

# UM CURSO DE ÁLGEBRA LINEAR

UFC  
Centro de Ciências  
Departamento de Matemática  
Plácido Andrade

July 27, 2003

em memória de  
Hugo Lourinho de Andrade

## PREFÁCIO

Está bastante disseminado nas Universidades brasileiras um segundo curso de Álgebra Linear, oferecido não apenas aos estudantes dos cursos de Matemática, mas muitas vezes também para os de Física, Economia, Engenharias e outros, seja na graduação ou pós-graduação.

Esse livro foi elaborado como um texto para aquele segundo curso. Embora não seja imprescindível é conveniente que o aluno já tenha feito um curso básico de Álgebra Linear ou que tenha conhecimentos de Geometria Analítica com tratamento vetorial. A apresentação feita em quatorze capítulos é auto-suficiente, exceto na parte que trata de polinômios, entretanto, a maioria absoluta das propriedades utilizadas sobre polinômios é do conhecimento dos estudantes desde o Ensino Médio.

O principal objetivo desse texto é estudar operadores lineares em espaços vetoriais de dimensão finita com ênfase nas representações matriciais. Os tópicos foram elaborados de modo que permitem adaptações a qualquer curso com carga horária entre 60h e 90h semestrais. Os capítulos 11, 12 e 13 devem ser omitidos pois são capítulos de referência. Dependendo do nível da turma e da carga horária semestral os capítulos 1,2 e 3 podem ser considerados capítulos de revisão.

Para maior clareza, os operadores normais são estudados separadamente, primeiro em espaços Euclidianos e posteriormente em espaços unitários. Ao longo do texto intercalamos dezenas de exercícios, fatos simples que serão utilizados no corpo da demonstração de uma proposição logo a seguir, ficando ao critério do expositor apresentá-los ou não. Como exercícios propostos segue uma lista de mais de 700 ítems.

Desejo agradecer aos colegas do Departamento de Matemática da Universidade Federal do Ceará pela receptividade com a qual fui recebido como seu novo membro. Em particular, registro o meu agradecimento ao Professor Celso Antônio Barbosa pela orientação datilográfica dada ao longo da elaboração desse texto.

Fortaleza, 26 de julho de 2003  
Plácido Francisco de Assis Andrade

# Índice

<b>1</b>	<b>Espaços vetoriais</b>	<b>1</b>
1.1	Espaços vetoriais . . . . .	1
1.2	Subespaços . . . . .	4
1.3	Conjunto de geradores . . . . .	7
1.4	Teorema da troca de Stainitz . . . . .	10
1.5	Dimensão . . . . .	13
1.6	Coordenadas de um vetor . . . . .	16
1.7	Soma direta . . . . .	18
<b>2</b>	<b>Transformações lineares</b>	<b>22</b>
2.1	Transformações lineares . . . . .	22
2.2	O espaço das transformações lineares . . . . .	26
2.3	Teorema do núcleo e da imagem . . . . .	30
2.4	Isomorfismos lineares . . . . .	33
2.5	Matriz de uma transformação linear . . . . .	35
2.6	Teorema da representação matricial . . . . .	37
2.7	Representação de isomorfismo . . . . .	41
2.8	Matriz mudança de coordenadas . . . . .	43
2.9	Subespaços invariantes . . . . .	46
<b>3</b>	<b>Espaço dual</b>	<b>48</b>
3.1	Funcionais lineares . . . . .	48
3.2	Base dual . . . . .	51
3.3	A adjunta . . . . .	53
3.4	Perpendicular de um conjunto . . . . .	54
3.5	Teorema do posto . . . . .	55
<b>4</b>	<b>Operadores e polinômios</b>	<b>58</b>
4.1	Polinômio minimal . . . . .	58
4.2	Polinômio característico . . . . .	61

4.3	Teorema de Cayley-Hamilton . . . . .	63
4.4	Sobre a fatoração do polinômio minimal . . . . .	66
4.5	Polinômio minimal de um vetor . . . . .	68
<b>5</b>	<b>Decomposição primária</b>	<b>70</b>
5.1	Teorema da decomposição primária . . . . .	70
5.2	Corolários . . . . .	74
5.3	Subespaços cíclicos . . . . .	79
5.4	Espaços cíclicos . . . . .	81
5.5	Sobre a decomposição cíclica . . . . .	82
5.6	Apêndice . . . . .	85
<b>6</b>	<b>Representação canônica (I)</b>	<b>88</b>
6.1	Autovalores e autovetores . . . . .	89
6.2	Operadores diagonalizáveis . . . . .	91
6.3	Operadores nilpotentes . . . . .	95
6.4	Decomposição cíclica II . . . . .	98
6.5	Representação de Jordan . . . . .	101
6.6	Exemplos . . . . .	102
6.7	Aplicação: decomposição $D + N$ . . . . .	105
<b>7</b>	<b>Espaços Euclidianos</b>	<b>109</b>
7.1	Produto interno . . . . .	109
7.2	Norma . . . . .	110
7.3	Ortogonalidade . . . . .	113
7.4	Espaços Euclidianos . . . . .	114
7.5	Representação de um funcional linear . . . . .	118
7.6	Operador transposto . . . . .	120
<b>8</b>	<b>Operadores normais</b>	<b>124</b>
8.1	Operadores normais . . . . .	124
8.2	Decomposição normal . . . . .	126
8.3	Operadores simétricos . . . . .	127
8.4	Decomposição cíclica normal . . . . .	130
8.5	Representação de operadores normais . . . . .	133
8.6	Operadores anti-simétricos . . . . .	136
8.7	Operadores ortogonais . . . . .	137
8.8	Espectro . . . . .	140
8.9	Operadores que preservam a estrutura . . . . .	142
8.10	Comparando estruturas Euclidianas . . . . .	144
8.11	Decomposição polar . . . . .	147

8.12	Isometrias . . . . .	149
<b>9</b>	<b>Espaços unitários</b>	<b>152</b>
9.1	Espaços unitários . . . . .	152
9.2	O operador adjunto . . . . .	154
9.3	Operadores normais . . . . .	157
9.4	Operadores autoadjuntos . . . . .	158
9.5	Operadores unitários . . . . .	159
9.6	Decomposição polar . . . . .	160
<b>10</b>	<b>Formas bilineares</b>	<b>162</b>
10.1	Formas bilineares . . . . .	162
10.2	Representação matricial . . . . .	163
10.3	Propriedades da representação . . . . .	165
10.4	Formas bilineares simétricas . . . . .	167
10.5	Bases g-ortonormais . . . . .	170
10.6	Teorema do índice . . . . .	173
10.7	Formas bilineares antisimétricas . . . . .	175
<b>11</b>	<b>Representação canônica (II)</b>	<b>179</b>
11.1	Complexificação . . . . .	179
11.2	Espaço A- simples . . . . .	181
11.3	Decomposição cíclica III . . . . .	184
11.4	Decomposição cíclica IV . . . . .	186
<b>12</b>	<b>Matrizes</b>	<b>192</b>
12.1	Matrizes . . . . .	192
12.2	Matrizes quadradas . . . . .	195
12.3	Matrizes normais . . . . .	198
12.4	Apêndice: Grupos, anéis e corpos . . . . .	200
<b>13</b>	<b>Determinantes</b>	<b>202</b>
13.1	Determinante . . . . .	202
13.2	Definição . . . . .	204
13.3	Existência . . . . .	206
13.4	Unicidade . . . . .	207
13.5	Propriedades . . . . .	209
13.6	Adjunta clássica . . . . .	211
13.7	Regra de Cramer . . . . .	213



# Capítulo 1

## Espaços vetoriais

Estudaremos duas noções matemáticas: uma estrutura algébrica chamada espaço vetorial e uma outra denominada transformação linear, ou seja, funções entre espaços vetoriais que possuem propriedades específicas. Os dois primeiros capítulos são dedicados ao estudo desses conceitos. Iniciaremos definindo espaço vetorial, uma das estruturas mais úteis e que surge naturalmente nas várias áreas do conhecimento, merecendo por isto uma sistematização mais extensa.

### 1.1 Espaços vetoriais

Um exemplo muito conhecido de espaço vetorial é  $\mathbb{R}^2$ , o conjunto dos pares ordenados  $(x_1, x_2)$  com  $x_i \in \mathbb{R}$ , equipado com uma operação de adição de dois elementos e uma operação de multiplicação de um elemento por um número real  $\lambda$ ,

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

A definição de espaço vetorial relaciona as propriedades algébricas dessas operações, propriedades que ocorrem não apenas nesse exemplo mas também em inúmeros outros, como veremos.

