A} = \ker A \qquad \text{e} \qquad \operatorname{im} \sqrt{A} = \operatorname{im} A.$$

Ex. 13.9 — Use a forma canônica real do Teorema 13.20 para provar:

1. o determinante de um operador normal real é o produto dos escalares λ dos blocos unidimensionais pelos números $a^2 + b^2$ dos blocos bidimensionais; conclua que o operador é invertível se, e somente se, nenhum bloco unidimensional é $[0]$;

2. ele é ortogonal se, e somente se, os blocos unidimensionais são $[1]$ ou $[-1]$ e, em cada bloco bidimensional, $a^2 + b^2 = 1$.

Ex. 13.10 — Na construção da decomposição polar, prove que

$$\operatorname{im} |T| = (\ker T)^{\perp}.$$

Descreva todas as escolhas possíveis para o fator unitário ou ortogonal U quando T é singular e conclua que esse fator é único se, e somente se, T é invertível.

Consolidação

Ex. 13.11 — Calcule uma decomposição em valores singulares da matriz

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

Determine explicitamente matrizes ortogonais Q e P e uma matriz Σ tais que $A = Q\Sigma P^{\mathrm{t}}$.

Ex. 13.12 — Prove que T e T^{*} têm os mesmos valores singulares positivos, com as mesmas multiplicidades. Relacione as bases singulares de um operador com as de sua adjunta.

Ex. 13.13 — Se

$$T\mathbf{v}_j = s_j \mathbf{u}_j \quad (1 \le j \le r)$$

é uma SVD de T, defina o operador $T^{+} : W \to V$ por

$$T^{+}\mathbf{u}_j = s_j^{-1}\mathbf{v}_j \quad (1 \le j \le r)$$

e por zero no complemento ortogonal de $\operatorname{im} T$. Prove que TT^{+} é a projeção ortogonal sobre $\operatorname{im} T$, que $T^{+}T$ é a projeção ortogonal sobre $(\ker T)^{\perp}$ e que

$$TT^{+}T = T, \qquad T^{+}TT^{+} = T^{+}.$$

O operador T^{+} é chamado de *pseudoinversa de Moore–Penrose*.

Ex. 13.14 — Seja $A = Q\Sigma P^{*}$ uma SVD e sejam $s_1 \ge \cdots \ge s_r > 0$ seus valores singulares. Para $k < r$, substitua em Σ os valores $s_{k+1}, \dots, s_r$ por zero e denote por A_k a matriz resultante. Prove que A_k tem posto k e minimiza $\|A - B\|_{\mathrm{F}}$ entre todas as matrizes B de posto no máximo k, onde

$$\|C\|_{\mathrm{F}}^2 = \sum_{i,j} |c_{ij}|^2.$$

Além disso, mostre que o valor mínimo é $\sqrt{s_{k+1}^2 + \cdots + s_r^2}$. *Sugestão.* Use a invariância da norma de Frobenius por multiplicação unitária. Se R é a projeção ortogonal sobre $\operatorname{im} B$, compare B com RA e use $\sum_j \|R\mathbf{u}_j\|^2 \le k$ para uma base de vetores singulares à esquerda.

Ex. 13.15 — Sejam S e T operadores normais em um espaço complexo e suponha que $ST = TS$. Prove que existe uma base ortonormal de autovetores simultâneos de S e T. Estenda o argumento a uma família finita de operadores normais que comutam dois a dois.

14
Formas Bilineares

Uma transformação linear descreve uma dependência linear em uma única variável. Muitas construções naturais, porém, recebem dois vetores: o produto de polinômios, o produto de matrizes, o determinante em dimensão dois e as várias noções de ortogonalidade são exemplos desse fenômeno. A condição adequada não é pedir linearidade no par de variáveis, mas linearidade em cada uma delas quando a outra permanece fixa.

Em dimensão finita, uma forma bilinear é determinada por uma matriz. A relação com o espaço dual permite descrever sua degenerescência. Estudaremos, em seguida, as formas simétricas, alternadas e hermitianas, e sua classificação por decomposições em blocos ortogonais elementares.

Salvo indicação em contrário, todos os espaços vetoriais serão considerados sobre um mesmo corpo $\mathbb{K}$. As formas hermitianas aparecerão apenas sobre $\mathbb{C}$, sempre com a convenção adotada no Capítulo 6: linearidade na primeira variável e linearidade conjugada na segunda.

Nesse capítulo

14.1 Aplicações Bilineares

Definição 14.1 — Aplicação Bilinear. Sejam V, W e Z espaços vetoriais sobre $\mathbb{K}$. Uma aplicação

$$F : V \times W \longrightarrow Z$$

é denominada **aplicação bilinear**, ou **mapa bilinear**, se for linear em cada variável separadamente. Explicitamente, para $\mathbf{v}_1, \mathbf{v}_2 \in V$, $\mathbf{w}_1, \mathbf{w}_2 \in W$ e $\lambda_1, \lambda_2 \in \mathbb{K}$,

$$F(\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2, \mathbf{w}) = \lambda_1 F(\mathbf{v}_1, \mathbf{w}) + \lambda_2 F(\mathbf{v}_2, \mathbf{w}),$$
$$F(\mathbf{v}, \lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2) = \lambda_1 F(\mathbf{v}, \mathbf{w}_1) + \lambda_2 F(\mathbf{v}, \mathbf{w}_2).$$

Linear em cada entrada
Uma aplicação bilinear é linear em cada variável quando a outra permanece fixa. Em geral, ela não é uma transformação linear definida no produto cartesiano $V \times W$: ao variar as duas entradas simultaneamente, surgem termos cruzados.

Fixar a segunda variável transforma F em uma aplicação linear de V em Z; fixar a primeira transforma F em uma aplicação linear de W em Z. Em particular,

$$F(\mathbf{v}, \vec{0}) = F(\vec{0}, \mathbf{w}) = \vec{0}$$

para todos $\mathbf{v} \in V$ e $\mathbf{w} \in W$.

Observação 14.2. Uma aplicação bilinear não é, em geral, uma transformação linear no espaço produto $V \times W$. De fato, ao expandirmos $F(\mathbf{v}_1 + \mathbf{v}_2, \mathbf{w}_1 + \mathbf{w}_2)$ aparecem também os termos cruzados $F(\mathbf{v}_1, \mathbf{w}_2)$ e $F(\mathbf{v}_2, \mathbf{w}_1)$. É justamente a presença desses termos que torna a bilinearidade diferente da linearidade usual.

Definição 14.3 — Forma Bilinear. Uma aplicação bilinear $g : V \times W \to \mathbb{K}$ é denominada **forma bilinear** em $V \times W$. Quando $V = W$, dizemos simplesmente que g é uma forma bilinear em V.
Denotaremos por $\mathrm{Bil}(V, W; Z)$ o conjunto das aplicações bilineares de $V \times W$ em Z e escreveremos $\mathrm{Bil}(V, W)$ quando $Z = \mathbb{K}$.

Exemplo 14.4 — Formas definidas por matrizes. Seja $A \in \mathscr{M}_{m,n}(\mathbb{K})$. A aplicação

$$g : \mathbb{K}^m \times \mathbb{K}^n \longrightarrow \mathbb{K}, \qquad g(\mathbf{x}, \mathbf{y}) = \mathbf{x}^{\mathrm{t}} A \mathbf{y},$$

é bilinear. Por exemplo, se

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

então

$$g(\mathbf{x}, \mathbf{y}) = x_1 y_1 + (2x_1 + 3x_2) y_2 + (-x_1 + 4x_2) y_3.$$

A expressão deixa visível que, mantendo uma das variáveis fixa, obtemos uma combinação linear das coordenadas da outra.

◁

Exemplo 14.5 — Multiplicação em uma álgebra. Se A é uma álgebra sobre $\mathbb{K}$, sua multiplicação

$$A \times A \longrightarrow A, \qquad (a, b) \longmapsto ab,$$

é bilinear. Isso vale, por exemplo, para a multiplicação de polinômios em $\mathbb{K}[x]$ e para a multiplicação de matrizes quadradas. Observe que o valor agora pertence à própria álgebra; portanto, trata-se de uma aplicação bilinear, mas não de uma forma bilinear.

◁

Exemplo 14.6 — Aplicações com valores vetoriais. Seja $\underline{z} = (\mathbf{z}_1, \ldots, \mathbf{z}_q)$ uma base de Z e sejam $g_1, \ldots, g_q \in \mathrm{Bil}(V, W)$. Então

$$F(\mathbf{v}, \mathbf{w}) = \sum_{k=1}^{q} g_k(\mathbf{v}, \mathbf{w}) \mathbf{z}_k$$

define uma aplicação bilinear de $V \times W$ em Z. Assim, uma aplicação bilinear com valores em um espaço de dimensão finita pode ser estudada coordenada por coordenada.

◁

Teorema 14.7.
Com as operações definidas ponto a ponto, $\mathrm{Bil}(V, W; Z)$ *é um espaço vetorial sobre* $\mathbb{K}$.

Demonstração. A aplicação nula é bilinear. Se $F, G \in \mathrm{Bil}(V, W; Z)$ e $c \in \mathbb{K}$, devemos verificar que $F+G$ e cF continuam lineares em cada variável. Para a primeira variável, por exemplo,

$$\begin{aligned}(F+G)(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2, \mathbf{w}) &= F(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2, \mathbf{w}) + G(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2, \mathbf{w}) \\ &= \lambda_1(F+G)(\mathbf{v}_1, \mathbf{w}) + \lambda_2(F+G)(\mathbf{v}_2, \mathbf{w}),\end{aligned}$$

e

$$(cF)(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2, \mathbf{w}) = \lambda_1(cF)(\mathbf{v}_1, \mathbf{w}) + \lambda_2(cF)(\mathbf{v}_2, \mathbf{w}).$$

Os mesmos cálculos na segunda variável completam a verificação. As demais propriedades de espaço vetorial são herdadas do espaço de todas as funções de $V \times W$ em Z. ■

Proposição 14.8.
Se $F : V \times W \to Z$ é bilinear e $T : Z \to U$ é linear, então $T \circ F : V \times W \to U$ é bilinear.

Demonstração. Para a primeira variável,

$$\begin{aligned}(T \circ F)(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2, \mathbf{w}) &= T\big(\lambda_1 F(\mathbf{v}_1, \mathbf{w}) + \lambda_2 F(\mathbf{v}_2, \mathbf{w})\big) \\ &= \lambda_1(T \circ F)(\mathbf{v}_1, \mathbf{w}) + \lambda_2(T \circ F)(\mathbf{v}_2, \mathbf{w}).\end{aligned}$$

A demonstração para a segunda variável é idêntica. ■

Teorema 14.9 (Critério por Coordenadas).
Seja $\underline{z} = (\mathbf{z}_1, \ldots, \mathbf{z}_q)$ uma base de Z e escreva

$$F(\mathbf{v}, \mathbf{w}) = \sum_{k=1}^{q} F_k(\mathbf{v}, \mathbf{w})\mathbf{z}_k.$$

Então $F : V \times W \to Z$ é bilinear se, e somente se, cada função $F_k : V \times W \to \mathbb{K}$ é uma forma bilinear.

Demonstração. Se $\pi_k : Z \to \mathbb{K}$ é o funcional que fornece a k-ésima coordenada na base $\underline{z}$, então $F_k = \pi_k \circ F$. Logo, se F é bilinear, a Proposição 14.8 mostra que cada F_k é bilinear.

Reciprocamente, se todas as funções F_k são bilineares, a linearidade em cada variável de F segue da linearidade de cada coordenada e da unicidade da representação na base $\underline{z}$. ■

14.2 Matrizes de Formas Bilineares

Como ocorre com as transformações lineares, uma forma bilinear em espaços de dimensão finita fica determinada por seus valores nos vetores de duas bases. A diferença é que agora há um índice para cada variável.

Congruência, não similaridade
Ao mudar a base de uma forma bilinear, sua matriz se transforma por congruência, $A \mapsto P^{\mathsf{T}}AP$. Para operadores lineares, a mudança ocorre por similaridade. A diferença reflete o fato de que a forma possui duas entradas vetoriais.

Definição 14.10 — Matriz de uma Forma Bilinear. Sejam $\underline{\mathrm{v}} = (\mathbf{v}_1, \dots, \mathbf{v}_m)$ uma base de V, $\underline{\mathrm{w}} = (\mathbf{w}_1, \dots, \mathbf{w}_n)$ uma base de W e $g \in \mathrm{Bil}(V, W)$. A **matriz de g em relação às bases** $\underline{\mathrm{v}}$ e $\underline{\mathrm{w}}$ é

$$[g]_{\underline{\mathrm{v}},\underline{\mathrm{w}}} = \big(g(\mathbf{v}_i, \mathbf{w}_j)\big)_{\substack{1\le i\le m\\ 1\le j\le n}}.$$

Essa matriz também é chamada de **matriz de Gram** de g nessas bases. Quando $V = W$ e a mesma base $\underline{\mathrm{v}}$ é usada nas duas variáveis, escrevemos simplesmente $[g]_{\underline{\mathrm{v}}}$.

Teorema 14.11 (Representação Matricial).
Se $G = [g]_{\underline{V},\underline{W}}$, então, para todos $\mathbf{v} \in V$ e $\mathbf{w} \in W$,

$$g(\mathbf{v}, \mathbf{w}) = [\mathbf{v}]_{\underline{\mathrm{v}}}^{\mathrm{t}} G [\mathbf{w}]_{\underline{W}}.$$

Reciprocamente, cada matriz $G \in \mathscr{M}_{m,n}(\mathbb{K})$ define uma única forma bilinear em $V \times W$ por essa fórmula.

Demonstração. Escreva

$$\mathbf{v} = \sum_{i=1}^{m} x_i \mathbf{v}_i \qquad \text{e} \qquad \mathbf{w} = \sum_{j=1}^{n} y_j \mathbf{w}_j.$$

A bilinearidade permite expandir uma variável de cada vez:

$$g(\mathbf{v}, \mathbf{w}) = \sum_{i=1}^{m} \sum_{j=1}^{n} x_i y_j \, g(\mathbf{v}_i, \mathbf{w}_j).$$

Essa soma dupla é precisamente o produto matricial $[\mathbf{v}]_{\underline{\mathrm{v}}}^{\mathrm{t}} G [\mathbf{w}]_{\underline{\mathrm{w}}}$.

Reciprocamente, dado G, o Exemplo 14.4 mostra que a fórmula define uma forma bilinear. Seus valores nos pares $(\mathbf{v}_i, \mathbf{w}_j)$ são as entradas de G, de modo que a forma é única. ∎

Exemplo 14.12. Considere $V = \mathbb{K}_2[x]$, com a base $\underline{\mathrm{p}} = (1, x, x^2)$, e a forma

$$g(p, q) = p(0)q(0) + p'(0)q'(0).$$

Se $p = a + bx + cx^2$ e $q = r + sx + tx^2$, então

$$g(p, q) = ar + bs = \begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} r \\ s \\ t \end{bmatrix}.$$

A matriz torna visível que o termo de grau dois não é detectado por g. Retomaremos essa observação ao estudar formas degeneradas.

◁

Corolário 14.13.
Se V, W e Z têm dimensões finitas, então

$$\dim \operatorname{Bil}(V, W; Z) = (\dim V)(\dim W)(\dim Z).$$

Em particular,

$$\dim \operatorname{Bil}(V, W) = (\dim V)(\dim W) \quad e \quad \dim \operatorname{Bil}(V, V) = (\dim V)^2.$$

Demonstração. Escolhidas bases de V, W e Z, uma aplicação bilinear $F : V \times W \to Z$ fica determinada pelos coeficientes das imagens $F(\mathbf{v}_i, \mathbf{w}_j)$ na base de Z. Há $(\dim V)(\dim W)(\dim Z)$ coeficientes, que podem ser escolhidos arbitrariamente. O critério por coordenadas garante que cada escolha produz uma aplicação bilinear. ■

Teorema 14.14 (Mudança de Bases).
Sejam $\underline{v}$ e $\underline{v}'$ bases de V, e $\underline{w}$ e $\underline{w}'$ bases de W. Defina

$$P = \mathrm{M}_{\underline{V}' \to \underline{V}}, \qquad Q = \mathrm{M}_{\underline{W}' \to \underline{W}}.$$

Se $G = [g]_{\underline{V},\underline{W}}$ e $G' = [g]_{\underline{v}',\underline{w}'}$, então

$$G' = P^{\mathrm{t}} G Q.$$

Demonstração. Pela definição das matrizes de mudança de base,

$$[\mathbf{v}]_{\underline{\mathrm{V}}} = P[\mathbf{v}]_{\underline{\mathrm{V}}'} \qquad \text{e} \qquad [\mathbf{w}]_{\underline{\mathrm{W}}} = Q[\mathbf{w}]_{\underline{\mathrm{W}}'}.$$

Substituindo essas relações na representação matricial de g, obtemos

$$\begin{aligned} g(\mathbf{v}, \mathbf{w}) &= [\mathbf{v}]_{\underline{\mathrm{V}}}^{\mathrm{t}} G [\mathbf{w}]_{\underline{\mathrm{W}}} \\ &= [\mathbf{v}]_{\underline{\mathrm{V}}'}^{\mathrm{t}} P^{\mathrm{t}} G Q [\mathbf{w}]_{\underline{\mathrm{W}}'}. \end{aligned}$$

Pela unicidade da matriz que representa a forma nas novas bases, $G' = P^{\mathrm{t}} G Q$. ■

Exemplo 14.15. Considere o produto escalar usual em $\mathbb{R}^2$ e a base canônica $\underline{\mathrm{E}}$. Na base

$$\underline{\mathrm{V}} = (\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 - \mathbf{e}_2),$$

a matriz de mudança de base é

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

Como $G = I_2$ na base canônica,

$$[g]_{\underline{\mathrm{V}}} = P^{\mathrm{t}} P = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

A forma não mudou; mudou apenas sua matriz. Os vetores da nova base são ortogonais, mas têm norma $\sqrt{2}$.

◁

Observação 14.16. Quando $V = W$ e usamos a mesma mudança de base nas duas variáveis, a transformação $G \mapsto P^{t}GP$ é chamada de **congruência**. Diferentemente da semelhança $A \mapsto P^{-1}AP$, ela é a mudança de coordenadas apropriada para formas bilineares.

14.3 Dualidade e Degenerescência

Fixar uma entrada
Uma forma bilinear B associa a cada vetor v o funcional $w \mapsto B(v, w)$. Desse modo, a forma pode ser estudada como uma transformação de V em W^*; degenerescência corresponde à perda de injetividade dessa transformação.

Uma forma bilinear determina uma transformação linear com valores no espaço dual: fixada uma das entradas, a outra permanece como argumento de um funcional. Essa correspondência relaciona o estudo das formas bilineares ao das transformações lineares.

Teorema 14.17 (Dualidade das Formas Bilineares).
Para quaisquer espaços vetoriais V e W, existem isomorfismos naturais

$$\mathrm{Bil}(V, W) \cong \mathrm{Hom}(V, W^*) \qquad e \qquad \mathrm{Bil}(V, W) \cong \mathrm{Hom}(W, V^*).$$

Esses isomorfismos não dependem de escolhas de bases nem da hipótese de dimensão finita.

Demonstração. A uma forma $g \in \mathrm{Bil}(V, W)$ associamos a transformação

$$g_L : V \longrightarrow W^*, \qquad g_L(\mathbf{v})(\mathbf{w}) = g(\mathbf{v}, \mathbf{w}).$$

Para cada $\mathbf{v}$, a função $g_L(\mathbf{v})$ é um funcional em W porque g é linear na segunda variável. A linearidade de g_L segue da linearidade de g na primeira variável.

Reciprocamente, se $T \in \mathrm{Hom}(V, W^*)$, definimos

$$g_T(\mathbf{v}, \mathbf{w}) = T(\mathbf{v})(\mathbf{w}).$$

A linearidade de T fornece a linearidade na primeira variável, e o fato de $T(\mathbf{v})$ ser um funcional fornece a linearidade na segunda. Logo, g_T é bilinear. As construções $g \mapsto g_L$ e $T \mapsto g_T$ são lineares e inversas entre si.

O segundo isomorfismo é obtido do mesmo modo, associando a g a

$$g_R : W \longrightarrow V^*, \qquad g_R(\mathbf{w})(\mathbf{v}) = g(\mathbf{v}, \mathbf{w}).$$

■

Definição 14.18 — Radicais. Se $g \in \mathrm{Bil}(V, W)$, definimos o **radical esquerdo** e o **radical direito** de g por

$$\mathrm{Rad}_E(g) = \{\mathbf{v} \in V \mid g(\mathbf{v}, \mathbf{w}) = 0 \text{ para todo } \mathbf{w} \in W\},$$
$$\mathrm{Rad}_D(g) = \{\mathbf{w} \in W \mid g(\mathbf{v}, \mathbf{w}) = 0 \text{ para todo } \mathbf{v} \in V\}.$$

Radical e isotropia
Um vetor no radical tem emparelhamento nulo com todos os vetores na outra entrada da forma. Essa condição difere da isotropia: um vetor pode satisfazer $B(v, v) = 0$ sem pertencer ao radical.

Pelo Teorema 14.17,

$$\mathrm{Rad}_E(g) = \ker g_L \qquad \text{e} \qquad \mathrm{Rad}_D(g) = \ker g_R.$$

Em particular, ambos são subespaços. Essa identificação caracteriza os radicais pelos vetores que anulam a forma quando colocados, respectivamente, na primeira ou na segunda entrada.

