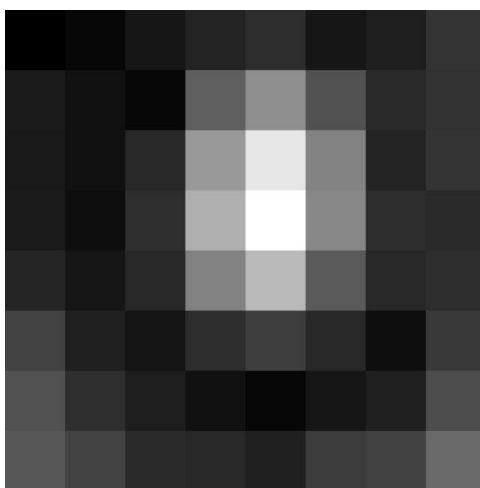


# ÁLGEBRA LINEAR ESSENCIAL



Ronaldo Freire de Lima



# Sumário

<b>1</b>	<b>Matrizes</b>	<b>1</b>
1.1	Matrizes e sua Álgebra . . . . .	1
1.2	Resolução de Sistemas Lineares . . . . .	6
1.3	Inversão e Semelhança de Matrizes . . . . .	14
<b>2</b>	<b>Espaços Vetoriais</b>	<b>19</b>
2.1	Espaços e Subespaços Vetoriais . . . . .	19
2.2	Dependência Linear . . . . .	26
2.3	Bases – Dimensão . . . . .	31
2.4	O Posto de uma Matriz . . . . .	38
2.5	Somas de Subespaços . . . . .	42
<b>3</b>	<b>Transformações Lineares</b>	<b>49</b>
3.1	Linearidade de Aplicações . . . . .	50
3.2	Transformações Lineares e Matrizes . . . . .	54
3.3	Mudança de Coordenadas . . . . .	58
3.4	Núcleo e Imagem . . . . .	60
3.5	Operações com Transformações Lineares . . . . .	65
3.6	Isomorfismos . . . . .	68
3.7	Aplicação aos Sistemas Lineares . . . . .	71
<b>4</b>	<b>Determinantes</b>	<b>79</b>
4.1	Determinantes em $\mathcal{M}(2)$ . . . . .	79
4.2	Determinantes em $\mathcal{M}(n)$ . . . . .	82
4.3	Propriedade Fundamental – Aplicações . . . . .	92

<b>5</b>	<b>Produto Interno</b>	<b>107</b>
5.1	Geometria Plana e Produto Escalar . . . . .	108
5.2	Espaços Vetoriais com Produto Interno . . . . .	110
5.3	Ortogonalidade . . . . .	115
5.4	Volume em Espaços Euclidianos . . . . .	122
<b>6</b>	<b>Decomposição de Operadores</b>	<b>131</b>
6.1	Considerações Iniciais . . . . .	132
6.2	Autovalores e Autovetores . . . . .	137
6.3	Polinômio Mínimo . . . . .	145
6.4	O Teorema da Decomposição Primária . . . . .	150
<b>7</b>	<b>Operadores em Espaços com Produto Interno</b>	<b>161</b>
7.1	A Adjunta de uma Transformação Linear . . . . .	162
7.2	Operadores Ortogonais . . . . .	165
7.3	Operadores Autoadjuntos . . . . .	170
7.4	Formas Quadráticas . . . . .	172
<b>Apêndices</b>		
<b>A</b>	<b>Decomposição Racional e o Teorema de Cayley–Hamilton</b>	<b>185</b>
<b>Referências</b>		<b>193</b>
<b>Índice Remissivo</b>		<b>195</b>

# *Prefácio*

Neste texto, destinado primordialmente a estudantes de graduação em matemática, nos propomos apresentar a teoria da Álgebra Linear dos espaços vetoriais reais de dimensão finita. Nosso intuito é o de abordar os tópicos dessa teoria, os quais são indispensáveis ao estudo de outras que lhe são subsequentes na grade curricular, tais como Análise (de várias variáveis), Equações Diferenciais e Geometria Diferencial. Sendo assim, nossas questões estarão ligadas aos aspectos estruturais dos espaços vetoriais, incluindo a classificação desses objetos, bem como de suas transformações.

Uma das características do texto, em contraste com seus congêneres, é a concisão, a qual deve-se à nossa intenção de apresentar a parte mais essencial da Álgebra Linear (daí, o título) de uma forma simples e direta; evitando, porém, a superficialidade. Assim, fomos impelidos a fazer uma abordagem um tanto alternativa, sobretudo no que diz respeito à apresentação de alguns teoremas fundamentais — notadamente, o Teorema da Decomposição Primária, o Teorema Espectral, e o Teorema da Decomposição Racional.

A disposição do conteúdo ao longo dos capítulos se dá da seguinte forma: No primeiro, consideramos brevemente as matrizes e o método de eliminação de Gauss para resolução de sistemas lineares. Nos Capítulos 2 e 3, introduzimos os espaços vetoriais e as transformações entre os mesmos, ditas lineares. Em seguida, no Capítulo 4, consideramos formalmente o determinante, enquanto função de matrizes, e no Capítulo 5, introduzimos o conceito de produto interno, enfatizando o seu caráter geométrico.

A partir do Capítulo 6, nos voltamos aos operadores lineares, com enfoque no método de decomposição. Ali, discutimos sobre o conceito de subespaço invariante, no qual se baseia esse método, bem como estabelecemos o Teo-

rema da Decomposição Primária, o qual desempenha um papel fundamental no desenvolvimento da teoria.

O Capítulo final trata dos operadores em espaços vetoriais munidos de produto interno, onde consideramos as isometrias lineares, isto é, os operadores ortogonais, assim como os operadores autoadjuntos, os quais constituem o objeto do célebre Teorema Espectral. Em seguida, abordamos as formas quadráticas à luz de suas relações com operadores autoadjuntos.

O texto se encerra no Apêndice A, onde apresentamos uma engenhosa demonstração do Teorema da Decomposição Racional, devida a H. G. Jacob, o qual aplicamos para obter uma versão generalizada do Teorema de Cayley-Hamilton.

Quanto aos pré-requisitos, cabe-nos mencionar que, para uma efetiva leitura do texto, faz-se necessária uma certa familiaridade com a Geometria Analítica, com os rudimentos da teoria de polinômios, e com certos aspectos da linguagem matemática formal, tais como o conceito de relação de equivalência e o método de indução matemática.

Por fim, agradecemos calorosamente aos colegas que leram versões preliminares do texto e contribuíram com valiosas sugestões.

Natal, Setembro de 2020

RFL

# 1

---

## *Matrizes*

Matrizes são, fundamentalmente, tabelas numéricas sobre as quais se definem certas operações algébricas. Tais operações concedem ao conjunto das matrizes (de um mesmo tipo) a estrutura de espaço vetorial, que, por sua vez, constitui o conceito fundamental da Álgebra Linear. Além disso, as matrizes se aplicam ao estudo dos sistemas lineares (conforme veremos neste capítulo), bem como desempenham um papel decisivo no estudo das transformações lineares, as quais são justamente as funções estudadas na Álgebra Linear. Desta forma, esses objetos permeiam toda a teoria que estudaremos, razão pela qual os introduziremos aqui, um tanto brevemente. Nosso propósito, na verdade, é o de considerar apenas os aspectos mais relevantes acerca de matrizes, os quais serão indispensáveis à nossa abordagem aos temas dos capítulos posteriores.

### **1.1 Matrizes e sua Álgebra**

Consideraremos conhecido e estabelecido o conjunto dos números reais,  $\mathbb{R}$ , juntamente com sua estrutura de corpo, a qual é formada por esse conjunto e suas operações usuais de adição e multiplicação. Essas operações, como se sabe, têm as propriedades fundamentais de comutatividade, associatividade e existência de elemento neutro e inverso. Além disso, a multiplicação é distributiva com respeito a adição.

Sejam  $m$  e  $n$  dois números inteiros positivos. Uma *matriz (real)* do tipo  $m \times n$  (lê-se:  $m$  por  $n$ ) é uma função  $A$  que, a cada par de inteiros  $(i, j) \in$

$\{1, \dots, m\} \times \{1, \dots, n\}$ , associa um número real  $a_{ij} = A(i, j)$ .

Considerando-se os nossos propósitos, uma matriz será melhor representada como uma tabela disposta em  $m$  linhas e  $n$  colunas, em que a  $i$ -ésima linha é formada pelos reais  $a_{i1}, \dots, a_{in}$ , e a  $j$ -ésima coluna é formada pelos reais  $a_{1j}, \dots, a_{mj}$ . Assim, escreveremos:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}.$$

Os reais  $a_{ij}$  que compõem uma matriz  $A$  são ditos as suas *entradas*. Indicaremos, também, uma matriz  $A$  do tipo  $m \times n$  e com entradas  $a_{ij}$  por

$$A = (a_{ij})_{m \times n}$$

ou, simplesmente,  $A = (a_{ij})$ . Quando  $m = n$ , a matriz  $A$  é dita *quadrada* de *ordem*  $n$ . O conjunto formado por suas entradas  $a_{ii}$  é dito, então, a *diagonal* de  $A$ . Uma matriz cujas entradas são todas iguais a zero chama-se *nula* e será denotada por  $\mathbf{0}$  (independentemente de seu tipo).

A matriz  $A = (a_{ij})_{3 \times 2}$ , em que  $a_{ij} = i + j$ , por exemplo, é:

$$A = \begin{bmatrix} 2 & 3 \\ 3 & 4 \\ 4 & 5 \end{bmatrix}.$$

Dadas matrizes do mesmo tipo,  $A = (a_{ij})_{m \times n}$  e  $B = (b_{ij})_{m \times n}$ , juntamente com um número real  $\lambda$ , definem-se a *adição*  $A + B$  e o *produto por escalar*  $\lambda A$  por:

$$A + B = (a_{ij} + b_{ij})_{m \times n} \quad \text{e} \quad \lambda A = (\lambda a_{ij})_{m \times n},$$

ou seja,

$$A + B = \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix} \quad \text{e} \quad \lambda A = \begin{bmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{bmatrix}.$$



Quando  $C = A + B$ , diz-se que  $C$  é a *soma* de  $A$  e  $B$ . Nesse contexto, os números reais são também chamados de *escalares* e, como no caso da multiplicação entre números reais, adota-se a notação  $-A$  para o produto  $(-1)A$ .

Considerando-se matrizes  $A, B$  e  $C$ , todas do mesmo tipo, e  $\lambda, \mu \in \mathbb{R}$ , verificam-se facilmente as seguintes igualdades:

- $A + B = B + A$  (comutatividade da adição);
- $A + (B + C) = (A + B) + C$  (associatividade da adição);
- $A + \mathbf{0} = A$  (elemento neutro da adição);
- $A + (-A) = \mathbf{0}$  (existência do *oposto*);
- $\lambda(\mu A) = (\lambda\mu)A$  (associatividade da multiplicação por escalar);
- $1.A = A$  (elemento neutro da multiplicação por escalar).

Além disso, a adição e multiplicação por escalar são distributivas, uma com respeito à outra, isto é:

- $\lambda(A + B) = \lambda A + \lambda B$ ;
- $(\lambda + \mu)A = \lambda A + \mu A$ .

Passemos, agora, à multiplicação de matrizes. Para tanto, consideremos  $A = (a_{ij})_{m \times n}$  e  $B = (b_{ij})_{n \times p}$ , e observemos que o número de colunas de  $A$  coincide com o número de linhas de  $B$ . Nesse caso, definimos o *produto*  $AB$  como a matriz  $P = (p_{ij})_{m \times p}$ , cujas entradas são:

$$p_{ij} = \sum_{k=1}^n a_{ik}b_{kj} = a_{i1}b_{1j} + \cdots + a_{in}b_{nj}.$$

Por exemplo:

$$\begin{bmatrix} 2 & 3 \\ 3 & 4 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 0 & -1 & 3 \\ 3 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 2.0 + 3.3 & 2.(-1) + 3.2 & 2.3 + 3.1 \\ 3.0 + 4.3 & 3.(-1) + 4.2 & 3.3 + 4.1 \\ 4.0 + 5.3 & 4.(-1) + 5.2 & 4.3 + 5.1 \end{bmatrix} = \begin{bmatrix} 9 & 4 & 9 \\ 12 & 5 & 13 \\ 15 & 6 & 17 \end{bmatrix}.$$

Cabe aqui um comentário acerca das operações com matrizes, sobretudo pela multiplicação, a qual define-se de uma forma um tanto inesperada, digamos. Conforme mencionamos à introdução, há uma associação natural entre matrizes e certos tipos de funções entre espaços vetoriais, chamadas de transformações lineares, sobre as quais são definidas operações de adição, multiplicação e multiplicação por escalar. Veremos, então, no Capítulo 3, que as operações sobre matrizes são definidas de modo que correspondam às respectivas operações com transformações lineares. No caso do produto de matrizes, há a justificativa adicional de que ele se adequa ao estudo dos sistemas lineares, conforme constataremos na próxima seção.

A multiplicação de matrizes é associativa, bem como distributiva (à direita e à esquerda) com respeito à adição. Mais especificamente, dadas matrizes  $A = (a_{ij})_{m \times n}$ ,  $B = (b_{ij})_{n \times p}$ ,  $C = (c_{ij})_{p \times q}$ ,  $D = (d_{ij})_{n \times p}$  e  $E = (e_{ij})_{m \times n}$ , são válidas as seguintes igualdades:

- $A(BC) = (AB)C$ ;
- $A(B + D) = AB + AD$ ;
- $(A + E)B = AB + EB$ .

Provemos, pois, a primeira delas. Com esse propósito, escrevamos:

- $AB = (a'_{ij})_{m \times p}$  e  $BC = (b'_{ij})_{n \times q}$ ;
- $A(BC) = (m_{ij})_{m \times q}$  e  $(AB)C = (n_{ij})_{m \times q}$ .

Assim, temos que:

$$\begin{aligned} m_{ij} &= \sum_{k=1}^n a_{ik} b'_{kj} = \sum_{k=1}^n \left( a_{ik} \sum_{l=1}^p b_{kl} c_{lj} \right) = \sum_{k=1}^p \left( \sum_{l=1}^n a_{il} b_{lk} \right) c_{kj} \\ &= \sum_{k=1}^p a'_{ik} c_{kj} = n_{ij}, \end{aligned}$$

donde se conclui que  $A(BC) = (AB)C$ .

**Observação 1.1.** Alertamos para o fato de que, quando duas matrizes  $A$  e  $B$  são quadradas e de mesma ordem, os dois produtos  $AB$  e  $BA$  estão bem definidos. No entanto, de modo geral, eles não são iguais, isto é, *o produto de matrizes quadradas de mesma ordem não é comutativo*. Quando, excepcionalmente, ocorre a igualdade  $AB = BA$ , dizemos que  $A$  e  $B$  *comutam*.

A associatividade do produto de matrizes, entretanto, nos permite definir, sem ambiguidade, a  $n$ -ésima potência de uma matriz  $A$  através da igualdade:

$$A^n = A.A \dots A \quad (n \text{ fatores}).$$

Vejamos, agora, alguns tipos especiais de matrizes quadradas.

**Exemplo 1.1** (MATRIZES TRIANGULARES E DIAGONAIS). Uma matriz quadrada  $A = (a_{ij})_{n \times n}$  é dita *triangular*, quando ocorre uma das possibilidades:

$$a_{ij} = 0 \quad \forall i > j \quad \text{ou} \quad a_{ij} = 0 \quad \forall i < j.$$

No primeiro caso, ela é dita *triangular superior* e, no segundo, *triangular inferior*. Uma matriz que é triangular superior e inferior é dita *diagonal*. Note que duas matrizes diagonais de mesma ordem sempre comutam.

**Exemplo 1.2** (MATRIZ IDENTIDADE). A matriz diagonal  $n \times n$  cujas entradas não nulas são todas iguais a 1 é chamada de *matriz identidade* de ordem  $n$ , a qual denota-se por  $I_n$  ou, simplesmente, por  $I$ . Tal matriz constitui o elemento neutro da multiplicação de matrizes quadradas, pois, para toda matriz quadrada  $A$  de ordem  $n$ , vale a igualdade:

$$AI = IA = A.$$

**Exemplo 1.3** (MATRIZES SIMÉTRICAS). Diz-se que uma matriz quadrada de ordem  $n$ ,  $A = (a_{ij})$ , é *simétrica*, quando  $a_{ij} = a_{ji} \quad \forall 1 \leq i, j \leq n$ . Evidentemente, toda matriz diagonal é simétrica. Um dos resultados mais importantes da Álgebra Linear, o Teorema Espectral, estabelece uma espécie de recíproca dessa propriedade, uma vez que, segundo esse teorema, toda matriz simétrica é *diagonalizável*. Voltaremos a essa questão ao final deste capítulo, onde tornaremos isso mais claro.

Dada uma matriz  $A = (a_{ij})_{m \times n}$ , a *transposta* de  $A$ , por definição, é a matriz  $A^* := (a_{ij}^*)_{n \times m}$ , em que  $a_{ij}^* = a_{ji}$ . Dito de forma simples, a transposta de uma matriz  $m \times n$  é a matriz  $n \times m$  obtida desta permutando-se suas linhas pelas colunas correspondentes.

Por exemplo, a transposta da matriz  $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$  é a matriz

$$A^* = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

Segue-se diretamente da definição de transposição que, para quaisquer matrizes  $A, B \in \mathcal{M}(m, n)$  e  $\lambda \in \mathbb{R}$ , tem-se

$$(A + \lambda B)^* = A^* + \lambda B^* \quad \text{e} \quad (A^*)^* = A. \quad (1.1)$$

Observe também que uma matriz quadrada é simétrica se, e somente se, é igual à sua transposta.

## 1.2 Resolução de Sistemas Lineares

Dados dois números naturais  $m, n$ ,  $mn$  números reais  $a_{ij}$ ,  $1 \leq i \leq m$ ,  $1 \leq j \leq n$ , e  $b_1, \dots, b_m \in \mathbb{R}$ , chama-se o conjunto de equações

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \quad (1.2)$$

de *sistema linear* nas variáveis  $x_1, \dots, x_n$ . Uma *solução* do sistema (1.2) é uma  $n$ -upla de números reais,  $(\lambda_1, \dots, \lambda_n)$ , que satisfaz cada uma das suas  $m$

equações, isto é,

$$\begin{cases} a_{11}\lambda_1 + a_{12}\lambda_2 + \cdots + a_{1n}\lambda_n = b_1 \\ a_{21}\lambda_1 + a_{22}\lambda_2 + \cdots + a_{2n}\lambda_n = b_2 \\ \vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \\ a_{m1}\lambda_1 + a_{m2}\lambda_2 + \cdots + a_{mn}\lambda_n = b_m \end{cases}.$$

O conjunto solução de um sistema linear é o conjunto formado por todas as suas soluções. Note que, considerando-se a matriz  $A = (a_{ij})_{m \times n}$ , bem como

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

tem-se que o sistema (1.2) equivale à seguinte equação matricial

$$AX = B. \tag{1.3}$$

Nesse caso, dizemos que  $A$  é a *matriz do sistema* (1.2). Quando  $B = \mathbf{0}$ , o sistema (1.2) é dito *homogêneo*. Observe que todo sistema homogêneo admite a solução  $X = \mathbf{0}$ , dita *trivial*.

Neste momento, o qual anunciamos na seção anterior, devemos observar que a igualdade (1.3) justifica a definição do produto de matrizes, uma vez que ele nos permite reduzir um sistema linear, com inúmeras equações, a uma única equação matricial. Como consequência, o estudo dos sistemas lineares simplifica-se consideravelmente, pois finda por resumir-se ao estudo das matrizes.

Dado  $n \in \mathbb{R}$ , relembremos que o espaço  $\mathbb{R}^n$  define-se como o produto cartesiano de  $n$  cópias de  $\mathbb{R}$ . Cada elemento de  $\mathbb{R}^n$  é, portanto, uma  $n$ -upla de números reais  $(x_1, \dots, x_n)$ , a qual é dita um *vetor* de  $\mathbb{R}^n$ . Para cada matriz  $A = (a_{ij})_{m \times n}$ , associamos à sua  $i$ -ésima linha e à sua  $j$ -ésima coluna, respectivamente, os seguintes vetores

$$\ell_i = (a_{i1}, a_{i2}, \dots, a_{in}) \in \mathbb{R}^n \quad \text{e} \quad v_j = (a_{1j}, a_{2j}, \dots, a_{mj}) \in \mathbb{R}^m.$$

Cada  $\ell_i$  é dito um *vetor linha* de  $A$ , e cada  $v_j$  um *vetor coluna* de  $A$ . Assim, representamos a matriz  $A$  das seguintes formas:

$$A = \begin{bmatrix} \ell_1 \\ \ell_2 \\ \vdots \\ \ell_m \end{bmatrix} = [ v_1 \quad v_2 \quad \cdots \quad v_n ].$$

Por exemplo,  $(1, 3, -1)$ ,  $(0, 1, 2) \in \mathbb{R}^3$  são os vetores linha da matriz

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

enquanto  $(1, 0)$ ,  $(3, 1)$ ,  $(-1, 2) \in \mathbb{R}^2$  são seus vetores coluna.

**Definição 1.1** (SISTEMAS LINEARES EQUIVALENTES). Dois sistemas lineares são ditos *equivalentes*, quando têm o mesmo conjunto solução.

No que se segue, introduziremos uma técnica de resolução de sistemas lineares que consiste em transformar um sistema dado em um sistema que lhe é equivalente e de solução imediata. Para tanto, dado um sistema linear  $AX = B$ , consideraremos a sua *matriz aumentada*  $(A|B)$ , que se obtém posicionando-se a *matriz coluna*  $B$  à direita da matriz  $A$ , isto é,

$$(A|B) = \begin{bmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{bmatrix}.$$

Determinaremos, então, o conjunto solução de um sistema linear, efetuando uma série de operações, ditas *elementares*, sobre as linhas de sua matriz aumentada. São elas:

- (E<sub>1</sub>) Troca de posição entre duas linhas  $\ell_i$  e  $\ell_j$ ;
- (E<sub>2</sub>) Multiplicação de uma linha  $\ell_i$  por um escalar  $\lambda \neq 0$ ;
- (E<sub>3</sub>) Substituição de uma linha  $\ell_j$  por  $\ell_j + \lambda\ell_i$ , sendo  $\lambda$  um real não nulo.

Denotaremos as operações  $\mathbf{E}_1$ ,  $\mathbf{E}_2$  e  $\mathbf{E}_3$ , respectivamente, por

$$\ell_i \leftrightarrow \ell_j \qquad \ell_i \rightarrow \lambda \ell_i \qquad \ell_j \rightarrow \ell_j + \lambda \ell_i.$$

**Definição 1.2** (EQUIVALÊNCIA POR LINHAS). Diz-se que uma matriz  $B$  é *linha equivalente* a uma matriz  $A$ , quando  $B$  é obtida de  $A$  efetuando-se nesta uma sequência de operações elementares.

É imediato que qualquer uma das operações elementares sobre uma matriz pode ser invertida e que essa inversão também é uma operação elementar. Especificamente, com a notação acima, as inversas de  $\mathbf{E}_1$ ,  $\mathbf{E}_2$  e  $\mathbf{E}_3$  são, respectivamente,

$$\ell_j \leftrightarrow \ell_i \qquad \ell_i \rightarrow \frac{1}{\lambda} \ell_i \qquad \ell_j \rightarrow \ell_j - \lambda \ell_i.$$

Logo, se  $B$  é linha equivalente à matriz  $A$ , então  $A$  é linha equivalente à matriz  $B$ . Melhor dizendo, a relação de linha-equivalência é simétrica. É fácil ver também que a relação de linha-equivalência é reflexiva e transitiva (verifique!), isto é, *a relação de linha-equivalência é uma relação de equivalência entre matrizes*.

**Proposição 1.1.** *Dois sistemas lineares  $AX = B$  e  $A'X = B'$  são equivalentes, se suas matrizes aumentadas  $(A|B)$  e  $(A'|B')$  são linha equivalentes.*

*Demonstração.* Cada linha da matriz aumentada  $(A|B)$  corresponde a uma equação do sistema  $AX = B$ . Daí, vê-se imediatamente que se  $(A'|B')$  é obtida de  $(A|B)$  efetuando-se nesta uma das duas primeiras operações elementares, então os sistemas correspondentes são equivalentes. Quanto à terceira operação elementar, basta ver que  $(\lambda_1, \dots, \lambda_n)$  é solução de

$$\begin{cases} a_{i1}x_1 + \dots + a_{in}x_n = b_i \\ a_{j1}x_1 + \dots + a_{jn}x_n = b_j \end{cases} \quad (1.4)$$

se, e somente se, é solução de

$$\begin{cases} a_{i1}x_1 + \dots + a_{in}x_n = b_i \\ (a_{j1} + \lambda a_{i1})x_1 + \dots + (a_{jn} + \lambda a_{in})x_n = b_j + \lambda b_i, \end{cases} \quad (1.5)$$

o que se verifica por substituição direta. ■

**Exemplo 1.4.** Consideremos o sistema

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

cujas matriz aumentada é

$$\begin{bmatrix} 1 & 2 & 1 \\ 3 & -1 & 2 \end{bmatrix}.$$

Efetuada-se nessa matriz a operação  $\ell_2 \rightarrow \ell_2 - 3\ell_1$ , obtém-se

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & -7 & -1 \end{bmatrix}.$$

Segue-se, então, da Proposição 1.1, que o sistema dado é equivalente a

$$\begin{cases} x + 2y = 1 \\ -7y = -1 \end{cases},$$

cujas única solução, claramente, é  $(5/7, 1/7)$ .

Note que, nesse exemplo, a operação realizada sobre a matriz aumentada do sistema resultou na eliminação da variável  $x$  da segunda equação, transformando-a numa equação linear de uma única variável, donde se obteve facilmente o conjunto solução. Esse processo de resolução de sistemas lineares nos leva naturalmente ao conceito de *matriz escalonada*, que introduzimos a seguir.

**Definição 1.3** (MATRIZ ESCALONADA). Diz-se que uma matriz  $A = (a_{ij})_{m \times n}$  é *escalonada*, quando cumpre as seguintes condições:

- i) O primeiro elemento não nulo de uma linha está à esquerda do primeiro elemento não-nulo da linha subsequente;
- ii) As linhas nulas, caso existam, estão abaixo das demais.



As matrizes,

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

por exemplo, são escalonadas, enquanto

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

não é uma matriz escalonada.

Não é difícil ver que, efetuando-se adequadamente operações elementares sobre as linhas de uma matriz, podemos transformá-la numa matriz escalonada (a este processo chamamos *escalonamento* da matriz), isto é, *toda matriz é linha equivalente a alguma matriz escalonada*.

Consideremos, por exemplo, a matriz

$$\begin{bmatrix} 1 & -2 & 3 & 1 \\ 2 & 1 & -1 & 2 \\ 4 & -3 & 5 & 4 \end{bmatrix}.$$

Para escaloná-la, operamos sobre suas linhas da seguinte maneira:

$$\begin{bmatrix} 1 & -2 & 3 & 1 \\ 2 & 1 & -1 & 2 \\ 4 & -3 & 5 & 4 \end{bmatrix} \xrightarrow{(1) \text{ e } (2)} \begin{bmatrix} 1 & -2 & 3 & 1 \\ 0 & 5 & -7 & 0 \\ 0 & 5 & -7 & 0 \end{bmatrix} \xrightarrow{(3) \text{ e } (4)} \begin{bmatrix} 1 & 0 & 1/5 & 1 \\ 0 & 5 & -7 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

em que (1), (2), (3), (4) denotam as seguintes operações

$$(1) \ell_2 \rightarrow \ell_2 - 2\ell_1 \quad (2) \ell_3 \rightarrow \ell_3 - 4\ell_1 \quad (3) \ell_1 \rightarrow \ell_1 + \frac{2}{5}\ell_2 \quad (4) \ell_3 \rightarrow \ell_3 - \ell_2.$$

Estamos, agora, em condições de aplicar o *Método de Gauss* para resolução de sistemas lineares. Ele consiste em escalonar a matriz aumentada  $(A|B)$  de um sistema linear  $AX = B$ , transformando-a numa matriz  $(A'|B')$ . Como

resultado, obtém-se o sistema  $A'X = B'$  que, pela Proposição 1.1, é equivalente ao sistema  $AX = B$ , e que é de fácil resolução, pois  $(A'|B')$  é escalonada.

Ilustraremos esse método considerando, inicialmente, o sistema linear

$$\begin{cases} x - 2y + 3z = 1 \\ 2x + y - z = 2 \\ 4x - 3y + 5z = 4 \end{cases} .$$

Observando-se que sua matriz aumentada é a matriz  $A$  do exemplo acima, temos que o sistema dado é equivalente a

$$\begin{cases} x + \frac{1}{5}z = 1 \\ 5y - 7z = 0 \\ 0 = 0 \end{cases} ,$$

cujo conjunto solução é

$$S = \left\{ \left( \frac{5-z}{5}, \frac{7z}{5}, z \right); z \in \mathbb{R} \right\} \subset \mathbb{R}^3.$$

Consideremos agora o sistema

$$(*) \begin{cases} 2x - y + 4t = 9 \\ x + y - z + 2t = 7 \\ -x + 2y + z - t = 3 \\ 4y - z + 3t = 13 \end{cases} .$$

Procedendo-se como indicado acima (verifique, em cada passagem, quais foram as operações elementares realizadas), temos

$$\begin{aligned}
 & \begin{bmatrix} 2 & -1 & 0 & 4 & 9 \\ 1 & 1 & -1 & 2 & 7 \\ -1 & 2 & 1 & -1 & 3 \\ 0 & 4 & -1 & 3 & 13 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & -1 & 2 & 7 \\ 2 & -1 & 0 & 4 & 9 \\ -1 & 2 & 1 & -1 & 3 \\ 0 & 4 & -1 & 3 & 13 \end{bmatrix} \rightarrow \\
 & \begin{bmatrix} 1 & 1 & -1 & 2 & 7 \\ 0 & -3 & 2 & 0 & -5 \\ 0 & 3 & 0 & 1 & 10 \\ 0 & 4 & -1 & 3 & 13 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -\frac{1}{3} & 2 & \frac{16}{3} \\ 0 & -3 & 2 & 0 & -5 \\ 0 & 0 & 2 & 1 & 5 \\ 0 & 0 & \frac{5}{3} & 3 & \frac{19}{3} \end{bmatrix} \rightarrow \\
 & \begin{bmatrix} 1 & 0 & -\frac{1}{3} & 2 & \frac{16}{3} \\ 0 & -3 & 2 & 0 & -5 \\ 0 & 0 & 2 & 1 & 5 \\ 0 & 0 & 0 & \frac{13}{3} & \frac{13}{3} \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -\frac{1}{3} & 2 & \frac{16}{3} \\ 0 & -3 & 2 & 0 & -5 \\ 0 & 0 & 2 & 1 & 5 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}.
 \end{aligned}$$

Dessa forma, o sistema (\*) é equivalente a

$$\begin{cases} x & - \frac{1}{3}z + 2t = \frac{16}{3} \\ & - 3y + 2z & = -5 \\ & & 2z + t = 5 \\ & & & t = 1 \end{cases}$$

cujo conjunto solução é  $S = \{(4, 3, 2, 1)\}$ .

Consideremos, finalmente, o sistema

$$\begin{cases} x + 3y = 5 \\ 2x + 6y = 8 \end{cases}.$$

Escalonando-se sua matriz aumentada, concluí-se facilmente que este é equivalente a

$$\begin{cases} x + 3y = 5 \\ 0 = -2 \end{cases},$$

cujo conjunto solução, evidentemente, é vazio.

No Capítulo 3, voltaremos à discussão sobre sistemas lineares. Lá, estabeleceremos condições sobre as matrizes  $A$  e  $B$  para que um sistema linear  $AX = B$  admita uma ou infinitas soluções.

### 1.3 Inversão e Semelhança de Matrizes

Nesta seção, introduziremos dois conceitos fundamentais da teoria das matrizes, os quais, dentre outras características, determinam uma relação de equivalência na classe das matrizes quadradas de mesma ordem. Vejamos, então, o primeiro deles.

**Definição 1.4** (MATRIZES INVERTÍVEIS). Diz-se que uma matriz quadrada  $A$  é *invertível*, quando existe uma matriz quadrada  $B$ , de mesma ordem que  $A$ , tal que  $AB = BA = I$ .

Sejam  $A, B \in \mathcal{M}(n)$  como na definição acima, e  $C \in \mathcal{M}(n)$ , tal que  $AC = CA = I$ . Multiplicando-se, à esquerda, ambos os membros da igualdade  $AC = I$  por  $B$ , obtém-se  $C = B$ . Logo, a matriz  $B$  é única, a qual é dita a *inversa* de  $A$ , e denotada por  $A^{-1}$ .

Observemos que, se  $A$  é uma matriz invertível, então todo sistema linear  $AX = B$  admite uma única solução, a saber,  $X = A^{-1}B$ . Assim, podemos afirmar que *matrizes invertíveis definem sistemas lineares determinados, isto é, que têm solução única*.

**Definição 1.5** (SEMELHANÇA DE MATRIZES). Dizemos que duas matrizes quadradas  $n \times n$ ,  $A$  e  $B$ , são *semelhantes*, e escrevemos  $A \sim B$ , quando existe uma matriz  $n \times n$  invertível,  $M$ , a qual cumpre a igualdade  $AM = MB$ .

É imediato que a semelhança entre matrizes é uma relação reflexiva e simétrica. Ademais, se  $A$  é semelhante a  $B$ , e  $B$  é semelhante a  $C$ , existem matrizes invertíveis,  $M, N$ , tais que  $AM = MB$  e  $BN = NC$ . Logo,  $A(MN) = (MN)C$ . Porém, sendo  $M$  e  $N$  invertíveis, temos que  $MN$  é invertível (vide Exercício 5), donde se conclui que  $A$  é semelhante a  $C$  e, portanto, que a semelhança de matrizes é também transitiva. Assim, denotando-se por  $\mathcal{M}(n)$  o conjunto das matrizes reais  $n \times n$ , vale o resultado seguinte.

**Proposição 1.2.** *A semelhança  $\sim$  é uma relação de equivalência em  $\mathcal{M}(n)$ .*

Assim como as operações em  $\mathcal{M}(n)$ , o conceito de semelhança é motivado pelas relações entre matrizes e transformações lineares. Nesse contexto, surgem

duas funções especiais, o *traço* e o *determinante*, as quais são invariantes em classes de matrizes semelhantes, isto é, quando duas matrizes são semelhantes, seus traços e determinantes coincidem. Estudaremos os determinantes no Capítulo 4 (vide, entretanto, Exercício 17). O traço, por outro lado, é um conceito mais simples e pode ser considerado agora.

**Definição 1.6** (TRAÇO DE UMA MATRIZ). Chama-se *traço* de uma matriz quadrada,  $A = (a_{ij})_{n \times n}$ , a soma das entradas de sua diagonal, isto é,

$$\text{traço } A = \sum_{i=1}^n a_{ii} = a_{11} + \cdots + a_{nn}.$$

Vejamos, então, que matrizes semelhantes têm o mesmo traço. Para tanto, tomemos matrizes  $n \times n$  arbitrárias,  $A$  e  $B$ , e observemos que

$$\text{traço}(AB) = \sum_{i=1}^n \left( \sum_{k=1}^n a_{ik} b_{ki} \right) = \sum_{k=1}^n \left( \sum_{i=1}^n b_{ki} a_{ik} \right) = \text{traço}(BA).$$

Suponhamos, agora, que  $A$  e  $B$  sejam semelhantes. Assim, existe uma matriz invertível,  $M$ , tal que  $A = MBM^{-1}$ . Logo,

$$\text{traço } A = \text{traço}(MBM^{-1}) = \text{traço}(MM^{-1}B) = \text{traço } B.$$

Vale, desta forma, a implicação:

$$A \sim B \Rightarrow \text{traço } A = \text{traço } B.$$

Diz-se que uma matriz quadrada  $A$  é *diagonalizável*, quando é semelhante a uma matriz diagonal. Conforme mencionamos anteriormente, o Teorema Espectral afirma que toda matriz simétrica é diagonalizável. A importância desse teorema, o qual estabeleceremos no Capítulo 7, deve-se ao fato de que as matrizes diagonais, sob diversos aspectos, são as mais simples. Isso pode ser constatado, por exemplo, quando multiplicamos matrizes diagonais, ou quando calculamos seus determinantes.

## Exercícios

### Seção 1.1

1. Determine (caso exista) a matriz  $A$ , tal que  $AB = C$ , em que

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

2. Mostre que se  $A$  é uma matriz triangular superior, então  $A^2$  é também uma matriz triangular superior.
3. Sejam  $A$  e  $B$  matrizes  $m \times n$  e  $n \times p$ , respectivamente. A implicação

$$AB = 0 \Rightarrow A = 0 \text{ ou } B = 0$$

é verdadeira ou falsa? No caso negativo, exiba um contraexemplo.

4. Encontre uma matriz  $A = (a_{ij})_{2 \times 2}$ , não nula, tal que  $A^2 = \mathbf{0}$ .
5. Sejam  $A$  e  $B$  matrizes  $m \times n$  e  $n \times p$ , respectivamente. Prove que

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

6. Uma matriz quadrada  $A$  é dita *antissimétrica*, quando  $A^* = -A$ . Prove que toda matriz quadrada se expressa como a soma de uma matriz simétrica com uma matriz antissimétrica.
7. Dado  $v = (x_1, \dots, x_n)$ , defina  $[v] = (x_{i1})_{n \times 1}$  por  $x_{i1} = x_i$ , isto é,

$$[v] = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Sejam  $A$  uma matriz  $m \times n$ , e  $B$  uma matriz  $n \times p$ , tais que

$$B = [v_1 \ v_2 \ \cdots \ v_p] \quad \text{e} \quad AB = [w_1 \ w_2 \ \cdots \ w_p].$$

Mostre que vale a igualdade:

$$[w_j] = A[v_j] \quad \forall j = 1, \dots, p.$$

Dito de forma simples, *as colunas de  $AB$  são as respectivas colunas de  $B$  multiplicadas, à esquerda, pela matriz  $A$ .*

### Seção 1.2

8. Determine o conjunto solução dos seguintes sistemas lineares:

$$\text{a) } \begin{cases} x - 2y - 3z = 0 \\ 3x + y - z = -1 \end{cases}$$

$$\text{b) } \begin{cases} x + y + z = 2 \\ 2x - y + 3z = 9 \\ x + 2y - z = -3 \end{cases}$$

9. Mostre que é vazio o conjunto solução do sistema linear:

$$\begin{cases} x_1 - 2x_2 + x_3 + 2x_4 = 1 \\ x_1 + x_2 - x_3 + x_4 = 2 \\ x_1 + 7x_2 - 5x_3 - x_4 = 3 \end{cases}.$$

10. Seja  $A$  a matriz  $4 \times 4$ :

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

Para quais matrizes coluna  $B$ , o sistema  $AX = B$  tem solução?

### Seção 1.3

11. Prove que, se  $A$  e  $B$  são matrizes  $n \times n$  e  $\lambda \in \mathbb{R}$ , então são verdadeiras as seguintes igualdades:

a)  $\text{traço}(A + B) = \text{traço} A + \text{traço} B$ ;

b)  $\text{traço}(\lambda A) = \lambda \text{traço} A$ .

12. Seja  $C$  uma matriz  $2 \times 2$ . Mostre que uma condição necessária e suficiente para que existam matrizes  $A$  e  $B$ ,  $2 \times 2$ , as quais cumprem

$$AB - BA = C,$$

é a de que  $C$  tenha traço nulo.

13. Prove que uma matriz quadrada  $A$  é invertível se, e somente se, sua transposta  $A^*$  é invertível, e que, no caso afirmativo,  $(A^*)^{-1} = (A^{-1})^*$ .

14. Mostre que, se  $A$  e  $B$  são matrizes  $n \times n$ , ambas invertíveis, então  $AB$  é invertível e vale a igualdade  $(AB)^{-1} = B^{-1}A^{-1}$ .

15. Seja  $A$  uma matriz quadrada, a qual cumpre  $A^2 - A + I = 0$ . Mostre que  $A$  é invertível.

16. Prove que uma matriz quadrada  $A$  não é invertível, em qualquer das seguintes ocorrências:

- a) As entradas de uma de suas linhas são todas nulas;
- b) As entradas de uma de suas colunas são todas nulas;
- c) Dois de seus vetores coluna são múltiplos, um do outro;
- d) Dois de seus vetores linha são múltiplos, um do outro.

Conclua que o produto de uma matriz  $n \times 1$  por uma matriz  $1 \times n$  é uma matriz  $n \times n$ , a qual não é invertível.

17. Suponha que  $f : \mathcal{M}(n) \rightarrow \mathbb{R}$  seja uma função, tal que  $f(I) = 1$  e

$$f(AB) = f(A)f(B) \quad \forall A, B \in \mathcal{M}(n).$$

Prove que:

a)  $A \sim B \Rightarrow f(A) = f(B)$ ;

b)  $A$  invertível  $\Rightarrow f(A) \neq 0$  e  $f(A^{-1}) = 1/f(A)$ .



# 2

---

## *Espaços Vetoriais*

Neste capítulo, introduzimos os espaços vetoriais, os quais constituem os principais objetos de nosso estudo. Um tal espaço, simplesmente, é uma estrutura formada por um conjunto e duas operações que possuem certas propriedades, ditas fundamentais. Os espaços euclidianos, os quais consideraremos frequentemente, admitem uma estrutura natural de espaço vetorial que os torna, nesse contexto, os exemplos padrão.

Estabeleceremos os conceitos básicos de subespaço vetorial, dependência linear, base e dimensão, e os aplicaremos para obter um critério de inversão de matrizes. Concluiremos, então, introduzindo o conceito de soma de subespaços vetoriais.

### 2.1 Espaços e Subespaços Vetoriais

Um *espaço vetorial (real)* é um conjunto  $V$  munido de duas operações:

$$\begin{array}{lcl} V \times V & \rightarrow & V \\ (v, w) & \mapsto & v + w \end{array} \quad \text{e} \quad \begin{array}{lcl} \mathbb{R} \times V & \rightarrow & V \\ (\lambda, v) & \mapsto & \lambda v \end{array},$$

chamadas, respectivamente, de *adição* e *multiplicação por escalar*, as quais têm as seguintes propriedades:

**P1:**  $u + v = v + u \quad \forall u, v \in V.$

**P2:**  $(u + v) + w = u + (v + w) \quad \forall u, v, w \in V.$

**P3:** Existe  $\mathbf{0} \in V$ , tal que  $v + \mathbf{0} = v \quad \forall v \in V$ .

**P4:** Para todo  $v \in V$ , existe  $-v \in V$ , tal que  $v + (-v) = \mathbf{0}$ .

**P5:**  $\lambda(\mu v) = (\lambda\mu)v$ .

**P6:**  $1.v = v \quad \forall v \in V$ .

**P7:**  $\lambda(u + v) = \lambda u + \lambda v \quad \forall \lambda \in \mathbb{R}, u, v \in V$ .

**P8:**  $(\lambda + \mu)v = \lambda v + \mu v \quad \forall \lambda, \mu \in \mathbb{R}, v \in V$ .

Nesse contexto, os elementos de  $V$  são chamados de *vetores*, enquanto os números reais são chamados de *escalares*. Um vetor  $u = v + w$  é dito a *soma* de  $v$  e  $w$ . Assim, as propriedades **(P1)** e **(P2)** expressam, respectivamente, a *comutatividade* e *associatividade* da adição de vetores. A propriedade **(P3)**, por sua vez, nos diz que essa adição possui um *elemento neutro*  $\mathbf{0} \in V$ , o qual é chamado de *vetor nulo* ou, simplesmente, de *zero* de  $V$  (vide Exercício 1). A propriedade **(P4)** estabelece que todo vetor de  $V$  possui um *oposto* com respeito à adição. As propriedades **(P5)** e **(P6)** correspondem à associatividade e existência de elemento neutro da multiplicação por escalar. Finalmente, as propriedades **(P7)** e **(P8)** expressam a distributividade da adição e da multiplicação por escalar, uma com respeito à outra.

Decorre diretamente das propriedades fundamentais da adição e multiplicação de números reais, que os conjuntos descritos nos quatro exemplos a seguir, juntamente com as correspondentes operações, são espaços vetoriais.

**Exemplo 2.1** (*Os espaços euclidianos*).  $\mathbb{R}^n = \{(x_1, x_2, \dots, x_n); x_i \in \mathbb{R}\}$ . Dados os vetores  $v = (x_1, \dots, x_n)$ ,  $w = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$  e o escalar  $\lambda \in \mathbb{R}$ , definem-se  $v + w$  e  $\lambda v$ , respectivamente, por

$$v + w = (x_1 + y_1, \dots, x_n + y_n) \quad \text{e} \quad \lambda v = (\lambda x_1, \dots, \lambda x_n).$$

Nesse espaço, o vetor nulo é  $\mathbf{0} = (0, \dots, 0)$ , e o oposto do vetor  $v = (x_1, \dots, x_n)$  é vetor  $-v = (-x_1, \dots, -x_n)$ .

**Exemplo 2.2** (*Os espaços de matrizes*).  $\mathcal{M}(m, n) = \{\text{matrizes reais } m \times n\}$  com as operações usuais da adição de matrizes e multiplicação de número real por matriz. O vetor nulo desse espaço é a matriz nula de  $\mathcal{M}(m, n)$ , isto é, aquela cujas entradas são todas nulas. O oposto da matriz  $A = (a_{ij})$  é a matriz  $-A = (-a_{ij})$ .

**Exemplo 2.3** (*O espaço dos polinômios*). O conjunto  $\mathcal{P}[t, \mathbb{R}]$  formado por todos os polinômios de variável  $t$  e coeficientes em  $\mathbb{R}$ . As operações de adição de polinômios e multiplicação de um número real por um polinômio fazem de  $\mathcal{P}[t, \mathbb{R}]$  um espaço vetorial cujo vetor nulo é o polinômio nulo. Um vetor  $v \in \mathcal{P}[t, \mathbb{R}]$  se escreve como  $v = a_0 + a_1t + \cdots + a_{n-1}t^{n-1} + a_nt^n$ ,  $a_i \in \mathbb{R}$ , e o seu oposto é o vetor  $-v = -a_0 - a_1t - \cdots - a_{n-1}t^{n-1} - a_nt^n$ .

**Exemplo 2.4** (*Os espaços de funções*). O conjunto  $\mathcal{F}(A, \mathbb{R})$  formado por todas as funções  $f : A \rightarrow \mathbb{R}$ , em que  $A$  é um conjunto arbitrário. Dados  $f, g \in \mathcal{F}(A, \mathbb{R})$ , definem-se  $f + g$  e  $\lambda f$ ,  $\lambda \in \mathbb{R}$ , por

$$\begin{array}{ccc} f + g : A & \rightarrow & \mathbb{R} \\ x & \mapsto & f(x) + g(x) \end{array} \quad \text{e} \quad \begin{array}{ccc} \lambda f : A & \rightarrow & \mathbb{R} \\ x & \mapsto & \lambda f(x) \end{array} .$$

A função identicamente nula de  $A$  em  $\mathbb{R}$  é o vetor nulo de  $\mathcal{F}(A, \mathbb{R})$  e, dado  $f \in \mathcal{F}(A, \mathbb{R})$ , seu oposto é a função  $-f$  que, a cada  $x \in A$ , associa o real  $-f(x)$ .

**Proposição 2.1.** *Dado um espaço vetorial  $V$ , valem, para quaisquer  $\lambda \in \mathbb{R}$  e  $v \in V$ , as seguintes igualdades:*

- i)  $0v = \mathbf{0}$ .
- ii)  $\lambda \mathbf{0} = \mathbf{0}$ .
- iii)  $-v = (-1)v$ .

*Demonstração.* De fato,  $0v = (0 + 0)v = 0v + 0v$ , isto é,  $0v = 0v + 0v$ . Adicionando-se  $-0v$  a ambos os membros dessa última igualdade, obtém-se  $0v = \mathbf{0}$ , o que prova (i). De modo inteiramente análogo, prova-se (ii). Da igualdade (i), temos que  $\mathbf{0} = 0v = (-1 + 1)v$ , donde  $(-1)v + v = \mathbf{0}$ . Adicionando-se,

então,  $-v$  a ambos os membros, obtém-se  $-v = (-1)v$ , o que prova a igualdade (iii) e conclui a demonstração. ■

Note que a propriedade (iii) na proposição acima implica na unicidade do vetor oposto e nos sugere definir a *diferença*

$$u - v = u + (-v).$$

**Definição 2.1** (SUBESPAÇO VETORIAL). Um subconjunto  $W \subset V$  de um espaço vetorial  $V$  é chamado de *subespaço vetorial* (ou simplesmente *subespaço*) de  $V$ , se cumpre as condições seguintes:

- i)  $\mathbf{0} \in W$ ;
- ii)  $u + v \in W$  sempre que  $u, v \in W$ ;
- iii)  $\lambda u \in W \forall \lambda \in \mathbb{R}, u \in W$ .

É imediato que  $V$  e  $\{\mathbf{0}\}$  são subespaços de  $V$ . Eles são ditos *subespaços triviais*. O subespaço  $\{\mathbf{0}\}$  é também dito *nulo*. Note que *todo subespaço vetorial é um espaço vetorial*.

**Exemplo 2.5.** Verifiquemos que  $W = \{(x, y) \in \mathbb{R}^2; x = y\}$  é um subespaço vetorial de  $\mathbb{R}^2$ . Com efeito, claramente,  $\mathbf{0} = (0, 0) \in W$ . Ademais, dados  $u = (x_1, y_1), v = (x_2, y_2) \in W$ , temos que  $x_1 = y_1$  e  $x_2 = y_2$ . Logo,  $x_1 + x_2 = y_1 + y_2$  e, para todo  $\lambda \in \mathbb{R}$ ,  $\lambda x_1 = \lambda y_1$ , donde  $u + v = (x_1 + x_2, y_1 + y_2) \in W$  e  $\lambda u = (\lambda x_1, \lambda y_2) \in W$ . Portanto,  $W$  é um subespaço vetorial de  $\mathbb{R}^2$ .

No exemplo acima, podemos observar que, geometricamente,  $W$  é uma reta  $r$  do plano  $\mathbb{R}^2$  (aquela de equação cartesiana  $y = x$ ) que passa pela origem do mesmo. Tomando-se uma equação paramétrica de  $r$ , vemos que

$$W = \{v \in \mathbb{R}^2; v = t(1, 1), t \in \mathbb{R}\},$$

isto é,  $W$  é o conjunto de todos os múltiplos<sup>(i)</sup> de  $(1, 1)$ . Dessa forma, podemos dizer que *o vetor  $(1, 1)$  determina o subespaço vetorial  $W$* .

---

<sup>(i)</sup>Um vetor  $w$  é um *múltiplo* de outro vetor  $v$ , se existe  $\lambda \in \mathbb{R}$ , tal que  $w = \lambda v$ .

**Exemplo 2.6.** Seja  $W$  o seguinte subconjunto do espaço  $\mathbb{R}^3$  :

$$W = \{(x, y, z) \in \mathbb{R}^3; 2x + y - z = 0\}.$$

É imediato que o vetor nulo  $\mathbf{0} = (0, 0, 0)$  pertence a  $W$ . Além disso, dados  $u = (x_1, y_1, z_1)$  e  $v = (x_2, y_2, z_2)$  em  $W$ , temos que  $2x_1 + y_1 - z_1 = 0$  e  $2x_2 + y_2 - z_2 = 0$ . Adicionando-se estas igualdades membro a membro, obtém-se  $2(x_1 + x_2) + (y_1 + y_2) - (z_1 + z_2) = 0$ . Além disso, multiplicando-se a primeira delas por  $\lambda \in \mathbb{R}$ , tem-se  $2(\lambda x_1) + (\lambda y_1) - (\lambda z_1) = 0$ . Dessa forma,  $u + v = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$  e  $\lambda u = (\lambda x_1, \lambda y_1, \lambda z_1)$  são vetores de  $W$ . Logo,  $W$  é um subespaço vetorial de  $\mathbb{R}^3$ .

É sabido que o conjunto  $W$  do exemplo anterior é um plano de  $\mathbb{R}^3$  que contém a origem  $(0, 0, 0)$ . Tomando-se, por exemplo, os vetores  $v_1 = (0, 1, 1)$  e  $v_2 = (1, 0, 2)$ , de  $W$ , obtém-se uma representação paramétrica para o mesmo, qual seja,  $W = \{v \in \mathbb{R}^3; v = \lambda_1 v_1 + \lambda_2 v_2, \lambda_1, \lambda_2 \in \mathbb{R}\}$ . Assim, podemos afirmar que *os vetores  $v_1 = (0, 1, 1)$  e  $v_2 = (1, 0, 2)$  determinam o subespaço  $W$* .

Os exemplos acima motivam introduzir, no contexto dos espaços vetoriais, o conceito de *subespaço gerado* por um certo subconjunto de vetores. Com essa terminologia, temos, no Exemplo 2.5, que  $W$  é o subespaço gerado pelo vetor  $(1, 1)$ , enquanto no Exemplo 2.6,  $W$  é o subespaço gerado pelos vetores  $(0, 1, 1)$  e  $(1, 0, 2)$ . No que se segue, tornaremos essa ideia mais precisa.

**Definição 2.2** (COMBINAÇÃO LINEAR). Diz-se que um vetor  $v$ , de um espaço vetorial  $V$ , é uma *combinação linear* dos vetores  $v_1, \dots, v_n \in V$ , quando existem escalares  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ , tais que

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n.$$

O vetor  $(1, 2) \in \mathbb{R}^2$ , por exemplo, é uma combinação linear dos vetores  $v_1 = (1, 0)$  e  $v_2 = (0, 1)$  de  $\mathbb{R}^2$ , pois  $(1, 2) = 1(1, 0) + 2(0, 1)$ . Note que todo vetor  $v = (x, y) \in \mathbb{R}^2$  tem essa propriedade, já que

$$(x, y) = x(1, 0) + y(0, 1) \quad \forall (x, y) \in \mathbb{R}^2.$$

**Exemplo 2.7.** Verifiquemos que, em  $\mathbb{R}^2$ , o vetor  $(2, -3)$  é combinação linear dos vetores  $(1, -1)$  e  $(2, 1)$ . Para isso, devemos encontrar escalares  $\lambda_1, \lambda_2$ , tais que  $(2, -3) = \lambda_1(1, -1) + \lambda_2(2, 1) = (\lambda_1 + 2\lambda_2, -\lambda_1 + \lambda_2)$ . Sendo assim, devemos ter  $\lambda_1 + 2\lambda_2 = 2$  e  $-\lambda_1 + \lambda_2 = -3$ , isto é,  $\lambda_1 = 8/3$  e  $\lambda_2 = -1/3$ . Logo, vale a seguinte igualdade:

$$(2, -3) = \frac{8}{3}(1, -1) - \frac{1}{3}(2, 1).$$

**Exemplo 2.8.** Consideremos em  $\mathcal{M}(2, 2)$  os vetores

$$v_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad v_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Dado um vetor  $v = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathcal{M}(2, 2)$ , tem-se

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

donde se conclui que todo vetor  $v \in \mathcal{M}(2, 2)$  é combinação linear dos vetores  $v_1, v_2, v_3, v_4$ .

**Exemplo 2.9.** Segue-se diretamente da definição de polinômio que todo vetor  $v \in \mathcal{P}[t, \mathbb{R}]$ , de grau  $n$ , é uma combinação linear dos vetores  $1, t, \dots, t^n$ .

**Observação 2.1.** O conceito de combinação linear pode ser usado para reinterpretar o produto de uma matriz  $A = (a_{ij})_{m \times n}$  por uma *matriz coluna*  $X = [v]$ , em que  $v = (x_1, \dots, x_n)$  é um vetor de  $\mathbb{R}^n$  (vide Exercício 7 – Cap. 1). Com efeito, indicando-se por  $v_1, \dots, v_n \in \mathbb{R}^n$  os vetores coluna de  $A$ , tem-se, pela definição de produto de matrizes, que

$$AX = A[v] = \begin{bmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{bmatrix} = x_1[v_1] + \dots + x_n[v_n].$$

Logo, o produto de uma matriz  $m \times n$ ,  $A$ , por uma matriz coluna  $n \times 1$ ,  $X$ , é uma combinação linear dos vetores coluna de  $A$ , cujos coeficientes são as correspondentes entradas de  $X$ .