**Definição 1.1.1** *Um espaço vetorial consiste de*

1. *Um conjunto  $V$  cujos elementos são chamados de vetores;*
2. *Um corpo  $\mathbb{K}$  cujos elementos são chamados de escalares;*
3. *Uma operação chamada de adição de vetores na qual cada par de vetores  $u, v \in V$  é associado ao vetor  $u + v \in V$ , chamado de soma de  $u$  e  $v$ , satisfazendo aos seguintes axiomas:*
  - a) *a adição é comutativa,  $u + v = v + u$ ;*

- b) a adição é associativa,  $(u + v) + w = u + (v + w)$ ;
- c) existe um único vetor  $o$  tal que  $v + o = v$  para todo  $v \in V$ ;
- d) para cada  $v \in V$  existe um único vetor  $-v \in V$  tal que  $v + (-v) = o$ ;
4. Uma operação chamada de multiplicação por escalar em que um vetor  $v \in V$  e um escalar  $\lambda \in \mathbb{K}$  são associados ao vetor  $\lambda v \in V$ , chamado de produto de  $v$  por  $\lambda$ , satisfazendo aos seguintes axiomas:
- a)  $1v = v$  para todo  $v \in V$ ;
- b) a multiplicação por escalar é associativa,  $\lambda_1(\lambda_2 v) = (\lambda_1 \lambda_2)v$ ;
- c) a multiplicação por escalar é distributiva em relação à adição de vetores,  $\lambda(u + v) = \lambda u + \lambda v$ ;
- d) a multiplicação por escalar é distributiva em relação à adição de escalares,  $(\lambda_1 + \lambda_2)v = \lambda_1 v + \lambda_2 v$ ;

Em resumo, um conjunto  $V$  é um *espaço vetorial sobre o corpo*  $\mathbb{K}$  se satisfaz as condições da definição acima. Quando não for necessário particularizar indicaremos por  $\mathbb{K}$  tanto o corpo dos números reais  $\mathbb{R}$  quanto o corpo dos números complexos  $\mathbb{C}$  e apenas estes dois corpos serão considerados. Se  $\mathbb{K} = \mathbb{R}$  diremos que  $V$  é um *espaço vetorial real* e quando  $\mathbb{K} = \mathbb{C}$  chamaremos  $V$  de *espaço vetorial complexo*. Inúmeras vezes omitiremos o corpo com o qual estamos trabalhando mas estará claro que o comentário é verdadeiro independente do corpo considerado. Um espaço vetorial  $V$  não é vazio pois contém pelo menos o vetor  $o$  chamado de *vetor nulo*. Se  $V$  consiste de um único elemento será chamado de *espaço trivial* e pela definição segue que  $V = \{o\}$ .

**Exemplo 1.1.1** 1) Dado um número inteiro  $n \geq 1$  denotaremos por  $\mathbb{K}^n$  o conjunto de todas as  $n$ -uplas, isto é, o conjunto de todas as seqüências de  $n$  termos  $u = (x_1, x_2, \dots, x_n)$  com  $x_i \in \mathbb{K}$ . Induzimos em  $\mathbb{K}^n$  uma estrutura de espaço vetorial sobre o corpo  $\mathbb{K}$  do seguinte modo: se  $u = (x_1, x_2, \dots, x_n)$ ,  $v = (y_1, y_2, \dots, y_n) \in \mathbb{K}^n$  e  $\lambda \in \mathbb{K}$ , definimos a adição de vetores e a multiplicação de um vetor por um escalar, respectivamente por

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

Não é difícil verificar que as duas operações induzem em  $\mathbb{K}^n$  uma estrutura de espaço vetorial sobre  $\mathbb{K}$ . Neste espaço o vetor nulo é o elemento  $o = (0, 0, \dots, 0)$  e  $-u = (-x_1, -x_2, \dots, -x_n)$ . Finalmente observamos que o corpo  $\mathbb{K}$  é naturalmente identificado com o espaço vetorial  $\mathbb{K}^1$ . Não distinguiremos uma 1-úpla ordenada  $(x_1)$  de um número  $x_1$ .

2) Seja  $\mathbb{K}[t]$  o conjunto de todos os polinômios na variável  $t$  com coeficientes no corpo  $\mathbb{K}$ . As usuais adição de polinômios e multiplicação de um polinômio por um escalar induzem em  $\mathbb{K}[t]$  uma estrutura de espaço vetorial sobre  $\mathbb{K}$ .

3) O primeiro exemplo pode ser generalizado. O *produto direto*, ou *produto Cartesiano*, de espaços vetoriais  $V_1, V_2, \dots, V_n$  sobre o mesmo corpo  $\mathbb{K}$  é o conjunto denotado por

$$V = V_1 \times V_2 \times \cdots \times V_n$$

e constituído por todas as  $n$ -uplas ordenadas  $u = (u_1, u_2, \dots, u_n)$  com  $u_i \in V_i$ . Do mesmo modo, induz-se uma estrutura de espaço vetorial sobre o corpo  $\mathbb{K}$ : dados os elementos  $u = (u_1, u_2, \dots, u_n)$ ,  $v = (v_1, v_2, \dots, v_n) \in V$  e um escalar  $\lambda \in \mathbb{K}$ , definimos a adição de dois vetores e a multiplicação de um vetor por um escalar, respectivamente, por

$$\begin{aligned} u + v &= (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n), \\ \lambda u &= (\lambda u_1, \lambda u_2, \dots, \lambda u_n). \end{aligned}$$

4) Sejam  $\Omega$  um conjunto não vazio e  $V$  um espaço vetorial sobre o corpo  $\mathbb{K}$ . O conjunto  $C(\Omega, V)$  formado por todas as funções  $f : \Omega \rightarrow V$  adquire uma estrutura de espaço vetorial sobre  $\mathbb{K}$  considerando as seguintes operações: para  $f, g \in C(\Omega, V)$  e  $\lambda \in \mathbb{K}$  define-se a função soma  $f + g$  e a função  $f$  multiplicada pelo escalar  $\lambda f$ , respectivamente, por

$$\begin{aligned} (f + g)(\omega) &= f(\omega) + g(\omega), \\ (\lambda f)(\omega) &= \lambda f(\omega), \end{aligned}$$

em que  $\omega \in \Omega$ . O vetor nulo é a função identicamente nula,  $f(\omega) = 0$  para todo  $\omega \in \Omega$ , e  $-f \equiv (-1)f$ . Quando escrevermos "as usuais operações de soma de funções e multiplicação de uma função por um escalar" estaremos nos referindo às duas operações aqui definidas.

5) Uma *matriz*  $m \times n$  com entradas em  $\mathbb{K}$  é uma sequência de escalares  $N = [a_{ij}]$  com  $a_{ij} \in \mathbb{K}$  organizada em  $m$  linhas e  $n$  colunas,

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

O primeiro índice de  $a_{ij}$  indica a linha na qual a entrada encontra-se e o segundo índice indica a coluna. Induzimos uma estrutura de espaço vetorial no conjunto das matrizes  $m \times n$  com entradas em  $\mathbb{K}$ , conjunto esse denotado por  $M(m \times n, \mathbb{K})$ , definindo a adição de matrizes e a multiplicação de uma matriz por um escalar, respectivamente, por

$$\begin{aligned} N + P &= [a_{ij} + b_{ij}], \\ \lambda N &= [\lambda a_{ij}], \end{aligned}$$

em que  $N = [a_{ij}]$ ,  $P = [b_{ij}] \in M(m \times n, \mathbb{K})$  e  $\lambda \in \mathbb{K}$ . O vetor nulo do espaço é a matriz identicamente nula (todas as entradas são iguais a zero) e  $-N = [-a_{ij}]$ . Caso o leitor deseje maiores detalhes indicamos o capítulo sobre matrizes.  $\square$

### Exercícios propostos 1.1.1

1. Demonstre as seguintes afirmações sobre um espaço vetorial  $V$ .
  - a)  $\lambda o = o$  para todo escalar  $\lambda \in \mathbb{K}$ .
  - b)  $\lambda v = o$  para algum escalar  $0 \neq \lambda \in \mathbb{K} \Leftrightarrow v = 0$ .
  - c)  $-v = (-1)v$  para todo vetor  $v \in V$ .
  - d)  $nv = \underbrace{v + v + \cdots + v}_{n \text{ vezes}}$  onde  $n$  é um inteiro.
  
2. Procure num livro de Cálculo os teoremas que garantem a existência de uma estrutura de espaço vetorial real nos seguintes conjuntos equipados com as operações de soma de funções e multiplicação de uma função por um escalar.
  - a) O conjunto  $C^0([a, b], \mathbb{R})$  formado por todas as funções contínuas  $f : [a, b] \rightarrow \mathbb{R}$ , em que  $[a, b] \subset \mathbb{R}$  é um intervalo.
  - b) O conjunto  $\mathfrak{R}([a, b], \mathbb{R})$  constituído por todas as funções  $f : [a, b] \rightarrow \mathbb{R}$  que são Riemann integráveis.
  - c) O conjunto  $\mathcal{S}(\mathbb{Z}, \mathbb{R})$  de todas sequências reais  $(a_1, a_2, a_3, \dots)$  cuja série  $\sum a_n$  converge (convergência simples).
  - d) O conjunto  $\ell^1(\mathbb{R})$  de todas as sequências reais  $(a_1, a_2, a_3, \dots)$  cuja série  $\sum |a_n|$  converge (convergência absoluta).
  
3. Seja  $V$  um espaço vetorial complexo. Podemos definir uma outra estrutura de espaço vetorial complexo em  $V$  da seguinte forma. A soma de vetores continua a mesma e a multiplicação de um vetor por um escalar  $\lambda \in \mathbb{C}$  é dada pela regra  $\lambda * v = \bar{\lambda}v$ . Mostre que de fato as operações induzem uma nova estrutura de espaço vetorial complexo sobre o mesmo conjunto  $V$ , chamada de *espaço vetorial conjugado*.

## 1.2 Subespaços

Destacamos um tipo especial de subconjunto de um espaço vetorial  $V$  sobre um corpo  $\mathbb{K}$ . Um subconjunto  $W \subset V$  é um *subespaço vetorial* se  $W$  é não vazio e um espaço vetorial sobre  $\mathbb{K}$  com respeito às operações de adição de vetores e multiplicação de um vetor por um escalar, ambas operações induzidas de  $V$ . Por comodidade, diremos apenas que  $W$  é um *subespaço* de  $V$ .

**Exemplo 1.2.1** 1) O subconjunto  $W \subset \mathbb{K}^n$  formado por todos os vetores  $v = (x_1, x_2, \dots, x_n)$  tais que  $x_n = 0$  é um subespaço.

2) O subconjunto  $\mathbb{K}_n[t] \subset \mathbb{K}[t]$  formado pelos polinômios  $p(t)$  com *grau*  $p(t) \leq n$  é um subespaço. Por convenção, o polinômio identicamente nulo tem grau  $-\infty$  e um polinômio constante não nulo tem grau zero.