Proposição 14.19 (Radicais e Matriz).
Suponha que V e W tenham dimensão finita e seja $G = [g]_{\underline{V},\underline{W}}$. Então

$$[\mathrm{Rad}_E(g)]_{\underline{V}} = \ker(G^{\mathrm{t}}) \qquad e \qquad [\mathrm{Rad}_D(g)]_{\underline{W}} = \ker G.$$

Consequentemente,

$$\dim \mathrm{Rad}_E(g) = \dim V - \mathrm{posto}\, G,$$
$$\dim \mathrm{Rad}_D(g) = \dim W - \mathrm{posto}\, G.$$

Demonstração. Um vetor $\mathbf{v}$ pertence ao radical esquerdo se, e somente se,

$$[\mathbf{v}]_{\underline{V}}^{\mathrm{t}} G [\mathbf{w}]_{\underline{W}} = 0$$

para toda coluna $[\mathbf{w}]_{\underline{W}}$. Isso equivale a $G^{\mathrm{t}}[\mathbf{v}]_{\underline{V}} = \vec{0}$. A afirmação sobre o radical direito é análoga. As fórmulas das dimensões seguem do teorema do núcleo e da imagem. ■

Exemplo 14.20. Considere a forma em $\mathbb{R}^2$ cuja matriz canônica é

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

Então $g(\mathbf{x}, \mathbf{y}) = x_1 y_2$ e

$$\mathrm{Rad}_E(g) = \langle \mathbf{e}_2 \rangle, \qquad \mathrm{Rad}_D(g) = \langle \mathbf{e}_1 \rangle.$$

Esse exemplo mostra que os dois radicais podem ser diferentes. No Exemplo 14.12, por outro lado, a matriz é simétrica e ambos os radicais são $\langle x^2 \rangle$.

◁

Definição 14.21 — Forma Não Degenerada. Uma forma $g \in \mathrm{Bil}(V, W)$ é **não degenerada** se

$$\mathrm{Rad}_E(g) = \{\vec{0}\} \qquad \text{e} \qquad \mathrm{Rad}_D(g) = \{\vec{0}\}.$$

Caso contrário, g é denominada **degenerada**.

Proposição 14.22.
Se V tem dimensão finita e g é uma forma bilinear em V, as seguintes afirmações são equivalentes:

a. *g é não degenerada;*
b. $\mathrm{Rad}_E(g) = \{\vec{0}\}$;
c. $\mathrm{Rad}_D(g) = \{\vec{0}\}$;
d. *$g_L : V \to V^*$ é um isomorfismo;*
e. *a matriz de g em alguma — e, portanto, em toda — base é invertível.*

Demonstração. Em uma base qualquer, os radicais correspondem aos núcleos de G^{t} e G. Como G é quadrada,

$$\ker G = \{\vec{0}\} \iff \det G \neq 0 \iff \ker G^{\mathsf{t}} = \{\vec{0}\}.$$

Além disso, $\ker g_L = \mathrm{Rad}_E(g)$ e $\dim V = \dim V^*$; assim, g_L é injetiva se, e somente se, é um isomorfismo. Isso prova todas as equivalências. ∎

Observação 14.23. Se $g \in \mathrm{Bil}(V, W)$ é não degenerada e V, W têm dimensão finita, então a Proposição 14.19 implica $\dim V = \mathrm{posto}\, G = \dim W$. Portanto, uma forma não degenerada entre dois espaços de dimensões diferentes não pode existir.

14.4 Simetria, Alternância e Formas Hermitianas

Até aqui as duas entradas de uma forma desempenharam papéis independentes. Quando ambas pertencem ao mesmo espaço, podemos compará-las e impor condições de simetria.

A característica dois é especial
Se $2 \neq 0$ no corpo, toda forma alternada é antissimétrica e as partes simétrica e antissimétrica podem ser obtidas dividindo por 2. Em característica 2, essa decomposição deixa de funcionar e as duas noções precisam ser tratadas com cuidado.

Definição 14.24 — Transposta de uma Forma. Se g é uma forma bilinear em V, sua **forma transposta** é

$$g^{\mathsf{t}}(\mathbf{u}, \mathbf{v}) = g(\mathbf{v}, \mathbf{u}).$$

Definição 14.25. Uma forma bilinear g em V é

a **simétrica** se $g(\mathbf{u}, \mathbf{v}) = g(\mathbf{v}, \mathbf{u})$;

b **antissimétrica** se $g(\mathbf{u}, \mathbf{v}) = -g(\mathbf{v}, \mathbf{u})$;

c **alternada** se $g(\mathbf{v}, \mathbf{v}) = 0$ para todo $\mathbf{v} \in V$.

Uma **forma simplética** é uma forma bilinear alternada e não degenerada.

Definição 14.26 — Forma Sesquilinear e Forma Hermitiana. Seja V um espaço vetorial complexo. Uma aplicação $h : V \times V \to \mathbb{C}$ é uma **forma sesquilinear** se for linear na primeira variável e linear conjugada na segunda:

$$h(\mathbf{u}, \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2) = \overline{\lambda_1} h(\mathbf{u}, \mathbf{v}_1) + \overline{\lambda_2} h(\mathbf{u}, \mathbf{v}_2).$$

A forma adjunta de h é

$$h^*(\mathbf{u}, \mathbf{v}) = \overline{h(\mathbf{v}, \mathbf{u})}.$$

Dizemos que h é **hermitiana** se $h^* = h$. Nesse caso, $h(\mathbf{v}, \mathbf{v})$ é real para todo $\mathbf{v} \in V$.

Observação 14.27. Formas hermitianas não são formas bilineares complexas: a segunda variável é conjugado-linear. Essa distinção é indispensável. Neste livro, a expressão *produto interno* continua reservada às formas simétricas positivas definidas sobre $\mathbb{R}$ e às formas hermitianas positivas definidas sobre $\mathbb{C}$, como no Capítulo 6. As definições de radical e de não degenerescência se estendem literalmente às formas sesquilineares; para uma forma hermitiana, os radicais esquerdo e direito coincidem.

Conjugação e simetria hermitiana
Sobre $\mathbb{C}$, uma forma hermitiana é linear na primeira entrada e conjugado-linear na segunda. A simetria conjugada garante que $B(v, v)$ seja real. Acrescentando a positividade definida, obtemos os produtos internos complexos.

Proposição 14.28.
Seja g uma forma bilinear em V.

a *Se g é alternada, então g é antissimétrica.*

b *Se $\mathrm{char}(\mathbb{K}) \neq 2$ e g é antissimétrica, então g é alternada.*

c *Se $\mathrm{char}(\mathbb{K}) = 2$, as condições de simetria e antissimetria coincidem, mas não implicam alternância.*

Demonstração. Se g é alternada, então

$$0 = g(\mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v}) = g(\mathbf{u},\mathbf{v}) + g(\mathbf{v},\mathbf{u}),$$

pois os termos diagonais são nulos. Logo, $g(\mathbf{u},\mathbf{v}) = -g(\mathbf{v},\mathbf{u})$.

Reciprocamente, se g é antissimétrica,

$$g(\mathbf{v},\mathbf{v}) = -g(\mathbf{v},\mathbf{v}),$$

e, portanto, $2g(\mathbf{v},\mathbf{v}) = 0$. Quando a característica é diferente de 2, segue que $g(\mathbf{v},\mathbf{v}) = 0$.

Em característica 2, temos $-1 = 1$, de modo que simetria e antissimetria expressam a mesma igualdade. A forma em $\mathbb{K}^2$ com matriz $\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ é alternada e não nula, mostrando por que a alternância deve ser mantida como uma condição separada. ■

Proposição 14.29 (Partes Simétrica e Antissimétrica).
Se $\mathrm{char}(\mathbb{K}) \neq 2$, toda forma bilinear g se decompõe de maneira única como

$$g = g_s + g_a, \qquad g_s = \frac{1}{2}(g + g^{\mathrm{t}}), \qquad g_a = \frac{1}{2}(g - g^{\mathrm{t}}),$$

onde g_s é simétrica e g_a é alternada.

Demonstração. As identidades $g_s^{\mathrm{t}} = g_s$ e $g_a^{\mathrm{t}} = -g_a$ seguem diretamente das definições; pela Proposição 14.28, g_a é alternada. Além disso, $g_s + g_a = g$.

Se $g = s + a$ com s simétrica e a antissimétrica, então $g^{\mathrm{t}} = s - a$. Somando e subtraindo as duas igualdades, obtemos necessariamente $s = g_s$ e $a = g_a$. ■

A função $q_g : V \to \mathbb{K}$ definida por $q_g(\mathbf{v}) = g(\mathbf{v},\mathbf{v})$ é a forma quadrática associada a g. Quando g é simétrica e $\mathrm{char}(\mathbb{K}) \neq 2$, a identidade de polarização

$$g(\mathbf{u},\mathbf{v}) = \frac{1}{2}\big(q_g(\mathbf{u}+\mathbf{v}) - q_g(\mathbf{u}) - q_g(\mathbf{v})\big)$$

mostra que os valores diagonais determinam toda a forma. Para uma forma bilinear arbitrária, esses valores enxergam apenas sua parte simétrica: a parte alternada desaparece em $g(\mathbf{v},\mathbf{v})$.

Teorema 14.30 (Caracterização Matricial).
Se $G = [g]_{\underline{v}}$, então:

a $[g^{\mathrm{t}}]_{\underline{v}} = G^{\mathrm{t}}$;

b g *é simétrica se, e somente se,* $G^{\mathrm{t}} = G$;

c g *é antissimétrica se, e somente se,* $G^{\mathrm{t}} = -G$;

d g *é alternada se, e somente se,* $G^{\mathrm{t}} = -G$ *e todas as entradas diagonais de* G *são nulas.*

Se h é sesquilinear e sua matriz H é definida pela fórmula

$$h(\mathbf{u}, \mathbf{v}) = [\mathbf{u}]_{\underline{v}}^{\mathrm{t}} H \, \overline{[\mathbf{v}]_{\underline{v}}},$$

então h é hermitiana se, e somente se, $\overline{H}^{\mathrm{t}} = H$.

Demonstração. A entrada (i, j) da matriz de g^{t} é $g^{\mathrm{t}}(\mathbf{v}_i, \mathbf{v}_j) = g(\mathbf{v}_j, \mathbf{v}_i)$, que é a entrada (j, i) de G. Isso prova o primeiro item, e os dois itens seguintes seguem imediatamente.

Se g é alternada, ela é antissimétrica e $g(\mathbf{v}_i, \mathbf{v}_i) = 0$ para todo i. Reciprocamente, se $G^{\mathrm{t}} = -G$ e sua diagonal é nula, então, para $\mathbf{x} = (x_1, \dots, x_n)^{\mathrm{t}}$, os termos $g_{ij}x_ix_j$ e $g_{ji}x_jx_i$ se cancelam aos pares, enquanto os termos diagonais são nulos. Logo $\mathbf{x}^{\mathrm{t}}G\mathbf{x} = 0$.

No caso sesquilinear, a mesma comparação de entradas mostra que a matriz da forma adjunta é $\overline{H}^{\mathrm{t}}$. ∎

Exemplo 14.31 — Uma forma integral. No espaço real $C[a, b]$, seja $w : [a, b] \to \mathbb{R}$ uma função contínua fixa. A expressão

$$g(f_1, f_2) = \int_a^b w(x) f_1(x) f_2(x)\, dx$$

define uma forma bilinear simétrica. Se $w(x) > 0$ em todo o intervalo, ela é positiva definida; se w muda de sinal, a forma continua simétrica, mas pode deixar de ser um produto interno.

◁

Exemplo 14.32 — Polarização do determinante. Para $A, B \in \mathscr{M}_{2,2}(\mathbb{K})$, defina

$$g(A, B) = \det(A + B) - \det A - \det B.$$

Escrevendo $A = (a_{ij})$ e $B = (b_{ij})$, obtemos

$$g(A, B) = a_{11}b_{22} + a_{22}b_{11} - a_{12}b_{21} - a_{21}b_{12}.$$

Portanto, g é uma forma bilinear simétrica. Além disso, $g(A, A) = 2 \det A$: a forma recupera, por polarização, a forma quadrática dada pelo determinante em dimensão dois.

◁

Exemplo 14.33 — Espaço de Minkowski. Em $\mathbb{R}^4$, a forma

$$g(\mathbf{x}, \mathbf{y}) = x_1y_1 + x_2y_2 + x_3y_3 - x_4y_4$$

é simétrica e não degenerada, mas não é positiva definida. Sua matriz canônica é $\mathrm{diag}(1, 1, 1, -1)$. Vetores não nulos podem satisfazer $g(\mathbf{x}, \mathbf{x}) = 0$; por exemplo, $(1, 0, 0, 1)$. Esses vetores são chamados de *isotrópicos* ou *tipo luz*.

◁

Exemplo 14.34 — Forma simplética padrão. Escreva os vetores de $\mathbb{K}^{2n}$ como pares $(\mathbf{x}, \vec{p})$, com $\mathbf{x}, \vec{p} \in \mathbb{K}^n$. A forma

$$\omega\big((\mathbf{x}, \vec{p}), (\mathbf{y}, \vec{q})\big) = \mathbf{x}^{\mathrm{t}}\vec{q} - \vec{p}^{\mathrm{t}}\mathbf{y}$$

tem matriz

$$J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}.$$

Ela é alternada e, como J é invertível, não degenerada. Portanto, ω é uma forma simplética.

◁

14.5 Ortogonalidade

Para uma forma bilinear arbitrária, a igualdade $g(\mathbf{u}, \mathbf{v}) = 0$ não implica necessariamente $g(\mathbf{v}, \mathbf{u}) = 0$. Nas formas simétricas, alternadas e hermitianas, entretanto, uma igualdade implica a outra. Nessas três classes, a ortogonalidade é uma relação simétrica.

Ortogonalidade e vetores isotrópicos
Para um produto interno positivo, $v \perp v$ implica $v = 0$. Para outras formas bilineares podem existir vetores não nulos ortogonais a si mesmos, chamados isotrópicos. Por isso, argumentos que dependem da positividade não se estendem diretamente a essas formas.

Definição 14.35 — Ortogonalidade e Complemento Ortogonal. Seja V munido de uma forma simétrica, alternada ou hermitiana. Dizemos que $\mathbf{u}, \mathbf{v} \in V$ são **ortogonais** se $g(\mathbf{u}, \mathbf{v}) = 0$, e escrevemos $\mathbf{u} \perp \mathbf{v}$.
Se $U \subseteq V$ é um subespaço, seu **complemento ortogonal** é

$$U^{\perp} = \{\mathbf{v} \in V \mid g(\mathbf{u}, \mathbf{v}) = 0 \text{ para todo } \mathbf{u} \in U\}.$$

O complemento ortogonal é sempre um subespaço. Além disso, $V^{\perp} = \mathrm{Rad}(g)$, onde, para as formas consideradas nesta seção, os radicais esquerdo e direito coincidem.

Definição 14.36. A restrição de g a U é a forma

$$g\restriction_{U\times U} : U \times U \longrightarrow \mathbb{K}.$$

Dizemos que U é um **subespaço não degenerado** se essa restrição for não degenerada. Equivalentemente,

$$U \cap U^{\perp} = \{\vec{0}\}.$$

Exemplo 14.37 — Uma reta isotrópica. Considere em $\mathbb{R}^2$ a forma

$$g(\mathbf{x}, \mathbf{y}) = x_1 y_1 - x_2 y_2$$

e o subespaço $U = \langle (1,1) \rangle$. Como $g((1,1),(1,1)) = 0$, temos $U \subseteq U^\perp$. Uma conta direta mostra que $U^\perp = U$. Assim, embora a forma em $\mathbb{R}^2$ seja não degenerada, sua restrição a U é degenerada e não existe decomposição $\mathbb{R}^2 = U \oplus U^\perp$. A hipótese de não degenerescência no próximo teorema é, portanto, essencial.

◁

Teorema 14.38 (Decomposição Ortogonal).
Seja V um espaço de dimensão finita munido de uma forma simétrica, alternada ou hermitiana.

a *Se $U \subseteq V$ é não degenerado, então*

$$V = U \oplus U^\perp.$$

b *Se a forma em V é não degenerada, então, para todo subespaço $U \subseteq V$,*

$$\dim U^\perp = \dim V - \dim U \qquad e \qquad (U^\perp)^\perp = U.$$

Demonstração. Considere a aplicação

$$T : V \longrightarrow U^*, \qquad T(\mathbf{v})(\mathbf{u}) = g(\mathbf{u}, \mathbf{v}).$$

No caso bilinear, T é linear; no caso hermitiano, é conjugado-linear. Em ambos os casos, seu núcleo é $U^\perp$ e vale o teorema do núcleo e da imagem.

Se a restrição de g a U é não degenerada, a restrição de T a U é uma bijeção de U em U^* — conjugado-linear no caso hermitiano. Logo, T é sobrejetiva e

$$\dim U^\perp = \dim \ker T = \dim V - \dim U.$$

Como $U \cap U^\perp = \{\vec{0}\}$, as dimensões mostram que $V = U \oplus U^\perp$.

Agora suponha que a forma em V seja não degenerada. A aplicação associada de V em V^* é uma bijeção; compondo-a com a restrição $V^* \to U^*$, vemos que T é sobrejetiva para qualquer U. Portanto, $\dim U^\perp = \dim V - \dim U$. A inclusão $U \subseteq (U^\perp)^\perp$ é imediata, e

$$\dim (U^\perp)^\perp = \dim V - \dim U^\perp = \dim U.$$

A inclusão entre subespaços de mesma dimensão é uma igualdade. ■

Observação 14.39. Em um espaço simplético, toda reta é degenerada, pois $\omega(\mathbf{v}, \mathbf{v}) = 0$. Por isso, a decomposição simplética começa com planos não degenerados, e não com retas. Essa diferença será decisiva na classificação.

14.6 Classificação

A congruência compara as matrizes de duas formas após mudanças de base; a noção intrínseca correspondente é a de isometria.

Invariantes por congruência
Classificar formas significa descobrir quais dados não mudam sob congruência. Sobre $\mathbb{R}$, formas simétricas não degeneradas são determinadas pelos números de quadrados positivos e negativos: essa é a assinatura da forma.

Definição 14.40 — Isometria. Sejam (V, g) e (W, h) espaços munidos de formas do mesmo tipo. Uma **isometria** de (V, g) em (W, h) é um isomorfismo linear $T : V \to W$ tal que

$$h(T\mathbf{u}, T\mathbf{v}) = g(\mathbf{u}, \mathbf{v})$$

para todos $\mathbf{u}, \mathbf{v} \in V$. Quando existe uma isometria, dizemos que as formas, ou os espaços correspondentes, são **isométricos**.

Se duas bases são relacionadas por uma matriz P, o Teorema 14.14 mostra que as matrizes de uma forma bilinear são relacionadas por

$$G' = P^{\mathrm{t}} G P.$$

No caso hermitiano, com nossa convenção de linearidade na primeira variável, a relação correspondente é

$$H' = P^{\mathrm{t}} H \overline{P}.$$

Assim, classificar formas equivale a classificar matrizes por congruência.

14.6.1 Os Blocos Elementares

Em dimensão um, uma forma bilinear simétrica tem a expressão

$$g(x\mathbf{v}, y\mathbf{v}) = a\,xy, \qquad a = g(\mathbf{v}, \mathbf{v}).$$

Se substituímos $\mathbf{v}$ por $c\mathbf{v}$, o coeficiente a é substituído por $c^2 a$. Portanto, sobre um corpo arbitrário, os blocos unidimensionais não nulos são classificados pelas classes de $\mathbb{K}^*/(\mathbb{K}^*)^2$. Em particular:

- sobre $\mathbb{R}$, todo bloco não nulo é isométrico a xy ou a $-xy$;
- sobre $\mathbb{C}$, todo bloco bilinear simétrico não nulo é isométrico a xy.

Para uma forma hermitiana em uma reta complexa,

$$h(z\mathbf{v}, w\mathbf{v}) = a\,z\overline{w},$$

onde $a = h(\mathbf{v}, \mathbf{v}) \in \mathbb{R}$. A troca $\mathbf{v} \mapsto c\mathbf{v}$ multiplica a por $|c|^2$; logo os blocos não nulos são $z\overline{w}$ e $-z\overline{w}$.

Uma forma alternada é nula em toda reta. Seu primeiro bloco não trivial tem dimensão dois:

$$\omega\big((x_1, x_2), (y_1, y_2)\big) = x_1 y_2 - x_2 y_1, \qquad J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Esse plano é não degenerado e será o bloco básico da geometria simplética.

Teorema 14.41 (Decomposição em Blocos).
Seja V um espaço vetorial de dimensão finita.

a *Se g é uma forma bilinear simétrica e $\operatorname{char}(\mathbb{K}) \neq 2$, existe uma base na qual*

$$[g] = \operatorname{diag}(a_1, \dots, a_s, 0_{n-s}), \qquad a_i \neq 0.$$

b *Se h é uma forma hermitiana em um espaço complexo, existe uma base na*

qual

$$[h] = \operatorname{diag}(I_p, -I_q, 0_r).$$

Nesse caso, $p + q + r = \dim V$.

c *Se* ω *é uma forma alternada, existe uma base na qual*

$$[\omega] = \operatorname{diag}(\underbrace{J, \ldots, J}_{s \text{ blocos}}, 0_r).$$

Nesse caso, $2s + r = \dim V$.