**Proposição 2.2** (COMBINAÇÕES LINEARES E SUBESPAÇOS). *Sejam  $V$  um espaço vetorial e  $v_1, \dots, v_n$  vetores de  $V$ . Considere o conjunto  $W$  formado por todas as combinações lineares desses vetores, isto é,*

$$W = \{v \in V; v = \lambda_1 v_1 + \dots + \lambda_n v_n, \lambda_i \in \mathbb{R}\}.$$

*Então,  $W$  é um subespaço vetorial de  $V$ .*

*Demonstração.* Temos que  $\mathbf{0} = 0v_1 + \dots + 0v_n$ , isto é,  $\mathbf{0} \in W$ . Tomemos  $u = \mu_1 v_1 + \dots + \mu_n v_n$  e  $v = \lambda_1 v_1 + \dots + \lambda_n v_n$  em  $W$ , e observemos que  $u + v = (\mu_1 + \lambda_1)v_1 + \dots + (\mu_n + \lambda_n)v_n \in W$ . Ademais, para todo  $\lambda \in \mathbb{R}$ ,  $\lambda v = (\lambda\lambda_1)v_1 + \dots + (\lambda\lambda_n)v_n \in W$ . Logo,  $W$  é um subespaço de  $V$ . ■

**Definição 2.3** (SUBESPAÇO GERADO). O subespaço  $W$  da Proposição 2.2 é chamado de *subespaço gerado* pelos vetores  $v_1, \dots, v_n$ , o qual denotaremos por  $\text{Ger}\{v_1, \dots, v_n\}$ . Assim,

$$\text{Ger}\{v_1, \dots, v_n\} = \{v \in V; v = \lambda_1 v_1 + \dots + \lambda_n v_n, \lambda_i \in \mathbb{R}\}.$$

Pelo Exemplo 2.6, bem como pelas observações feitas após o mesmo, temos que  $\text{Ger}\{(0, 1, 1), (1, 0, 2)\} = \{(x, y, z) \in \mathbb{R}^3; 2x + y - z = 0\} \subset \mathbb{R}^3$ . Segue-se também de considerações anteriores que  $\text{Ger}\{(1, 0), (0, 1)\} = \mathbb{R}^2$ . Note ainda que os vetores  $v_1, v_2, v_3, v_4$  do Exemplo 2.8 geram  $\mathcal{M}(2, 2)$ , isto é,  $\text{Ger}\{v_1, v_2, v_3, v_4\} = \mathcal{M}(2, 2)$ .

Essencialmente, a única operação entre conjuntos que é fechada na classe dos subespaços vetoriais é a interseção, conforme a proposição seguinte.

**Proposição 2.3** (INTERSEÇÃO DE SUBESPAÇOS). *Sejam  $U$  e  $W$  subespaços de um espaço vetorial  $V$ . Então,  $U \cap W$  é não vazio e constitui um subespaço de  $V$ .*

*Demonstração.* Temos, pela definição de subespaço vetorial, que  $\mathbf{0} \in U \cap W$ . Logo,  $U \cap W \neq \emptyset$ . Sejam  $u, w \in U \cap W$  e  $\lambda \in \mathbb{R}$ . Em particular,  $u, w \in U$ , o que nos dá  $u + w \in U$  e  $\lambda u \in U$ , já que  $U$  é um subespaço de  $V$ . Analogamente,  $u + w \in W$  e  $\lambda u \in W$ . Dessa forma,  $u + w \in U \cap W$  e  $\lambda u \in U \cap W$ , donde se conclui que  $U \cap W$  é um subespaço vetorial de  $V$ . ■

**Proposição 2.4.** *Seja  $\{v_1, \dots, v_n\}$  um subconjunto finito de um espaço vetorial  $V$ . Se  $w \in \text{Ger}\{v_1, \dots, v_n\}$ , então  $\text{Ger}\{v_1, \dots, v_n, w\} = \text{Ger}\{v_1, \dots, v_n\}$ .*

*Demonstração.* Dado  $v = \lambda_1 v_1 + \dots + \lambda_n v_n \in \text{Ger}\{v_1, \dots, v_n\}$ , podemos escrever  $v = \lambda_1 v_1 + \dots + \lambda_n v_n + 0w$ , donde  $v \in \text{Ger}\{v_1, \dots, v_n, w\}$ . Logo,  $\text{Ger}\{v_1, \dots, v_n, w\} \supset \text{Ger}\{v_1, \dots, v_n\}$ .

Agora, uma vez que  $w \in \text{Ger}\{v_1, \dots, v_n\}$ , existem escalares  $\lambda_1, \lambda_2, \dots, \lambda_n$ , tais que  $w = \lambda_1 v_1 + \dots + \lambda_n v_n$ . Dado, então,  $v \in \text{Ger}\{v_1, \dots, v_n, w\}$ , sejam  $\mu_1, \dots, \mu_{n+1}$  escalares tais que  $v = \mu_1 v_1 + \dots + \mu_n v_n + \mu_{n+1} w$ . Assim, temos

$$\begin{aligned} v &= \mu_1 v_1 + \dots + \mu_n v_n + \mu_{n+1} (\lambda_1 v_1 + \dots + \lambda_n v_n) \\ &= (\mu_1 + \mu_{n+1} \lambda_1) v_1 + \dots + (\mu_n + \mu_{n+1} \lambda_n) v_n, \end{aligned}$$

donde se infere que  $v \in \text{Ger}\{v_1, \dots, v_n\}$ . Logo, vale a inclusão

$$\text{Ger}\{v_1, \dots, v_n, w\} \subset \text{Ger}\{v_1, \dots, v_n\},$$

e, portanto,  $\text{Ger}\{v_1, \dots, v_n, w\} = \text{Ger}\{v_1, \dots, v_n\}$ , como queríamos demonstrar. ■

## 2.2 Dependência Linear

No que diz respeito a subespaços gerados, a Proposição 2.4 da seção anterior nos leva a concluir que, num conjunto  $\mathcal{A}$  de geradores de um tal subespaço, pode haver vetores “obsoletos”, isto é, o conjunto obtido de  $\mathcal{A}$  por exclusão desses vetores gera o mesmo subespaço que  $\mathcal{A}$ . Nosso objetivo, agora, será o de determinar condições que, ao serem verificadas, nos permitirão decidir se há ou não vetores obsoletos num conjunto de geradores de um subespaço vetorial. Isso nos leva à noção de *dependência linear* de vetores.

Consideremos, inicialmente, o espaço  $\mathbb{R}^2$ .

**Lema 2.1.** *Sejam  $v = (a, b)$  e  $w = (c, d)$  vetores não nulos de  $\mathbb{R}^2$ . Então,  $v$  é múltiplo de  $w$  se, e somente se,  $ad - bc = 0$ .*



*Demonstração.* Com efeito, se existe  $\lambda \in \mathbb{R}$  satisfazendo  $v = \lambda w$ , então  $\lambda \neq 0$ , pois  $v, w \neq 0$ . Além disso, devemos ter  $a = \lambda c$  e  $b = \lambda d$ . Logo,

$$ad - bc = (\lambda c)d - (\lambda d)c = 0.$$

Reciprocamente, suponhamos que  $ad - bc = 0$  se cumpra. Como  $v$  e  $w$  são não nulos, temos que  $a \neq 0$  ou  $b \neq 0$  e  $c \neq 0$  ou  $d \neq 0$ . Suponhamos, sem perda de generalidade, que  $d, b \neq 0$  e façamos  $\lambda = b/d$ . Nesse caso, segue-se da hipótese que  $a = \lambda c$ . Assim,

$$v = (a, b) = (\lambda c, \lambda d) = \lambda(c, d) = \lambda w,$$

donde se infere que  $v$  é múltiplo de  $w$ . ■

**Proposição 2.5.** *Sejam  $v = (a, b)$  e  $w = (c, d)$  vetores não nulos de  $\mathbb{R}^2$ . Então,  $\text{Ger}\{v, w\} = \mathbb{R}^2$  se, e somente se,  $v$  não é múltiplo de  $w$ .*

*Demonstração.* Suponhamos que  $\text{Ger}\{v, w\} = \mathbb{R}^2$ . Então, para todo vetor  $u = (x, y) \in \mathbb{R}^2$ , existem escalares  $\lambda, \mu$  satisfazendo  $u = \lambda v + \mu w$ . Daí, tem-se  $x = \lambda a + \mu c$  e  $y = \lambda b + \mu d$ . Multiplicando-se a primeira destas equações por  $-b$ , a segunda por  $a$  e adicionando-as, obtemos  $-bx + ay = \mu(ad - bc)$ . Façamos, agora,  $u = (-b, a)$ . Teremos, então,  $a^2 + b^2 = \mu(ad - bc)$ . No entanto,  $v = (a, b) \neq 0$ , isto é,  $a^2 + b^2 > 0$ . Segue-se que  $ad - bc \neq 0$  e, pelo Lema 2.1, que  $v$  não é múltiplo de  $w$ .

Reciprocamente, suponhamos que  $v = (a, b)$  não seja múltiplo do vetor  $w = (c, d)$ , isto é, que  $ad - bc \neq 0$ . Dado então  $u = (x, y) \in \mathbb{R}^2$ , tomando-se

$$\lambda = \frac{dx - cy}{ad - bc} \quad \text{e} \quad \mu = \frac{ay - bx}{ad - bc},$$

verifica-se facilmente que  $u = \lambda v + \mu w$ , isto é,  $u \in \text{Ger}\{v, w\}$ . Uma vez que  $u$  é arbitrário, segue-se que  $\text{Ger}\{v, w\} = \mathbb{R}^2$ . ■

Segue-se das Proposições 2.4 e 2.5 que, se  $v, w \in \mathbb{R}^2$  são não-nulos e um é múltiplo do outro, então  $\text{Ger}\{v, w\} = \text{Ger}\{v\} = \text{Ger}\{w\} \subsetneq \mathbb{R}^2$ . Nesse caso,  $v$

e  $w$  satisfazem a relação  $v - \lambda w = 0$  para algum escalar  $\lambda$ . Por outro lado, se  $\text{Ger}\{v, w\} = \mathbb{R}^2$ , a única relação entre os vetores  $v$  e  $w$  do tipo

$$\lambda v + \mu w = \mathbf{0} \quad (2.1)$$

é aquela em que  $\lambda = \mu = 0$ . De fato, se tivéssemos, digamos,  $\lambda \neq 0$ , resultaria de (2.1) que  $v = -\frac{\mu}{\lambda}w$ , o que contradiria a Proposição 2.5.

Essas considerações motivam a definição que se segue.

**Definição 2.4** (DEPENDÊNCIA LINEAR). Seja  $\mathcal{A} = \{v_1, \dots, v_n\}$  um conjunto de  $n$  vetores de um espaço vetorial  $V$ . Uma combinação linear desses vetores,

$$\lambda_1 v_1 + \dots + \lambda_n v_n = v,$$

é dita *nula*, se o vetor resultante  $v$  é o vetor nulo. Uma combinação linear nula  $\lambda_1 v_1 + \dots + \lambda_n v_n = \mathbf{0}$  é dita *trivial*, se todos os escalares  $\lambda_i$  são nulos, caso contrário, ela é dita *não trivial*. Diz-se, então, que o conjunto  $\mathcal{A}$  é *linearmente dependente* (LD), quando existe uma combinação linear nula e não trivial de seus elementos, caso contrário, ele é dito *linearmente independente* (LI).

Sejam  $v$  e  $w$  vetores de um espaço vetorial  $V$ . Se  $v$  é múltiplo de  $w$ , então  $\{v, w\}$  é LD pois, nesse caso, existe um escalar  $\lambda$ , tal que  $v = \lambda w$ , isto é,  $v - \lambda w = \mathbf{0}$ . Já o conjunto  $\{(1, 0), (0, 1)\} \subset \mathbb{R}^2$  é LI. Com efeito, considerando-se uma combinação linear nula de seus elementos,  $\lambda(1, 0) + \mu(0, 1) = (0, 0)$ , teremos  $(\lambda, \mu) = (0, 0)$ , implicando  $\lambda = \mu = 0$ . Assim, a única combinação linear nula possível de  $(1, 0)$  e  $(0, 1)$  é a trivial. Logo,  $\{(1, 0), (0, 1)\}$  é LI.

**Exemplo 2.10.** Consideremos  $\mathcal{A} = \{(1, 2, -1), (1, 0, 1), (0, -1, 1)\} \subset \mathbb{R}^3$  e verifiquemos se esse conjunto é LI ou LD. Para tanto, tomemos uma combinação linear nula de seus elementos

$$x(1, 2, -1) + y(1, 0, 1) + z(0, -1, 1) = (0, 0, 0).$$

Daí, resulta o seguinte sistema linear

$$\begin{cases} x + y & = 0 \\ 2x & - z = 0 \\ -x + y + z & = 0 \end{cases} .$$

Resolvendo-o, verifica-se que qualquer terno  $(x, -x, 2x)$  é uma solução do mesmo. Tomando-se, então,  $x \neq 0$ , obtém-se uma combinação linear nula e não trivial dos vetores de  $\mathcal{A}$ , qual seja,

$$x(1, 2, -1) - x(1, 0, 1) + 2x(0, -1, 1) = \mathbf{0}.$$

Logo,  $\mathcal{A}$  é linearmente dependente.

Procedendo-se como no exemplo acima, obtém-se facilmente o resultado seguinte (vide, também, Exercício 11).

**Proposição 2.6.** *Dado um subconjunto finito  $\mathcal{A} = \{w_1, \dots, w_n\} \subset \mathbb{R}^m$ , seja  $M = [w_1 \cdots w_n]$  a matriz  $m \times n$  cujos vetores coluna são os elementos de  $\mathcal{A}$ . Então,  $\mathcal{A}$  é LI se, e somente se, o sistema linear homogêneo  $MX = \mathbf{0}$  admite apenas a solução trivial.*

**Exemplo 2.11.** Sejam  $f, g : \mathbb{R} \rightarrow \mathbb{R}$  as função cosseno e seno, isto é,  $f(x) = \cos x$  e  $g(x) = \sin x$ . O conjunto  $\mathcal{A} = \{f, g\}$  é LI no espaço de funções  $\mathcal{F}(\mathbb{R}, \mathbb{R})$ . De fato, dada uma combinação linear nula de  $f$  e  $g$ ,  $\lambda f + \mu g = 0$ , devemos ter, para todo  $x \in \mathbb{R}$ ,

$$\lambda \cos x + \mu \sin x = 0.$$

Fazendo-se  $x = 0$  e, em seguida,  $x = \pi/2$ , obtém-se  $\lambda = \mu = 0$ . Logo,  $\mathcal{A}$  é um conjunto LI de vetores de  $\mathcal{F}(\mathbb{R}, \mathbb{R})$ .

Verifiquemos, nas proposições seguintes, algumas propriedades de conjuntos LI e LD.

**Proposição 2.7.** *Sejam  $\mathcal{A} \subset \mathcal{B} \subset \mathcal{C}$  subconjuntos finitos de um espaço vetorial  $V$ . Então, são verdadeiras as seguintes afirmações:*

- i)  $\mathcal{A}$  é LI, se  $\mathcal{B}$  é LI;
- ii)  $\mathcal{C}$  é LD, se  $\mathcal{B}$  é LD;

*isto é, todo subconjunto de um conjunto LI é LI, e todo conjunto que contém um conjunto LD é LD.*

*Demonstração.* Fazemos  $\mathcal{A} = \{v_1, \dots, v_k\}$ ,  $\mathcal{B} = \{v_1, \dots, v_k, v_{k+1}, \dots, v_m\}$  e  $\mathcal{C} = \{v_1, \dots, v_k, v_{k+1}, \dots, v_m, v_{m+1}, \dots, v_n\}$ , em que  $k \leq m \leq n$ . Suponhamos que  $\mathcal{B}$  seja LI e tomemos uma combinação linear nula dos vetores de  $\mathcal{A}$ ,  $\lambda_1 v_1 + \dots + \lambda_k v_k = 0$ . Se algum dos  $\lambda_i$ 's fosse diferente de zero, teríamos uma combinação linear nula não trivial dos elementos de  $\mathcal{B}$ , a saber,

$$\lambda_1 v_1 + \dots + \lambda_k v_k + 0v_{k+1} + \dots + 0v_m = \mathbf{0},$$

contrariando o fato de  $\mathcal{B}$  ser LI. Segue-se que cada  $\lambda_i$  é nulo e, portanto, que  $\mathcal{A}$  é LI, o que prova (i). Quanto ao item (ii), suponhamos que  $\mathcal{B}$  seja LD. Então, existe uma combinação linear nula e não trivial dos vetores de  $\mathcal{B}$ ,  $\lambda_1 v_1 + \dots + \lambda_m v_m = 0$ . Dessa forma,

$$\lambda_1 v_1 + \dots + \lambda_m v_m + 0v_{m+1} + \dots + 0v_n = \mathbf{0}$$

é uma combinação linear nula e não trivial dos vetores de  $\mathcal{C}$ , donde se infere que  $\mathcal{C}$  é LD. ■

**Proposição 2.8.** *Um subconjunto  $\mathcal{A} \subset V$  de um espaço vetorial  $V$  é LD se, e somente se, um de seus vetores se expressa como combinação linear dos demais.*

*Demonstração.* Suponhamos que  $\mathcal{A} = \{v_1, \dots, v_n\}$  seja LD e tomemos uma combinação linear nula e não trivial dos vetores de  $\mathcal{A}$ ,

$$\lambda_1 v_1 + \dots + \lambda_n v_n = \mathbf{0}. \quad (2.2)$$

Então, para algum  $i \in \{1, 2, \dots, n\}$ , tem-se  $\lambda_i \neq 0$ . Daí e da igualdade (2.2), obtém-se

$$v_i = -\frac{\lambda_1}{\lambda_i} v_1 - \dots - \frac{\lambda_{i-1}}{\lambda_i} v_{i-1} - \frac{\lambda_{i+1}}{\lambda_i} v_{i+1} \dots - \frac{\lambda_n}{\lambda_i} v_n,$$

isto é,  $v_i$  é combinação linear de  $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$ .

Agora, se para algum  $i \in \{1, 2, \dots, n\}$ , tem-se

$$v_i = \lambda_1 v_1 + \dots + \lambda_{i-1} v_{i-1} + \lambda_{i+1} v_{i+1} + \dots + \lambda_n v_n,$$

obtem-se daí a combinação linear nula e não trivial

$$\lambda_1 v_1 + \cdots + \lambda_{i-1} v_{i-1} - v_i + \lambda_{i+1} v_{i+1} + \cdots + \lambda_n v_n = \mathbf{0},$$

donde  $\mathcal{A}$  é LD. ■

O conceito de dependência linear pode ser facilmente estendido a subconjuntos infinitos de um espaço vetorial  $V$ . Basta dizer que um tal conjunto é LI se qualquer um de seus subconjuntos finitos é LI no sentido da Definição 2.4. Caso contrário, diz-se que este conjunto é LD. Deixamos a cargo do leitor a tarefa de verificar que o conjunto

$$\mathcal{A} = \{1, t, t^2, \dots, t^n, t^{n+1}, \dots\}$$

é LI no espaço dos polinômios,  $\mathcal{P}[t, \mathbb{R}]$ .

## 2.3 Bases – Dimensão

Nesta seção, consideraremos subconjuntos de espaços vetoriais, os quais chamamos de *base*, que possuem duas características: são linearmente independentes e geram o espaço vetorial que os contém. Constataremos, então, que duas bases finitas quaisquer de um espaço vetorial  $V$  têm, necessariamente, o mesmo número de elementos, o qual chamamos *dimensão* de  $V$ .

Muitas propriedades de um espaço vetorial são verificadas através de suas bases. Isto se dá, especialmente, quando do estudo de aplicações entre espaços vetoriais, conforme constataremos.

**Definição 2.5** (BASE – DIMENSÃO FINITA). Um subconjunto  $\mathcal{B} \subset V$  de um espaço vetorial  $V$  é dito uma *base* do mesmo, se

- i)  $\mathcal{B}$  é LI;
- ii)  $\text{Ger } \mathcal{B} = V$ .

Diz-se, então, que  $V$  tem *dimensão finita*, quando contém uma base finita. Caso contrário, diz-se que  $V$  tem *dimensão infinita*.

Para estabelecer o principal resultado desta seção (Teorema 2.1), usaremos a proposição abaixo, que refina o resultado da Proposição 2.8.

**Proposição 2.9.** *Seja  $\mathcal{A} = \{v_1, \dots, v_n\}$  um conjunto de vetores não nulos de um espaço vetorial  $V$ . Então,  $\mathcal{A}$  é LD se, e somente se, para algum  $k \in \{2, \dots, n\}$ , o vetor  $v_k$  é uma combinação linear de  $v_1, \dots, v_{k-1}$ .*

*Demonstração.* Suponhamos que  $\mathcal{A}$  seja LD. Defina  $k \in \{2, \dots, n\}$  como o menor inteiro para o qual o conjunto  $\{v_1, v_2, \dots, v_k\}$  seja LD (note que, se  $v_1$  fosse o vetor nulo, um tal  $k$  não existiria). Dessa forma, existe uma combinação linear nula e não trivial de  $v_1, \dots, v_k$ ,

$$\lambda_1 v_1 + \dots + \lambda_k v_k = \mathbf{0}.$$

Além disso, devemos ter  $\lambda_k \neq 0$ , caso contrário  $\lambda_1 v_1 + \dots + \lambda_{k-1} v_{k-1} = \mathbf{0}$  seria uma combinação linear nula e não trivial de  $v_1, \dots, v_{k-1}$ , o que, juntamente com a Proposição 2.8, implicaria que estes vetores são LD, contrariando, dessa forma, a definição de  $k$ . Segue-se que  $\lambda_k \neq 0$  e, portanto, que

$$v_k = -\frac{\lambda_1}{\lambda_k} v_1 - \dots - \frac{\lambda_{k-1}}{\lambda_k} v_{k-1},$$

isto é,  $v_k$  é combinação linear de  $v_1, \dots, v_{k-1}$ .

Reciprocamente, suponhamos que, para algum  $k \in \{2, \dots, n\}$ , tenha-se  $v_k \in \text{Ger}\{v_1, \dots, v_{k-1}\}$ . Nesse caso, pela Proposição 2.8,  $\{v_1, \dots, v_{k-1}, v_k\}$  é LD. Logo, pela Proposição 2.7-(ii),  $\mathcal{A} = \{v_1, \dots, v_k, \dots, v_n\}$  é LD. ■

**Corolário 2.1.** *Sejam  $V$  um espaço vetorial e  $\mathcal{A} = \{v_1, \dots, v_n\} \subset V$  um subconjunto LI. Nessas condições, se  $v \in V - \text{Ger } \mathcal{A}$ , então  $\{v_1, \dots, v_n, v\}$  é LI.*

*Demonstração.* Pela Proposição 2.9, nenhum dos vetores  $v_i$  é combinação linear dos anteriores, uma vez que  $\mathcal{A}$  é LI. Além disso, por hipótese,  $v$  não é combinação linear dos vetores de  $\mathcal{A}$ . Logo, pela Proposição 2.9,  $\{v_1, \dots, v_n, v\}$  é linearmente independente. ■

O teorema a seguir, cujo argumento é devido a P. Halmos, nos permitirá introduzir um dos conceitos fundamentais da Álgebra Linear, o de *dimensão*.

**Teorema 2.1.** *Seja  $V$  um espaço vetorial de dimensão finita. Então, a quantidade de elementos de um conjunto LI, de  $V$ , nunca excede a de um conjunto de geradores desse espaço. Mais precisamente, se  $\mathcal{A} = \{v_1, \dots, v_n\} \subset V$  é LI, e  $\mathcal{B} = \{w_1, \dots, w_m\}$  gera  $V$ , então,  $m \geq n$ .*

*Demonstração.* Consideremos o conjunto  $\mathcal{B}' = \{v_n, w_1, \dots, w_m\}$ . Uma vez que  $\mathcal{B}$  gera  $V$ , temos, em particular, que  $v_n \in \text{Ger}\{w_1, \dots, w_m\}$ . Logo, pela Proposição 2.8,  $\mathcal{B}'$  é LD. Então, pela Proposição 2.9, um dos vetores  $w_i$ , de  $\mathcal{B}'$ , é uma combinação linear dos anteriores a ele em  $\mathcal{B}'$ . Segue-se, então, da Proposição 2.4 e do fato de  $\mathcal{B}'$  gerar  $V$  (já que  $\mathcal{B}$  gera  $V$ ), que

$$\mathcal{B}_1 = \{v_n, w_1, \dots, w_{i-1}, w_{i+1}, \dots, w_m\}$$

gera  $V$ . Consideremos agora o conjunto

$$\mathcal{B}'_1 = \{v_{n-1}, v_n, w_1, \dots, w_{i-1}, w_{i+1}, \dots, w_m\}.$$

Como anteriormente,  $\mathcal{B}'_1$  é LD e gera  $V$ . Sendo assim, podemos proceder de modo análogo e obter

$$\mathcal{B}_2 = \{v_{n-1}, v_n, w_1, \dots, w_{i-1}, w_{i+1}, \dots, w_{j-1}, w_{j+1}, \dots, w_m\},$$

tal que  $\text{Ger } \mathcal{B}_2 = V$ .

Repetimos esse processo (de se retirar um elemento de  $\mathcal{B}$  a cada elemento de  $\mathcal{A}$  acrescentado) até que não haja mais  $v_i$ 's a acrescentar (o que ocorrerá se  $m \geq n$ ) ou  $w_i$ 's a retirar (o que ocorrerá se  $m \leq n$ ).

Dessa forma, se tivéssemos  $m < n$ , fazendo-se  $k = (n - m) + 1$ , terminaríamos com um subconjunto próprio de  $\mathcal{A}$ ,  $\{v_k, \dots, v_n\}$ , o qual geraria  $V$ . Em particular, cada um dos vetores  $v_1, \dots, v_{k-1}$  seria uma combinação linear de  $v_k, \dots, v_n$ , contrariando o fato de  $\mathcal{A}$  ser LI. Segue-se desta contradição que devemos ter, necessariamente,  $m \geq n$ , como queríamos demonstrar. ■

**Corolário 2.2.** *Dois bases quaisquer de um espaço vetorial de dimensão finita têm a mesma quantidade de vetores.*

*Demonstração.* Com efeito, suponha que  $\mathcal{A}$  e  $\mathcal{B}$  sejam bases distintas de um tal espaço vetorial, com  $m$  e  $n$  elementos, respectivamente. Uma vez que, ambos,  $\mathcal{A}$  e  $\mathcal{B}$ , são LI e geram  $V$ , pelo Teorema 2.1, devemos ter  $m \geq n$  e  $n \geq m$ , donde  $m = n$ . ■

**Definição 2.6** (DIMENSÃO). Seja  $V$  um espaço vetorial de dimensão finita. A *dimensão* de  $V$ , que será denotada por  $\dim V$ , define-se como o número de elementos de uma qualquer de suas bases. O espaço vetorial nulo, por convenção, tem dimensão 0.

Os vetores

$$e_1 = (1, 0, \dots, 0), e_2 = (0, 1, 0, \dots, 0), \dots, e_n = (0, 0, \dots, 1),$$

claramente, formam uma base de  $\mathbb{R}^n$ , dita a *base canônica* desse espaço. Segue-se que  $\dim \mathbb{R}^n = n$ . É fácil ver também que os vetores  $v_1, v_2, v_3, v_4$ , definidos no Exemplo 2.8, constituem uma base de  $\mathcal{M}(2, 2)$ , donde se infere que  $\dim \mathcal{M}(2, 2) = 4$ . De modo geral, tem-se  $\dim \mathcal{M}(m, n) = mn$ .

Dada uma base  $\mathcal{B} = \{v_1, \dots, v_n\}$  de um espaço vetorial  $V$  de dimensão  $n$ , para todo  $v \in V$ , os escalares  $\lambda_1, \dots, \lambda_n$  que cumprem a igualdade

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n$$

são únicos. Eles são ditos as *coordenadas* de  $v$  com relação à base  $\mathcal{B}$ . Para constatarmos isso, suponhamos que

$$v = \mu_1 v_1 + \dots + \mu_n v_n.$$

Nesse caso, devemos ter

$$\lambda_1 v_1 + \dots + \lambda_n v_n = \mu_1 v_1 + \dots + \mu_n v_n,$$

donde  $(\lambda_1 - \mu_1)v_1 + \dots + (\lambda_n - \mu_n)v_n = \mathbf{0}$ . Uma vez que  $\mathcal{B}$  é LI, segue-se que  $\lambda_1 - \mu_1 = \dots = \lambda_n - \mu_n = 0$ . Logo,  $\lambda_i = \mu_i \forall i \in \{1, \dots, n\}$ .

**Exemplo 2.12.** O conjunto  $\mathcal{B} = \{(1, 1), (-1, 1)\}$  é uma base de  $\mathbb{R}^2$  (verifique!). Dado um vetor  $(x, y) \in \mathbb{R}^2$ , determinemos suas coordenadas com respeito a



essa base. Assim, buscamos escalares  $\lambda, \mu \in \mathbb{R}$ , os quais cumpram  $(x, y) = \lambda(1, 1) + \mu(-1, 1)$ , isto é,  $\lambda$  e  $\mu$  são as variáveis do sistema:

$$\begin{cases} \lambda - \mu = x \\ \lambda + \mu = y \end{cases},$$

cujas soluções são  $\lambda = (x + y)/2$  e  $\mu = (x - y)/2$ . Dessa forma, temos

$$(x, y) = \frac{x + y}{2}(1, 1) + \frac{x - y}{2}(-1, 1).$$

**Exemplo 2.13.** No espaço  $\mathbb{R}^3$ , as retas que contêm a origem  $\mathbf{0} = (0, 0, 0)$  são os subespaços de dimensão 1, enquanto os planos que contêm  $\mathbf{0}$  são seus subespaços de dimensão 2. Esses são todos os subespaços não triviais de  $\mathbb{R}^3$ .

**Exemplo 2.14.** Seja  $W \subset \mathcal{M}(2, 2)$  o conjunto das matrizes quadradas de ordem 2 que têm traço nulo. Então, todo elemento de  $W$  se escreve como:

$$\begin{bmatrix} a & b \\ c & -a \end{bmatrix}, a, b, c \in \mathbb{R}.$$

Além disso, vale a seguinte igualdade:

$$\begin{bmatrix} a & b \\ c & -a \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Logo,  $W$  é o subespaço de  $\mathcal{M}(2, 2)$  gerado pelo conjunto

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}.$$

Pode-se verificar facilmente que  $\mathcal{B}$  é também LI. Dessa forma,  $\mathcal{B}$  é uma base de  $W$ , donde se infere que  $\dim W = 3$ .

De forma análoga, mostra-se que o conjunto formado pelas matrizes quadradas de ordem  $n$  que têm traço nulo é um subespaço de  $\mathcal{M}(n, n)$ , cuja dimensão é  $n - 1$ .

**Proposição 2.10.** *Seja  $V$  um espaço vetorial de dimensão  $n > 0$ . Então, são verdadeiras as seguintes afirmações:*

- i) *Todo subconjunto finito de  $V$  com mais de  $n$  vetores é LD;*
- ii) *Todo subconjunto LI de  $V$  que possui  $n$  vetores é uma base de  $V$ ;*
- iii) *Todo subespaço  $W$  de  $V$  é de dimensão finita, e cumpre  $\dim W \leq \dim V$ , em que a igualdade ocorre se, e somente se,  $V = W$ ;*
- iv) *Para todo subconjunto LI,  $\{v_1, \dots, v_k\} \subset V$ ,  $1 \leq k < n$ , existem vetores  $v_{k+1}, \dots, v_n$ , tais que  $\{v_1, \dots, v_k, \dots, v_n\}$  constitui uma base de  $V$  (isto é, todo subconjunto LI, de  $V$ , pode ser “completado” de modo a formar uma base).*

*Demonstração.* (i) Seja  $\mathcal{A}$  um subconjunto finito de  $V$  com mais de  $n$  vetores. Se  $\mathcal{A}$  fosse LI, teríamos uma contradição com o Teorema 2.1, pois toda base de  $V$  tem  $n$  elementos e gera  $V$ . Logo,  $\mathcal{A}$  é LD.

(ii) Consideremos um conjunto LI,  $\mathcal{B} = \{v_1, \dots, v_n\} \subset V$ . Se  $\mathcal{B}$  não gerasse  $V$ , existiria um vetor  $v \in V$ , tal que  $v \notin \text{Ger } \mathcal{B}$ . Pelo Corolário 2.1, o conjunto  $\{v_1, \dots, v_n, v\}$  seria LI, contradizendo (i). Dessa forma,  $\mathcal{B}$  gera  $V$  e, portanto, é uma base desse espaço.

(iii) Se o subespaço  $W$  não fosse de dimensão finita, ou tivesse dimensão finita e maior que  $\dim V$ , existiria um conjunto LI,  $\{w_1, \dots, w_m\} \subset W$ , tal que  $m > n$ , o que contradiz (i). Assim,  $W$  tem dimensão finita e  $\dim W \leq \dim V$ . Ocorrendo a igualdade, tem-se, por (ii), que toda base de  $W$  é também uma base de  $V$ , implicando que  $V = W$ . A recíproca é imediata.

(iv) O conjunto  $\{v_1, \dots, v_k\} \subset V$ , sendo LI, não pode gerar  $V$ , senão teríamos uma base de  $V$  com menos de  $n$  vetores. Logo, existe  $v_{k+1} \in V - \text{Ger}\{v_1, \dots, v_k\}$ , donde  $\{v_1, \dots, v_k, v_{k+1}\}$  é LI. Procedendo-se indutivamente, obtém-se um conjunto LI,  $\{v_1, \dots, v_k, \dots, v_n\}$ , o qual, por (ii), constitui uma base de  $V$ . ■

**Exemplo 2.15.** Consideremos  $\mathcal{B} = \{(1, -1, 1), (0, 1, 2), (1, 0, -2)\} \subset \mathbb{R}^3$  e denotemos por  $A$  a matriz cujas colunas são os vetores de  $\mathcal{B}$ . Resolvendo-se o sistema linear homogêneo associado,  $AX = \mathbf{0}$ , verifica-se que o mesmo admite apenas a solução trivial, donde se infere que  $\mathcal{B}$  é LI. Uma vez que a dimensão de  $\mathbb{R}^3$  é 3, segue-se da Proposição 2.10-(ii) que  $\mathcal{B}$  é uma base de  $\mathbb{R}^3$ .

Adotemos a seguinte notação: Um vetor  $v$  de um espaço vetorial  $V$ , cujas coordenadas com respeito a uma base  $\mathcal{B} = \{v_1, \dots, v_n\}$  de  $V$  são  $x_1, \dots, x_n$ , será indicado por  $(x_1, \dots, x_n)_{\mathcal{B}}$ , isto é, valem, por definição, as igualdades:

$$v = (x_1, \dots, x_n)_{\mathcal{B}} = x_1v_1 + \dots + x_nv_n.$$

Como vimos fazendo, omitiremos a indicação de  $\mathcal{B}$  quando se tratar da base canônica de  $\mathbb{R}^n$ .

**Exemplo 2.16.** Constatamos no Exemplo 2.12 que um vetor  $v$  de  $\mathbb{R}^2$ , cujas coordenadas com respeito à base canônica são  $x$  e  $y$ , tem  $\frac{x+y}{2}$  e  $\frac{x-y}{2}$  como coordenadas com respeito à base  $\mathcal{B} = \{(1, 1), (-1, 1)\}$ . Nesse caso, escrevemos

$$v = (x, y) = \left( \frac{x+y}{2}, \frac{x-y}{2} \right)_{\mathcal{B}}.$$

Nos será conveniente, também, associarmos a  $v$  uma matriz coluna  $[v]_{\mathcal{B}}$ , cujas entradas são suas coordenadas com respeito à base  $\mathcal{B}$ , isto é,

$$[v]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

No caso em que  $V = \mathbb{R}^n$  e  $\mathcal{B}$  é a sua base canônica, denotamos  $[v]_{\mathcal{B}}$  simplesmente por  $[v]$  (vide Exercício 7 – Cap. 1).

Segue-se diretamente da unicidade das coordenadas de um vetor com respeito a uma base, bem como das propriedades da adição e multiplicação por escalar em  $V$ , que a correspondência

$$v \in V \leftrightarrow [v]_{\mathcal{B}} \in \mathcal{M}(n, 1)$$

é biunívoca, e cumpre as seguintes igualdades:

- $[u + v]_{\mathcal{B}} = [u]_{\mathcal{B}} + [v]_{\mathcal{B}}$ ;
- $[\lambda u]_{\mathcal{B}} = \lambda[u]_{\mathcal{B}}$ .

## 2.4 O Posto de uma Matriz

Nesta seção, temos como objetivo desenvolver um método que nos permita extrair uma base de um conjunto de geradores de um subespaço  $W$  de um espaço de dimensão finita,  $V$ . Esse processo nos conduzirá naturalmente ao conceito de *posto* de uma matriz. Uma vez que usaremos coordenadas, nos limitaremos ao caso  $V = \mathbb{R}^n$ .

Dada uma matriz  $A = (a_{ij})_{m \times n}$ , o subespaço  $\mathcal{L}_A \subset \mathbb{R}^n$ , gerado pelos vetores linha de  $A$ , é dito o *espaço linha* de  $A$ , enquanto o subespaço  $\mathcal{C}_A \subset \mathbb{R}^m$ , gerado pelos vetores coluna de  $A$ , é chamado de *espaço coluna* de  $A$ .

Consideremos, por exemplo, a matriz

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

Seus espaços linha e coluna são  $\mathcal{L}_A = \text{Ger} \{(1, 2, -1), (2, -1, 1)\} \subset \mathbb{R}^3$  e  $\mathcal{C}_A = \text{Ger} \{(1, 2), (2, -1), (-1, 1)\} \subset \mathbb{R}^2$ . Note que,

$$\dim \mathcal{L}_A = \dim \mathcal{C}_A = 2.$$

No caso da matriz

$$B = \begin{bmatrix} 2 & -4 \\ 1 & -2 \\ 3 & -6 \end{bmatrix},$$

temos que  $\mathcal{L}_B = \text{Ger} \{(2, -4), (1, -2), (3, -6)\} = \text{Ger} \{(1, -2)\} \subset \mathbb{R}^2$ , enquanto  $\mathcal{C}_B = \text{Ger} \{(2, 1, 3), (-4, -2, -6)\} = \text{Ger} \{(2, 1, 3)\} \subset \mathbb{R}^3$ . Em particular,  $\dim \mathcal{L}_B = \dim \mathcal{C}_B = 1$ .

Os exemplos acima sugerem o resultado seguinte, cuja peculiaridade reside no fato de os espaços linha e coluna de uma matriz  $m \times n$ , com  $m \neq n$ , serem subespaços de espaços euclidianos distintos.

**Proposição 2.11.** *O espaço linha e o espaço coluna de uma matriz arbitrária  $A$  têm mesma dimensão, a qual chama-se posto de  $A$ .*

*Demonstração.* Consideremos uma matriz  $A = (a_{ij})_{m \times n}$ , e designemos os seus vetores linha e coluna, respectivamente, por  $u_1, \dots, u_m$  e  $v_1, \dots, v_n$ , isto é,

$$A = \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} = [v_1 \ \cdots \ v_n].$$

Tomemos uma base  $\mathcal{B} = \{w_1, \dots, w_k\}$ , de  $\mathcal{L}_A$ , e escrevamos

$$w_p = (\lambda_{p1}, \dots, \lambda_{pn}) = \sum_{j=1}^n \lambda_{pj} e_j, \quad p = 1, \dots, k, \quad (2.3)$$

em que  $\{e_1, \dots, e_n\}$  é a base canônica de  $\mathbb{R}^n$ . Designemos por  $\mu_{ip}$ ,  $p = 1, \dots, k$ , as coordenadas de  $u_i$  com respeito a  $\mathcal{B}$ . Assim, temos que,

$$u_i = \sum_{p=1}^k \mu_{ip} w_p = \sum_{p=1}^k \mu_{ip} \left( \sum_{j=1}^n \lambda_{pj} e_j \right) = \sum_{j=1}^n \left( \sum_{p=1}^k \mu_{ip} \lambda_{pj} \right) e_j. \quad (2.4)$$

Uma vez que

$$u_i = (a_{i1}, \dots, a_{in}) = \sum_{j=1}^n a_{ij} e_j,$$

segue-se da unicidade das coordenadas de  $u_i$  com respeito à base canônica, que

$$a_{ij} = \sum_{p=1}^k \mu_{ip} \lambda_{pj} \quad \forall i \in \{1, \dots, m\}, \quad j \in \{1, \dots, n\}. \quad (2.5)$$

Definindo-se, então, para cada  $p \in \{1, \dots, k\}$ , o vetor

$$w'_p = \sum_{i=1}^m \mu_{ip} e_i,$$

a igualdade (2.5) nos dá

$$v_j = \sum_{p=1}^k \lambda_{pj} w'_p.$$

Logo,  $\mathcal{C}_A \subset \text{Ger} \{w'_1, \dots, w'_k\}$ , donde

$$\dim \mathcal{C}_A \leq k = \dim \mathcal{L}_A. \quad (2.6)$$

Finalmente, aplicando-se (2.6) à transposta  $A^*$  da matriz  $A$ , obtém-se

$$\dim \mathcal{C}_A = \dim \mathcal{L}_{A^*} \geq \dim \mathcal{C}_{A^*} = \dim \mathcal{L}_A,$$

o que implica  $\dim \mathcal{L}_A = \dim \mathcal{C}_A$ , como desejado. ■

**Proposição 2.12.** *Os vetores linha não nulos de uma matriz escalonada são linearmente independentes. Em particular, eles formam uma base do espaço linha dessa matriz.*

*Demonstração.* De fato, sejam  $u_1, \dots, u_r \subset \mathbb{R}^n$  os vetores linha não nulos de uma matriz escalonada  $A = (a_{ij})_{m \times n}$ . Nesse caso, se  $a_{1k}$ ,  $1 \leq k \leq n$ , é a primeira coordenada não nula de  $u_1$ , então a  $k$ -ésima coordenada dos demais vetores,  $u_2, \dots, u_r$ , é igual a zero. Assim,  $u_1$  não é combinação linear de  $u_2, \dots, u_r$ . Analogamente,  $u_2$  não é combinação linear de  $u_3, \dots, u_r$  e assim por diante, isto é, considerando-se os vetores linha de  $A$  na ordem

$$u_r, u_{r-1}, \dots, u_1,$$

tem-se que nenhum deles se escreve como combinação linear dos anteriores. O resultado segue-se, então, da Proposição 2.9. ■

**Proposição 2.13.** *Sejam  $A = (a_{ij})_{m \times n}$  e  $B = (b_{ij})_{m \times n}$  matrizes linha equivalentes. Então,  $\mathcal{L}_A = \mathcal{L}_B$ . Em particular, o posto de  $A$  é igual ao de  $B$ .*

*Demonstração.* Seja  $L = \{u_1, \dots, u_m\} \subset \mathbb{R}^n$  o conjunto formado pelas linhas não nulas de  $A$ . Evidentemente, a permutação de linhas, ou a multiplicação de uma linha por um escalar não nulo, não altera o espaço  $\mathcal{L}_A = \text{Ger } L$ . Consideremos, então, o conjunto  $L' = \{u_1, \dots, u_{j-1}, u_j + \lambda u_i, u_{j+1}, \dots, u_m\}$ , obtido de  $L$  substituindo-se o seu  $j$ -ésimo vetor  $u_j$  por  $u_j + \lambda u_i$ , em que  $\lambda$  é um escalar não nulo, e  $1 \leq i \neq j \leq m$ . Devemos, dessa forma, mostrar que

$\text{Ger } L' = \text{Ger } L$ . Para tanto, observemos que, dado  $v \in \text{Ger } L$ , existem escalares  $\mu_1, \dots, \mu_m$ , tais que

$$\begin{aligned} v &= \mu_1 u_1 + \dots + \mu_i u_i + \dots + \mu_j u_j + \dots + \mu_m u_m \\ &= \mu_1 u_1 + \dots + (\mu_i + \lambda \mu_j - \lambda \mu_j) u_i + \dots + \mu_j u_j + \dots + \mu_m u_m \\ &= \mu_1 u_1 + \dots + (\mu_i - \lambda \mu_j) u_i + \dots + \mu_j (u_j + \lambda u_i) + \dots + \mu_m u_m, \end{aligned}$$

donde  $v \in \text{Ger } L'$ , isto é,  $\text{Ger } L \subset \text{Ger } L'$ .

Fazendo-se, agora,  $u'_j = u_j + \lambda u_i$ , tem-se que  $L$  e  $L'$  se escrevem como:

- $L' = \{u_1, \dots, u_{j-1}, u'_j, u_{j+1}, \dots, u_m\}$ ;
- $L = \{u_1, \dots, u_i, \dots, u'_j - \lambda u_i, \dots, u_m\}$ .

Logo, os argumentos acima se aplicam igualmente, implicando que  $\text{Ger } L' \subset \text{Ger } L$  e, portanto, que  $\text{Ger } L' = \text{Ger } L$ , como queríamos demonstrar. ■

**Exemplo 2.17.** Determinemos, através dos resultados acima, uma base para o subespaço  $W$ , de  $\mathbb{R}^4$ , gerado pelo seguinte conjunto:

$$L = \{(1, 0, 2, -1), (3, -2, 1, 1), (4, -2, 3, 0), (2, -2, -1, 2)\}.$$

Para tanto, basta considerarmos a matriz  $A$  cujas linhas são os vetores de  $L$ , e verificar que a mesma é linha equivalente à matriz escalonada

$$E = \begin{bmatrix} 1 & 0 & 2 & -1 \\ 0 & -2 & -5 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Logo,  $\mathcal{B} = \{(1, 0, 2, -1), (0, -2, -5, 4)\}$  é uma base de  $\mathcal{L}_E = \mathcal{L}_A = W$ . Em particular,  $\dim W = 2$ .

Vejam, na proposição seguinte, como caracterizar matrizes invertíveis através do conceito de posto.

**Proposição 2.14.** *São equivalentes as seguintes afirmações acerca de uma matriz quadrada  $A \in \mathcal{M}(n, n)$ :*

- i)  $A$  é invertível;
- ii) Para toda matriz  $B \in \mathcal{M}(n, 1)$ , o sistema linear  $AX = B$  admite uma única solução;
- iii)  $A$  tem posto  $n$ .

*Demonstração.* (i)  $\Rightarrow$  (ii): Dada  $B \in \mathcal{M}(n, 1)$ , sendo  $A$  invertível, temos que  $X = A^{-1}B$  é, claramente, a única solução de  $AX = B$ .

(ii)  $\Rightarrow$  (i): Seja  $N = [w_1 \cdots w_n] \in \mathcal{M}(n)$  uma matriz arbitrária. Por hipótese, para cada  $i = 1, \dots, n$ , existe um único  $v_i \in \mathbb{R}^n$ , tal que  $A[v_i] = [w_i]$ . Logo,  $M = [v_1 \cdots v_n]$  é a única matriz de  $\mathcal{M}(n)$  que satisfaz a igualdade  $AM = N$ . Dito de outra forma, a equação matricial

$$AY = N, \quad N \in \mathcal{M}(n), \quad (2.7)$$

admite sempre uma única solução  $Y \in \mathcal{M}(n)$ . Em particular, se  $N = A$ , então  $Y = I$ . Agora, fazendo-se  $N = I$ , temos que existe  $M \in \mathcal{M}(n, n)$ , tal que  $AM = I$ . Escrevendo-se, então,  $I' = MA$ , e multiplicando-se à esquerda por  $A$ , obtém-se  $AI' = A$ , implicando que  $I' = I$ . Logo,  $AM = MA = I$ , donde  $A$  é invertível e  $M = A^{-1}$ .

(ii)  $\Rightarrow$  (iii): Por hipótese, o sistema  $AX = \mathbf{0}$  admite apenas a solução trivial. Logo, pela Proposição 2.6, os  $n$  vetores coluna de  $A$  são LI, donde se infere que o posto de  $A$  é  $n$ .

(iii)  $\Rightarrow$  (ii): Escrevendo-se  $A = [v_1 \cdots v_n]$ , tem-se que  $\mathcal{B} = \{v_1, \dots, v_n\}$  é uma base de  $\mathbb{R}^n$ , já que, por hipótese,  $A$  tem posto  $n$ . Dada  $B \in \mathcal{M}(n, 1)$ , seja  $w \in \mathbb{R}^n$ , tal que  $B = [w]$ . Fazendo-se, então,  $C = [w]_{\mathcal{B}}$ , tem-se, pelas considerações da Observação 2.1, que  $C$  é a única solução do sistema  $AX = B$ . ■

## 2.5 Somas de Subespaços

Sejam  $V$  um espaço vetorial e  $U, W \subset V$  subespaços de  $V$ . Definimos a soma de  $U$  e  $W$ ,  $U + W$ , como o subconjunto de  $V$  formado por todas as somas



do tipo  $u + w$ , em que  $u \in U$  e  $w \in W$ , isto é,

$$U + W = \{v \in V; v = u + w, u \in U, w \in W\}.$$

No caso em que  $U \cap W = \{\mathbf{0}\}$ , dizemos que essa soma é *direta* e a denotamos por  $U \oplus W$ .

**Proposição 2.15.** *Sejam  $U$  e  $W$  subespaços de um espaço vetorial  $V$ . Então,  $U + W$  é um subespaço de  $V$ .*

*Demonstração.* Temos que  $\mathbf{0} \in U \cap W$ . Logo,  $\mathbf{0} = \mathbf{0} + \mathbf{0} \in U + W$ . Tomemos  $v, v' \in U + W$ . Então, existem  $u, u' \in U$  e  $w, w' \in W$ , tais que  $v = u + w$  e  $v' = u' + w'$ . Uma vez que  $U$  e  $W$  são subespaços de  $V$ , temos que  $u + u' \in U$ , e  $w + w' \in W$ . Além disso, dado  $\lambda \in \mathbb{R}$ , tem-se  $\lambda u \in U$  e  $\lambda w \in W$ . Logo,  $v + v' = (u + u') + (w + w') \in U + W$  e  $\lambda v = \lambda u + \lambda w \in U + W$ , donde se conclui que  $U + W$  é um subespaço vetorial de  $V$ . ■

**Exemplo 2.18.** Sejam  $U = \text{Ger}\{(1, 0)\}$  e  $W = \text{Ger}\{(1, 1)\}$  os subespaços de  $\mathbb{R}^2$  gerados, respectivamente, por  $(1, 0)$  e  $(1, 1)$ . Claramente,  $\{(1, 0), (1, 1)\}$  é uma base de  $\mathbb{R}^2$ . Logo, cada  $v \in \mathbb{R}^2$  se escreve como  $v = \lambda(1, 0) + \mu(1, 1)$ . Fazendo-se  $\lambda(1, 0) = u \in U$  e  $\mu(1, 1) = w \in W$ , temos que  $v \in U + W$ . Além disso,  $U \cap W = \{\mathbf{0}\}$ . Logo,  $\mathbb{R}^2 = U \oplus W$ .

**Exemplo 2.19.** Considere, em  $\mathbb{R}^3$ , os vetores  $v_1 = (1, 0, 0)$ ,  $v_2 = (0, 1, 0)$  e  $v_3 = (1, 1, -1)$ , juntamente com os subespaços  $U = \text{Ger}\{v_1, v_2\}$  e  $W = \text{Ger}\{v_1, v_3\}$ . Pode-se verificar facilmente que  $\mathbb{R}^3 = U + W$  e que  $U \cap W = \text{Ger}\{v_1\} \neq \{\mathbf{0}\}$ . Em particular, essa soma não é direta.

**Proposição 2.16.** *Sejam  $U$  e  $W$  subespaços de um espaço vetorial  $V$ . Então, se  $U \cap W = \{\mathbf{0}\}$ , são verdadeiras as seguintes afirmações:*

- i) *Para todo  $v \in U \oplus W$ , existem únicos vetores  $u \in U$  e  $w \in W$ , tais que  $v = u + w$ ;*
- ii)  *$\dim(U \oplus W) = \dim U + \dim W$ , se  $V$  tiver dimensão finita.*

*Demonstração.* Dado  $v \in U \oplus W$ , suponhamos que

$$v = u + w \quad \text{e} \quad v = u' + w', \quad u, u' \in U, \quad w, w' \in W.$$

Daí, temos que  $u + w = u' + w'$ , isto é,  $u - u' = w' - w$ . Uma vez que  $U$  e  $W$  são subespaços de  $V$ , devemos ter  $u - u' \in U$  e  $w' - w \in W$ . Porém  $U \cap W = \{\mathbf{0}\}$ . Logo,  $u - u' = w' - w = \mathbf{0}$ , o que nos dá  $u = u'$  e  $w = w'$ , provando, assim, a asserção (i).

Quanto a (ii), tomemos bases  $\mathcal{A} = \{u_1, \dots, u_m\}$  e  $\mathcal{B} = \{w_1, \dots, w_n\}$ , de  $U$  e  $W$ , respectivamente. Provemos, então, que  $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$  é uma base de  $U \oplus W$ , donde se seguirá o resultado.

É imediato que  $\text{Ger } \mathcal{C} \subset U \oplus W$ . Para a inclusão contrária, consideremos  $u \in U$  e  $w \in W$ , e os escalares  $\lambda_1, \dots, \lambda_m$  e  $\mu_1, \dots, \mu_n$ , tais que

$$u = \sum_{i=1}^m \lambda_i u_i \quad \text{e} \quad w = \sum_{j=1}^n \mu_j w_j.$$

Assim, temos que

$$u + w = \sum_{i,j=1}^{m,n} (\lambda_i u_i + \mu_j w_j) \in \text{Ger } \mathcal{C},$$

ou seja,  $U \oplus W \subset \text{Ger } \mathcal{C}$ . Dessa forma,  $U \oplus W = \text{Ger } \mathcal{C}$ . Finalmente, tomando-se uma combinação linear nula dos vetores de  $\mathcal{C}$ ,

$$\lambda_1 u_1 + \dots + \lambda_m u_m + \mu_1 w_1 + \dots + \mu_n w_n = \mathbf{0},$$

devemos ter, pelo resultado do item (i),

$$\lambda_1 u_1 + \dots + \lambda_m u_m = \mathbf{0} \quad \text{e} \quad \mu_1 w_1 + \dots + \mu_n w_n = \mathbf{0},$$

pois  $\lambda_1 u_1 + \dots + \lambda_m u_m \in U$  e  $\mu_1 w_1 + \dots + \mu_n w_n \in W$ . Porém  $\{u_1, \dots, u_m\}$  e  $\{w_1, \dots, w_n\}$  são bases. Em particular, estes conjuntos são LI. Logo,

$$\lambda_1 = \dots = \lambda_m = \mu_1 = \dots = \mu_n = 0.$$

Dessa forma,  $\mathcal{C}$  é LI e, portanto, constitui uma base de  $U \oplus W$ . ■

## Exercícios

### Seção 2.1

1. Mostre a unicidade do vetor nulo de um espaço vetorial arbitrário.
2. Verifique, em cada um dos itens abaixo, se o conjunto  $W \subset V$  é um subespaço do espaço vetorial  $V$ .
  - a)  $W = \{(x, y) \in \mathbb{R}^2; x = -y\}$ ,  $V = \mathbb{R}^2$ ;
  - b)  $W = \{(x, y) \in \mathbb{R}^2; y = x^2\}$ ,  $V = \mathbb{R}^2$ ;
  - c)  $W = \{(x, y, z) \in \mathbb{R}^3; x + y + z \geq 0\}$ ,  $V = \mathbb{R}^3$ ;
  - d)  $W = \{(x, y, z) \in \mathbb{R}^3; z = 0\}$ ,  $V = \mathbb{R}^3$ ;
  - e)  $W = \{A \in \mathcal{M}(n); A \text{ é diagonal}\}$ ,  $V = \mathcal{M}(n)$ ;
  - f)  $W = \{X \in \mathcal{M}(n, 1); AX = \mathbf{0}, A \in \mathcal{M}(m, n)\}$ ,  $V = \mathcal{M}(n, 1)$ ;
  - g)  $W = \{p \in \mathcal{P}[\mathbb{R}, t]; p = \mathbf{0} \text{ ou grau}(p) \leq n, n \in \mathbb{N}\}$ ,  $V = \mathcal{P}[\mathbb{R}, t]$ ;
  - h)  $W = \{f \in \mathcal{F}(\mathbb{R}, \mathbb{R}); f(x) = f(-x)\}$ ,  $V = \mathcal{F}(\mathbb{R}, \mathbb{R})$ .
  - i)  $W = \{f \in \mathcal{F}(\mathbb{R}, \mathbb{R}); f(x) = -f(-x)\}$ ,  $V = \mathcal{F}(\mathbb{R}, \mathbb{R})$ .
3. Sejam  $U$  e  $W$  subespaços de um espaço vetorial  $V$ . Mostre que  $U \cup W$  é um subespaço de  $V$  se, e somente se,  $U \subset W$  ou  $W \subset U$ .
4. Determine se, em  $\mathbb{R}^4$ , o vetor  $(4, -2, 5, -1)$  está no subespaço gerado pelos vetores  $(1, -1, 0, 1)$ ,  $(0, 1, 1, -1)$  e  $(1, -1, 1, 0)$ .
5. Seja  $\mathcal{A}$  um subconjunto arbitrário (não necessariamente finito) de um espaço vetorial  $V$ . Mostre que o conjunto de todas as combinações lineares finitas de vetores de  $\mathcal{A}$  é um subespaço de  $V$ , o qual denotamos, como no caso finito, por  $\text{Ger } \mathcal{A}$ . Mostre que  $\text{Ger } \mathcal{A}$  coincide com a interseção de todos os subespaços de  $V$  que contêm  $\mathcal{A}$ .
6. Dada uma matriz  $A \in \mathcal{M}(m, n)$ , prove que o conjunto solução do sistema linear homogêneo  $AX = \mathbf{0}$  é um subespaço vetorial de  $\mathbb{R}^n$ .