3) Fixado um vetor  $v$  de um espaço vetorial  $V$ , o conjunto  $W_v = \{\lambda v; \lambda \in \mathbb{K}\}$  é um subespaço. Chamaremos de subespaço trivial ao subconjunto formado apenas pelo vetor nulo.

4) Seja  $C(\Omega, V)$  espaço de todas as funções de um conjunto  $\Omega$  no espaço vetorial  $V$ . Fixado  $\omega_0 \in \Omega$ , o subconjunto  $C_{\omega_0}(\Omega, V) = \{f : \Omega \rightarrow V; f(\omega_0) = 0\}$  é um subespaço.  $\square$

Apresentaremos a seguir um critério de fácil aplicação para determinarmos quando um subconjunto  $W$  é um subespaço de um espaço vetorial  $V$ .

**Proposição 1.2.1** *Um subconjunto não vazio  $W \subset V$  é um subespaço se, e somente se, para cada par de vetores  $u, v \in W$  e para cada escalar  $\lambda \in \mathbb{K}$ , o vetor  $u + \lambda v \in W$ .*

**Demonstração** Assuma que  $W$  é um subespaço de  $V$ . É claro que se  $u, v$  são vetores de  $W$  e  $\lambda \in \mathbb{K}$ , então  $u + \lambda v \in W$ . Vejamos agora a suficiência. Mostremos inicialmente que o vetor nulo pertence a  $W$ . Para isso, escolhamos qualquer vetor  $v$  em  $W$  e escrevamos o vetor nulo como  $o = (-1)v + v$ . Da hipótese segue que  $o \in W$ . Por outro lado,  $\lambda v + o = \lambda v$ , para qualquer escalar  $\lambda \in \mathbb{K}$ , de onde concluímos que  $W$  é fechado em relação à multiplicação de um vetor por um escalar. Logo, o subconjunto também contém todos os vetores  $-v$  se  $v \in W$ . Finalmente, se  $u$  e  $v$  são dois elementos de  $W$ , então  $u + v = u + 1v \in W$ , mostrando que  $W$  é fechado em relação à operação de adição de vetores. Os outros axiomas são válidos pois já são válidos em  $V$ .  $\square$

**Exercício 1.2.1** Prove: a interseção de dois subespaços de  $V$  é um subespaço.  $\square$

Dentre os vários métodos para construímos subespaços vejamos um que terá desdobramentos nas próximas seções. Para isso, precisaremos da definição de combinação linear, um conceito básico para o nosso estudo.

**Definição 1.2.1** *Seja  $V$  um espaço vetorial sobre o corpo  $\mathbb{K}$ . Diremos que o vetor  $u \in V$  é uma combinação linear dos vetores  $v_1, v_2, \dots, v_n \in V$  se existem escalares  $x_1, x_2, \dots, x_n \in \mathbb{K}$  tais que  $u = x_1v_1 + x_2v_2 + \dots + x_nv_n$ .*

Note que uma combinação linear envolve apenas um número finito de vetores!

**Proposição 1.2.2** *Seja  $\varepsilon$  um subconjunto não vazio do espaço vetorial  $V$ . O subconjunto  $W \subset V$  formado por todos os vetores que são combinações lineares de vetores de  $\varepsilon$  é um subespaço. Além disso,  $W$  é o menor subespaço contendo  $\varepsilon$ .*

**Demonstração** Com efeito, se  $u, v \in W$ , por definição desse subespaço, existem vetores  $u_1, \dots, u_m, v_1, \dots, v_n \in \varepsilon$  e escalares  $x_1, \dots, x_m, y_1, \dots, y_n \in \mathbb{K}$  tais que  $u = x_1u_1 + \dots + x_mu_m$  e  $v = y_1v_1 + \dots + y_nv_n$ . Agora, se  $\lambda \in \mathbb{K}$ , então

$$u + \lambda v = x_1u_1 + x_2u_2 + \dots + x_mu_m + \lambda y_1v_1 + \lambda y_2v_2 + \dots + \lambda y_nv_n.$$

Desde que  $u + \lambda v$  é uma combinação linear de vetores de  $\varepsilon$  temos mostrado que  $u + \lambda v \in W$ . Pela proposição anterior segue que  $W$  é um subespaço de  $V$ . Por outro lado, se  $W_0$  é um outro subespaço que contém o conjunto  $\varepsilon$  é evidente que ele contém todas as combinações lineares de vetores de  $\varepsilon$ . Portanto  $W \subset W_0$ , concluindo que  $W$  é o menor subespaço contendo o conjunto  $\varepsilon$ .  $\square$

### Exercícios propostos 1.2.1

1. Prove que a interseção de uma família  $\{W_i\}_{i \in \mathcal{I}}$  de subespaços de  $V$  é um subespaço.
2. Sejam  $W_1$  e  $W_2$  subespaços de um espaço vetorial  $V$  sobre o corpo  $\mathbb{K}$ . Verifique quais dos conjuntos são subespaços de  $V$ .
  - a)  $W = W_1 \cup W_2$ .
  - b)  $W = W_1 \times W_2$ .
  - c)  $W = \{w_1 + w_2; w_i \in W_i\}$ .
3. Quais subconjuntos são subespaços do  $\mathbb{R}^3$ ?
  - a)  $W = \{(x, y, z) \in \mathbb{R}^3; x - y + z = 0\}$ .
  - b)  $W = \{(x, y, z) \in \mathbb{R}^3; x - y + z = 1\}$ .
  - c)  $W = \{(x, y, z) \in \mathbb{R}^3; x - y + z = 0 \text{ e } x + y + z = 0\}$ .
  - d)  $W = \{a(1, 1, 0) + b(0, 1, 1); a, b \in \mathbb{R}\}$ .
4. O vetor  $v = (1, 2, 3)$  pertence ao subespaço das combinações lineares do conjunto  $\varepsilon = \{(i, 1, 1), (1, 2 + 3i, 4)\} \subset \mathbb{C}^3$ ? E o vetor  $v = (2 + i, 5 + 3i, 9)$ ?
5. Sejam  $\varepsilon_1$  e  $\varepsilon_2$  dois conjuntos não vazios do espaço vetorial  $V$ .
  - a) Mostre que se  $\varepsilon_1 \subset \varepsilon_2$  então o espaço das combinações lineares de  $\varepsilon_1$  está contido no espaço das combinações lineares de  $\varepsilon_2$ .
  - b) Se os espaços das combinações lineares dos dois conjuntos são iguais, necessariamente os conjuntos são iguais?
6. Dados  $\varepsilon_1 = \{(1, 1, 0), (0, 1, 1)\}$  e  $\varepsilon_2 = \{(1, 2, 1), (1, 0, -1), (3, 4, 1)\}$ , subconjuntos do  $\mathbb{R}^3$ , mostre que os espaços das combinações lineares dos conjuntos são iguais.

7. Se  $W_1$  e  $W_2$  são subespaços de  $V$  tais que  $W_1 \cup W_2$  é um subespaço, mostre que um dos subespaços contém o outro.
8. Consulte algum livro de Cálculo e verifique quais são os teoremas que validam as seguintes afirmações.
- O espaço das funções contínuas,  $C^0([a, b], \mathbb{R})$ , é um subespaço do espaço das funções Riemann integráveis,  $\mathcal{R}([a, b], \mathbb{R})$ .
  - O espaço das sequências que convergem absolutamente  $\ell^1(\mathbb{R})$  é um subespaço do espaço das sequências que convergem simplesmente  $\mathcal{S}(\mathbb{Z}, \mathbb{R})$ .
9. Seja  $W$  um subespaço do espaço vetorial  $V$  sobre  $\mathbb{K}$ . Defina uma relação entre os vetores de  $V$  por:  $u \sim v \pmod{W} \Leftrightarrow u - v \in W$ .
- Verifique que  $\sim$  é uma *relação de equivalência* (simétrica, reflexiva e transitiva).
  - Denote por  $\bar{u}$  a classe de equivalência que contém  $u \in V$  e por  $V/W$  o conjunto formado por todas as classes. Defina as operações em  $V/W$ :

$$\bar{u} + \bar{v} = \overline{u + v} \quad \text{e} \quad \lambda \bar{u} = \overline{\lambda u},$$

onde  $\lambda$  é um escalar em  $\mathbb{K}$ . Mostre que as operações definem uma estrutura de espaço vetorial sobre  $\mathbb{K}$  no conjunto  $V/W$ . Com essa estrutura  $V/W$  será chamado de *espaço quociente*.

### 1.3 Conjunto de geradores

Tendo em vista a Proposição 2.5 da seção anterior na qual foi contruído o menor subespaço contendo todas as combinações lineares de um conjunto de vetores, estabeleceremos agora um conceito baseado naquela idéia.

Diremos que um conjunto de vetores  $\varepsilon \subset W$  é um *conjunto de geradores* de um subespaço  $W \subset V$ , se todo vetor de  $W$  é uma combinação linear de vetores de  $\varepsilon$ . Em outras palavras, dado um vetor  $w \in W$  existem vetores  $v_1, v_2, \dots, v_n \in \varepsilon$  e escalares  $x_1, x_2, \dots, x_n \in \mathbb{K}$  tais que  $w = x_1 v_1 + x_2 v_2 + \dots + x_n v_n$ . Por simplicidade, diremos que  $W$  é gerado por  $\varepsilon$  ou que os vetores de  $\varepsilon$  geram  $W$ .

Questionar sobre a existência de um conjunto de geradores para um espaço vetorial não tem consequências das mais proveitosas pois a resposta é sim, basta considerar  $\varepsilon = V$ . Uma questão mais relevante é examinar a existência de um conjunto de geradores possuindo determinadas propriedades, como ser finito ou não, enumerável, ortogonal, etc. É nesta direção que caminharemos.