Demonstração. A ideia é a mesma nos três casos: encontrar um pequeno subespaço não degenerado, separá-lo por meio do Teorema 14.38 e aplicar indução ao complemento ortogonal.

No caso simétrico, suponha que g não seja nula. A identidade de polarização mostra que existe $\mathbf{v} \in V$ com $g(\mathbf{v}, \mathbf{v}) \neq 0$; caso contrário, g seria identicamente nula. A reta $U = \langle \mathbf{v} \rangle$ é não degenerada. Logo $V = U \oplus U^{\perp}$, e repetimos o argumento em $U^{\perp}$. Quando a restrição restante se torna nula, completamos a base com uma base de seu radical. Isso produz a matriz diagonal do item **a**.

Para uma forma hermitiana não nula, também existe $\mathbf{v}$ com $h(\mathbf{v}, \mathbf{v}) \neq 0$. De fato, se todos os valores diagonais fossem nulos, as expansões de $h(\mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v})$ e $h(\mathbf{u}+i\mathbf{v}, \mathbf{u}+i\mathbf{v})$ mostrariam, respectivamente, que as partes real e imaginária de $h(\mathbf{u}, \mathbf{v})$ são nulas. A reta gerada por $\mathbf{v}$ é, portanto, não degenerada. Após multiplicar $\mathbf{v}$ por um escalar adequado, fazemos $h(\mathbf{v}, \mathbf{v})$ igual a 1 ou -1. Decompomos novamente $V = U \oplus U^{\perp}$ e aplicamos indução.

Finalmente, suponha que ω seja alternada e não nula. Existem $\mathbf{u}, \mathbf{v}$ com $\omega(\mathbf{u}, \mathbf{v}) \neq 0$. Reescalando $\mathbf{u}$, podemos supor que $\omega(\mathbf{u}, \mathbf{v}) = 1$. Os vetores $\mathbf{u}, \mathbf{v}$ são linearmente independentes, e a matriz da restrição ao plano $U = \langle \mathbf{u}, \mathbf{v} \rangle$ é J; em particular, U é não degenerado. Assim, $V = U \oplus U^{\perp}$, e a indução fornece os demais blocos. A parte restante, quando nula, é o radical. ■

Corolário 14.42.
Todo espaço simplético de dimensão finita tem dimensão par.

Demonstração. Se a forma alternada é não degenerada, não há bloco nulo na decomposição do Teorema 14.41. Como cada bloco J tem dimensão dois, $\dim V = 2s$. ■

14.6.2 Invariantes e Teorema de Inércia

A decomposição do teorema anterior não é única: há muitas bases ortogonais possíveis. O número de blocos de cada tipo, contudo, não depende da base.

Definição 14.43 — Índices de Inércia. Para uma forma simétrica real ou hermitiana complexa, escrevemos

$$[g] = \operatorname{diag}(I_p, -I_q, 0_r)$$

em uma base adequada, com $p+q+r = \dim V$. Os números p, q e r são chamados, respectivamente, de **índice positivo**, **índice negativo** e **índice de nulidade**. A tripla (p,q,r) é a **inércia** da forma. Quando $r = 0$, é comum chamar o par (p,q) de **assinatura**; outra convenção chama $p-q$ de assinatura.

Teorema 14.44 (Teorema de Inércia).
Em dimensão finita:

- a *uma forma bilinear simétrica real é determinada, a menos de isometria, por sua inércia (p,q,r);*
- b *uma forma hermitiana complexa é determinada, a menos de isometria, por sua inércia (p,q,r);*
- c *uma forma bilinear simétrica complexa é determinada, a menos de isometria, por sua dimensão e seu posto;*
- d *uma forma alternada sobre $\mathbb{K}$ é determinada, a menos de isometria, por sua dimensão e seu posto, sendo este sempre par.*

Demonstração. A existência das formas normais segue do Teorema 14.41. No caso real simétrico, cada coeficiente diagonal não nulo pode ser normalizado para 1 ou -1. No caso hermitiano, os coeficientes diagonais são reais e a mesma normalização é possível. No caso bilinear simétrico complexo, todo número complexo não nulo possui raiz quadrada, de modo que todos os coeficientes não nulos podem ser normalizados para 1. A forma alternada já aparece em blocos J.

Resta justificar a unicidade de p e q nos dois primeiros casos. Considere uma decomposição

$$V = V_+ \oplus V_- \oplus V_0$$

na qual a forma é positiva definida em V_+, negativa definida em V_- e nula em V_0. Se $L \subseteq V$ é um subespaço no qual a forma é positiva definida, a projeção $L \to V_+$ é injetiva. De fato, um vetor não nulo no núcleo dessa projeção pertenceria a $V_- \oplus V_0$ e teria valor quadrático menor ou igual a zero, uma contradição. Logo, $\dim L \le \dim V_+ = p$. Como V_+ tem dimensão p, esse número é a maior dimensão possível de um subespaço positivo definido e, portanto, é intrínseco à forma. O mesmo argumento aplicado a $-g$ mostra que q também é intrínseco. Por fim, $r = \dim \operatorname{Rad}(g)$.

Nos itens c e d, o posto determina o número de blocos não nulos: um por coordenada no caso simétrico complexo e metade do posto no caso alternado. A dimensão determina então o tamanho do bloco nulo. Formas com os mesmos invariantes possuem a mesma matriz normal e, por isso, são isométricas. ∎

Exemplo 14.45 — Cálculo da inércia. Considere em $\mathbb{R}^3$ a forma quadrática

$$q(x,y,z) = 2x^2 + 2xy + 2y^2 - z^2.$$

Completando o quadrado,

$$q(x,y,z) = 2\left(x + \frac{y}{2}\right)^2 + \frac{3}{2}y^2 - z^2.$$

A mudança de coordenadas é invertível e exibe duas direções positivas e uma negativa. Portanto, a forma bilinear simétrica associada tem inércia $(2, 1, 0)$. O Teorema 14.44 garante que nenhuma outra mudança de base pode alterar esses três números.

◁

A representação por matrizes e a congruência permitem formular a classificação das formas em coordenadas. A decomposição ortogonal reduz suas matrizes a blocos elementares, cujos invariantes distinguem as classes de formas, independentemente das bases escolhidas.

No Capítulo 15, essa passagem da bilinearidade para a linearidade ganhará uma realização concreta. O produto tensorial transforma o Teorema 14.17 na identificação natural

$$\mathrm{Bil}(V, W) \cong (V \otimes W)^*,$$

reunindo formas bilineares, dualidade e transformações lineares em uma mesma construção.

15

Produtos Tensoriais e Álgebra Multilinear

Uma aplicação linear recebe um vetor por vez. Muitas operações naturais, entretanto, dependem linearmente de várias entradas: o produto de polinômios, a multiplicação de matrizes, o determinante e as formas bilineares estudadas no Capítulo 14 são exemplos fundamentais. O produto tensorial é a construção que permite tratar essas operações multilineares por meio das ferramentas usuais da álgebra linear.

A propriedade que caracteriza o produto tensorial é a correspondência entre aplicações bilineares e transformações lineares. Uma aplicação bilinear

$$B : V \times W \longrightarrow Z$$

passará a ser representada por uma única transformação linear

$$\widetilde{B} : V \otimes W \longrightarrow Z, \qquad \widetilde{B}(\mathbf{v} \otimes \mathbf{w}) = B(\mathbf{v}, \mathbf{w}).$$

O símbolo $\mathbf{v} \otimes \mathbf{w}$ não representa um par ordenado nem um produto numérico. Ele é um vetor de um novo espaço, construído precisamente para que a igualdade acima faça sentido.

Neste capítulo trabalharemos sempre sobre um mesmo corpo $\mathbb{K}$. Quando tratarmos da decomposição entre partes simétrica e alternada, suporemos $\mathrm{char}(\mathbb{K}) \neq 2$. Salvo indicação em contrário, os espaços envolvidos serão de dimensão finita; a construção do produto tensorial, contudo, não depende dessa hipótese.

Nesse capítulo

15.1 Aplicações Multilineares

O ponto de partida é ampliar a bilinearidade para qualquer número finito de variáveis.

Definição 15.1 — Aplicação Multilinear. Sejam $V_1, \ldots, V_r$ e Z espaços vetoriais sobre $\mathbb{K}$. Uma aplicação

$$F : V_1 \times \cdots \times V_r \longrightarrow Z$$

é denominada **aplicação multilinear**, ou r-linear, se for linear em cada variável

quando as demais permanecem fixas. Explicitamente, para $1 \le k \le r$,

$$\begin{aligned} &F(\mathbf{v}_1, \dots, a\mathbf{u} + b\mathbf{u}', \dots, \mathbf{v}_r) \\ &\qquad = aF(\mathbf{v}_1, \dots, \mathbf{u}, \dots, \mathbf{v}_r) + bF(\mathbf{v}_1, \dots, \mathbf{u}', \dots, \mathbf{v}_r). \end{aligned}$$

Quando $Z = \mathbb{K}$, a aplicação é chamada de **forma multilinear**.

Definição 15.2 — Aplicação Alternada. Uma aplicação r-linear $F : V^r \to Z$ é **alternada** quando $F(\mathbf{v}_1, \dots, \mathbf{v}_r) = \vec{0}$ sempre que duas entradas coincidem. Para $r = 2$, recuperamos a noção introduzida na Definição 14.25.

Se F é alternada, trocar duas entradas troca o sinal de seu valor. De fato, expandindo por multilinearidade uma entrada duplicada, obtemos

$$0 = F(\dots, \mathbf{u} + \mathbf{v}, \dots, \mathbf{u} + \mathbf{v}, \dots) = F(\dots, \mathbf{u}, \dots, \mathbf{v}, \dots) + F(\dots, \mathbf{v}, \dots, \mathbf{u}, \dots).$$

Em característica diferente de 2, vale a recíproca; em característica 2, ela pode falhar, como já vimos na Proposição 14.28.

Para $r = 1$, recuperamos as transformações lineares; para $r = 2$, recuperamos as aplicações bilineares. Como na situação bilinear, uma aplicação multilinear se anula sempre que uma de suas entradas é o vetor nulo.

Exemplo 15.3 — Produto de polinômios. Para inteiros não negativos $m_1, \dots, m_r$, a aplicação

$$\mathbb{K}_{m_1}[x] \times \cdots \times \mathbb{K}_{m_r}[x] \longrightarrow \mathbb{K}_{m_1+\cdots+m_r}[x], \qquad (p_1, \dots, p_r) \longmapsto p_1 \cdots p_r,$$

é multilinear. Fixados todos os fatores exceto um, resta a multiplicação por um polinômio fixo, que é uma transformação linear.

◁

Exemplo 15.4 — Uma forma integral. No espaço $C([a,b])$ das funções contínuas, a aplicação

$$\Phi(f_1, \dots, f_r) = \int_a^b f_1(t) \cdots f_r(t)\, dt$$

é uma forma r-linear. A observação seguinte muda o domínio e o contradomínio: em $C^1([a,b])$, a aplicação

$$(f_1, \dots, f_r) \longmapsto (f_1 \cdots f_r)'$$

também é r-linear, pela regra de Leibniz, mas seu valor pertence a $C([a,b])$, e não ao corpo de escalares. Trata-se, portanto, de uma aplicação multilinear que não é uma forma.

◁

Exemplo 15.5 — Determinante. Se $V = \mathbb{K}^n$, então

$$(\mathbf{v}_1, \dots, \mathbf{v}_n) \longmapsto \det \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix}$$

é uma forma n-linear. Além disso, ela é alternada: seu valor é zero quando duas entradas coincidem. Essa propriedade voltará a aparecer na construção das potências exteriores.

◁

Exemplo 15.6 — Multiplicação sucessiva de matrizes. Se os tamanhos são compatíveis, a aplicação

$$\mathscr{M}_{m,n}(\mathbb{K}) \times \mathscr{M}_{n,p}(\mathbb{K}) \times \mathscr{M}_{p,q}(\mathbb{K}) \longrightarrow \mathscr{M}_{m,q}(\mathbb{K}), \qquad (A, B, C) \longmapsto ABC,$$

é trilinear. A associatividade da multiplicação permite escrever ABC sem ambiguidade, mas não é responsável pela trilinearidade: esta decorre da distributividade em cada fator.

◁

Linear em cada entrada
Ao verificar a multilinearidade, alteramos uma entrada por vez. Variar várias entradas simultaneamente produz termos mistos; por isso uma aplicação multilinear não é, em geral, linear no produto cartesiano dos domínios.

Em dimensão finita, os valores de uma aplicação multilinear nos vetores de bases determinam todos os seus valores.

Teorema 15.7 (Extensão Multilinear).
Para cada k, seja $\underline{v}_k = (\mathbf{v}_{k1}, \dots, \mathbf{v}_{kn_k})$ uma base de V_k. Escolhidos arbitrariamente vetores

$$\mathbf{z}_{i_1,\dots,i_r} \in Z, \qquad 1 \le i_k \le n_k,$$

existe uma única aplicação multilinear $F : V_1 \times \cdots \times V_r \to Z$ tal que

$$F(\mathbf{v}_{1i_1}, \dots, \mathbf{v}_{ri_r}) = \mathbf{z}_{i_1,\dots,i_r}.$$

Demonstração. Escreva cada entrada na base correspondente:

$$\mathbf{x}_k = \sum_{i_k=1}^{n_k} a_{ki_k} \mathbf{v}_{ki_k}.$$

Se F existir, sua multilinearidade obrigará

$$F(\mathbf{x}_1, \dots, \mathbf{x}_r) = \sum_{i_1=1}^{n_1} \cdots \sum_{i_r=1}^{n_r} a_{1i_1} \cdots a_{ri_r}\, \mathbf{z}_{i_1,\dots,i_r}.$$

Essa fórmula define uma aplicação multilinear com os valores prescritos e, como qualquer outra aplicação com a mesma propriedade teria de satisfazer a mesma expansão, também prova a unicidade. ■

Exemplo 15.8 — Reconstrução pelas coordenadas. Uma forma trilinear $F : (\mathbb{R}^2)^3 \to \mathbb{R}$ fica determinada por oito números $c_{ijk} = F(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k)$, com $i, j, k \in \{1, 2\}$. Se $\mathbf{x} = (x_1, x_2)$, $\mathbf{y} = (y_1, y_2)$ e $\mathbf{z} = (z_1, z_2)$, então

$$F(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \sum_{i,j,k=1}^{2} c_{ijk} x_i y_j z_k.$$

Assim, os vários índices não introduzem uma nova espécie de álgebra: eles apenas registram uma coordenada para cada entrada.

◁

15.2 O Problema que o Produto Tensorial Resolve

Considere uma aplicação bilinear $B : V \times W \to Z$. Embora $V \times W$ seja um espaço vetorial, B não é em geral linear nesse espaço. De fato,

$$B(\mathbf{v}_1 + \mathbf{v}_2, \mathbf{w}_1 + \mathbf{w}_2)$$

contém quatro termos, enquanto a linearidade no par produziria apenas dois. Queremos substituir $V \times W$ por um espaço no qual a bilinearidade de B se torne linearidade ordinária.

Definição 15.9 — Produto Tensorial. Sejam V e W espaços vetoriais sobre $\mathbb{K}$. Um **produto tensorial** de V e W é um par $(V \otimes W, \tau)$, formado por um espaço vetorial $V \otimes W$ e uma aplicação bilinear

$$\tau : V \times W \longrightarrow V \otimes W,$$

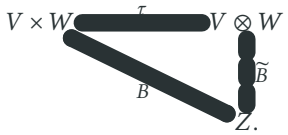
com a seguinte propriedade: para todo espaço vetorial Z e toda aplicação bilinear $B : V \times W \to Z$, existe uma única transformação linear $\widetilde{B} : V \otimes W \to Z$ tal que

$$B = \widetilde{B} \circ \tau.$$

Escrevemos

$$\mathbf{v} \otimes \mathbf{w} := \tau(\mathbf{v}, \mathbf{w}).$$

No diagrama, a propriedade afirma que toda aplicação bilinear definida em $V \times W$ se fatora de maneira única pelo produto tensorial:



A correspondência $B \mapsto \widetilde{B}$ é bijetiva: sua inversa envia $L : V \otimes W \to Z$ em $L \circ \tau$. Ela também é linear, pois a unicidade da fatoração implica

$$\widetilde{aB + bC} = a\widetilde{B} + b\widetilde{C}.$$

Portanto, a propriedade universal pode ser escrita como o isomorfismo natural

$$\mathrm{Hom}(V \otimes W, Z) \cong \mathrm{Bil}(V, W; Z).$$

O lado esquerdo envolve apenas transformações lineares; o lado direito reúne as aplicações bilineares que desejamos compreender.

O que é universal?
O espaço $V \otimes W$ não depende da aplicação bilinear particular que se pretende estudar. Uma mesma construção lineariza simultaneamente todas as aplicações bilineares definidas em $V \times W$.

Exemplo 15.10 — Uma forma bilinear se torna um funcional. Se $g : V \times W \to \mathbb{K}$ é uma forma bilinear, existe um único funcional

$$\tilde{g} : V \otimes W \longrightarrow \mathbb{K}$$

tal que $\tilde{g}(\mathbf{v} \otimes \mathbf{w}) = g(\mathbf{v}, \mathbf{w})$. Para o produto escalar usual em $\mathbb{R}^n$,

$$\tilde{g}(\mathbf{e}_i \otimes \mathbf{e}_j) = \delta_{ij}.$$

Logo, se $z = \sum_{i,j} a_{ij} \mathbf{e}_i \otimes \mathbf{e}_j$, então $\tilde{g}(z) = \sum_i a_{ii}$.

◁

Exemplo 15.11 — Multiplicação de polinômios. A multiplicação

$$\mu : \mathbb{K}_m[x] \times \mathbb{K}_n[x] \longrightarrow \mathbb{K}_{m+n}[x]$$

é bilinear e, portanto, induz uma transformação linear

$$\tilde{\mu} : \mathbb{K}_m[x] \otimes \mathbb{K}_n[x] \longrightarrow \mathbb{K}_{m+n}[x], \qquad p \otimes q \longmapsto pq.$$

Por exemplo,

$$\tilde{\mu}\big(1 \otimes x^2 + 2x \otimes x - x^2 \otimes 1\big) = x^2 + 2x^2 - x^2 = 2x^2.$$

◁

Exemplo 15.12 — Multiplicação de matrizes. A aplicação $(A, B) \mapsto AB$ induz uma transformação linear

$$\mathscr{M}_{m,n}(\mathbb{K}) \otimes \mathscr{M}_{n,p}(\mathbb{K}) \longrightarrow \mathscr{M}_{m,p}(\mathbb{K}).$$

Isso não torna a multiplicação de matrizes comutativa. O produto tensorial lineariza uma operação; ele não acrescenta propriedades que a operação original não possuía.

◁

Teorema 15.13 (Unicidade do Produto Tensorial).
Dois produtos tensoriais de V e W são canonicamente isomorfos. Mais precisamente, se (T, τ) e (T', τ') satisfazem a Definição 15.9, existe um único isomorfismo $S : T \to T'$ tal que $S \circ \tau = \tau'$.

Demonstração. Pela universalidade de (T, τ) aplicada a τ', existe uma única transformação linear $S : T \to T'$ tal que $S \circ \tau = \tau'$. Analogamente, existe uma única transformação linear $R : T' \to T$ tal que $R \circ \tau' = \tau$.

Temos

$$(RS) \circ \tau = R \circ \tau' = \tau = \mathrm{I}_T \circ \tau.$$

A unicidade na propriedade universal implica $RS = \mathrm{I}_T$. Da mesma forma, $SR = \mathrm{I}_{T'}$. Logo, S é um isomorfismo, com inversa R, e sua própria unicidade já foi obtida na primeira aplicação da propriedade universal. ■

Unicidade a menos de isomorfismo
Duas construções do produto tensorial podem representar seus elementos de modos diferentes. Existe, porém, um único isomorfismo entre elas compatível com as aplicações bilineares canônicas, isto é, que leva cada tensor puro ao tensor correspondente na outra construção.

Esse resultado permite falar em *o* produto tensorial sem escolher uma realização particular. Ainda precisamos mostrar que ele existe.

15.3 Construção e Cálculo com Tensores

A construção seguinte começa com símbolos formais e impõe exatamente as relações necessárias para que o mapa $(\mathbf{v}, \mathbf{w}) \mapsto \mathbf{v} \otimes \mathbf{w}$ seja bilinear.

Seja $F(V \times W)$ o espaço vetorial livre sobre o conjunto $V \times W$: concretamente, seus elementos são combinações lineares finitas de símbolos $[\mathbf{v}, \mathbf{w}]$, e esses símbolos formam

uma base. Seja R o subespaço gerado pelos vetores

$$\begin{aligned}[\mathbf{v}_1+\mathbf{v}_2,\mathbf{w}]-[\mathbf{v}_1,\mathbf{w}]-[\mathbf{v}_2,\mathbf{w}],\\ [\mathbf{v},\mathbf{w}_1+\mathbf{w}_2]-[\mathbf{v},\mathbf{w}_1]-[\mathbf{v},\mathbf{w}_2],\\ [a\mathbf{v},\mathbf{w}]-a[\mathbf{v},\mathbf{w}],\\ [\mathbf{v},a\mathbf{w}]-a[\mathbf{v},\mathbf{w}],\end{aligned}$$

onde os vetores e o escalar a são arbitrários. Definimos

$$V\otimes W := F(V\times W)/R$$

e denotamos por $\mathbf{v}\otimes\mathbf{w}$ a classe de $[\mathbf{v},\mathbf{w}]$ nesse quociente.