7. Encontre um sistema linear homogêneo cujo conjunto solução seja o subespaço de  $\mathbb{R}^4$  gerado pelos vetores  $(-1, 0, 1, 2)$ ,  $(3, 4, -2, 5)$  e  $(1, 4, 0, 9)$ .
8. Prove a recíproca da Proposição 2.4.

### Seção 2.2

9. Determine três vetores linearmente dependentes em  $\mathbb{R}^3$ , tais que dois quaisquer deles sejam linearmente independentes.
10. Para que valores de  $\lambda$  os vetores  $(\lambda, 1, 0)$ ,  $(1, \lambda, 1)$  e  $(0, 1, \lambda)$  são linearmente dependentes em  $\mathbb{R}^3$ ?
11. Demonstre a Proposição 2.6 com base na Observação 2.1.
12. Sejam  $u, v, w$  vetores LI de um espaço vetorial  $V$ . Determine se o conjunto  $\{u + v, u + w, v + w\}$  é LI ou LD.

### Seção 2.3

13. Em cada item abaixo, encontre uma base para o subespaço  $W \subset \mathbb{R}^n$ .
  - a)  $W = \text{Ger} \{(1, -4, 3), (3, -14, 17), (0, -2, 5)\} \subset \mathbb{R}^3$ ;
  - b)  $W = \text{Ger} \{(1, 2, 3, 4), (5, 6, 7, 8), (9, 10, 11, 12)\} \subset \mathbb{R}^4$ ;
14. Mostre que cada um dos seguintes conjuntos é um subespaço de  $\mathcal{M}(n)$  e, em seguida, determine sua dimensão.
  - a)  $\{A \in \mathcal{M}(n); A \text{ é triangular superior}\}$ ;
  - b)  $\{A \in \mathcal{M}(n); A \text{ é triangular inferior}\}$ ;
  - c)  $\{A \in \mathcal{M}(n); A \text{ é simétrica}\}$ ;
  - d)  $\{A \in \mathcal{M}(n); A \text{ é antissimétrica}\}$ ;
  - e)  $\{A \in \mathcal{M}(n); A \text{ tem traço nulo}\}$ .
15. Exiba bases de  $\mathcal{M}(2)$ , bem como de  $\mathcal{M}(3)$ , cujas matrizes sejam todas invertíveis.

**Seção 2.4**

16. Prove que  $\mathcal{L}_A = \mathcal{C}_A$ , em que  $A$  é a matriz

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

17. Seja  $A = (a_{ij})$  uma matriz quadrada de ordem  $n \geq 3$ . Suponha que existam reais não nulos  $\lambda, \mu$ , tais que

$$a_{ij} = \lambda i + \mu j \quad \forall 1 \leq i, j \leq n.$$

Prove que  $A$  não é invertível e conclua que vale o mesmo para a matriz

$$\begin{bmatrix} 11 & 12 & 13 & 14 \\ 21 & 22 & 23 & 24 \\ 31 & 32 & 33 & 34 \\ 41 & 42 & 43 & 44 \end{bmatrix}.$$

18. Seja  $A$  uma matriz  $5 \times 5$ , tal que seus vetores linha  $u_i, i = 1, \dots, 5$ , satisfazem a seguinte relação:

$$u_1 + u_2 - 2u_4 + 3u_5 = \mathbf{0}.$$

Encontre uma matriz  $C$ , de posto 3, tal que  $CA$  tenha linhas  $u_1, u_4$  e  $\mathbf{0}$ .

19. Dada uma matriz  $A = (a_{ij})_{2 \times 2}$ , prove que todo sistema linear  $AX = B$  admite uma única solução se, e somente se,  $a_{11}a_{22} - a_{12}a_{21} \neq 0$ .

**Seção 2.5**

20. Prove que, se  $U$  e  $W$  são subespaços de um espaço vetorial  $V$  de dimensão finita, então  $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$ .
21. Seja  $\mathcal{B} = \{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$  uma base de um espaço vetorial  $V$ . Defina,  $U = \text{Ger}(v_1, \dots, v_k)$  e  $W = \text{Ger}(v_{k+1}, \dots, v_n)$ , e mostre que  $V = U \oplus W$ .
22. Sejam  $W_p$  e  $W_i$  os subespaços de  $\mathcal{F}(\mathbb{R}, \mathbb{R})$  definidos nos itens (h) e (i) do Exercício 2. Prove que  $\mathcal{F}(\mathbb{R}, \mathbb{R}) = W_p \oplus W_i$ .



# 3

---

## *Transformações Lineares*

No capítulo anterior, vimos que os espaços vetoriais constituem uma determinada estrutura, a qual envolve fundamentalmente as operações de adição de vetores e multiplicação de escalar por vetor. Neste, vamos estudar as funções entre espaços vetoriais, ditas *transformações lineares*, as quais preservam essas operações.

Através das transformações lineares, estabelece-se uma relação de equivalência entre espaços vetoriais, dita *isomorfismo*. Constataremos, então, que dois espaços vetoriais de dimensão finita são equivalentes (isto é, *isomorfos*) se, e somente se, têm mesma dimensão. Em particular, todo espaço vetorial de dimensão finita é isomorfo a algum espaço euclidiano  $\mathbb{R}^n$ .

A propriedade fundamental de uma transformação linear é a de ser determinada por seus valores numa base, pois, a partir dela, verifica-se a existência de uma íntima relação entre as transformações lineares e as matrizes. Mais especificamente, para cada transformação linear entre espaços de dimensão finita com bases predeterminadas, existe uma única matriz a ela associada. Como consequência, o estudo das transformações lineares reduz-se, em muitos aspectos, ao das matrizes. Esse fato, na verdade, é uma das características mais marcantes da Álgebra Linear, conforme verificaremos neste e nos capítulos subsequentes.

### 3.1 Linearidade de Aplicações

As funções entre espaços vetoriais (reais) são designadas como tal, quando o seu contradomínio é o conjunto dos números reais. Caso contrário, elas são chamadas de *transformações* ou *aplicações*. As transformações lineares entre espaços vetoriais são aquelas que preservam as operações de adição e produto por escalar, conforme a definição seguinte.

**Definição 3.1** (TRANSFORMAÇÃO LINEAR). Sejam  $V$  e  $W$  espaços vetoriais. Uma aplicação  $T : V \rightarrow W$  é dita uma *transformação linear* se, para quaisquer  $u, v \in V$  e  $\lambda \in \mathbb{R}$ , são válidas as seguintes igualdades:

- $T(u + v) = T(u) + T(v)$ ;
- $T(\lambda u) = \lambda T(u)$ .

Uma tal transformação em que  $W = \mathbb{R}$  é dita um *funcional linear* de  $V$ , e transformações lineares de um espaço vetorial nele mesmo são chamadas também de *operadores lineares*.

Segue-se da definição que *toda transformação linear*  $T : V \rightarrow W$  *leva o vetor nulo de*  $V$  *no vetor nulo de*  $W$ . De fato,  $T(\mathbf{0}) = T(0\mathbf{0}) = 0.T(\mathbf{0}) = \mathbf{0}$ .

Dados espaços vetoriais  $V$  e  $W$ , verifica-se facilmente que são lineares a *aplicação identidade* de  $V$  e a *aplicação nula* de  $V$  em  $W$ , definidas por

$$\begin{array}{ccc} I : V & \rightarrow & V \\ v & \mapsto & v \end{array} \quad \text{e} \quad \begin{array}{ccc} T : V & \rightarrow & W \\ v & \mapsto & \mathbf{0} \end{array} .$$

Temos, também, que a *homotetia* de fator  $a \in \mathbb{R}$ , definida por

$$\begin{array}{ccc} T : V & \rightarrow & V \\ v & \mapsto & av \end{array} ,$$

é um operador linear. Com efeito, dados  $u, v \in V$  e  $\lambda \in \mathbb{R}$ , tem-se

- $T(u + v) = a(u + v) = au + av = T(u) + T(v)$ ;
- $T(\lambda u) = a(\lambda u) = \lambda(au) = \lambda T(u)$ .



Observemos que, se  $T : V \rightarrow W$  for uma transformação linear, então

$$T(u + \lambda v) = T(u) + \lambda T(v) \quad \forall u, v \in V, \lambda \in \mathbb{R}. \quad (3.1)$$

Reciprocamente, a validade de (3.1) implica que  $T$  é linear. Com efeito, tomando-se  $\lambda = 1$ , tem-se  $T(u + v) = T(u) + T(v)$  e, tomando-se  $u = \mathbf{0}$ , tem-se  $T(\lambda v) = \lambda T(v)$ . Assim, podemos adotar a condição (3.1) como critério de linearidade de aplicações.

**Exemplo 3.1** (PROJEÇÕES). Suponha que  $U$  e  $W$  sejam subespaços de  $V$ , tais que  $V = U \oplus W$ . Nesse caso, segue-se da Proposição 2.16 que cada  $v \in V$  se escreve de modo único como  $v = u + w$ , em que  $u \in U$  e  $w \in W$ . Assim, ficam bem definidas as *projeções* de  $V$  sobre  $U$  e  $W$ , as quais são definidas, respectivamente, por

$$P_U : \begin{array}{ccc} V & \rightarrow & U \\ u + w & \mapsto & u \end{array} \quad \text{e} \quad P_W : \begin{array}{ccc} V & \rightarrow & W \\ u + w & \mapsto & w \end{array}.$$

Dados  $v = u + w$  e  $v' = u' + w'$  em  $V$ , e  $\lambda \in \mathbb{R}$ , temos que

$$P_U(v + \lambda v') = P_U((u + \lambda u') + (w + \lambda w')) = u + \lambda u' = P_U(v) + \lambda P_U(v'),$$

donde se conclui que  $P_U$  é linear. De modo análogo, verifica-se que a projeção  $P_W$ , igualmente, é linear.

**Exemplo 3.2.** A função  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , definida por  $f(x, y) = 2x + y$ , é um funcional linear, pois, dados  $u = (x_1, y_1)$ ,  $v = (x_2, y_2) \in \mathbb{R}^2$ , e  $\lambda \in \mathbb{R}$ , tem-se

$$\begin{aligned} f(u + \lambda v) &= f(x_1 + \lambda x_2, y_1 + \lambda y_2) \\ &= 2(x_1 + \lambda x_2) + (y_1 + \lambda y_2) \\ &= (2x_1 + y_1) + \lambda(2x_2 + y_2) \\ &= f(u) + \lambda f(v). \end{aligned}$$

**Exemplo 3.3.** Consideremos agora a aplicação  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ , em que

$$T(x, y) = (x - y, x + y, y).$$

Dados  $u = (x_1, y_1)$ ,  $v = (x_2, y_2) \in \mathbb{R}^2$ , tem-se

$$\begin{aligned} T(u + \lambda v) &= T(x_1 + \lambda x_2, y_1 + \lambda y_2) \\ &= ((x_1 + \lambda x_2) - (y_1 + \lambda y_2), (x_1 + \lambda x_2) + (y_1 + \lambda y_2), y_1 + \lambda y_2) \\ &= (x_1 - y_1, x_1 + y_1, y_1) + \lambda(x_2 - y_2, x_2 + y_2, y_2) \\ &= T(u) + \lambda T(v), \end{aligned}$$

donde se infere que  $T$  é uma transformação linear.

**Exemplo 3.4.** A aplicação  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , em que  $T(x, y) = (x^2, xy)$ , não é linear. Com efeito, dados  $u = (x, y) \in \mathbb{R}^2$  e  $\lambda \in \mathbb{R}$ , tem-se

$$T(\lambda u) = T(\lambda x, \lambda y) = (\lambda^2 x^2, \lambda^2 xy) = \lambda^2 T(u).$$

Assim, se  $\lambda$  for diferente de zero e de 1,  $T(\lambda u) \neq \lambda T(u)$ , donde  $T$  não é linear.

**Exemplo 3.5** (LINEARIDADE DO TRAÇO). Conforme verificamos anteriormente, para quaisquer matrizes quadradas  $A, B$  e  $\lambda \in \mathbb{R}$ , valem as seguintes igualdades:

$$\text{traço}(A + B) = \text{traço}(A) + \text{traço}(B) \quad \text{e} \quad \text{traço}(\lambda A) = \lambda \text{traço}(A).$$

Segue-se, então, desse fato que

$$\begin{aligned} f : \mathcal{M}(n) &\rightarrow \mathbb{R} \\ A &\mapsto \text{traço}(A) \end{aligned}$$

é um funcional linear de  $\mathcal{M}(n)$ .

Verifiquemos agora a notável propriedade das transformações lineares, que é a de ser determinada por seus valores numa base arbitrária. Mais especificamente, vale o resultado seguinte.

**Proposição 3.1.** *Sejam  $V$  e  $W$  espaços vetoriais de dimensão finita, e  $\mathcal{B} = \{v_1, \dots, v_n\}$  uma base de  $V$ . Então, dados  $n$  vetores  $w_1, \dots, w_n$  em  $W$ , existe uma única transformação linear  $T : V \rightarrow W$ , tal que*

$$T(v_i) = w_i \quad \forall i = 1, \dots, n. \quad (3.2)$$

*Demonstração.* Provemos, inicialmente, a existência de  $T$ . Dado  $u \in V$ , sejam  $x_1, \dots, x_n$  suas coordenadas com respeito à base  $\mathcal{B}$ , isto é,

$$u = (x_1, \dots, x_n)_{\mathcal{B}} = x_1v_1 + \dots + x_nv_n.$$

Defina, então,  $T : V \rightarrow W$  por

$$T(u) = T(x_1, \dots, x_n)_{\mathcal{B}} = \sum_{i=1}^n x_iw_i.$$

É imediato que  $T(v_i) = w_i \forall i = 1, \dots, n$ . Além disso, para quaisquer  $v = (y_1, \dots, y_n)_{\mathcal{B}} \in V$  e  $\lambda \in \mathbb{R}$ , tem-se

$$\begin{aligned} T(u + \lambda v) &= T(x_1 + \lambda y_1, \dots, x_n + \lambda y_n)_{\mathcal{B}} \\ &= \sum_{i=1}^n (x_i + \lambda y_i)w_i = \sum_{i=1}^n x_iw_i + \lambda \sum_{i=1}^n y_iw_i \\ &= T(u) + \lambda T(v), \end{aligned}$$

donde  $T$  é linear.

Suponhamos, agora, que exista uma transformação linear  $T' : V \rightarrow W$ , tal que  $T'(v_i) = w_i \forall i = 1, \dots, n$ . Nesse caso, para todo  $u = (x_1, \dots, x_n)_{\mathcal{B}}$  in  $V$ , segue-se da linearidade de  $T'$  que

$$T'(u) = T' \left( \sum_{i=1}^n x_i v_i \right) = \sum_{i=1}^n x_i T'(v_i) = \sum_{i=1}^n x_i w_i = T(u),$$

o que prova a unicidade de  $T$  com respeito à igualdade (3.2) e conclui, dessa forma, a demonstração. ■

**Exemplo 3.6.** Consideremos a base canônica de  $\mathbb{R}^3$ ,  $\mathcal{B} = \{e_1, e_2, e_3\}$ , e determinemos a transformação linear  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ , tal que

$$T(e_1) = (-1, 2), \quad T(e_2) = (1, 1) \quad \text{e} \quad T(e_3) = (0, 1).$$

Temos que,

$$\begin{aligned} T(x, y, z) &= T(xe_1 + ye_2 + ze_3) = xT(e_1) + yT(e_2) + zT(e_3) \\ &= x(-1, 2) + y(1, 1) + z(0, 1), \end{aligned}$$

isto é,

$$T(x, y, z) = (-x + y, 2x + y + z) \quad \forall (x, y, z) \in \mathbb{R}^3.$$

## 3.2 Transformações Lineares e Matrizes

O teorema seguinte estabelece a relação fundamental entre transformações lineares e matrizes, a qual mencionamos à introdução deste capítulo. Do mesmo, infere-se que, fixando-se bases em espaços vetoriais de dimensão finita,  $V$  e  $W$ , existe uma correspondência biunívoca entre as transformações lineares de  $V$  em  $W$  e as matrizes reais  $m \times n$ , em que  $n = \dim V$  e  $m = \dim W$ .

**Teorema 3.1** (TRANSFORMAÇÕES LINEARES E MATRIZES). *Sejam  $V$  e  $W$  espaços vetoriais de dimensões  $n$  e  $m$ , respectivamente, e  $\mathcal{B} \subset V$ ,  $\mathcal{B}' \subset W$  bases desses espaços. Então, dada  $A \in \mathcal{M}(m, n)$ , existe uma única transformação linear*

$$T_A : V \rightarrow W \\ v \mapsto T_A(v) ,$$

a qual cumpre a igualdade

$$[T_A(v)]_{\mathcal{B}'} = A[v]_{\mathcal{B}} \quad \forall v \in V. \quad (3.3)$$

*Reciprocamente, para toda transformação linear  $T : V \rightarrow W$ , existe uma única matriz  $A \in \mathcal{M}(m, n)$ , tal que  $T = T_A$ .*

*Demonstração.* Uma vez que as relações

$$v \in V \leftrightarrow [v]_{\mathcal{B}} \in \mathcal{M}(n, 1) \quad \text{e} \quad w \in W \leftrightarrow [w]_{\mathcal{B}'} \in \mathcal{M}(m, 1)$$

são biunívocas, temos que a igualdade (3.3), na verdade, define a aplicação  $T_A$ , isto é, para cada  $v \in V$ , existe um único vetor  $T_A(v) \in W$ , para o qual a igualdade (3.3) se cumpre.

Resta-nos, pois, provar que  $T_A$  é linear. Para tanto, tomemos  $u, v \in V$  e  $\lambda \in \mathbb{R}$ , e observemos que

$$\begin{aligned} [T_A(u + \lambda v)]_{\mathcal{B}'} &= A[u + \lambda v]_{\mathcal{B}} = A([u]_{\mathcal{B}} + [\lambda v]_{\mathcal{B}}) = A[u]_{\mathcal{B}} + \lambda A[v]_{\mathcal{B}} \\ &= [T_A(u)]_{\mathcal{B}'} + \lambda [T_A(v)]_{\mathcal{B}'} = [T_A(u) + \lambda T_A(v)]_{\mathcal{B}'} . \end{aligned}$$

Dessa forma,

$$T_A(u + \lambda v) = T_A(u) + \lambda T_A(v) \quad \forall u, v \in V, \lambda \in \mathbb{R},$$

onde  $T_A$  é linear.

A fim de verificar a unicidade de  $T_A$  com respeito à igualdade (3.3), tomemos uma transformação linear  $T : V \rightarrow W$ , a qual cumpre

$$[T(v)]_{\mathcal{B}'} = A[v]_{\mathcal{B}} \quad \forall v \in V.$$

Nesse caso, para todo  $v \in V$ , temos que  $[T(v)]_{\mathcal{B}'} = [T_A(v)]_{\mathcal{B}'}$  e, portanto,  $T(v) = T_A(v) \forall v \in V$ , isto é,  $T = T_A$ .

Reciprocamente, suponhamos que  $T : V \rightarrow W$  seja uma transformação linear, escrevamos  $\mathcal{B} = \{v_1, \dots, v_n\} \subset V$ ,  $\mathcal{B}' = \{w_1, \dots, w_m\} \subset W$ , e consideremos a matriz  $A = (a_{ij}) \in \mathcal{M}(m, n)$ , cuja  $j$ -ésima coluna é formada pelas coordenadas do vetor  $Tv_j$  com respeito à base  $\mathcal{B}'$ . Mais precisamente, as entradas  $a_{ij}$  são definidas pela igualdade

$$Tv_j = \sum_{i=1}^m a_{ij}w_i = a_{1j}w_1 + \dots + a_{mj}w_m, \quad 1 \leq j \leq n. \quad (3.4)$$

Assim, pelas considerações da Observação 2.1 e pela linearidade de  $T$ , para um dado  $v = x_1v_1 + \dots + x_nv_n \in V$ , tem-se

$$A[v]_{\mathcal{B}} = x_1[T(v_1)]_{\mathcal{B}'} + \dots + x_n[T(v_n)]_{\mathcal{B}'} = [T(x_1v_1 + \dots + x_nv_n)]_{\mathcal{B}'} = [T(v)]_{\mathcal{B}'}$$

Daí, e da unicidade de  $T_A$  com respeito à igualdade (3.3), segue-se que  $T = T_A$ .

Finalmente, suponhamos que exista  $B \in \mathcal{M}(m, n)$ , tal que  $T = T_B$ . Então,  $T_A = T_B$ , donde, para todo  $v \in V$ , tem-se  $A[v]_{\mathcal{B}} = B[v]_{\mathcal{B}}$ . Fazendo-se  $v = v_j$ , segue-se dessa última igualdade que as colunas correspondentes de  $A$  e  $B$  coincidem e, portanto, que  $A = B$ . Isso conclui a demonstração. ■

Nas condições do teorema acima, dizemos que  $A$  é a *matriz de  $T = T_A$  com respeito às bases  $\mathcal{B}$  e  $\mathcal{B}'$* , e a indicamos por  $[T]_{\mathcal{B}'}^{\mathcal{B}}$ . Sendo assim, a igualdade (3.3) assume a seguinte forma

$$[T(v)]_{\mathcal{B}'} = [T]_{\mathcal{B}'}^{\mathcal{B}}[v]_{\mathcal{B}}.$$

No caso particular em que  $V = W$  e  $\mathcal{B} = \mathcal{B}'$ , escrevemos  $[T]_{\mathcal{B}}$ , ao invés de  $[T]_{\mathcal{B}}^{\mathcal{B}}$ . Além disso, omitiremos a indicação das bases  $\mathcal{B}$  e  $\mathcal{B}'$ , isto é, escreveremos simplesmente  $[T]$ , quando  $V$  e  $W$  forem espaços euclidianos, e  $\mathcal{B}$  e  $\mathcal{B}'$  suas respectivas bases canônicas.

**Observação 3.1.** Convém mencionar que, com a notação acima, em coordenadas com respeito à base  $\mathcal{B}'$ ,  $T(v_1), \dots, T(v_n)$  são justamente os vetores coluna de  $[T]_{\mathcal{B}'}$ .

**Exemplo 3.7.** Tomemos, com a notação do Teorema 3.1,  $V = \mathbb{R}^2$ ,  $W = \mathbb{R}^3$ , e suponhamos que  $\mathcal{B} \subset \mathbb{R}^2$  e  $\mathcal{B}' \subset \mathbb{R}^3$  sejam suas respectivas bases canônicas. Nessas condições, a transformação  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ , em que

$$[T] = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix},$$

cumpre, para  $v = (x, y) \in \mathbb{R}^2$ ,

$$[T(v)] = [T][v] = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x - y \\ x + y \\ y \end{bmatrix},$$

donde  $T$  é definida por

$$T(x, y) = (x - y, x + y, y).$$

Note que  $T$  é justamente a transformação linear do Exemplo 3.3.

**Exemplo 3.8.** Dada a transformação linear  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ ,

$$T(x, y, z) = (x + \pi y + ez, \sqrt{2}x - (\log 3)y + \sqrt{3}z),$$

obtenhamos sua matriz  $[T] = (a_{ij})_{2 \times 3}$  com respeito às bases canônicas de  $\mathbb{R}^3$  e  $\mathbb{R}^2$ . Para tanto, observemos que, pela igualdade (3.4), basta obtermos as imagens dos vetores da base canônica de  $\mathbb{R}^3$ , pois os mesmos determinarão as colunas de  $[T]$ . Assim, uma vez que

$$T(1, 0, 0) = (1, \sqrt{2}), \quad T(0, 1, 0) = (\pi, -\log 3), \quad T(0, 0, 1) = (e, \sqrt{3}),$$

temos que  $[T] = \begin{bmatrix} 1 & \pi & e \\ \sqrt{2} & -\log 3 & \sqrt{3} \end{bmatrix}$ .

Convém mencionar que, com a notação introduzida no fim da Seção 2.3, infere-se do Teorema 3.1 que toda transformação linear  $T : V \rightarrow W$  se exprime de maneira única como

$$T(x_1, \dots, x_n)_{\mathcal{B}} = (a_{11}x_1 + \dots + a_{1n}x_n, \dots, a_{m1}x_1 + \dots + a_{mn}x_n)_{\mathcal{B}'},$$

em que  $[T]_{\mathcal{B}}^{\mathcal{B}'} = (a_{ij})_{m \times n}$ .

**Exemplo 3.9.** Tomemos as bases  $\mathcal{B} = \{(1, 0, 1), (0, 1, -1), (1, 2, -3)\} \subset \mathbb{R}^3$  e  $\mathcal{B}' = \{(1, 0), (1, -1)\} \subset \mathbb{R}^2$ , e consideremos a transformação linear  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ , cuja matriz com respeito a essas bases é:

$$[T]_{\mathcal{B}}^{\mathcal{B}'} = \begin{bmatrix} 2 & -1 & 2 \\ 4 & 1 & -2 \end{bmatrix}.$$

Assim, em coordenadas com respeito a  $\mathcal{B}$  e  $\mathcal{B}'$ ,  $T$  se escreve como:

$$T(x, y, z)_{\mathcal{B}} = (2x - y + 2z, 4x + y - 2z)_{\mathcal{B}'}$$

A fim de expressar  $T$  em coordenadas com respeito às bases canônicas de  $\mathbb{R}^3$  e  $\mathbb{R}^2$ , podemos proceder como no Exemplo 2.12 e obter as seguintes relações:

- $(x, y, z) = \left(\frac{x+y+z}{2}, -x + 2y + z, \frac{x-y-z}{2}\right)_{\mathcal{B}}$ .
- $(x, y)_{\mathcal{B}'} = (x + y, -y)$ .

Logo, valem as igualdades:

$$\begin{aligned} T(x, y, z) &= T\left(\frac{x+y+z}{2}, -x + 2y + z, \frac{x-y-z}{2}\right)_{\mathcal{B}} \\ &= (3x - 2y - z, 5y + 4z)_{\mathcal{B}'} \\ &= (3x + 3y + 3z, -5y - 4z). \end{aligned}$$

Em particular,

$$[T] = \begin{bmatrix} 3 & 3 & 3 \\ 0 & -5 & 4 \end{bmatrix}.$$

**Exemplo 3.10.** Seja  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  o operador linear, tal que

$$[T] = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix},$$

isto é,  $T(x, y) = (x + 2y, 4x + 3y)$ . Consideremos a base  $\mathcal{B} = \{(1, -1), (1, 2)\} \subset \mathbb{R}^2$  e determinemos a matriz de  $T$  com respeito a  $\mathcal{B}$ . Para tanto, como no exemplo anterior, buscamos inicialmente a relação entre as coordenadas com respeito à base canônica e as coordenadas com respeito à base  $\mathcal{B}$ , obtendo

$$(x, y)_{\mathcal{B}} = (x + y, -x + 2y) \quad \text{e} \quad (x, y) = \left( \frac{2x - y}{3}, \frac{x + y}{3} \right)_{\mathcal{B}}. \quad (3.5)$$

Logo,

$$T(x, y)_{\mathcal{B}} = T(x + y, -x + 2y) = (-x + 5y, x + 10y) = (-x, 5y)_{\mathcal{B}},$$

donde se infere que

$$[T]_{\mathcal{B}} = \begin{bmatrix} -1 & 0 \\ 0 & 5 \end{bmatrix}.$$

Note que  $\text{traço}[T] = \text{traço}[T]_{\mathcal{B}}$  (vide último parágrafo da próxima seção).

### 3.3 Mudança de Coordenadas

Nos dois últimos exemplos da seção anterior, constatamos que, em coordenadas, uma transformação linear  $T : V \rightarrow W$  pode se expressar de formas distintas, a depender das bases consideradas em  $V$  e  $W$ . No caso do operador do Exemplo 3.10, vê-se ainda que a mudança da base canônica de  $\mathbb{R}^2$  para a base  $\mathcal{B}$  tornou a expressão de  $T$  mais simples, pois sua matriz com respeito a  $\mathcal{B}$  resultou ser diagonal.

Essas considerações nos levam a buscar um método efetivo de se obter as relações entre dois sistemas de coordenadas, quando os mesmos são relativos a bases distintas. Com esse propósito, tomemos bases  $\mathcal{B}, \mathcal{B}'$  de um espaço vetorial  $V$ , e denotemos por  $I$  a aplicação identidade de  $V$ . Nessas condições,



a matriz  $[I]_{\mathcal{B}}^{\mathcal{B}'}$  é dita a *matriz de mudança* da base  $\mathcal{B}$  para a base  $\mathcal{B}'$ . Note que essa nomenclatura deve-se ao fato de que, para todo  $v \in V$ ,

$$[v]_{\mathcal{B}'} = [Iv]_{\mathcal{B}'} = [I]_{\mathcal{B}}^{\mathcal{B}'} [v]_{\mathcal{B}},$$

isto é, a multiplicação de  $[I]_{\mathcal{B}}^{\mathcal{B}'}$  por  $[v]_{\mathcal{B}}$  produz o efeito de mudar as coordenadas de  $v$  da base  $\mathcal{B}$  para a base  $\mathcal{B}'$ .

**Exemplo 3.11.** Considere as bases  $\mathcal{B} = \{(1, 0), (0, 1)\}$  e  $\mathcal{B}' = \{(1, -1), (1, 2)\}$  de  $\mathbb{R}^2$ . Pondo-se

$$(1, 0) = x_1(1, -1) + y_1(1, 2) \quad \text{e} \quad (0, 1) = x_2(1, -1) + y_2(1, 2),$$

obtem-se facilmente  $x_1 = 2/3$ ,  $y_1 = 1/3$  e  $x_2 = -y_2 = -1/3$ . Logo,

$$[I]_{\mathcal{B}}^{\mathcal{B}'} = \begin{bmatrix} 2/3 & -1/3 \\ 1/3 & 1/3 \end{bmatrix}.$$

Assim, dado  $v = (x, y) \in \mathbb{R}^2$ ,

$$[v]_{\mathcal{B}'} = [I]_{\mathcal{B}}^{\mathcal{B}'} [v]_{\mathcal{B}} = \begin{bmatrix} 2/3 & -1/3 \\ 1/3 & 1/3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2x - y \\ x + y \end{bmatrix},$$

isto é (compare com (3.5)),

$$v = (x, y) = \left( \frac{2x - y}{3}, \frac{x + y}{3} \right)_{\mathcal{B}'}$$

**Proposição 3.2** (SEMELHANÇA ENTRE AS MATRIZES DE UM OPERADOR). *Sejam  $\mathcal{B}$  e  $\mathcal{B}'$  bases de um espaço vetorial  $V$  de dimensão finita. Então, para todo operador linear  $T : V \rightarrow V$ , fazendo-se  $M = [I]_{\mathcal{B}}^{\mathcal{B}'}$ , tem-se*

$$[T]_{\mathcal{B}} = M^{-1}[T]_{\mathcal{B}'}M.$$

*Em particular,  $[T]_{\mathcal{B}}$  e  $[T]_{\mathcal{B}'}$  são matrizes semelhantes.*

*Demonstração.* Dado  $v \in V$ , tem-se  $M[v]_{\mathcal{B}} = [v]_{\mathcal{B}'}$  e, portanto,  $M^{-1}[v]_{\mathcal{B}'} = [v]_{\mathcal{B}}$ . Além disso,  $[T(v)]_{\mathcal{B}} = [T]_{\mathcal{B}}[v]_{\mathcal{B}}$  e  $[T(v)]_{\mathcal{B}'} = [T]_{\mathcal{B}'}[v]_{\mathcal{B}'}$ . Logo,

$$(M^{-1}[T]_{\mathcal{B}'}M)[v]_{\mathcal{B}} = M^{-1}[T]_{\mathcal{B}'}[v]_{\mathcal{B}'} = M^{-1}[T(v)]_{\mathcal{B}'} = [T(v)]_{\mathcal{B}} = [T]_{\mathcal{B}}[v]_{\mathcal{B}},$$

donde se infere que  $[T]_{\mathcal{B}} = M^{-1}[T]_{\mathcal{B}'}M$ , como desejado. ■

Como sabemos, matrizes semelhantes têm traços iguais. Assim, a proposição acima nos permite definir o *traço* de um operador linear  $T : V \rightarrow V$  como o traço da matriz  $[T]_{\mathcal{B}}$ , em que  $\mathcal{B}$  é uma base arbitrária de  $V$ .

### 3.4 Núcleo e Imagem

Dada uma transformação linear  $T : V \rightarrow W$ , associamos a ela dois subespaços, um de  $V$ , outro de  $W$ , os quais designamos, respectivamente, por *núcleo* e *imagem* de  $T$ . Mostraremos que o comportamento de uma transformação linear, no que diz respeito a injetividade ou sobrejetividade, é relativo às propriedades desses subespaços.

**Definição 3.2** (NÚCLEO – IMAGEM). Seja  $T : V \rightarrow W$  uma transformação linear entre dois espaços vetoriais  $V$  e  $W$ . O *núcleo* de  $T$ , denotado por  $\text{Nuc } T$ , é o conjunto formado por todos os vetores de  $V$  cuja imagem, por  $T$ , é o vetor nulo de  $W$ . A *imagem* de  $T$ , denotada por  $\text{Im } T$ , é o conjunto dos vetores de  $W$  que são imagem, por  $T$ , dos vetores de  $V$ . Em símbolos,

$$\text{Nuc } T = \{v \in V ; T(v) = \mathbf{0}\} \quad \text{e} \quad \text{Im } T = \{w \in W ; w = T(v), v \in V\}.$$

Dado um conjunto arbitrário  $\Omega \subset V$ , denota-se por  $T(\Omega)$  o conjunto formado por todos os elementos de  $W$  que são imagens de elementos de  $\Omega$ , isto é,  $T(\Omega) = \{w \in W ; w = T(v), v \in \Omega\}$ . Em particular,  $\text{Im } T = T(V)$ .

**Proposição 3.3.** *Dada uma transformação linear  $T : V \rightarrow W$ , os conjuntos  $\text{Nuc } T$  e  $\text{Im } T$  são subespaços de  $V$  e  $W$ , respectivamente.*

*Demonstração.* Sendo  $T$  uma aplicação linear, tem-se  $T(\mathbf{0}) = \mathbf{0}$ . Logo,  $\text{Nuc } T$  e  $\text{Im } T$  contêm os vetores nulos dos respectivos espaços em que estão contidos. Além disso, dados  $u, v \in \text{Nuc } T$  e  $\lambda \in \mathbb{R}$ , tem-se  $T(u + \lambda v) = T(u) + \lambda T(v) = \mathbf{0}$ , donde  $u + \lambda v \in \text{Nuc } T$ . Logo,  $\text{Nuc } T$  é um subespaço de  $V$ .

Consideremos agora  $z, w \in \text{Im } T$  e  $\lambda \in \mathbb{R}$ . Existem, então,  $u, v \in V$  satisfazendo  $T(u) = z$  e  $T(v) = w$ . Assim,  $T(u + \lambda v) = T(u) + \lambda T(v) = z + \lambda w$ , implicando que  $z + \lambda w \in \text{Im } T$  e, portanto, que  $\text{Im } T$  é um subespaço de  $W$ . ■

**Definição 3.3** (POSTO – NULIDADE). Dada uma transformação linear  $T$ , chama-se de *posto* de  $T$  a dimensão de sua imagem, e de *nulidade* de  $T$  a dimensão de seu núcleo.

A proposição seguinte é de grande valia na determinação da imagem de uma transformação linear.

**Proposição 3.4.** *Sejam  $V$  um espaço vetorial e  $\mathcal{B} = \{v_1, \dots, v_n\}$  uma base de  $V$ . Então, se  $T : V \rightarrow W$  é uma transformação linear, tem-se*

$$\text{Im } T = \text{Ger } \{T(v_1), \dots, T(v_n)\}.$$

*Demonstração.* Seja  $w = T(v) \in \text{Im } T \subset W$ , em que  $v = (x_1, \dots, x_n)_{\mathcal{B}} \in V$ . Sendo assim, tem-se

$$w = T(v) = x_1 T(v_1) + \dots + x_n T(v_n) \in \text{Ger } \{T(v_1), \dots, T(v_n)\},$$

ou seja,  $\text{Im } T \subset \text{Ger } \{T(v_1), \dots, T(v_n)\}$ .

Temos que  $\text{Im } T$  é um subespaço de  $W$  e que  $\{T(v_1), \dots, T(v_n)\} \subset \text{Im } T$ . Logo,  $\text{Ger } \{T(v_1), \dots, T(v_n)\} \subset \text{Im } T$  e, portanto,

$$\text{Im } T = \text{Ger } \{T(v_1), \dots, T(v_n)\},$$

como desejávamos provar. ■

**Exemplo 3.12.** Determinemos o núcleo e a imagem da transformação linear  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ , dada por  $T(x, y, z) = (x + y + 2z, x - y - 3z)$ . Para tanto, observemos que a igualdade  $T(x, y, z) = \mathbf{0}$  equivale ao sistema linear

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

Resolvendo-o, concluímos que

$$\text{Nuc } T = \{(x, -5x, 2x) \in \mathbb{R}^3; x \in \mathbb{R}\} = \text{Ger}\{(1, -5, 2)\} \subset \mathbb{R}^3.$$

Já a imagem de  $T$ , pela Proposição 3.4, é o espaço gerado pelos vetores  $T(1, 0, 0) = (1, 1)$ ,  $T(0, 1, 0) = (1, -1)$  e  $T(0, 0, 1) = (2, -3)$ , donde se conclui que  $\text{Im } T = \mathbb{R}^2$ . Em particular,  $T$  é sobrejetiva.

**Exemplo 3.13.** Tomemos agora a transformação linear  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ , definida por  $T(x, y) = (x - 4y, 3x + y, x - 7y)$ . Procedendo-se de forma análoga à do exemplo anterior, conclui-se que  $\text{Nuc } T = \{\mathbf{0}\}$ . Temos ainda que  $T(1, 0) = (1, 3, 1)$  e  $T(0, 1) = (-4, 1, -7)$ . Logo, pela Proposição 3.4, tem-se  $\text{Im } T = \text{Ger}\{(1, 3, 1), (-4, 1, -7)\}$ . Em particular, o posto de  $T$  é 2, pois  $(1, 3, 1)$  e  $(-4, 1, -7)$  são, claramente, LI.

**Exemplo 3.14.** Seja  $f : \mathcal{M}(n, n) \rightarrow \mathbb{R}$ ,  $n > 1$ , o funcional linear traço, introduzido no Exemplo 3.5. O núcleo de  $f$  é, dessa forma, composto pelas matrizes de  $\mathcal{M}(n, n)$  que têm traço nulo. Portanto,  $\dim(\text{Nuc } f) = n - 1$  (vide Exemplo 2.14). Temos, também, que  $f$  é sobrejetivo, pois, dado  $a \in \mathbb{R}$ ,  $a \neq 0$ , a matriz  $A \in \mathcal{M}(n, n)$  cujas entradas  $a_{ij}$  são todas nulas, exceto por  $a_{11} = a$ , é tal que  $f(A) = a$ .

**Proposição 3.5** (INJETIVIDADE DE TRANSFORMAÇÕES LINEARES). *São equivalentes as seguintes afirmações a respeito de uma transformação linear  $T : V \rightarrow W$ :*

- i)  $T$  é injetiva;
- ii)  $\text{Nuc } T = \{\mathbf{0}\}$ ;
- iii)  $T$  leva vetores LI de  $V$  em vetores LI de  $W$ .

*No caso em que  $V$  tem dimensão finita, cada uma das afirmações acima é equivalente a:*

- iv) *Existe uma base  $\mathcal{B} = \{v_1, \dots, v_n\}$  de  $V$ , tal que  $\{T(v_1), \dots, T(v_n)\}$  é LI em  $W$ .*

*Demonstração.* (i)  $\Rightarrow$  (ii). Seja  $v \in \text{Nuc } T$ . Então  $T(v) = \mathbf{0} = T(\mathbf{0})$ . Uma vez que estamos supondo que  $T$  é injetiva, devemos ter  $v = \mathbf{0}$ , isto é,  $\text{Nuc } T = \{\mathbf{0}\}$ .

(ii)  $\Rightarrow$  (iii). Tomemos vetores  $v_1, \dots, v_n \in V$  e suponhamos que eles sejam LI. Devemos provar que  $T(v_1), \dots, T(v_n)$  são LI em  $W$ . Para tanto, tomemos uma combinação linear nula desses vetores,  $\lambda_1 T(v_1) + \dots + \lambda_n T(v_n) = \mathbf{0}$ . Daí e da linearidade de  $T$ , tem-se  $T(\lambda_1 v_1 + \dots + \lambda_n v_n) = \mathbf{0}$ , donde  $v = \lambda_1 v_1 + \dots + \lambda_n v_n \in$

$\text{Nuc } T$ . Pela hipótese, devemos ter  $v = \mathbf{0}$ , o que nos dá  $\lambda_1 = \dots = \lambda_n = 0$ , já que os vetores  $v_1, \dots, v_n$  são LI. Logo, os vetores  $T(v_1), \dots, T(v_n)$  são LI.

(iii)  $\Rightarrow$  (i). Sejam  $v_1, v_2 \in V$ , tais que  $T(v_1) = T(v_2)$ . Nesse caso,  $T(v_1 - v_2) = \mathbf{0}$ . Se tivéssemos  $v = v_1 - v_2 \neq \mathbf{0}$ , teríamos que a imagem do conjunto LI,  $\{v\} \subset V$ , seria o conjunto LD,  $\{\mathbf{0}\} \subset W$ , contrariando a hipótese. Segue-se que  $v_1 = v_2$  e, portanto, que  $T$  é injetiva.

Suponhamos agora que  $V$  tenha dimensão finita. Nesse caso, a implicação (iii)  $\Rightarrow$  (iv) é imediata. Vejamos, então, que (iv)  $\Rightarrow$  (ii). Para tanto, tomemos  $v = (x_1, \dots, x_n)_{\mathcal{B}} \in \text{Nuc } T \subset V$ . Então,

$$\mathbf{0} = T(v) = T(x_1v_1 + \dots + x_nv_n) = x_1T(v_1) + \dots + x_nT(v_n).$$

Daí e da independência linear de  $T(v_1), \dots, T(v_n)$ , segue-se que  $x_1 = \dots = x_n = 0$ , o que nos dá  $v = \mathbf{0}$ . Isso prova que  $\text{Nuc } T = \{\mathbf{0}\}$  e conclui, dessa forma, a demonstração. ■

Infere-se da Proposição 3.5 que a transformação linear do Exemplo 3.13 é injetiva, enquanto as dos Exemplos 3.12 e 3.14 não o são. Ainda com respeito a esses exemplos, observa-se que, em todos eles, a dimensão do domínio da transformação linear considerada é a soma de sua nulidade com seu posto. Provemos, agora, que esse fato é geral.

**Teorema 3.2** (DO NÚCLEO E DA IMAGEM). *Seja  $T : V \rightarrow W$  uma transformação linear. Então, se  $V$  tem dimensão finita, vale a igualdade:*

$$\dim V = \dim(\text{Nuc } T) + \dim(\text{Im } T). \quad (3.6)$$

*Demonstração.* Suponhamos, inicialmente, que  $\text{Nuc } T = \{\mathbf{0}\}$ , e tomemos uma base  $\mathcal{B} = \{v_1, \dots, v_n\}$  de  $V$ . Pelas Proposições 3.4 e 3.5, temos que o conjunto  $\mathcal{B}' = \{T(v_1), \dots, T(v_n)\}$  gera  $\text{Im } T$  e é LI. Logo,  $\mathcal{B}'$  é uma base de  $\text{Im } T$ , donde a igualdade (3.6) se cumpre.

Suponhamos agora que  $\dim(\text{Nuc } T) = k > 0$ . Pela Proposição 2.10-(iv), existe uma base  $\mathcal{B} = \{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$  de  $V$ , tal que seus primeiros  $k$  vetores,  $v_1, \dots, v_k$ , formam uma base de  $\text{Nuc } T$ . Além disso, pondo-se  $V_0 = \text{Ger } \{v_{k+1}, \dots, v_n\}$ , tem-se  $V_0 \cap \text{Nuc } T = \{\mathbf{0}\}$ , pois  $V = \text{Nuc } T \oplus V_0$  (vide

Exercício 21 – Cap. 2). Em particular, pela Proposição 3.5,  $T|_{V_0} : V_0 \rightarrow W$  é injetiva. Porém, pela Proposição 3.4, o conjunto

$$\{T(v_1), \dots, T(v_k), \dots, T(v_{k+1}), \dots, T(v_n)\}$$

gera  $\text{Im } T$ . Uma vez que  $T(v_1) = \dots = T(v_k) = \mathbf{0}$ , segue-se que

$$\mathcal{B}' = \{T(v_{k+1}), \dots, T(v_n)\},$$

igualmente, gera  $\text{Im } T$ . No entanto,  $\mathcal{B}'$  é a imagem, pela aplicação injetiva  $T|_{V_0}$ , do conjunto LI  $\{v_{k+1}, \dots, v_n\}$ . Ainda pela Proposição 3.5,  $\mathcal{B}'$  é LI, donde  $\mathcal{B}'$  é uma base de  $\text{Im } T$ . Dessa forma,

$$\dim V = n = k + (n - k) = \dim(\text{Nuc } T) + \dim(\text{Im } T),$$

como queríamos demonstrar. ■

**Corolário 3.1.** *Sejam  $V$  um espaço vetorial de dimensão finita e  $T : V \rightarrow W$  uma transformação linear. São válidas as seguintes asserções:*

- i) *Se  $T$  é bijetiva, então  $W$  tem dimensão finita e  $\dim V = \dim W$ ;*
- ii) *Se  $\dim V = \dim W$ , então, uma condição necessária e suficiente para que  $T$  seja bijetiva, é a de que seja injetiva ou sobrejetiva.*

*Demonstração.* Suponha que  $T$  seja bijetiva. Nesse caso, tem-se  $W = \text{Im } T$  e  $\text{Nuc } T = \{\mathbf{0}\}$ . Em particular,  $W$  tem dimensão finita. Além disso, por (3.6),  $\dim V = \dim(\text{Nuc } T) + \dim(\text{Im } T) = \dim W$ , o que prova (i).

Suponhamos agora que  $\dim V = \dim W$ . Se  $T$  for bijetiva, por definição, será injetiva e sobrejetiva. Reciprocamente, se  $T$  for injetiva, teremos  $\dim(\text{Nuc } T) = 0$  e, então, por (3.6),  $\dim W = \dim V = \dim(\text{Im } T)$ , donde  $T$  será sobrejetiva e, portanto, bijetiva.

Finalmente, supondo-se que  $T$  seja sobrejetiva, segue-se de (3.6) que

$$\dim(\text{Nuc } T) = \dim V - \dim(\text{Im } T) = \dim V - \dim W = 0,$$

isto é,  $T$  é injetiva. Logo, bijetiva. Isso prova (ii) e conclui a demonstração. ■

Segue-se do Corolário 3.1-(ii) que o operador linear  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ , tal que

$$[T] = \begin{bmatrix} 3 & -1 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & -1 \end{bmatrix},$$

é bijetivo. Para verificarmos isso, basta observarmos que os vetores coluna de  $[T]$ , os quais geram  $\text{Im } T$ , são LI. Logo, eles constituem uma base de  $\mathbb{R}^3$ , donde se infere que  $\text{Im } T = \mathbb{R}^3$ . Assim,  $T$  é sobrejetivo. Logo, pelo Corolário 3.1-(ii),  $T$  é bijetivo.

**Exemplo 3.15** (BIJETIVIDADE DA TRANSPOSIÇÃO DE MATRIZES). Seja

$$T : \mathcal{M}(m, n) \rightarrow \mathcal{M}(n, m)$$

a transposição de matrizes, isto é,  $T(X) = X^*$ ,  $X \in \mathcal{M}(m, n)$ . Temos que  $T$  é linear (vide igualdade (1.1)). Além disso,  $T(X) = \mathbf{0}$  se, e só se,  $X^* = \mathbf{0}$  ou, equivalentemente,  $X = \mathbf{0}$ . Logo,  $\text{Nuc } T = \{\mathbf{0}\}$ , donde  $T$  é injetiva. Uma vez que  $\dim \mathcal{M}(m, n) = \dim \mathcal{M}(n, m) = mn$ , segue-se do Corolário 3.1-(ii) que  $T$  é bijetiva.

### 3.5 Operações com Transformações Lineares

Introduziremos agora as operações de adição e produto de transformações lineares, e constataremos que, quando os espaços envolvidos são de dimensão finita, essas operações determinam uma álgebra para transformações lineares, a qual é idêntica àquela das matrizes. Com esse propósito, consideremos espaços vetoriais  $V$  e  $W$ , e denotemos por  $\mathcal{L}(V, W)$  o conjunto formado por todas as transformações lineares de  $V$  em  $W$ , isto é,

$$\mathcal{L}(V, W) = \{T : V \rightarrow W ; T \text{ é linear}\}.$$

Dadas  $T_1, T_2 \in \mathcal{L}(V, W)$  e  $\lambda \in \mathbb{R}$ , definimos a *adição*  $T_1 + T_2 : V \rightarrow W$  e o *produto por escalar*  $\lambda T_1 : V \rightarrow W$  por

$$(T_1 + T_2)(v) = T_1(v) + T_2(v) \quad \text{e} \quad (\lambda T_1)(v) = \lambda T_1(v).$$

É imediato que, para quaisquer  $T_1, T_2 \in \mathcal{L}(V, W)$  e  $\lambda \in \mathbb{R}$ , tem-se  $T_1 + T_2 \in \mathcal{L}(V, W)$  e  $\lambda T_1 \in \mathcal{L}(V, W)$ . É também de fácil verificação que essas operações têm as propriedades fundamentais (comutatividade, associatividade, etc.) que conferem à  $\mathcal{L}(V, W)$  uma estrutura de espaço vetorial (verifique!).

Tomando-se bases  $\mathcal{B} \subset V$  e  $\mathcal{B}' \subset W$ , para quaisquer  $T_1, T_2 \in \mathcal{L}(V, W)$ ,  $v \in V$  e  $\lambda \in \mathbb{R}$ , tem-se

$$\begin{aligned} [(T_1 + \lambda T_2)(v)]_{\mathcal{B}'} &= [T_1(v) + \lambda T_2(v)]_{\mathcal{B}'} = [T_1(v)]_{\mathcal{B}'} + \lambda [T_2(v)]_{\mathcal{B}'} \\ &= [T_1]_{\mathcal{B}}^{\mathcal{B}'} [v]_{\mathcal{B}} + \lambda [T_2]_{\mathcal{B}}^{\mathcal{B}'} [v]_{\mathcal{B}} = ([T_1]_{\mathcal{B}}^{\mathcal{B}'} + \lambda [T_2]_{\mathcal{B}}^{\mathcal{B}'}) [v]_{\mathcal{B}}. \end{aligned}$$

Segue-se, então, dessa última igualdade, juntamente com a unicidade da matriz de uma transformação linear com respeito a bases predeterminadas, o resultado seguinte.

**Proposição 3.6** (A MATRIZ DA SOMA É A SOMA DAS MATRIZES). *Considere transformações lineares  $T_1, T_2 \in \mathcal{L}(V, W)$ , em que  $V$  e  $W$  são espaços vetoriais de dimensões  $n$  e  $m$ , respectivamente. Então, para quaisquer bases  $\mathcal{B} \subset V$ ,  $\mathcal{B}' \subset W$ , e todo  $\lambda \in \mathbb{R}$ , vale a igualdade:*

$$[T_1 + \lambda T_2]_{\mathcal{B}}^{\mathcal{B}'} = [T_1]_{\mathcal{B}}^{\mathcal{B}'} + \lambda [T_2]_{\mathcal{B}}^{\mathcal{B}'}$$

Em particular, a aplicação

$$\begin{aligned} \mathcal{L}(V, W) &\rightarrow \mathcal{M}(m, n) \\ T &\mapsto [T]_{\mathcal{B}}^{\mathcal{B}'} \end{aligned}$$

é uma transformação linear.

Vejamos agora o comportamento das transformações lineares com respeito à composição de aplicações. Para tanto, tomemos espaços vetoriais  $U, V, W$ , e aplicações lineares  $T_1 : U \rightarrow V$ ,  $T_2 : V \rightarrow W$ . Verifiquemos, então, que a composta  $T_2 \circ T_1 : U \rightarrow W$  é linear. Com efeito, dados  $u, v \in U$  e  $\lambda \in \mathbb{R}$ , tem-se

$$\begin{aligned} (T_2 \circ T_1)(u + \lambda v) &= T_2(T_1(u + \lambda v)) = T_2(T_1(u) + \lambda T_1(v)) \\ &= T_2(T_1(u)) + \lambda T_2(T_1(v)) = (T_2 \circ T_1)(u) + \lambda (T_2 \circ T_1)(v), \end{aligned}$$



donde  $T_2 \circ T_1$  é linear.

Diz-se, então, que  $T = T_2 \circ T_1 : U \rightarrow W$  é o *produto* de  $T_2$  e  $T_1$  e escreve-se  $T = T_2 T_1$ . Dado um operador linear  $T : V \rightarrow V$ , o produto  $TT \dots T$ , com  $k$  fatores, denota-se também por  $T^k$ . Essa nomenclatura justifica-se pelo resultado seguinte que, grosso modo, nos diz que a matriz do produto de transformações lineares é o produto das matrizes correspondentes.