**Definição 1.3.1** *Seja  $V$  um espaço vetorial sobre o corpo  $\mathbb{K}$ . Diremos que o conjunto  $\varepsilon \subset V$  é linearmente dependente se existem distintos vetores  $v_1, v_2, \dots, v_n \in \varepsilon$  e existem escalares não todos nulos  $x_1, x_2, \dots, x_n \in \mathbb{K}$  tais que*

$$o = x_1v_1 + x_2v_2 + \cdots + x_nv_n.$$

Um subconjunto  $\delta \subset V$  que não é linearmente dependente será dito *linearmente independente*. A última definição merece alguns comentários, pois o conceito nela contido é capital para a compreensão da estrutura de espaço vetorial.

1. O conjunto vazio  $\delta = \{ \}$  é linearmente independente por vacuidade. Não podemos exibir vetores distintos de  $\delta$ .
2. Qualquer conjunto  $\varepsilon \subset V$  contendo o vetor nulo é linearmente dependente pois é possível formar a combinação linear  $o = x_1o$  com um escalar não nulo  $x_1 \in \mathbb{K}$ .
3. O conjunto  $\delta \subset V$  é linearmente independente se para toda coleção de distintos vetores  $v_1, v_2, \dots, v_n \in \delta$  a combinação linear

$$o = x_1v_1 + x_2v_2 + \cdots + x_nv_n$$

implica que  $x_1 = x_2 = \cdots = x_n = 0$ . Esta afirmação será bastante utilizada!

4. Um subconjunto de um conjunto linearmente independente é linearmente independente.
5. Se  $\varepsilon \subset V$  admite um subconjunto linearmente dependente então  $\varepsilon$  é linearmente dependente.

Por abuso de linguagem é comum referir-se aos elementos de um conjunto  $\varepsilon$  dizendo que os seus vetores são linearmente independentes (ou linearmente dependentes se for o caso) em lugar de aplicar o conceito ao conjunto.

**Exemplo 1.3.1** O subconjunto  $\delta = \{u, v, w\} \subset \mathbb{K}^3$  em que  $u = (1, 0, 0)$ ,  $v = (1, 1, 0)$  e  $w = (1, 1, 1)$ , é linearmente independente pois a combinação linear

$$\begin{aligned} (0, 0, 0) &= x_1u + x_2v + x_3w \\ &= (x_1 + x_2 + x_3, x_2 + x_3, x_3), \end{aligned}$$

implica que  $0 = x_1 = x_2 = x_3$ . □

**Exercício 1.3.1** Demonstre as afirmações sobre um espaço vetorial  $V$ .

1. Sejam  $W \subset V$  um subespaço e  $\delta \subset W$  um conjunto linearmente independente. Se  $v_0 \notin W$  então o conjunto  $\delta \cup \{v_0\}$  é linearmente independente.

2. Seja  $\varepsilon \subset V$  um conjunto de geradores de  $V$ . Assuma que  $v_0 \in \varepsilon$  é uma combinação linear de outros vetores de  $\varepsilon$ . Então  $\varepsilon$  é linearmente dependente e o conjunto  $\varepsilon_1$  obtido de  $\varepsilon$  por supressão de  $v_0$  ainda é um conjunto de geradores de  $V$ .  $\square$

Finalizaremos esta seção com a definição de base de um espaço vetorial.

**Definição 1.3.2** *Uma base  $\beta$  de um espaço vetorial  $V$  é um conjunto de geradores linearmente independente.*

**Exemplo 1.3.2** 1) Chamaremos de *base canônica de  $\mathbb{K}^n$*  ao conjunto de vetores  $\alpha = \{e_1, e_2, \dots, e_n\}$  em que

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

De fato,  $\alpha$  é uma base. A combinação linear

$$\begin{aligned} (0, 0, 0) &= x_1 e_1 + x_2 e_2 + \dots + x_n e_n \\ &= (x_1, x_2, \dots, x_n), \end{aligned}$$

implica que  $x_1 = x_2 = \dots = x_n = 0$ , mostrando a independência linear de  $\alpha$ . Por outro lado, dado o vetor  $v = (x_1, x_2, \dots, x_n) \in \mathbb{K}^n$  podemos expressá-lo como uma combinação linear dos vetores da base na forma  $v = x_1 e_1 + x_2 e_2 + \dots + x_n e_n$ . Isto significa que  $\alpha$  é um conjunto de geradores.

2) O conjunto de polinômios  $\beta = \{t^n\}_{n \geq 0}$  é uma base para o espaço vetorial dos polinômios  $\mathbb{K}[t]$ .

3) Um espaço trivial tem conjunto de geradores mas não admite uma base.  $\square$

### Exercícios propostos 1.3.1

- Quais subconjuntos são linearmente independentes?
  - $\varepsilon = \{(1, 0, 1), (0, 1, 1)\} \subset \mathbb{R}^3$ .
  - $\varepsilon = \{(1, 1, 1), (i, i, i)\} \subset \mathbb{C}^3$ .
  - $\varepsilon = \{(1, 0, 1)\} \subset \mathbb{R}^3$ .
  - $\varepsilon = \{(1, 0, 1), (0, 1, 1), (1, 1, 2)\} \subset \mathbb{R}^3$ .
- Seja  $V$  um espaço vetorial sobre o corpo  $\mathbb{K}$ . Demonstre as afirmações.
  - Se  $u, v \in V$  são linearmente dependentes então um deles é múltiplo do outro.
  - Se  $\varepsilon$  é um conjunto de geradores de  $V$  e  $\lambda \in \mathbb{K}$  é um escalar não nulo, então o conjunto  $\lambda\varepsilon = \{\lambda v; v \in \varepsilon\}$  é também um conjunto de geradores.
  - Se  $W_1$  e  $W_2$  são subespaços de  $V$  tais que  $W_1 \cap W_2 = \{0\}$ . Então quaisquer dois vetores não nulos  $v_1 \in W_1$  e  $v_2 \in W_2$  são linearmente independentes.
  - Se  $\delta \subset V$  é linearmente independente, então  $\delta$  é uma base para o espaço de suas combinações lineares.

3. Denote por  $C^1(\mathbb{R}, \mathbb{R})$  o espaço vetorial de todas funções  $f : \mathbb{R} \rightarrow \mathbb{R}$  cujas derivadas são contínuas. Mostre que as funções  $g_1(t) = \cos t$  e  $g_2(t) = \sin t$  são linearmente independentes. Quais os resultados de Análise que garantem a existência de uma estrutura de espaço vetorial em  $C^1(\mathbb{R}, \mathbb{R})$ ?
4. Utilize desenvolvimento de Taylor para mostrar que  $\beta = \{(t-1)^n\}_{n \geq 0}$  é uma base do espaço dos polinômios  $\mathbb{K}[t]$ .

## 1.4 Teorema da troca de Stainitz

Neste ponto, a Teoria de espaços vetoriais bifurca-se em duas linhas de estudos: se o espaço admite ou não admite um conjunto finito de geradores. Nos restringiremos ao exame do primeiro caso, isto é, aos chamados *espaços vetoriais de dimensão finita*. Espaços vetoriais de dimensão infinita merecem um estudo em separado pois, em geral, os espaços mais interessantes são aqueles que também estão equipados com uma estrutura topológica. Para iniciarmos o estudo é conveniente introduzir terminologia apropriada.

Uma sequência finita num espaço vetorial  $V$  é uma aplicação  $s : \{1, 2, \dots, n\} \rightarrow V$ . Como é usual, escreveremos  $v_i$  para designar o termo  $s(i)$  e indicaremos a sequência por  $s = (v_1, v_2, \dots, v_n)$ . Não devemos confundir a sequência com o conjunto formado pelos vetores na imagem de  $s$ . Observe que quando uma sequência de  $n$  vetores tem repetições o conjunto imagem da sequência possui um número menor de vetores.

Seja  $\varepsilon \subset V$  um conjunto finito com  $n$  elementos,  $n > 0$ . *Ordenar* o conjunto  $\varepsilon$  é escolher uma função injetiva (e sobrejetiva)  $s : \{1, 2, \dots, n\} \rightarrow \varepsilon$ . Nesse caso o conjunto imagem de  $s$  tem  $n$  elementos. Feito isto, indicamos o conjunto por  $\varepsilon = \{v_1, v_2, \dots, v_n\}$  chamando-o de *conjunto ordenado*.

Seja  $\{\varepsilon_i\}_{i=1}^k$  uma família de subconjuntos finitos de  $V$  dois a dois disjuntos. Uma *união ordenada* desses conjuntos é um conjunto ordenado construído do seguinte modo: primeiro ordenamos os elementos de  $\varepsilon_1$ , em seguida ordenamos os elementos de  $\varepsilon_1 \cup \varepsilon_2$ , respeitando-se a ordem de  $\varepsilon_1$ , depois ordenamos  $\varepsilon_1 \cup \varepsilon_2 \cup \varepsilon_3$  respeitando-se a ordem de  $\varepsilon_1 \cup \varepsilon_2$ , etc.. Denotaremos uma união ordenada por  $\varepsilon = \overrightarrow{\bigcup} \varepsilon_i$ .

Quando os conjuntos já estão ordenados a união ordenada é construída com os mesmos procedimentos e respeitando-se a ordem de cada  $\varepsilon_i$ . Demonstraremos uma propriedade de um conjunto linearmente dependente que será utilizada no estudo de geradores de um espaço vetorial.

**Lema 1.4.1** *Suponha que  $\varepsilon = \{v_1, v_2, \dots, v_n\}$  é um conjunto ordenado de vetores não nulos do espaço vetorial  $V$ . As seguintes afirmações são equivalentes.*

a) O conjunto  $\varepsilon$  é linearmente dependente.

b) Existe um vetor  $v_i$  que é uma combinação linear dos vetores  $v_1, v_2, \dots, v_{i-1}$ .