Teorema 15.14 (Existência do Produto Tensorial).
O espaço $F(V\times W)/R$, com a aplicação

$$\tau(\mathbf{v},\mathbf{w})=\mathbf{v}\otimes\mathbf{w},$$

satisfaz a propriedade universal do produto tensorial.

Demonstração. Pela definição de R, as quatro famílias de relações acima tornam-se igualdades no quociente. Logo, τ é bilinear.

Seja agora $B : V\times W\to Z$ uma aplicação bilinear. Como $F(V\times W)$ é livre sobre $V\times W$, existe uma única transformação linear

$$L : F(V\times W)\longrightarrow Z,\qquad L([\mathbf{v},\mathbf{w}])=B(\mathbf{v},\mathbf{w}).$$

A bilinearidade de B mostra que cada gerador de R pertence a $\ker L$. Assim, $R\subseteq\ker L$, e a Propriedade Universal do Quociente, Teorema 4.17, fornece uma única transformação linear

$$\widetilde{B} : F(V\times W)/R\longrightarrow Z$$

com $\widetilde{B}(\mathbf{v}\otimes\mathbf{w})=B(\mathbf{v},\mathbf{w})$. Como os tensores $\mathbf{v}\otimes\mathbf{w}$ geram o quociente, essa condição também garante a unicidade de $\widetilde{B}$. ∎

Da construção obtemos imediatamente as regras de cálculo fundamentais:

$$(\mathbf{u}+\mathbf{v})\otimes\mathbf{w}=\mathbf{u}\otimes\mathbf{w}+\mathbf{v}\otimes\mathbf{w}, \tag{15.1}$$

$$\mathbf{v}\otimes(\mathbf{w}+\mathbf{z})=\mathbf{v}\otimes\mathbf{w}+\mathbf{v}\otimes\mathbf{z}, \tag{15.2}$$

$$(a\mathbf{v})\otimes\mathbf{w}=a(\mathbf{v}\otimes\mathbf{w})=\mathbf{v}\otimes(a\mathbf{w}). \tag{15.3}$$

Em particular,

$$\vec{0}\otimes\mathbf{w}=\mathbf{v}\otimes\vec{0}=\vec{0}.$$

Exemplo 15.15 — Uma expansão completa. Se $\underline{\mathrm{E}}=(\mathbf{e}_1,\mathbf{e}_2)$ e $\underline{\mathrm{F}}=(\vec{f}_1,\vec{f}_2,\vec{f}_3)$, então

$$\begin{aligned}&(2\mathbf{e}_1-\mathbf{e}_2)\otimes(\vec{f}_1+3\vec{f}_2-\vec{f}_3)\\ &\quad=2\mathbf{e}_1\otimes\vec{f}_1+6\mathbf{e}_1\otimes\vec{f}_2-2\mathbf{e}_1\otimes\vec{f}_3\\ &\qquad-\mathbf{e}_2\otimes\vec{f}_1-3\mathbf{e}_2\otimes\vec{f}_2+\mathbf{e}_2\otimes\vec{f}_3.\end{aligned}$$

A expansão se comporta como um produto distributivo, mas cada termo $\mathbf{e}_i\otimes\vec{f}_j$ é um vetor do novo espaço.

◁

Observação 15.16 — O produto tensorial não é multiplicação. Não se pode cancelar fatores em uma soma de tensores nem distribuir uma soma que não esteja dentro de um dos fatores. Por exemplo,

$$\mathbf{u} \otimes \mathbf{w} + \mathbf{v} \otimes \mathbf{z}$$

normalmente não pode ser escrito como um único tensor puro. A notação $\otimes$ sugere um produto porque é bilinear, não porque possua todas as propriedades da multiplicação de números.

Exemplo 15.17 — O corpo como unidade. A aplicação $\mu : \mathbb{K} \times V \to V$ dada por $\mu(a, \mathbf{v}) = a\mathbf{v}$ é bilinear. Ela induz

$$\tilde{\mu} : \mathbb{K} \otimes V \longrightarrow V, \qquad a \otimes \mathbf{v} \longmapsto a\mathbf{v}.$$

Sua inversa é $\mathbf{v} \mapsto 1 \otimes \mathbf{v}$. Portanto,

$$\mathbb{K} \otimes V \cong V \cong V \otimes \mathbb{K}.$$

◁

Proposição 15.18 (Quando um Tensor Puro é Nulo).
Para $\mathbf{v} \in V$ e $\mathbf{w} \in W$,

$$\mathbf{v} \otimes \mathbf{w} = \vec{0} \quad \Longleftrightarrow \quad \mathbf{v} = \vec{0} \text{ ou } \mathbf{w} = \vec{0}.$$

Demonstração. Uma das implicações segue das regras de cálculo. Para a recíproca, suponha $\mathbf{v} \neq \vec{0}$ e $\mathbf{w} \neq \vec{0}$. Escolha funcionais $f \in V^*$ e $g \in W^*$ tais que $f(\mathbf{v}) = g(\mathbf{w}) = 1$. A forma bilinear $(\mathbf{x}, \mathbf{y}) \mapsto f(\mathbf{x})g(\mathbf{y})$ induz um funcional em $V \otimes W$ que assume o valor 1 em $\mathbf{v} \otimes \mathbf{w}$. Logo, $\mathbf{v} \otimes \mathbf{w} \neq \vec{0}$. ■

15.4 Bases e Tensores Puros

O próximo resultado descreve uma base de $V \otimes W$ a partir de bases de V e de W. Ele permite calcular a dimensão do produto tensorial e escrever seus elementos em coordenadas.

Teorema 15.19 (Base do Produto Tensorial).
Se $\underline{v} = (\mathbf{v}_1, \ldots, \mathbf{v}_m)$ é uma base de V e $\underline{w} = (\mathbf{w}_1, \ldots, \mathbf{w}_n)$ é uma base de W, então

$$\underline{v} \otimes \underline{w} = \left(\mathbf{v}_i \otimes \mathbf{w}_j \,:\, 1 \leq i \leq m,\ 1 \leq j \leq n\right)$$

é uma base de $V \otimes W$. Consequentemente,

$$\dim(V \otimes W) = \dim V \dim W.$$

Demonstração. As regras de cálculo mostram que todo tensor puro é combinação linear dos vetores $\mathbf{v}_i \otimes \mathbf{w}_j$. Como os tensores puros geram $V \otimes W$, essa família gera o espaço.

Para provar a independência linear, sejam $\mathbf{v}_1^*, \dots, \mathbf{v}_m^*$ e $\mathbf{w}_1^*, \dots, \mathbf{w}_n^*$ as bases duais. Para cada par (k, ℓ), a forma bilinear

$$B_{k\ell}(\mathbf{v}, \mathbf{w}) = \mathbf{v}_k^*(\mathbf{v})\mathbf{w}_\ell^*(\mathbf{w})$$

induz um funcional $\widetilde{B}_{k\ell}$ em $V \otimes W$. Se

$$\sum_{i,j} a_{ij}\mathbf{v}_i \otimes \mathbf{w}_j = \vec{0},$$

aplicar $\widetilde{B}_{k\ell}$ fornece $a_{k\ell} = 0$. Como isso vale para todo par, a família é linearmente independente. ■

Assim, todo $z \in V \otimes W$ possui uma expressão única

$$z = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\mathbf{v}_i \otimes \mathbf{w}_j.$$

Os coeficientes a_{ij} podem ser organizados em uma matriz. Essa observação fornece nosso exemplo recorrente.

Exemplo 15.20 — $\mathbb{R}^2 \otimes \mathbb{R}^3$ como espaço de matrizes. Sejam $\underline{\mathrm{E}} = (\mathbf{e}_1, \mathbf{e}_2)$ e $\underline{\mathrm{F}} = (\vec{f}_1, \vec{f}_2, \vec{f}_3)$ as bases canônicas. A aplicação bilinear

$$\mathbb{R}^2 \times \mathbb{R}^3 \longrightarrow \mathscr{M}_{2,3}(\mathbb{R}), \qquad (\mathbf{x}, \mathbf{y}) \longmapsto \mathbf{x}\mathbf{y}^{\mathrm{t}},$$

induz um isomorfismo

$$\mathbb{R}^2 \otimes \mathbb{R}^3 \longrightarrow \mathscr{M}_{2,3}(\mathbb{R})$$

que leva $\mathbf{e}_i \otimes \vec{f}_j$ à matriz elementar E_{ij}. Portanto,

$$\sum_{i=1}^{2}\sum_{j=1}^{3} a_{ij}\mathbf{e}_i \otimes \vec{f}_j \longleftrightarrow \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}.$$

A expansão do Exemplo 15.15 corresponde à matriz

$$\begin{bmatrix} 2 & 6 & -2 \\ -1 & -3 & 1 \end{bmatrix},$$

que tem posto um porque é o produto de um vetor-coluna por um vetor-linha.

◁

Definição 15.21 — Tensor Puro e Posto Tensorial. Um tensor da forma $\mathbf{v} \otimes \mathbf{w}$ é chamado de **tensor puro**, **tensor simples** ou tensor decomponível. O **posto tensorial** de $z \in V \otimes W$ é o menor inteiro r para o qual existem vetores $\mathbf{v}_k \in V$ e $\mathbf{w}_k \in W$ tais que

$$z = \sum_{k=1}^{r} \mathbf{v}_k \otimes \mathbf{w}_k.$$

Por convenção, o tensor nulo tem posto zero.

Proposição 15.22 (Posto Tensorial, Matrizes e SVD).
No caso de dois fatores, o posto tensorial de $z \in V \otimes W$ coincide com o posto de sua matriz de coeficientes em quaisquer bases. Se V e W são espaços com produto interno real e as bases são ortonormais, a decomposição em valores singulares fornece uma expressão mínima

$$z = \sum_{k=1}^{r} s_k \mathbf{u}_k \otimes \mathbf{w}_k,$$

na qual $s_k > 0$ e as famílias $(\mathbf{u}_k)$ e $(\mathbf{w}_k)$ são ortonormais.

Demonstração. Na identificação com matrizes, um tensor puro $\mathbf{u} \otimes \mathbf{w}$ corresponde ao produto de um vetor-coluna por um vetor-linha e, portanto, tem posto no máximo um. Assim, uma soma de r tensores puros produz uma matriz de posto no máximo r. Reciprocamente, uma matriz de posto r é soma de r matrizes de posto um; pela identificação inversa, o tensor é soma de r tensores puros. Os dois postos são, portanto, iguais.

Quando as bases são ortonormais, aplicamos o Teorema 13.36 à matriz de coeficientes. A decomposição $A = Q\Sigma P^{\mathrm{t}}$ se traduz exatamente na expressão exibida, com os vetores singulares como fatores. A minimalidade decorre da primeira parte. ■

O produto interno dos fatores induz em $V \otimes W$ o produto interno caracterizado por

$$\langle \mathbf{u} \otimes \mathbf{w}, \mathbf{u}' \otimes \mathbf{w}' \rangle = \langle \mathbf{u}, \mathbf{u}' \rangle \langle \mathbf{w}, \mathbf{w}' \rangle.$$

Assim, os termos da decomposição anterior são ortogonais. Essa é a ligação entre o posto tensorial e a decomposição em valores singulares estudada no Capítulo 13. Para três ou mais fatores, a situação se torna consideravelmente mais delicada: não existe uma forma canônica análoga à SVD.

Exemplo 15.23 — Nem todo tensor é puro. Em $\mathbb{R}^2 \otimes \mathbb{R}^2$, considere

$$z = \mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_2.$$

Sob a identificação com $\mathscr{M}_{2,2}(\mathbb{R})$, esse tensor corresponde à matriz identidade. Como essa matriz tem posto dois, z não pode ser escrito na forma $\mathbf{u} \otimes \mathbf{v}$.

Em contraste,

$$\mathbf{e}_1 \otimes \mathbf{e}_1 + 2\mathbf{e}_1 \otimes \mathbf{e}_2 + 3\mathbf{e}_2 \otimes \mathbf{e}_1 + 6\mathbf{e}_2 \otimes \mathbf{e}_2 = (\mathbf{e}_1 + 3\mathbf{e}_2) \otimes (\mathbf{e}_1 + 2\mathbf{e}_2)$$

é puro; sua matriz de coeficientes tem posto um.

◁

Puros geram, mas não formam uma base
Todo tensor é uma soma finita de tensores puros. Entretanto, os tensores puros não constituem um subespaço: a soma de dois deles pode ter posto dois, como no Exemplo 15.23.

Proposição 15.24 (Critério de Igualdade por Formas Bilineares).
Dois tensores $z, z' \in V \otimes W$ são iguais se, e somente se,

$$\widetilde{B}(z) = \widetilde{B}(z')$$

para toda forma bilinear $B : V \times W \to \mathbb{K}$.

Demonstração. Uma implicação é imediata. Para a outra, escreva $z - z'$ na base $\mathbf{v}_i \otimes \mathbf{w}_j$ e use as formas B_{ij} construídas na prova do Teorema 15.19. Elas recuperam individualmente todos os coeficientes de $z - z'$. ■

15.5 Transformações Lineares e Produtos Tensoriais

Se $S : V \to V'$ e $T : W \to W'$ são transformações lineares, a aplicação

$$V \times W \longrightarrow V' \otimes W', \qquad (\mathbf{v}, \mathbf{w}) \longmapsto S(\mathbf{v}) \otimes T(\mathbf{w}),$$

é bilinear. A propriedade universal produz a transformação seguinte.

Definição 15.25 — Produto Tensorial de Transformações. O **produto tensorial** de S e T é a única transformação linear

$$S \otimes T : V \otimes W \longrightarrow V' \otimes W'$$

que satisfaz

$$(S \otimes T)(\mathbf{v} \otimes \mathbf{w}) = S(\mathbf{v}) \otimes T(\mathbf{w}).$$

Proposição 15.26 (Propriedades Funtoriais).
Para transformações de domínios e contradomínios compatíveis,

a $\mathrm{I}_V \otimes \mathrm{I}_W = \mathrm{I}_{V \otimes W}$;

b $(S_2 \circ S_1) \otimes (T_2 \circ T_1) = (S_2 \otimes T_2) \circ (S_1 \otimes T_1)$;

c *se S e T são isomorfismos, então $S \otimes T$ é um isomorfismo e*

$$(S \otimes T)^{-1} = S^{-1} \otimes T^{-1};$$

d *em dimensão finita,*

$$\operatorname{posto}(S \otimes T) = \operatorname{posto}(S)\operatorname{posto}(T).$$

Demonstração. Nos três primeiros itens, basta verificar as igualdades nos tensores puros, que geram o domínio. Por exemplo,

$$\begin{aligned} \big((S_2 \otimes T_2) \circ (S_1 \otimes T_1)\big)(\mathbf{v} \otimes \mathbf{w}) &= S_2(S_1(\mathbf{v})) \otimes T_2(T_1(\mathbf{w})) \\ &= \big((S_2 \circ S_1) \otimes (T_2 \circ T_1)\big)(\mathbf{v} \otimes \mathbf{w}). \end{aligned}$$

Para o último item, observe que

$$\operatorname{im}(S \otimes T) = \operatorname{im}(S) \otimes \operatorname{im}(T),$$

pois ambos os lados são gerados pelos tensores $S(\mathbf{v}) \otimes T(\mathbf{w})$. O Teorema 15.19 fornece então $\dim \operatorname{im}(S \otimes T) = \dim \operatorname{im}(S) \dim \operatorname{im}(T)$, que é a igualdade desejada. ■

Em bases, o produto tensorial de transformações aparece como o produto de Kronecker de matrizes.

Definição 15.27 — Produto de Kronecker. Se $A = (a_{ij}) \in \mathscr{M}_{m,n}(\mathbb{K})$ e $B \in \mathscr{M}_{p,q}(\mathbb{K})$, o **produto de Kronecker** de A e B é a matriz em blocos

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}.$$

Proposição 15.28 (Matriz de $S \otimes T$).
Sejam $A = [S]_{\underline{V},\underline{V}'}$ e $B = [T]_{\underline{W},\underline{W}'}$. Ordene a base $\underline{V} \otimes \underline{W}$ lexicograficamente, isto é,

$$\mathbf{v}_1 \otimes \mathbf{w}_1, \dots, \mathbf{v}_1 \otimes \mathbf{w}_n, \mathbf{v}_2 \otimes \mathbf{w}_1, \dots, \mathbf{v}_m \otimes \mathbf{w}_n,$$

e faça o mesmo com $\underline{V}' \otimes \underline{W}'$. Nessas bases, a matriz de $S \otimes T$ é o produto de Kronecker $A \otimes B$.

Demonstração. A coluna de índice (j, ℓ) é o vetor de coordenadas de

$$\begin{aligned}(S \otimes T)(\mathbf{v}_j \otimes \mathbf{w}_\ell) &= S\mathbf{v}_j \otimes T\mathbf{w}_\ell \\ &= \sum_{i,k} a_{ij} b_{k\ell}\, \mathbf{v}'_i \otimes \mathbf{w}'_k.\end{aligned}$$

Sua entrada de índice (i, k) é $a_{ij}b_{k\ell}$, exatamente a entrada correspondente do bloco $a_{ij}B$. ∎

Exemplo 15.29 — Cálculo de um produto de Kronecker. Para

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

temos

$$A \otimes B = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 3 & -1 & 6 & -2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & -1 \end{bmatrix}.$$

A organização em blocos registra simultaneamente a ação de A no primeiro fator e a ação de B no segundo.

◁

Os isomorfismos seguintes são canônicos; suas fórmulas nos tensores puros determinam completamente as transformações.

Teorema 15.30 (Regras Estruturais).
Existem isomorfismos naturais

$$\begin{aligned}(V \otimes W) \otimes U &\cong V \otimes (W \otimes U), \\ V \otimes W &\cong W \otimes V, \\ (V \oplus W) \otimes U &\cong (V \otimes U) \oplus (W \otimes U), \\ \mathbb{K} \otimes V &\cong V.\end{aligned}$$

Eles são dados, respectivamente, por

$$\begin{aligned}(\mathbf{v}\otimes\mathbf{w})\otimes\mathbf{u} &\longmapsto \mathbf{v}\otimes(\mathbf{w}\otimes\mathbf{u}),\\ \mathbf{v}\otimes\mathbf{w} &\longmapsto \mathbf{w}\otimes\mathbf{v},\\ (\mathbf{v},\mathbf{w})\otimes\mathbf{u} &\longmapsto (\mathbf{v}\otimes\mathbf{u},\mathbf{w}\otimes\mathbf{u}),\\ a\otimes\mathbf{v} &\longmapsto a\mathbf{v}.\end{aligned}$$

Demonstração. A associatividade requer duas aplicações sucessivas da propriedade universal. Fixado $\mathbf{u}\in U$, a aplicação $(\mathbf{v},\mathbf{w})\mapsto \mathbf{v}\otimes(\mathbf{w}\otimes\mathbf{u})$ é bilinear e induz uma transformação $L_{\mathbf{u}} : V\otimes W\to V\otimes(W\otimes U)$. A dependência de $L_{\mathbf{u}}$ em $\mathbf{u}$ é linear; logo $(z,\mathbf{u})\mapsto L_{\mathbf{u}}(z)$ é bilinear e induz a transformação do primeiro isomorfismo. Repetindo o argumento com os parênteses invertidos, obtemos sua inversa.

Nos outros três casos, cada fórmula e sua fórmula inversa provêm diretamente de aplicações bilineares. As composições fixam todos os tensores puros; como estes geram os espaços envolvidos, são as identidades. No terceiro isomorfismo, escrevemos os elementos da soma direta externa $V\oplus W$ como pares $(\mathbf{v},\mathbf{w})$. ■

Graças ao primeiro isomorfismo, escreveremos produtos com três ou mais fatores sem indicar parênteses. O segundo isomorfismo não afirma que $\mathbf{v}\otimes\mathbf{w}=\mathbf{w}\otimes\mathbf{v}$: em geral, esses vetores pertencem a espaços diferentes. Ele afirma que existe uma identificação natural entre os dois espaços.

15.6 Dualidade, Operadores e Contrações

O produto tensorial permite descrever relações entre espaços duais, formas bilineares e operadores por meio de isomorfismos independentes de bases.

Teorema 15.31 (Tensores como Transformações Lineares).
Se V tem dimensão finita, a transformação

$$\Phi : V^*\otimes W\longrightarrow \mathrm{Hom}(V,W),\qquad \Phi(f\otimes\mathbf{w})(\mathbf{v})=f(\mathbf{v})\mathbf{w},$$

é um isomorfismo.