**Proposição 3.7** (A MATRIZ DO PRODUTO É O PRODUTO DAS MATRIZES). *Sejam  $U, V$  e  $W$  espaços vetoriais de dimensão finita, e  $T_1 : U \rightarrow V, T_2 : V \rightarrow W$  aplicações lineares. Então, para quaisquer bases  $\mathcal{A} \subset U, \mathcal{B} \subset V$  e  $\mathcal{C} \subset W$ , tem-se:*

$$[T_2 T_1]_{\mathcal{A}}^{\mathcal{C}} = [T_2]_{\mathcal{B}}^{\mathcal{C}} [T_1]_{\mathcal{A}}^{\mathcal{B}}. \quad (3.7)$$

*Demonstração.* Dado  $u \in U$ , considere os vetores

$$v = T_1(u) \in V \quad \text{e} \quad w = T_2(v) = (T_2 T_1)(u) \in W.$$

Nessas condições, temos

$$[(T_2 T_1)u]_{\mathcal{C}} = [T_2(v)]_{\mathcal{C}} = [T_2]_{\mathcal{B}}^{\mathcal{C}} [v]_{\mathcal{B}} = [T_2]_{\mathcal{B}}^{\mathcal{C}} [T_1(u)]_{\mathcal{B}} = ([T_2]_{\mathcal{B}}^{\mathcal{C}} [T_1]_{\mathcal{A}}^{\mathcal{B}}) [u]_{\mathcal{A}},$$

o que implica a igualdade (3.7). ■

Uma vez que transformações lineares entre espaços de dimensão finita são univocamente associadas a matrizes, segue-se das Proposições 3.6 e 3.7 que a álgebra das transformações lineares tem as mesmas propriedades da álgebra das matrizes. Mais precisamente, se  $T_1, T_2, T_3$  são transformações lineares entre espaços de dimensão finita, valem, sempre que os produtos e adições mencionados estejam definidos, as seguintes igualdades:

- i)  $T_1 \mathbf{0} = \mathbf{0} T_1 = \mathbf{0}$ ;
- ii)  $T_1 I = I T_1 = T_1$ , em que  $I$  é a identidade;
- iii)  $T_1(T_2 + T_3) = T_1 T_2 + T_1 T_3$ ;
- iv)  $(T_1 + T_2) T_3 = T_1 T_3 + T_2 T_3$ ;

$$\text{v)} (T_1 T_2) T_3 = T_1 (T_2 T_3);$$

$$\text{vi)} (\lambda T_1) T_2 = T_1 (\lambda T_2) = \lambda (T_1 T_2) \quad \forall \lambda \in \mathbb{R}.$$

Convém mencionar que, se  $T_1, T_2 : V \rightarrow V$  são operadores distintos, tem-se, em geral,  $T_1 T_2 \neq T_2 T_1$ , isto é, assim como o produto de matrizes, *o produto de operadores lineares não é comutativo*. Quando ocorre  $T_1 T_2 = T_2 T_1$ , dizemos que  $T_1$  e  $T_2$  *comutam*. Segue-se, portanto, da Proposição 3.7, que, para qualquer base  $\mathcal{B}$  de  $V$ ,  $T_1$  e  $T_2$  comutam se, e somente se, as matrizes  $[T_1]_{\mathcal{B}}$  e  $[T_2]_{\mathcal{B}}$  comutam.

**Exemplo 3.16.** Sejam  $T_1, T_2, T_3 : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  os operadores lineares dados por

$$T_1(x, y) = (-y, x), \quad T_2(x, y) = (-x, y) \quad \text{e} \quad T_3(x, y) = (2x, y).$$

Temos que suas respectivas matrizes com respeito à base canônica de  $\mathbb{R}^2$  são

$$[T_1] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad [T_2] = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{e} \quad [T_3] = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}.$$

Daí, verifica-se facilmente que  $[T_1]$  não comuta com  $[T_2]$  ou  $[T_3]$ , e que  $[T_2]$  e  $[T_3]$  comutam. Logo, valem as mesmas afirmações para as transformações lineares correspondentes.

## 3.6 Isomorfismos

Estudaremos agora as transformações lineares bijetivas, chamadas de *isomorfismos*. Os isomorfismos, dentre outras características, destacam-se pela propriedade de determinarem uma relação de equivalência na classe dos espaços vetoriais.

**Proposição 3.8** (LINEARIDADE DA INVERSA). *Seja  $T : V \rightarrow W$  uma aplicação linear bijetiva. Então, a inversa de  $T$ ,  $T^{-1} : W \rightarrow V$ , é linear.*

*Demonstração.* Dados  $w_1, w_2 \in W$  e  $\lambda \in \mathbb{R}$ , sejam  $v_1 = T^{-1}(w_1)$  e  $v_2 = T^{-1}(w_2)$ . Uma vez que  $T$  é linear, tem-se  $T(v_1 + \lambda v_2) = T(v_1) + \lambda T(v_2) = w_1 + \lambda w_2$ . Dessa forma,

$$T^{-1}(w_1 + \lambda w_2) = v_1 + \lambda v_2 = T^{-1}(w_1) + \lambda T^{-1}(w_2),$$

donde  $T^{-1}$  é linear. ■

Uma aplicação linear bijetiva  $T : V \rightarrow W$ , como mencionamos acima, é dita um *isomorfismo*. Nesse caso, pela Proposição 3.8,  $T^{-1} : W \rightarrow V$  é também um isomorfismo. Nessas condições, podemos dizer que os espaços  $V$  e  $W$  são *isomorfos*, e indicar essa relação por  $V \sim W$ . Assim,  $\sim$  é simétrica. Além disso, observando-se que a aplicação identidade de um espaço vetorial é um isomorfismo, infere-se que  $\sim$  é reflexiva. Por fim, uma vez que uma composição de bijeções é também uma bijeção, conclui-se que  $\sim$  é também transitiva.

Em suma, vale o resultado seguinte.

**Proposição 3.9.** *A relação de isomorfismo é uma equivalência na classe dos espaços vetoriais.*

**Exemplo 3.17.** A aplicação

$$\begin{aligned} T : \mathcal{P}_2 &\rightarrow \mathbb{R}^3 \\ at^2 + bt + c &\mapsto (a, b, c) \end{aligned}$$

é linear e bijetiva (verifique!), donde se conclui que o espaço dos polinômios de grau menor que, ou igual a, 2 é isomorfo a  $\mathbb{R}^3$ . Analogamente, o espaço  $\mathcal{P}_n$  é isomorfo a  $\mathbb{R}^{n+1}$ .

Mais geralmente, vale o seguinte resultado.

**Proposição 3.10.** *Dois espaços vetoriais de dimensão finita são isomorfos se, e somente se, têm a mesma dimensão.*

*Demonstração.* Sejam  $V$  e  $W$  espaços vetoriais de dimensão finita. Se  $V \sim W$ , segue-se do Corolário 3.1-(i) que  $\dim V = \dim W$ .

Suponhamos, então, que  $\dim V = \dim W = n$ , e tomemos bases  $\mathcal{B} = \{v_1, \dots, v_n\}$  e  $\mathcal{B}' = \{w_1, \dots, w_n\}$ , de  $V$  e  $W$ , respectivamente. Pela Proposição 3.1, existe uma única transformação linear  $T : V \rightarrow W$  que satisfaz  $T(v_i) = w_i \forall i \in \{1, 2, \dots, n\}$ . Esse fato, juntamente com os resultados da Proposição 3.5 e do Corolário 3.1-(ii), nos leva à conclusão de que  $T$  é bijetiva e, portanto, um isomorfismo. ■

**Corolário 3.2.** *Todo espaço vetorial de dimensão  $n$  é isomorfo ao espaço  $\mathbb{R}^n$ .*

**Proposição 3.11.** *Sejam  $V$  e  $W$  espaços vetoriais de mesma dimensão finita. Então, dadas bases  $\mathcal{B} \subset V$  e  $\mathcal{B}' \subset W$ , tem-se que uma transformação linear  $T : V \rightarrow W$  é um isomorfismo se, e somente se, a matriz  $[T]_{\mathcal{B}}^{\mathcal{B}'}$  é invertível. No caso afirmativo, vale a igualdade:*

$$([T]_{\mathcal{B}}^{\mathcal{B}'})^{-1} = [T^{-1}]_{\mathcal{B}'}^{\mathcal{B}}.$$

*Demonstração.* Suponhamos que  $T : V \rightarrow W$  seja um isomorfismo. Nesse caso, designando-se por  $I = W \rightarrow W$  a aplicação identidade de  $W$ , temos que  $TT^{-1} = I$ . Logo, pela Proposição 3.7,

$$[I]_{\mathcal{B}'} = [TT^{-1}]_{\mathcal{B}'} = [T]_{\mathcal{B}}^{\mathcal{B}'} [T^{-1}]_{\mathcal{B}'}^{\mathcal{B}}.$$

Uma vez que  $[I]_{\mathcal{B}'}$  é matriz identidade de ordem  $n = \dim V = \dim W$ , segue-se que  $[T]_{\mathcal{B}}^{\mathcal{B}'}$  é invertível e que  $([T]_{\mathcal{B}}^{\mathcal{B}'})^{-1} = [T^{-1}]_{\mathcal{B}'}^{\mathcal{B}}$ .

Reciprocamente, supondo-se que  $[T]_{\mathcal{B}}^{\mathcal{B}'}$  seja invertível, temos que o núcleo de  $T$  é trivial, donde  $T$  é injetiva e, portanto, bijetiva, já que  $V$  e  $W$  têm mesma dimensão. Com efeito, se  $T(v) = \mathbf{0}$ , então,  $[\mathbf{0}]_{\mathcal{B}'} = [T(v)]_{\mathcal{B}'} = [T]_{\mathcal{B}}^{\mathcal{B}'} [v]_{\mathcal{B}}$ . Logo,  $[v]_{\mathcal{B}} = ([T]_{\mathcal{B}}^{\mathcal{B}'})^{-1} [\mathbf{0}]_{\mathcal{B}'} = [\mathbf{0}]_{\mathcal{B}}$ , donde  $v = \mathbf{0}$  e, portanto,  $\text{Nuc}(T) = \{\mathbf{0}\}$ . ■

**Corolário 3.3.** *Sejam  $T \in \mathcal{L}(V)$  um operador linear num espaço  $V$  de dimensão finita, e  $\mathcal{B}$  uma base arbitrária de  $V$ . Então,  $T$  é invertível se, e somente se,  $[T]_{\mathcal{B}}$  é invertível. No caso afirmativo, tem-se  $[T^{-1}]_{\mathcal{B}} = [T]_{\mathcal{B}}^{-1}$ .*

Decorre diretamente do Teorema 3.1 e da Proposição 3.6 o resultado seguinte.

**Teorema 3.3.** *Sejam  $V$  e  $W$  espaços vetoriais de dimensões  $n$  e  $m$ , respectivamente. Dadas bases  $\mathcal{B} \subset V$  e  $\mathcal{B}' \subset W$ , a aplicação*

$$\begin{aligned} \mathcal{L}(V, W) &\rightarrow \mathcal{M}(m, n) \\ T &\mapsto [T]_{\mathcal{B}'}^{\mathcal{B}} \end{aligned}$$

que associa a cada transformação linear  $T \in \mathcal{L}(V, W)$  a sua matriz com respeito a  $\mathcal{B}$  e  $\mathcal{B}'$ , é um isomorfismo. Em particular,  $\dim \mathcal{L}(V, W) = mn$ .

Se  $V$  e  $W$  são como no teorema acima, segue-se dele e do Corolário 3.2 que

$$\mathcal{L}(V, W) \sim \mathcal{M}(m, n) \sim \mathbb{R}^{mn}.$$

## 3.7 Aplicação aos Sistemas Lineares

Nesta seção que encerra o capítulo, faremos uma breve aplicação dos resultados obtidos ao estudo dos sistemas lineares. Em particular, estabeleceremos condições sobre as matrizes  $A$  e  $B$  de um sistema linear  $AX = B$  para que o mesmo admita solução. Antes, consideremos o seguinte resultado.

**Lema 3.1.** *Suponha que o sistema linear  $AX = B$  admita uma solução particular  $X_p$ . Então,  $\bar{X}$  é solução de  $AX = B$  se, e somente se,  $\bar{X} = X_h + X_p$ , em que  $X_h$  é uma solução do sistema homogêneo  $AX = \mathbf{0}$ . Em particular,  $X_p$  é uma solução única de  $AX = B$  se, e somente se,  $AX = \mathbf{0}$  admite apenas a solução trivial.*

*Demonstração.* Suponhamos que  $X_h$  seja uma solução de  $AX = \mathbf{0}$ . Então, para  $\bar{X} = X_h + X_p$ , teremos

$$A\bar{X} = A(X_h + X_p) = AX_h + AX_p = AX_p = B,$$

donde  $\bar{X}$  será uma solução de  $AX = B$ .

Reciprocamente, suponha que  $\bar{X}$  seja uma solução de  $AX = B$ . Nesse caso, fazendo-se  $X_h = \bar{X} - X_p$ , verifica-se igualmente que  $X_h$  é uma solução do sistema homogêneo  $AX = \mathbf{0}$ . ■

Conforme constatamos anteriormente, para toda matriz  $A \in \mathcal{M}(m, n)$ , o conjunto solução do sistema linear homogêneo  $AX = 0$  é um subespaço vetorial de  $\mathbb{R}^n$ , o qual, naturalmente, corresponde ao núcleo da transformação linear associada,  $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ . Segue-se, portanto, do Lema 3.1, que o conjunto solução do sistema linear  $AX = B$  é a *variedade afim*:

$$\text{Nuc } T_A + v_p = \{v + v_p \in \mathbb{R}^n ; v \in \text{Nuc } T_A\},$$

em que  $v_p \in \mathbb{R}^n$  satisfaz  $A[v_p] = B$  (Fig. 3.1).

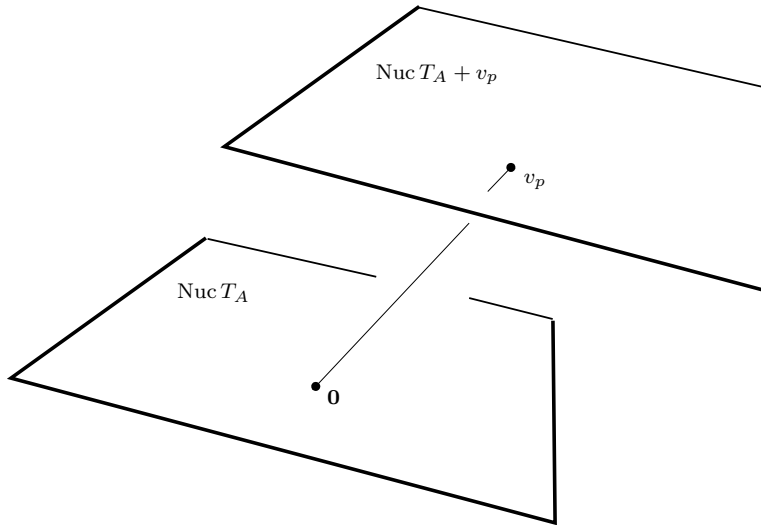


Figura 3.1: O conjunto solução de  $AX = B$  como um subespaço afim de  $\mathbb{R}^n$ .

**Teorema 3.4** (SOLUBILIDADE DE SISTEMAS LINEARES). *Dadas matrizes  $A \in \mathcal{M}(m, n)$  e  $B \in \mathcal{M}(m, 1)$ , o sistema linear  $AX = B$  admite solução se, e somente se,*

$$\text{posto } A = \text{posto } [A | B]. \quad (3.8)$$

*No caso afirmativo, a solução é única se, e somente se, posto  $A = n$ .*

*Demonstração.* Sejam  $w_1, \dots, w_n \in \mathbb{R}^m$  os vetores coluna de  $A$ . Fazendo-se  $B = [w]$ ,  $w \in \mathbb{R}^m$ , pelo Teorema 3.1, existe uma solução de  $AX = B$  se, e

somente se,  $w \in \text{Im } T_A$ , o que ocorre se, e somente se,  $w \in \text{Ger } \{w_1, \dots, w_n\}$  (vide Proposição 3.4 e Observação 3.1). Uma vez que  $[A|B] = [w_1 \cdots w_n w]$ , temos que essa última condição é claramente equivalente à igualdade (3.8).

Pelo Lema 3.1, se existe uma solução particular  $X_p$  de  $AX = B$ , então essa solução é única se, e somente se, o sistema homogêneo associado admite apenas a solução trivial. Esse fato, porém, é equivalente à independência linear dos vetores coluna de  $A$  (vide Proposição 2.6), isto é, à igualdade posto  $A = n$ . ■

## Exercícios

### Seção 3.1

1. Verifique, em cada item, se a aplicação  $T$  é ou não uma transformação linear.
  - a)  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ ,  $T(x, y) = (3x - 2y, x - y, -2x + 2y)$ ;
  - b)  $T : \mathbb{R}^4 \rightarrow \mathbb{R}^2$ ,  $T(x, y, z, w) = (xy, zw)$ ;
  - c)  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $T(x, y, z) = (x + y - 2z, x - 2y + z, x - z)$ ;
  - d)  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $T(x, y) = (x + y - 2, 4x - y/2, x + 9y)$ .
2. Sejam  $T : V \rightarrow W$  uma transformação linear e  $W_0 \subset W$  um subespaço vetorial de  $W$ . Prove que  $T^{-1}(W_0) := \{v \in V ; T(v) \in W_0\}$  é um subespaço vetorial de  $V$ .
3. Seja  $T : V \rightarrow W$  uma transformação linear. Prove que, se o conjunto  $\{T(v_1), \dots, T(v_k)\}$  é LI em  $W$ , então  $\{v_1, \dots, v_k\}$  é LI em  $V$ .
4. Sejam  $V_0, V_1 \subset V$  subespaços de  $V$ , e  $T : V \rightarrow W$  uma transformação linear. Mostre que  $T(V_0 + V_1) = T(V_0) + T(V_1)$ .
5. Sejam  $T : V \rightarrow V$  um operador linear e  $\lambda \in \mathbb{R}$ . Prove que:
  - a)  $V_\lambda = \{v \in V ; T(v) = \lambda v\} \subset V$  é um subespaço de  $V$ ;
  - b) Se  $v_1 \in V_{\lambda_1}$  e  $v_2 \in V_{\lambda_2}$  são não nulos e  $\lambda_1 \neq \lambda_2$ , então  $v_1$  e  $v_2$  são LI.

**Seção 3.2**

6. Seja  $A$  uma matriz quadrada de ordem  $n$  que possui linhas ou colunas nulas. Prove que a transformação linear  $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , tal que  $A = [T]$ , não é injetiva ou sobrejetiva.
7. Seja  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  um operador linear cuja matriz  $[T]$  é simétrica. Prove que existem reais  $\lambda_1, \lambda_2$ , e uma base de  $\mathbb{R}^2$ ,  $\mathcal{B} = \{v_1, v_2\}$ , tais que  $Tv_1 = \lambda_1 v_1$  e  $Tv_2 = \lambda_2 v_2$ .
8. Sejam  $A$  e  $B$  matrizes  $n \times n$ , e  $V$  um espaço vetorial de dimensão  $n$ . Prove que, se  $A$  e  $B$  são semelhantes, então existem bases  $\mathcal{B}, \mathcal{B}'$  de  $V$ , e um operador linear  $T : V \rightarrow V$ , tais que  $[T]_{\mathcal{B}} = A$  e  $[T]_{\mathcal{B}'} = B$ .
9. Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  um funcional linear não identicamente nulo. Prove que  $f$  é sobrejetivo e que existe uma base  $\mathcal{B}$  de  $\mathbb{R}^n$ , tal que

$$f(x_1, \dots, x_n)_{\mathcal{B}} = ax_1, \quad a \in \mathbb{R}.$$

Exiba, então, a matriz de  $f$  com respeito à base  $\mathcal{B}$ .

10. Dada  $A \in \mathcal{M}(2, 2)$ , seja  $T : \mathcal{M}(2, 2) \rightarrow \mathcal{M}(2, 2)$  a aplicação definida por  $T(X) = AX$ ,  $X \in \mathcal{M}(2, 2)$ . Prove que  $T$  é linear e determine a matriz de  $T$  com respeito à base canônica de  $\mathcal{M}(2, 2)$ .

**Seção 3.3**

11. Considere, em  $\mathbb{R}^3$ , as bases

$$\mathcal{B} = \{(1, 1, 0), (-1, 1, 1), (0, 1, 2)\} \quad \text{e} \quad \mathcal{B}' = \{(2, 1, 1), (0, 0, 1), (1, -1, 1)\},$$

e determine as matrizes de mudança de base,  $[I]_{\mathcal{B}'}^{\mathcal{B}}$  e  $[I]_{\mathcal{B}}^{\mathcal{B}'}$ . Em seguida, determine as coordenadas de  $v = (1, 2, 3)_{\mathcal{B}}$  com respeito a  $\mathcal{B}'$ , bem como as coordenadas de  $w = (1, 2, 3)_{\mathcal{B}'}$  com respeito a  $\mathcal{B}$ .

12. Sejam  $\mathcal{B}$  e  $\mathcal{B}'$  como no exercício anterior, e  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  o operador linear definido por  $T(x, y, z) = (x - y + z, 5x + 2y - z, x + y + 3z)$ . Determine as matrizes  $[T]_{\mathcal{B}}$  e  $[T]_{\mathcal{B}'}$  e compare os seus respectivos traços.



**Seção 3.4**

13. Prove que existe uma transformação linear  $T : \mathbb{R}^4 \rightarrow \mathbb{R}^4$  que satisfaz:
- $\text{Nuc } T = \text{Ger} \{(1, 0, 1, 0), (-1, 0, 0, 1)\}$ ;
  - $\text{Im } T = \text{Ger} \{(1, -1, 0, 2), (0, 1, -1, 0)\}$ .
14. Determine o núcleo e a imagem de cada transformação do Exercício 1 que é linear. Em seguida, investigue o comportamento destas transformações lineares no que diz respeito à injetividade e sobrejetividade.
15. Determine o núcleo e a imagem da transformação linear  $T : \mathcal{P}_2 \rightarrow \mathcal{P}_3$  definida pela igualdade:

$$T(at^2 + bt + c) = \frac{a}{3}t^3 + \frac{b}{2}t^2 + ct.$$

16. Sejam  $T_1, T_2 \in \mathcal{L}(V, V)$  operadores lineares de um espaço vetorial  $V$  de dimensão finita. Prove que:
- Existe  $T \in \mathcal{L}(V, V)$  satisfazendo  $T_1 = T_2T \Leftrightarrow \text{Im } T_1 \subset \text{Im } T_2$ ;
  - Existe  $T \in \mathcal{L}(V, V)$  satisfazendo  $T_1 = TT_2 \Leftrightarrow \text{Nuc } T_2 \subset \text{Nuc } T_1$ .
17. Sejam  $V$  e  $W$  espaços vetoriais de dimensão finita, e  $W_0 \subset W$  um subespaço de  $W$ . Dada uma transformação linear  $T : V \rightarrow W$ , prove que  $\dim(T^{-1}(W_0)) \geq \dim V - \dim W + \dim W_0$ .
18. Sejam  $T_1, T_2 : V \rightarrow V$  operadores lineares num espaço vetorial  $V$  de dimensão finita. Prove que:
- $\text{Nuc } T_2 \subset \text{Nuc } (T_1T_2)$ ;
  - $T_2(\text{Nuc } (T_1T_2)) \subset \text{Nuc } (T_1)$ ;

e conclua que  $\dim(\text{Nuc } (T_1T_2)) \leq \dim(\text{Nuc } T_1) + \dim(\text{Nuc } T_2)$ . Generalize para um número finito de operadores  $T_1, \dots, T_k$ .

**Seção 3.5**

19. Considere transformações lineares  $T_1, T_2 \in \mathcal{L}(U, V)$  e  $T_3 \in \mathcal{L}(V, W)$ . Mostre, então, que:
- posto  $(T_1 + T_2) \leq \text{posto}(T_1) + \text{posto}(T_2)$ ;
  - posto  $(T_3 T_1) \leq \min\{\text{posto}(T_1), \text{posto}(T_3)\}$ .
20. Seja  $T : V \rightarrow V$  um operador linear. Suponha que exista  $v \in V$  satisfazendo  $T^k(v) = \mathbf{0}$  e  $T^{k-1}(v) \neq \mathbf{0}$  para algum  $k \in \mathbb{N}$ . Prove que o conjunto  $\{v, T(v), T^2(v), \dots, T^{k-1}(v)\}$  é LI.
21. Seja  $T : V \rightarrow V$  uma transformação linear injetiva. Prove que, se  $T_1, T_2 \in \mathcal{L}(V, V)$  satisfazem  $TT_1 = TT_2$ , então  $T_1 = T_2$ .
22. Seja  $T : V \rightarrow V$  um operador linear. Prove que  $T^2 = \mathbf{0}$  se, e somente se,  $\text{Im} T \subset \text{Nuc} T$ .

### Seção 3.6

23. Considere um isomorfismo  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , e suponha que  $a, b, c, d \in \mathbb{R}^2$  sejam vértices de uma paralelogramo. Mostre que vale o mesmo para  $T(a), T(b), T(c)$  e  $T(d)$ .
24. Diz-se que um operador  $T : V \rightarrow V$  é *nilpotente*, quando existe  $k \in \mathbb{N}$ , tal que  $T^k = \mathbf{0}$ . Mostre que, se  $T : V \rightarrow V$  é nilpotente, então  $T - I : V \rightarrow V$  é um isomorfismo, em que  $I$  denota o operador identidade de  $V$ .
25. Aplique o Corolário 3.3 para mostrar que, dadas matrizes  $A$  e  $B$  em  $\mathcal{M}(n)$ , se  $AB = I$ , então  $BA = I$ , isto é,  $A$  e  $B$  são invertíveis, sendo uma a inversa da outra.
26. Dadas matrizes  $A, B \in \mathcal{M}(n, n)$ , seja  $T : \mathcal{M}(n, n) \rightarrow \mathcal{M}(n, n)$  a transformação linear definida por  $T(X) = AX - XB$ . Mostre que, se  $A$  e  $B$  são semelhantes, então  $T$  não é um isomorfismo.
27. Sejam  $V$  e  $W$  espaços vetoriais, tais que  $\dim V = n$  e  $\dim W = m$ . Dadas bases  $\mathcal{B}_1, \mathcal{B}_2 \subset V$ , e  $\mathcal{B}'_1, \mathcal{B}'_2 \subset W$ , e uma transformação linear  $T : V \rightarrow W$ ,

mostre que existem matrizes invertíveis,  $M \in \mathcal{M}(m, m)$  e  $N \in \mathcal{M}(n, n)$ , tais que  $M[T]_{\mathcal{B}_1}^{\mathcal{B}'_1} = [T]_{\mathcal{B}_2}^{\mathcal{B}'_2} N$ .



# 4

---

## *Determinantes*

Neste capítulo, estudaremos os determinantes. O determinante de uma matriz quadrada  $A$  é um número real, a ela associado, cujas características refletem certas propriedades de  $A$ . Estabelece-se, por exemplo, que  $A$  será invertível se, e somente se, seu determinante for não nulo.

Conforme constataremos nos capítulos posteriores, o determinante desempenha também um papel fundamental na teoria de operadores lineares em espaços de dimensão finita, bem como na conceituação de volume em espaços euclidianos. Como consequência, determinantes têm uma grande incidência em diversas teorias da matemática, especialmente em Análise e Geometria.

Assim, para cada  $n \in \mathbb{N}$ , definiremos uma *função determinante* em  $\mathcal{M}(n)$ . Verificaremos, então, suas principais propriedades e, através das mesmas, obteremos alguns resultados fundamentais, incluindo-se o supracitado critério de inversão de matrizes.

### **4.1 Determinantes em $\mathcal{M}(2)$**

Consideremos uma matriz quadrada de segunda ordem,  $A = (a_{ij})_{2 \times 2}$ , e observemos que, pela Proposição 2.5, um sistema linear do tipo

$$AX = B$$

admite uma única solução se, e somente se, as entradas  $a_{ij}$  de  $A$  satisfazem

$$a_{11}a_{22} - a_{12}a_{21} \neq 0. \tag{4.1}$$

Em particular, a condição (4.1) é equivalente à invertibilidade de  $A$  e, portanto, à invertibilidade do operador linear  $T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ , tal que  $[T] = A$ .

Esses fatos sugerem o estudo da função  $\mathfrak{D} : \mathcal{M}(2) \rightarrow \mathbb{R}$ , definida por

$$\mathfrak{D}(A) = a_{11}a_{22} - a_{12}a_{21}, \quad A = (a_{ij})_{2 \times 2}. \quad (4.2)$$

Para tanto, identificaremos  $\mathcal{M}(2)$  com  $\mathbb{R}^2 \times \mathbb{R}^2$  através da correspondência

$$A = (a_{ij}) \in \mathcal{M}(2) \cong (v_1, v_2) \in \mathbb{R}^2 \times \mathbb{R}^2,$$

em que  $v_1 = (a_{11}, a_{21})$  e  $v_2 = (a_{12}, a_{22})$  são os vetores coluna de  $A$ .

Verifica-se, então, que a função  $\mathfrak{D}$  tem as seguintes propriedades, as quais são válidas para quaisquer  $v_1, v_2, v_3 \in \mathbb{R}^2$  e  $\lambda \in \mathbb{R}$  :

- i)  $\mathfrak{D}(v_1 + \lambda v_2, v_3) = \mathfrak{D}(v_1, v_3) + \lambda \mathfrak{D}(v_2, v_3)$ ;
- ii)  $\mathfrak{D}(v_1, v_2 + \lambda v_3) = \mathfrak{D}(v_1, v_2) + \lambda \mathfrak{D}(v_1, v_3)$ ;
- iii)  $\mathfrak{D}(v_1, v_1) = 0$ ;
- iv)  $\mathfrak{D}(e_1, e_2) = 1$ .

Assim,  $\mathfrak{D}$  é *bilinear* (propriedades (i) e (ii)), *alternada* (propriedade (iii)) e vale 1 na matriz identidade de  $\mathcal{M}(2)$  (propriedade (iv)). Além disso, como consequência das três primeiras propriedades, tem-se que  $\mathfrak{D}$  é *antissimétrica*, isto é, para quaisquer  $v_1, v_2 \in \mathbb{R}^2$ , vale a igualdade

$$\text{v) } \mathfrak{D}(v_1, v_2) = -\mathfrak{D}(v_2, v_1).$$

De fato, segue-se de (i), (ii) e (iii) que:

$$\begin{aligned} 0 &= \mathfrak{D}(v_1 + v_2, v_1 + v_2) \\ &= \mathfrak{D}(v_1, v_1) + \mathfrak{D}(v_1, v_2) + \mathfrak{D}(v_2, v_1) + \mathfrak{D}(v_2, v_2) \\ &= \mathfrak{D}(v_1, v_2) + \mathfrak{D}(v_2, v_1). \end{aligned}$$

Suponhamos agora que  $\mathfrak{D}' : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$  seja uma função que cumpra as propriedades (i)–(iv) acima e, conseqüentemente, a propriedade (v). Nessas

condições, dados  $v_1 = (a_{11}, a_{21})$  e  $v_2 = (a_{12}, a_{22})$  em  $\mathbb{R}^2$ , tem-se

$$\begin{aligned}\mathfrak{D}'(v_1, v_2) &= \mathfrak{D}'(a_{11}e_1 + a_{21}e_2, a_{12}e_1 + a_{22}e_2) \\ &= a_{11}a_{12}\mathfrak{D}'(e_1, e_1) + a_{11}a_{22}\mathfrak{D}'(e_1, e_2) \\ &\quad + a_{21}a_{12}\mathfrak{D}'(e_2, e_1) + a_{21}a_{22}\mathfrak{D}'(e_2, e_2) \\ &= a_{11}a_{22} - a_{12}a_{21} \\ &= \mathfrak{D}(v_1, v_2).\end{aligned}$$

Segue-se dessas considerações que a única função definida em  $\mathbb{R}^2 \times \mathbb{R}^2$  que é bilinear, alternada e que vale 1 em  $(e_1, e_2)$  é a função  $\mathfrak{D}$  definida em (4.2), a qual denomina-se *determinante*.

Dada uma matriz  $A = (a_{ij})_{2 \times 2}$  em  $\mathcal{M}(2)$ , denota-se o determinante de  $A$  por  $\det A$  e escreve-se

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Temos, então, o seguinte resultado:

**Proposição 4.1.** *São equivalentes as seguintes afirmações acerca de uma matriz  $A \in \mathcal{M}(2)$ :*

- *Para toda matriz coluna  $B \in \mathcal{M}(2, 1)$ , o sistema linear  $AX = B$  admite uma única solução;*
- *$A$  é invertível;*
- $\det A \neq 0$ .

Segue-se da Proposição 4.1 que a matriz  $A = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}$ , contrariamente à matriz  $B = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$ , é invertível, pois  $\det A = 1 \neq 0$  e  $\det B = 0$ .

## 4.2 Determinantes em $\mathcal{M}(n)$

A fim de estender para  $\mathcal{M}(n)$  a noção de determinante introduzida na seção anterior, identifiquemos esse espaço com  $\mathbb{R}^n \times \cdots \times \mathbb{R}^n$  ( $n$  cópias) através da seguinte correspondência:

$$A = (a_{ij})_{n \times n} \in \mathcal{M}(n) \cong (v_1, \dots, v_n) \in \mathbb{R}^n \times \cdots \times \mathbb{R}^n,$$

em que, para cada  $j \in \{1, \dots, n\}$ ,  $v_j = (a_{1j}, \dots, a_{nj})$  é o  $j$ -ésimo vetor coluna da matriz  $A$ .

**Definição 4.1** ( $n$ -FORMA). Uma função  $f: \mathbb{R}^n \times \cdots \times \mathbb{R}^n \rightarrow \mathbb{R}$  é dita  $n$ -linear ou, equivalentemente, uma  $n$ -forma em  $\mathbb{R}^n$ , se, para todo índice  $j \in \{1, \dots, n\}$ , vale a igualdade

$$f(v_1, \dots, \lambda v_j + w_j, \dots, v_n) = \lambda f(v_1, \dots, v_j, \dots, v_n) + f(v_1, \dots, w_j, \dots, v_n)$$

quaisquer que sejam  $\lambda \in \mathbb{R}$  e  $v_1, \dots, v_j, w_j, \dots, v_n \in \mathbb{R}^n$ .

**Exemplo 4.1.** A função

$$f(v_1, \dots, v_n) = a_{11} \cdots a_{nn}, \quad v_j = (a_{1j}, \dots, a_{nj}) \in \mathbb{R}^n,$$

é uma  $n$ -forma em  $\mathbb{R}^n$ . Com efeito, dados  $j \in \{1, \dots, n\}$  e  $\lambda \in \mathbb{R}$ , tem-se, para quaisquer  $v = (x_1, \dots, x_n)$ ,  $w = (y_1, \dots, y_n) \in \mathbb{R}^n$ , que

$$\begin{aligned} f(v_1, \dots, v + \lambda w, \dots, v_n) &= a_{11} \cdots (x_j + \lambda y_j) \cdots a_{nn} \\ &= a_{11} \cdots x_j \cdots a_{nn} + \lambda a_{11} \cdots y_j \cdots a_{nn}, \end{aligned}$$

donde se conclui que

$$f(v_1, \dots, v + \lambda w, \dots, v_n) = f(v_1, \dots, v, \dots, v_n) + \lambda f(v_1, \dots, w, \dots, v_n).$$

De modo análogo, prova-se que, dados  $i_1, \dots, i_n \in \{1, \dots, n\}$ , dois a dois distintos, a função  $f: \mathbb{R}^n \times \cdots \times \mathbb{R}^n \rightarrow \mathbb{R}$ , definida por

$$f(v_1, \dots, v_n) = a_{i_1 1} a_{i_2 2} \cdots a_{i_n n}, \quad v_j = (a_{1j}, \dots, a_{nj}),$$

é uma  $n$ -forma em  $\mathbb{R}^n$ .



**Definição 4.2** ( $n$ -FORMA ALTERNADA – ANTISSIMÉTRICA). Diz-se que uma  $n$ -forma  $\mathbf{f}$  em  $\mathbb{R}^n$  é *alternada* se, para quaisquer  $n$  vetores  $v_1, \dots, v_n \in \mathbb{R}^n$ , tem-se

$$\mathbf{f}(v_1, \dots, v_n) = 0$$

sempre que  $v_i = v_j$  para algum par de índices  $i \neq j \in \{1, \dots, n\}$ . Diz-se que  $\mathbf{f}$  é *antissimétrica*, se

$$\mathbf{f}(v_1, \dots, v_i, \dots, v_j, \dots, v_n) = -\mathbf{f}(v_1, \dots, v_j, \dots, v_i, \dots, v_n),$$

quaisquer que sejam  $i, j \in \{1, \dots, n\}$ .

Dadas duas  $n$ -formas em  $\mathbb{R}^n$ ,  $\mathbf{f}$  e  $\mathbf{g}$ , verifica-se facilmente que, para todo  $\lambda \in \mathbb{R}$ , a função  $\mathbf{f} + \lambda\mathbf{g}$  é uma  $n$ -forma em  $\mathbb{R}^n$ . Além disso,  $\mathbf{f} + \lambda\mathbf{g}$  é alternada, se  $\mathbf{f}$  e  $\mathbf{g}$  o forem. Daí, segue-se que

$$\mathcal{F}(\mathbb{R}^n) = \{n\text{-formas em } \mathbb{R}^n\} \text{ e } \mathcal{A}(\mathbb{R}^n) = \{n\text{-formas alternadas em } \mathbb{R}^n\}$$

são espaços vetoriais, sendo que  $\mathcal{A}(\mathbb{R}^n)$ , evidentemente, é um subespaço de  $\mathcal{F}(\mathbb{R}^n)$ .

**Proposição 4.2.** *Sejam  $\mathbf{f}$  uma  $n$ -forma alternada em  $\mathbb{R}^n$  e  $\{v_1, \dots, v_n\} \subset \mathbb{R}^n$  um conjunto LD. Então,  $\mathbf{f}(v_1, \dots, v_n) = 0$ .*

*Demonstração.* Com efeito, pela  $n$ -linearidade de  $\mathbf{f}$ , se um dos vetores  $v_i$  for nulo, teremos  $\mathbf{f}(v_1, \dots, v_n) = 0$ . Caso contrário, pela Proposição 2.9, existem  $k \in \{2, \dots, n\}$  e  $\lambda_1, \dots, \lambda_{k-1} \in \mathbb{R}$ , tais que  $v_k = \lambda_1 v_1 + \dots + \lambda_{k-1} v_{k-1}$ . Logo,

$$\mathbf{f}(v_1, \dots, v_k, \dots, v_n) = \mathbf{f}(v_1, \dots, v_{k-1}, \lambda_1 v_1 + \dots + \lambda_{k-1} v_{k-1}, v_{k+1}, \dots, v_n) = 0,$$

uma vez que  $\mathbf{f}$  é  $n$ -linear e alternada. ■

**Proposição 4.3.** *Uma  $n$ -forma em  $\mathbb{R}^n$  é alternada se, e somente se, é antissimétrica.*

*Demonstração.* Seja  $f$  uma  $n$ -forma alternada em  $\mathbb{R}^n$ . Nesse caso, dados  $i, j \in \{1, \dots, n\}$  e  $v_1, \dots, v_i, \dots, v_j, \dots, v_n \in \mathbb{R}^n$ , tem-se

$$\begin{aligned} 0 &= f(v_1, \dots, v_i + v_j, \dots, v_i + v_j, \dots, v_n) \\ &= f(v_1, \dots, v_i, \dots, v_i, \dots, v_n) + f(v_1, \dots, v_i, \dots, v_j, \dots, v_n) \\ &\quad + f(v_1, \dots, v_j, \dots, v_i, \dots, v_n) + f(v_1, \dots, v_j, \dots, v_j, \dots, v_n) \\ &= f(v_1, \dots, v_i, \dots, v_j, \dots, v_n) + f(v_1, \dots, v_j, \dots, v_i, \dots, v_n), \end{aligned}$$

donde  $f$  antissimétrica.

Reciprocamente, suponhamos que  $f$  seja uma  $n$ -forma antissimétrica em  $\mathbb{R}^n$ . Então, para quaisquer  $v_1, \dots, v_n, v \in \mathbb{R}^n$ , vale a igualdade

$$f(v_1, \dots, v, \dots, v, \dots, v_n) = -f(v_1, \dots, v, \dots, v, \dots, v_n).$$

Logo,  $f(v_1, \dots, v, \dots, v, \dots, v_n) = 0$ , isto é,  $f$  é alternada. ■

## 4.2.1 Existência do Determinante

**Definição 4.3** (FUNÇÃO DETERMINANTE). Uma  $n$ -forma alternada  $\mathfrak{D}$  em  $\mathbb{R}^n$ , a qual cumpre  $\mathfrak{D}(e_1, \dots, e_n) = 1$ , é dita uma *função determinante*.

Claramente, a função identidade é a única função determinante em  $\mathbb{R}$ . Conforme vimos na seção anterior, existe também uma única função determinante em  $\mathbb{R}^2 \times \mathbb{R}^2 \cong \mathcal{M}(2)$ .

Vejam os caso  $n = 3$ , isto é, suponhamos que  $\mathfrak{D} : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$  seja uma função determinante. Escrevendo-se  $v_j = \sum_{i=1}^3 a_{ij}e_i$  e valendo-se das propriedades de  $\mathfrak{D}$ , tem-se

$$\begin{aligned} \mathfrak{D}(v_1, v_2, v_3) &= \mathfrak{D} \left( \sum_{i_1=1}^3 a_{i_1 1} e_{i_1}, \sum_{i_2=1}^3 a_{i_2 2} e_{i_2}, \sum_{i_3=1}^3 a_{i_3 3} e_{i_3} \right) \\ &= a_{11} a_{22} a_{33} \mathfrak{D}(e_1, e_2, e_3) + a_{11} a_{32} a_{23} \mathfrak{D}(e_1, e_3, e_2) \\ &\quad + a_{21} a_{12} a_{33} \mathfrak{D}(e_2, e_1, e_3) + a_{21} a_{32} a_{13} \mathfrak{D}(e_2, e_3, e_1) \\ &\quad + a_{31} a_{12} a_{23} \mathfrak{D}(e_3, e_1, e_2) + a_{31} a_{22} a_{13} \mathfrak{D}(e_3, e_2, e_1). \end{aligned}$$

Porém, por definição,  $\mathfrak{D}(e_1, e_2, e_3) = 1$  e, pela Proposição 4.3,  $\mathfrak{D}$  é antissimétrica. Logo, pondo-se  $A \cong (v_1, v_2, v_3) \in \mathcal{M}(3)$ , tem-se

$$\mathfrak{D}(A) = a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13}.$$

Observemos agora que essa última igualdade pode ser escrita como

$$\begin{aligned} \mathfrak{D}(A) &= a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{21}(a_{12}a_{33} - a_{32}a_{13}) + a_{31}(a_{12}a_{23} - a_{22}a_{13}) \\ &= a_{11} \det A_{11} - a_{21} \det A_{21} + a_{31} \det A_{31}, \end{aligned} \quad (4.3)$$

em que  $A_{1j}$ ,  $j \in \{1, 2, 3\}$ , é a matriz  $2 \times 2$  obtida de  $A$  por eliminação de suas primeira linha e  $j$ -ésima coluna.

A igualdade (4.3) sugere fortemente a construção de funções determinante por indução, conforme o faremos na demonstração do Teorema 4.1. Antes, porém, consideremos o lema seguinte.

**Lema 4.1.** *Seja  $\mathfrak{f}$  uma  $n$ -forma em  $\mathbb{R}^n$  com a seguinte propriedade: Para todo  $j \in \{2, \dots, n\}$  e quaisquer  $v_1, \dots, v_{j-1}, v_j, \dots, v_n$ , tem-se*

$$\mathfrak{f}(v_1, \dots, v_{j-1}, v_j, \dots, v_n) = 0$$

*sempre que  $v_{j-1} = v_j$ . Então,  $\mathfrak{f}$  é alternada.*

*Demonstração.* O argumento usado na primeira parte da demonstração da Proposição 4.3 aplicado à  $n$ -forma  $\mathfrak{f}$  prova que

$$\mathfrak{f}(v_1, \dots, v_{j-1}, v_j, \dots, v_n) = -\mathfrak{f}(v_1, \dots, v_j, v_{j-1}, \dots, v_n),$$

quaisquer que sejam  $j \in \{1, \dots, n\}$  e  $v_1, \dots, v_{j-1}, v_j, \dots, v_n \in \mathbb{R}^n$ . Além disso, toda  $n$ -upla  $(v_1, \dots, v_i, \dots, v_j, \dots, v_n)$  de vetores de  $\mathbb{R}^n$  pode ser transformada na  $n$ -upla  $(v_1, \dots, v_i, v_j, \dots, v_n)$  através de permutações sucessivas de vetores adjacentes, isto é, do tipo  $v_{k-1}, v_k$ . Segue-se, então, da igualdade acima, que

$$\mathfrak{f}(v_1, \dots, v_i, \dots, v_j, \dots, v_n) = (-1)^p \mathfrak{f}(v_1, \dots, v_i, v_j, \dots, v_n),$$

em que  $p$  é o número de permutações realizadas. Em particular, se  $v_i = v_j$ , tem-se  $\mathfrak{f}(v_1, \dots, v_i, \dots, v_j, \dots, v_n) = 0$ , donde se conclui que  $\mathfrak{f}$  é alternada. ■

**Teorema 4.1** (EXISTÊNCIA DO DETERMINANTE EM  $\mathcal{M}(n)$ ). *Para todo  $n \in \mathbb{N}$ , existe uma função determinante em  $\mathcal{M}(n)$ .*

*Demonstração.* O resultado é imediato para  $n = 1$ . Suponhamos, então, que, dado  $n \geq 2$ , exista uma função determinante  $\mathfrak{D}$  em  $\mathcal{M}(n-1)$ , e definamos a função  $\mathfrak{D}_1$  em  $\mathcal{M}(n)$  por

$$\mathfrak{D}_1(A) = \sum_{j=1}^n (-1)^{1+j} a_{1j} \mathfrak{D}(A_{1j}), \quad (4.4)$$

isto é

$$\mathfrak{D}_1(A) = a_{11} \mathfrak{D}(A_{11}) - a_{12} \mathfrak{D}(A_{12}) + \cdots + (-1)^{1+n} a_{1n} \mathfrak{D}(A_{1n}),$$

em que  $A_{1j}$ ,  $j \in \{1, \dots, n\}$ , é a matriz  $(n-1) \times (n-1)$  obtida de  $A$  eliminando-se sua primeira linha e sua  $j$ -ésima coluna.

Mostremos, pois, que  $\mathfrak{D}_1$  é uma função determinante em  $\mathcal{M}(n)$ , o que, juntamente com o Princípio da Indução, nos levará à conclusão do teorema.

Com esse propósito, consideremos a aplicação linear  $T : \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$ , definida por  $T(x_1, \dots, x_n) = (x_2, \dots, x_n)$ , e observemos que, para cada  $j \in \{1, \dots, n\}$ , vale

$$A \cong (v_1, \dots, v_n) \Rightarrow A_{1j} \cong (T(v_1), \dots, T(v_{j-1}), T(v_{j+1}), \dots, T(v_n)).$$

Em particular, a aplicação

$$A \cong (v_1, \dots, v_n) \in \mathcal{M}(n) \mapsto A_{1j} \in \mathcal{M}(n-1)$$

é linear com respeito a cada uma das variáveis  $v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n$ . Consequentemente, uma vez que  $\mathfrak{D}$  é  $(n-1)$ -linear, vale o mesmo para a função

$$A \cong (v_1, \dots, v_n) \in \mathcal{M}(n) \mapsto \mathfrak{f}_j(v_1, \dots, v_n) = a_{1j} \mathfrak{D}(A_{1j}). \quad (4.5)$$

Além disso, tomando-se  $\lambda \in \mathbb{R}$  e  $v'_j = (a'_{1j}, \dots, a'_{nj})$  em  $\mathbb{R}^n$ , tem-se

$$\begin{aligned} \mathfrak{f}_j(v_1, \dots, v_j + \lambda v'_j, \dots, v_n) &= (a_{1j} + \lambda a'_{1j}) \mathfrak{D}(A_{1j}) \\ &= \mathfrak{f}_j(v_1, \dots, v_j, \dots, v_n) + \lambda \mathfrak{f}_j(v_1, \dots, v'_j, \dots, v_n), \end{aligned}$$

ou seja,  $f_j$  é linear, também, com respeito a  $v_j$ .

Segue-se que, para todo  $j = 1, \dots, n$ , a função  $f_j$  é  $n$ -linear. Portanto,  $\mathfrak{D}$ , sendo uma combinação linear dessas funções, é  $n$ -linear.

Tomemos agora  $A \cong (v_1, \dots, v_n) \in \mathcal{M}(n)$  e suponhamos que, para algum  $k \in \{2, \dots, n\}$ , tenha-se  $v_{k-1} = v_k$ . Nesse caso, para todo  $j \in \{1, \dots, n\}$  diferente de  $k-1$  e de  $k$ , a matriz  $A_{1j}$  tem duas colunas iguais. Além disso, as matrizes  $A_{1(k-1)}$  e  $A_{1k}$  são iguais. Logo,

$$\mathfrak{D}_1(v_1, \dots, v_n) = (-1)^k a_{1(k-1)} \mathfrak{D}(A_{1(k-1)}) + (-1)^{k+1} a_{1k} \mathfrak{D}(A_{1k}) = 0.$$

Dessa igualdade e do Lema 4.1, conclui-se que  $\mathfrak{D}$  é alternada.

Por fim, verifica-se por substituição direta que  $\mathfrak{D}_1(e_1, \dots, e_n) = 1$ , donde se infere que  $\mathfrak{D}_1$  é uma função determinante em  $\mathcal{M}(n)$ , como queríamos demonstrar. ■

Através de um argumento análogo ao da demonstração acima, prova-se que, para todo  $i \in \{1, \dots, n\}$ , a função

$$\mathfrak{D}_i(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \mathfrak{D}(A_{ij}), \quad A \in \mathcal{M}(n), \quad (4.6)$$

é uma função determinante em  $\mathcal{M}(n)$ , dita o *desenvolvimento de Laplace* de  $A$  com respeito à sua  $i$ -ésima linha.

Consideremos, por exemplo, a matriz

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

Temos que:

- $\mathfrak{D}_1(A) = 1 \cdot \begin{vmatrix} 2 & -3 \\ 0 & 1 \end{vmatrix} + 1 \cdot \begin{vmatrix} 1 & -3 \\ 2 & 1 \end{vmatrix} + 0 \cdot \begin{vmatrix} 1 & 2 \\ 2 & 0 \end{vmatrix} = 9;$
- $\mathfrak{D}_2(A) = - \begin{vmatrix} -1 & 0 \\ 0 & 1 \end{vmatrix} + 2 \cdot \begin{vmatrix} 1 & 0 \\ 2 & 1 \end{vmatrix} + 3 \cdot \begin{vmatrix} 1 & -1 \\ 2 & 0 \end{vmatrix} = 9;$

$$\bullet \mathfrak{D}_3(A) = 2 \cdot \begin{vmatrix} -1 & 0 \\ 2 & -3 \end{vmatrix} - 0 \cdot \begin{vmatrix} 1 & 0 \\ 1 & -3 \end{vmatrix} + 1 \cdot \begin{vmatrix} 1 & -1 \\ 1 & 2 \end{vmatrix} = 9;$$

o que sugere a unicidade da função determinante, tema da próxima seção.

## 4.2.2 Unicidade do Determinante

No que se segue, constataremos que as  $n$  expressões (4.6) definem a mesma função determinante, o que decorrerá do fato de que, na verdade, para todo  $n \in \mathbb{N}$ , existe uma única função determinante em  $\mathcal{M}(n)$ . A demonstração desse fato nos levará a fazer algumas considerações sobre permutações.

**Definição 4.4** (PERMUTAÇÃO – TRANSPOSIÇÃO). Dado  $n \in \mathbb{N}$ , uma *permutação* (de grau  $n$ ) é uma função bijetiva  $\sigma : N_n \rightarrow N_n$ , em que  $N_n = \{1, \dots, n\}$ . Uma permutação  $\sigma$  de grau  $n$  é dita uma *transposição*, se existem  $i, j \in N_n$ , tais que:

- $i \neq j$ ;
- $\sigma(i) = j$  e  $\sigma(j) = i$ ;
- $\sigma(k) = k \forall k \in N_n - \{i, j\}$ .

Adotaremos a seguinte notação: Dada uma permutação  $\sigma : N_n \rightarrow N_n$ , escreveremos  $\sigma(1) = \sigma_1, \dots, \sigma(n) = \sigma_n$ . Quando  $\sigma$  e  $\tau$  forem, ambas, permutações de grau  $n$ , denotaremos a composta  $\sigma \circ \tau$  como o produto  $\sigma\tau$ . Evidentemente, o produto de permutações de grau  $n$  é também uma permutação de grau  $n$ .

Dito de forma simples, uma permutação de grau  $n$  nada mais é que uma reordenação da sequência  $1, \dots, n$ . A partir desse ponto de vista, é fácil ver que toda permutação se expressa como um produto de transposições, isto é, qualquer permutação preestabelecida pode ser obtida, reordenando-se os elementos de  $N_n$  de dois em dois. Assim, para toda permutação  $\sigma : N_n \rightarrow N_n$ , existem transposições  $\tau_1, \dots, \tau_k$ , tais que

$$\sigma = \tau_1 \dots \tau_k. \tag{4.7}$$

À igualdade (4.7), chamaremos uma *representação* de  $\sigma$ .

É claro que, em geral, uma permutação admite mais de uma representação. No entanto, o número  $k$  de transposições de uma representação qualquer de uma permutação  $\sigma$  é sempre par ou sempre ímpar, ou seja, a paridade de  $k$  depende de  $\sigma$ , não da representação em si. Curiosamente, podemos demonstrar facilmente esse fato, de caráter puramente combinatório, usando a teoria de determinantes. Com efeito, seja  $\mathfrak{D}$  uma função determinante em  $\mathcal{M}(n)$ . Uma vez que  $\mathfrak{D}$  é antissimétrica, se  $\sigma$  for uma permutação de grau  $n$  representada como em (4.7), teremos que

$$\mathfrak{D}(e_{\sigma_1}, \dots, e_{\sigma_n}) = (-1)^k \mathfrak{D}(e_1, \dots, e_n) = (-1)^k.$$

Dessa forma,  $k$  será par ou ímpar, conforme tenha-se  $\mathfrak{D}(e_{\sigma_1}, \dots, e_{\sigma_n}) = 1$  ou  $\mathfrak{D}(e_{\sigma_1}, \dots, e_{\sigma_n}) = -1$ , respectivamente.

**Definição 4.5** (PARIDADE DE PERMUTAÇÕES). Diz-se que uma permutação é *par*, quando é representada por um número par de transposições. Caso contrário, ela é dita *ímpar*. Define-se o  *sinal* de uma permutação  $\sigma$ ,  $\text{sgn } \sigma$ , como 1, para permutações pares, e  $-1$ , para permutações ímpares.

Tomemos agora uma  $n$ -forma alternada  $\mathfrak{f}$  em  $\mathbb{R}^n$ . Pondo-se, para cada índice  $j \in \{1, \dots, n\}$ ,

$$v_j = (a_{1j}, \dots, a_{nj}) = \sum_{i=1}^n a_{ij} e_i \in \mathbb{R}^n,$$

e procedendo-se como o fizemos no caso particular  $n = 3$ , concluí-se, igualmente, que  $\mathfrak{f}(v_1, \dots, v_n)$  é uma soma de  $n!$  parcelas do tipo

$$a_{\sigma_1 1} a_{\sigma_2 2} \dots a_{\sigma_n n} \mathfrak{f}(e_{\sigma_1}, e_{\sigma_2}, \dots, e_{\sigma_n}),$$

em que cada  $\sigma$  é uma permutação de grau  $n$ . Uma vez que  $\mathfrak{f}$  é antissimétrica, essa expressão nada mais é que

$$(\text{sgn } \sigma) a_{\sigma_1 1} a_{\sigma_2 2} \dots a_{\sigma_n n} \mathfrak{f}(e_1, e_2, \dots, e_n),$$

isto é, dada  $A \in \mathcal{M}(n)$ , tem-se

$$f(A) = \sum_{\sigma \in \Lambda} (\text{sgn } \sigma) a_{\sigma_1 1} a_{\sigma_2 2} \cdots a_{\sigma_n n} f(I), \quad (4.8)$$

em que  $\Lambda = \{\text{permutações de grau } n\}$ .

Segue-se da igualdade (4.8) que existe uma única função determinante em  $\mathcal{M}(n)$ , a qual denotamos por  $\det$ , isto é,

$$\det A = \sum_{\sigma \in \Lambda} (\text{sgn } \sigma) a_{\sigma_1 1} a_{\sigma_2 2} \cdots a_{\sigma_n n}, \quad A = (a_{ij}) \in \mathcal{M}(n). \quad (4.9)$$

Em suma, tem-se o resultado seguinte.

**Teorema 4.2** (UNICIDADE DO DETERMINANTE). *Para todo  $n \in \mathbb{N}$ , existe uma única função determinante em  $\mathcal{M}(n)$ , a qual denota-se por  $\det$  e define-se através da igualdade (4.9). Além disso, se  $f$  for uma  $n$ -forma alternada, para toda matriz  $A \in \mathcal{M}(n)$ , valerá a seguinte igualdade:*

$$f(A) = (\det A) f(I).$$

Em particular,  $\dim \mathcal{A}(\mathbb{R}^n) = 1$ .