**Demonstração** a)  $\Rightarrow$  b) Como os vetores  $v_1, v_2, \dots, v_n$  são não nulos e linearmente dependentes, existe um menor inteiro  $i$ ,  $2 \leq i \leq n$ , tal que  $v_1, v_2, \dots, v_i$  são linearmente dependentes. Considere escalares não todos nulos  $x_1, x_2, \dots, x_i \in \mathbb{K}$  satisfazendo a equação  $o = x_1v_1 + x_2v_2 + \dots + x_iv_i$ . Necessariamente teremos o coeficiente  $x_i \neq 0$ , caso contrário os  $i - 1$  primeiros elementos seriam linearmente dependentes contrariando a definição de  $i$ . Sendo assim, temos

$$v_i = \frac{-x_1}{x_i}v_1 + \dots + \frac{-x_{i-1}}{x_i}v_{i-1},$$

concluindo a demonstração da implicação.

b)  $\Rightarrow$  a). Se algum vetor  $v_i$  é uma combinação linear dos  $i - 1$  primeiros vetores, isto é,  $v_i = x_1v_1 + \dots + x_{i-1}v_{i-1}$ , obrigatoriamente existe algum coeficiente  $x_j \neq 0$  com  $1 \leq j \leq i - 1$ , caso contrário  $v_i = o$ , contrariando a hipótese de  $\varepsilon$  não conter o vetor nulo. Isto implica que o subconjunto  $\{v_1, \dots, v_i\} \subset \varepsilon$  é linearmente dependente e por conseguinte concluímos que  $\varepsilon$  é linearmente dependente.  $\square$

O próximo resultado sobre espaços vetoriais admitindo um conjunto finito de geradores compara numérica e conceitualmente as noções de conjunto linearmente independente, base e conjunto de geradores. Utilizaremos o símbolo  $\sharp$  para indicar o número cardinal de um conjunto.

**Teorema 1.4.1 (Teorema da troca de Steinitz)** *Seja  $V$  um espaço vetorial não trivial sobre o corpo  $\mathbb{K}$ . Suponha que  $\delta \subset V$  é um conjunto linearmente independente e que  $\varepsilon \subset V$  é um conjunto finito de geradores. Então existe um subconjunto  $\gamma \subset \varepsilon$  tal que a união  $\beta = \delta \cup \gamma$  é uma base de  $V$ . Além disto valem as desigualdades  $\sharp\delta \leq \sharp\beta \leq \sharp\varepsilon$ .*

**Demonstração** A demonstração será feita em duas etapas. Sem perda de generalidade podemos assumir que o conjunto de geradores  $\varepsilon$  não contém o vetor nulo pois se o teorema é verdadeiro para o conjunto de geradores  $\varepsilon_0 \subset \varepsilon$  obtido por supressão do vetor nulo, ele é também verdadeiro para  $\varepsilon$ .

Antes de tudo ordenemos o conjunto de geradores  $\varepsilon = \{v_1, v_2, \dots, v_n\}$ . Por clareza vamos supor que  $\delta$  contenha pelo menos um elemento, caso ele seja vazio iremos diretamente à Segunda etapa da demonstração.

*Primeira etapa.* Escolhido um elemento  $u_1 \in \delta$  considere a união ordenada

$$\varepsilon_1 = \{u_1\} \vec{\cup} \{v_1, v_2, \dots, v_n\}.$$

O conjunto  $\varepsilon_1$  é um conjunto de geradores pois contém  $\varepsilon$  e é também linearmente dependente desde que  $u_1$  é uma combinação linear de outros vetores com coeficientes não todos nulos. Seja  $w$  o primeiro vetor de  $\varepsilon_1$  que é uma combinação linear de seus antecessores. É evidente que  $w$  não pode ser  $u_1$ , logo  $w = v_{j_1}$  para algum  $j_1$ . O conjunto

$$\delta_1 = \{u_1\} \vec{\cup} \{v_1, v_2, \dots, \widehat{v}_{j_1}, \dots, v_n\}$$

é um conjunto de geradores (o sinal  $\widehat{\phantom{x}}$  indica que o vetor foi suprimido). Verifiquemos esta afirmação. Um vetor  $v \in V$  é uma combinação linear de  $v_1, v_2, \dots, v_n$ . Como  $v_{j_1}$  é uma combinação de  $u_1, v_1, v_2, \dots, v_{j_1-1}$ , a substituição de  $v_{j_1}$  na combinação linear anterior por esta última expressa  $v$  como uma combinação linear dos vetores de  $\delta_1$ , mostrando a afirmação.

Suponhamos que exista um outro elemento  $u_2 \in \delta$ . Considere o conjunto ordenado

$$\varepsilon_2 = \{u_1, u_2\} \vec{\cup} \{v_1, v_2, \dots, \widehat{v}_{j_1}, \dots, v_n\}.$$

Por um lado, como  $\varepsilon_2$  contém o conjunto de geradores  $\delta_1$  ele é também um conjunto de geradores e, por outro lado,  $\varepsilon_2$  é linearmente dependente pois  $u_2$  é uma combinação linear dos outros vetores do conjunto. Novamente, seja  $w$  o primeiro vetor de  $\varepsilon_2$  que é uma combinação linear dos seus antecessores. Como  $\{u_1, u_2\}$  é linearmente independente podemos afirmar que  $w = v_{j_2}$ . Para aliviar a notação vamos assumir que  $j_1 < j_2$  embora esta exigência não tenha importância alguma na demonstração. Examinemos o conjunto ordenado

$$\delta_2 = \{u_1, u_2\} \vec{\cup} \{v_1, v_2, \dots, \widehat{v}_{j_1}, \dots, \widehat{v}_{j_2}, \dots, v_n\}.$$

Pela repetição do argumento anterior sobre combinações lineares, mostramos que  $\delta_2$  é um conjunto de geradores.

Note que o processo garante que os vetores de  $\varepsilon$  não serão suprimidos antes que todos os vetores de  $\delta$  estejam na lista, pois supor que  $\#\delta > \#\varepsilon = n$  implica num absurdo, isto é, implica que existe um vetor  $u_{n+1} \in \delta$  que é uma combinação linear dos vetores de  $\delta_n = \{u_1, u_2, \dots, u_n\}$  contrariando a hipótese de independência linear de  $\delta$ . Portanto,  $m = \#\delta \leq \#\varepsilon$  e na  $m$ -ésima etapa temos

$$\delta_m = \{u_1, \dots, u_m\} \vec{\cup} \{v_1, \dots, \widehat{v}_{j_1}, \dots, \widehat{v}_{j_2}, \dots, \widehat{v}_{j_m}, \dots, v_n\}.$$

Por construção  $\delta_m$  é um conjunto de geradores e se for linearmente independente a demonstração está terminada, basta considerar

$$\beta = \delta \cup \gamma \quad \text{onde} \quad \gamma = \{v_1, \dots, \widehat{v}_{j_1}, \dots, \widehat{v}_{j_2}, \dots, \widehat{v}_{j_m}, \dots, v_n\}.$$

*Segunda etapa.* Suponhamos que já acrescentamos todos os elementos de  $\delta$  e que  $\delta_m$  é linearmente dependente. Seja  $w$  o primeiro vetor de  $\delta_m$  que é uma combinação linear dos antecessores. Pelos mesmos argumentos utilizados concluímos que  $w =$

$v_{j_{m+1}}$ . Eliminando este vetor obtemos o conjunto ordenado

$$\delta_{m+1} = \{u_1, \dots, u_m\} \vec{\cup} \{v_1, \dots, \widehat{v}_{j_1}, \dots, \widehat{v}_{j_2}, \dots, \widehat{v}_{j_{m+1}}, \dots, v_n\}$$

que ainda é um conjunto de geradores. Continuando o processo iremos eliminando elementos de  $\varepsilon$  enquanto tivermos um conjunto linearmente dependente. Sendo o número de etapas finito em algum momento obteremos um conjunto  $\beta$  de geradores linearmente independente contendo todos os elementos de  $\delta$ . A menos de uma contagem de números de elementos dos conjuntos envolvidos a demonstração está completa.  $\square$

Fixaremos uma idéia que está implícita no Teorema da troca. Diz-se que um conjunto  $\delta \subset V$  pode ser *estendido a uma base de  $V$* , se existe uma base  $\beta \subset V$  contendo  $\delta$ .

## 1.5 Dimensão

Nesta seção estudaremos espaços vetoriais que admitem um conjunto finito de geradores, chamados mais apropriadamente de espaços vetoriais de dimensão finita. Como veremos na sequência, o Teorema da troca permite que associemos a cada um deles um inteiro positivo chamado de dimensão do espaço.

**Definição 1.5.1** *Um espaço vetorial  $V$  sobre o corpo  $\mathbb{K}$  é de dimensão finita se  $V$  admite uma base com um número finito de vetores.*

Um espaço vetorial trivial não admite uma base, portanto não satisfaz as condições exigidas na definição. Ressaltamos que num enunciado no qual está explicitado a expressão "dimensão finita" os espaços vetoriais triviais estarão excluídos das hipóteses. Feita esta ressalva passemos às consequências do Teorema da troca.

**Corolário 1.5.1** *Todas as bases de um espaço vetorial de dimensão finita  $V$  têm o mesmo número de vetores.*

**Demonstração** Por definição, existe pelo menos uma base finita  $\beta_1$  de  $V$ . Consideremos uma outra base  $\beta_2$ . Como  $\beta_1$  é um conjunto de geradores e  $\beta_2$  é linearmente independente, segue do Teorema da troca que  $\#\beta_2 \leq \#\beta_1$ , significando que  $\beta_2$  é finito. Com o mesmo argumento segue que  $\#\beta_1 \leq \#\beta_2$ .  $\square$

**Definição 1.5.2** *A dimensão de um espaço vetorial de dimensão finita  $V$  é o número de elementos de uma de suas bases.*

O Corolário anterior garante que a dimensão de  $V$  está bem definida. Indicaremos por  $\dim V$  a dimensão de um espaço vetorial de dimensão finita. Se  $V$  é trivial eventualmente será conveniente indicar o escalar zero por  $\dim V$ .