Demonstração. A aplicação $(f,\mathbf{w})\mapsto(\mathbf{v}\mapsto f(\mathbf{v})\mathbf{w})$ é bilinear e, portanto, induz Φ. Se $\underline{\mathrm{v}}=(\mathbf{v}_1,\dots,\mathbf{v}_m)$ é uma base de V, com base dual $(\mathbf{v}_1^*,\dots,\mathbf{v}_m^*)$, e $\underline{\mathrm{w}}=(\mathbf{w}_1,\dots,\mathbf{w}_n)$ é uma base de W, então $\Phi(\mathbf{v}_i^*\otimes\mathbf{w}_j)$ é a transformação que leva $\mathbf{v}_i$ a $\mathbf{w}_j$ e anula os demais vetores da base. Essas transformações são as que correspondem às matrizes elementares e, pelo Teorema 3.47, formam uma base de $\mathrm{Hom}(V,W)$. Assim, Φ leva uma base em uma base e é um isomorfismo. ■

Exemplo 15.32 — Operadores de posto um. Um tensor puro $f \otimes \mathbf{w} \in V^* \otimes W$ corresponde ao operador

$$\mathbf{v} \longmapsto f(\mathbf{v})\mathbf{w},$$

cuja imagem está contida na reta gerada por $\mathbf{w}$. Se $f \neq 0$ e $\mathbf{w} \neq \vec{0}$, esse operador tem posto um. Assim, decompor uma transformação linear como soma de operadores de posto um equivale a decompor o tensor correspondente como soma de tensores puros.

◁

Exemplo 15.33 — A matriz de uma transformação reaparece. Se $T : V \to W$ satisfaz

$$T(\mathbf{v}_i) = \sum_{j=1}^{n} a_{ji}\mathbf{w}_j,$$

o tensor correspondente a T é

$$\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ji}\mathbf{v}_i^* \otimes \mathbf{w}_j.$$

Os mesmos coeficientes formam a matriz de T; o produto tensorial apenas revela o objeto intrínseco que essa matriz representa depois da escolha de bases.

◁

Há também um isomorfismo canônico

$$(V \otimes W)^* \cong \mathrm{Bil}(V, W),$$

que envia um funcional λ à forma $(\mathbf{v}, \mathbf{w}) \mapsto \lambda(\mathbf{v} \otimes \mathbf{w})$. Em dimensão finita, o Teorema 15.31 e o Teorema 14.17 fornecem a cadeia

$$V^* \otimes W^* \cong \mathrm{Hom}(V, W^*) \cong \mathrm{Bil}(V, W) \cong (V \otimes W)^*.$$

Definição 15.34 — Contração. A forma bilinear de avaliação

$$V^* \times V \longrightarrow \mathbb{K}, \qquad (f, \mathbf{v}) \longmapsto f(\mathbf{v}),$$

induz a transformação linear

$$\mathrm{contr} : V^* \otimes V \longrightarrow \mathbb{K}, \qquad f \otimes \mathbf{v} \longmapsto f(\mathbf{v}),$$

denominada **contração**.

Contrair é emparelhar
Uma contração reúne um fator em V^* e um fator em V por meio da avaliação $f(\mathbf{v})$. Em coordenadas, isso corresponde a somar sobre um índice repetido; para operadores, essa soma é o traço.

Sob o isomorfismo $V^* \otimes V \cong \mathrm{Hom}(V, V)$, a contração é exatamente o traço.

Proposição 15.35 (Contração e Traço).
Se $T \in \mathrm{Hom}(V, V)$ corresponde ao tensor

$$t_T = \sum_{k=1}^{r} f_k \otimes \mathbf{v}_k,$$

então

$$\mathrm{tr}(T) = \mathrm{contr}(t_T) = \sum_{k=1}^{r} f_k(\mathbf{v}_k).$$

Demonstração. Para um operador de posto um $T(\mathbf{x}) = f(\mathbf{x})\mathbf{v}$, o caso $\mathbf{v} = \vec{0}$ é imediato. Se $\mathbf{v} \neq \vec{0}$, escolha uma base cujo primeiro vetor seja $\mathbf{v}$. A soma dos elementos diagonais da matriz de T é $f(\mathbf{v})$. O resultado geral segue pela linearidade do traço e da contração. ■

Exemplo 15.36 — O tensor identidade. Se $\underline{\mathrm{v}} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ tem base dual $(\mathbf{v}_1^*, \dots, \mathbf{v}_n^*)$, o operador identidade corresponde ao tensor

$$\mathrm{I}_V \longleftrightarrow \sum_{i=1}^{n} \mathbf{v}_i^* \otimes \mathbf{v}_i.$$

Embora a expressão use uma base, o tensor obtido não depende dela: ele é o único tensor que corresponde ao operador identidade. Sua contração é $\dim V$.

◁

15.7 Potências Tensoriais e Tipos de Tensores

Os isomorfismos associativos permitem definir

$$V^{\otimes r} = \underbrace{V \otimes \cdots \otimes V}_{r \text{ fatores}}.$$

Se $\dim V = n$ e $\underline{\mathrm{E}} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$, então os vetores

$$\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_r}, \qquad 1 \leq i_1, \dots, i_r \leq n,$$

formam uma base de $V^{\otimes r}$; portanto,

$$\dim V^{\otimes r} = n^r.$$

Teorema 15.37 (Propriedade Universal Multilinear).
Para todo espaço vetorial Z, existe um isomorfismo natural

$$\mathrm{Hom}(V_1 \otimes \cdots \otimes V_r, Z) \cong \mathrm{Mult}(V_1, \dots, V_r; Z).$$

Ele associa a L a aplicação multilinear

$$(\mathbf{v}_1, \dots, \mathbf{v}_r) \longmapsto L(\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_r).$$

Demonstração. Aplicamos repetidamente a propriedade universal bilinear e o isomorfismo associativo do Teorema 15.30. Equivalentemente, podemos repetir a construção por quociente do Teorema 15.14, impondo linearidade em cada uma das r entradas. Em ambos os casos, os tensores puros $\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_r$ geram o domínio e garantem a unicidade. ■

Definição 15.38 — Tensores Covariantes e Contravariantes. Um **tensor de tipo** (p, q) em V é um elemento de

$$\mathcal{T}_q^p(V) = V^{\otimes p} \otimes (V^*)^{\otimes q}.$$

Os fatores em V são chamados de contravariantes e os fatores em V^*, de covariantes. Alguns autores invertem a ordem dos índices; por isso, a definição do espaço deve sempre acompanhar a notação.

Os índices registram os fatores
No tipo (p, q), os números contam quantas cópias de V e de V^* aparecem, não a dimensão do espaço nem as coordenadas de um tensor. A convenção varia entre autores; a fórmula de $\mathcal{T}_q^p(V)$ elimina a ambiguidade.

Exemplo 15.39 — Objetos já conhecidos como tensores. Com a convenção anterior:

- **a** vetores são tensores de tipo $(1, 0)$;
- **b** funcionais são tensores de tipo $(0, 1)$;
- **c** formas bilineares são tensores de tipo $(0, 2)$;
- **d** operadores lineares são tensores de tipo $(1, 1)$;
- **e** formas r-lineares são tensores de tipo $(0, r)$.

O vocabulário novo reúne objetos que já apareceram separadamente ao longo do livro.

◁

Exemplo 15.40 — Coordenadas de um tensor de tipo $(1, 2)$. Se $\underline{\mathrm{E}} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ e $(\mathbf{e}_1^*, \dots, \mathbf{e}_n^*)$ é a base dual, todo $T \in \mathcal{T}_2^1(V)$ se escreve de modo único como

$$T = \sum_{i,j,k=1}^{n} T_{jk}^{i}\, \mathbf{e}_i \otimes \mathbf{e}_j^* \otimes \mathbf{e}_k^*.$$

Os índices superiores registram fatores em V e os inferiores, fatores em V^*.

Se $\mathbf{e}'_j = \sum_i P_{ij}\mathbf{e}_i$ e $Q = P^{-1}$, então $(\mathbf{e}'_j)^* = \sum_i Q_{ji}\mathbf{e}_i^*$, e as coordenadas mudam segundo

$$T'^{i}_{jk} = \sum_{a,b,c=1}^{n} Q_{ia} P_{bj} P_{ck}\, T^{a}_{bc}.$$

Um índice superior transforma-se por Q e um inferior, por P. Essa lei não é uma definição de tensor: é consequência da mudança simultânea das bases de V e de V^*.

◁

A avaliação entre um fator V^* e um fator V permite contrair qualquer tensor de tipo (p, q), produzindo um tensor de tipo $(p-1, q-1)$. Em termos intrínsecos, usamos os isomorfismos de troca e associatividade para aproximar os dois fatores escolhidos e aplicamos $\mathrm{contr} : V^* \otimes V \to \mathbb{K}$. Como todas essas transformações são canônicas, a operação não depende de bases. Em coordenadas, ela corresponde a igualar um índice superior e um inferior e somar sobre esse índice.

15.8 Simetria, Alternância e Potências Exteriores

Considere o operador de troca

$$\sigma : V \otimes V \longrightarrow V \otimes V, \qquad \sigma(\mathbf{u} \otimes \mathbf{v}) = \mathbf{v} \otimes \mathbf{u}.$$

Suponha nesta primeira parte que $\operatorname{char}(\mathbb{K}) \neq 2$. Como $\sigma^2 = \mathrm{I}_{V\otimes V}$, seus autovalores possíveis são 1 e -1, que são distintos nessa característica.

Definição 15.41 — Tensores Simétricos e Antissimétricos. Um tensor $z \in V \otimes V$ é **simétrico** se $\sigma(z) = z$ e **antissimétrico** se $\sigma(z) = -z$. Denotamos os subespaços correspondentes por $\operatorname{Sym}^2(V)$ e $\operatorname{Alt}^2(V)$.

Os operadores

$$\operatorname{Sym}(z) = \frac{1}{2}(z + \sigma z), \qquad \operatorname{Alt}(z) = \frac{1}{2}(z - \sigma z)$$

são projeções complementares.

Teorema 15.42 (Decomposição Simétrica e Alternada).
Se $\operatorname{char}(\mathbb{K}) \neq 2$, *então*

$$V \otimes V = \operatorname{Sym}^2(V) \oplus \operatorname{Alt}^2(V).$$

Se $\dim V = n$, *então*

$$\dim \operatorname{Sym}^2(V) = \frac{n(n+1)}{2}, \qquad \dim \operatorname{Alt}^2(V) = \frac{n(n-1)}{2}.$$

Demonstração. Todo tensor admite a decomposição

$$z = \frac{1}{2}(z + \sigma z) + \frac{1}{2}(z - \sigma z).$$

A primeira parcela é simétrica e a segunda é antissimétrica. Um tensor que pertence aos dois subespaços satisfaz $z = -z$, portanto $2z = 0$ e $z = 0$.

Para uma base $(\mathbf{e}_1, \dots, \mathbf{e}_n)$, escreva $z = \sum_{i,j} a_{ij}\mathbf{e}_i \otimes \mathbf{e}_j$. A igualdade $\sigma(z) = z$ equivale a $a_{ij} = a_{ji}$; portanto, a parte simétrica é gerada por

$$\mathbf{e}_i \otimes \mathbf{e}_i \quad \text{e} \quad \mathbf{e}_i \otimes \mathbf{e}_j + \mathbf{e}_j \otimes \mathbf{e}_i, \qquad i < j.$$

Analogamente, $\sigma(z) = -z$ equivale a $a_{ij} = -a_{ji}$ e, como $2a_{ii} = 0$, a parte antissimétrica é gerada por

$$\mathbf{e}_i \otimes \mathbf{e}_j - \mathbf{e}_j \otimes \mathbf{e}_i, \qquad i < j.$$

Em cada lista, a comparação dos coeficientes na base $(\mathbf{e}_i \otimes \mathbf{e}_j)$ mostra imediatamente a independência linear. As listas são, portanto, bases, e sua contagem fornece as dimensões. ∎

Essa decomposição é especial do grau dois. Se, em potências maiores, chamarmos de totalmente simétricos os tensores fixados por todas as permutações e de totalmente

antissimétricos os que mudam pelo sinal da permutação, em geral esses dois subespaços não esgotam $V^{\otimes r}$. Por exemplo, para $V = \mathbb{R}^2$,

$$\dim V^{\otimes 3} = 8, \qquad \dim \operatorname{Sym}^3(V) = 4, \qquad \dim \operatorname{Alt}^3(V) = 0.$$

Logo, $V^{\otimes 3} \neq \operatorname{Sym}^3(V) \oplus \operatorname{Alt}^3(V)$.

Para graus maiores, é mais natural caracterizar diretamente o objeto que lineariza aplicações alternadas.

Definição 15.43 — Potência Exterior. A r-ésima **potência exterior** de V é um espaço $\bigwedge^r V$ equipado com uma aplicação alternada r-linear

$$V^r \longrightarrow \bigwedge^r V, \qquad (\mathbf{v}_1, \dots, \mathbf{v}_r) \longmapsto \mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_r,$$

universal para aplicações alternadas: toda aplicação alternada $F : V^r \to Z$ se escreve de modo único como

$$F(\mathbf{v}_1, \dots, \mathbf{v}_r) = \widetilde{F}(\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_r)$$

para alguma transformação linear $\widetilde{F} : \bigwedge^r V \to Z$.

Teorema 15.44 (Existência da Potência Exterior).
A potência exterior $\bigwedge^r V$ existe para todo $r \geq 1$.

Demonstração. Seja R_r o subespaço de $V^{\otimes r}$ gerado pelos tensores puros $\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_r$ nos quais duas entradas são iguais. Tome

$$\bigwedge^r V = V^{\otimes r} / R_r$$

e denote por $\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_r$ a classe de $\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_r$. A aplicação que envia uma r-upla nessa classe é multilinear e alternada.

Se $F : V^r \to Z$ é multilinear e alternada, o Teorema 15.37 fornece uma transformação linear $L : V^{\otimes r} \to Z$ com $L(\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_r) = F(\mathbf{v}_1, \dots, \mathbf{v}_r)$. A alternância implica $R_r \subseteq \ker L$; pelo Teorema 4.17, L se fatora de modo único pelo quociente. Essa é a propriedade universal da Definição 15.43. ∎

Como a aplicação universal é alternada,

$$\mathbf{v} \wedge \mathbf{v} = 0 \qquad \text{e} \qquad \mathbf{u} \wedge \mathbf{v} = -\mathbf{v} \wedge \mathbf{u}.$$

A segunda igualdade é a regra de troca deduzida após a Definição 15.2. Quando $\operatorname{char}(\mathbb{K}) \neq 2$, a aplicação

$$\mathbf{u} \wedge \mathbf{v} \longmapsto \frac{1}{2}(\mathbf{u} \otimes \mathbf{v} - \mathbf{v} \otimes \mathbf{u})$$

identifica canonicamente $\bigwedge^2 V$ com o subespaço $\operatorname{Alt}^2(V)$ definido acima.

Teorema 15.45 (Base das Potências Exteriores).
Se $\underline{E} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ é uma base de V, então

$$\mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_r}, \qquad 1 \le i_1 < \cdots < i_r \le n,$$

formam uma base de $\bigwedge^r V$. Em particular,

$$\dim \bigwedge^r V = \binom{n}{r}.$$

Se $r > n$, então $\bigwedge^r V = \{\vec{0}\}$.

Demonstração. Pela multilinearidade, os produtos exteriores dos vetores da base geram $\bigwedge^r V$. Um produto com índice repetido é nulo; trocando fatores, podemos ordenar os índices ao custo do sinal da permutação. Logo, os vetores indicados geram.

Para a independência linear, para cada sequência $I = (i_1 < \cdots < i_r)$ considere a forma alternada que associa a $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ o determinante da submatriz formada pelas coordenadas nas linhas $i_1, \dots, i_r$. A transformação linear induzida em $\bigwedge^r V$ vale 1 no vetor de índice I e 0 nos demais vetores da família. Portanto, todos os coeficientes de uma relação linear são nulos. ■

Definição 15.46 — Álgebra Exterior. Adotamos $\bigwedge^0 V = \mathbb{K}$ e reunimos todas as potências exteriores em

$$\bigwedge V = \bigoplus_{r=0}^{n} \bigwedge^r V.$$

O produto exterior de elementos puros é definido por concatenação,

$$(\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_p) \wedge (\mathbf{w}_1 \wedge \cdots \wedge \mathbf{w}_q) = \mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_p \wedge \mathbf{w}_1 \wedge \cdots \wedge \mathbf{w}_q,$$

e estendido bilinearmente. O espaço graduado $\bigwedge V$, com esse produto, é a **álgebra exterior** de V.

Se $\alpha \in \bigwedge^p V$ e $\beta \in \bigwedge^q V$ são homogêneos, mover os p fatores de α através dos q fatores de β fornece

$$\alpha \wedge \beta = (-1)^{pq} \beta \wedge \alpha.$$

Além disso, o Teorema 15.45 e o binômio de Newton dão

$$\dim \bigwedge V = \sum_{r=0}^{n} \binom{n}{r} = 2^n.$$

Observação 15.47 — Uma origem comum. A **álgebra tensorial** $\mathcal{T}(V) = \bigoplus_{r \ge 0} V^{\otimes r}$ usa a concatenação de tensores como produto. A álgebra exterior é seu quociente pelas relações $\mathbf{v} \otimes \mathbf{v} = 0$; a álgebra simétrica $\mathrm{Sym}(V)$ é o quociente pelas relações $\mathbf{u} \otimes \mathbf{v} - \mathbf{v} \otimes \mathbf{u} = 0$. Assim, as construções simétrica e exterior nascem da mesma álgebra, mas impõem relações diferentes.

Exemplo 15.48 — Bivetores e menores de ordem dois. Em $\mathbb{R}^3$, sejam $\mathbf{u} = (u_1, u_2, u_3)$ e $\mathbf{v} = (v_1, v_2, v_3)$. Então

$$\begin{aligned}\mathbf{u} \wedge \mathbf{v} = {} & (u_1 v_2 - u_2 v_1)\mathbf{e}_1 \wedge \mathbf{e}_2 \\ & + (u_1 v_3 - u_3 v_1)\mathbf{e}_1 \wedge \mathbf{e}_3 \\ & + (u_2 v_3 - u_3 v_2)\mathbf{e}_2 \wedge \mathbf{e}_3.\end{aligned}$$

Os coeficientes são os menores de ordem dois da matriz com colunas $\mathbf{u}$ e $\mathbf{v}$. O bivetor é nulo exatamente quando essas colunas são linearmente dependentes.

◁

Exemplo 15.49 — Área orientada no plano. Em $\mathbb{R}^2$,

$$(a\mathbf{e}_1 + b\mathbf{e}_2) \wedge (c\mathbf{e}_1 + d\mathbf{e}_2) = (ad - bc)\mathbf{e}_1 \wedge \mathbf{e}_2.$$

O determinante aparece como a única coordenada de $\bigwedge^2 \mathbb{R}^2$, que é unidimensional. Seu módulo é a área do paralelogramo; o sinal registra a orientação.

◁

A aplicação $V^r \to \bigwedge^r W$ que envia $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ em $T\mathbf{v}_1 \wedge \cdots \wedge T\mathbf{v}_r$ é multilinear e alternada. Pela propriedade universal, ela induz uma única transformação linear

$$\bigwedge^r T : \bigwedge^r V \longrightarrow \bigwedge^r W, \qquad \mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_r \longmapsto T\mathbf{v}_1 \wedge \cdots \wedge T\mathbf{v}_r.$$

Teorema 15.50 (Determinante como Ação no Volume).
Se $\dim V = n$ *e* $T \in \operatorname{Hom}(V, V)$, *então* $\bigwedge^n V$ *é unidimensional e*

$$\bigwedge^n T = (\det T)\, \mathrm{I}_{\bigwedge^n V}.$$

Demonstração. Escolha uma base $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ de V. Pelo Teorema 15.45, $\mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_n$ é uma base de $\bigwedge^n V$, e basta calcular a imagem desse vetor. Escrevendo $A = [T]_{\underline{\mathrm{E}}} = (a_{ij})$, temos $T\mathbf{e}_j = \sum_i a_{ij}\mathbf{e}_i$; logo, a multilinearidade fornece

$$T\mathbf{e}_1 \wedge \cdots \wedge T\mathbf{e}_n = \sum_{i_1, \dots, i_n = 1}^{n} a_{i_1,1} \cdots a_{i_n,n}\, \mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_n}.$$

Pela alternância, as parcelas com algum índice repetido se anulam. Restam as que correspondem a uma permutação $\sigma \in S_n$, e reordenar os fatores custa $\operatorname{sgn}(\sigma)$. Portanto,

$$\begin{aligned}T\mathbf{e}_1 \wedge \cdots \wedge T\mathbf{e}_n &= \left(\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} \right) \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_n \\ &= (\det A)\, \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_n,\end{aligned}$$

pela fórmula de Leibniz. Como $\det A = \det T$, concluímos que $\bigwedge^n T = (\det T)\, \mathrm{I}_{\bigwedge^n V}$. ∎

O determinante, introduzido como uma função de matrizes, admite assim uma descrição intrínseca: é o escalar pelo qual o operador induzido age na potência exterior de grau máximo, que tem dimensão um. Essa descrição relaciona a construção multilinear às propriedades do determinante estudadas anteriormente.

Exercícios

Ex. 15.1 — Verifique diretamente que a aplicação

$$F : (\mathbb{R}^2)^3 \to \mathbb{R}, \qquad F(\mathbf{x},\mathbf{y},\mathbf{z}) = x_1 y_2 z_1 + x_2 y_1 z_2,$$

é trilinear. Determine seus oito coeficientes na base canônica.

Ex. 15.2 — Mostre que $(\mathbf{u},\mathbf{v}) \mapsto \mathbf{u}+\mathbf{v}$ não é bilinear, exceto no espaço vetorial nulo.

Ex. 15.3 — Expanda

$$(\mathbf{e}_1 - 2\mathbf{e}_2) \otimes (3\vec{f}_1 + \vec{f}_2 - 4\vec{f}_3)$$

na base $\mathbf{e}_i \otimes \vec{f}_j$.