Dada uma matriz  $A = (a_{ij}) \in \mathcal{M}(n)$ , observemos que

$$a_{\sigma_1 1} a_{\sigma_2 2} \cdots a_{\sigma_n n} = a_{1\sigma_1^{-1}} a_{2\sigma_2^{-1}} \cdots a_{n\sigma_n^{-1}} \quad \forall \sigma \in \Lambda.$$

Além disso, se  $\sigma$  se expressa como em (4.7), então,  $\sigma^{-1} = \tau_k^{-1} \cdots \tau_1^{-1}$ , donde  $\text{sgn } \sigma = \text{sgn } \sigma^{-1}$ . Dessas considerações e da igualdade (4.9), segue-se que

$$\det A = \det A^* \quad \forall A \in \mathcal{M}(n). \quad (4.10)$$

Dada uma matriz  $A = (a_{ij}) \in \mathcal{M}(n)$ , escreveremos

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$



Observemos que, pela  $n$ -linearidade do determinante, e também do fato desta função ser alternada, vale, para quaisquer  $v_1, \dots, v_n \in \mathbb{R}^n$ , a seguinte igualdade:

$$\det(v_1, \dots, v_i, \dots, v_j, \dots, v_n) = \det(v_1, \dots, v_i, \dots, \lambda v_i + v_j, \dots, v_n),$$

isto é, dada uma matriz  $A = [v_1 \cdots v_n]$ , seu determinante não se altera quando substituimos uma coluna por sua soma com um múltiplo de outra. O mesmo se aplica às linhas de  $A$ , devido à igualdade (4.10).

Vejam, agora, que o cálculo do determinante de certas matrizes pode ser mais facilmente efetuado quando essa propriedade é considerada.

**Exemplo 4.2.** Provemos a seguinte igualdade:

$$\det A = \begin{vmatrix} t & 0 & 0 & a \\ -1 & t & 0 & b \\ 0 & -1 & t & c \\ 0 & 0 & -1 & d \end{vmatrix} = a + bt + ct^2 + dt^3.$$

Para tanto, efetuem uma sequência de operações sobre as linhas  $\ell_i$  de  $A$ , as quais serão substituídas pelas linhas  $\ell'_i$  da seguinte forma:

$$\ell_3 \rightarrow \ell_3 + t\ell_4 = \ell'_3 \quad \ell_2 \rightarrow \ell_2 + t\ell'_3 = \ell'_2 \quad \ell_1 \rightarrow \ell_1 + t\ell'_2 = \ell'_1,$$

o que implicará na igualdade:

$$\det A = \begin{vmatrix} 0 & 0 & 0 & a + bt + ct^2 + dt^3 \\ -1 & 0 & 0 & b + ct + dt^2 \\ 0 & -1 & 0 & c + dt \\ 0 & 0 & -1 & d \end{vmatrix} = \begin{vmatrix} 0 & 0 & 0 & p_1(t) \\ -1 & 0 & 0 & p_2(t) \\ 0 & -1 & 0 & p_3(t) \\ 0 & 0 & -1 & p_4(t) \end{vmatrix}.$$

Agora, denotando-se as colunas desse último determinante por  $v_i$ , e substituindo-se sucessivamente a última coluna  $v_3$  por  $v_3 + p_i v_i$ ,  $i = 1, 2, 3$ , obtém-se

$$\det A = \begin{vmatrix} 0 & 0 & 0 & p_1(t) \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{vmatrix}.$$

Por fim, após 3 permutações das colunas e multiplicação de três delas por -1, concluímos que:

$$\det A = \begin{vmatrix} p_1(t) & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} = p_1(t) = a + bt + ct^2 + dt^3.$$

Com um argumento inteiramente análogo, obtém-se a igualdade:

$$\begin{vmatrix} t & 0 & 0 & \cdots & 0 & a_0 \\ -1 & t & 0 & \cdots & 0 & a_1 \\ 0 & -1 & t & \cdots & 0 & a_2 \\ 0 & 0 & -1 & \ddots & 0 & \vdots \\ \vdots & \vdots & \vdots & \ddots & t & \vdots \\ 0 & 0 & 0 & \cdots & -1 & a_{n-1} \end{vmatrix} = a_0 + a_1 t + \cdots + a_{n-1} t^{n-1}, \quad (4.11)$$

à qual nos referiremos posteriormente.

### 4.3 Propriedade Fundamental – Aplicações

Estabeleceremos nessa seção a propriedade fundamental dos determinantes, qual seja, que o determinante do produto de duas matrizes (quadradas e de mesma ordem) é igual ao produto de seus determinantes. Aplicaremos essa propriedade para mostrar que uma matriz quadrada é invertível se, e somente se, seu determinante é não nulo, o que nos permitirá definir o determinante de um operador linear de um espaço de dimensão finita. Obteremos, também, uma fórmula de inversão para matrizes, bem como um método para calcular o posto de uma matriz arbitrária (não necessariamente quadrada). Discutiremos, em seguida, sobre matrizes diagonais por blocos, e concluiremos, então, apresentando um resultado sobre polinômios, o qual envolve o conceito de determinante, conhecido como *critério de Sylvester*.

**Teorema 4.3.** *Dadas matrizes  $A, B \in \mathcal{M}(n)$ , tem-se*

$$\det(AB) = \det A \det B.$$

*Demonstração.* Fixemos  $A \in \mathcal{M}(n)$  e consideremos a função  $f: \mathcal{M} \rightarrow \mathbb{R}$ , dada por  $f(B) = \det(AB)$ . Observemos, então, que se  $B \cong (v_1, \dots, v_n)$ , tem-se  $AB \cong (Av_1, \dots, Av_n)$ , isto é,  $f(v_1, \dots, v_n) = \det(Av_1, \dots, Av_n)$ . Daí, conclui-se facilmente que  $f$  é  $n$ -linear e alternada. Logo, pelo Teorema 4.2, tem-se  $f(B) = \det(B)f(I)$ , donde  $\det(AB) = \det(B)\det(A)$ , como desejado. ■

Sejam  $A \in \mathcal{M}(n)$  uma matriz invertível e  $A^{-1}$  sua inversa. Segue-se da igualdade  $AA^{-1} = I$  e do Teorema 4.3 que  $\det A$  e  $\det A^{-1}$  são não nulos e cumprem  $\det A = 1/\det A^{-1}$ . Por outro lado, se  $A \cong (v_1, \dots, v_n)$  não for invertível, conforme vimos no capítulo anterior, o posto de  $A$  será menor que  $n$ . Em particular, os vetores coluna de  $A$  serão LD. Logo, pela Proposição 4.2, teremos  $\det A = 0$ .

Vale, portanto, o seguinte resultado.

**Proposição 4.4.** *Uma condição necessária e suficiente para que uma matriz  $A \in \mathcal{M}(n)$  seja invertível, é a de que seu determinante seja não nulo. No caso afirmativo, valerá a igualdade*

$$\det A = \frac{1}{\det A^{-1}}.$$

**Exemplo 4.3.** Conforme constatamos anteriormente, a matriz

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

tem determinante igual a  $9 \neq 0$ . Logo,  $A$  é invertível e  $\det A^{-1} = 1/9$ .

Consideremos um espaço vetorial  $V$  de dimensão finita, e bases  $\mathcal{B}, \mathcal{B}' \subset V$ . Temos, pela Proposição 3.2, que as respectivas matrizes de um operador linear  $T: V \rightarrow V$  com respeito a  $\mathcal{B}$  e  $\mathcal{B}'$ ,  $[T]_{\mathcal{B}}$  e  $[T]_{\mathcal{B}'}$ , são semelhantes, isto é, existe uma matriz invertível  $M$  (a matriz de mudança de  $\mathcal{B}$  para  $\mathcal{B}'$ ), tal que

$$[T]_{\mathcal{B}} = M^{-1}[T]_{\mathcal{B}'}M. \quad (4.12)$$

Segue-se, portanto, de (4.12), do Teorema 4.3 e da Proposição 4.4 que

$$\det[T]_{\mathcal{B}} = \det[T]_{\mathcal{B}'},$$

igualdade esta que justifica a definição a seguir.

**Definição 4.6** (DETERMINANTES DE OPERADORES). O *determinante* de um operador linear  $T : V \rightarrow V$  num espaço vetorial  $V$  de dimensão finita, denotado por  $\det T$ , é definido por

$$\det T = \det [T]_{\mathcal{B}},$$

em que  $\mathcal{B}$  é uma base arbitrária de  $V$ .

Observe que, pela Proposição 4.4, *um operador*  $T : V \rightarrow V$  ( $\dim V < \infty$ ) *é invertível se, e somente se,  $\det T \neq 0$ .*

### 4.3.1 Inversão de Matrizes

Dada uma matriz  $A \in \mathcal{M}(n)$ , seja  $C = (c_{ij}) \in \mathcal{M}(n)$  a matriz tal que

$$c_{ij} = (-1)^{i+j} \det A_{ij}. \quad (4.13)$$

Cada entrada  $c_{ij}$  é dita um *cofator* de  $A$ , e  $C$  é dita, então, a *matriz dos cofatores* de  $A$ . Define-se a *adjunta clássica* de  $A$  como a matriz  $\text{adj } A = C^*$ , isto é,

$$\text{adj } A = (d_{ij})_{n \times n}, \quad d_{ij} = (-1)^{i+j} \det A_{ji}.$$

Verifiquemos que

$$A(\text{adj } A) = (\det A)I \quad \forall A \in \mathcal{M}(n). \quad (4.14)$$

De fato, escrevendo-se  $B = A(\text{adj } A)$ , tem-se, para todo  $i \in \{1, \dots, n\}$ , que

$$b_{ii} = \sum_{k=1}^n a_{ik} d_{ki} = \sum_{k=1}^n (-1)^{i+k} a_{ik} \det A_{ik} = \det A. \quad (4.15)$$

Agora, dados  $1 \leq i \neq j \leq n$ , seja  $A' = (a'_{ij}) \in \mathcal{M}(n)$  a matriz obtida de  $A$  substituindo-se sua  $j$ -ésima linha por sua  $i$ -ésima linha. Assim,  $A'$  terá

duas linhas iguais e, portanto,  $\det A' = 0$ . Note também que  $A'_{jk} = A_{jk} \forall k \in \{1, \dots, n\}$ . Dessa forma,

$$b_{ij} = \sum_{k=1}^n a_{ik} d_{kj} = \sum_{k=1}^n (-1)^{j+k} a_{ik} \det A_{jk} = \sum_{k=1}^n (-1)^{j+k} a'_{jk} \det A'_{jk} = \det A' = 0.$$

Combinando-se essa última igualdade com (4.15), obtém-se (4.14). Em particular, se  $A$  for invertível, valerá a igualdade

$$A^{-1} = \frac{1}{\det A} (\text{adj } A). \quad (4.16)$$

Retomando-se a matriz

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

do Exemplo 4.3, a qual verificamos ser invertível, pode-se obter facilmente sua inversa por meio da fórmula (4.16). Com efeito, calculando-se os cofatores  $(c_{ij})$  de  $A$  através de (4.13) e lembrando-se que  $\det A = 9$ , obtém-se

$$A^{-1} = \frac{1}{9} \begin{bmatrix} 2 & 1 & 3 \\ -7 & 1 & 3 \\ -4 & -2 & 3 \end{bmatrix}.$$

### 4.3.2 Determinação do Posto de uma Matriz

Consideremos a matriz

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 5 & 1 & 0 \\ 0 & 0 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

cujos posto, claramente, é igual a 3. Observemos que, excluindo-se sua última linha e suas duas últimas colunas, obtém-se a seguinte *submatriz*  $3 \times 3$  de  $A$ :

$$A' = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix},$$

cujo determinante é igual a  $1 \neq 0$  e que, portanto, é invertível. Em particular,  $\text{posto}(A') = 3$ . Note que a escolha da linha e colunas excluídas foi determinada pelo fato de o espaço linha de  $A$  ser gerado pelos três primeiros vetores linha, e seu espaço coluna ser gerado pelos três primeiros vetores coluna.

De modo geral, se o posto de uma matriz não nula  $A \in \mathcal{M}(m, n)$  for  $p$ , teremos, pela Proposição 2.11, que  $\dim \mathcal{L}_A = \dim \mathcal{C}_A = p$ . Nessas condições, excluindo-se, de  $A$ ,  $m-p$  vetores linha e  $n-p$  vetores coluna, convenientemente escolhidos, determina-se uma submatriz  $A' \in \mathcal{M}(p)$ , cujo posto é  $p$  e que, portanto, tem determinante não nulo.

Por fim, observemos que se  $A'$  for uma submatriz de  $A$  e  $\alpha'$  for um subconjunto LI do conjunto formado por seus vetores coluna, então o conjunto  $\alpha$ , formado pelos vetores coluna correspondentes de  $A$ , será também LI. Logo, o posto de uma submatriz de  $A$  nunca é maior que o de  $A$ . Dessa forma, se o posto de uma matriz  $A$  for igual a  $p$ , nenhuma submatriz quadrada de  $A$ , de ordem maior que  $p$ , pode ter determinante não nulo.

Vale, portanto, o resultado seguinte.

**Proposição 4.5.** *O posto de uma matriz  $A \in \mathcal{M}(m, n)$  é o maior inteiro  $p$  com a propriedade de que existe uma submatriz  $A' \in \mathcal{M}(p)$ , de  $A$ , tal que  $\det A' \neq 0$ .*

**Exemplo 4.4.** O posto da matriz

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

é 2, pois, designando-se por  $A' \in \mathcal{M}(2)$  a submatriz de  $A$  obtida por exclusão de sua terceira linha, tem-se  $\det A' = 1 \neq 0$ .

### 4.3.3 Matrizes Diagonais por Blocos

Do ponto de vista operacional, as matrizes mais simples são as diagonais. Por exemplo, dadas duas matrizes diagonais  $A, B \in \mathcal{M}(n)$ , tem-se que  $AB$  é

também uma matriz diagonal, a qual é obtida multiplicando-se os elementos correspondentes das diagonais de  $A$  e  $B$ , isto é,

$$AB = \begin{bmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{bmatrix} \begin{bmatrix} b_{11} & & \\ & \ddots & \\ & & b_{nn} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & & \\ & \ddots & \\ & & a_{nn}b_{nn} \end{bmatrix},$$

em que as entradas omitidas são todas nulas.

Nesse contexto, um outro fato relevante é o de que o determinante de uma matriz diagonal é simplesmente o produto de suas entradas:

$$\det A = \begin{vmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{vmatrix} = a_{11} \dots a_{nn}.$$

Vejam agora que essas propriedades se estendem às matrizes *diagonais por blocos*. Com este fim, comecemos considerando as seguintes matrizes

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix},$$

cujos produto é:

$$AB = \begin{bmatrix} 3 & 5 & 0 \\ 4 & 7 & 0 \\ 0 & 0 & 6 \end{bmatrix}.$$

Observemos que, destacando-se os “blocos diagonais” de  $A$  e  $B$ , pode-se efetuar esse produto “bloco a bloco”, como no caso das matrizes diagonais. Mais precisamente, tem-se:

$$AB = \begin{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \\ \\ 2 \end{bmatrix} \begin{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \\ \\ 3 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \\ \\ 2 \cdot 3 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 3 & 5 \\ 4 & 7 \end{bmatrix} \\ \\ 6 \end{bmatrix},$$

em que, como anteriormente, as entradas omitidas são todas nulas.

Analogamente, para as matrizes

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

tem-se

$$\begin{aligned} AB &= \begin{bmatrix} \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} & & & \\ & & & & \\ & & \begin{bmatrix} -3 & 1 \\ 2 & 2 \end{bmatrix} & & \\ & & & & \\ & & & & \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\ & & & & & & & & \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \end{bmatrix} \\ &= \begin{bmatrix} \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} & & & & \\ & & & & & & & & \\ & & & \begin{bmatrix} -3 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} & & & & \\ & & & & & & & & \end{bmatrix} \\ &= \begin{bmatrix} \begin{bmatrix} 0 & 0 \\ 5 & 5 \end{bmatrix} & & & & \\ & & & & \\ & & & \begin{bmatrix} -4 & 5 \\ 0 & 2 \end{bmatrix} & \\ & & & & \end{bmatrix}, \end{aligned}$$

isto é,

$$AB = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 5 & 5 & 0 & 0 \\ 0 & 0 & -4 & 5 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

Mais geralmente, suponhamos que  $A \in \mathcal{M}(n)$  seja tal que

$$A = \begin{bmatrix} A_1 & & & \\ & \ddots & & \\ & & & A_k \end{bmatrix},$$



em que, para cada  $i = 1, \dots, k$ ,  $A_i$  é uma submatriz quadrada de  $A$  de ordem  $n_i$ ,  $1 \leq n_i < n$ , cuja diagonal é parte da diagonal de  $A$ . Além disso,

$$n_1 + \dots + n_k = n.$$

Nessas condições, diremos que cada submatriz  $A_i$  é um *bloco diagonal* de ordem  $n_i$  de  $A$ , e que  $A$  é uma matriz *diagonal por blocos*.

Consideremos agora uma matriz diagonal por blocos

$$B = \begin{bmatrix} B_1 & & \\ & \ddots & \\ & & B_k \end{bmatrix},$$

e suponhamos que  $A$  e  $B$  sejam do *mesmo tipo*, isto é, para todo  $i = 1, \dots, n$ ,  $A_i$  e  $B_i$  têm a mesma ordem  $n_i$ . Nesse caso, segue-se diretamente da definição de produto de matrizes que  $AB$  é diagonal por blocos, e que cada um de seus blocos é o produto dos blocos correspondentes de  $A$  e  $B$ , isto é,

$$AB = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{bmatrix} \begin{bmatrix} B_1 & & \\ & \ddots & \\ & & B_k \end{bmatrix} = \begin{bmatrix} A_1 B_1 & & \\ & \ddots & \\ & & A_k B_k \end{bmatrix}.$$

Em particular, para todo  $m \in \mathbb{N}$ , tem-se

$$A^m = \begin{bmatrix} A_1^m & & \\ & \ddots & \\ & & A_k^m \end{bmatrix}.$$

Provemos, agora, a seguinte igualdade:

$$\det \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{bmatrix} = \det A_1 \dots \det A_k. \quad (4.17)$$

Suponhamos, inicialmente, que  $A_i = I_i \forall i = 2, \dots, k$ , em que  $I_i$  é a matriz identidade de ordem  $n_i$ . Nesse caso, denotando-se os  $n_1$  primeiros vetores coluna de  $A$  por  $v_1, \dots, v_{n_1}$ , temos que

$$A = [v_1 \dots v_{n_1} e_{n_1+1} \dots e_n].$$



isto é,

$$\begin{aligned}
 A &= \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{bmatrix} \\
 &= \begin{bmatrix} A_1 & & & \\ & I_2 & & \\ & & \ddots & \\ & & & I_k \end{bmatrix} \begin{bmatrix} I_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & I_k \end{bmatrix} \cdots \begin{bmatrix} I_1 & & & \\ & I_2 & & \\ & & \ddots & \\ & & & A_k \end{bmatrix},
 \end{aligned}$$

e aplicando-se a propriedade fundamental dos determinantes, obtém-se

$$\det A = \det A_1 \dots \det A_k,$$

como desejávamos provar.

### 4.3.4 Critério de Sylvester

Em conclusão, aplicaremos o Teorema 4.3 para estabelecer um elegante resultado sobre fatoração de polinômios, devido a James Sylvester (1814–1897), o qual envolve determinantes<sup>(i)</sup>.

Dados dois polinômios  $f, g \in \mathcal{P}[t, \mathbb{R}]$ , em que

$$f(t) = a_m t^m + \dots + a_1 t + a_0 \quad \text{e} \quad g(t) = b_n t^n + \dots + b_1 t + b_0, \quad a_m, b_n \neq 0,$$

associamos aos mesmos uma matriz  $(m+n) \times (m+n)$ , dita a *matriz de Sylvester* do par  $(f, g)$ , a qual se define como

$$S(f, g) = \begin{bmatrix} a_m & \cdots & a_0 & & & \\ & \ddots & & & \ddots & \\ & & a_m & \cdots & & a_0 \\ b_n & & \cdots & b_0 & & \\ & \ddots & & & \ddots & \\ & & b_n & \cdots & & b_0 \end{bmatrix} \left. \begin{array}{l} \vphantom{\begin{bmatrix} a_m \\ \vdots \\ a_m \\ b_n \\ \vdots \\ b_n \end{bmatrix}} \right\} n \text{ linhas} \\ \left. \vphantom{\begin{bmatrix} a_m \\ \vdots \\ a_m \\ b_n \\ \vdots \\ b_n \end{bmatrix}} \right\} m \text{ linhas}$$

---

<sup>(i)</sup>Vide: H. Scala, *An application of determinants*. American Mathematical Monthly **78** (1971) 889–890.

Na matriz  $S(f, g)$ , em cada linha, as entradas não especificadas são todas iguais a zero, enquanto as especificadas, como indicado, são os coeficientes de  $f$  ( $n$  primeiras linhas) ou de  $g$  ( $m$  últimas linhas). Note que, escrevendo-se  $S(f, g) = (s_{ij})$ , os coeficientes de  $f$  e  $g$  são dispostos ao longo das linhas de  $S(f, g)$  de modo que, se  $s_{ij} = a_k$  (respec.,  $s_{ij} = b_k$ ), então  $s_{(i+1)(j+1)} = a_k$  (respec.,  $s_{(i+1)(j+1)} = b_k$ ),  $1 \leq i, j < m + n$

Por fim, lembremos que todo polinômio  $f \in \mathcal{P}[t, \mathbb{R}]$  admite uma fatoração única, a menos da ordem dos fatores, do tipo

$$f = f_1 \dots f_k,$$

em que cada fator  $f_i$  é um polinômio irredutível (isto é, que não admite fatoração em  $\mathcal{P}[t, \mathbb{R}]$ ) de grau 1 ou 2. Nesse contexto, diz-se que  $f, g \in \mathcal{P}[t, \mathbb{R}]$  têm um *fator comum*, se um dos fatores de  $f$  for igual a um dos fatores de  $g$ .

**Teorema 4.4** (Critério de Sylvester). *Dois polinômios  $f, g \in \mathcal{P}[t, \mathbb{R}]$  têm um fator comum se, e somente se,  $\det S(f, g) = 0$ .*

*Demonstração.* Consideremos a seguinte matriz  $(m + n) \times (m + n)$ ,

$$A = \begin{bmatrix} t^{n+m-1} & 0 & 0 & \cdots & 0 & 0 \\ t^{n+m-2} & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ t & 0 & 0 & \cdots & 1 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix},$$

e observemos que  $\det A = t^{n+m-1}$ . Além disso, pondo-se  $S = S(f, g)$ , tem-se

$$SA = \begin{bmatrix} t^{n-1}f(t) & a_{m-1} & a_{m-2} & \cdots & 0 \\ t^{n-2}f(t) & a_m & a_{m-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ f(t) & \cdots & & \cdots & a_0 \\ t^{n-1}g(t) & b_{m-1} & b_{m-2} & \cdots & 0 \\ t^{n-2}g(t) & b_m & b_{m-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ g(t) & \cdots & & \cdots & b_0 \end{bmatrix}.$$

Logo, considerando-se o desenvolvimento de Laplace de  $SA$  relativo à sua primeira coluna, obtém-se

$$\det(SA) = p(t)f(t) + q(t)g(t),$$

em que  $p, q \in \mathcal{P}[t, \mathbb{R}]$  são polinômios cujos graus não excedem  $n - 1$  e  $m - 1$ , respectivamente. Porém, pelo Teorema 4.3,  $\det(SA) = \det(S) \det A$ , donde

$$t^{n+m-1} \det S = p(t)f(t) + q(t)g(t). \quad (4.18)$$

Assim, se  $\det S = 0$ , teremos  $p(t)f(t) = -q(t)g(t)$  e, então,  $pf$  e  $-qg$  terão os mesmos fatores. Em particular, todos os fatores de  $f$  dividirão  $q$  ou  $g$ . Porém, o grau de  $f$  é maior que o de  $q$ . Logo, algum fator de  $f$  não dividirá  $q$  e, portanto, dividirá  $g$ , isto é,  $f$  e  $g$  terão um fator comum.

Suponhamos agora que  $f$  e  $g$  tenham um fator comum  $r$ . Nesse caso, segue-se de (4.18) que  $t^{n+m-1} \det S = \lambda(t)r(t)$ , em que  $\lambda \in \mathcal{P}[t, \mathbb{R}]$ . Se  $\lambda$  for o polinômio nulo, então  $\det S = 0$ . Caso contrário, devemos ter  $r(t) = ct^k$ , em que  $k = 1$  ou  $k = 2$ , e  $c \neq 0$ . Em particular, sendo  $r$  um fator de  $f$  e de  $g$ , tem-se  $a_0 = b_0 = 0$ , o que nos dá, igualmente,  $\det S = 0$ . ■

Aplicando-se o critério de Sylvester aos polinômios

$$f(t) = t^3 - 2t^2 + 3t - 2 \quad \text{e} \quad g(t) = 2t^2 - t - 1,$$

conclui-se que os mesmos têm um fator comum, pois sua matriz de Sylvester,

$$S(f, g) = \begin{bmatrix} 1 & -2 & 3 & -2 & 0 \\ 0 & 1 & -2 & 3 & -2 \\ 2 & -1 & -1 & 0 & 0 \\ 0 & 2 & -1 & -1 & 0 \\ 0 & 0 & 2 & -1 & -1 \end{bmatrix},$$

tem determinante nulo (verifique!).

## Exercícios

### Seções 4.1 e 4.2

1. Exiba matrizes  $A, B \in \mathcal{M}(2)$ , tais que

$$\det A = \det(A + I) \quad \text{e} \quad \det B = \det(B - I).$$

Prove que não existe uma matriz  $A \in \mathcal{M}(2)$ , tal que

$$\det A = \det(A + I) = \det(A - I).$$

2. Prove que o determinante de uma matriz triangular é igual ao produto das entradas de sua diagonal.
3. Seja  $A$  uma matriz  $n \times n$  com mais de  $n^2 - n$  entradas nulas. Mostre que  $\det A = 0$ .
4. Seja  $A = (a_{ij}) \in \mathcal{M}(n)$  uma matriz quadrada cujas entradas  $a_{ij}$  são números inteiros. Suponha que a soma das entradas de cada linha de  $A$  seja igual a  $k \in \mathbb{Z}$ , isto é,

$$\sum_{j=1}^n a_{ij} = k \quad \forall i \in \{1, \dots, n\}.$$

Denote por  $T \in \mathcal{L}(\mathbb{R}^n)$  o operador de  $\mathbb{R}^n$ , tal que  $[T] = A$ . Prove que, nessas condições:

- $k$  divide  $\det A$ ;
- Existe  $v \in \mathbb{R}^n$ , tal que  $T(v) = kv$ .

5. Seja  $A = (a_{ij}) \in \mathcal{M}(3)$  uma matriz simétrica, tal que  $a_{ii} = 1 \forall i \in \{1, 2, 3\}$  e  $|a_{ij}| \leq 1 \forall i \neq j \in \{1, 2, 3\}$ . Mostre que  $\det A \leq 1$ .
6. Prove que, para quaisquer reais  $a, b, c$ , vale a igualdade

$$\begin{vmatrix} 1 & 1 & 1 \\ a_1 & a_2 & a_3 \\ a_1^2 & a_2^2 & a_3^2 \end{vmatrix} = (a_2 - a_1)(a_3 - a_1)(a_3 - a_2).$$

Estabeleça, então, através do princípio da indução, que a *matriz de Vandermonde*,

$$A = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ a_1 & a_2 & a_3 & \cdots & a_n \\ a_1^2 & a_2^2 & a_3^2 & \cdots & a_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1^{n-1} & a_2^{n-1} & a_3^{n-1} & \cdots & a_n^{n-1} \end{bmatrix},$$

satisfaz

$$\det A = \prod_{i>j} (a_i - a_j).$$

7. Seja  $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$  um operador linear. Mostre que

$$\det(T(v_1), \dots, T(v_n)) = \lambda \det(v_1, \dots, v_n) \quad \forall v_1, \dots, v_n \in \mathbb{R}^n,$$

em que  $\lambda = \det(T(e_1), \dots, T(e_n))$ .

### Seção 4.3

8. Seja  $A$  uma matriz  $n \times n$  *antissimétrica*, isto é,  $A^* = -A$ . Prove que  $\det A = 0$ , se  $n$  for ímpar.
9. Diz-se que uma matriz  $A \in \mathcal{M}(n)$  é *ortogonal*, quando  $AA^* = I$ . Prove que o determinante de uma matriz ortogonal é sempre 1 ou  $-1$ .
10. Mostre que, quaisquer que sejam  $a, b, c, \theta \in \mathbb{R}$ , as matrizes

$$A = \begin{bmatrix} \cos \theta & -\text{sen } \theta \\ \text{sen } \theta & \cos \theta \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$$

não são semelhantes, a menos que se tenha  $A = B = I$ .

11. Sejam  $A, B \in \mathcal{M}(n)$  matrizes linha equivalentes. Mostre que  $\det A = 0$  se, e somente se,  $\det B = 0$ . Se  $\det A \neq 0$ , sob que condições teremos  $\det A = \det B$ ?
12. Prove que  $\text{adj } A^* = (\text{adj } A)^* \quad \forall A \in \mathcal{M}(n)$ .

13. Sejam  $A, B \in \mathcal{M}(n)$  matrizes invertíveis. Prove que:

a)  $\text{adj}(AB) = (\text{adj } B)(\text{adj } A)$ ;

b)  $(\text{adj } A)B = B(\text{adj } A)$ , se  $AB = BA$ .

14. (*Regra de Cramer*) Seja  $A \cong (v_1, \dots, v_n) \in \mathcal{M}(n)$  uma matriz invertível. Dado  $b \in \mathbb{R}^n$ , mostre que a (única) solução do sistema  $AX = B$ ,  $\mathcal{M}(n, 1) \ni B \cong b$ , é  $X \cong (x_1, \dots, x_n)$ , tal que

$$x_j = \frac{\det A_j}{\det A}, \quad A_j \cong (v_1, \dots, v_{j-1}, b, v_{j+1}, \dots, v_n).$$

15. Dada uma matriz  $A \in \mathcal{M}(n)$ , defina  $T : \mathcal{M}(n) \rightarrow \mathcal{M}(n)$  por

$$T(X) = AX - XA.$$

Prove que  $T$  é linear e que  $\det T = 0$ .

16. Seja  $A \cong (v_1, \dots, v_n) \in \mathcal{M}(n)$ ,  $n > 1$ , tal que, para quaisquer vetores  $w_1, \dots, w_n \in \mathbb{R}^n$ , tenha-se

$$\sum_{j=1}^n \det(v_1, v_2, \dots, v_{j-1}, w_j, v_{j+1}, \dots, v_n) = 0.$$

Prove que o posto de  $A$  é menor que  $n - 1$ .



# 5

---

## *Produto Interno*

Uma característica fundamental dos espaços vetoriais é a de admitir estruturas geométricas. Dito de outra forma, em todo espaço vetorial podem ser introduzidas as noções de *distância* e *ângulo* entre vetores, as quais generalizam aquelas da Geometria Euclidiana. Isso é feito através de uma função com propriedades especiais, denominada *produto interno*.

Assim, um produto interno é um acessório que, agregado a um espaço vetorial, enriquece sua estrutura, tornando-o um objeto matemático de grande interesse. Por exemplo, a geometria de uma superfície de  $\mathbb{R}^3$ , tal como uma esfera, um cilindro ou uma quádriga, é totalmente determinada pelo produto interno que se associa a cada um de seus planos tangentes, os quais são subespaços vetoriais de  $\mathbb{R}^3$ . Teorias mais avançadas, como Geometria Riemanniana, valem-se de ideias análogas. Em suma, parafraseando o célebre astrônomo Johannes Kepler, que afirmava: “Onde há matéria, há Geometria”; declaramos: Onde há Geometria, há produto interno.

No que se segue, discutiremos sobre o conceito de produto interno em espaços vetoriais, e de seu conseqüente, o de ortogonalidade. Como motivação, iniciaremos mostrando que a Geometria Euclidiana Plana é totalmente determinada por um produto interno, dito o produto escalar. Concluiremos, então, introduzindo a noção de volume (de paralelepípedos) em espaços euclidianos.

## 5.1 Geometria Plana e Produto Escalar

Consideremos, no plano  $\mathbb{R}^2$ , as noções de comprimento e ângulo da Geometria Euclidiana, as quais, como sabemos, relacionam-se através das lendárias igualdades trigonométricas em triângulos retângulos. Do ponto de vista da Álgebra Linear, o fato notável, o qual fundamenta a teoria que discutiremos nas próximas seções deste capítulo, é que essas noções e, portanto, toda a Geometria Euclidiana Plana, são determinadas por uma única função, dita o *produto escalar*. Vejamos, então, como isto se dá.

O *produto escalar*  $\langle v, w \rangle$  entre os vetores  $v = (x_1, x_2)$  e  $w = (y_1, y_2)$  em  $\mathbb{R}^2$  define-se através da expressão

$$\langle v, w \rangle = x_1 y_1 + x_2 y_2.$$

Representando-se vetores em  $\mathbb{R}^2$  como segmentos de reta orientados e com origem em  $\mathbf{0} = (0, 0)$ , segue-se diretamente do Teorema de Pitágoras que o comprimento de um vetor  $v = (x_1, x_2) \in \mathbb{R}^2$ , o qual chamamos *norma* de  $v$  e denotamos por  $\|v\|$ , expressa-se como

$$\|v\| = \sqrt{x_1^2 + x_2^2}.$$

Igualmente, se  $w = (y_1, y_2)$ , o comprimento do segmento de reta determinado pelas extremidades de  $v$  e  $w$ , dito a *distância* entre  $v$  e  $w$  e denotado por  $\text{dist}(v, w)$ , satisfaz (vide Fig. 5.1):

$$\text{dist}(v, w) = \|v - w\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}.$$

Observemos que, para qualquer vetor  $v = (x_1, x_2) \in \mathbb{R}^2$ , tem-se

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

Em particular, vale a igualdade

$$\text{dist}(v, w) = \sqrt{\langle v - w, v - w \rangle}.$$

Consequentemente, em  $\mathbb{R}^2$ , distância (ou comprimento) é um conceito que se exprime completamente através do produto escalar. Vejamos que o mesmo

acontece com o conceito de ângulo. Mais precisamente, verifiquemos que o ângulo  $\theta$  entre dois vetores LI,  $v = (x_1, x_2)$  e  $w = (y_1, y_2)$ , satisfaz

$$\cos \theta = \frac{\langle v, w \rangle}{\|v\| \|w\|}. \quad (5.1)$$

Em particular,  $v$  e  $w$  são ortogonais (isto é, perpendiculares entre si) se, e somente se,  $\langle v, w \rangle = 0$ .

Para obtermos (5.1), designemos por  $\theta_v$  e  $\theta_w$  os ângulos (orientados no sentido anti-horário) entre o eixo  $x_1$ , das abscissas, e os vetores  $v$  e  $w$ , respectivamente. Assim, admitindo-se  $\theta_v > \theta_w$ , tem-se que  $\theta$  satisfaz  $\theta = \theta_v - \theta_w$ , donde  $\cos \theta = \cos \theta_v \cos \theta_w + \sin \theta_v \sin \theta_w$ . Logo (vide Figura 5.1),

$$\cos \theta = \frac{x_1}{\|v\|} \frac{y_1}{\|w\|} + \frac{x_2}{\|v\|} \frac{y_2}{\|w\|} = \frac{\langle v, w \rangle}{\|v\| \|w\|}.$$

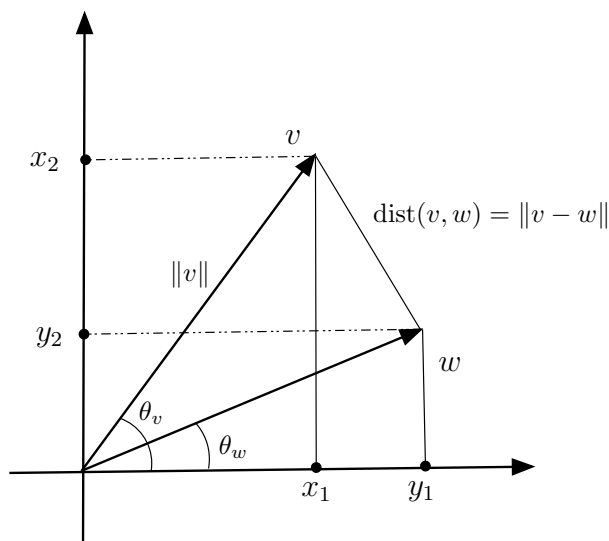


Figura 5.1

Segue-se das considerações acima que o produto escalar, de fato, determina a Geometria Eucliana Plana, no sentido de que os conceitos de comprimento

e ângulo podem ser introduzidos a partir daquele. Na próxima seção, veremos que o produto escalar é um caso particular de uma classe de funções definidas em espaços vetoriais, chamadas de produtos internos. Tais funções, conforme constataremos, generalizam o produto escalar e, analogamente, definem geometrias nos espaços vetoriais em que estão definidas.

## 5.2 Espaços Vetoriais com Produto Interno

Seja  $V$  um espaço vetorial. Uma função  $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$  é dita um *produto interno* em  $V$ , quando é *bilinear*, *simétrica* e *positiva definida*. Nessa ordem, essas propriedades significam que, para quaisquer  $u, v, w \in V$  e  $\lambda \in \mathbb{R}$ , tem-se:

- $\langle u + \lambda v, w \rangle = \langle u, w \rangle + \lambda \langle v, w \rangle$  e  $\langle u, v + \lambda w \rangle = \langle u, v \rangle + \lambda \langle u, w \rangle$ ;
- $\langle v, w \rangle = \langle w, v \rangle$ ;
- $\langle v, v \rangle \geq 0$ , ocorrendo a igualdade se, e só se,  $v = 0$ .

Um espaço vetorial  $V$  munido de um produto interno  $\langle \cdot, \cdot \rangle$  será denotado pelo par  $(V, \langle \cdot, \cdot \rangle)$ .

**Exemplo 5.1** (PRODUTO INTERNO CANÔNICO DE  $\mathbb{R}^n$ ). Sejam  $v = (x_1, \dots, x_n)$  e  $w = (y_1, \dots, y_n)$  vetores de  $\mathbb{R}^n$ . Verifica-se facilmente que o *produto escalar*

$$\langle v, w \rangle = x_1 y_1 + \dots + x_n y_n$$

define em  $\mathbb{R}^n$  um produto interno, dito *canônico*.

Advertimos que, *sempre que considerado, o espaço  $\mathbb{R}^n$  estará implicitamente munido de seu produto interno canônico, salvo menção em contrário.*

**Exemplo 5.2.** Dadas matrizes  $A, B \in \mathcal{M}(m, n)$ , observemos que o produto  $A^*B$  é uma matriz quadrada de ordem  $n$ . Logo, a função

$$\langle A, B \rangle = \text{traço}(A^*B)$$

está bem definida. Além disso, dadas matrizes  $A, B, C \in \mathcal{M}(m, n)$ , e  $\lambda \in \mathbb{R}$ , segue-se das propriedades operatórias das matrizes que

$$(A + \lambda B)^* C = (A^* + \lambda B^*) C = A^* C + \lambda B^* C.$$

Logo (vide Exercício 11),

$$\begin{aligned} \langle A + \lambda B, C \rangle &= \text{traço}(A^* C + \lambda B^* C) \\ &= \text{traço}(A^* C) + \lambda \text{traço}(B^* C) = \langle A, C \rangle + \lambda \langle B, C \rangle. \end{aligned}$$

Analogamente, verifica-se que  $\langle A, B + \lambda C \rangle = \langle A, B \rangle + \lambda \langle A, C \rangle$ , donde se conclui que  $\langle \cdot, \cdot \rangle$  é bilinear. Ademais,

$$\langle A, B \rangle = \text{traço}(A^* B) = \text{traço}(A^* B)^* = \text{traço}(B^* A) = \langle B, A \rangle,$$

isto é,  $\langle \cdot, \cdot \rangle$  é também simétrica.

Por fim, verifiquemos que  $\langle \cdot, \cdot \rangle$  é positiva definida. Para tanto, escrevamos  $A = (a_{ij})$  e observemos que, nesse caso,

$$\langle A, A \rangle = \text{traço}(A^* A) = \sum_{j=1}^n \sum_{i=1}^m a_{ij}^2,$$

donde se infere que  $\langle A, A \rangle > 0$ , a menos que a matriz  $A$  seja nula. Portanto,  $\langle \cdot, \cdot \rangle$  é positiva definida e, conseqüentemente,  $\langle \cdot, \cdot \rangle$  é um produto interno em  $\mathcal{M}(m, n)$ .

Tomemos, por exemplo, as seguintes matrizes em  $\mathcal{M}(3, 2)$ :

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

Nesse caso, temos que

$$\langle A, B \rangle = \text{traço}(A^* B) = \text{traço} \begin{bmatrix} -5 & -3 \\ 4 & -4 \end{bmatrix} = -9.$$

Seja  $V$  um espaço vetorial munido de um produto interno  $\langle \cdot, \cdot \rangle$ . A *norma* de um vetor  $v \in V$  determinada por  $\langle \cdot, \cdot \rangle$  é o número real  $\|v\|$  definido por

$$\|v\| = \sqrt{\langle v, v \rangle}. \quad (5.2)$$

A partir da norma, definimos a *distância* entre dois vetores  $v, w \in V$  por

$$\text{dist}(v, w) = \|v - w\|. \quad (5.3)$$

Observemos que, no espaço  $\mathbb{R}^n$  (munido do produto interno canônico), a distância entre dois vetores  $v = (x_1, \dots, x_n)$  e  $w = (y_1, \dots, y_n)$  corresponde à distância euclidiana usual em  $\mathbb{R}^n$ , isto é,

$$\text{dist}((x_1, \dots, x_n), (y_1, \dots, y_n)) = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}.$$

Assim, os vetores  $v = (1, -1, 2)$  e  $w = (2, 0, -1)$  de  $\mathbb{R}^3$  têm normas  $\|v\| = \sqrt{6}$  e  $\|w\| = \sqrt{5}$ , e a distância entre os mesmos é

$$\text{dist}(v, w) = \|v - w\| = \|(-1, -1, 3)\| = \sqrt{11}.$$

**Exemplo 5.3.** Sejam  $A, B \in \mathcal{M}(3, 2)$  as matrizes do Exemplo 5.2. Temos que

$$\|A\|^2 = \langle A, A \rangle = \text{traço}(A^*A) = \text{traço} \begin{bmatrix} 18 & 6 \\ 6 & 5 \end{bmatrix} = 23,$$

donde  $\|A\| = \sqrt{23}$ . Analogamente, obtém-se  $\|B\| = \sqrt{10}$ .

A relação fundamental entre o produto interno e a norma por ele definida expressa-se através da célebre Desigualdade de Cauchy-Schwarz, a qual estabeleceremos a seguir.

**Teorema 5.1** (DESIGUALDADE DE CAUCHY-SCHWARZ). *Seja  $V$  um espaço vetorial munido de um produto interno  $\langle \cdot, \cdot \rangle$ . Então,*

$$|\langle v, w \rangle| \leq \|v\| \|w\| \quad \forall v, w \in V, \quad (5.4)$$

*ocorrendo a igualdade se, e somente se,  $v$  e  $w$  são múltiplos um do outro.*

*Demonstração.* O resultado é imediato quando  $v$  ou  $w$  é o vetor nulo. Suponhamos, então, que  $v$  e  $w$  sejam ambos não nulos e definamos a função:

$$f(t) = \langle v - tw, v - tw \rangle, \quad t \in \mathbb{R}.$$

Temos que  $f(t) \geq 0 \forall t \in \mathbb{R}$ , e que  $f(t) = 0$  para algum  $t \in \mathbb{R}$  se, e somente se,  $v = tw$ , pois o produto interno é positivo definido. Além disso, pela bilinearidade do produto interno,

$$f(t) = \|v\|^2 - 2\langle v, w \rangle t + \|w\|^2 t^2,$$

isto é,  $f$  é uma função quadrática não negativa de variável  $t$ . Em particular, o discriminante de  $f$  é não positivo. Logo,

$$4\langle v, w \rangle^2 - 4\|v\|^2\|w\|^2 \leq 0,$$

o que implica (5.4). Temos também que o discriminante de  $f$  é nulo se, e somente se,  $f$  possui um único zero, donde se infere que ocorre a igualdade em (5.4) se, e somente se,  $v = tw$  para algum  $t \in \mathbb{R}$ . ■

A desigualdade de Cauchy-Schwarz tem inúmeras consequências. A primeira que destacaremos é a prova do cumprimento da desigualdade triangular por normas advindas de produtos internos, conforme a proposição seguinte.

**Proposição 5.1** (PROPRIEDADES FUNDAMENTAIS DA NORMA). *A norma  $\|\cdot\|$  advinda de um produto interno  $\langle \cdot, \cdot \rangle$  de um espaço vetorial  $V$  tem as seguintes propriedades, as quais são válidas para quaisquer  $v, w \in V$  e  $\lambda \in \mathbb{R}$ :*

- i)  $\|v\| \geq 0$  e  $\|v\| = 0 \Leftrightarrow v = 0$  (positividade);
- ii)  $\|\lambda v\| = |\lambda|\|v\|$  (homogeneidade);
- iii)  $\|v + w\| \leq \|v\| + \|w\|$  (desigualdade triangular).

*Demonstração.* As propriedades (i) e (ii) decorrem imediatamente da definição de norma e da positividade e bilinearidade do produto interno. Já a desigualdade triangular, ela decorre da Desigualdade de Cauchy-Schwarz. De fato, dados vetores  $v, w \in V$ , temos  $\|v + w\|^2 = \langle v + w, v + w \rangle$ . Logo,

$$\|v + w\|^2 = \|v\|^2 + 2\langle v, w \rangle + \|w\|^2 \leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2,$$

donde se obtém (iii). ■

A desigualdade de Cauchy-Schwarz também nos permite introduzir a noção de ângulo entre vetores em qualquer espaço vetorial  $V$  munido de um produto interno  $\langle \cdot, \cdot \rangle$ . Mais precisamente, dados vetores não nulos  $v, w \in V$ , definimos o *ângulo*  $\sphericalangle(v, w) \in [0, \pi]$  entre  $v$  e  $w$  por (compare com (5.1)):

$$\sphericalangle(v, w) := \arccos \frac{\langle v, w \rangle}{\|v\| \|w\|}. \quad (5.5)$$

Note que, pela Desigualdade de Cauchy-Schwarz,

$$-1 \leq \frac{\langle v, w \rangle}{\|v\| \|w\|} \leq 1,$$

de modo que  $\sphericalangle(v, w)$  está bem definido.

Calculemos o ângulo entre os vetores  $v = (1, -1, 2)$  e  $w = (2, 0, -1)$  de  $\mathbb{R}^3$ , considerados acima. Temos que,

$$\sphericalangle(v, w) = \arccos \frac{\langle v, w \rangle}{\|v\| \|w\|} = \arccos \frac{0}{\|\sqrt{6}\| \|\sqrt{5}\|} = 0,$$

isto é,  $\sphericalangle(v, w) = \pi/2$ .

**Exemplo 5.4.** Tomemos, uma vez mais, as matrizes  $A$  e  $B$  do Exemplo 5.2. Considerando-se também os resultados do Exemplo 5.3, temos que

$$\sphericalangle(A, B) = \arccos \frac{\langle A, B \rangle}{\|A\| \|B\|} = \arccos \frac{-9}{\sqrt{23}\sqrt{10}} \approx \frac{7\pi}{10}.$$

A desigualdade de Cauchy-Schwarz pode ser aplicada para se obter desigualdades numéricas não triviais, conforme ilustrado no exemplo a seguir.

**Exemplo 5.5.** Verifiquemos que, para quaisquer reais  $a, b \geq 1$ , tem-se

$$\sqrt{a^2 - 1} + \sqrt{b^2 - 1} \leq ab.$$

Com efeito, tomando-se em  $\mathbb{R}^2$  o produto interno canônico e aplicando-se a desigualdade de Cauchy-Schwarz aos vetores  $v = (\sqrt{a^2 - 1}, 1)$  e  $w = (1, \sqrt{b^2 - 1})$ , obtém-se

$$\sqrt{a^2 - 1} + \sqrt{b^2 - 1} = \langle v, w \rangle \leq \|v\| \|w\| = ab.$$



## 5.3 Ortogonalidade

Em Geometria Euclidiana Plana, o conceito de ortogonalidade está envolvido em vários resultados fundamentais dessa teoria; notadamente, no Teorema de Pitágoras. Considerando-se (5.5), este conceito se estende naturalmente ao contexto dos espaços vetoriais, conforme a definição seguinte.

**Definição 5.1** (ORTOGONALIDADE). Seja  $(V, \langle \cdot, \cdot \rangle)$  um espaço vetorial com produto interno  $\langle \cdot, \cdot \rangle$ . Dizemos que  $v, w \in V$  são *ortogonais*, quando  $\langle v, w \rangle = 0$ .

Dados vetores  $v, w \in (V, \langle \cdot, \cdot \rangle)$ , tem-se

$$\|v + w\|^2 = \langle v + w, v + w \rangle = \|v\|^2 + 2\langle v, w \rangle + \|w\|^2.$$

Vale, portanto, o seguinte resultado.

**Teorema 5.2.** *Dois vetores  $v, w \in (V, \langle \cdot, \cdot \rangle)$  são ortogonais se, e somente se, cumprem a igualdade  $\|v + w\|^2 = \|v\|^2 + \|w\|^2$ .*

Note que a parte “somente se” do teorema acima constitui uma versão do Teorema de Pitágoras para espaços vetoriais, já que, em  $\mathbb{R}^2$ , quando  $v$  e  $w$  são ortogonais,  $v$ ,  $w$ , e  $v + w$  são os lados (ou vértices, se vistos como pontos) de um triângulo retângulo.

**Definição 5.2** (ORTOGONALIDADE DE CONJUNTOS). Um subconjunto de  $(V, \langle \cdot, \cdot \rangle)$  é dito *ortogonal*, quando seus elementos são, dois a dois, ortogonais. Um subconjunto ortogonal cujos vetores são todos unitários (isto é, de norma 1) é dito *ortonormal*.

Observe que todo conjunto ortogonal  $\mathcal{C} = \{v_1, \dots, v_k\} \subset (V, \langle \cdot, \cdot \rangle)$  cujos vetores são não nulos é LI. A fim de verificarmos essa propriedade, consideremos uma combinação linear nula desses vetores, isto é,

$$\lambda_1 v_1 + \dots + \lambda_i v_i + \dots + \lambda_k v_k = \mathbf{0}.$$

Para um dado  $i \in \{1, \dots, k\}$ , tomando-se o produto interno por  $v_i$  em ambos os membros da igualdade acima, obtém-se  $\lambda_i \langle v_i, v_i \rangle = 0$ , o que nos dá  $\lambda_i = 0$ , provando, assim, a independência linear de  $\mathcal{C}$ .

No que se segue, através de um bem conhecido método denominado *ortogonalização de Gram-Schmidt*, mostraremos que todo espaço vetorial  $(V, \langle \cdot, \cdot \rangle)$  de dimensão finita admite uma base ortonormal. O fundamento desse método é o conceito de *projecção ortogonal*, que ora introduzimos.

**Definição 5.3** (PROJEÇÃO ORTOGONAL). Dados  $v, w \in (V, \langle \cdot, \cdot \rangle)$ ,  $w \neq \mathbf{0}$ , o vetor

$$P_w(v) = \frac{\langle v, w \rangle}{\langle w, w \rangle} w$$

denomina-se a *projecção ortogonal* de  $v$  sobre (o espaço gerado por)  $w$ .

O motivo da nomenclatura, na definição acima, reside no fato de o vetor  $v - P_w(v)$  ser ortogonal a  $w$  (vide Fig. 5.2). Com efeito,

$$\langle v - P_w(v), w \rangle = \langle v, w \rangle - \langle P_w(v), w \rangle = \langle v, w \rangle - \frac{\langle v, w \rangle}{\langle w, w \rangle} \langle w, w \rangle = 0.$$

Acrescentamos, ainda, que a sentença entre parênteses deve-se à igualdade

$$P_{\lambda w}(v) = P_w(v) \quad \forall \lambda \neq 0,$$

a qual é de verificação imediata.

Observemos que, nas condições do parágrafo anterior, se  $v$  e  $w$  forem LI, então  $\{v - P_w(v), w\}$  será uma base ortogonal do subespaço  $\text{Ger}\{v, w\} \subset V$ . Nesse caso, multiplicando-se cada um desses vetores pelos inversos de suas respectivas normas, obtém-se uma base ortonormal  $\mathcal{B}$  de  $\text{Ger}\{v, w\}$ . Em suma, a partir da base  $\{v, w\}$  de  $\text{Ger}\{v, w\}$ , obtivemos uma base ortogonal desse subespaço substituindo-se o vetor  $v$  por  $v - P_w(v)$ . Em seguida, “normalizamos” essa base ortogonal a fim de obter a base ortonormal  $\mathcal{B}$ .

**Exemplo 5.6.** Consideremos em  $\mathbb{R}^3$  os vetores  $v = (1, 0, -1)$  e  $w = (1, 1, 0)$ , e apliquemos ortogonalização de Gram-Schmidt para obter uma base ortonormal de  $\text{Ger}\{v, w\}$ . Temos que  $\langle v, w \rangle / \langle w, w \rangle = 1/2$ , donde

$$P_w(v) = \frac{\langle v, w \rangle}{\langle w, w \rangle} w = (1/2, 1/2, 0)$$

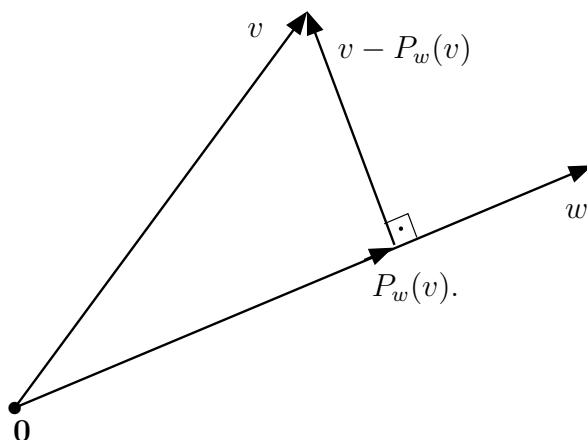


Figura 5.2: Projeção de  $v$  sobre  $\text{Ger}\{w\}$ .

e, portanto,  $v - P_w(v) = (1/2, -1/2, -1)$ . Logo,  $\{(1/2, -1/2, -1), (1, 1, 0)\}$  é uma base ortogonal de  $\text{Ger}\{v, w\}$ . Normalizando-a, obtemos a base ortonormal

$$\mathcal{B} = \{(\sqrt{6}/6, -\sqrt{6}/6, -\sqrt{2}/\sqrt{3}), (\sqrt{2}/2, \sqrt{2}/2, 0)\}.$$

Procedendo-se indutivamente, podemos aplicar ortogonalização de Gram-Schmidt a um conjunto arbitrário de vetores LI em  $(V, \langle \cdot, \cdot \rangle)$ . Em particular, a uma base. Para constatar isso, basta observarmos que, se  $\{v_1, \dots, v_k\}$  é uma base ortogonal de um subespaço  $W$  de  $V$ , e  $v \in V - W$ , então o vetor

$$w_k = \lambda_1 v_1 + \dots + \lambda_k v_k, \quad \lambda_i = \frac{\langle v_i, v \rangle}{\langle v_i, v_i \rangle},$$

satisfaz a igualdade (verifique!)

$$\langle v - w_k, v_i \rangle = 0 \quad \forall i \in 1, \dots, k.$$

Logo,  $\{v_1, \dots, v_{k-1}, v - w_k\}$  é uma base ortogonal do subespaço de  $V$ , o qual é gerado por esses vetores.

Das considerações acima, segue-se o seguinte resultado.