**Exemplo 1.5.1** 1) O produto cartesiano  $\mathbb{K}^n$  tem dimensão  $n$  desde que a base canônica  $\alpha = \{e_1, e_2, \dots, e_n\} \subset \mathbb{K}^n$  possui  $n$  elementos.

2) O espaço das matrizes  $m \times n$  com entradas no corpo  $\mathbb{K}$  tem dimensão  $mn$  pois o conjunto  $\alpha = \{E_1^1, \dots, E_i^j, \dots, E_m^n\} \subset M(m \times n, \mathbb{K})$  é uma base. Nesta notação  $E_i^j$  é a matriz cuja entrada  $ij$  é igual a 1 e todas as outras entradas são iguais a zero. Chamaremos  $\alpha$  de *base canônica de  $M(m \times n, \mathbb{K})$* .  $\square$

**Exercício 1.5.1** Utilize o Teorema da troca para demonstrar as seguintes afirmações sobre um subconjunto  $\varepsilon$  de um espaço vetorial  $V$  de dimensão finita.

- a) Se  $\#\varepsilon > \dim V$  então  $\varepsilon$  é um conjunto linearmente dependente.
- b) Se  $\#\varepsilon < \dim V$  então  $\varepsilon$  não é um conjunto de geradores.
- c) Se  $\#\varepsilon = \dim V$  e  $\varepsilon$  é linearmente independente então  $\varepsilon$  é uma base de  $V$ .
- d) Se  $\#\varepsilon = \dim V$  e  $\varepsilon$  é um conjunto de geradores então  $\varepsilon$  é uma base de  $V$ .  $\square$

Como outra aplicação do Teorema da troca examinaremos a relação entre a dimensão de  $V$  e a dimensão de seus subespaços.

**Corolário 1.5.2** *Seja  $W$  um subespaço não trivial de um espaço vetorial de dimensão finita  $V$ . Então  $W$  tem dimensão finita e  $\dim W \leq \dim V$ . Mais ainda,  $\dim W = \dim V$  se, e somente se,  $W = V$ .*

**Demonstração** Denote por  $\wp$  a classe formada por todos os subconjuntos de  $W$  que são linearmente independentes. Como  $W$  é um subespaço não trivial, é possível escolher um vetor não nulo  $v_0 \in W$  e formar o conjunto linearmente independente  $\varepsilon = \{v_0\} \subset W$ . Isso mostra que a classe  $\wp$  não é constituída apenas pelo conjunto vazio. Observe que um conjunto  $\varepsilon \subset W$  que é linearmente independente em  $W$  é linearmente independente em  $V$ . Pelo Teorema da troca,  $\varepsilon$  deve ter no máximo  $n$  elementos em que  $n = \dim V$ . Escolhido  $\beta_0 \in \wp$  um conjunto com o máximo número de vetores, mostremos que ele é uma base de  $W$ . Por definição de base, só precisamos verificar que  $\beta_0$  é um conjunto de geradores. Seja  $W_0$  o espaço das combinações lineares de  $\beta_0$ . É claro que  $W_0 \subset W$ . Por absurdo, vamos supor que  $W_0 \subsetneq W$ . Sendo assim, podemos escolher um vetor não nulo  $v_0 \in W$  com  $v_0 \notin W_0$

e contruir o conjunto linearmente independente  $\beta_0 \cup \{v_0\} \in \varnothing$ . Isso contraria a maximalidade de  $\beta_0$ , logo,  $\beta_0$  é uma base de  $W$  com no máximo  $n$  elementos, demonstrando a desigualdade  $\dim W \leq \dim V$ . Deixaremos aos cuidados do leitor mostrar que  $\dim W = \dim V$  se, e somente se,  $W = V$ .  $\square$

### Exercícios propostos 1.5.1

- Para cada item encontre uma base para o espaço das combinações lineares de  $\varepsilon$  e estenda-a a uma base do espaço
  - $\varepsilon = \{(1, 0, 1), (0, 1, 1)\} \subset \mathbb{R}^3$ .
  - $\varepsilon = \{(1, 0, 1), (2, 1, 3), (0, 1, 1)\} \subset \mathbb{R}^3$ .
  - $\varepsilon = \{(1, 1, 1), (i, i, i)\} \subset \mathbb{C}^3$ .
  - $\varepsilon = \{(1, 0, 1), (0, 1, 1), (1, 1, 2)\} \subset \mathbb{R}^3$ .
- Calcule as dimensões dos subespaços complexos e dê uma base para cada um.
  - $W = \{(x_1, x_2, x_2) \in \mathbb{C}^3; x_1 - x_2 = 0\}$ .
  - $W = \{(x_1, x_2, x_2) \in \mathbb{C}^3; x_1 - x_2 = 0, x_2 - x_3 = 0\}$ .
  - $W = \{a(1, 0, 1) + b(0, 1, 1) + c(1, -1, 0); a, b, c \in \mathbb{C}\}$ .
- Mostre as afirmações abaixo sobre um espaço vetorial  $V$  de dimensão finita.
  - Qualquer vetor não nulo de  $V$  é elemento de alguma base.
  - Se  $\varepsilon \subset V$  é um conjunto finito de geradores, então existe uma base  $\beta \subset \varepsilon$ .
  - Se  $\delta \subset V$  pode ser estendido a uma base então  $\delta$  é linearmente independente.
  - Se  $\delta \subset V$  é linearmente independente, então  $\delta$  está contido numa base.
- Definimos a *soma dos subespaços*  $V_1, V_2, \dots, V_k$  de um espaço vetorial  $V$  como sendo o subconjunto  $V_1 + V_2 + \dots + V_k = \{v_1 + v_2 + \dots + v_k; v_i \in V_i\}$ .
  - Demonstre que a soma de subespaços é um subespaço.
  - Utilize o Teorema da troca para provar que se  $V$  tem dimensão finita então  $\dim(V_1 + V_2 + \dots + V_k) \leq \dim V_1 + \dim V_2 + \dots + \dim V_k$ .
- Seja  $V$  um espaço vetorial sobre o corpo  $\mathbb{K}$ . Definimos a *diagonal do produto cartesiano*  $V \times V \times \dots \times V$  como sendo o conjunto  $\Delta = \{(v, v, \dots, v), v \in V\}$ .
  - Prove que a diagonal  $\Delta$  do produto cartesiano é um subespaço.
  - Se  $V$  tem dimensão  $n$ , qual a dimensão de  $\Delta$ ?
- Equipando um conjunto com uma estrutura de espaço vetorial, a dimensão do novo espaço depende do corpo considerado. Por exemplo, mostre que com as operações algébricas usuais  $\mathbb{C}^2$  tem dimensão dois sobre  $\mathbb{C}$  e tem dimensão quatro sobre  $\mathbb{R}$ .
- O corpo dos números reais é um espaço vetorial de dimensão um sobre ele mesmo! Quantos subespaços próprios de  $\mathbb{R}$  existem?

8. Todos os vetores de um subespaço próprio  $W \subsetneq R^2$  são múltiplos de um mesmo vetor. Demonstre este fato.
9. Sejam  $W$  um subespaço próprio não trivial de um espaço vetorial de dimensão finita  $V$  e  $\delta \subset W$  uma base de  $W$ . Prove que  $\delta$  pode ser estendido a uma base de  $V$  e que os vetores acrescentados não pertencem ao subespaço  $W$ .
10. O espaço dos polinômios  $\mathbb{K}[t]$  não tem dimensão finita! Caso esta afirmação seja verdadeira, demonstre-a.
11. Dois subespaços  $W_1$  e  $W_2$  de um espaço vetorial  $V$  são *transversais* quando  $V = W_1 + W_2$ . Quando  $V$  tem dimensão finita e  $W_1$  e  $W_2$  são transversais, vale a igualdade  $\dim W_1 + \dim W_2 = \dim V + \dim W_1 \cap W_2$ ?
12. Considere o espaço das matrizes quadradas  $M(n, \mathbb{K})$ . Responda as seguintes perguntas justificando as respostas.
  - a) Existe uma base de  $M(n, \mathbb{K})$  formada só por elementos não invertíveis?
  - b) Existe uma base de  $M(n, \mathbb{K})$  formada só por elementos invertíveis?
13. Prove que um espaço vetorial de dimensão finita  $V$  não é uma união finita de subespaços próprios.
14. Mostre que se  $V$  é um espaço vetorial de dimensão  $n$  e  $W \subset V$  é um subespaço de dimensão  $k$ , então o espaço quociente  $V/W$  tem dimensão  $n - k$ .

## 1.6 Coordenadas de um vetor

Dada uma base  $\beta$  de um espaço vetorial  $V$  de dimensão  $n$ , será sempre vantajoso ordená-la e escolhida uma ordem,  $\beta$  será chamada de base ordenada. Com isso, podemos falar em  $i$ -ésimo elemento da base ou em  $i$ -ésima *coordenada do vetor* relativo à  $\beta$ , terminologia que passaremos a explicar. Digamos que  $\beta = \{v_1, v_2, \dots, v_n\}$ . Como sabemos, dado um vetor  $v \in V$  existem escalares  $x_1, x_2, \dots, x_n \in \mathbb{K}$  tais que  $v = x_1v_1 + x_2v_2 + \dots + x_nv_n$ . Um fato importante é a unicidade dos coeficientes desta combinação linear. Suponha que expressemos o mesmo vetor como  $v = y_1v_1 + y_2v_2 + \dots + y_nv_n$ . Por subtração das duas combinações lineares obtemos

$$0 = (x_1 - y_1)v_1 + (x_2 - y_2)v_2 + \dots + (x_n - y_n)v_n.$$

A independência linear da base implica nas igualdades  $x_1 = y_1, x_2 = y_2, \dots, x_n = y_n$ . Devido a esta propriedade diremos que os únicos escalares  $x_1, x_2, \dots, x_n \in \mathbb{K}$  tais que  $v = x_1v_1 + x_2v_2 + \dots + x_nv_n$  são as coordenadas do vetor  $v$  na base ordenada  $\beta = \{v_1, v_2, \dots, v_n\}$ . Por motivos que serão explicitados nas próximas seções é conveniente organizar a sequência de escalares numa matriz, a saber,

$$[v]_{\beta} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

e referir-se à  $[v]_{\beta}$  como sendo a *matriz das coordenadas do vetor  $v$*  na base ordenada  $\beta$ . Observe que a matriz depende da base e da ordem da base.