Ex. 15.4 — Prove diretamente, usando a propriedade universal, que $V \otimes \mathbb{K} \cong V$.

Ex. 15.5 — Em $\mathbb{R}^2 \otimes \mathbb{R}^2$, determine quais dos tensores seguintes são puros:

1. $\mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_1 \otimes \mathbf{e}_2$;
2. $\mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_2$;
3. $\mathbf{e}_1 \otimes \mathbf{e}_1 + 2\mathbf{e}_1 \otimes \mathbf{e}_2 + 2\mathbf{e}_2 \otimes \mathbf{e}_1 + 4\mathbf{e}_2 \otimes \mathbf{e}_2$.

Fundamentos

Ex. 15.6 — Se $z \in V \otimes W$ tem posto tensorial r e $a \neq 0$, prove que az também tem posto r.

Ex. 15.7 — Se $S : V \to V'$ e $T : W \to W'$ são sobrejetivas, prove que $S \otimes T$ é sobrejetiva. Faça o mesmo para injetividade em dimensão finita.

Ex. 15.8 — Se $A \in \mathscr{M}_{n,n}(\mathbb{K})$ e $B \in \mathscr{M}_{q,q}(\mathbb{K})$, a fórmula geral é $\det(A \otimes B) = \det(A)^q \det(B)^n$. Calcule $A \otimes B$ e verifique a fórmula no caso $n = q = 2$, para

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

Ex. 15.9 — Descreva explicitamente o isomorfismo

$$(V \oplus W) \otimes U \cong (V \otimes U) \oplus (W \otimes U)$$

e escreva sua inversa.

Ex. 15.10 — Seja $T : \mathbb{R}^2 \to \mathbb{R}^3$ representada na base canônica por

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

Escreva o tensor de $(\mathbb{R}^2)^* \otimes \mathbb{R}^3$ correspondente a T.

Ex. 15.11 — Prove diretamente que o traço do operador de posto um $\mathbf{x} \mapsto f(\mathbf{x})\mathbf{v}$ é $f(\mathbf{v})$.

Consolidação

Ex. 15.12 — Para $\mathbf{u} = (1, 2, 0)$ e $\mathbf{v} = (0, 1, 3)$, calcule $\mathbf{u} \wedge \mathbf{v}$ na base

$$(\mathbf{e}_1 \wedge \mathbf{e}_2, \mathbf{e}_1 \wedge \mathbf{e}_3, \mathbf{e}_2 \wedge \mathbf{e}_3).$$

Ex. 15.13 — Mostre que $\mathbf{u}_1 \wedge \cdots \wedge \mathbf{u}_r = 0$ se os vetores $\mathbf{u}_1, \ldots, \mathbf{u}_r$ são linearmente dependentes. Prove a recíproca usando o Teorema 15.45.

Ex. 15.14 — Se $T : V \to V$ tem posto menor que r, prove que $\bigwedge^r T = 0$.

Ex. 15.15 — Seja $A \in \mathscr{M}_{n,n}(\mathbb{K})$ e, para $1 \le i \le n$, defina

$$\omega_i = (-1)^{i-1} \mathbf{e}_1 \wedge \cdots \wedge \widehat{\mathbf{e}_i} \wedge \cdots \wedge \mathbf{e}_n.$$

Mostre que a matriz de $\bigwedge^{n-1} A$ na base $(\omega_1, \ldots, \omega_n)$ é a matriz dos cofatores de A. Conclua que sua transposta é a matriz adjugada de A.

Appendices

A
Teoria de Conjuntos

Nesse capítulo

Dados dois conjuntos A e B, diremos que A é **subconjunto** de B, e escreveremos $A \subset B$, se cada elemento de A pertence a B, ou seja, se $x \in A$ implica $x \in B$.

Diremos que A é **igual** a B, e escreveremos $A = B$, se A e B possuem os mesmos elementos, ou seja, se $A \subset B$ e $B \subset A$.

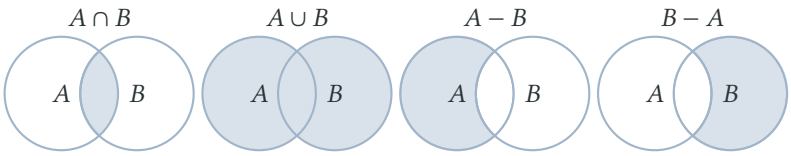
O **conjunto vazio** será denotado por $\emptyset$.

Definição A.1. Definimos a **união**, a **intersecção** e a **diferença** de dois conjuntos A e B, respectivamente, por

$$A \cup B = \{x : x \in A \text{ ou } x \in B\},$$

$$A \cap B = \{x : x \in A \text{ e } x \in B\},$$

$$A \backslash B = \{x : x \in A \text{ e } x \notin B\}.$$



Se estamos considerando subconjuntos de um conjunto fixo X, então conjunto $X \backslash A$ é denominado de **complementar** de A em X, e é denotado por A^c.

Dado um conjunto X, o **conjunto das partes** $\mathscr{P}(X)$ é o conjunto formado pelos subconjuntos de X, ou seja,

$$\mathscr{P}(X) = \{A : A \subset X\}.$$

Definição A.2 — Produto Cartesiano. O **produto cartesiano** $X \times Y$ de dois conjuntos X e Y é o conjunto dos pares ordenados $(x,\ y)$ tais que $x \in X$ e $y \in Y$.
O produto cartesiano $X_1 \times \cdots \times X_n$ de n conjuntos $X_1, \ldots, X_n$ é o conjunto das ênuplas $(x_1, \ldots, x_n)$ tais que $x_i \in X_i$ para $i = 1, \ldots, n$.
Escreveremos X^n em lugar de $X \times \cdots \times X$ (n vezes).

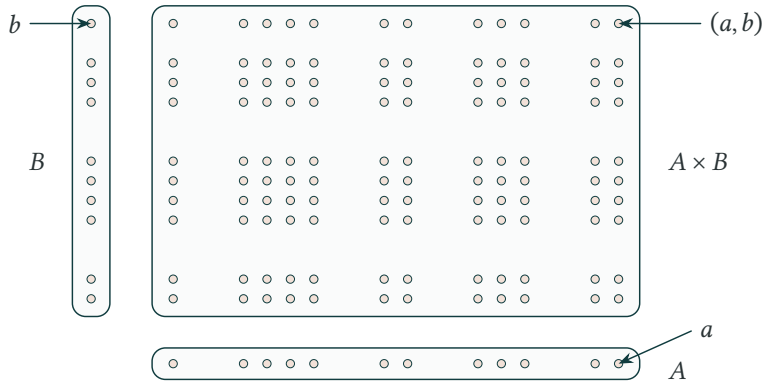


Figura A.1 Produto Cartesiano de dois conjuntos

Definição A.3. Uma **função** ou **aplicação** f de X em Y, denotada por $f : X \to Y$, é uma regra que associa a cada elemento $x \in X$ um único elemento $f(x) \in$ Y. Nesse caso o conjunto X é denominado de domínio de f e o conjunto Y é denominado de contradomínio de f.

Definição A.4. Uma função f é dita

a **injetiva** se $f(x_1) = f(x_2)$ implica $x_1 = x_2$.

b **sobrejetiva** se para todo $y \in Y$ existe $x \in X$ tal que $f(x) = y$.

c **bijetiva** se é injetiva e sobrejetiva.

Se $f : X \to Y$ é bijetiva, a **função inversa** $f^{-1} : Y \to X$ é definida por $f^{-1}(y) = x$ se $f(x) = y$.

Dados $A \subset X$ e $B \subset Y$, a **imagem** de A e a **imagem inversa** de B são respectivamente os conjuntos

$$f(A) = \{y \in Y : y = f(x) \text{ para algum } x \in A\}, \tag{A.1}$$

$$f^{-1}(B) = \{x \in X : f(x) \in B\}. \tag{A.2}$$

Dadas duas aplicações $f : X \to Y$ e $g : Y \to Z$, a **aplicação composta** $g \circ f : X \to Z$ é definida por $(g \circ f)(x) = g(f(x))$ para todo $x \in X$.

Definição A.5. □ $g : B \to A$ é uma inversa à esquerda de $f : A \to B$ se, e somente se, $g(f(a)) = a$ para todos $a \in A$

□ $h : B \to A$ é uma inversa à direita de $f : A \to B$ se, e somente se, $f(h(b)) = b$ para todos os $b \in B$

Teorema A.6. □ *Toda função que possui uma inversa à esquerda é injetiva. Reciprocamente, toda função injetiva com domínio não vazio possui uma inversa à esquerda.*

□ *Toda função que possui uma inversa à direita é sobrejetiva. Admitindo o*

> *Axioma da Escolha, toda função sobrejetiva possui uma inversa à direita.*

Demonstração. Se $g \circ f = \mathrm{I}_X$ e $f(x_1) = f(x_2)$, aplicar g dá $x_1 = x_2$. Na recíproca, fixe $x_0 \in X$ e defina $g(y)$ como a única pré-imagem de y quando $y \in f(X)$, e como x_0 caso contrário. Então $g \circ f = \mathrm{I}_X$.

Se $f \circ h = \mathrm{I}_Y$, cada $y \in Y$ é a imagem de $h(y)$, logo f é sobrejetiva. Reciprocamente, os conjuntos $f^{-1}(\{y\})$, para $y \in Y$, são não vazios; o Axioma da Escolha permite escolher $h(y)$ em cada um deles, e então $f \circ h = \mathrm{I}_Y$. ■

A.1 Operações com famílias de conjuntos

Nesta seção, lidaremos com *famílias (ou classes) de conjuntos*, isto é, conjuntos cujos elementos são, por sua vez, também conjuntos. Queremos estender a essa situação algumas operações entre conjuntos, assim como descrever algumas propriedades.

Seja $\mathscr{F}$ uma família de conjuntos, i.e. $\mathscr{F} = \{A_\iota\}_{\iota \in J}$ onde J é um qualquer conjunto de índices e cada A_ι é um conjunto.

Exemplo A.7.

$$A_i = (0, i] \subset \mathbb{R}$$

com $i \in \mathbb{N}$

◁

Exemplo A.8.

$$B_x = (x, x^2] \subset \mathbb{R}$$

com $x \in \mathbb{R}, x > 2$

◁

A **união** dos conjuntos da família $\mathscr{F}$ é o conjunto formado pelos elementos que pertencem a *ao menos um* dos conjuntos de $\mathscr{F}$, i.e.

$$\bigcup_{\iota \in J} A_\iota = \{x \mid x \in A_\jmath \text{ para algum } \jmath \in J\}$$

A **intersecção** dos conjuntos da família $\mathscr{F}$ é o conjunto formado pelos elementos que pertencem a *todos* os conjuntos de $\mathscr{F}$, i.e.

$$\bigcap_{\iota \in J} A_\iota = \{x \mid x \in A_\jmath \text{ para todo } \jmath \in J\}$$

Dentre as propriedades mais importantes, destacamos as seguintes: dada uma família $\mathscr{F} = \{A_\iota\}_{\iota \in J}$ de conjuntos e dado um conjunto qualquer B, tem-se:

$$B \cap \left(\bigcup_{\iota \in J} A_\iota \right) = \bigcup_{\iota \in J} (B \cap A_\iota)$$

$$B \cup \left(\bigcap_{\imath \in J} A_\imath \right) = \bigcap_{\imath \in J} (B \cup A_\imath)$$

Além disso, se $\mathbb{U}$ é um conjunto que contém todos os conjuntos $A_\imath$, então, tomando o complementar relativamente a $\mathbb{U}$, tem-se:

$$(\bigcup_{\imath \in J} A_\imath)^{\mathrm{c}} = \bigcap_{\imath \in J} A_\imath^{\mathrm{c}}$$

$$(\bigcap_{\imath \in J} A_\imath)^{\mathrm{c}} = \bigcup_{\imath \in J} A_\imath^{\mathrm{c}}$$

Utilizaremos a seguinte convenção:

$$\bigcup_{A \in \emptyset} A = \emptyset$$
$$\bigcap_{A \in \emptyset} A = X.$$

Produto Cartesiano

Como primeiro passo, vejamos como definir o produto cartesiano de uma quantidade qualquer (mas finita) de conjuntos. Dados n conjuntos não vazios $A_1, A_2, \dots, A_n$, o **produto cartesiano** $A_1 \times A_2 \times \cdots \times A_n$ é o conjunto dos elementos na forma $(x_1, x_2, \dots, x_n)$, onde para todo $1 \leq \imath \leq n$ tem-se que $x_\imath \in A_\imath$. Em símbolos:

$$A_1 \times A_2 \times \cdots \times A_n = \{(x_1, x_2, \dots, x_n) \,|\, x_\imath \in A_\imath, 1 \leq \imath \leq n\}.$$

Os elementos na forma $(x_1, x_2, \dots, x_n)$ são denominados de *n-upla ordenada* (que se lê "ênupla" ordenada).

Note-se que o produto cartesiano de n conjuntos é muito semelhante ao produto cartesiano de dois conjuntos, só diferindo, de fato, pelo número de conjuntos envolvidos.

Nosso propósito, agora, é contemplar famílias quaisquer de conjuntos, eventualmente infinitas. Para tanto, não é difícil perceber que a descrição acima não é adequada. Para chegar a um outro modo de tratar o produto cartesiano, pode ser útil revermos, sob outro olhar, o produto cartesiano que nos é já conhecido (vamos considerar o caso mais simples, com somente dois conjuntos). Dados dois conjuntos não vazios A_1 e A_2 (o uso de índices aqui é proposital), podemos identificar um par ordenado (x_1, x_2) do produto cartesiano $A_1 \times A_2$ com a função $f : \{1, 2\} \to (A_1 \cup A_2)$ dada por

$$f(1) = x_1 \quad \text{e} \quad f(2) = x_2$$

Essa representação associa a cada índice a coordenada correspondente do par ordenado. A função registra, portanto, a escolha de um elemento de cada conjunto. A mesma descrição pode ser usada quando o conjunto de índices é arbitrário, sem depender da notação de pares ou de ênuplas.

A vantagem dessa linguagem, porém, está no fato de permitir que se defina o produto cartesiano para uma família qualquer de conjuntos. De fato, seja dada uma família de

conjuntos

$$\mathscr{F} = \{A_\iota\}_{\iota \in J}$$

onde J é um qualquer conjunto de índices. O **produto cartesiano** dos conjuntos da família $\mathscr{F}$ é o conjunto das funções

$$x : J \to \bigcup_{\iota \in J} A_\iota$$

tais que $x(\jmath) \in A_\jmath$ para todo $\jmath \in J$. Em símbolos:

$$\prod_{\iota \in J} A_\iota = \{x : J \to \bigcup_{\iota \in J} A_\iota \mid x(\jmath) \in A_\jmath, \ \forall\ \jmath \in J\}.$$

Escreveremos x_ι em lugar de $x(\iota)$, para todo $x \in \prod_{\iota \in J} A_\iota$ e todo $\iota \in J$. Para cada $\jmath \in J$, a projeção na $\jmath$-ésima coordenada é a função

$$\begin{aligned} \pi_\jmath : \prod_{\iota \in J} A_\iota &\longrightarrow A_\jmath, \\ x &\longmapsto x_\jmath. \end{aligned}$$

Cada elemento desse produto é usualmente denotado por $(x_\iota)_{\iota \in J}$.

Mesmo que todo X_i seja não vazio, não é óbvio que o produto $\prod_{i \in I} X_i$ seja não vazio. Isto é consequência do seguinte axioma.

Axioma A.9 — Axioma da Escolha. Seja $\{X_i : i \in I\}$ uma família não vazia de conjuntos disjuntos não vazios. Então existe uma função $f : I \to \bigcup_{i \in I} X_i$ tal que $f(i) \in X_i$ para todo $i \in I$. A função f é denominada função escolha.

O Axioma da Escolha é um princípio fundamental da Teoria dos Conjuntos. Ele foi usado implicitamente durante muito tempo, antes de Zermelo o formular explicitamente. Sua aceitação gerou controvérsia, sobretudo por objeções construtivistas como as de Brouwer: o ponto não era uma contradição lógica, mas a possibilidade de afirmar a existência de uma escolha sem fornecer um procedimento para construí-la. Mais tarde, Gödel estabeleceu a consistência relativa do Axioma da Escolha com os demais axiomas usuais da teoria dos conjuntos, e Cohen mostrou que ele não pode ser demonstrado a partir desses axiomas, admitida a consistência deles. Assim, a escolha é hoje tratada explicitamente como um axioma adicional, amplamente utilizado na matemática.

> **Proposição A.10.**
> *Seja $\{X_i : i \in I\}$ uma família não vazia de conjuntos não vazios. Então o produto cartesiano $\prod_{i \in I} X_i$ é não vazio.*

Demonstração. Se os conjuntos X_i fossem disjuntos, a conclusão seria consequência imediata do Axioma da Escolha. No caso geral definamos $Y_i = X_i \times \{i\}$ para todo $i \in I$. É claro que $\{Y_i : i \in I\}$ é uma família não vazia de conjuntos disjuntos não vazios. Pelo Axioma da Escolha existe uma função $f : I \to \bigcup_{i \in I} Y_i$ tal que $f(i) \in Y_i$ para todo $i \in I$. Podemos escrever $f(i) = (x_i, i)$, com $x_i \in X_i$ para todo $i \in I$. Se definimos $x(i) = x_i$ para todo $i \in I$, então $x \in \prod_{i \in I} X_i$. ∎

A.2 Relações

Definição A.11. Uma **relação** R num conjunto X é um subconjunto R de $X \times X$. Denotaremos por xRy se $(x,\ y) \in R$.

Uma relação R em X é dita:

a reflexiva se xRx para todo $x \in X$.

b simétrica se xRy implica yRx.

c transitiva se xRy e yRz implicam xRz.

Diremos que R é uma relação de equivalência se R é reflexiva, simétrica e transitiva.

A.2.1 Relação de Equivalência

Usualmente escrevemos

$$a \sim b$$

ao invés de dizer que (a, b) pertence à relação $\sim$.

Definição A.12. Uma relação $\sim$ definida em um conjunto X é uma *relação de equivalência* se satisfizer, para todo $a, b, c \in X$,

1 $a \sim a$.

2 $a \sim b \Rightarrow b \sim a$.

3 $a \sim b$ e $b \sim c \Rightarrow a \sim c$.

Definição A.13. Dada uma relação de equivalência $\sim$ em um conjunto X, para todo $x \in X$, definimos a **classe de equivalência** de x por $\sim$ como

$$[x] = \{y \in X \ :\ y \sim x\}.$$

Observe que, pela reflexividade de $\sim$, $x \in [x]$ para todo $x \in X$.

Proposição A.14.
Seja $\sim$ uma relação de equivalência no conjunto X. Então, para $x, y \in X$, são equivalentes:

a $y \in [x]$;

b $[x] = [y]$;

Demonstração. Sejam $x, y \in X$. Se $y \in [x]$, $y \sim x$, então para todo $z \in [y]$, i.e., $z \sim y$, temos, por transitividade, $z \sim x$, então $z \in [x]$. Por outro lado, pela simetria de $\sim$, temos $x \sim y$, portanto, para todo $z \in [x]$, temos, novamente por transitividade, que $z \in [y]$. Assim, $[x] = [y]$. Por outro lado, se $[x] = [y]$, então $y \in [y] = [x]$. ■

Definição A.15. O elemento $y \in [x]$ é denominado **representante da classe**.

Exemplo A.16. Seja $A = \{a, b, c\}$, e defina $\sim$ no conjunto das partes de A, $\mathscr{P}(A)$, por $X \sim Y$ se, e somente se, $|X| = |Y|$. É simples mostrar que $\sim$ é uma relação de equivalência em $\mathscr{P}(A)$, sob a qual $\mathscr{P}(A)$ tem exatamente 4 classes de equivalência distintas:

$$\begin{aligned}
[\varnothing] &= \{\varnothing\},\\
[\{a\}] = [\{b\}] = [\{c\}] &= \{\{a\}, \{b\}, \{c\}\},\\
[\{a, b\}] = [\{a, c\}] = [\{b, c\}] &= \{\{a, b\}, \{a, c\}, \{b, c\}\}, \text{ e}\\
[A] = \{A\} &= \big\{\{a, b, c\}\big\}.
\end{aligned}$$

◁

Definição A.17. Seja X um conjunto. Então uma coleção de subconjuntos $P = \{X_i\}_{i \in I}$ é uma partição de X se cada $X_i \neq \varnothing$ e cada elemento de X estiver em exatamente um X_i. Em outras palavras, $P = \{X_i\}_{i \in I}$ é uma partição de X se, e somente se:

a $X_i \neq \varnothing$;

b X_i são mutuamente disjuntos: $X_i \cap X_j = \varnothing$ para $i \neq j$ em I; e

c $X = \bigcup_{i \in I} X_i$.

Os conjuntos X_i são chamados de células da partição.

Exemplo A.18. O conjunto {1,2,3 } possui 5 partições: a saber,

$$\{\{1, 2, 3\}\}, \{\{1, 2\}, \{3\}\}, \{\{1, 3\}, \{2\}\}, \{\{2, 3\}, \{1\}\} \text{ e } \{\{1\}, \{2\}, \{3\}\}.$$

A primeira partição que mencionamos tem uma célula, as próximas três têm duas células e a última tem três células.