**Proposição 5.2** (ORTOGONALIZAÇÃO DE GRAM-SCHMIDT). *Seja  $\{v_1, \dots, v_n\}$  uma base de um espaço vetorial  $(V, \langle, \rangle)$ . Então, o conjunto  $\{v'_1, \dots, v'_n\} \subset V$  definido pelas relações de recorrência:*

$$v'_1 = v_1$$

$$v'_{i+1} = v_{i+1} - \sum_{k=1}^i \frac{\langle v_{i+1}, v'_k \rangle}{\langle v'_k, v'_k \rangle} v'_k, \quad i \in \{1, \dots, n-1\},$$

*constitui uma base ortogonal de  $V$ . Em particular, o conjunto*

$$\mathcal{B} = \left\{ \frac{v'_1}{\|v'_1\|}, \dots, \frac{v'_n}{\|v'_n\|} \right\} \subset V$$

*é uma base ortonormal de  $V$ .*

Seja  $\mathcal{B} = \{u_1, \dots, u_n\}$  uma base ortonormal de  $(V, \langle, \rangle)$ . Dados vetores  $v = (x_1, \dots, x_n)_{\mathcal{B}}$ ,  $w = (y_1, \dots, y_n)_{\mathcal{B}} \in V$ , temos que

$$\langle v, w \rangle = \langle x_1 u_1 + \dots + x_n u_n, y_1 u_1 + \dots + y_n u_n \rangle = x_1 y_1 + \dots + x_n y_n,$$

isto é, em coordenadas com respeito a uma base ortonormal, o produto interno em  $(V, \langle, \rangle)$  expressa-se como o produto interno canônico de  $\mathbb{R}^n$ . Deste simples fato, segue-se que todo funcional linear  $f \in \mathcal{L}(V, \mathbb{R})$  exprime-se como um produto interno por um vetor fixo, conforme o teorema seguinte.

**Teorema de Riesz.** *Seja  $(V, \langle, \rangle)$  um espaço vetorial de dimensão finita munido de um produto interno  $\langle, \rangle$ . Então, para todo funcional linear  $f \in \mathcal{L}(V, \mathbb{R})$ , existe um único vetor  $a \in V$ , tal que*

$$f(v) = \langle a, v \rangle \quad \forall v \in V.$$

*Demonstração.* Tome uma base ortonormal  $\mathcal{B} = \{u_1, \dots, u_n\}$ , de  $V$ , e defina o vetor  $a = f(u_1)u_1 + \dots + f(u_n)u_n$ . Assim, dado  $v = x_1 u_1 + \dots + x_n u_n \in V$ , tem-se  $f(v) = x_1 f(u_1) + \dots + x_n f(u_n) = \langle a, v \rangle$ .

Suponhamos, agora, que exista  $b \in V$ , tal que  $f(v) = \langle b, v \rangle \quad \forall v \in V$ . Nesse caso, para todo  $v \in V$ , teremos  $\langle a, v \rangle = \langle b, v \rangle$ , isto é,

$$\langle a - b, v \rangle = 0 \quad \forall v \in V.$$

Fazendo-se, na igualdade acima,  $v = a - b$ , conclui-se imediatamente que  $a = b$  e, portanto, que  $a$  é único. ■

A fim de estender o conceito de projeção ortogonal, tomemos um vetor  $v \in (V, \langle, \rangle)$  e um subespaço não trivial  $W \subset V$ . Verifiquemos, então, que existe um único vetor  $w \in W$ , tal que  $v - w$  é ortogonal a todos os vetores do subespaço  $W$ .

Para tanto, valendo-nos da Proposição 5.2, fixemos uma base ortonormal  $\{u_1, \dots, u_k\}$  de  $W$ . Definindo-se, então,

$$w = \sum_{i=1}^k \langle v, u_i \rangle u_i, \quad (5.6)$$

verifica-se facilmente que  $\langle v - w, u_i \rangle = 0$  para todo  $i = 1, \dots, k$ . Logo, pela bilinearidade do produto interno, temos que  $v - w$  é ortogonal a todos os vetores de  $W$ .

Suponhamos agora que exista um vetor  $w' \in W$ , tal que  $v - w'$  seja ortogonal a todos os vetores de  $W$ . Nesse caso, escrevendo-se  $w' = \lambda_1 u_1 + \dots + \lambda_k u_k$ , para cada  $i = 1, \dots, k$ , tem-se

$$0 = \langle v - w', u_i \rangle = \langle v, u_i \rangle - \langle w', u_i \rangle = \langle v, u_i \rangle - \lambda_i,$$

donde  $\lambda_i = \langle v, u_i \rangle$  e, portanto,  $w' = w$ .

O vetor  $w$  em (5.6) é dito a *projeção ortogonal* de  $v$  sobre  $W$ , e é denotado por  $P_W(v)$ . Dessa forma, fica definida a *aplicação projeção ortogonal* de  $V$  em  $W$ ,  $P_W : V \rightarrow W$ , dada por

$$P_W(v) = \langle v, u_1 \rangle u_1 + \dots + \langle v, u_k \rangle u_k, \quad v \in V.$$

Note que, pela bilinearidade do produto interno,  $P_W$  é linear! Além disso, essa aplicação tem as seguintes propriedades, cuja verificação deixamos a cargo do leitor (vide Exercício 14).

- $P_W \circ P_W = P_W$  (*idempotência*);
- $\langle P_W(v), v' \rangle = \langle v, P_W(v') \rangle \quad \forall v, v' \in V$  (*simetria*);

- $\langle P_W(v), v \rangle \geq 0 \forall v \in V$  (*positividade*);
- $\|P_W(v)\| \leq \|v\| \forall v \in V$  e  $\|P_W(v)\| = \|v\| \Leftrightarrow v \in W$  (*semi-contratilidade*).

O conceito de ortogonalidade nos permite decompor um espaço vetorial  $(V, \langle \cdot, \cdot \rangle)$  numa soma direta de subespaços, os quais são ortogonais entre si, o que nos leva à seguinte definição.

**Definição 5.4** (COMPLEMENTO ORTOGONAL). Dado um subespaço vetorial  $W$  de  $(V, \langle \cdot, \cdot \rangle)$ , chama-se o conjunto

$$W^\perp = \{v \in V; \langle v, w \rangle = 0 \forall w \in W\}$$

de *complemento ortogonal* de  $W$  em  $V$ .

**Proposição 5.3** (PROPRIEDADES DO COMPLEMENTO ORTOGONAL). *O complemento ortogonal  $W^\perp$  de um subespaço  $W \subset (V, \langle \cdot, \cdot \rangle)$  é um subespaço vetorial de  $V$ , o qual tem as seguintes propriedades:*

- i)  $V = W \oplus W^\perp$ ;
- ii)  $W^{\perp\perp} = W$ .

*Demonstração.* É imediato que  $\mathbf{0} \in W^\perp$ . Além disso, dados  $v, v' \in W^\perp$ ,  $\lambda \in \mathbb{R}$ , e  $w \in W$ , tem-se  $\langle v + \lambda v', w \rangle = \langle v, w \rangle + \lambda \langle v', w \rangle = 0$ , donde se infere que  $W^\perp$  é um subespaço vetorial de  $V$ .

A fim de provar as asserções (i) e (ii), tomemos uma base ortonormal de  $W$ ,  $\{u_1, \dots, u_k\}$ . Usando ortogonalização de Gram-Schmidt, essa base pode ser completada para uma base ortonormal de  $V$ ,  $\{u_1, \dots, u_k, u_{k+1}, \dots, u_n\}$ . Assim, pondo-se  $V' = \text{Ger}\{u_{k+1}, \dots, u_n\}$ , tem-se  $V = W \oplus V'$ . Resta-nos, pois, mostrar que  $V' = W^\perp$ . A inclusão  $V' \subset W^\perp$  é imediata. Ademais, se  $v \in W^\perp$ , tem-se  $\langle v, u_i \rangle = 0 \forall i = 1, \dots, k$ , donde  $v \in V'$ . Logo,  $V' = W^\perp$ , o que prova (i).

Quanto à afirmação (ii), observemos que, pelo estabelecido em (i),

$$W \oplus W^\perp = V = W^\perp \oplus W^{\perp\perp},$$

o que nos dá  $\dim W^{\perp\perp} = \dim W$  (vide Proposição 2.16). Além disso, pelo provado em (i), dado  $v \in W^{\perp\perp}$ , existem únicos  $w \in W$  e  $w^\perp \in W^\perp$ , tais que  $v = w + w^\perp$ . Porém,  $0 = \langle v, w^\perp \rangle = \langle w^\perp, w^\perp \rangle$ , donde  $w^\perp = \mathbf{0}$ , isto é,  $v = w$ . Segue-se que  $W^{\perp\perp} \subset W$  e, portanto, que  $W = W^{\perp\perp}$ , uma vez que esses subespaços têm dimensões iguais. ■

Ilustremos a força conceitual do produto interno e seus correlatos estabelecendo uma propriedade do espaço euclidiano  $\mathbb{R}^n$  visto como espaço métrico, isto é, como um conjunto de pontos com uma distância (no caso, a euclidiana) bem definida. Quando  $n = 2$ , essa propriedade reflete o fato de que o maior (em cardinalidade) conjunto de pontos equidistantes do plano euclidiano é constituído pelos vértices de um triângulo equilátero.

**Proposição 5.4** (PONTOS EQUIDISTANTES EM  $\mathbb{R}^n$ ). *Seja  $\mathfrak{C}$  um subconjunto do espaço  $\mathbb{R}^n$  cujos pontos são equidistantes, isto é, existe  $\epsilon > 0$ , tal que*

$$\text{dist}(p_1, p_2) = \epsilon \quad \forall p_1, p_2 \in \mathfrak{C}.$$

*Então,  $\mathfrak{C}$  não contém mais que  $n + 1$  pontos.*

*Demonstração.* Suponhamos, por absurdo, que existam  $n + 2$  pontos equidistantes em  $\mathfrak{C}$ ,  $p_0, \dots, p_{n+1}$ , e tomemos um sistema de coordenadas ortogonais em  $\mathbb{R}^n$ , tal que  $p_0$  seja a origem  $\mathbf{0}$ . Dotemos  $\mathbb{R}^n$ , então, de sua estrutura de espaço vetorial com produto interno canônico, de modo que os pontos  $p_1, \dots, p_{n+1}$  passam a ser vetores de  $\mathbb{R}^n$ , os quais passaremos a denotar por  $v_1, \dots, v_{n+1}$ .

Mostremos que, nas condições dadas, esses vetores são LI, contradizendo que  $\dim \mathbb{R}^n = n$ , e provando, desta forma, o resultado. Para tanto, tomemos uma combinação linear nula de  $v_1, \dots, v_{n+1}$ ,

$$a_1 v_1 + \dots + a_{n+1} v_{n+1} = \mathbf{0}. \quad (5.7)$$

Podemos supor, sem perda de generalidade, que  $\epsilon = 1$ , de modo que

$$\|v_i\|^2 = \|v_j\|^2 = \|v_i - v_j\|^2 = 1 \quad \forall i \neq j \in \{1, \dots, n+1\}.$$

Dessas igualdades, e da bilinearidade do produto interno, obtém-se, para  $i \neq j$ ,  $\langle v_i, v_j \rangle = 1/2$ . Assim, para cada  $i \in \{1, \dots, n+1\}$ , tomando-se o produto interno por  $v_i$  em ambos os membros da igualdade (5.7), chega-se a

$$a_i + \frac{1}{2}(a_1 + \dots + a_{i-1} + a_{i+1} + \dots + a_{n+1}) = 0.$$

Adicionando-se  $a_i/2$  a ambos os membros dessa última igualdade, e denotando-se por  $S$  a soma de todos os coeficientes  $a_i$ , obtém-se  $a_i = -S$ . Logo,

$$S = \sum_{i=1}^{n+1} a_i = -(n+1)S,$$

donde  $S = 0$  e, portanto  $a_i = 0$ . Desta forma, os vetores  $v_1, \dots, v_{n+1}$  são LI, o que constitui a nossa anunciada contradição. ■

## 5.4 Volume em Espaços Euclidianos

No que diz respeito a medição de objetos do espaço euclidiano  $\mathbb{R}^3$ , as “medidas” fundamentais são: comprimento, área e volume. Através da teoria que desenvolvemos neste capítulo e no anterior, podemos estender essas noções para o espaço  $\mathbb{R}^n$ , isto é, para todo  $k \in \{1, \dots, n\}$ , podemos definir um “volume  $k$ -dimensional” de subconjuntos de  $\mathbb{R}^n$ .

A introdução de um conceito de volume para subconjuntos arbitrários de  $\mathbb{R}^n$  é uma questão bastante delicada e não será tratada neste texto. Na verdade, esse problema é o objeto de uma vasta teoria, dita, *da medida*. No entanto, o primeiro passo na direção da construção do conceito geral de volume, o que constitui aqui nosso propósito, é o de estabelecê-lo para paralelepípedos de  $\mathbb{R}^n$ .

Assim como o fizemos com as noções de comprimento e ângulo, usaremos a Geometria Euclidiana Plana como guia. Mais precisamente, consideraremos o conceito de área de paralelogramo à luz da Álgebra Linear.

Com esse intuito, tomemos vetores linearmente independentes

$$v = (a, b) \text{ e } w = (c, d) \in \mathbb{R}^2,$$



denotemos por  $\mathcal{P}(v, w)$  o paralelogramo de lados  $v$  e  $w$ , e provemos que

$$|\mathcal{P}(v, w)| = |\det(v, w)|, \quad (5.8)$$

em que  $|\mathcal{P}(v, w)|$  denota a área de  $\mathcal{P}(v, w)$ , a ser entendida aqui como o produto do comprimento de um de seus lados, dito a base, pela altura correspondente. Tomando-se  $w$  como base de  $\mathcal{P}(v, w)$ , a altura será a norma da projeção de  $v$  sobre o vetor  $w^\perp = (-d, c)$ , pois  $\langle w^\perp, w \rangle = 0$  (Fig. 5.4).

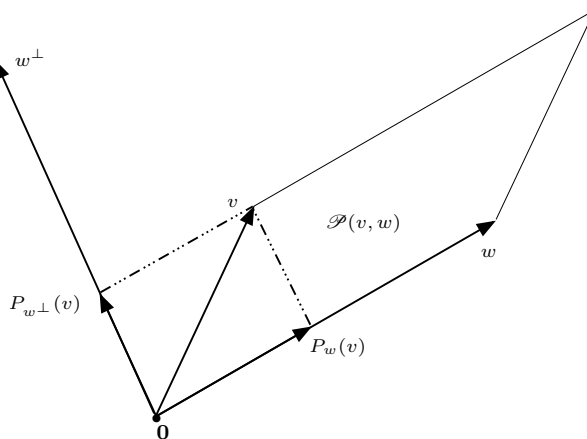


Figura 5.3

Logo, uma vez que  $\|w\| = \|w^\perp\|$  e  $\langle w^\perp, v \rangle = -ad + bc = -\det(v, w)$ , tem-se

$$|\mathcal{P}(v, w)| = \|w\| \|P_{w^\perp}(v)\| = \|w\| \frac{|\langle v, w^\perp \rangle|}{\langle w^\perp, w^\perp \rangle} \|w^\perp\| = |\det(v, w)|.$$

Uma outra expressão da área de  $\mathcal{P}(v, w)$  pode ser obtida a partir da igualdade  $v = P_w(v) + P_{w^\perp}(v)$ . De fato, pelo Teorema de Pitágoras, tem-se  $\|v\|^2 = \|P_w(v)\|^2 + \|P_{w^\perp}(v)\|^2$ . Desta forma,

$$\|v\|^2 = \frac{\langle v, w \rangle^2}{\langle w, w \rangle^2} \|w\|^2 + \frac{\langle v, w^\perp \rangle^2}{\langle w^\perp, w^\perp \rangle^2} \|w^\perp\|^2 = \frac{\langle v, w \rangle^2}{\langle w, w \rangle} + \frac{|\det(v, w)|^2}{\langle w, w \rangle}.$$

Daí e de (5.8), obtém-se

$$|\mathcal{P}(v, w)|^2 = \langle v, v \rangle \langle w, w \rangle - \langle v, w \rangle^2 = \det \begin{bmatrix} \langle v, v \rangle & \langle v, w \rangle \\ \langle w, v \rangle & \langle w, w \rangle \end{bmatrix}.$$

A matriz na última igualdade acima é dita a matriz *de Gram* dos vetores  $v$  e  $w$ , e seu determinante é denotado por  $g(v, w)$ . Assim, temos

$$|\mathcal{P}(v, w)| = \sqrt{g(v, w)}. \quad (5.9)$$

Note que podemos expressar o comprimento de um vetor  $v \in \mathbb{R}^2$  através da matriz de Gram  $1 \times 1$ ,  $[\langle v, v \rangle]$ , uma vez que  $\|v\| = \sqrt{\langle v, v \rangle} = \sqrt{g(v)}$ . Convém mencionar também que podemos descrever  $\mathcal{P}(v, w)$  como

$$\mathcal{P}(v, w) = \{\lambda v + \mu w; \lambda, \mu \in [0, 1]\}.$$

**Exemplo 5.7.** Dados  $v = (3, -1), w = (2, 2) \in \mathbb{R}^2$ , temos

$$|\mathcal{P}(v, w)| = |\det(v, w)| = 8.$$

Observemos também que

$$g(v, w) = \det \begin{bmatrix} \langle v, v \rangle & \langle v, w \rangle \\ \langle w, v \rangle & \langle w, w \rangle \end{bmatrix} = \det \begin{bmatrix} 10 & 4 \\ 4 & 8 \end{bmatrix} = 64 = |\mathcal{P}(v, w)|^2,$$

confirmando, assim, a igualdade (5.9).

Mais geralmente, o *paralelepípedo  $k$ -dimensional* de  $\mathbb{R}^n$  determinado por  $k$  vetores linearmente independentes,  $v_1, \dots, v_k$ , é definido por:

$$\mathcal{P}(v_1, \dots, v_k) = \{\lambda_1 v_1 + \dots + \lambda_k v_k; \lambda_i \in [0, 1] \forall i = 1, \dots, k\}.$$

Pondo-se  $W_{k-1} = \text{Ger}\{v_1, \dots, v_{k-1}\}$ , o *volume  $k$ -dimensional* do paralelepípedo  $\mathcal{P}(v_1, \dots, v_k)$ , a ser denotado por  $|\mathcal{P}(v_1, \dots, v_k)|$ , é definido indutivamente da seguinte forma (Fig. 5.4):

$$|\mathcal{P}(v_1, \dots, v_k)| = \begin{cases} \|v_1\|, & \text{se } k = 1. \\ |\mathcal{P}(v_1, \dots, v_{k-1})| \|v_k - P_{W_{k-1}}(v_k)\|, & \text{se } k > 1. \end{cases}$$

Consideremos, então, a *matriz de Gram* de  $v_1, \dots, v_k$ ,

$$G(v_1, \dots, v_k) = (\langle v_i, v_j \rangle)_{k \times k},$$

e denotemos seu determinante por  $g(v_1, \dots, v_k)$ .

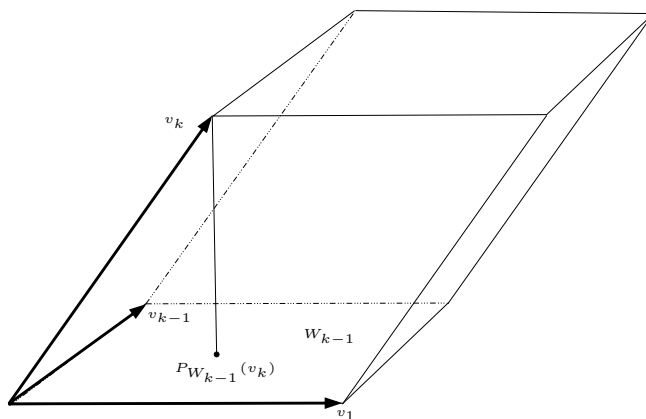


Figura 5.4: Paralelepípedo  $\mathcal{P}(v_1, \dots, v_k)$ .

**Proposição 5.5.** Para todo paralelepípedo  $\mathcal{P}(v_1, \dots, v_k) \subset \mathbb{R}^n$ , vale a igualdade

$$|\mathcal{P}(v_1, \dots, v_k)| = \sqrt{g(v_1, \dots, v_k)}. \tag{5.10}$$

*Demonstração.* Temos que a igualdade (5.10) é trivial para  $k = 1$ . Suponhamos, por indução, que ela seja verdadeira para  $k \in \mathbb{N}$ . Pondo-se, então,  $P_{W_k}(v_{k+1}) = \lambda_1 v_1 + \dots + \lambda_k v_k$  e  $w_{k+1} = v_{k+1} - P_{W_k}(v_{k+1})$ , tem-se

$$v_{k+1} = \sum_{j=1}^k \lambda_j v_j + w_{k+1}.$$

Assim, denotando-se por  $c_1, \dots, c_{k+1}$  os vetores coluna da matriz de Gram

$G(v_1, \dots, v_{k+1})$ , tem-se

$$\begin{aligned}
 c_{k+1} &= \sum_{i=1}^{k+1} \langle v_{k+1}, v_i \rangle e_i = \sum_{i=1}^{k+1} \left\langle \sum_{j=1}^k \lambda_j v_j + w_{k+1}, v_i \right\rangle e_i \\
 &= \sum_{i=1}^{k+1} \left\langle \sum_{j=1}^k \lambda_j v_j, v_i \right\rangle e_i + \sum_{i=1}^{k+1} \langle w_{k+1}, v_i \rangle e_i \\
 &= \sum_{j=1}^k \sum_{i=1}^{k+1} \langle \lambda_j v_j, v_i \rangle e_i + \langle w_{k+1}, v_{k+1} \rangle e_{k+1} \\
 &= \sum_{j=1}^k \lambda_j c_j + \|w_{k+1}\|^2 e_{k+1}.
 \end{aligned}$$

Daí, observando-se que  $\det(c_1, \dots, c_k, e_{k+1}) = g(v_1, \dots, v_k)$ , e lembrando-se que  $\det$  é uma forma multilinear alternada, obtém-se

$$\begin{aligned}
 g(v_1, \dots, v_{k+1}) &= \det(c_1, \dots, c_k, c_{k+1}) \\
 &= \det \left( c_1, \dots, c_k, \sum_{j=1}^k \lambda_j c_j + \|w_{k+1}\|^2 e_{k+1} \right) \\
 &= \|w_{k+1}\|^2 g(v_1, \dots, v_k) = \|w_{k+1}\|^2 |\mathcal{P}(v_1, \dots, v_k)|^2 \\
 &= |\mathcal{P}(v_1, \dots, v_{k+1})|^2.
 \end{aligned}$$

O resultado, portanto, segue-se do Princípio da Indução. ■

Note que, na Proposição 5.5, provamos também que  $g(v_1, \dots, v_k) > 0$  sempre que  $v_1, \dots, v_k$  são LI.

**Corolário 5.1.** *Para todo paralelepípedo  $n$ -dimensional  $\mathcal{P}(v_1, \dots, v_n) \subset \mathbb{R}^n$ , vale a seguinte igualdade:*

$$|\mathcal{P}(v_1, \dots, v_n)| = |\det(v_1, \dots, v_n)|. \quad (5.11)$$

*Demonstração.* Basta observarmos que, se  $A$  é a matriz  $n \times n$  cujos vetores linha são  $v_1, \dots, v_n$ , então  $G(v_1, \dots, v_n) = AA^*$  (vide Exercício 12). Logo,

$$|\mathcal{P}(v_1, \dots, v_n)|^2 = \det G(v_1, \dots, v_n) = (\det A)^2,$$

donde se obtém (5.11). ■

**Exemplo 5.8.** Consideremos os vetores de  $\mathbb{R}^3$ ,  $v_1 = (1, 0, 1)$ ,  $v_2 = (2, -1, 1)$  e  $v_3 = (1, 1, -1)$ . Um cálculo direto nos dá  $g(v_1, v_2) = 3$ ,  $g(v_1, v_3) = 6$  e  $g(v_2, v_3) = 18$ . Logo, as áreas dos correspondentes paralelogramos são:

$$|\mathcal{P}(v_1, v_2)| = \sqrt{3}, \quad |\mathcal{P}(v_1, v_3)| = \sqrt{6} \quad \text{e} \quad |\mathcal{P}(v_2, v_3)| = 3\sqrt{2}.$$

Quanto ao volume do paralelepípedo determinado por  $v_1$ ,  $v_2$  e  $v_3$ , temos

$$|\mathcal{P}(v_1, v_2, v_3)| = |\det(v_1, v_2, v_3)| = |3| = 3.$$

## Exercícios

### Seção 5.2

1. Sejam  $\langle \cdot, \cdot \rangle_1$  e  $\langle \cdot, \cdot \rangle_2$  produtos internos num espaço vetorial  $V$  e  $\lambda$  um número real positivo. Mostre que

$$\langle v, w \rangle = \langle v, w \rangle_1 + \lambda \langle v, w \rangle_2, \quad v, w \in V,$$

é um produto interno em  $V$ .

2. Seja  $T : V \rightarrow (W, \langle \cdot, \cdot \rangle)$  uma aplicação linear injetiva. Mostre que:
  - i) A função  $\langle u, v \rangle_0 = \langle T(u), T(v) \rangle$ ,  $u, v \in V$ , define um produto interno  $\langle \cdot, \cdot \rangle_0$  em  $V$ , dito *induzido* por  $T$ .
  - ii) O produto interno definido em  $\mathcal{M}(m, n)$  no Exemplo 5.2 é aquele induzido pelo isomorfismo natural  $T : \mathcal{M}(m, n) \rightarrow (\mathbb{R}^{mn}, \langle \cdot, \cdot \rangle)$ , em que  $\langle \cdot, \cdot \rangle$  é o produto interno canônico de  $\mathbb{R}^{mn}$ .
3. Mostre que a norma  $\| \cdot \|$  de um espaço vetorial  $(V, \langle \cdot, \cdot \rangle)$  tem as seguintes propriedades, válidas para quaisquer  $v, w \in V$  :

- i)  $\|v+w\|^2 + \|v-w\|^2 = 2(\|v\|^2 + \|w\|^2)$  (*identidade do paralelogramo*);

- ii)  $|\|v\| - \|w\|| \leq \|v - w\|$  (*segunda desigualdade triangular*).

4. Mostre que, para todo espaço vetorial  $(V, \langle \cdot, \cdot \rangle)$ , vale a seguinte *identidade de polarização*:

$$\langle v, w \rangle = \frac{1}{2} (\|v\|^2 + \|w\|^2 - \|v - w\|^2) \quad \forall v, w \in V,$$

isto é, todo produto interno é determinado pela norma a ele associada.

5. Mostre que a função distância  $\text{dist}$  definida em (5.3) tem as seguintes propriedades, válidas para quaisquer  $u, v, w \in (V, \langle \cdot, \cdot \rangle)$ :

- i)  $\text{dist}(v, w) \geq 0$  e  $\text{dist}(v, w) = 0 \Leftrightarrow v = w$  (*positividade*);
- ii)  $\text{dist}(v, w) = \text{dist}(w, v)$  (*simetria*);
- iii)  $\text{dist}(u, w) \leq \text{dist}(u, v) + \text{dist}(v, w)$  (*desigualdade triangular*).

6. Dado  $c \in \mathbb{R}$ , considere a função  $f(x) = (x + c)^2 / (x^2 + 1)$ ,  $x \in \mathbb{R}$ . Use a desigualdade de Cauchy-Schwarz para provar que

$$f(x) \leq (1 + c)^2 \quad \forall x \in \mathbb{R}.$$

7. Sejam  $u$  e  $v$  vetores LI num espaço vetorial  $(V, \langle \cdot, \cdot \rangle)$ . Determine o vetor  $w$  de  $\text{Ger}\{u, v\}$ , o qual é paralelo à bissetriz do ângulo  $\sphericalangle(u, v)$ , isto é,  $w$  é tal que  $\sphericalangle(u, w) = \sphericalangle(w, v)$ .
8. Considere o exercício anterior e determine o incentro (ponto de encontro das bissetrizes) do triângulo de  $\mathbb{R}^2$  de vértices  $v_1 = (0, 0)$ ,  $v_2 = (1, 1)$  e  $v_3 = (-2, 3)$ .

### Seção 5.3

9. Seja  $\{v_1, \dots, v_k\}$  uma base ortonormal de um subespaço  $W \subset (V, \langle \cdot, \cdot \rangle)$ . Prove a *desigualdade de Bessel*:

$$\sum_{i=1}^k \langle v, v_i \rangle^2 \leq \|v\|^2 \quad \forall v \in V.$$

10. Mostre que o operador linear  $J : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  definido por

$$J(x_1, x_2) = (-x_2, x_1)$$

tem as seguintes propriedades, as quais são válidas para quaisquer vetores  $v, w \in \mathbb{R}^2$ .

i)  $\langle J(v), v \rangle = 0$  ( $J$  é a rotação positiva de ângulo  $\pi/2$ );

ii)  $\langle J(v), J(w) \rangle = \langle v, w \rangle$  ( $J$  preserva produto interno);

iii)  $\|J(v)\| = \|v\|$  ( $J$  preserva norma).

11. Sejam  $W_1$  e  $W_2$  subespaços de  $(V, \langle \cdot, \cdot \rangle)$ . Mostre que:

i)  $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$ ;

ii)  $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$ .

Conclua que:

$$V = W_1 \oplus W_2 \Leftrightarrow V = W_1^\perp \oplus W_2^\perp.$$

12. Seja  $A$  uma matriz  $m \times n$ ,  $v_1, \dots, v_m \in \mathbb{R}^n$  seus vetores linha, e  $W$  o espaço gerado por  $v_1, \dots, v_m$ . Dado  $v \in \mathbb{R}^n$ , mostre que:

$$A[v] = \begin{bmatrix} \langle v, v_1 \rangle \\ \vdots \\ \langle v, v_m \rangle \end{bmatrix}.$$

Conclua que são válidas as seguintes igualdades:

i)  $AA^* = (\langle v_i, v_j \rangle)_{m \times m}$ .

ii)  $\text{Nuc}(T_A) = W^\perp$ .

13. Seja  $\mathcal{B} = \{u_1, \dots, u_n\}$  uma base de  $(V, \langle \cdot, \cdot \rangle)$ . Suponha que, para todo  $v = (x_1, \dots, x_n)_{\mathcal{B}} \in V$ , tenha-se  $\|v\|^2 = x_1^2 + \dots + x_n^2$ . Mostre que  $\mathcal{B}$  é ortonormal.

14. Verifique as propriedades da projeção ortogonal  $P_W : V \rightarrow W$  listadas na Seção 5.3.
15. Seja  $\mathcal{B} = \{u_1, \dots, u_n\}$  uma base ortonormal de  $(V, \langle, \rangle)$ . Dado um operador linear  $T : V \rightarrow V$ , mostre que a matriz de  $T$  com respeito à base  $\mathcal{B}$ ,  $A = (a_{ij})_{n \times n}$ , satisfaz:

$$a_{ij} = \langle Tu_i, u_j \rangle.$$

16. Seja  $W \subset (V, \langle, \rangle)$  um subespaço não trivial e  $v \in V - W$ . Mostre que:

$$\|v - P_W(v)\| < \|v - w\| \quad \forall w \in W - P_W(v),$$

isto é,  $P_W(v)$  é o ponto de  $W$  que está mais próximo de  $v$ .

17. Seja  $\Pi$  o plano de  $\mathbb{R}^3$  de equação  $x - 2y + z = 0$ . Considere o exercício anterior e determine o ponto de  $\Pi$  que está mais próximo de  $v = (1, 2, -1)$ .

#### Seção 5.4

18. Seja  $\{v_1, \dots, v_n\}$  uma base de um espaço vetorial  $(V, \langle, \rangle)$ . Dados  $n$  escalares  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ , prove que existe um único vetor  $v \in V$ , tal que

$$\langle v, v_i \rangle = \lambda_i \quad \forall i = 1, \dots, n.$$

19. Seja  $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$  um operador linear. Mostre que, para todo paralelepípedo  $n$ -dimensional  $\mathcal{P}(v_1, \dots, v_n)$ , tem-se

$$|T(\mathcal{P}(v_1, \dots, v_n))| = |\det T| |\mathcal{P}(v_1, \dots, v_n)|.$$

20. Considere uma base  $\mathcal{B} = \{v_1, \dots, v_n\}$  de  $\mathbb{R}^n$ . Denote por  $W_k$  o subespaço gerado por  $v_1, \dots, v_k$ , e ponha  $w_{k+1} = v_{k+1} - P_{W_k}(v_{k+1})$ ,  $1 \leq k \leq n - 1$ . Mostre que, nessas condições, vale a igualdade:

$$|\mathcal{P}(v_1, \dots, v_n)| = \|v_1\| \|w_2\| \dots \|w_n\|.$$

Conclua, então, a validade da *desigualdade de Hadamard*:

$$|\det(v_1, \dots, v_n)| \leq \|v_1\| \|v_2\| \dots \|v_n\|,$$

em que a igualdade ocorre se, e somente se, a base  $\mathcal{B}$  é ortogonal.



# 6

---

## *Decomposição de Operadores*

A partir deste capítulo, nos dedicaremos exclusivamente ao estudo dos operadores lineares em espaços vetoriais de dimensão finita. Nesse contexto, a abordagem mais eficaz envolve a decomposição do operador dado em “suboperadores” mais simples, num sentido que logo deixaremos claro. Para tanto, conforme constataremos, é suficiente que o espaço vetorial onde se define o operador possa ser decomposto em subespaços, os quais sejam invariantes pela ação do mesmo.

Uma propriedade notável dos operadores lineares em espaços de dimensão finita, a qual estabeleceremos ao longo deste e dos próximos capítulos, é a de que uma tal decomposição sempre pode ser feita, sendo o caso ótimo o dos operadores *diagonalizáveis*, isto é, aqueles em que todos os subespaços da decomposição têm dimensão 1.

Neste capítulo, após algumas considerações sobre o método de decomposição de operadores, introduziremos os conceitos de autovalor e autovetor, bem como estabeleceremos suas propriedades fundamentais. Faremos o mesmo com o conceito de polinômio mínimo e encerraremos apresentando nosso primeiro resultado de decomposição de operadores, conhecido como *Teorema da Decomposição Primária*, o qual será devidamente aplicado na caracterização de operadores diagonalizáveis, assim como em outros resultados fundamentais que apresentaremos posteriormente.

## 6.1 Considerações Iniciais

Conforme discutimos anteriormente, os operadores lineares mais simples são aqueles que podem ser representados por matrizes diagonais, pois toda a informação sobre o operador que está contida na matriz pode ser imediatamente obtida por inspeção direta da mesma.

Consideremos, por exemplo, o operador linear  $T \in \mathcal{L}(\mathbb{R}^3)$ , cuja matriz com respeito à base canônica de  $\mathbb{R}^3$ ,  $\{e_1, e_2, e_3\}$ , é:

$$[T] = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Examinando-a, conclui-se que o conjunto  $\{(2, 0, 0), (0, 3, 0)\}$  constitui uma base do conjunto imagem de  $T$ , e que o núcleo de  $T$  é gerado pelo vetor  $e_3 = (0, 0, 1)$ .

Ainda com respeito ao operador  $T$ , atentemos para seguinte fato:

$$T(e_1) = 2e_1, \quad T(e_2) = 3e_2 \quad \text{e} \quad T(e_3) = 0e_3,$$

de modo que a imagem de cada vetor  $e_i$  é um múltiplo de  $e_i$ . É imediato que a recíproca dessa última asserção é verdadeira, isto é, se um operador linear  $T \in \mathcal{L}(\mathbb{R}^3)$  é tal que a imagem de cada vetor  $e_i$  da base canônica é um múltiplo de si mesmo, então a matriz de  $T$  com respeito a essa base é diagonal.

Mais precisamente, se

$$T(e_1) = \lambda_1 e_1, \quad T(e_2) = \lambda_2 e_2 \quad \text{e} \quad T(e_3) = \lambda_3 e_3, \quad (6.1)$$

tem-se que a matriz  $[T]$  é dada por

$$[T] = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}.$$

Note que, na presença das igualdades (6.1), cada subespaço unidimensional  $W_i = \text{Ger} \{e_i\}$  é invariante pelo operador  $T$ , no sentido de que vale a implicação

$$v \in W_i \Rightarrow T(v) \in W_i.$$

De fato, fazendo-se  $v = x_i e_i$ , tem-se

$$T(v) = T(x_i e_i) = x_i T(e_i) = x_i \lambda_i e_i \in W_i.$$

Desta forma, para cada  $i = 1, 2, 3$ , os operadores

$$T_i = T|_{W_i} : W_i \rightarrow W_i \in \mathcal{L}(W_i) \quad (6.2)$$

estão bem definidos.

Assim, existem 3 subespaços unidimensionais,  $W_1$ ,  $W_2$  e  $W_3$ , todos invariantes por  $T$ , tais que

$$\mathbb{R}^3 = W_1 \oplus W_2 \oplus W_3.$$

Nesse caso, como sabemos, todo vetor  $v \in \mathbb{R}^3$  se escreve de forma única como soma de vetores dos subespaços  $W_i$ , isto é, para cada  $v \in \mathbb{R}^3$ , existem únicos vetores  $w_i \in W_i$ ,  $i = 1, 2, 3$ , tais que:

$$v = w_1 + w_2 + w_3.$$

Daí, e das igualdades 6.2, tem-se:

$$T(v) = T(w_1 + w_2 + w_3) = T(w_1) + T(w_2) + T(w_3) = T_1(w_1) + T_2(w_2) + T_3(w_3).$$

Indicaremos essa relação entre o operador  $T$  e suas restrições  $T_i$  aos respectivos subespaços invariantes por:

$$T = T_1 \oplus T_2 \oplus T_3.$$

Consideremos, agora, o operador  $T \in \mathcal{L}(\mathbb{R}^3)$ , tal que

$$[T] = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Nesse caso, temos

$$T(e_1) = e_2, \quad T(e_2) = -e_1 \quad \text{e} \quad T(e_3) = 2e_3.$$

Claramente,  $\text{Ger}\{e_3\}$  é invariante por  $T$ , e nenhum dos subespaços  $\text{Ger}\{e_1\}$  ou  $\text{Ger}\{e_2\}$  o é. No entanto, o subespaço bidimensional  $\text{Ger}\{e_1, e_2\}$  é invariante por  $T$ , pois dado  $v = x_1e_1 + x_2e_2 \in \text{Ger}\{e_1, e_2\}$ , tem-se

$$T(v) = x_1T(e_1) + x_2T(e_2) = x_1e_2 - x_2e_1 \in \text{Ger}\{e_1, e_2\}.$$

Fazendo-se, agora,

$$W_1 = \text{Ger}\{e_1, e_2\} \text{ e } W_2 = \text{Ger}\{e_3\},$$

tem-se que

$$\mathbb{R}^3 = W_1 \oplus W_2$$

e que os operadores

$$T_1 = T|_{W_1} : W_1 \rightarrow W_1 \text{ e } T_2 = T|_{W_2} : W_2 \rightarrow W_2 \quad (6.3)$$

estão bem definidos.

Além disso, como no exemplo anterior, dado  $v \in \mathbb{R}^3$ , existem únicos  $w_1 \in W_1$  e  $w_2 \in W_2$ , tais que  $v = w_1 + w_2$ , donde  $T(v) = T_1(w_1) + T_2(w_2)$ . Igualmente, expressamos essa relação entre  $T$ ,  $T_1$  e  $T_2$  por

$$T = T_1 \oplus T_2.$$

Observemos que  $[T]$  é diagonal por blocos (vide Seção 4.3.3), e que seu bloco diagonal superior  $2 \times 2$  é a matriz de  $T_1$  com respeito à base  $\{e_1, e_2\}$  de  $W_1$ ,

$$[T_1] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Analogamente, a matriz de  $T_2$  com respeito à base  $\{e_3\}$  de  $W_2$  é  $[2]$ , a qual corresponde o bloco diagonal inferior  $1 \times 1$  de  $[T]$ .

Em suma,

$$T = T_1 \oplus T_2 \quad \text{e} \quad [T] = \begin{bmatrix} [T_1] & \\ & [T_2] \end{bmatrix}.$$

Como último exemplo, consideremos o operador  $T \in \mathcal{L}(\mathbb{R}^6)$ , cuja matriz, também diagonal em blocos, é:

$$[T] = \left[ \begin{array}{ccc|cc} 1 & 2 & 3 & & \\ 0 & 1 & 4 & & \\ 5 & 7 & 6 & & \\ \hline & & & 1 & 2 \\ & & & 2 & 1 \\ \hline & & & & & 1 \end{array} \right] = \begin{bmatrix} [T_1] & & \\ & [T_2] & \\ & & [T_3] \end{bmatrix},$$

em que  $T_i = T|_{W_i}$  e

$$W_1 = \text{Ger} \{e_1, e_2, e_3\}, \quad W_2 = \text{Ger} \{e_4, e_5\} \quad \text{e} \quad W_3 = \text{Ger} \{e_6\}$$

são os subespaços invariantes por  $T$ . Assim, com a notação acima,

$$T = T_1 \oplus T_2 \oplus T_3.$$

De modo geral, dado um operador linear  $T \in \mathcal{L}(V)$ , dizemos que um subespaço  $W \subset V$  é *invariante por  $T$* , ou, equivalentemente,  *$T$ -invariante*, quando  $T(W) \subset W$ , isto é, quando vale a implicação:

$$v \in W \Rightarrow T(v) \in W.$$

Quando  $W \subset V$  é  $T$ -invariante, temos que a restrição de  $T$  a  $W$ , a qual será denotada por  $T_W$ , é um operador linear bem definido em  $\mathcal{L}(W)$ . Desta forma, em total analogia com os exemplos precedentes, uma decomposição de  $V$  numa soma direta de subespaços invariantes,

$$V = W_1 \oplus \cdots \oplus W_k,$$

induz uma decomposição de  $T \in \mathcal{L}(V)$  numa soma direta de operadores,

$$T = T_{W_1} \oplus \cdots \oplus T_{W_k},$$

tendo essa igualdade o seguinte significado:

$$v = w_1 + \cdots + w_k \in V \Rightarrow T(v) = T_{W_1}(w_1) + \cdots + T_{W_k}(w_k), \quad w_i \in W_i.$$

Considerando-se as decomposições acima, tomemos uma base de  $V$ ,

$$\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k,$$

tal que  $\mathcal{B}_i$  seja uma base de  $W_i$ . Com respeito a  $\mathcal{B}$ , a matriz de  $T$  assume, então, a forma diagonal em blocos:

$$[T]_{\mathcal{B}} = \begin{bmatrix} [T_{W_1}]_{\mathcal{B}_1} & & & \\ & [T_{W_2}]_{\mathcal{B}_2} & & \\ & & \ddots & \\ & & & [T_{W_k}]_{\mathcal{B}_k} \end{bmatrix}. \quad (6.4)$$

Em particular (vide igualdade (4.17)),

$$\det T = \det T_{W_1} \dots \det T_{W_k}. \quad (6.5)$$

A este ponto, deve estar clara a ideia por trás da decomposição de um operador linear  $T$ , como vimos fazendo: Estudar o comportamento de  $T$  analisando cada uma das componentes  $T_{W_i}$  da decomposição. Por exemplo, a igualdade (6.5) nos diz que  $T$  é invertível se, e somente se, cada  $T_i$  é invertível. Note que, um operador que compõe  $T$ ,  $T_{W_i}$ , será tão mais simples quanto menor for a dimensão de  $W_i$ , sendo a situação ideal aquela em que cada  $W_i$  é unidimensional, como em nosso primeiro exemplo. Tais operadores são ditos *diagonalizáveis*, conforme mencionamos à introdução.

Ao longo deste capítulo e dos próximos, constataremos que nem todo operador linear é diagonalizável. No entanto, quando  $V$  tem dimensão finita ( $\geq 3$ ), todo operador  $T \in \mathcal{L}(V)$  admite decomposições em soma diretas de operadores, como as que fizemos acima.

Na Seção 6.4, consideraremos um tipo de decomposição de um operador, dita *primária*, a qual, dentre outros benefícios, nos permitirá estabelecer condições necessárias e suficientes para que um operador num espaço de dimensão finita seja diagonalizável.

Por fim, advertimos que, daqui por diante, não consideraremos espaços vetoriais de dimensão infinita, de modo que  $V$  *indicará sempre um espaço vetorial (real) de dimensão finita*  $n \geq 1$ .

## 6.2 Autovalores e Autovetores

Nesta seção, investigaremos os subespaços invariantes de um operador arbitrário  $T \in \mathcal{L}(V)$ , os quais são associados a um determinado tipo de vetor de  $V$  — que pode ou não existir — dito *autovetor* de  $T$ . Os operadores diagonalizáveis, conforme veremos, são justamente aqueles cujos autovetores constituem uma base de  $V$ .

A título de motivação, consideremos o operador linear  $T \in \mathcal{L}(\mathbb{R}^2)$ , tal que

$$[T] = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix},$$

e observemos que os vetores  $v_1 = (1, 1)$  e  $v_2 = (1, -1)$  satisfazem

$$T(1, 1) = (3, 3) = 3(1, 1) \quad \text{e} \quad T(1, -1) = (-1, 1) = -(1, -1),$$

isto é,  $T(v_i)$  é paralelo a  $v_i$ ,  $i = 1, 2$ . Os vetores  $v_i$  são ditos, então, *autovetores* de  $T$ , enquanto os escalares 3 e  $-1$  são ditos os *autovalores* de  $T$  (Fig. 6.1). Observemos ainda que  $v_1$  e  $v_2$  são LI (na verdade, ortogonais) e, portanto, definem uma base  $\mathcal{B}$  de  $\mathbb{R}^2$  formada por autovetores de  $T$ . Em particular, a matriz de  $T$  com respeito a essa base é dada por:

$$[T]_{\mathcal{B}} = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix},$$

a qual é mais simples que  $[T]$ , por ser uma matriz diagonal.

O mesmo ocorre com o operador linear  $T \in \mathcal{L}(\mathbb{R}^3)$  cuja matriz  $[T]$  é:

$$[T] = \begin{bmatrix} 14 & -13 & 8 \\ -13 & 14 & 8 \\ 8 & 8 & -7 \end{bmatrix}.$$

Com efeito, fazendo-se  $v_1 = (1, -1, 0)$ ,  $v_2 = (1, 1, 1)$  e  $v_3 = (1, 1, -2)$ , por meio de um cálculo direto, verifica-se que

$$T(v_1) = 27v_1, \quad T(v_2) = 9v_2 \quad \text{e} \quad T(v_3) = -15v_3,$$

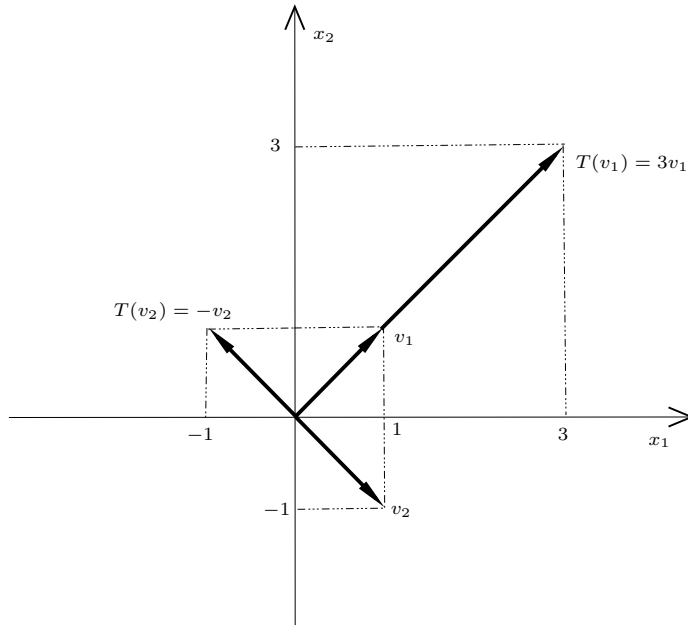


Figura 6.1: Autovetores de  $T(x_1, x_2) = (x_1 + 2x_2, 2x_2 + x_1)$ .

isto é,  $\mathcal{B} = \{v_1, v_2, v_3\}$  é uma base (ortogonal) de  $\mathbb{R}^3$  formada por autovetores de  $T$ . Logo,

$$[T]_{\mathcal{B}} = \begin{bmatrix} 27 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & -15 \end{bmatrix}.$$

Observemos que, nos dois exemplos acima, a matriz  $[T]$  é simétrica e a base de autovetores  $\mathcal{B}$  é ortogonal. No próximo capítulo, veremos que esse comportamento é típico de uma classe especial de operadores, ditos *autoadjuntos*.

De modo geral, dado um operador linear  $T \in \mathcal{L}(V)$ , diz-se que  $v \in V - \{\mathbf{0}\}$  é um *autovetor* de  $T$ , quando existe um escalar  $\lambda \in \mathbb{R}$ , dito o *autovalor* de  $T$  associado ao autovetor  $v$ , o qual cumpre:

$$T(v) = \lambda v.$$

Equivalentemente à definição da seção anterior, podemos dizer que  $T$  é *diagonalizável*, quando existe uma base  $\mathcal{B}$ , de  $V$ , formada por autovetores de



$T$ . Nesse caso, diz-se também que  $\mathcal{B}$  *diagonaliza*  $T$ , pois a matriz de  $T$  com respeito à base  $\mathcal{B}$  é diagonal, sendo suas entradas os respectivos autovalores de  $T$ , isto é, quando a base  $\mathcal{B} = \{v_1, \dots, v_n\} \subset V$  satisfaz  $T(v_i) = \lambda_i v_i$ , tem-se

$$[T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

Suponha que  $v, w \in V$  sejam autovetores de  $T : V \rightarrow V$  associados ao autovalor  $\lambda$ . Nesse caso, dado  $\mu \in \mathbb{R}$ , temos que

$$T(v + \mu w) = T(v) + \mu T(w) = \lambda v + \mu(\lambda w) = \lambda(v + \mu w),$$

isto é,  $v + \mu w$  é também um autovetor de  $T$  associado ao autovalor  $\lambda$ . Daí, segue-se o resultado seguinte.

**Proposição 6.1.** *Dados  $T \in \mathcal{L}(V)$  e  $\lambda \in \mathbb{R}$ , tem-se que o conjunto*

$$W_{\lambda} = \{v \in V; T(v) = \lambda v\} \subset V$$

*é um subespaço vetorial de  $V$ , dito o autoespaço de  $T$  associado a  $\lambda$ .*

Note que, embora o vetor nulo esteja excluído da definição de autovetor, dado  $T \in \mathcal{L}(V)$ , é imediato que  $\mathbf{0} \in W_{\lambda}$ , isto é,

$$W_{\lambda} = \{\text{autovetores de } T \text{ associados a } \lambda\} \cup \{\mathbf{0}\} \subset V.$$

Segue-se diretamente da definição que todo autoespaço  $W_{\lambda}$  como acima é  $T$ -invariante e que vale a seguinte igualdade:

$$T_{W_{\lambda}} = \lambda I_{W_{\lambda}}. \quad (6.6)$$

Dado um operador linear  $T \in \mathcal{L}(V)$ , coloquemo-nos o problema de determinar, caso existam, os autovalores e autovetores de  $T$ . Iniciemos observando que, dados  $v \in V - \{\mathbf{0}\}$  e  $\lambda \in \mathbb{R}$ , são evidentes as equivalências:

$$T(v) = \lambda v \Leftrightarrow (\lambda I - T)v = \mathbf{0} \Leftrightarrow \text{Nuc}(\lambda I - T) \neq \{\mathbf{0}\}. \quad (6.7)$$

Logo,  $\lambda \in \mathbb{R}$  será autovalor de  $T$  se, e somente se,

$$\det(\lambda I - T) = 0. \quad (6.8)$$

Assim, a igualdade (6.8) é uma equação algébrica de variável  $\lambda$ , cujas soluções são autovalores de  $T$ . Uma vez determinados esses autovalores, obtém-se os autovetores correspondentes reconsiderando-se as equivalências (6.7).

A fim de ilustrar esse método, retomemos o operador  $T \in \mathcal{L}(\mathbb{R}^2)$ , tal que

$$[T] = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix},$$

o qual consideramos no começo da seção. Nesse caso, temos

$$[\lambda I - T] = \begin{bmatrix} \lambda - 1 & -2 \\ -2 & \lambda - 1 \end{bmatrix}.$$

Logo,  $\det(\lambda I - T) = (\lambda - 1)^2 - 4$ , donde (6.8) é a equação quadrática

$$\lambda^2 - 2\lambda - 3 = 0,$$

cujas raízes são  $\lambda_1 = 3$  e  $\lambda_2 = -1$ .

Considerando-se a igualdade  $(\lambda I - T)v = \mathbf{0}$  da equivalência (6.7) com  $\lambda = 3$  e  $v = (x_1, x_2)$ , tem-se que o autoespaço de  $T$  associado ao autovalor  $\lambda_1 = 3$  coincide com o conjunto solução do seguinte sistema homogêneo:

$$\begin{cases} 2x_1 - 2x_2 = 0 \\ -2x_1 + 2x_2 = 0 \end{cases},$$

o que, claramente, nos dá:

$$W_{\lambda_1} = \{(x_1, x_2) \in \mathbb{R}^2; x_1 = x_2\} = \text{Ger} \{(1, 1)\}.$$

De modo inteiramente análogo, obtém-se

$$W_{\lambda_2} = \{(x_1, x_2) \in \mathbb{R}^2; x_1 = -x_2\} = \text{Ger} \{(1, -1)\}.$$

A igualdade (6.8) nos leva naturalmente a introduzir o *polinômio característico* de  $T \in \mathcal{L}(V)$ , o qual se define como

$$c(t) = \det(tI - T), \quad t \in \mathbb{R}.$$

Das considerações acima, e do fato de todo polinômio de grau ímpar possuir uma raiz real, segue-se o resultado seguinte.

**Proposição 6.2.** *Um escalar  $\lambda \in \mathbb{R}$  é autovalor de um operador  $T \in \mathcal{L}(V)$  se, e só se,  $\lambda$  é raiz do polinômio característico de  $T$ . Em particular,  $T$  possui pelo menos um autovalor — e, portanto, um autovetor — se a dimensão de  $V$  for ímpar.*

Convém observar que  $c$  é um polinômio mônico de grau  $n$  em  $t$ , isto é,

$$c(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1t + a_0, \quad a_0, \dots, a_{n-1} \in \mathbb{R},$$

o que pode ser constatado pelo cálculo direto do determinante

$$\det[tI - T]_{\mathcal{B}} = \begin{vmatrix} t - a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & t - a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & t - a_{nn} \end{vmatrix},$$

em que  $\mathcal{B}$  é uma base arbitrária de  $V$  e  $[T]_{\mathcal{B}} = (a_{ij})_{n \times n}$ . Nesse caso, os coeficientes  $a_k$  do polinômio característico  $c$  são dados por expressões que envolvem as entradas  $a_{ij}$  de  $[T]_{\mathcal{B}}$ . Acrescentamos ainda que, para  $n = 2$ , verifica-se a seguinte igualdade (vide Exercício 2):

$$c(t) = t^2 - (\text{traço } T)t + \det T.$$

**Proposição 6.3.** *Sejam  $v_1, \dots, v_k \in V$  autovetores de um operador linear  $T \in \mathcal{L}(V)$ , cujos correspondentes autovalores,  $\lambda_1, \dots, \lambda_k$ , são distintos entre si. Então,  $\{v_1, \dots, v_k\}$  é LI.*

*Demonstração.* Provemos por indução sobre  $k$ . O resultado é óbvio para  $k = 1$ . Suponhamo-lo, então, verdadeiro para  $k - 1$  e provemo-lo verdadeiro para  $k$ . Para tanto, consideremos uma combinação linear nula de  $v_1, \dots, v_k$ ,

$$\mu_1 v_1 + \dots + \mu_k v_k = \mathbf{0}. \quad (6.9)$$

Aplicando-se  $T$  a ambos os membros dessa igualdade, obtém-se

$$\mu_1 \lambda_1 v_1 + \dots + \mu_k \lambda_k v_k = \mathbf{0}. \quad (6.10)$$

Multiplicando-se ambos os membros da igualdade (6.9) por  $-\lambda_k$  e adicionando-se a (6.10), chega-se a

$$(\lambda_1 - \lambda_k) \mu_1 v_1 + \dots + (\lambda_{k-1} - \lambda_k) \mu_{k-1} v_{k-1} = \mathbf{0}.$$

Pela hipótese de indução,  $v_1, \dots, v_{k-1}$  são LI. Logo,

$$(\lambda_1 - \lambda_k) \mu_1 = \dots = (\lambda_{k-1} - \lambda_k) \mu_{k-1} = 0,$$

o que nos dá  $\mu_i = 0 \forall i = 1, \dots, k - 1$ , já que, também por hipótese,  $\lambda_i - \lambda_k \neq 0$  para tais valores de  $i$ . Daí e de (6.9), tem-se  $\mu_k v_k = \mathbf{0}$ , donde  $\mu_k = 0$ , isto é,  $v_1, \dots, v_k$  são LI, como desejávamos provar. ■

**Corolário 6.1.** *Sejam  $\lambda_1, \lambda_2$  autovalores distintos de um operador linear  $T \in \mathcal{L}(V)$ . Então  $W_{\lambda_1} \cap W_{\lambda_2} = \{\mathbf{0}\}$ .*

*Demonstração.* Suponha, por absurdo, que exista um vetor não nulo  $v \in W_{\lambda_1} \cap W_{\lambda_2}$ . Nesse caso, associando-se  $v$  a  $\lambda_1$  e  $2v$  a  $\lambda_2$ , temos que  $\{v, 2v\}$  é um conjunto LD de autovetores associados a autovalores distintos, contrariando a Proposição 6.3. Logo,  $W_{\lambda_1} \cap W_{\lambda_2} = \{\mathbf{0}\}$ . ■

Segue-se diretamente das Proposições 6.2 e 6.3 o resultado seguinte, o qual estabelece uma condição suficiente sobre o polinômio característico de um operador para que o mesmo seja diagonalizável.