**Exemplo 1.6.1** 1) O vetor  $v = (x_1, x_2, \dots, x_n) \in \mathbb{K}^n$  é escrito na base canônica  $\alpha = \{e_1, e_2, \dots, e_n\}$  como  $v = x_1e_1 + x_2e_2 + \dots + x_n e_n$ . Por definição, a matriz de  $v$  na base ordenada  $\alpha$  é a matriz  $n \times 1$

$$[v]_{\alpha} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

2) O conjunto ordenado  $\gamma = \{u_1, u_2\}$  em que  $u_1 = (1, 1)$  e  $u_2 = (-3, 2)$  é uma base ordenada do  $\mathbb{R}^2$ . Como o vetor  $v = (-2, 8)$  é expresso pela combinação linear  $v = 4(1, 1) + 2(-3, 2)$ , a matriz das coordenadas de  $v$  na base ordenada  $\gamma$  é

$$[v]_{\gamma} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}.$$

Se trocarmos a ordem da base  $\gamma$ , a matriz de  $v$  na nova base ordenada será uma matriz obtida por permutação das estradas da matriz  $[v]_{\gamma}$ .  $\square$

Tomamos conhecimento de vários espaços vetoriais que possuem uma base e no desdobramento do texto tomaremos conhecimento de outros exemplos. Portanto podemos desenvolver a teoria sem o risco de estarmos falando do vazio, trabalhando com objetos que não existem. Nossa situação é mais confortável, todo espaço vetorial não trivial admite uma base. A demonstração deste resultado pode ser vistas em S. Lang [10].

### Exercícios propostos 1.6.1

1. Dê as matrizes das coordenadas dos vetores  $v = (1 - 2, 2)$  e  $e_2 = (0, 1, 0) \in \mathbb{R}^3$ 
  - a) na base canônica  $\alpha \subset \mathbb{R}^3$ ;
  - b) na base  $\beta = \{(1, 0, 0), (1, 1, 0), (1, 1, 1)\} \subset \mathbb{R}^3$ ;
  - c) na base  $\gamma = \{(3, 4, -4), (8, 7, -8), (10, 10, -11)\} \subset \mathbb{R}^3$ .
2. Assuma que  $\varepsilon = \{v_1, v_2, \dots, v_n\} \subset V$  é um conjunto de geradores linearmente dependente que não contém o vetor nulo. Exiba o vetor nulo como duas combinações lineares distintas.

## 1.7 Soma direta

Definimos a *soma dos subespaços*  $V_1, V_2, \dots, V_k$  de um espaço vetorial  $V$  por

$$V_1 + V_2 + \dots + V_k = \{v_1 + v_2 + \dots + v_k; v_i \in V_i\}.$$

A soma de subespaços é um subespaço contendo cada  $V_i$  e no caso de cada um deles ser de dimensão finita, segue do Teorema da troca a desigualdade

$$\dim(V_1 + V_2 + \dots + V_k) \leq \dim V_1 + \dim V_2 + \dots + \dim V_k.$$

Diremos que a soma  $W = V_1 + V_2 + \dots + V_k$  é uma *soma direta* quando cada vetor  $v \in W$  é expresso de modo único como  $w = v_1 + v_2 + \dots + v_k$  com  $v_i \in V_i$ . Indicaremos uma soma direta por  $W = V_1 \oplus V_2 \oplus \dots \oplus V_k$ .

**Exemplo 1.7.1** 1) Para construir uma soma direta para um espaço vetorial de dimensão finita  $V$  na qual um dado subespaço  $V_1$  é uma das parcelas,  $V = V_1 \oplus V_2$ , é suficiente escolher uma base ordenada  $\beta_1 = \{v_1, \dots, v_k\}$  de  $V_1$ , estendê-la à uma base ordenada  $\beta = \{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$  de  $V$  e considerar o subespaço  $V_2$  gerado pelos  $n - k$  últimos vetores de  $\beta$ . Desde que um vetor  $v \in V$  é expresso por

$$v = (x_1 v_1 + \dots + x_k v_k) + (x_{k+1} v_{k+1} + \dots + x_n v_n),$$

concluimos que  $V = V_1 + V_2$ . A soma direta segue da unicidade da expressão.

2) Outras decomposições em soma direta podem ser construídas a partir da base ordenada  $\beta \subset V$ . Por exemplo,  $V = V_1 \oplus \dots \oplus V_n$ , em que  $V_i$  é o subespaço unidimensional gerado pelo vetor  $v_i \in \beta$ .

3) Observamos que  $\mathbb{R}^3 = V_1 + V_2$ , em que  $V_1 = \{(x_1, x_2, x_3); x_2 = 0\}$  e  $V_2 = \{(x_1, x_2, x_3); x_3 = 0\}$  pois qualquer vetor  $(x_1, x_2, x_3) \in \mathbb{R}^3$  é uma combinação linear da forma  $(x_1, x_2, x_3) = (x_1, 0, x_3) + (0, x_2, 0)$  com o primeiro vetor em  $V_1$  e o segundo vetor em  $V_2$ . Entretanto, a soma não é direta pois o mesmo vetor também é expresso por  $(x_1, x_2, x_3) = (0, 0, x_3) + (x_1, x_2, 0)$  com o primeiro vetor em  $V_1$  e o segundo vetor em  $V_2$ .  $\square$

Existem vários critérios para detetar quando uma soma de subespaços é uma soma direta. A utilização de um ou outro é mais ou menos conveniente dependendo da circunstância.

**Proposição 1.7.1** *Sejam  $V_1, V_2, \dots, V_k$ ,  $k \geq 2$ , subespaços de um espaço vetorial de dimensão finita  $V$  tais que  $V = V_1 + \dots + V_k$ . As seguintes afirmações são equivalentes.*

a)  $V = V_1 \oplus V_2 \oplus \dots \oplus V_k$ .

b)  $V_i \cap \{V_1 + \dots + V_{i-1}\} = \{o\}$  para todo  $2 \leq i \leq k$ .

c) Se  $\beta_i$  é uma base de  $V_i$ ,  $1 \leq i \leq k$ , então a união  $\beta = \cup \beta_i$  é uma base de  $V$ .

d)  $\dim V = \dim V_1 + \dim V_2 + \dots + \dim V_k$ .

**Demonstração** a)  $\Rightarrow$  b) Um vetor  $v$  na interseção  $V_i \cap \{V_1 + \dots + V_{i-1}\}$  pode ser expresso por duas combinações lineares

$$\begin{aligned} v &= v_1 + \dots + v_{i-1}, \\ v &= o + \dots + o + \underbrace{v}_{\in V_i} + \dots + o. \end{aligned}$$

Por unicidade da expressão segue que  $v_1 = \dots = v_{i-1} = o$ , implicando que  $v = o$ .

b)  $\Rightarrow$  c) Seja  $\beta_i = \{v_{i1}, v_{i2}, \dots, v_{ik_i}\}$  uma base ordenada de  $V_i$ . É claro que a união ordenada  $\beta = \vec{\cup} \beta_i$  é um conjunto ordenado de geradores para o subespaço  $V = V_1 + V_2 + \dots + V_k$ . Precisamos mostrar que  $\beta$  é linearmente independente. Por absurdo, suponha que  $\beta$  é linearmente dependente. Sendo assim, escolha  $i_0$ , o menor inteiro para o qual a base  $\beta_{i_0}$  contém um vetor  $v_{i_0 j_0}$  que é uma combinação linear dos seus antecessores em  $\beta_1 \vec{\cup} \beta_2 \vec{\cup} \dots \vec{\cup} \beta_{i_0}$ . Note que o fato de  $\beta_1$  ser uma base de  $V_1$ , garante que  $i_0 \geq 2$ . Pela escolha feita, é possível expressar uma combinação linear para  $v_{i_0 j_0}$  na forma

$$v_{i_0 j_0} = \underbrace{\sum_{n=1}^{j_0-1} a_{i_0 n} v_{i_0 n}}_{\in V_{i_0}} + \underbrace{\sum_{j=1}^{i_0-1} \sum_{n=1}^{k_j} a_{jn} v_{jn}}_{\in V_{i_0-1}}.$$

Isto significa que o vetor  $v \in V$ ,

$$v = \underbrace{v_{i_0 j_0}}_{\in V_{i_0}} - \underbrace{\sum_{n=1}^{j_0-1} a_{i_0 n} v_{i_0 n}}_{\in V_{i_0}} = \underbrace{\sum_{n=1}^{k_1} a_{1n} v_{1n}}_{\in V_1} + \underbrace{\sum_{n=1}^{k_2} a_{2n} v_{2n}}_{\in V_2} + \dots + \underbrace{\sum_{n=1}^{k_{i_0-1}} a_{i_0-1,n} v_{i_0-1,n}}_{\in V_{i_0-1}},$$

pertence à interseção  $V_{i_0} \cap \{V_1 + \dots + V_{i_0-1}\} = \{o\}$ , de onde concluímos que

$$v_{i_0 j_0} = \sum_{n=1}^{j_0-1} a_{i_0 n} v_{i_0 n}.$$

Logo, existe um vetor não nulo na base  $\beta_{i_0} = \{v_{i_0 1}, v_{i_0 2}, \dots, v_{i_0 k_{i_0}}\}$  que é uma combinação linear dos outros vetores desta base, evidentemente uma contradição. Daí segue que  $\beta$  é linearmente independente, como desejávamos demonstrar.