◁

Teorema A.19.
Seja X um conjunto. Então:

1 *Se $\sim$ é uma relação de equivalência em X, então o conjunto de todas as classes de equivalência de X por $\sim$ é uma partição de X;*

2 *Se P for uma partição de X, então a relação em X definida por $x \sim y$ se e somente x estiver na mesma célula de P que y é uma relação de equivalência em X.*

Observe que, em cada caso, as células da partição são as classes de equivalência do conjunto na relação de equivalência correspondente.

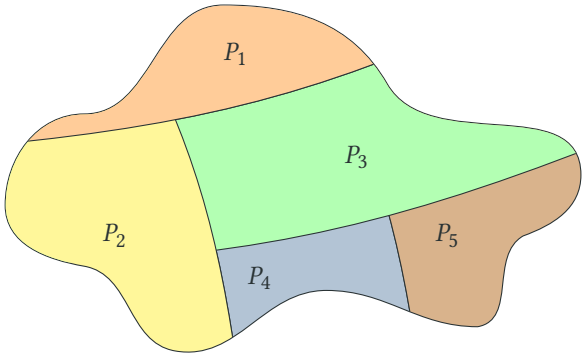


Figura A.2 Uma partição induz uma relação de equivalência e uma relação de equivalência induz uma partição.

Demonstração. Seja $\sim$ uma relação de equivalência em X. Claramente, as classes de equivalência de X por $\sim$ são conjuntos não vazios cuja união é X. Portanto, basta mostrar $Y \cap Y' = \emptyset$ para todo par de classes de equivalência $Y \neq Y'$ de X por $\sim$.

Sejam Y, Y' classes de equivalência de X por $\sim$ que não são disjuntas. Existe um elemento $z \in Y \cap Y'$. Então, $[z] = Y$ e $[z] = Y'$ e então $Y = Y'$. Portanto, se $Y \neq Y'$, $Y \cap Y' = \emptyset$.

Finalmente, a volta é direta. ■

Notação A.20. Dada uma relação de equivalência $\sim$ em X, denotamos por $X/\sim$ o conjunto das classes de equivalência de $\sim$. A **projeção natural** de X em $X/\sim$ é a aplicação dada por

$$\begin{array}{rccc} \pi: & X & \to & X/\sim \\ & x & \mapsto & [x] \end{array}.$$

A.2.2 Relações de Ordem e o Lema de Zorn

Definição A.21. Uma **relação de ordem parcial** num conjunto X é uma relação $\leq$ em X com as seguintes propriedades:

- [a] $x \leq x$ para todo $x \in X$ ($\leq$ é reflexiva);
- [b] se $x \leq y$ e $y \leq x$, então $x = y$ ($\leq$ é antissimétrica);
- [c] se $x \leq y$ e $y \leq z$, então $x \leq z$ ($\leq$ é transitiva).

Neste caso diremos que X é um **conjunto parcialmente ordenado**.
Diremos que $\leq$ é uma relação de ordem total se além de verificar [a], [b] e [c], também verifica

- [d] dados $x, y \in X$, tem-se que $x \leq y$ ou $y \leq x$.

Neste caso diremos que X é um **conjunto totalmente ordenado**.

Exemplos A.22.

a Se X é um conjunto, então a relação de inclusão é uma relação de ordem parcial em $\mathscr{P}(X)$.

b A relação de ordem usual em $\mathbb{R}$ é uma relação de ordem total.

◁

Definição A.23. Seja X um conjunto parcialmente ordenado, e seja $A \subset X$.

a Se existir $a_0 \in A$ tal que $a_0 \leq a$ para todo $a \in A$, diremos que a_0 é o **elemento mínimo** de A. De maneira análoga definimos **elemento máximo**.

b Se existir $a_0 \in A$ tal que $a = a_0$ sempre que $a \in A$ e $a \leq a_0$, diremos que a_0 é um **elemento minimal** de A. De maneira análoga definimos **elemento maximal**.

c Se existir $c \in X$ tal que $c \leq a$ para todo $a \in A$, diremos que A é **limitado inferiormente** e que c é uma **cota inferior** de A. De maneira análoga definimos **conjunto limitado superiormente** e **cota superior**.

d Diremos que A é uma **cadeia** em X se A é totalmente ordenado sob a relação de ordem parcial induzida por X.

e Diremos que A é **bem ordenado** se cada subconjunto não vazio de A possui um elemento mínimo.

Exemplos A.24.

a $\mathbb{N}$, com a ordem usual, é um conjunto bem ordenado.

b $\mathbb{R}$, com a ordem usual, é um conjunto totalmente ordenado que não é bem ordenado: o intervalo aberto (a, b) não possui elemento mínimo. Ser bem ordenado é uma propriedade do par formado pelo conjunto e pela ordem escolhida. Portanto, isso não contradiz o Teorema de Zermelo abaixo, que afirma a existência de alguma boa ordem em $\mathbb{R}$, e não que a ordem usual seja uma boa ordem.

◁

Como observa William Timothy Gowers, o Lema de Zorn é especialmente útil quando um objeto é construído por etapas, nenhuma etapa encerra a construção e nada parece impedir que ela continue.

A seguir apresentaremos dois resultados equivalentes ao Axioma da Escolha.

Lema A.25 (de Zorn).
Seja X um conjunto parcialmente ordenado não vazio tal que cada cadeia em X é limitada superiormente. Então X possui pelo menos um elemento maximal.

Nesse texto utilizaremos o Lema de Zorn para demonstrar que todo espaço vetorial possui base.

Teorema A.26 (Teorema de Zermelo).
Todo conjunto não vazio pode ser bem ordenado.

Teorema A.27.
As seguintes afirmações são equivalentes:

a *O Axioma da escolha.*

b *O Lema de Zorn.*

c *O Teorema de Zermelo.*

A demonstração completa dessas equivalências exige um desenvolvimento mais longo de teoria dos conjuntos e será omitida. Neste livro, o Lema de Zorn será a ferramenta principal derivada do Axioma da Escolha; em particular, ele será usado para provar que todo espaço vetorial possui uma base.

Uma observação bem-humorada atribuída a Jerry Bona resume o contraste intuitivo entre as três formulações: o Axioma da Escolha parece evidente, o princípio da boa ordenação parece falso e o Lema de Zorn parece indecifrável. O teorema mostra que essa diferença é apenas de aparência.

A.3 Cardinalidade

Podemos comparar quantidades sem contar numericamente. Numa sala de cinema, por exemplo, cada ingresso vendido corresponde a uma única poltrona. Se todos os ingressos foram vendidos e cada pessoa possui um ingresso, essas correspondências mostram que há tantas pessoas quanto poltronas. Em linguagem matemática, estamos usando uma bijeção. Se soubéssemos apenas que cada pessoa possui uma poltrona distinta, teríamos uma injeção do conjunto de pessoas no conjunto de poltronas e concluiríamos apenas que não faltam lugares. Essa ideia continua válida para conjuntos infinitos.

Definição A.28. Dizemos que dois conjuntos A e B têm a mesma cardinalidade, denotado por $|A| = |B|$, se existe uma função bijetiva $f\colon A \to B$.

Um conjunto A é dito **finito** se $A = \emptyset$ ou se existe $n \in \mathbb{N}$, $n \geq 1$, e uma bijeção $\{1, \dots, n\} \to A$. Um conjunto é dito **infinito** caso contrário.

Escreveremos $|A| < \infty$ para indicar que A é finito e $|A| = \infty$ para indicar que é infinito. Essas expressões são abreviações, não igualdades entre cardinais e o símbolo ∞.

Exemplo A.29. Os conjuntos $A = \{n \in \mathbb{Z} : 0 \leq n \leq 5\}$ e $B = \{n \in \mathbb{Z} : -5 \leq n \leq 0\}$ tem a mesma cardinalidade porque existe uma função bijetiva $f : A \to B$ dada por $f(n) = -n$.

◁

Teorema A.30.
Existe uma bijeção $f\colon \mathbb{N} \to \mathbb{Z}$. Portanto $|\mathbb{N}| = |\mathbb{Z}|$.

Demonstração. Tomando $\mathbb{N} = \{1, 2, 3, \ldots\}$, defina

$$f(1) = 0, \qquad f(2k) = k, \qquad f(2k+1) = -k \quad (k \geq 1).$$

Cada inteiro aparece exatamente uma vez entre os valores de f; portanto, f é uma bijeção. Se adotarmos a convenção $0 \in \mathbb{N}$, basta deslocar os índices em uma unidade. ■

A igualdade entre as cardinalidades de $\mathbb{N}$ e $\mathbb{Z}$ leva à comparação com outros conjuntos infinitos. Consideremos agora $\mathbb{N}$ e $\mathbb{R}$.

De fato, $|\mathbb{N}| \neq |\mathbb{R}|$. Esse fato foi descoberto por Georg Cantor (1845–1918), que criou um argumento engenhoso para mostrar que não há funções sobrejetivas $f : \mathbb{N} \to \mathbb{R}$. Em particular, não pode haver uma bijeção entre esses conjuntos.

Teorema A.31.
Não existe nenhuma bijeção $f : \mathbb{N} \to \mathbb{R}$. Portanto $|\mathbb{N}| \neq |\mathbb{R}|$.

Demonstração. Suponha que todos os números do intervalo $(0, 1)$ apareçam numa lista $x_1, x_2, \ldots$. Escolha para cada x_n sua expansão decimal que não termina numa sequência infinita de algarismos 9. Construa $y = 0, y_1 y_2 \ldots$ escolhendo $y_n = 1$ se o n-ésimo algarismo decimal de x_n for diferente de 1, e $y_n = 2$ caso contrário. Então $y \in (0, 1)$ e difere de x_n no n-ésimo algarismo para todo n. Logo y não aparece na lista, uma contradição. Portanto $(0, 1)$, e assim $\mathbb{R}$, não pode ser enumerado por $\mathbb{N}$. ■

Esse resultado mostra que conjuntos infinitos podem ter cardinalidades diferentes.

A.3.1 Conjuntos Enumeráveis e não Enumeráveis

Definição A.32. Seja A um conjunto. Dizemos que A é infinito enumerável se $|\mathbb{N}| = |A|$, isto é, se existe uma bijeção $\mathbb{N} \to A$. O conjunto A é não enumerável se A for infinito e $|\mathbb{N}| \neq |A|$, ou seja, se A for infinito e não existir bijeção $\mathbb{N} \to A$.

Assim $\mathbb{Z}$ é infinito enumerável, mas $\mathbb{R}$ é não enumerável.

A relação de ter a mesma cardinalidade é uma relação de equivalência. Contudo, não definiremos a cardinalidade de A como a coleção de todos os conjuntos equipotentes a A: essa coleção é, em geral, uma classe própria, e não um conjunto. Neste texto, a notação $|A| = |B|$ significará sempre que existe uma bijeção $A \to B$; de modo análogo, $|A| \leq |B|$ significará que existe uma injeção $A \to B$. Uma construção formal de cardinais como representantes canônicos exigiria a teoria dos ordinais, que não usaremos.

Definição A.33. A cardinalidade dos números naturais é denotada como $\aleph_0$. Ou seja, $|\mathbb{N}| = \aleph_0$. Assim, qualquer conjunto infinito enumerável tem cardinalidade $\aleph_0$.

Uma consequência simples de ser enumerável é a seguinte:

Proposição A.34.
Um conjunto A é infinito enumerável se, e somente se, seus elementos puderem ser organizados, sem repetição, em uma lista infinita $a_1, a_2, a_3, a_4, \ldots$

Demonstração. Uma bijeção $f : \mathbb{N} \to A$ produz a lista $f(1), f(2), \ldots$. Reciprocamente, uma lista sem repetição que contém todos os elementos de A define a bijeção $f(n) = a_n$. ■

Teorema A.35.
O conjunto $\mathbb{Q}$ dos números racionais é infinito enumerável.

Demonstração. Todo racional pode ser escrito de modo único como p/q, com $p \in \mathbb{Z}$, $q \in \mathbb{N}$, $q \geq 1$ e $\text{mdc}(|p|, q) = 1$. Para enumerá-los, percorra, para cada $s = 1, 2, \ldots$, os pares (p, q) que satisfazem $|p| + q = s$ e conserve apenas os pares coprimos. Em cada etapa há somente um número finito de pares e todo racional aparece em alguma etapa. Isso produz uma lista sem repetição de $\mathbb{Q}$. ■

Corolário A.36.
Dados n conjuntos infinitos enumeráveis $A_1, A_2, A_3, \ldots, A_n$, com $n \geq 2$, o produto cartesiano $A_1 \times A_2 \times A_3 \times \cdots \times A_n$ também é infinito enumerável.

Demonstração. Se $A = \{a_1, a_2, \ldots\}$ e $B = \{b_1, b_2, \ldots\}$, enumeramos $A \times B$ pelas diagonais: primeiro os pares com $i + j = 2$, depois os que têm $i + j = 3$, e assim por diante. Cada par (a_i, b_j) aparece exatamente uma vez. Logo o produto de dois conjuntos enumeráveis é enumerável; o caso de n fatores segue por indução. ■

Teorema A.37.
Se A e B são ambos infinitos enumeráveis, então $A \cup B$ é infinito enumerável.

Demonstração. Intercale enumerações dos dois conjuntos: $a_1, b_1, a_2, b_2, \ldots$. Essa sequência contém todos os elementos de $A \cup B$. Eliminando cada ocorrência que repita um termo anterior, obtemos uma lista sem repetição. Ela continua infinita, pois contém todos os elementos do conjunto infinito A, e portanto enumera $A \cup B$. ■

Um resultado fundamental para nós é que se trocarmos cada ponto de A por um conjunto finito de pontos a cardinalidade não é aumentada.

Lema A.38.
Sejam A e B conjuntos, com A infinito. Se, para cada $x \in A$, temos um subconjunto finito $I_x \subseteq B$, então

$$\left|\bigcup_{x\in A} I_x\right| \leq |A|\,.$$

Demonstração. Admitindo o Axioma da Escolha, fixe uma boa ordem em A e enumere cada conjunto finito I_x. Obtemos uma injeção de $\bigcup_{x\in A} I_x$ em $A\times\mathbb{N}$: para cada elemento, escolhemos o menor índice x cuja enumeração o contém e registramos também sua posição nessa enumeração. O fato padrão de aritmética cardinal $|A \times \mathbb{N}| = |A|$, válido para todo conjunto infinito A sob o Axioma da Escolha, conclui a prova. ■

B

Matrizes e Sistemas Lineares

Matrizes são tabelas de escalares que permitem representar transformações lineares e sistemas de equações. Este apêndice reúne sua notação e as operações elementares. O método de eliminação de Gauss será desenvolvido por meio de transformações que preservam o conjunto de soluções do sistema.

Nesse capítulo

B.1 Matrizes

Sejam m, n dois inteiros positivos. Uma **matriz** $m \times n$ sobre $\mathbb{K}$ é uma tabela de valores $a_{ij} \in \mathbb{K}$, com $1 \leq i \leq m$ e $1 \leq j \leq n$, agrupados em m linhas e n colunas. Ela será representada em uma das duas formas

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{1n} \\ a_{21} & a_{22} & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{mn} \end{bmatrix} \text{ ou } A = \begin{pmatrix} a_{11} & a_{12} & a_{1n} \\ a_{21} & a_{22} & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{mn} \end{pmatrix},$$

onde a_{ij} é a entrada na interseção da i-ésima linha com a j-ésima coluna. Escreveremos $A = [a_{ij}]$ para indicar a matriz cujas entradas são os a_{ij}. Usaremos $A[i, :] = \text{Lin}_i(A)$ para indicar a i-ésima linha de A, e $A[:, j] = \text{Col}_j(A)$ para indicar sua j-ésima coluna.

O conjunto das matrizes $m \times n$ com entradas em $\mathbb{K}$ é denotado por $\mathscr{M}_{m,n}(\mathbb{K})$. Se $A \in \mathscr{M}_{m,n}(\mathbb{K})$, sua entrada (i, j) será indicada por $A_{i,j}$ ou a_{ij}. A matriz identidade $n \times n$ será denotada por I_n. Os elementos do corpo $\mathbb{K}$ serão denominados escalares.

Os elementos de $\mathscr{M}_{m,1}(\mathbb{K})$ serão ditos **vetores coluna** e os de $\mathscr{M}_{1,n}(\mathbb{K})$, **vetores linha**. Eles serão denotados por letras em negrito, como $\mathbf{v}, \vec{a}, \ldots$.

A **diagonal principal** de uma matriz $A \in \mathscr{M}_{m,n}(\mathbb{K})$ é a sequência das entradas $A_{11}, A_{22}, \ldots, A_{kk}$, onde $k = \min\{m, n\}$.

Definição B.1. Diz-se que duas matrizes $A = [a_{ij}], B = [b_{ij}] \in \mathscr{M}_{m,n}(\mathbb{K})$ são iguais se $a_{ij} = b_{ij}$ para todos os índices i e j.

Em outras palavras, duas matrizes são consideradas iguais se tiverem a mesma ordem e suas entradas correspondentes forem iguais.

Definição B.2. Sejam $A = [a_{ij}], B = [b_{ij}] \in \mathscr{M}_{m,n}(\mathbb{K})$ e $k \in \mathbb{K}$.

1. A soma de A e B, denotada por $A + B$, é a matriz $C = [c_{ij}] \in \mathscr{M}_{m,n}(\mathbb{K})$ com $c_{ij} = a_{ij} + b_{ij}$ para todos os i, j.
2. O produto do escalar k por A, denotado por kA, é a matriz $kA = [ka_{ij}]$.

Proposição B.3.
Sejam $A, B, C \in \mathscr{M}_{m,n}(\mathbb{K})$ e $k, l \in \mathbb{K}$. Então

1. $A + B = B + A$ *(comutatividade).*
2. $(A + B) + C = A + (B + C)$ *(associatividade).*
3. $k(lA) = (kl)A$.
4. $(k + l)A = kA + lA$.

Demonstração. 1 Sejam $A = [a_{ij}]$ e $B = [b_{ij}]$. Então, por definição

$$A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}] = [b_{ij} + a_{ij}] = [b_{ij}] + [a_{ij}] = B + A$$

pois a adição no corpo é comutativa. O item 2 segue, entrada a entrada, da associatividade da adição em $\mathbb{K}$. Para os dois últimos itens, em cada posição (i, j) usamos, respectivamente,

$$k(la_{ij}) = (kl)a_{ij} \quad \text{e} \quad (k + l)a_{ij} = ka_{ij} + la_{ij}.$$

■

A matriz nula $0_{m\times n}$ é o elemento neutro da adição em $\mathscr{M}_{m,n}(\mathbb{K})$. A inversa aditiva de $A = [a_{ij}]$ é $-A = [-a_{ij}] = (-1)A$. Em particular, com as operações acima, $\mathscr{M}_{m,n}(\mathbb{K})$ é um espaço vetorial sobre $\mathbb{K}$.

Definição B.4. Sejam $A = [a_{ij}] \in \mathscr{M}_{m,n}(\mathbb{K})$ e $B = [b_{ij}] \in \mathscr{M}_{n,r}(\mathbb{K})$. O produto de A e B, denotado por AB, é a matriz $C = [c_{ij}] \in \mathscr{M}_{m,r}(\mathbb{K})$ definida por

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

Assim, o produto AB só está definido quando o número de colunas de A é igual ao número de linhas de B. Em geral, mesmo quando ambos os produtos fazem sentido, AB e BA não precisam ser iguais.

Proposição B.5.
Para matrizes de tamanhos compatíveis e $c \in \mathbb{K}$, valem

$$(AB)C = A(BC), \qquad A(B+C) = AB + AC, \qquad (A+B)C = AC + BC,$$

$$c(AB) = (cA)B = A(cB), \qquad \mathrm{I}_m A = A = A\,\mathrm{I}_n \quad \text{se } A \in \mathscr{M}_{m,n}(\mathbb{K}).$$

Demonstração. As identidades distributivas e as identidades envolvendo escalares seguem entrada a entrada. Para a associatividade, se $A \in \mathscr{M}_{m,n}(\mathbb{K})$, $B \in \mathscr{M}_{n,r}(\mathbb{K})$ e $C \in \mathscr{M}_{r,s}(\mathbb{K})$, então

$$((AB)C)_{ij} = \sum_{\ell=1}^{r}\sum_{k=1}^{n} a_{ik}b_{k\ell}c_{\ell j} = \sum_{k=1}^{n}\sum_{\ell=1}^{r} a_{ik}b_{k\ell}c_{\ell j} = (A(BC))_{ij}.$$

A afirmação sobre a identidade é imediata da definição do produto. ■

Definição B.6. A **transposta** de $A \in \mathscr{M}_{m,n}(\mathbb{K})$ é a matriz $A^{\mathrm{t}} \in \mathscr{M}_{n,m}(\mathbb{K})$ definida por

$$(A^{\mathrm{t}})_{i,j} = A_{j,i}$$

Uma matriz $A \in \mathscr{M}_{n,n}(\mathbb{K})$ é dita **simétrica** se $A = A^{\mathrm{t}}$ e **antissimétrica** se $A^{\mathrm{t}} = -A$.