**Proposição 6.4.** *Seja  $T \in \mathcal{L}(V)$  um operador definido num espaço vetorial  $V$  de dimensão  $n$ . Se o polinômio característico de  $T$  tiver  $n$  raízes reais distintas, então  $T$  será diagonalizável.*

Diante da Proposição 6.4, no que diz respeito à diagonalização de um operador  $T$ , cabe-nos indagar o que ocorre quando o polinômio característico  $c$  tem raízes com multiplicidade. Abordaremos essa questão no exemplo seguinte.

**Exemplo 6.1.** Sejam  $T_1, T_2 \in \mathcal{L}(\mathbb{R}^3)$  os operadores lineares cujas matrizes com respeito à base canônica de  $\mathbb{R}^3$  são:

$$[T_1] = \begin{bmatrix} 2 & 2 & -1 \\ 1 & 0 & 1 \\ 2 & -2 & 3 \end{bmatrix} \quad \text{e} \quad [T_2] = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 4 \\ 3 & -6 & -4 \end{bmatrix}.$$

Calculando-se os determinantes  $\det(tI - T_1)$  e  $\det(tI - T_2)$ , conclui-se que  $T_1$  e  $T_2$  têm o mesmo polinômio característico, qual seja,

$$c(t) = (t - 1)(t - 2)^2.$$

Em particular,  $\lambda_1 = 1$  e  $\lambda_2 = 2$  são ambos autovalores de  $T_1$  e  $T_2$ .

Designemos por  $W_{\lambda_i}^j$  o autoespaço de  $T_j$  associado a  $\lambda_i$ ,  $i, j = 1, 2$ . Procedendo-se como anteriormente, obtém-se para o operador  $T_1$ ,

$$W_{\lambda_1}^1 = \text{Ger} \{(1, -2, -3)\} \quad \text{e} \quad W_{\lambda_2}^1 = \text{Ger} \{(0, 1, 2)\},$$

enquanto para o operador  $T_2$ ,

$$W_{\lambda_1}^2 = \text{Ger} \{(3, -9, 11)\} \quad \text{e} \quad W_{\lambda_2}^2 = \text{Ger} \{(2, 1, 0), (2, 0, 1)\}.$$

Assim, qualquer conjunto LI de autovetores de  $T_1$  tem, no máximo, 2 vetores e, portanto, não constitui uma base de  $\mathbb{R}^3$ . Por outro lado, o conjunto

$$\mathcal{B} = \{(3, -9, 11), (2, 1, 0), (2, 0, 1)\},$$

claramente, é uma base de  $\mathbb{R}^3$  formada por autovetores de  $T_2$ , implicando que esse operador é diagonalizável e que sua matriz com respeito à base  $\mathcal{B}$  é:

$$[T_2]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Para melhor compreendermos a diferença de comportamento dos operadores  $T_1$  e  $T_2$ , observemos que

$$\dim W_{\lambda_1}^1 + \dim W_{\lambda_2}^1 = 2 < 3 = \dim \mathbb{R}^3 = \dim W_{\lambda_1}^2 + \dim W_{\lambda_2}^2,$$

donde se infere que

$$W_{\lambda_1}^2 \oplus W_{\lambda_2}^2 = \mathbb{R}^3 \supsetneq W_{\lambda_1}^1 \oplus W_{\lambda_2}^1.$$

Note que o fato dessas somas serem diretas deve-se à Proposição 6.3.

O exemplo acima sugere a proposição seguinte.

**Proposição 6.5.** *Suponha que  $\lambda_1, \dots, \lambda_k$  sejam os autovalores distintos de um operador linear  $T \in \mathcal{L}(V)$ . Então, as seguintes afirmações são equivalentes:*

- $T$  é diagonalizável;
- $V = W_{\lambda_1} \oplus \dots \oplus W_{\lambda_k}$ .

*Demonstração.* Suponhamos que  $T$  seja diagonalizável. Então, existe uma base  $\mathcal{B}$ , de  $V$ , formada por autovetores de  $T$ . Nesse caso, para cada  $i = 1, \dots, k$ , o conjunto  $\mathcal{B} \cap W_{\lambda_i}$  é não vazio. Caso contrário, pela Proposição 6.3, para todo vetor não nulo  $v \in W_{\lambda_i}$ ,  $\{v\} \cup \mathcal{B}$  seria um conjunto LI contendo uma base de  $V$ . Logo,  $V$  é a soma dos autoespaços  $W_{\lambda_i}$ . Porém, pelo Corolário 6.1,  $W_{\lambda_i} \cap W_{\lambda_j} = \{\mathbf{0}\}$ ,  $i \neq j$ , donde  $V = W_{\lambda_1} \oplus \dots \oplus W_{\lambda_k}$ .

Suponhamos, agora, que  $V = W_{\lambda_1} \oplus \dots \oplus W_{\lambda_k}$  e, para cada  $i = 1, \dots, k$ , tomemos uma base  $\mathcal{B}_i$  de  $W_{\lambda_i}$ . Nesse caso, segue-se diretamente da Proposição 6.3 que  $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$  é uma base de autovetores de  $T$ , donde  $T$  é diagonalizável. ■

Na proposição seguinte, abordaremos a seguinte questão, a qual surge naturalmente no contexto dos operadores diagonalizáveis: Sob que condições existe uma base que diagonalize simultaneamente dois tais operadores?

**Proposição 6.6.** *Sejam  $T, S \in \mathcal{L}(V)$  dois operadores diagonalizáveis que comutam. Então, existe uma base de  $V$  que diagonaliza  $T$  e  $S$  simultaneamente.*

*Demonstração.* Demonstremos o resultado por indução sobre  $n = \dim V$ . Para  $n = 1$ , não há o que se provar. Suponhamos, então, que o resultado seja verdadeiro para todo espaço vetorial de dimensão  $\leq n - 1$ . Nessas condições, consideremos um autoespaço  $W_\lambda$  de  $T$ , o qual, como sabemos, é  $T$ -invariante. Vejamos que  $W_\lambda$  é também  $S$ -invariante. De fato, dado  $v \in W_\lambda$ , tem-se  $T(v) = \lambda v$ . Daí, e da comutatividade de  $T$  e  $S$ , tem-se

$$TS(v) = ST(v) = \lambda S(v),$$

isto é,  $S(v) \in W_\lambda$ . Logo,  $S(W_\lambda) \subset W_\lambda$  e, portanto,  $W_\lambda$  é  $S$ -invariante.

Se  $W_\lambda = V$ , então  $T = \lambda I$ . Nesse caso, qualquer base que diagonalize  $S$  irá diagonalizar  $T$ . Caso contrário, cada autoespaço  $W_{\lambda_i}$  de  $T$  terá dimensão  $\leq n - 1$  e, pelo discutido acima, será invariante por ambos,  $T$  e  $S$ . Aplicando-se, então, a cada autoespaço  $W_{\lambda_i}$  a hipótese de indução, obtém-se uma base  $\mathcal{B}_i$ , de  $W_{\lambda_i}$ , a qual diagonaliza as restrições de  $T$  e  $S$  a  $W_{\lambda_i}$ . Dessa forma, uma vez que  $V = W_{\lambda_1} \oplus \cdots \oplus W_{\lambda_k}$ , temos que  $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$  é uma base de  $V$  que diagonaliza  $T$  e  $S$  simultaneamente. ■

## 6.3 Polinômio Mínimo

Na seção anterior, vimos que as propriedades do polinômio característico de um operador incidem em seu comportamento no que diz respeito a diagonalização. No entanto, tais propriedades estabelecem apenas condições suficientes e, portanto, não caracterizam operadores diagonalizáveis. Dito de outra forma, não há condições necessárias sobre o polinômio característico de um operador diagonalizável. Este lapso é remediado por um outro polinômio associado a  $T$ , dito *mínimo*, sobre o qual podem-se estabelecer tais condições (vide Teorema 6.1 na seção seguinte).

Introduziremos agora o conceito de polinômio mínimo de um operador e verificaremos suas propriedades mais fundamentais. Posteriormente, veremos que o polinômio mínimo tem uma íntima relação com o polinômio característico. Constataremos, também, que os polinômios mínimos desempenham um papel fundamental em decomposição de operadores.

Dado um operador linear  $T \in \mathcal{L}(V)$ , como sabemos, estão bem definidas quaisquer combinações lineares de potências de  $T$ . Assim, dado um polinômio

$$p(t) = a_0 + a_1t + \cdots + a_k t^k \in \mathcal{P}[t, \mathbb{R}],$$

definimos o operador linear  $p(T) \in \mathcal{L}(V)$  por

$$p(T) = a_0I + a_1T + \cdots + a_k T^k.$$

De forma inteiramente análoga, definimos  $p(A) \in \mathcal{M}(n)$  para uma matriz quadrada  $A \in \mathcal{M}(n)$ . Em particular, pelas Proposições 3.6 e 3.7, para qualquer base  $\mathcal{B}$  de  $V$ , vale a seguinte igualdade:

$$[p(T)]_{\mathcal{B}} = p([T]_{\mathcal{B}}).$$

**Definição 6.1** (POLINÔMIO ANULADOR). Diz-se que um polinômio  $p \in \mathcal{P}[t, \mathbb{R}]$  é um *anulador* de  $T$ , ou, equivalentemente, que  $p$  *anula*  $T$ , quando  $p(T)$  é o operador nulo de  $V$ , isto é,  $p(T)(v) = \mathbf{0} \forall v \in V$ .

Verifiquemos que o polinômio  $p(t) = t^2 + 1$  anula o operador  $T \in (\mathbb{R}^2)$ , cuja matriz com respeito à base canônica é:

$$[T] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

De fato, temos que

$$p([T]) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^2 + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

donde  $p(T) = \mathbf{0}$ , isto é,  $p$  anula  $T$ .

No exemplo acima, o leitor atento pode ter notado que  $p$  é precisamente o polinômio característico de  $T$ . Um clássico resultado da Álgebra Linear, conhecido como Teorema de Cayley–Hamilton, afirma que esse fenômeno é típico, isto é, *todo operador linear é anulado por seu polinômio característico*. Estabeleceremos esse teorema no Apêndice A (vide também Exercício 12).



Por ora, verifiquemos que o conjunto dos polinômios que anulam um operador nunca é vazio. Para tanto, tomemos um operador não nulo  $T \in \mathcal{L}(V)$ , e lembremos que  $\dim \mathcal{L}(V) = n^2$ . Logo, o conjunto  $\{I, T, T^2, \dots, T^{n^2}\}$ , por conter  $n^2 + 1$  operadores, é necessariamente LD em  $\mathcal{L}(V)$ . Assim, existem escalares  $a_0, \dots, a_{n^2}$ , não todos nulos, para os quais tem-se

$$a_0I + a_1T + \dots + a_{n^2}T^{n^2} = \mathbf{0}. \quad (6.11)$$

Essa igualdade, porém, nos diz que o polinômio

$$p(t) = a_0 + a_1t + \dots + a_{n^2}t^{n^2}$$

anula o operador linear  $T \in \mathcal{L}(V)$ , mostrando que, de fato, o conjunto dos anuladores de  $T$  é não vazio. Desta forma, faz sentido introduzir o seguinte conceito:

**Definição 6.2** (POLINÔMIO MÍNIMO). O *polinômio mínimo* de um operador linear  $T \in \mathcal{L}(V)$  é o polinômio mônico de menor grau que anula  $T$ , o qual será denotado por  $m$ .

Verifiquemos que o polinômio mínimo está bem definido, isto é, que ele é único. Antes, relembremos alguns fatos acerca de polinômios.

Primeiramente, um polinômio

$$p(t) = a_0 + a_1t + \dots + a_nt^n \in \mathcal{P}[t, \mathbb{R}]$$

é dito de *grau*  $n$ , quando  $a_n \neq 0$ , e *mônico*, quando  $a_n = 1$ . Dados polinômios  $p_1, p_2 \in \mathcal{P}[t, \mathbb{R}]$ , diz-se que  $p_1$  é um *divisor* de  $p_2$ , quando existe  $q \in \mathcal{P}[t, \mathbb{R}]$ , tal que  $p_2 = qp_1$ . Um polinômio  $p$  cujos únicos divisores são o próprio  $p$  e polinômio constante 1 é dito *primo* ou *irredutível*.

Nesse contexto, apelaremos aos seguintes teoremas fundamentais:

**Algoritmo da Divisão.** Dados polinômios  $p_1, p_2 \in \mathcal{P}[t, \mathbb{R}]$ , existem únicos polinômios  $q, r \in \mathcal{P}[t, \mathbb{R}]$ , os quais cumprem a igualdade

$$p_2 = qp_1 + r,$$

em que  $r = 0$  ou  $\text{grau}(r) < \text{grau}(p_1)$ .

**Teorema da Fatoração Única.** Todo polinômio  $\mathcal{P}[t, \mathbb{R}]$  se expressa de forma única como um produto de polinômios primos.

Considerando o Teorema da Fatoração Única, dizemos que dois polinômios são *primos entre si*, quando não possuem fatores primos comuns.

Por fim, enfatizamos que, pelo Teorema Fundamental da Álgebra, um polinômio primo  $p \in \mathcal{P}[t, \mathbb{R}]$  tem grau 1 ou 2, sendo de uma das seguintes formas:

$$p(t) = at + b, \quad a \neq 0, \quad b \in \mathbb{R}, \quad \text{ou} \quad p(t) = at^2 + bt + c, \quad a \neq 0, \quad b^2 - 4ac < 0.$$

A unicidade do polinômio mínimo de um operador  $T \in \mathcal{L}(V)$  segue-se do algoritmo da divisão. Com efeito, suponha que  $m_1, m_2 \in \mathcal{P}[t, \mathbb{R}]$  sejam polinômios mínimos de  $T$ . Pelo algoritmo da divisão, existem  $q, r \in \mathcal{P}[t, \mathbb{R}]$ , com  $r = 0$  ou grau( $r$ ) < grau( $m_1$ ), tais que  $m_2 = m_1q + r$ . Se tivéssemos  $r \neq 0$ , então, para alguma constante  $\mu > 0$ ,  $\mu r$  seria mônico e anulava  $T$ , já que  $\mu r = \mu(p_2 - p_1q)$ . Isso, porém, contradiria a minimalidade de  $m_1$ . Assim, devemos ter  $r = 0$ . Daí, e do fato de  $m_1$  e  $m_2$  serem ambos mônicos e de mesmo grau, segue-se que  $q = 1$  e, portanto, que  $m_1 = m_2$ .

Com o mesmo tipo de argumento, obtém-se o seguinte resultado:

**Proposição 6.7.** *Sejam  $T \in \mathcal{L}(V)$  e  $p \in \mathcal{P}[t, \mathbb{R}]$  um anulador de  $T$ . Então, o polinômio mínimo de  $T$  divide  $p$ .*

*Demonstração.* Sejam  $q, r \in \mathcal{P}[t, \mathbb{R}]$ , com  $r = 0$  ou grau( $r$ ) < grau( $m$ ), tais que  $p = qm + r$ . Procedendo-se como acima, conclui-se que a minimalidade de  $m$  seria violada se tivéssemos  $r \neq 0$ . Logo,  $p = qm$  e, portanto, o polinômio mínimo  $m$  divide o anulador  $p$ . ■

**Exemplo 6.2.** Reconsideremos o operador  $T \in \mathcal{L}(\mathbb{R}^2)$ , tal que

$$[T] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Conforme constatamos,  $p(t) = t^2 + 1$  é um anulador de  $T$ . Logo, o polinômio mínimo de  $T$  é  $p$  ou um polinômio de grau 1,  $m(t) = t - \lambda$ . Veja que, no último caso, da igualdade  $m(T) = \mathbf{0}$ , teríamos  $T(v) = \lambda v$  para todo  $v \in V$ , isto é,  $T = \lambda I$ . Por outro lado, é evidente que  $[T]$  não é um múltiplo da matriz identidade de  $\mathcal{M}(2)$ . Logo, o polinômio mínimo de  $T$  é  $p$ .

**Exemplo 6.3.** Seja  $T \in \mathcal{L}(\mathbb{R}^3)$ , tal que

$$[T] = \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix},$$

em que  $a, b \neq 0$ . O polinômio característico de  $T$ , como pode ser facilmente verificado, é  $c(t) = (t - 1)^3$ . Já adiantamos que o polinômio característico sempre anula o operador correspondente, de modo que  $c$  anula  $T$ . Como no exemplo anterior, pode-se ver que  $[T]$  não é um múltiplo da matriz identidade de  $\mathcal{M}(3)$ , pois  $a$  e  $b$  são não nulos. Logo, o polinômio mínimo de  $T$  tem grau maior que 1. Além disso, pode-se verificar que  $([T] - [I])^2 = \mathbf{0}$ , donde se infere que o polinômio mínimo de  $T$  é  $m(t) = (t - 1)^2$ .

Fixemos, agora,  $v \in V$  e  $T \in \mathcal{L}(V)$ . Em analogia com o discutido nos parágrafos precedentes, observemos que o conjunto  $\{v, T(v), \dots, T^n(v)\}$  é LD em  $V$ , uma vez que o mesmo possui  $n + 1$  vetores e  $n = \dim V$ . Assim, existem  $a_0, \dots, a_n \in \mathbb{R}$ , não todos nulos, tais que

$$a_0v + a_1T(v) + \dots + a_nT^n(v) = \mathbf{0}.$$

Nesse caso, dizemos que o polinômio

$$p(t) = a_0v + a_1t + \dots + a_nt^n$$

anula  $T$  em  $v$ . Isso nos leva ao seguinte conceito:

**Definição 6.3.** Dados  $v \in V$  e  $T \in \mathcal{L}(V)$ , o *polinômio mínimo* de  $T$  em  $v$ , o qual denotaremos por  $m_v$ , é o polinômio mônico de menor grau que anula  $T$  em  $v$ .

De modo análogo ao feito acima, isto é, considerando-se o Algoritmo da Divisão, verifica-se que o polinômio mínimo de um operador num vetor é único. O mesmo se aplica à proposição seguinte:

**Proposição 6.8.** *Sejam  $v \in V$ ,  $T \in \mathcal{L}(V)$  e  $p \in \mathcal{P}[t, \mathbb{R}]$  um anulador de  $T$  em  $v$ . Então, o polinômio mínimo de  $T$  em  $v$  divide  $p$ .*

Segue-se da Proposição 6.8 que, para todo operador  $T \in \mathcal{L}(V)$ , e todo  $v \in V$ ,  $m_v$  divide o polinômio mínimo  $m$ . Com efeito, sendo  $m$  um anulador de  $T$ , ele será, em particular, um anulador de  $T$  em  $v$ .

**Exemplo 6.4.** Vimos que  $e_1$  e  $e_2$  são autovetores do operador  $T \in \mathcal{L}(\mathbb{R}^3)$  do Exemplo 6.3. Logo,  $m_{e_1} = m_{e_2} = t - 1$ . Temos, ainda, que  $T(e_3) = (a, b, 1)$ , donde se infere que  $e_3$  não é autovetor de  $T$ , pois  $a, b \neq 0$ . Logo, o polinômio mínimo de  $T$  em  $e_3$  é igual ao polinômio mínimo de  $T$ , isto é,  $m_{e_3}(t) = m(t) = (t - 1)^2$ .

Concluindo esta seção, convém mencionarmos que, através do isomorfismo natural entre  $\mathcal{L}(V)$  e  $\mathcal{M}(n)$ , os conceitos e resultados obtidos aqui, concernentes a operadores, podem ser atribuídos a matrizes. Assim, os autovalores e autovetores de uma matriz  $A$  são aqueles do operador  $T$  a ela associado. Analogamente,  $A$  será dita diagonalizável se  $T$  for diagonalizável, e assim por diante.

## 6.4 O Teorema da Decomposição Primária

Estabeleceremos, agora, o anunciado Teorema da Decomposição Primária, segundo o qual, para todo operador linear  $T \in \mathcal{L}(V)$ , existe uma decomposição de  $V$  em subespaços  $T$ -invariantes, os quais são determinados pelos fatores primos do polinômio mínimo de  $T$ . Como primeira aplicação desse teorema, caracterizaremos os operadores diagonalizáveis através de seus polinômios mínimos.

Na demonstração do Teorema da Decomposição Primária, nos será útil o seguinte resultado.

**Lema 6.1.** *Sejam  $T \in \mathcal{L}(V)$  um operador linear e  $p, q \in \mathcal{P}[t, \mathbb{R}]$  polinômios mônicos. Então, são válidas as seguintes afirmações:*

- i)  $p(T)$  e  $q(T)$  comutam;
- ii) Para todo subespaço  $T$ -invariante  $W \subset V$ , vale a igualdade  $p(T_W) = p(T)|_W$ .

*Demonstração.* (i) Suponhamos que  $p, q \in \mathcal{P}[t, \mathbb{R}]$  se escrevam como:

- $p(t) = a_0 + a_1t + \cdots + a_{n-1}t^{n-1} + a_nt^n;$
- $q(t) = b_0 + b_1t + \cdots + b_{m-1}t^{m-1} + b_mt^m.$

Assim, uma vez que duas potências quaisquer de  $T$  comutam, tem-se

$$p(T)q(T) = \sum_{i=0}^n a_i T^i \sum_{j=0}^m b_j T^j = \sum_{j=0}^m b_j T^j \sum_{i=0}^n a_i T^i = q(T)p(T),$$

em que estamos considerando  $T^0 = I$ .

(ii) Uma vez que  $W$  é  $T$ -invariante, verifica-se facilmente por indução que, para todo  $i \in \mathbb{N}$ ,  $W$  é  $T^i$ -invariante, isto é,  $T^i(w) \in W \forall w \in W$ . Em particular,  $T^i(w) = T_W^i(w) \forall w \in W$ . Assim,

$$p(T_W)w = \sum_{i=0}^n a_i T_W^i(w) = \sum_{i=0}^n a_i T^i(w) = p(T)w.$$

donde se infere que  $p(T_W) = p(T)|_W$  ■

**Teorema da Decomposição Primária.** Sejam  $V$  um espaço vetorial de dimensão finita e  $T \in \mathcal{L}(V)$ . Suponha que o polinômio mínimo de  $T$  seja dado por

$$m(t) = (p_1(t))^{s_1} \cdots (p_k(t))^{s_k},$$

em que  $p_1, \dots, p_k$  são polinômios mônicos irredutíveis e distintos entre si. Nessas condições,  $V$  admite a seguinte decomposição:

$$V = W_1 \oplus \cdots \oplus W_k,$$

em que  $W_i$  denota o núcleo de  $p_i(T)^{s_i}$ . Além disso, para todo  $i = 1, \dots, k$ , são válidas as seguintes afirmações:

- i)  $W_i$  é  $T$ -invariante;
- ii) O polinômio mínimo de  $T_{W_i}$  é  $p_i^{s_i}$ .

*Demonstração.* Observemos, inicialmente, que  $W_i \neq \{\mathbf{0}\} \forall i = 1, \dots, k$ . Supondo, por absurdo, que  $W_i = \{\mathbf{0}\}$ , temos que  $p_i(T)^{s_i} : V \rightarrow V$  é um isomorfismo. Pelo Lema 6.1-(i), os operadores  $p_j(T)^{s_j}$  comutam. Então, podemos supor, sem perda de generalidade, que  $i = k$ . Agora, existe  $w \in V$ , tal que  $p_1(T)^{s_1} \dots p_{k-1}(T)^{s_{k-1}}(w) \neq \mathbf{0}$ . Caso contrário,  $p_1^{s_1} \dots p_{k-1}^{s_{k-1}}$  seria o polinômio mínimo de  $T$ . Porém, sendo  $p_k(T)^{s_k}$  um isomorfismo, existe  $v \in V$ , tal que  $p_k(T)^{s_k}(v) = w$ . Logo,

$$m(T)(v) = (p_1(T))^{s_1} \dots (p_k(T))^{s_k}(v) \neq \mathbf{0},$$

o que, obviamente, é uma contradição. Assim,  $W_i \neq \{\mathbf{0}\} \forall i = 1, \dots, k$ .

Suponhamos, agora, que exista um vetor não nulo  $w \in W_i \cap W_j$ ,  $i \neq j$ . Nesse caso,  $p_i^{s_i}$  e  $p_j^{s_j}$  anulam  $T$  em  $w$  e, portanto,  $m_w$  divide ambos, o que é absurdo, visto que  $p_i^{s_i}$  e  $p_j^{s_j}$  são primos entre si. Logo,  $W_i \cap W_j = \{\mathbf{0}\} \forall i \neq j \in \{1, \dots, k\}$ . Além disso, temos

$$V = \text{Nuc } m(T) = \text{Nuc } ((p_1(T))^{s_1} \dots (p_k(T))^{s_k}).$$

Desta forma (vide Exercício 18 – Cap. 3),

$$\dim V = \dim (\text{Nuc } m(T)) \leq \dim W_1 + \dots + \dim W_k = \dim (W_1 \oplus \dots \oplus W_k),$$

donde se infere que  $V = W_1 \oplus \dots \oplus W_k$ .

Segue-se da comutatividade de  $T$  com  $p_i(T)^{s_i}$  que cada subespaço  $W_i$  é  $T$ -invariante. De fato, dado  $v \in W_i$ , tem-se

$$p_i(T)^{s_i}(T(v)) = T(p_i(T)^{s_i}(v)) = T(\mathbf{0}) = \mathbf{0},$$

donde  $T(v) \in W_i$ . Em particular,  $T = T_{W_1} \oplus \dots \oplus T_{W_k}$ .

Provemos que o polinômio mínimo de  $T_{W_i}$ , o qual denotaremos por  $m_i$ , é  $p_i^{s_i}$ . Podemos supor, sem perda de generalidade, que  $i = 1$ . Primeiramente, observemos que, do Lema 6.1-(ii), e do fato de  $W_1$  ser o núcleo de  $p_1^{s_1}(T)$ , segue-se que  $p_1^{s_1}(T_{W_1}) = p_1^{s_1}(T)|_{W_1} = \mathbf{0}$ , isto é,  $p_1^{s_1}$  é um anulador de  $T_{W_1}$ . Logo, pela Proposição 6.7,  $m_1$  divide  $p_1^{s_1}$ .

Definamos, então, o polinômio  $p = m_1 p_2^{s_2} \dots p_k^{s_k}$  e vejamos que o mesmo anula  $T$ . De fato, tomando-se  $v = w_1 + \dots + w_k \in V$ ,  $w_i \in W_i$ , e lembrando-se que  $p_j^{s_j}(T)(w_j) = \mathbf{0}$ , tem-se

$$\begin{aligned} p(T)(v) &= m_1(T)p_2^{s_2}(T) \dots p_k^{s_k}(T)(w_1 + \dots + w_k) \\ &= m_1(T)p_2^{s_2}(T) \dots p_k^{s_k}(T)(w_1 + \dots + w_{k-1}) + \\ &+ m_1(T)p_2^{s_2}(T) \dots p_k^{s_k}(T)(w_k) \\ &= m_1(T)p_2^{s_2}(T) \dots p_k^{s_k}(T)(w_1 + \dots + w_{k-1}). \end{aligned}$$

Repetindo-se esse procedimento  $k - 1$  vezes, chega-se facilmente a:

$$\begin{aligned} p(T)(v) &= m_1(T)p_2^{s_2}(T) \dots p_k^{s_k}(T)(w_1) = p_2^{s_2}(T) \dots p_k^{s_k}(T)m_1(T)(w_1) \\ &= p_2^{s_2}(T) \dots p_k^{s_k}(T)m_1(T_{W_1})(w_1) = \mathbf{0}. \end{aligned}$$

Desta forma,  $p_i^{s_i}$  é o polinômio mínimo de  $T_{W_i}$ , pois, se não o fosse, o polinômio  $p$  seria um anulador de  $T$  com grau menor que o do polinômio mínimo  $m$ . Isso conclui a nossa demonstração. ■

Com a notação do teorema acima, observemos que, se  $p_i(t) = t - \lambda_i$ , então  $\lambda_i$  é um autovalor de  $T$ . Com efeito, sendo  $p_i^{s_i}$  o polinômio mínimo de  $T_{W_i}$ , temos, em particular, que

$$0 = \det(T_{W_i} - \lambda_i I_i)^{s_i} = (\det(T_{W_i} - \lambda_i I_i))^{s_i},$$

em que  $I_i$  denota a aplicação identidade de  $W_i$ . Desta forma,  $\det(T_{W_i} - \lambda_i I_i) = 0$  e, portanto, existe  $w_i \in W_i - \{\mathbf{0}\}$ , tal que

$$\mathbf{0} = (T_{W_i} - \lambda_i I_i)w_i = (T - \lambda_i I)w_i,$$

donde  $w_i$  é um autovetor de  $T$  com autovalor  $\lambda_i$ .

O mesmo argumento aplicado ao caso em que  $p_i(t) = t^2 + bt + c$ , nos garante a existência de  $w_i \in W_i - \{\mathbf{0}\}$ , tal que

$$(T^2 + bT + cI)(w_i) = \mathbf{0},$$

isto é,  $T^2(w_i) = -bT(w_i) - cw_i$ . Essa última igualdade nos diz que  $\text{Ger}\{w_i, Tw_i\}$  é invariante por  $T$ . Além disso, esse subespaço tem dimensão 2, isto é,  $w_i$  e

$T(w_i)$  são vetores LI. Caso contrário, existiria  $\lambda_i \in \mathbb{R}$ , tal que  $T(w_i) = \lambda_i w_i$ , o que nos daria, então,

$$\mathbf{0} = (T^2 + bT + cI)(w_i) = (\lambda_i^2 + b\lambda_i + c)w_i,$$

donde  $\lambda_i$  seria uma raiz de  $p_i$ , contradizendo o fato desse polinômio ser primo.

Segue-se, dessas considerações, o resultado seguinte:

**Proposição 6.9.** *As seguintes afirmações a respeito de um operador linear  $T \in \mathcal{L}(V)$  com polinômio mínimo  $m$  são verdadeiras:*

- *As raízes de  $m$  são autovalores de  $T$ ;*
- *$V$  possui pelo menos um subespaço  $T$ -invariante de dimensão 1 ou 2.*

Apliquemos agora o Teorema da Decomposição Primária nesta elegante caracterização dos operadores diagonalizáveis:

**Teorema 6.1** (CARACTERIZAÇÃO DE OPERADORES DIAGONALIZÁVEIS). *Um operador linear  $T \in \mathcal{L}(V)$  é diagonalizável se, e somente se, existem  $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ , dois a dois distintos, tais que:*

$$m(t) = (t - \lambda_1) \dots (t - \lambda_k). \quad (6.12)$$

No caso afirmativo,  $\lambda_1, \dots, \lambda_k$  são os autovalores de  $T$ .

*Demonstração.* Suponha que  $T$  seja diagonalizável e que seus autovalores (distintos) sejam  $\lambda_1, \dots, \lambda_k$ . Nesse caso, pela Proposição 6.5,  $V$  se decompõe como  $V = W_{\lambda_1} \oplus \dots \oplus W_{\lambda_k}$ . Consequentemente,

$$T = \lambda_1 I_1 \oplus \dots \oplus \lambda_k I_k,$$

em que  $I_i$  denota a restrição da aplicação identidade de  $V$  a  $W_{\lambda_i}$ . Assim, dado um polinômio  $p \in \mathcal{P}[t, \mathbb{R}]$ , temos que (vide Exercícios 13 e 14):

$$p(T) = p(\lambda_1)I_1 \oplus \dots \oplus p(\lambda_k)I_k,$$



onde  $p(T) = \mathbf{0}$  se, e somente se,  $p(\lambda_i) = 0 \forall i = 1, \dots, k$ . Certamente, o polinômio  $p$  de menor grau que satisfaz essa condição é:

$$p(t) = (t - \lambda_1) \dots (t - \lambda_k),$$

o que prova (6.12).

Suponhamos, agora, que a igualdade (6.12) seja válida. Uma vez que o núcleo de  $T - \lambda_i I$  coincide com o auto espaço  $W_{\lambda_i}$ , segue-se do Teorema da Decomposição Primária que

$$V = W_{\lambda_1} \oplus \dots \oplus W_{\lambda_k}.$$

Logo, pela Proposição 6.5, o operador  $T$  é diagonalizável e  $\lambda_1, \dots, \lambda_k$  são os seus autovalores. ■

Seja  $T \in \mathcal{L}(\mathbb{R}^3)$ , tal que

$$[T] = \begin{bmatrix} 2 & 2 & -1 \\ 1 & 0 & 1 \\ 2 & -2 & 3 \end{bmatrix},$$

o qual consideramos no Exemplo 6.1. Lá, constatamos que  $T$  não é diagonalizável, e que seu polinômio característico é  $c(t) = (t - 2)^2(t - 1)$ . Logo, pelo Teorema 6.1, o polinômio mínimo de  $T$  é  $m(t) = c(t)$ .

Efetuem a decomposição primária de  $T$ . Para tanto, devemos determinar o núcleo dos operadores  $(T - 2I)^2$  e  $T - I$ . Tomando-se a matriz de  $T - 2I$  e resolvendo-se, para  $v = (x, y, z) \in \mathbb{R}^3$ , a equação  $[T - 2I]^2[v] = \mathbf{0}$ , obtém-se

$$W_1 = \text{Nuc}(T - 2I)^2 = \text{Ger}\{(1, 0, 0), (0, 1, 2)\}.$$

Já o núcleo de  $T - I$  é o autoespaço associado ao autovalor 1, o qual verificamos ser o espaço gerado por  $(1, -2, -3)$ . Logo,

$$W_2 = \text{Nuc}(T - I) = \text{Ger}\{(1, -2, -3)\}.$$

Segue-se do Teorema da Decomposição Primária que  $T = T_{W_1} \oplus T_{W_2}$ . Observando-se que:

- $T_{W_1}(1, 0, 0) = T(1, 0, 0) = (2, 1, 2) = 2(1, 0, 0) + (0, 1, 2)$ ;
- $T_{W_1}(0, 1, 2) = T(0, 1, 2) = (0, 2, 4) = 0(1, 0, 0) + 2(0, 1, 2)$ ;

tem-se que a matriz de  $T_{W_1}$  com respeito à base  $\mathcal{B}_1 = \{(1, 0, 0), (0, 1, 2)\}$ , do subespaço invariante  $W_1$ , é dada por:

$$[T_{W_1}]_{\mathcal{B}_1} = \begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}.$$

Logo, a matriz de  $T$  com respeito à base  $\mathcal{B} = \{(1, 0, 0), (0, 1, 2), (1, -2, -3)\}$ , de  $\mathbb{R}^3$ , é:

$$[T]_{\mathcal{B}} = \begin{bmatrix} [T_{W_1}]_{\mathcal{B}_1} & & \\ & [T_{W_2}]_{\mathcal{B}_2} & \\ & & \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

em que  $\mathcal{B}_2 = \{(1, -2, -3)\} \subset W_2$ .

## Exercícios

### Seções 6.1 e 6.2

1. Em cada item abaixo, é dada a matriz do operador  $T \in \mathcal{L}(\mathbb{R}^n)$  com respeito à base canônica de  $\mathbb{R}^n$ , em que  $n$  é a ordem da matriz dada. Determine, em cada caso, os autovalores e autoespaços de  $T$ .

i)  $\begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$

ii)  $\begin{bmatrix} 4 & 0 & 1 \\ -2 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$

iii)  $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$

2. Sejam  $V$  um espaço vetorial bidimensional e  $T \in \mathcal{L}(V)$ . Mostre que o polinômio característico de  $T$  é

$$c(t) = t^2 - (\text{traço } T)t + \det T.$$

3. Mostre que o operador  $T \in \mathcal{L}(\mathbb{R}^2)$ , tal que

$$[T] = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

é diagonalizável.

4. Suponha que o operador linear  $T : V \rightarrow V$  seja invertível e que tenha um autovalor  $\lambda$ . Mostre que  $\lambda \neq 0$  e que  $1/\lambda$  é um autovalor de  $T^{-1}$ .
5. Seja  $T \in \mathcal{L}(\mathbb{R}^2)$ , tal que

$$[T] = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}.$$

Mostre que o subespaço  $W_1 = \text{Ger}\{e_1\}$  é  $T$ -invariante. Existe um subespaço  $W_2 \subset \mathbb{R}^2$ , o qual seja  $T$ -invariante e satisfaça  $\mathbb{R}^2 = W_1 \oplus W_2$ ?

6. Seja  $T \in \mathcal{L}(V)$  um operador linear, tal que todos os vetores não nulos de  $V$  são autovetores de  $T$ . Mostre que  $T = \lambda I$ , para algum  $\lambda \in \mathbb{R}$ .
7. Dados  $T, S \in \mathcal{L}(V)$ , prove que  $TS$  e  $ST$  têm os mesmos autovalores.
8. Seja  $T \in \mathcal{L}(V)$  um operador linear de posto 1. Mostre que uma e somente uma das possibilidades ocorre:
- $T$  é diagonalizável;
  - Existe  $k \in \mathbb{N}$ , tal que  $T^k = \mathbf{0}$ .
9. Mostre que uma matriz  $A \in \mathcal{M}(n)$  é diagonalizável se, e somente se, é semelhante a uma matriz diagonal de  $\mathcal{M}(n)$ . Em seguida, verifique se

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

é ou não diagonalizável.

### Seções 6.3 e 6.4

10. Dados  $T \in \mathcal{L}(V)$  e  $p \in \mathcal{P}[t, \mathbb{R}]$ , prove que, se  $\lambda$  é autovalor de  $T$ , então  $p(\lambda)$  é autovalor de  $p(T)$ .
11. Sejam  $T, S \in \mathcal{L}(V)$  e  $c \in \mathcal{P}[t, \mathbb{R}]$  o polinômio característico de  $S$ . Considere o exercício anterior e mostre que, se  $c(T)$  é invertível, então nenhum autovalor de  $T$  é autovalor de  $S$ .
12. (*Teorema de Cayley-Hamilton em dimensão 2*). Sejam  $V$  e  $T$  como no Exercício 2. Mostre que o polinômio característico  $c$  anula  $T$ .
13. Dado  $p \in \mathcal{P}[t, \mathbb{R}]$ , mostre que  $p(\lambda I) = p(\lambda)I$ , em que  $I$  é o operador identidade de  $V$ .
14. Seja  $T \in \mathcal{L}(V)$  um operador linear que admite uma decomposição:

$$V = W_1 \oplus \cdots \oplus W_k, \quad T = T_{W_1} \oplus \cdots \oplus T_{W_k}.$$

Dado  $p \in \mathcal{P}[t, \mathbb{R}]$ , considere a representação matricial (6.4), bem como as regras operacionais de matrizes diagonais por blocos, e mostre que

$$p(T) = p(T_{W_1}) \oplus \cdots \oplus p(T_{W_k}).$$

15. Dados  $T \in \mathcal{L}(V)$  e  $p \in \mathcal{P}[t, \mathbb{R}]$ , mostre que o núcleo de  $p(T)$  é um subespaço  $T$ -invariante de  $V$ .
16. Determine os polinômios mínimo e característico do operador identidade e do operador nulo de um espaço vetorial  $V$ .
17. Seja  $T \in \mathcal{L}(V)$ , tal que  $T^k = \mathbf{0}$  para algum  $k \in \mathbb{N}$ . Mostre que  $T^n = \mathbf{0}$ ,  $n = \dim V$ .
18. Mostre que  $T \in \mathcal{L}(V)$  é invertível se, e somente se, o termo constante de seu polinômio mínimo é não nulo.

19. Decomponha, segundo o Teorema da Decomposição Primária, o operador do Exemplo 6.3.
20. Sejam  $T \in \mathcal{L}(V)$  um operador diagonalizável e  $W \subset V$  um subespaço  $T$ -invariante. Considere o Teorema 6.1 e mostre que  $T_W \in \mathcal{L}(W)$  é diagonalizável.



# 7

---

## *Operadores em Espaços com Produto Interno*

Dando prosseguimento ao estudo dos operadores lineares em espaços de dimensão finita, iniciado no capítulo anterior, neste, os consideraremos no contexto dos espaços munidos de produto interno.

O fato de que produtos internos definem geometrias em espaços vetoriais, nos leva a indagar sobre existência e possibilidades de caracterização de operadores lineares que preservem as propriedades geométricas dos espaços onde estão definidos, isto é, que sejam isometrias. Tais operadores, os quais constataremos existirem em abundância, são chamados de *ortogonais*.

Neste contexto, surgem também os operadores ditos *autoadjuntos*, os quais têm a propriedade que mais se deseja de um operador linear — traduzida no célebre Teorema Espectral — que é a de ser diagonalizável por uma base ortonormal.

Faremos, neste capítulo, uma breve consideração dos operadores ortogonais, bem como dos operadores autoadjuntos, estabelecendo as suas propriedades mais fundamentais. Em seguida, através dos operadores autoadjuntos, introduziremos as formas quadráticas, as quais são funções definidas em espaços vetoriais, e que se relacionam com objetos matemáticos de diversas naturezas, incluindo curvas, superfícies e equações diferenciais.

## 7.1 A Adjunta de uma Transformação Linear

Os operadores que consideraremos neste capítulo definem-se a partir de suas relações com um outro operador, a cada um deles associado, dito o *adjunto*, o qual introduziremos nesta seção. Na verdade, essa associação pode ser feita com qualquer transformação linear, conforme o teorema seguinte.

Advertimos que, neste capítulo, como no anterior, todos os espaços vetoriais mencionados são, implicitamente, de dimensão finita. Lembramos ainda que  $(V, \langle \cdot, \cdot \rangle)$  denota o espaço vetorial  $V$  munido de um produto interno  $\langle \cdot, \cdot \rangle$ .

**Teorema 7.1** (EXISTÊNCIA E UNICIDADE DA ADJUNTA). *Sejam  $(V, \langle \cdot, \cdot \rangle)$  e  $(W, \langle \cdot, \cdot \rangle)$  espaços vetoriais de dimensão finita munidos de produto interno. Dada uma transformação linear  $T \in \mathcal{L}(V, W)$ , existe uma única transformação linear  $T^* \in \mathcal{L}(W, V)$ , a qual cumpre a seguinte igualdade:*

$$\langle T(v), w \rangle = \langle v, T^*(w) \rangle \quad \forall v \in V, w \in W. \quad (7.1)$$

*Demonstração.* Sejam  $\mathcal{B} = \{v_1, \dots, v_n\} \subset V$  e  $\mathcal{B}' = \{w_1, \dots, w_m\} \subset W$  bases ortonormais de  $V$  e  $W$ , respectivamente. Defina  $T^* : W \rightarrow V$  como a única transformação linear de  $\mathcal{L}(W, V)$  para a qual se tem (vide Proposição 3.1)

$$T^*(w_i) = \sum_{k=1}^n \langle T(v_k), w_i \rangle v_k, \quad i \in \{1, \dots, m\}.$$

Nesse caso, para cada  $i \in \{1, \dots, m\}$  e  $j \in \{1, \dots, n\}$ , tem-se

$$\langle v_j, T^*(w_i) \rangle = \left\langle v_j, \sum_{k=1}^n \langle T(v_k), w_i \rangle v_k \right\rangle = \langle T(v_j), w_i \rangle,$$

de modo que (7.1) é satisfeita para os vetores das bases  $\mathcal{B}$  e  $\mathcal{B}'$ . Daí, da linearidade de  $T$  e  $T^*$ , e da bilinearidade do produto interno, segue-se que (7.1) é satisfeita para quaisquer  $v \in V$  e  $w \in W$ .

Suponhamos agora que  $\tilde{T} \in \mathcal{L}(W, V)$  satisfaça:

$$\langle T(v), w \rangle = \langle v, \tilde{T}(w) \rangle \quad \forall v \in V, w \in W.$$



Então, para todo  $i \in \{1, \dots, m\}$ , tem-se

$$\tilde{T}(w_i) = \sum_{k=1}^n \langle \tilde{T}(w_i), v_k \rangle v_k = \sum_{k=1}^n \langle w_i, T(v_k) \rangle v_k = T^*(w_i),$$

donde se conclui que  $\tilde{T} = T^*$  e, portanto, que  $T^*$  é única. ■

A transformação linear  $T^* \in \mathcal{L}(V, W)$  no teorema acima é dita a *adjunta* de  $T \in \mathcal{L}(V, W)$ . Considerando-se as bases ortonormais  $\mathcal{B} \subset V$  e  $\mathcal{B}' \subset W$  como na demonstração, tem-se

$$[T^*]_{\mathcal{B}'}^{\mathcal{B}} = ([T]_{\mathcal{B}}^{\mathcal{B}'})^*. \quad (7.2)$$

Dito de forma simples, *com respeito a bases ortonormais de  $V$  e  $W$ , a matriz de  $T^*$  é a transposta da matriz de  $T$* . De fato, pondo-se

$$[T]_{\mathcal{B}}^{\mathcal{B}'} = (a_{ij})_{m \times n} \quad \text{e} \quad [T^*]_{\mathcal{B}'}^{\mathcal{B}} = (a_{ij}^*)_{n \times m},$$

tem-se

$$T(v_j) = \sum_{k=1}^m a_{kj} w_k \quad \text{e} \quad T^*(w_i) = \sum_{k=1}^n a_{ki}^* v_k.$$

Logo, fixando-se  $i$  e  $j$ , e tomando-se o produto interno em ambos os membros das equações acima por  $w_i$  e  $v_j$ , respectivamente, obtém-se

$$a_{ij} = \langle T(v_j), w_i \rangle = \langle v_j, T^*(w_i) \rangle = a_{ji}^*,$$

o que implica (7.2).

Segue-se das propriedades da transposição de matrizes e da igualdade (7.2) que a adjunção de transformações lineares tem as seguintes propriedades:

- i)  $(T^*)^* = T \quad \forall T \in \mathcal{L}(V, W)$ ;
- ii)  $(\lambda T + Z)^* = \lambda T^* + Z^* \quad \forall T, Z \in \mathcal{L}(V, W), \lambda \in \mathbb{R}$ ;
- iii)  $(ZT)^* = T^*Z^* \quad \forall T \in \mathcal{L}(V, W), Z \in \mathcal{L}(W, U)$ .

Note que, pela propriedade (ii), a aplicação

$$T \in \mathcal{L}(V, W) \mapsto T^* \in \mathcal{L}(W, V)$$

é linear.

Segue-se também das propriedades da adjunção que, dado um polinômio

$$p(t) = a_0 + a_1t + \cdots + a_nt^n \in \mathcal{P}[t, \mathbb{R}],$$

tem-se, para todo  $T \in \mathcal{L}(V, \langle, \rangle)$ , que

$$\begin{aligned} p(T^*) &= a_0I + a_1T^* + \cdots + a_n(T^n)^* \\ &= (a_0I + a_1T + \cdots + a_n(T^n))^* = (p(T))^*. \end{aligned} \quad (7.3)$$

Em particular, o polinômio mínimo de  $T$  anula  $T^*$  e vice-versa, donde se obtém o seguinte resultado.

**Proposição 7.1.** *O polinômio mínimo de todo operador  $T \in \mathcal{L}(V, \langle, \rangle)$  é igual ao polinômio mínimo de seu adjunto  $T^* \in \mathcal{L}(V, \langle, \rangle)$ .*

**Exemplo 7.1.** Seja  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  a transformação linear

$$T(x_1, x_2) = (x_1 - 2x_2, 3x_1 + x_2, x_1).$$

Uma vez que

$$[T] = \begin{bmatrix} 1 & -2 \\ 3 & 1 \\ 1 & 0 \end{bmatrix},$$

temos que

$$[T^*] = [T]^* = \begin{bmatrix} 1 & 3 & 1 \\ -2 & 1 & 0 \end{bmatrix}.$$

Logo,  $T^* : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  é dada por

$$T^*(x_1, x_2, x_3) = (x_1 + 3x_2 + x_3, -2x_1 + x_2).$$

Encerremos esta seção estabelecendo uma propriedade clássica do operador adjunto.

**Proposição 7.2.** *Sejam  $T \in (V, \langle \cdot, \cdot \rangle)$  um operador linear e  $W \subset V$  um subespaço  $T$ -invariante. Então  $W^\perp \subset V$  é  $T^*$ -invariante.*

*Demonstração.* Dado  $w^\perp \in W^\perp$ , temos, para todo  $w \in W$ , que  $\langle T(w), w^\perp \rangle = 0$ , pois  $W$  é  $T$ -invariante. Logo,

$$0 = \langle T(w), w^\perp \rangle = \langle w, T^*w^\perp \rangle,$$

donde se infere que  $T^*(w^\perp) \in W^\perp$  e, portanto, que  $W^\perp$  é  $T^*$ -invariante. ■

## 7.2 Operadores Ortogonais

Nesta seção, introduziremos os operadores ortogonais, os quais, conforme aludimos à introdução, são as isometrias lineares de um espaço vetorial munido de produto interno.

**Definição 7.1** (OPERADOR ORTOGONAL). Diz-se que um operador linear  $T \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$  é *ortogonal*, quando satisfaz  $TT^* = I$ .

Assim, um operador linear ortogonal é invertível, tendo seu adjunto como operador inverso (vide Exercício 25 – Cap. 3).

Vejamos que um operador ortogonal pode ser caracterizado de diversas maneiras, conforme a proposição seguinte.

**Proposição 7.3** (CARACTERIZAÇÃO DE OPERADORES ORTOGONAIS). *As seguintes afirmações a respeito de um operador linear  $T \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$  são equivalentes:*

- i)  $T$  é ortogonal;
- ii)  $T$  preserva produto interno, isto é,  $\langle T(v), T(w) \rangle = \langle v, w \rangle \forall v, w \in V$ ;
- iii)  $T$  preserva norma, isto é,  $\|T(v)\| = \|v\| \forall v \in V$ ;
- iv)  $T$  preserva distância, isto é,  $\text{dist}(T(v), T(w)) = \text{dist}(v, w) \forall v, w \in V$ .

*Demonstração.* Dado  $T \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$ , tem-se, para quaisquer  $v, w \in V$ , que

$$\langle T(v), T(w) \rangle = \langle v, T^*T(w) \rangle,$$

donde se infere que  $T$  preserva produto interno se, e somente se,  $T^*T = I$ . Logo (i) e (ii) são equivalentes.

Segue-se diretamente da identidade de polarização (vide Exercício 4 – Cap. 5) que  $T$  preserva norma se, e somente se, preserva produto interno, donde se infere que (ii) e (iii) são equivalentes.

Suponhamos que (iii) se cumpra. Então,

$$\text{dist}(T(v), T(w)) = \|T(v) - T(w)\| = \|T(v - w)\| = \|v - w\| = \text{dist}(v, w),$$

donde  $T$  preserva distância. Reciprocamente, se  $T$  preserva distância, para todo  $v \in V$ , tem-se

$$\|T(v)\| = \text{dist}(T(v), T(\mathbf{0})) = \text{dist}(v, \mathbf{0}) = \|v\|,$$

isto é,  $T$  preserva norma. Isto prova que (iii) e (iv) são equivalentes e conclui, desta forma, a demonstração. ■

Uma aplicação  $(V, \langle \cdot, \cdot \rangle) \rightarrow (V, \langle \cdot, \cdot \rangle)$  que preserva distância é dita uma *isometria* de  $(V, \langle \cdot, \cdot \rangle)$ . Segue-se, portanto, da Proposição 7.3, que *os operadores ortogonais  $T \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$  são as únicas isometrias lineares de  $(V, \langle \cdot, \cdot \rangle)$ .*

Uma outra consequência imediata da Proposição 7.3 é a de que todo operador ortogonal  $T \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$  leva bases ortonormais em bases ortonormais. Em particular, os vetores coluna da matriz de  $T$  com respeito a uma base ortonormal  $\mathcal{B}$ , de  $V$ , é também uma base ortonormal de  $V$ . Nesse caso, dizemos que a matriz  $[T]_{\mathcal{B}}$  é *ortogonal*.

**Exemplo 7.2** (OPERADORES ORTOGONAIS EM ESPAÇOS BIDIMENSIONAIS). Considere um espaço vetorial bidimensional  $(V, \langle \cdot, \cdot \rangle)$  e um operador ortogonal  $T$  definido em  $V$ . Dada uma base ortonormal  $\mathcal{B} \subset V$ , suponha que

$$[T]_{\mathcal{B}} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Uma vez que os vetores coluna dessa matriz são ortonormais, existe  $\theta \in [0, 2\pi)$ , tal que  $(a, c)_{\mathcal{B}} = (\cos \theta, \sin \theta)_{\mathcal{B}}$  e  $(b, d)_{\mathcal{B}} = \pm(-\sin \theta, \cos \theta)_{\mathcal{B}}$ , isto é,

$$[T]_{\mathcal{B}} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad \text{ou} \quad [T]_{\mathcal{B}} = \begin{bmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{bmatrix}.$$

Observe que, no primeiro caso, o polinômio característico de  $T$  é  $c(t) = t^2 + 1$ , o qual não possui raízes reais. Em particular,  $T$  não é diagonalizável. Já no segundo caso, o polinômio característico é  $c(t) = t^2 - 1$ , cujas raízes são 1 e  $-1$ , donde  $T$  é diagonalizável.

Acrescentamos ainda que, quando  $V = \mathbb{R}^2$ , no primeiro caso,  $T(v)$  corresponde a uma rotação do vetor  $v$  de um ângulo  $\theta$  no sentido anti-horário. De fato, supondo-se, sem perda de generalidade, que  $\mathcal{B}$  é a base canônica de  $\mathbb{R}^n$ , para todo vetor não nulo  $v = (x, y) \in \mathbb{R}^2$ , tem-se

$$\langle T(v), v \rangle = \langle (x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta), (x, y) \rangle = (x^2 + y^2) \cos \theta.$$

Desta forma, uma vez que  $x^2 + y^2 = \|v\|^2$  e  $\|T(v)\| = \|v\|$ , fazendo-se  $\phi = \angle(T(v), v)$ , tem-se

$$\cos \phi = \frac{\langle T(v), v \rangle}{\|T(v)\| \|v\|} = \frac{\|v\|^2 \cos \theta}{\|v\|^2} = \cos \theta,$$

donde  $\theta = \phi$ , se  $\theta \in (0, \pi)$ , ou  $\theta = 2\pi - \phi$ , caso contrário (Fig. 7.1).

Veremos, agora, que um operador ortogonal pode ser decomposto numa soma de operadores como os do Exemplo 7.2. Com esse fim, consideremos as propriedades dos operadores ortogonais destacadas na proposição seguinte.

**Proposição 7.4.** *As seguintes afirmações acerca de um operador ortogonal  $T \in \mathcal{L}(V, \langle, \rangle)$  são verdadeiras:*

- i) *Se  $\lambda \in \mathbb{R}$  é um autovalor de  $T$ , então  $\lambda = \pm 1$ ;*
- ii) *Dois autovetores de  $T$  associados a autovalores distintos são ortogonais;*
- iii) *Se  $W \subset V$  é um subespaço  $T$ -invariante, então  $W^\perp$  é  $T$ -invariante.*

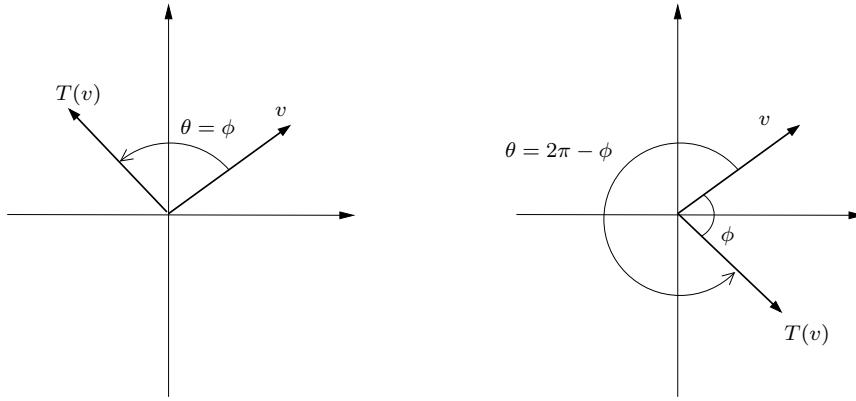


Figura 7.1: A rotação de um ângulo  $\theta > 0$  em  $\mathbb{R}^2$ .

*Demonstração.* Suponha que  $T(v) = \lambda v$  para algum  $v \in V - \{\mathbf{0}\}$ . Daí, uma vez que  $T$  preserva norma, obtém-se

$$\|v\| = \|T(v)\| = |\lambda|\|v\|,$$

donde  $|\lambda| = 1$  e, portanto,  $\lambda = \pm 1$ . Isso prova (i).

Se  $v, w \in V$  são tais que  $T(v) = v$  e  $T(w) = -w$ , então,

$$\langle v, w \rangle = \langle T(v), T(w) \rangle = -\langle v, w \rangle,$$

donde  $\langle v, w \rangle = 0$ , o que prova (ii).

Para provarmos (iii), observemos inicialmente que, sendo  $W$  um subespaço  $T$ -invariante, e  $T$  um operador invertível, temos que  $T_W : W \rightarrow W$  é injetivo e, portanto, bijetivo. Em particular,  $T^*(W) = T^{-1}(W) = W$ . Assim, para quaisquer  $w \in W$  e  $v \in W^\perp$ , tem-se  $T^*(w) \in W$ , donde

$$\langle T(v), w \rangle = \langle v, T^*(w) \rangle = 0,$$

isto é,  $T(v) \in W^\perp$ . Logo,  $W^\perp$  é  $T$ -invariante. ■

**Teorema 7.2.** *Para todo operador ortogonal  $T \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$ , existe uma base ortonormal de  $V$ ,  $\mathcal{B}$ , tal que  $[T]_{\mathcal{B}}$  é uma matriz diagonal em blocos da seguinte*

forma:

$$[T]_{\mathcal{B}} = \begin{bmatrix} I_{k_1} & & & & \\ & -I_{k_2} & & & \\ & & A_{\theta_1} & & \\ & & & \ddots & \\ & & & & A_{\theta_r} \end{bmatrix}, \quad (7.4)$$

em que  $I_{k_i}$  denota a matriz identidade de ordem  $k_i$ , e  $A_{\theta_i}$  é dada por:

$$A_{\theta_i} = \begin{bmatrix} \cos \theta_i & -\text{sen } \theta_i \\ \text{sen } \theta_i & \cos \theta_i \end{bmatrix}.$$

*Demonstração.* Sejam  $\overline{W}_1$  e  $\overline{W}_2$  os autoespaços definidos pelos autovalores 1 e  $-1$ , respectivamente. Ponha  $W = \overline{W}_1 \oplus \overline{W}_2$ , e considere o complemento ortogonal de  $W$ ,  $W^\perp$ . Uma vez que  $W$  é  $T$ -invariante, pela Proposição 7.4(ii),  $W^\perp$  é  $T$ -invariante.

Observemos que  $W^\perp$  não contém subespaços  $T$ -invariantes de dimensão 1, pois, pela Proposição 7.4(i), todos os autovetores de  $T$  estão em  $W$ . No entanto, pela Proposição 6.9,  $W^\perp$  contém um subespaço bidimensional,  $W_1$ , o qual é  $T$ -invariante. Uma vez que  $W_1$  não contém autovetores de  $T$ , temos, pelas considerações do Exemplo 7.2, que a matriz de  $T_{W_1}$  com respeito a uma base ortonormal de  $W_1$  é  $A_{\theta_1}$ , para algum  $\theta_1 \in \mathbb{R}$ .

Agora, basta notarmos que  $W_1^\perp$  é  $T$ -invariante, de modo que, procedendo indutivamente, obtemos um número finito de subespaços bidimensionais  $T$ -invariantes,  $W_1, \dots, W_r$ , os quais, juntamente com  $\overline{W}_1$  e  $\overline{W}_2$ , satisfazem:

$$V = \overline{W}_1 \oplus \overline{W}_2 \oplus W_1 \oplus \dots \oplus W_r.$$

Observando-se que os subespaços dessas decomposições são ortogonais entre si, segue-se dessa última igualdade que, tomando-se bases ortonormais em cada um dos subespaços  $W_i$  e  $\overline{W}_i$  e unindo-as, obtém-se uma base  $\mathcal{B}$ , de  $V$ , tal que a matriz  $[T]_{\mathcal{B}}$  é como em (7.4), sendo  $k_1$  e  $k_2$  as dimensões dos autoespaços  $\overline{W}_1$  e  $\overline{W}_2$ , respectivamente. ■

No teorema acima, deve-se observar que a decomposição do operador ortogonal  $T$  é a mais geral possível, isto é, para um determinado  $T$ , pode ocorrer de alguns dos subespaços  $\bar{W}_i$  ou  $W_i$  não fazerem parte da decomposição de  $V$ .

### 7.3 Operadores Autoadjuntos

Nesta seção, faremos nossas considerações a respeito dos operadores autoadjuntos, os quais, conforme anunciamos, protagonizam o célebre Teorema Espectral.

**Definição 7.2** (OPERADOR AUTOADJUNTO). Um operador  $T \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$  é dito *autoadjunto*, quando  $T = T^*$ , isto é, quando satisfaz

$$\langle T(v), w \rangle = \langle v, T(w) \rangle \quad \forall v, w \in V.$$

**Exemplo 7.3.** Seja  $\mathcal{B} \subset (V, \langle \cdot, \cdot \rangle)$  uma base ortonormal. Segue-se da igualdade (7.2) que um operador  $T \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$  será autoadjunto se, e somente se, a matriz  $[T]_{\mathcal{B}}$  for simétrica.

**Exemplo 7.4.** Dado um subespaço não trivial  $W \subset (V, \langle \cdot, \cdot \rangle)$ , podemos considerar a projeção ortogonal de  $V$  sobre  $W$  como um operador  $P_W : V \rightarrow V$ , cuja imagem é  $W$ . Desse ponto de vista, segue-se da simetria da projeção ortogonal que  $P_W$  é autoadjunto (vide Cap. 5).

**Proposição 7.5.** *Sejam  $v_1, v_2$  autovetores de um operador autoadjunto  $T \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$ , cujos correspondentes autovalores,  $\lambda_1$  e  $\lambda_2$ , são distintos. Então,  $v_1$  é ortogonal a  $v_2$ .*

*Demonstração.* Uma vez que  $T$  é autoadjunto, temos que

$$0 = \langle T(v_1), v_2 \rangle - \langle v_1, T(v_2) \rangle = (\lambda_1 - \lambda_2) \langle v_1, v_2 \rangle,$$

donde se infere que  $\langle v_1, v_2 \rangle = 0$ , já que  $\lambda_1 \neq \lambda_2$ . ■

Verifiquemos duas propriedades dos operadores autoadjuntos, as quais aplicaremos na demonstração do Teorema Espectral.



**Lema 7.1.** *As seguintes afirmações a respeito de um operador autoadjunto  $T \in (V, \langle \cdot, \cdot \rangle)$  são verdadeiras:*

- i) *Se  $T^k = \mathbf{0}$  para algum  $k \in \mathbb{N}$ , então  $T = \mathbf{0}$ ;*
- ii) *Se  $p(t) = t^2 + bt + c$  é primo (isto é,  $b^2 - 4c < 0$ ), então  $p(T)$  é invertível.*

*Demonstração.* (i) O resultado é evidente para  $k = 1$ . Suponhamo-lo verdadeiro para  $k \in \mathbb{N}$  e provemo-lo verdadeiro para  $k + 1$ , donde o resultado se seguirá do Princípio da Indução.

Suponhamos, então, que  $T^{k+1} = \mathbf{0}$ . Assim, para todo  $v \in V$ , tem-se

$$0 = \langle T^{k+1}(v), T^{k-1}(v) \rangle = \langle T^k(v), T^k(v) \rangle,$$

donde  $T^k(v) = \mathbf{0}$ , isto é,  $T^k = \mathbf{0}$ . Logo, pela hipótese de indução,  $T = \mathbf{0}$ .

(ii) Para todo vetor  $v \in V$ , tem-se  $\langle T^2(v), v \rangle = \langle T(v), T(v) \rangle = \|T(v)\|^2$ . Supondo-se  $v \neq 0$ , segue-se daí e da Desigualdade de Cauchy-Schwarz que

$$\begin{aligned} \langle p(T)v, v \rangle &= \langle T^2v + bTv + cv, v \rangle = \|Tv\|^2 + b\langle Tv, v \rangle + c\|v\|^2 \\ &\geq \|Tv\|^2 - |b|\|Tv\|\|v\| + c\|v\|^2 \\ &= \left( \|Tv\|^2 - \frac{|b|\|v\|}{2} \right)^2 + \left( c - \frac{b^2}{4} \right) \|v\|^2 > 0, \end{aligned}$$

o que nos dá  $p(T)v \neq 0$ . Logo, o núcleo de  $p(T)$  é trivial e, portanto,  $p(T) \in \mathcal{L}(V)$  é um operador invertível. ■

**Teorema Espectral.** Todo operador autoadjunto  $T \in \mathcal{L}(V)$  é diagonalizável. Consequentemente, existe uma base ortonormal de  $V$ , a qual é formada por autovetores de  $T$ .

*Demonstração.* Consideremos o polinômio mínimo de  $T$ ,

$$m = p_1^{s_1} \dots p_k^{s_k},$$

em que  $p_1, \dots, p_k \in \mathcal{P}[t, \mathbb{R}]$  são primos, distintos, e de grau 1 ou 2. Pelo Teorema da Decomposição Primária, para cada  $i = 1, \dots, k$ , temos que

$$W_i = \text{Nuc}(p_i(T))^{s_i}$$

é não trivial e  $T$ -invariante. Além disso,  $p_i^{s_i}$  é o polinômio mínimo de  $T_i = T_{W_i}$ , donde  $(p_i(T_i))^{s_i} = \mathbf{0}$ . Porém, sendo  $T$  um operador autoadjunto, e  $W_i$  um subespaço  $T$ -invariante, temos que  $T_i \in \mathcal{L}(W_i)$  é autoadjunto. Logo,  $p_i(T_i)$  é autoadjunto (vide (7.3)). Desta forma, pelo Lema 7.1-(i),  $p_i(T_i) = \mathbf{0}$ , isto é,  $p_i$  é o polinômio mínimo de  $T_i$ . Temos, também, pelo Lema 7.1-(ii), que  $p_i$  não pode ser de grau 2, já que o núcleo de  $p_i(T)$  é não trivial.

Segue-se das considerações acima que existem  $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ , tais que

$$m(t) = (t - \lambda_1) \dots (t - \lambda_k).$$

Logo, pelo Teorema 6.1,  $T$  é diagonalizável e os espaços  $T$ -invariantes  $W_i$  são, na verdade, os autoespaços  $W_{\lambda_i}$ .

Agora, Pela Proposição 7.5, os autoespaços  $W_{\lambda_i}$  são ortogonais entre si. Sendo assim, tomando-se uma base ortonormal  $\mathcal{B}_i$  em  $W_{\lambda_i}$ , tem-se que  $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$  é uma base ortonormal de  $V$  formada por autovetores de  $T$ , como queríamos demonstrar. ■

Na próxima seção, onde estudaremos as formas quadráticas, aplicaremos o Teorema Espectral, bem como outras propriedades dos operadores autoadjuntos.

## 7.4 Formas Quadráticas

Existe uma classe de funções polinomiais de duas variáveis reais, ditas formas quadráticas, as quais surgem naturalmente em diversos contextos da Geometria e da Análise. Uma tal função define-se como

$$Q(x, y) = ax^2 + 2bxy + cy^2, \quad (x, y) \in \mathbb{R}^2,$$

em que  $a, b$  e  $c$  são constantes reais.

Temos, por exemplo, que uma cônica de  $\mathbb{R}^2$  pode ser descrita como um conjunto de pontos que satisfaz uma equação do tipo

$$Q(x, y) = 1.$$

A depender de  $a$ ,  $b$  e  $c$ , a cônica será uma elipse, parábola ou hipérbole.

Já em análise, as formas quadráticas desempenham um papel fundamental no estudo de máximos e mínimos de funções de duas variáveis, uma vez que a derivada segunda de uma tal função (num ponto) é uma forma quadrática.

Nos dois casos citados acima, a propriedade de  $Q$  que determinará o tipo de cônica, ou o tipo de ponto crítico, máximo ou mínimo, está ligada à sua “positividade”. Para que isso fique mais claro, consideremos o operador autoadjunto  $T \in \mathcal{L}(\mathbb{R}^2)$  definido pela matriz

$$[T] = \begin{bmatrix} a & b \\ b & c \end{bmatrix},$$

e observemos que  $Q$  e  $T$  relacionam-se através da igualdade

$$Q(x, y) = \langle T(x, y), (x, y) \rangle.$$

Tomemos, então, uma base ortonormal  $\mathcal{B} = \{e'_1, e'_2\} \subset \mathbb{R}^2$  que diagonaliza  $T$ . Assim, dado  $v = (x', y')_{\mathcal{B}} \in \mathbb{R}^2$ , tem-se

$$Q(v) = Q(x', y')_{\mathcal{B}} = x'T(e'_1) + y'T(e'_2) = x'\lambda_1 e'_1 + y'\lambda_2 e'_2,$$

em que  $\lambda_1$  e  $\lambda_2$  são os autovalores de  $T$ . Logo,

$$Q(x', y')_{\mathcal{B}} = \lambda_1(x')^2 + \lambda_2(y')^2. \quad (7.5)$$

Segue-se de (7.5) que a equação  $Q(v) = 1$  representará uma:

- Elipse, se  $\lambda_1$  e  $\lambda_2$  forem positivos;
- Hipérbole, se  $\lambda_1\lambda_2 < 0$ ;
- Parábola, se  $\lambda_1 > 0$  e  $\lambda_2 = 0$  ou  $\lambda_1 = 0$  e  $\lambda_2 > 0$ .

Os casos remanescentes, por exemplo  $\lambda_1$  e  $\lambda_2$  negativos, são ditos *degenerados* e não resultam em cônicas.

Segue-se também da igualdade (7.5) que  $Q$  será uma função positiva, se  $\lambda_1$  e  $\lambda_2$  forem ambos positivos. Nesse caso, se  $Q$  for a derivada segunda de

uma função num ponto crítico, este será necessariamente um mínimo local. Analogamente, se  $\lambda_1$  e  $\lambda_2$  forem ambos negativos, o ponto crítico será um máximo local. Por fim, se  $\lambda_1\lambda_2 < 0$ , o ponto crítico não será mínimo ou máximo local.

De modo inteiramente análogo, podemos introduzir o conceito de forma quadrática num espaço vetorial arbitrário  $(V, \langle \cdot, \cdot \rangle)$ :

**Definição 7.3** (FORMA QUADRÁTICA). Seja  $T \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$  um operador autoadjunto. A *forma quadrática*  $Q$  associada a  $T$  é a função  $Q : V \rightarrow \mathbb{R}$ , a qual se define pela igualdade:

$$Q(v) = \langle T(v), v \rangle, \quad v \in V.$$

Uma forma quadrática  $Q$  (ou o operador  $T$  associado) é dita:

- i) *Positiva*, se  $Q(v) > 0 \forall v \neq \mathbf{0}$ ;
- ii) *Negativa*, se  $Q(v) < 0 \forall v \neq \mathbf{0}$ ;
- iii) *Indefinida*, se  $Q(v_1) > 0$  e  $Q(v_2) < 0$  para algum par  $v_1, v_2 \in V$ .

Em (i) e (ii), valendo as desigualdades com  $\geq$  e  $\leq$ , diz-se que  $Q$  é *não negativa* e *não positiva*, respectivamente.

Dada uma forma quadrática  $Q(v) = \langle T(v), v \rangle$ ,  $v \in V$ , considere uma base ortonormal  $\mathcal{B} = \{v_1, \dots, v_n\}$ , de  $V$ , que diagonaliza  $T$ , e denote por  $\lambda_1, \dots, \lambda_n$  os correspondentes autovalores. Nessas condições, dado

$$v = (x_1, \dots, x_n)_{\mathcal{B}} = x_1v_1 + \dots + x_nv_n \in V,$$

tem-se que  $Q$  se expressa como:

$$\begin{aligned} Q(v) &= Q(x_1, \dots, x_n)_{\mathcal{B}} = \langle T(x_1v_1 + \dots + x_nv_n), x_1v_1 + \dots + x_nv_n \rangle \\ &= \lambda_1x_1^2 + \dots + \lambda_nx_n^2. \end{aligned}$$

Daí, segue-se o resultado seguinte.

**Proposição 7.6.** *Seja  $Q : V \rightarrow \mathbb{R}$  uma forma quadrática definida por um operador autoadjunto  $T \in \mathcal{L}(V)$ , cujos autovalores são  $\lambda_1, \dots, \lambda_n$ . Então, são verdadeiras as seguintes afirmações:*

- $Q$  é positiva se, e somente se,  $\lambda_i > 0 \forall i = 1, \dots, n$ ;
- $Q$  é negativa se, e somente se,  $\lambda_i < 0 \forall i = 1, \dots, n$ ;
- $Q$  é indefinida se, e somente se,  $\lambda_i > 0$  e  $\lambda_j < 0$  para algum par de índices  $i \neq j$ .

Mantendo a notação acima, definamos

$$\lambda_m = \min\{\lambda_1, \dots, \lambda_n\} \quad \text{e} \quad \lambda_M = \max\{\lambda_1, \dots, \lambda_n\},$$

e observemos que, para todo  $v \in V$  satisfazendo  $\|v\|^2 = x_1^2 + \dots + x_n^2 = 1$ , são válidas as seguintes desigualdades:

$$Q(v_m) = \lambda_m \leq \lambda_1 x_1^2 + \dots + \lambda_n x_n^2 \leq \lambda_M = Q(v_M),$$

donde se obtém a proposição seguinte, a qual tem grande relevância na teoria de hipersuperfícies regulares do espaço euclidiano  $\mathbb{R}^n$ .

**Proposição 7.7.** *Seja  $Q : V \rightarrow \mathbb{R}$  uma forma quadrática definida por um operador autoadjunto  $T \in \mathcal{L}(V)$ , cujos autovalores são  $\lambda_1, \dots, \lambda_n$ . Então, valem as igualdades:*

$$\min_{\|v\|=1} Q(v) = \min\{\lambda_1, \dots, \lambda_n\} \quad \text{e} \quad \max_{\|v\|=1} Q(v) = \max\{\lambda_1, \dots, \lambda_n\}.$$

Destaquemos, na proposição seguinte, uma propriedade fundamental dos operadores não negativos.

**Proposição 7.8.** *Seja  $T \in \mathcal{L}(V)$  um operador linear não negativo. Então, existe um único operador não negativo  $R \in \mathcal{L}(V)$ , tal que  $R^2 = T$ .*

*Demonstração.* Sejam  $\lambda_1, \dots, \lambda_n$  os autovalores (não negativos) de  $T$ , e  $\mathcal{B}$  uma base ortonormal de  $V$  que diagonaliza  $T$ . Defina  $R \in \mathcal{L}(V)$  como o operador cuja matriz com respeito a  $\mathcal{B}$  é:

$$[R]_{\mathcal{B}} = \begin{bmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n} \end{bmatrix}.$$

Claramente,  $R$  é não negativo e  $R^2 = T$ , o que prova a existência de  $R$ .

Suponhamos, agora, que exista um operador não negativo  $S \in \mathcal{L}(V)$ , tal que  $S^2 = T$ . Nesse caso, uma vez que  $S^2$  comuta com  $S$ , tem-se que  $T$  comuta com  $S$ . Logo, pela Proposição 6.6, existe uma base  $\mathcal{B}$ , de  $V$ , que diagonaliza  $T$  e  $S$  simultaneamente. Daí, e da igualdade  $S^2 = T$ , segue-se que a matriz  $[S]_{\mathcal{B}}$  é, necessariamente, igual a  $[R]_{\mathcal{B}}$ , donde  $S = R$ . Logo,  $R$  é único. ■

**Definição 7.4** (RAIZ QUADRADA DE OPERADORES). O operador  $R \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$ , descrito no enunciado da Proposição 7.8, é dito a *raiz quadrada* de  $T$ , e é denotado por  $\sqrt{T}$ .

Um operador autoadjunto  $T \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$  determina também uma função bilinear simétrica  $B : V \times V \rightarrow \mathbb{R}$ , dita a *forma bilinear* associada a  $T$ , a qual é definida pela igualdade:

$$B(v, w) = \langle T(v), w \rangle, \quad (v, w) \in V \times V.$$

Note que, nesse contexto, o produto interno nada mais é que a forma bilinear associada ao operador identidade de  $V$ .

A bilinearidade de  $B$  segue-se diretamente da linearidade de  $T$  e da bilinearidade do produto interno, enquanto a simetria de  $B$  decorre do fato do operador  $T$  ser autoadjunto. Ademais, uma vez que o produto interno  $\langle \cdot, \cdot \rangle$  e o operador  $T$  são determinados por seus valores numa base qualquer de  $V$ , o mesmo vale para  $B$ . Mais precisamente, vale o resultado seguinte, cujos detalhes da demonstração deixamos a encargo do leitor (vide Exercício 19).

**Proposição 7.9.** *Seja  $\mathcal{B} = \{v_1, \dots, v_n\}$  uma base de  $(V, \langle \cdot, \cdot \rangle)$ . Então, dados  $n^2$  escalares  $a_{ij} \in \mathbb{R}$  satisfazendo  $a_{ij} = a_{ji}$ ,  $1 \leq i, j \leq n$ , existe uma única forma bilinear simétrica  $B : V \times V \rightarrow \mathbb{R}$ , a qual cumpre a igualdade*

$$B(v_i, v_j) = a_{ij}, \quad 1 \leq i, j \leq n.$$

Dado um operador autoadjunto  $T \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$ , a forma bilinear  $B$  e a forma quadrática  $Q$  relacionam-se através das seguintes igualdades, as quais são de verificação imediata:

- i)  $Q(v) = B(v, v) \forall v \in V$ ;
- ii)  $B(v, w) = \frac{1}{2} (Q(v + w) - Q(v) - Q(w)) \forall (v, w) \in V \times V$ .

Assim,  $B$  determina  $Q$ , pela igualdade (i), e  $Q$  determina  $B$ , pela igualdade (ii), dita uma *identidade de polarização*.

**Definição 7.5** (MATRIZ POSITIVA). Dizemos que uma matriz  $A \in \mathcal{M}(n)$  é *positiva*, quando é simétrica, e o operador autoadjunto associado,  $T_A \in \mathcal{L}(\mathbb{R}^n)$ , é positivo. Analogamente, define-se matriz *negativa*.

Note que, identificando-se uma matriz  $1 \times 1$  com sua única entrada, temos que a forma quadrática  $Q$  determinada por  $T_A$  se expressa como:

$$Q(v) = [v]^* A [v] \quad \forall v \in \mathbb{R}^n.$$

Dada uma matriz  $A = (a_{ij}) \in \mathcal{M}(n)$ , para cada  $k = 1, \dots, n$ , consideremos a submatriz de  $A$ ,  $A_k = (a_{ij})_{k \times k}$ , tal que  $1 \leq i, j \leq k$ , como abaixo.

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1k} & & a_{1n} \\ \vdots & & \vdots & & \vdots \\ & A_k & & & \\ a_{k1} & \cdots & a_{kk} & & \vdots \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & \cdots & \cdots & a_{nn} \end{bmatrix}.$$

Através das submatrizes  $A_k$ , pode-se estabelecer um interessante critério de positividade da matriz  $A$ , conforme o resultado seguinte.

**Proposição 7.10** (CRITÉRIO DE POSITIVIDADE PARA MATRIZES SIMÉTRICAS). *Uma matriz  $A \in \mathcal{M}(n)$  é positiva se, e somente se, para cada  $k \in \{1, \dots, n\}$ , o determinante da submatriz  $A_k$ , descrita acima, é positivo.*

*Demonstração.* Suponhamos que  $A$  seja positiva. Então, todos os autovalores de  $A$  são positivos, de modo que  $\det A = \det A_n > 0$ . Ademais, para  $1 \leq k < n$ ,  $A_k$  é uma matriz simétrica de  $\mathcal{M}(k)$ . Logo, identificando-se com  $\mathbb{R}^k$  o subespaço de  $\mathbb{R}^n$  gerado por  $e_1, \dots, e_k$ , tem-se

$$[v]^* A_k [v] = [v]^* A [v] > 0 \quad \forall v \in \mathbb{R}^k,$$

donde  $A_k$  é positiva. Desta forma, assim como  $A_n$ , cada submatriz  $A_k$  tem determinante positivo.

Reciprocamente, suponhamos que  $\det A_k > 0 \forall k \in \{1, \dots, n\}$ , e provemos, por indução sobre  $n$ , que  $A$  é positiva. Para  $n = 1$ , o resultado é imediato. Suponhamo-lo verdadeiro, então, para  $n - 1 \geq 1$ . Nesse caso, a submatriz  $A_{n-1}$  é positiva. Em particular, para todo vetor não nulo  $v$  de  $\mathbb{R}^{n-1}$  (considerado aqui como o espaço gerado por  $\{e_1, \dots, e_{n-1}\}$ ), tem-se  $[v]^* A[v] = [v]^* A_{n-1}[v] > 0$ , donde se infere que  $A$  não é negativa.

Uma vez que, por hipótese,  $\det A = \det A_n > 0$ , temos que todos os autovalores de  $A$  são não nulos e, pelas considerações do parágrafo anterior, nem todos negativos. Claramente,  $A$  não pode ter apenas um autovalor negativo. Suponhamos, então, que  $A$  possua dois autovalores negativos, e denotemos por  $W$  o subespaço bidimensional de  $\mathbb{R}^n$  gerado pelos respectivos autovetores. Note que, para todo  $v \in W$ ,  $[v]^* A[v] < 0$ .

Agora, basta notar que a interseção de  $W$  com  $\mathbb{R}^{n-1} \subset \mathbb{R}^n$  é não trivial, uma vez que a soma das dimensões desses subespaços excede  $n$ . Isso, porém, contradiz a positividade de  $A_{n-1}$ . Logo,  $A$  é positiva. ■

**Exemplo 7.5.** Considere as matrizes simétricas:

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 5 & 1 \\ 1 & 1 & 3 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 1 & -1 \\ 2 & -1 & 3 \end{bmatrix}.$$

Temos que  $\det A_1 = 1$ ,  $\det A_2 = 1$  e  $\det A_3 = \det A = 1$ . Logo,  $A$  é positiva. Quanto à matriz  $B$ , temos que  $\det B_3 = \det B = -7$ , donde  $B$  não é positiva.

## Exercícios

### Seção 7.1

1. Dado  $T \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$ , suponha que  $T(v) = \lambda v$  e  $T^*(w) = \mu w$ . Mostre que,  $\lambda = \mu$  ou  $v$  é ortogonal a  $w$ .



2. Dada  $M \in \mathcal{M}(n)$ , defina  $T_M \in \mathcal{L}(\mathcal{M}(n))$  por  $T_M(A) = MA$ . Mostre que, considerando-se em  $\mathcal{M}(n)$  o produto interno  $\langle A, B \rangle = \text{traço}(A^*B)$ , vale a igualdade  $T_M^* = T_{M^*}$ .
3. Dado  $T \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$ , mostre que, se  $\text{traço}(T^*T) = 0$ , então  $T = \mathbf{0}$ .
4. Dado um operador linear  $T \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$ , mostre que  $T^*(V) = (\text{Nuc } T)^\perp$ .
5. Mostre que dois operadores  $T, S \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$  comutam se, e somente se,  $T^*$  e  $S^*$  comutam.
6. Seja  $\{v_1, \dots, v_n\}$  uma base ortonormal de  $(V, \langle \cdot, \cdot \rangle)$ . Dado  $T \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$ , mostre que vale a seguinte igualdade:

$$\sum_{i=1}^n \|Tv_i\|^2 = \sum_{i=1}^n \|T^*v_i\|^2.$$

### Seção 7.2

7. Mostre que, se  $T : V \rightarrow V$  é um operador ortogonal, então  $|\det T| = 1$ .
8. Mostre que um operador  $T \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$  é ortogonal se, e somente se, existe uma base ortonormal  $\mathcal{B}$ , de  $V$ , tal que  $T(\mathcal{B})$  é uma base ortonormal de  $V$ .
9. Seja  $T \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$  um operador ortogonal. Mostre que, se a dimensão de  $V$  for ímpar, então existe um vetor não nulo  $v \in V$ , tal que  $T^2(v) = v$ .
10. Seja  $\mathcal{B} = \{v_1, \dots, v_n\}$  uma base ortonormal de  $(V, \langle \cdot, \cdot \rangle)$ . Suponha que  $T \in \mathcal{L}(V, \langle \cdot, \cdot \rangle)$  cumpra  $\|T(v_i)\| = 1 \forall i = 1, \dots, n$ . Pergunta-se:  $T$  é, necessariamente, um operador ortogonal?

### Seção 7.3

11. Mostre que não existe um operador autoadjunto  $T \in \mathcal{L}(\mathbb{R}^3)$ , tal que  $T(1, 1, 1) = (2, 2, 2)$  e  $T(1, 1, -1) = (-2, -2, 2)$ .

12. Dado um operador autoadjunto  $T \in \mathcal{L}(V)$ , mostre que existe um único operador autoadjunto  $R \in \mathcal{L}(V)$ , tal que  $R^3 = T$ . (Compare com o resultado da Proposição 7.8).
13. Sejam  $T, S \in \mathcal{L}(V)$  operadores autoadjuntos. Mostre que  $TS$  e  $ST$  são autoadjuntos se, e somente se,  $T$  e  $S$  comutam.
14. Mostre que  $\mathcal{S}(V) = \{T \in \mathcal{L}(V); T \text{ é autoadjunto}\}$  é um subespaço vetorial de  $\mathcal{L}(V)$ . Qual a dimensão de  $\mathcal{S}(V)$ ?
15. Seja  $T \in \mathcal{L}(V, \langle, \rangle)$  um operador linear *idempotente*, isto é,  $T$  satisfaz  $T^2 = T$ . Mostre que  $T$  é autoadjunto se, e somente se,  $T$  e  $T^*$  comutam.
16. Seja  $T \in \mathcal{L}(\mathbb{R}^3)$  um operador ortogonal e autoadjunto. Prove que  $T$  é  $\pm I$ , uma reflexão com respeito a um plano de  $\mathbb{R}^2$  pela origem, ou uma rotação de  $180^\circ$  em torno de uma reta de  $\mathbb{R}^3$  pela origem.
17. Sejam  $\lambda_1, \dots, \lambda_k$  os autovalores distintos de um operador autoadjunto  $T \in \mathcal{L}(V, \langle, \rangle)$ . Denote por  $P_i$  a projeção ortogonal de  $V$  sobre o autoespaço  $W_{\lambda_i}$ . Mostre que valem as seguintes igualdades:

$$\text{i) } T = \lambda_1 P_1 \oplus \dots \oplus \lambda_k P_k;$$

$$\text{ii) } I = P_1 \oplus \dots \oplus P_k;$$

$$\text{iii) } P_i P_j = \mathbf{0} \text{ para quaisquer } i \neq j \in \{1, \dots, k\}.$$

(Essas igualdades constituem o que se pode chamar de uma versão alternativa do Teorema Espectral).

### Seção 7.4

18. Para cada uma das formas quadráticas abaixo, determine a natureza da cônica de equação  $Q(x, y) = 1$ .
- i)  $Q(x, y) = x^2 + 2xy + 3y^2$ ;
- ii)  $Q(x, y) = x^2 + xy + y^2$ ;
- iii)  $Q(x, y) = 2x^2 - 2xy + y^2$ .

19. Demonstre a Proposição 7.9.
20. Mostre que, para toda transformação linear  $T : V \rightarrow W$ , tem-se que  $T^*T \in \mathcal{L}(V)$  é um operador não negativo.
21. Seja  $T \in \mathcal{L}(V)$  um operador autoadjunto positivo. Mostre que  $T$  é invertível e que  $T^{-1}$  é positivo.
22. Seja  $T \in \mathcal{L}(V)$  um operador autoadjunto. Mostre que existe  $c \in \mathbb{R}$ , tal que o operador  $T + cI$  é positivo.
23. Sejam  $T \in \mathcal{L}(V)$  um operador não negativo e  $B$  e  $Q$  as formas bilinear e quadrática associadas a  $T$ . Mostre que

$$|B(v, w)| \leq \sqrt{Q(v)}\sqrt{Q(w)} \quad \forall v, w \in V.$$

(Compare com a Desigualdade de Cauchy–Schwarz).

24. Sejam  $A, B \in \mathcal{M}(n)$  matrizes positivas. Mostre que  $A + B$  e  $A^k$ ,  $k \in \mathbb{N}$ , são positivas.
25. Seja  $T : V \rightarrow V$  um operador linear, tal que  $T = SR$ , em que  $S$  é ortogonal e  $R > 0$ . Mostre que  $R = \sqrt{T^*T}$ .



# *Apêndices*



# A

---

## *Decomposição Racional e o Teorema de Cayley–Hamilton*

Neste apêndice, estabeleceremos um resultado fundamental da teoria de decomposição de operadores, o Teorema de Decomposição Racional. Em seguida, o aplicaremos na demonstração de uma versão generalizada do Teorema de Cayley–Hamilton, o qual citamos anteriormente. Convém mencionar que a prova que daremos para o Teorema da Decomposição Racional, devida a H. G. Jacob<sup>(i)</sup>, tem as virtudes da simplicidade e da elegância, as quais nos motivaram a adaptá-la ligeiramente e, então, apresentá-la aqui.

### **O Teorema da Decomposição Racional**

**Definição A.1.** Dados  $T \in \mathcal{L}(V)$  e  $v \in V$ , chamamos o conjunto

$$Z(v, T) = \{p(T)(v) ; p \in \mathcal{P}[t, \mathbb{R}]\}$$

de subespaço  $T$ -cíclico gerado por  $v$ .

---

<sup>(i)</sup>Vide: Jacob, H. G.: Another proof of the rational decomposition theorem. Amer. Math. Monthly. 1131–1134 (1973).

Verifica-se sem dificuldade que  $Z(v, T)$  é, de fato, um subespaço vetorial de  $V$ , o qual é  $T$ -invariante. Vejamos, na proposição seguinte, outras propriedades deste subespaço.

**Proposição A.1.** *Dados  $T \in \mathcal{L}(V)$  e  $v \in V$ , valem as seguintes afirmações:*

- i)  $\dim Z(v, T) = \text{grau}(m_v)$  ;
- ii)  $Z(v, T) = \text{Ger}\{v, T(v), \dots, T^{g-1}(v)\}$ ,  $g = \text{grau}(m_v)$  ;
- iii) *O polinômio mínimo de  $T_{Z(v, T)}$  é  $m_v$  ;*
- iv) *O polinômio característico de  $T_{Z(v, T)}$  é  $m_v$  .*

*Demonstração.* Dado um polinômio  $p \in \mathcal{P}[t, \mathbb{R}]$ , pelo Algoritmo da Divisão, existem polinômios  $q, r \in \mathcal{P}[t, \mathbb{R}]$ , tais que

$$p = qm_v + r,$$

em que  $r = 0$  ou  $\text{grau}(r) < \text{grau}(m_v)$ . Daí, tem-se

$$p(T)(v) = q(T)m_v(T)(v) + r(T)(v) = r(T)(v).$$

Se  $r = 0$ , então  $p(T)(v) = r(T)(v) = \mathbf{0} \in \text{Ger}\{v, T(v), \dots, T^{g-1}(v)\}$ . O mesmo vale, se  $\text{grau}(r) < \text{grau}(m_v) = g$ , pois, nesse caso,  $r(T)(v)$  se expressa, necessariamente, como uma combinação linear dos vetores  $v, T(v), \dots, T^{g-1}(v)$ . Logo, todo elemento  $p(T)(v) \in Z(v, T)$  pertence a  $\text{Ger}\{v, T(v), \dots, T^{g-1}(v)\}$ . Além disso,  $\{v, T(v), \dots, T^{g-1}(v)\}$  é LI. Caso contrário, existiria um polinômio anulador de  $T$  em  $v$ , cujo grau seria menor que o de  $m_v$ . Assim, ficam provadas as afirmações (i) e (ii).

Quanto a (iii), observemos que  $m_v$  anula  $T_{Z(v, T)}$ . Com efeito, dado um vetor  $w = p(T)(v) \in Z(v, T)$ , segue-se da  $T$ -invariância de  $Z(v, T)$  que

$$m_v(T_{Z(v, T)})(w) = m_v(T)(w) = m_v(T)p(T)(v) = p(T)m_v(T)(v) = \mathbf{0}.$$

Consequentemente, o polinômio mínimo de  $T_{Z(v, T)}$  é  $m_v$ , pois, de outra forma, o polinômio mínimo de  $T_{Z(v, T)}$  seria um anulador de  $v$  de grau menor que o de  $m_v$ . Isso prova (iii).



Para provarmos (iv), designemos por  $\mathcal{B}$  a base  $\{v, T(v), \dots, T^{g-1}(v)\}$  de  $Z(v, T)$ . Os  $g$  vetores coluna da matriz de  $T_{Z(v, T)}$  com respeito a  $\mathcal{B}$  são

$$T(v), T^2(v), \dots, T^g(v).$$

Além disso, fazendo-se

$$m_v(t) = a_0 + a_1 t + \dots + a_{g-1} t^{g-1} + t^g,$$

e levando-se em conta que  $m_v(T)(v) = 0$ , obtém-se

$$T^g(v) = -a_0 - a_1 T(v) - \dots - a_{g-1} T^{g-1}(v).$$

Segue-se dessas considerações que

$$[T_{Z(v, T)}]_{\mathcal{B}} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ 0 & 0 & 1 & \ddots & 0 & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{g-1} \end{bmatrix}.$$

Conseqüentemente,

$$p_{T_{Z(v, T)}}(t) = \det(tI - T_{Z(v, T)}) = \begin{vmatrix} t & 0 & 0 & \cdots & 0 & a_0 \\ -1 & t & 0 & \cdots & 0 & a_1 \\ 0 & -1 & t & \cdots & 0 & a_2 \\ 0 & 0 & -1 & \ddots & 0 & \vdots \\ \vdots & \vdots & \vdots & \ddots & t & \vdots \\ 0 & 0 & 0 & \cdots & -1 & t + a_{g-1} \end{vmatrix},$$

donde (vide (4.11)),

$$p_{T_{Z(v, T)}}(t) = a_0 + a_1 t + \dots + a_{g-2} t^{g-2} + (t + a_{g-1}) t^{g-1} = m_v,$$

o que mostra (iv) e encerra nossa demonstração. ■

**Observação A.1.** Atentemos para o seguinte fato, o qual decorre diretamente dos itens (i) e (ii) da proposição acima: Dado  $v \in V$ , se  $k \geq \text{grau}(m_v)$ , tem-se que o subespaço  $W = \text{Ger}\{v, T(v), \dots, T^{k-1}(v)\}$  é igual a  $Z(v, T)$ . Em particular,  $W$  é  $T$ -invariante.

**Teorema A.1.** *Seja  $T \in \mathcal{L}(V)$  um operador linear cujo polinômio mínimo é  $m$ . Então, existe um vetor não nulo  $v \in V$ , tal que  $m_v = m$ . Em particular, valem as igualdades:*

$$Z(v, T) = \text{Ger}\{v, T(v), \dots, T^{g-1}(v)\} \quad e \quad m' = m_v = m,$$

em que  $g = \text{grau}(m)$  e  $m'$  é o polinômio mínimo de  $T_{Z(v, T)}$ .

*Demonstração.* Consideremos a decomposição  $V = W_1 \oplus \dots \oplus W_k$  dada no Teorema da Decomposição Primária. Fixemos  $i \in \{1, \dots, k\}$ , e observemos que existe um vetor não nulo  $v_i \in W_i$ , tal que  $\{v_i, T_i(v_i), \dots, T_i^{g_i-1}(v_i)\}$  é LI, em que  $T_i = T_{W_i}$  e  $g_i = r_i s_i$  é o grau do polinômio mínimo de  $T_i$ ,  $m_i = p_i^{s_i}$ .

Com efeito, se não fosse o caso, para todo  $v \in W_i$ , existiriam escalares  $a_0, \dots, a_{g_i-1} \in \mathbb{R}$ , não todos nulos, tais que o polinômio

$$p(t) = a_0 + \dots + a_{g_i-1} t^{g_i-1}$$

anuladoria  $T_i$  em  $v$ . Daí, teríamos  $\text{grau}(m_v) \leq g_i - 1 < g_i$ . Em particular, existiria  $s_v < s_i$ , tal que  $m_v(t) = (p_i(t))^{s_v}$ , pois  $m_v$  é um divisor de  $m_i$ . Consequentemente, teríamos que o polinômio  $(p_i(t))^{s_i-1}$ , cujo grau é menor que  $g_i$ , anularia  $T_i$  em  $v$  para todo  $v \in W_i$ , contradizendo, assim, a minimalidade de  $m_i$ .

Temos que  $m_i$  anula  $T_i$  em  $v_i$ . Logo,  $m_{v_i}$  divide  $m_i$ . Suponha que valha a desigualdade  $g'_i = \text{grau}(m_{v_i}) < \text{grau}(m_i) = g_i$ . Nesse caso, escrevendo-se  $m_{v_i}(t) = a_0 + a_1 t + \dots + a_{g'_i-1} t^{g'_i-1} + t^{g'_i}$ , tem-se

$$\mathbf{0} = m_{v_i}(T)(v_i) = a_0 v + a_1 T(v_i) + a_{g'_i-1} T^{g'_i-1}(v_i) + \dots + T^{g'_i}(v_i),$$

o que contradiz o fato de o conjunto  $\{v_i, T(v_i), \dots, T^{g'_i}(v_i)\}$  ser LI (por ser um subconjunto de  $\{v_i, T(v_i), \dots, T^{g_i-1}(v_i)\}$ ). Logo,  $\text{grau}(m_{v_i}) = \text{grau}(m_i)$  e, portanto,  $m_{v_i} = m_i \quad \forall i = 1, \dots, k$ , isto é,

$$m = m_{v_1} \dots m_{v_k}.$$

Agora, observemos que

$$\mathbf{0} = m_v(T)(v) = m_v(T_1)(v_1) + \cdots + m_v(T_k)(v_k),$$

o que nos dá  $m_v(T_i)(v_i) = \mathbf{0} \forall i = 1, \dots, k$ . Segue-se que  $m_v$  é um anulador de cada  $T_i$  em  $v_i$ , donde  $m_v$  divide  $m_{v_i}$ . Uma vez que cada  $m_{v_i}$  é um fator primo de  $m$  e  $m_v$  divide  $m$ , segue-se que

$$m_v = m_{v_1} \dots m_{v_k} = m.$$

A ultima asserção do enunciado é uma consequência direta do Teorema A.1 e da igualdade  $m_v = m$ . ■

**Teorema da Decomposição Racional.** Para todo operador linear  $T \in \mathcal{L}(V)$ , existem vetores não nulos  $v_1, \dots, v_k \in V$ , tais que

$$V = Z(v_1, T) \oplus \cdots \oplus Z(v_k, T).$$

Ademais, para todo  $i = 1, \dots, k-1$ ,  $m_{v_{i+1}}$  divide  $m_{v_i}$ .

*Demonstração.* Seja  $v \in V$  um vetor não nulo satisfazendo  $m_v = m$ , cuja existência é assegurada pelo Teorema A.1. Se  $V = Z(v, T)$ , nada há a ser provado. Caso contrário, ponha  $g = \text{grau}(m_v) = \text{grau}(m)$ , e considere vetores  $v_{g+1}, \dots, v_n \in V$ , tais que

$$\mathcal{B} = \{v, T(v), \dots, T^{g-1}(v), v_{g+1}, \dots, v_n\}$$

seja uma base de  $V$ . Note que os primeiros  $g$  vetores de  $\mathcal{B}$ ,  $v, T(v), \dots, T^{g-1}(v)$ , formam uma base de  $Z(v, T)$ .

Façamos  $W = \text{Ger}\{v_{g+1}, \dots, v_n\}$  e tomemos o produto interno  $\langle \cdot, \cdot \rangle$  de  $V$ , com respeito ao qual  $\mathcal{B}$  é uma base ortonormal (vide Proposição ??). Nesse caso, temos que  $W = Z(v, T)^\perp$ .

Consideremos  $w = T^{g-1}(v)$ , e definamos o conjunto

$$W^* = \text{Ger}\{w, T^*(w), \dots, (T^*)^{g-1}(w)\},$$

em que  $T^* \in \mathcal{L}(V)$  é a adjunta de  $T$ . Uma vez que, pela Proposição 7.1,  $m = m_v$  é também o polinômio mínimo de  $T^*$ , tem-se

$$g = \text{grau}(m) \geq \dim Z(w, T^*),$$

donde  $W^* = Z(w, T^*)$ . Logo,  $W^*$  é  $T^*$ -invariante (vide Observação A.1).

Verifiquemos, então, que  $W^*$  cumpre

$$\dim W^* = g \quad \text{e} \quad W^* \cap W = \{\mathbf{0}\}. \quad (\text{A.1})$$

Para tanto, suponhamos, por absurdo, que existam constantes  $a_0, \dots, a_s$ , com  $a_s \neq 0$  e  $0 \leq s \leq g-1$ , tais que

$$w' = a_0 w + a_1 T^*(w) + \dots + a_s (T^*)^s(w) \in W.$$

Uma vez que  $W$  é  $T^*$ -invariante, temos que  $(T^*)^{g-s-1}(w') \in W$ . Logo,

$$\begin{aligned} 0 &= \langle (T^*)^{g-s-1}(w'), v \rangle \\ &= \langle a_0 (T^*)^{g-s-1} w + a_1 (T^*)^{g-s}(w) + \dots + a_s (T^*)^{g-1}(w), v \rangle \\ &= \langle w, a_0 T^{g-s-1}(v) + a_1 T^{g-s}(v) + \dots + a_s T^{g-1}(v) \rangle \\ &= \langle T^{g-1}(v), a_0 T^{g-s-1}(v) + a_1 T^{g-s}(v) + \dots + a_s T^{g-1}(v) \rangle = a_s, \end{aligned}$$

o que, evidentemente, é uma contradição.

Note que o argumento acima mostra não somente que  $W \cap W^* = \{\mathbf{0}\}$ , mas também que  $\dim W^* = g$ . Com efeito, fazendo-se  $w' = 0$  em (A.1), temos que os escalares  $a_i$ , tais como especificados, não existem, donde se infere que os vetores  $w, T^*(w), \dots, (T^*)^{g-1}(w)$  são LI.

Segue-se que as igualdades (A.1) se cumprem e, em particular, que

$$V = W \oplus W^*.$$

Agora, sendo  $W^*$   $T^*$ -invariante, temos, pela Proposição 7.2, que o subespaço  $W' = (W^*)^\perp$  é  $T$ -invariante. Além disso (vide Exercício 11 – Cap. 5),

$$V = W \oplus W^* = W^\perp \oplus (W^*)^\perp = Z(v, T) \oplus W'.$$

Em particular, o polinômio mínimo de  $T_{W'}$  divide  $m$ .

Uma vez que  $W'$  é  $T$ -invariante, nós podemos escrever  $v = v_1$ , e repetir o argumento acima para concluir que existem  $v_2 \in V - \{\mathbf{0}\}$  e um subespaço  $T$ -invariante  $W'' \subset V$ , tais que

$$V = Z(v_1, T) \oplus Z(v_2, T) \oplus W'',$$

em que o polinômio mínimo de  $T_{W'}$  é  $m_{v_2}$ . Desta forma,  $m_{v_2}$  divide  $m_{v_1} = m$ .

O resultado segue-se imediatamente das considerações acima e do Princípio da Indução. ■

**Teorema de Cayley-Hamilton (generalizado).** O polinômio mínimo  $m$  de todo operador  $T \in \mathcal{L}(V)$  divide seu polinômio característico  $c$ . Além disso, se

$$m(t) = (p_1(t))^{s_1} \dots (p_q(t))^{s_q}, \quad (\text{A.2})$$

sendo  $p_1, \dots, p_q$  os fatores primos e distintos de  $m$ , então

$$c(t) = (p_1(t))^{n_1/r_1} \dots (p_q(t))^{n_q/r_q},$$

em que  $n_i$  é a dimensão de  $W_i = \text{Nuc}(p_i^{s_i}(T))$  e  $r_i = \text{grau}(p_i)$ .

*Demonstração.* Sejam  $v_1, \dots, v_k$  como no Teorema da Decomposição Racional, de modo que

$$V = Z(v_1, T) \oplus \dots \oplus Z(v_k, T) \quad \text{e} \quad T = T_1 \oplus \dots \oplus T_k,$$

em que  $T_i = T_{Z(v_i, T)}$ . Denotando-se por  $c_i$  o polinômio característico de  $T_i$ , e aplicando-se a igualdade (6.5) ao operador  $T - tI$ , conclui-se que  $c = c_1 \dots c_k$ . Daí, lembrando-se que  $m = m_{v_1}$ , tem-se

$$c = c_1 \dots c_k = m_{v_1} \dots m_{v_k} = m(m_{v_2} \dots m_{v_k}), \quad (\text{A.3})$$

em que, nas duas últimas igualdades, usamos a Proposição A.1.

Temos, portanto, que  $m$  divide  $c$ . Além disso, como  $m_{v_{i+1}}$  divide  $m_{v_i}$ , segue-se de (A.3) que cada fator primo de  $(m_{v_2} \dots m_{v_k})$  é um fator primo de  $m$ , donde se infere que cada fator primo de  $c$  é também um fator primo de  $m$ .

Pelo Teorema da Decomposição Primária, temos que  $V = W_1 \oplus \cdots \oplus W_q$ , e que o polinômio mínimo de  $T_{W_i}$  é  $m_i = p_i^{s_i}$ . Como já provado,  $m_i$  divide o polinômio característico de  $T_{W_i}$ , o qual denotamos por  $c'_i$ . Juntando-se a isso o fato de que o grau de  $c'_i$  é igual a  $\dim W_i = n_i$ , tem-se  $c'_i = p_i^{n_i/r_i}$ , donde

$$c = c'_1 \cdots c'_q = p_1^{n_1/r_1} \cdots p_q^{n_q/r_q},$$

como queríamos demonstrar. ■

A versão clássica do Teorema de Cayley–Hamilton afirma que todo operador linear (definido num espaço vetorial de dimensão finita) é anulado por seu polinômio característico. Esse fato, claramente, segue-se do teorema acima.

# *Referências Bibliográficas*

- [1] AXLER, S. (1997). *Linear Algebra Done Right*. Springer.
- [2] BROIDA, J. G. e WILLIAMSON, S. G. (2012). *Comprehensive Introduction to Linear Algebra*. Creative Commons CC0 1.0. Disponível no sítio: [cseweb.ucsd.edu/~gill/CILASite/](http://cseweb.ucsd.edu/~gill/CILASite/).
- [3] DE SOUZA, P. N. e SILVA, J-N. (2004). *Berkeley Problems in Mathematics*. Springer.
- [4] HALMOS, P. R. (1987). *Finite Dimensional Vector Spaces*. Springer.
- [5] HOFFMAN, K. e KUNZE, R. (1971). *Linear Algebra*. Prentice Hall.
- [6] LIMA, E. L. (2009). *Álgebra Linear*. IMPA.





# *Índice Remissivo*

- ângulo, 115
- adição
  - de matrizes, 2
  - de transformações lineares, 65
  - de vetores, 19
- aplicação
  - identidade, 50
  - nula, 50
- autoespaço, 141
- autovalor, 140
- autovetor, 140
- base, 31
  - canônica, 34
- bloco diagonal, 99
- cofator, 94
- combinação linear, 23
  - nula, 28
  - trivial, 28
- complemento ortogonal, 121
- comutatividade
  - de operadores, 68
- comutatividade de matrizes, 5
- conjunto
  - linearmente dependente, 28
  - linearmente independente, 28
  - ortogonal, 117
  - ortonormal, 117
- conjunto solução, 7
- coordenadas, 34
- critério de Sylvester, 92
- desenvolvimento de Laplace, 87
- desigualdade
  - de Bessel, 130
  - de Hadamard, 132
  - triangular, 129
- determinante, 81
  - de um operador, 94
- diagonal, 2
- diferença de vetores, 22
- dimensão, 34
- dimensão finita, 31
- distância, 108, 113
- elemento neutro, 20
- escalar, 3, 20
- espaço coluna, 38
- espaço linha, 38
- espaço vetorial, 19
- espaços vetoriais
  - isomorfos, 69
- forma

- $n$ -linear, 82
- alternada, 83
- antissimétrica, 83
- bilinear, 178
- quadrática, 176
  - indefinida, 176
  - negativa, 176
  - não negativa, 176
  - não positiva, 176
  - positiva, 176
- funcional linear, 50
- função
  - alternada, 80
  - antissimétrica, 80
  - bilinear, 80
  - determinante, 84
- homotetia, 50
- identidade
  - de polarização, 129, 179
  - do paralelogramo, 129
- imagem, 60
- isometria, 168
- isomorfismo, 69
- matriz, 1
  - adjunta clássica, 94
  - antissimétrica, 16, 105
  - aumentada, 8
  - coluna, 8, 24
  - de Gram, 125, 126
  - de mudança de base, 59
  - de Sylvester, 101
  - de transformação linear, 55
  - de um sistema linear, 7
  - de Vandermonde, 105
  - diagonal, 5
  - diagonal por blocos, 99
  - diagonalizável, 5, 15
  - dos cofatores, 94
  - escalonada, 10
  - identidade, 5
  - inversa, 14
  - invertível, 14
  - negativa, 179
  - nula, 2
  - ortogonal, 105, 168
  - positiva, 179
  - quadrada, 2
  - simétrica, 5
  - transposta, 6
  - triangular, 5
    - inferior, 5
    - superior, 5
- matrizes
  - linha equivalentes, 9
  - semelhantes, 14
- multiplicação
  - por escalar, 19
- método de Gauss, 11
- norma, 108, 113
- nulidade, 61
- núcleo, 60
- operador
  - autoadjunto, 172

- diagonalizável, 138, 140
- idempotente, 182
- nilpotente, 76
- ortogonal, 167
- operador linear, 50
- operações elementares, 8
- oposto, 20
  
- paralelepípedo, 126
- permutação, 88
  - par, 89
  - ímpar, 89
- polinômio
  - anulador, 148
    - em um vetor, 151
  - característico, 143
  - mínimo
    - em um vetor, 151
- polinômio
  - divisor, 149
  - irreduzível, 149
  - mínimo, 149
  - mônico, 149
  - primo, 149
- polinômios
  - primos entre si, 150
- posto
  - de transformação linear, 61
  - de uma matriz, 38
- produto
  - de matrizes, 3
  - de transformações lineares, 67
  - por escalar, 2, 65
  - produto escalar, 108, 110
  - produto interno, 110
    - canônico, 110
    - induzido, 128
  - projeção, 51
    - ortogonal, 117, 121
  - raiz quadrada, 178
  - regra de Cramer, 106
  - representação, 89
  - sistema linear, 6
    - homogêneo, 7
  - sistemas lineares
    - equivalentes, 8
  - solução de sistema linear, 6
  - solução trivial, 7
  - soma
    - de matrizes, 3
    - de subespaços, 42
    - de vetores, 20
    - direta, 43
  - subespaço
    - $T$ -cíclico, 187
    - gerado, 25
    - invariante, 137
    - nulo, 22
    - vetorial, 22
  - subespaços
    - triviais, 22
  - transformação linear, 50
    - adjunta, 165
  - transposição, 88

traço, 15  
  de um operador, 60

variedade afim, 72

vetor, 20  
  nulo, 20

vetor coluna, 8

vetor linha, 8

vetores  
  ortogonais, 116

volume, 126