Proposição B.7.
Sejam $A, B \in \mathscr{M}_{m,n}(\mathbb{K})$ e $r \in \mathbb{K}$. Então

a $(A^{\mathrm{t}})^{\mathrm{t}} = A$

b $(A+B)^{\mathrm{t}} = A^{\mathrm{t}} + B^{\mathrm{t}}$

c $(rA)^{\mathrm{t}} = rA^{\mathrm{t}}$;

d *se $A \in \mathscr{M}_{m,n}(\mathbb{K})$ e $B \in \mathscr{M}_{n,r}(\mathbb{K})$, então $(AB)^{\mathrm{t}} = B^{\mathrm{t}}A^{\mathrm{t}}$.*

Particionamento e multiplicação de matrizes Seja M uma matriz de tamanho $m \times n$. Se $B \subseteq \{1, \dots, m\}$ e $C \subseteq \{1, \dots, n\}$, a submatriz $M[B, C]$ é obtida mantendo as linhas com índice em B e as colunas com índice em C. Assim, $M[B, C]$ tem tamanho $|B| \times |C|$.

Suponha que $M \in \mathscr{M}_{m,n}(\mathbb{K})$ e $N \in \mathscr{M}_{n,k}(\mathbb{K})$. Sejam $\mathscr{P} = \{B_1, \dots, B_p\}$ uma partição de $\{1, \dots, m\}$, $\mathscr{Q} = \{C_1, \dots, C_q\}$ uma partição de $\{1, \dots, n\}$ e $\mathscr{R} = \{D_1, \dots, D_r\}$ uma partição de $\{1, \dots, k\}$.

A multiplicação pode ser realizada com blocos, desde que suas dimensões sejam compatíveis. Com as partições acima, temos

$$(MN)[B_i, D_j] = \sum_{h=1}^{q} M[B_i, C_h]N[C_h, D_j].$$

Quando as partições em questão contêm apenas blocos de elemento único, essa é precisamente a fórmula usual para multiplicação de matrizes

$$(MN)_{i,j} = \sum_{h=1}^{n} M_{i,h}N_{h,j}.$$

Matrizes de bloco Será conveniente introduzir o dispositivo notacional de uma matriz de blocos. Se $B_{i,j}$ forem matrizes dos tamanhos apropriados, então por uma matriz de blocos

$$M = \begin{bmatrix} B_{1,1} & B_{1,2} & \cdots & B_{1,n} \\ \vdots & \vdots & & \vdots \\ B_{m,1} & B_{m,2} & \cdots & B_{m,n} \end{bmatrix}$$

queremos dizer a matriz cuja submatriz superior esquerda é $B_{1,1}$, e assim por diante. Portanto, os $B_{i,j}$ são submatrizes de M, e não entradas.

Exemplo B.8.

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \left[\begin{array}{cc|ccc} 1 & 2 & 3 & 2 & 1 \\ 3 & 1 & 2 & 2 & 2 \\ 0 & 1 & 2 & 1 & 2 \\ \hline 1 & 2 & 1 & 1 & 2 \\ 2 & 3 & 0 & 2 & 3 \end{array}\right]$$

Nesse caso

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 2 & 2 \\ 2 & 1 & 2 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \quad D = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 3 \end{bmatrix}$$

◁

Podemos, em particular, escrever a matriz como uma justaposição de suas colunas. Para isso, usaremos a notação

$$A = \left[\begin{array}{c|c|c|c} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{array}\right] = [A[:,1] \mid A[:,2] \mid \cdots \mid A[:,n]] = [A\mathbf{e}_1 \mid A\mathbf{e}_2 \mid \cdots \mid A\mathbf{e}_n]$$

sendo $\mathbf{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$ o vetor com 0 em todas as entradas exceto a i-ésima que é 1.

Uma matriz da forma

$$M = \begin{bmatrix} B_{1,1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & B_{m,m} \end{bmatrix}$$

onde cada $B_{i,i}$ é quadrada e 0 é uma submatriz nula, é dita uma matriz diagonal por blocos.

B.2 Sistemas de equações lineares

Definição B.9. Um **sistema de m equações lineares** nas n variáveis $x_1, x_2, \ldots, x_n$ é um conjunto de equações da forma

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{B.1}$$

onde $a_{ij}, b_i \in \mathbb{K}$, para $1 \leq i \leq m$ e $1 \leq j \leq n$. Um sistema linear é dito **homogêneo** se $b_1 = \cdots = b_m = 0$, e **não homogêneo** caso contrário.

Definição B.10. Sejam

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad \vec{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

Então, o sistema (B.1) pode ser escrito de modo compacto como $A\mathbf{x} = \vec{b}$. A matriz A é a **matriz de coeficientes**, e $[A \mid \vec{b}]$ é a **matriz aumentada** do sistema.

Na matriz aumentada $[A \mid \vec{b}]$, a coluna j corresponde à variável x_j, enquanto a última coluna contém os termos independentes. Essa leitura é útil porque as operações efetuadas nas equações podem ser realizadas diretamente nas linhas da matriz aumentada.

Definição B.11. Uma **solução** de $A\mathbf{x} = \vec{b}$ é um vetor $\mathbf{x}_0$ tal que $A\mathbf{x}_0 = \vec{b}$. O conjunto de todas as soluções é o **conjunto solução** do sistema.

Por exemplo, o conjunto solução de $A\mathbf{x} = \vec{b}$, com $A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 4 & 2 \\ 4 & 1 & 2 \end{bmatrix}$ e $\vec{b} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$, é $\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$.

Definição B.12. Um sistema linear é **consistente** se admite ao menos uma solução e **inconsistente** se não admite solução.

Todo sistema homogêneo $A\mathbf{x} = \vec{0}$ é consistente, pois $\vec{0}$ é uma solução. Em contraste, o sistema $x + y = 2$, $2x + 2y = 3$ é inconsistente.

Definição B.13. O **sistema homogêneo associado** a $A\mathbf{x} = \vec{b}$ é o sistema $A\mathbf{x} = \vec{0}$.

Proposição B.14.
Considere o sistema linear homogêneo $A\mathbf{x} = \vec{0}$.

1. *O vetor* $\mathbf{x} = \vec{0}$ *é sempre uma solução, denominada solução trivial.*
2. *Seja* $\mathbf{u} \neq \vec{0}$ *uma solução de* $A\mathbf{x} = \vec{0}$. *Então* $\mathbf{y} = c\mathbf{u}$ *também é solução para todo* $c \in \mathbb{K}$. *Uma solução não nula é dita uma solução não trivial. Se o sistema admitir uma solução não trivial e o corpo for infinito, então* $A\mathbf{x} = \vec{0}$ *possui infinitas soluções.*
3. *Sejam* $\mathbf{u}_1, \dots, \mathbf{u}_k$ *soluções de* $A\mathbf{x} = \vec{0}$. *Então* $\sum_{i=1}^{k} a_i\mathbf{u}_i$ *também é solução, para toda escolha de* $a_i \in \mathbb{K}$, $1 \leq i \leq k$.

Exemplo B.15. Seja $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Então, $\mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ é uma solução não trivial de $A\mathbf{x} = \vec{0}$.

◁

Se $\mathbf{u}$ e $\mathbf{v}$ são soluções de $A\mathbf{x} = \vec{b}$, então $A(\mathbf{u} - \mathbf{v}) = \vec{0}$: duas soluções diferem por um elemento de $\ker A$. Reciprocamente, somar a uma solução qualquer elemento de $\ker A$ produz outra solução. Portanto, se $\mathbf{x}_0$ é uma solução particular, o conjunto solução é o espaço afim

$$\{\mathbf{x}_0 + \mathbf{x}_h \mid \mathbf{x}_h \in \ker A\}.$$

B.3 Operações elementares

A eliminação de Gauss consiste em substituir um sistema por outros, cada vez mais simples, sem modificar suas soluções. As transformações permitidas são as operações elementares por linhas.

Definição B.16. Seja $A \in \mathscr{M}_{m,n}(\mathbb{K})$. As **operações elementares por linhas** são:

1. E_{ij}: trocar as linhas i e j;
2. $E_k(c)$, com $c \neq 0$: multiplicar a linha k por c;
3. $E_{ij}(c)$: substituir a linha i por $A[i,\ :] + cA[j,\ :]$, onde $i \neq j$ e $c \in \mathbb{K}$.

Cada uma dessas operações pode ser desfeita por outra operação do mesmo tipo: E_{ij} é sua própria inversa, enquanto as inversas de $E_k(c)$ e $E_{ij}(c)$ são, respectivamente, $E_k(c^{-1})$ e $E_{ij}(-c)$.

Definição B.17. Duas matrizes são **equivalentes por linhas** se uma pode ser obtida da outra mediante uma sequência finita de operações elementares por linhas. Dois sistemas são equivalentes por linhas quando suas matrizes aumentadas o são.

Proposição B.18.
Sistemas lineares equivalentes por linhas possuem o mesmo conjunto solução.

Demonstração. Basta considerar uma operação. Trocar duas equações ou multiplicar uma delas por um escalar não nulo preserva suas soluções. No terceiro tipo, substituímos a equação i por sua soma com c vezes a equação j. Todo vetor que satisfaz as equações originais satisfaz a nova; a recíproca segue aplicando $E_{ij}(-c)$. Uma sequência finita dessas operações preserva, portanto, o conjunto solução. ■

Definição B.19. Uma matriz $E \in \mathscr{M}_{n,n}(\mathbb{K})$ é **elementar** se for obtida aplicando uma operação elementar por linhas à identidade I_n. As três famílias são

1. $E_{ij} = \mathrm{I}_n - \mathbf{e}_i\mathbf{e}_i^{\mathsf{t}} - \mathbf{e}_j\mathbf{e}_j^{\mathsf{t}} + \mathbf{e}_i\mathbf{e}_j^{\mathsf{t}} + \mathbf{e}_j\mathbf{e}_i^{\mathsf{t}}$;
2. $E_k(c) = \mathrm{I}_n + (c-1)\mathbf{e}_k\mathbf{e}_k^{\mathsf{t}}$, para $c \neq 0$;
3. $E_{ij}(c) = \mathrm{I}_n + c\mathbf{e}_i\mathbf{e}_j^{\mathsf{t}}$, para $i \neq j$ e $c \in \mathbb{K}$.

Multiplicar A à esquerda por uma matriz elementar produz exatamente o efeito da operação por linhas correspondente.

Proposição B.20.
Se as operações que levam A a uma matriz equivalente por linhas B correspondem, nessa ordem, a $E_1, \dots, E_k$, então

$$B = E_k \cdots E_2 E_1 A.$$

Demonstração. Após a primeira operação obtemos E_1A; após a segunda, E_2E_1A; e, prosseguindo, chegamos a $E_k \cdots E_1 A$. ■

Definição B.21. Em cada linha não nula de uma matriz, a entrada não nula mais à esquerda é o **pivô** da linha. Uma coluna que contém um pivô é uma **coluna pivotal**.

Por exemplo, as entradas 3 e 2 são os pivôs de

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

logo, as colunas 2 e 3 são pivotais.

Definição B.22. Uma matriz está na **forma escalonada por linhas** se:

1. as linhas nulas estão abaixo das linhas não nulas;
2. o pivô de cada linha não nula, a partir da segunda, está à direita do pivô da linha anterior;
3. todas as entradas abaixo de cada pivô são nulas.

Ela está na **forma escalonada reduzida por linhas** se, além disso, cada pivô é 1 e é a única entrada não nula de sua coluna.

Exemplo B.23 — Eliminação de Gauss. Considere o sistema

$$\begin{aligned} x + 2y - z &= 3,\\ 2x + 5y + z &= 7,\\ -x - y + 2z &= -2. \end{aligned}$$

Escalonamos sua matriz aumentada, registrando ao lado de cada seta as operações realizadas:

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 2 & 5 & 1 & 7 \\ -1 & -1 & 2 & -2 \end{array}\right] \xrightarrow[L_3 \leftarrow L_3 + L_1,]{L_2 \leftarrow L_2 - 2L_1,} \left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & 1 & 3 & 1 \\ 0 & 1 & 1 & 1 \end{array}\right]$$

$$\xrightarrow{L_3 \leftarrow L_3 - L_2,} \left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & -2 & 0 \end{array}\right]$$

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

Na última passagem dividimos a terceira linha por -2 e eliminamos as entradas acima dos pivôs. A forma reduzida mostra imediatamente que a única solução é $(x, y, z) = (1, 1, 0)$.

◁

O escalonamento também torna visíveis as outras possibilidades. Uma linha $[0 \ \cdots \ 0 \mid c]$, com $c \neq 0$, representa a equação impossível $0 = c$ e revela que o sistema é inconsistente. Se não há uma linha desse tipo e alguma coluna de variável não contém pivô, a variável correspondente é livre e o sistema possui mais de uma solução. Por exemplo,

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

tem $z = t$ livre e soluções $(x, y, z) = (1 + 5t, 1 - 2t, t)$, com $t \in \mathbb{K}$.

Teorema B.24.
Para a equivalência por linhas em $\mathscr{M}_{m,n}(\mathbb{K})$, valem as propriedades:

- **a** *A equivalência por linhas é uma relação de equivalência.*
- **b** *Cada matriz A é equivalente a uma única matriz R na forma escalonada reduzida por linhas.*
- **c** *Uma matriz quadrada A é invertível se, e somente se, sua forma reduzida é I_n; equivalentemente, A é produto de matrizes elementares.*

Se $E_1, \ldots, E_k$ são as matrizes elementares das operações que reduzem A a R, então $R = E_k \cdots E_1 A$. Como as matrizes elementares são invertíveis e suas inversas também são elementares, essa identidade pode ser revertida.

O número de pivôs da forma escalonada de A é o **posto** de A. Como as operações por linhas preservam as relações entre as equações, obtemos o critério

$$A\mathbf{x} = \vec{b} \text{ é consistente} \iff \operatorname{posto} A = \operatorname{posto}[A \mid \vec{b}].$$

Quando o sistema é consistente, a solução é única exatamente quando todas as n colunas de variáveis são pivotais, isto é, quando posto $A = n$.

Definição B.25. Seja $A = [a_{ij}] \in \mathscr{M}_{n,n}(\mathbb{K})$.

1. A é **diagonal** se $a_{ij} = 0$ para $i \neq j$; nesse caso, escrevemos $A = \operatorname{diag}(a_{11}, \ldots, a_{nn})$.
2. A é **escalar** se $A = \alpha\, \mathrm{I}_n$ para algum $\alpha \in \mathbb{K}$.
3. A é **triangular superior** se $a_{ij} = 0$ para $i > j$, e **triangular inferior** se $a_{ij} = 0$ para $i < j$.

Exemplo B.26. As matrizes

$$\begin{bmatrix} 0 & 1 & 4 \\ 0 & 3 & -1 \\ 0 & 0 & -2 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

são, respectivamente, triangular superior e triangular inferior. Toda matriz diagonal — em particular, 0 e I_n — é dos dois tipos.

◁

O escalonamento permite determinar a consistência e a unicidade das soluções pela posição dos pivôs. Também fornece métodos para calcular bases, postos e inversas de matrizes.

Exercícios

Ex. B.1 — Calcule AB e BA, quando definidos, para

$$A = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 3 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & 1 \\ 0 & -1 \\ 4 & 2 \end{bmatrix}.$$

Explique por que os dois produtos têm tamanhos diferentes.

Ex. B.2 — Se $A = [A[:,1] \mid \cdots \mid A[:,n]]$ e $\mathbf{x} = (x_1, \ldots, x_n)^{\mathrm{t}}$, prove que

$$A\mathbf{x} = x_1 A[:,1] + \cdots + x_n A[:,n].$$

Interprete, a partir dessa identidade, o conjunto de soluções de $A\mathbf{x} = \vec{b}$.

Ex. B.3 — Escalone a matriz aumentada e resolva o sistema

$$\begin{aligned} x - y + 2z &= 4, \\ 2x + y + z &= 1, \\ 3x + 3z &= 5. \end{aligned}$$

Indique os pivôs em cada etapa.

Ex. B.4 — Determine, conforme o valor de $a \in \mathbb{R}$, se o sistema

$$x + y + z = 1, \qquad 2x + 2y + 2z = 2, \qquad x + y + az = 0$$

é inconsistente ou possui infinitas soluções. Parametrize o conjunto solução sempre que ele existir.

Ex. B.5 — Encontre a forma escalonada reduzida de

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

e determine uma base para $\ker A$.

Ex. B.6 — Escreva as matrizes elementares que realizam, nessa ordem, as operações $L_2 \leftarrow L_2 - 3L_1$, $L_1 \leftrightarrow L_3$ e $L_3 \leftarrow 2L_3$ em uma matriz com três linhas. Verifique diretamente a ordem do produto que representa a sequência das três operações.

Ex. B.7 — Prove que operações elementares por linhas não alteram o posto de uma matriz. Conclua que matrizes equivalentes por linhas têm o mesmo posto.

Ex. B.8 — Se $A \in \mathscr{M}_{n,n}(\mathbb{K})$, descreva um procedimento baseado em operações por linhas que decide se A é invertível e, quando for, calcula A^{-1}. Aplique-o a $A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$.

Bibliografia

[1] Tom M Apostol. *Calculus, volume II: multi-variable calculus and linear algebra, with applications to differential equations and probability*. John Wiley & Sons, 1969.

[2] Sheldon Jay Axler. *Linear algebra done right*, volume 2. Springer, 1997.

[3] Thomas Scott Blyth and E F Robertson. *Further linear algebra*. Springer Science & Business Media, 2013.

[4] William Clough Brown and William C Brown. *A second course in linear algebra*. Wiley, 1988.

[5] Ruel Vance Churchill et al. *Fourier series and boundary value problems*. McGraw-Hill, 1941.

[6] Flávio Ulhoa Coelho and Mary Lílian Lourenco. *Curso de Álgebra Linear*. Edusp, 2001.

[7] Morton L Curtis. *Abstract linear algebra*. Springer Science & Business Media, 2012.

[8] Douglas R Farenick. *Algebras of linear transformations*. Springer Science & Business Media, 2012.

[9] Jonathan S Golan. *The Linear Algebra a Beginning Graduate Student Ought to Know*. Springer, 2007.

[10] Frederick P Greenleaf and Sophie Marques. *Linear Algebra I*, volume 30. American Mathematical Soc., 2020.

[11] Frederick P Greenleaf and Sophie Marques. *Linear Algebra II*, volume 30. American Mathematical Soc., 2020.

[12] Werner Greub. *Multilinear Algebra*. Springer-Verlag, 1978.

[13] Werner H Greub. *Linear algebra*, volume 23. Springer Science & Business Media, 2012.

[14] Paul R Halmos. *Linear algebra problem book*. Number 16 in Dolciani Mathematical Expositions. Mathematical Association of America, 1995.

[15] Paul R Halmos. *Finite-dimensional vector spaces*. Courier Dover Publications, 2017.

[16] Kenneth Hoffman and Ray Kunze. *Linear Algebra*. Prentice-Hall, 1971.

[17] Leslie Hogben. *Handbook of linear algebra*. CRC press, 2006.

[18] Nathan Jacobson. *Lectures in Abstract Algebra: II. Linear Algebra*, volume 31. Springer Science & Business Media, 2013.

[19] Anthony W Knapp. *Basic algebra*. Springer Science & Business Media, 2007.

[20] Alexandra I Kostrikin. *Exercises in Algebra: A Collection of Exercises, in Algebra, Linear Algebra and Geometry*, volume 6. CRC Press, 1996.

[21] Alexei I Kostrikin and Yu I Manin. *Linear Algebra and Geometry*. Gordon and Breach Science Publishers, 1989.

[22] Erwin Kreyszig. *Introductory functional analysis with applications*, volume 1. Wiley, 1978.

[23] Carlos S Kubrusly. *Elements of operator theory*. Springer, 2011.

[24] Ramji Lal. *Algebra 2: Linear Algebra, Galois Theory, Representation Theory, Group Extensions and Schur Multiplier*. Springer, 2017.

[25] Elon Lages Lima. *Álgebra Linear*. IMPA, 2006.

[26] George Mackiw. A note on the equality of the column, and row rank of a matrix. *Mathematics Magazine*, 68(4):285, 1995.

[27] Gregory H Moore. The axiomatization of linear algebra: 1875-1940. *Historia Mathematica*, 22(3):262–303, 1995.

[28] M Thamban Nair and Arindama Singh. *Linear Algebra*. Springer, 2018.

[29] James M Ortega. *Matrix theory: A second course*. Springer Science & Business Media, 2013.

[30] Peter Petersen. *Linear Algebras*. Springer Science & Business Media, 2012.

[31] Steven Roman, S Axler, and FW Gehring. *Advanced linear algebra*, volume 3. Springer, 2005.

[32] Harvey E Rose. *Linear algebra: a pure mathematical approach*. Springer Science & Business Media, 2002.

[33] Igor R Shafarevich and Alexey O Remizov. *Linear algebra and geometry*. Springer Science & Business Media, 2012.

[34] G Shilov. *Linear Algebra*. Dover Books on Advanced Mathematics, 1977.

[35] Kazimierz Szymiczek. *Bilinear algebra: An introduction to the algebraic theory of quadratic forms*, volume 7. CRC Press, 1997.

[36] Steven H Weintraub. *Jordan canonical form: theory and practice*. Morgan & Claypool Publishers, 2009.

[37] Steven H Weintraub. *A guide to advanced linear algebra*. Number 44 in Dolciani Mathematical Expositions. Mathematical Association of America, 2011.

[38] Mark Wildon. A short proof of the existence of jordan normal form, 2011. Nota expositiva disponível na página do autor.

[39] Yisong Yang. *A concise text on advanced linear algebra*. Cambridge University Press, 2014.

Índice

Espaços Vetoriais