$c) \Rightarrow d)$  Seja  $\beta_i$  uma base de  $V_i$ . Por hipótese a união  $\beta = \cup \beta_i$  é uma base de  $V$ , garantindo que  $\beta_i \cap \beta_j = \{ \}$  para  $i \neq j$ . Como os conjuntos são dois a dois disjuntos, contando os elementos da união temos

$$\dim V = \# \beta = \sum_{i=1}^k \# \beta_i = \sum_{i=1}^k \dim V_i.$$

$d) \Rightarrow a)$  Se  $\beta_i$  é uma base de  $V_i$ . É claro que a união  $\beta = \cup \beta_i$  é um conjunto de geradores de  $V = V_1 + V_2 + \dots + V_k$ , Pelo Teorema da troca sabemos que  $\# \beta \geq \dim V$ . Pela hipótese segue a desigualdade

$$\dim V = \sum_{i=1}^k \dim V_i = \sum_{i=1}^k \# \beta_i \geq \# \beta.$$

Potanto,  $\# \beta = \dim V$ . Sendo  $\beta$  um conjunto de geradores com a cardinalidade igual a dimensão de  $V$ , garantimos que  $\beta$  é uma base de  $V$ . Finalmente, como cada vetor  $v \in V$  é expresso de maneira única na forma

$$v = \underbrace{\sum_{n \geq 1} a_{1n} v_{1n}}_{\in V_1} + \underbrace{\sum_{n \geq 1} a_{2n} v_{2n}}_{\in V_2} + \dots + \underbrace{\sum_{n \geq 1} a_{kn} v_{kn}}_{\in V_k},$$

obtemos a decomposição em soma direta  $V = V_1 \oplus V_2 \oplus \dots \oplus V_k$ . □

### Exercícios propostos 1.7.1

1. Verifique que  $\mathbb{R}^3 = W_1 + W_2$  e determine quais das somas são soma direta.
  - a)  $W_1 = \{(x_1, x_2, x_3); x_1 + x_2 + x_3 = 0\}$  e  $W_2 = \{(x_1, x_2, x_3); x_1 = x_2\}$ .
  - b)  $W_1 = \{(x_1, x_2, x_3); x_1 + x_2 = 0\}$  e  $W_2 = \{(x_1, x_2, x_3); x_1 = x_2 \text{ e } x_1 = x_3\}$ .
2. Sejam  $W_1$  e  $W_2$  subespaços do espaço vetorial de dimensão finita  $V$ .
  - a) Demonstre ou dê contra-exemplo para a afirmação: "se existem bases  $\beta_1$  e  $\beta_2$  de  $W_1$  e  $W_2$ , respectivamente, tais que  $\beta = \beta_1 \cup \beta_2$  é uma base de  $V$ , então  $V = W_1 \oplus W_2$ ".
  - b) Mostre que se  $V = W_1 + W_2$  e  $W_1 \cap W_2 = \{0\}$ , então  $V = W_1 \oplus W_2$ .
  - c) Assuma que  $V = W_1 \oplus W_2$ . Prove que em cada classe do espaço quociente  $V/W_2$  pode ser representado por um único vetor de  $W_1$ .
3. Sejam  $V_1, V_2, \dots, V_k$  subespaços do espaço vetorial de dimensão finita  $V$  tais que  $V = V_1 \oplus V_2 \oplus \dots \oplus V_k$ .

- a) Prove ou dê contra-exemplo para a afirmação: se  $W \subset V$  é um subespaço então  $W = (V_1 \cap W) \oplus (V_2 \cap W) \oplus \cdots \oplus (V_k \cap W)$ .
- b) Mostre que  $V_1 \cap V_i = \{o\}$ , para todo  $2 \leq i \leq k$ . A recíproca é verdadeira?
4. Demonstre que  $M(n, \mathbb{K})$ , o espaço das matrizes quadradas  $n \times n$ , decompõe-se em uma soma direta  $M(n, \mathbb{K}) = \mathcal{S} \oplus \mathcal{A}$ , na qual a primeira parcela é o subespaço das *matrizes simétricas* e a outra parcela é o subespaço das *matrizes anti-simétricas*. Calcule a dimensão de cada parcela.
5. Demonstre que  $C^0(\mathbb{R}, \mathbb{R})$ , o espaço das funções contínuas da reta para a reta, decompõe-se em uma soma direta na qual uma das parcelas é o subespaço das *funções pares* e a outra parcela é o subespaço das *funções ímpares*.
6. Assuma que um espaço vetorial de dimensão finita  $V$  decompõe-se em  $V = W_1 \oplus W_2$  e  $V = U_1 \oplus U_2$ . Mostre que se  $W_i \subset U_i$  então  $W_1 = U_1$  e  $W_2 = U_2$ .
7. Sejam  $W_1, W_2, \dots, W_r$  subespaços próprios de um espaço vetorial  $V$  de dimensão  $n$ . Suponha que  $\dim W_i = k < n$  para todo  $i = 1, 2, \dots, r$ . Prove que existe um subespaço  $W$  de dimensão  $n - k$  tal que  $V = W \oplus W_i$  para qualquer  $i$ .

## Capítulo 2

# Transformações lineares

Apresentaremos nesse capítulo o conceito de transformação linear, o outro objeto principal de nosso estudo. Recordamos que os únicos corpos aqui considerados são os corpos dos números reais e o dos complexos. Os resultados e definições das duas primeiras seções desse capítulo são válidos para espaços vetoriais de dimensão finita ou não.

### 2.1 Transformações lineares

Chamaremos de *transformação linear* a uma aplicação  $A : V \rightarrow W$ , em que  $V$  e  $W$  são espaços vetoriais sobre o mesmo corpo  $\mathbb{K}$ , possuindo as propriedades:

$$\left\{ \begin{array}{l} \text{tl 1) } A(\lambda v) = \lambda A(v), \\ \text{tl 2) } A(v + w) = A(v) + A(w), \end{array} \right.$$

para quaisquer vetores  $v, w \in V$  e qualquer escalar  $\lambda \in \mathbb{K}$ . Ao longo deste texto utilizaremos o termo "aplicação" como sinônimo de "função" e ficará subentendido que ao dizermos que  $A$  é uma transformação de  $V$  em  $W$  estaremos sempre supondo que os espaços vetoriais estão definidos sobre o mesmo corpo, isto evitará repetições enfadonhas.

Uma transformação linear possui duas propriedades básicas, a saber,  $A(o) = o$  e  $A(-v) = -A(v)$  qualquer que seja  $v \in V$ . Para demonstrá-las basta considerar  $\lambda = 0$  na propriedade tl.1 e  $w = -v$  na propriedade tl.2. Finalmente, uma transformação linear  $A : V \rightarrow V$  é também chamada de *operador linear*. Apresentemos exemplos para ilustrar o conceito.

**Exemplo 2.1.1** 1) A aplicação  $Id_V : V \rightarrow V$ ,  $Id(v) = v$ , chamada de *aplicação identidade*, é uma transformação linear do espaço vetorial  $V$  em  $V$ . Chamaremos

de transformação linear identicamente nula de  $V$  em  $W$  a aplicação  $A(v) = 0$  para todo  $v \in V$ .

2) Não é difícil verificar que a aplicação  $A : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ ,  $A(x_1, x_2) = (-x_1, 3x_2)$ , é um operador linear, isto é feito com os seguintes cálculos que são procedimentos padrões. Considere dois vetores  $v = (x_1, x_2)$  e  $w = (y_1, y_2)$  em  $\mathbb{C}^2$  e um escalar  $\lambda \in \mathbb{C}$ . Calculemos

$$\begin{aligned} A(v+w) &= A(x_1+y_1, x_2+y_2) & A(\lambda v) &= A(\lambda x_1, \lambda x_2) \\ &= (-x_1-y_1, 3x_2+3y_2) & &= (-\lambda x_1, 3\lambda x_2) \\ &= (-x_1, 3x_2) + (-y_1, 3y_2) & &= \lambda(-x_1, 3x_2) \\ &= A(v) + A(w), & &= \lambda A(x_1, x_2). \end{aligned}$$

3) A aplicação  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ ,  $A(x_1, x_2) = (x_1 + x_2, x_2, x_1 + x_2)$ , é uma transformação linear. Verifique!

4) Fixado  $\lambda_0 \in \mathbb{K}$ , um escalar do corpo sobre o qual o espaço vetorial  $V$  está definido, a aplicação  $A : V \rightarrow V$ ,  $A(v) = \lambda_0 v$ , é um operador linear chamado de *homotetia*.

5) Suponha que seja dada uma decomposição do espaço vetorial  $V$  em uma soma direta, digamos  $V = V_1 \oplus V_2 \oplus \dots \oplus V_k$ . A *projeção sobre  $V_i$  ao longo das outras parcelas* é a aplicação

$$\pi_i : V \rightarrow V, \quad \pi_i(v) = v_i,$$

em que  $v = v_1 + v_2 + \dots + v_k$  é a decomposição do vetor determinada pela soma direta. A aplicação  $\pi$  é uma transformação linear.

6) A derivada usual de uma função diferenciável induz um operador linear no espaço dos polinômios  $D : \mathbb{K}[t] \rightarrow \mathbb{K}[t]$ ,  $D(p)(t) = p'(t)$ .  $\square$

A cada transformação linear  $A : V \rightarrow W$  podemos associar dois conjuntos,

$$\begin{aligned} \text{Im } A &= \{A(v) \in W; v \in V\}, \\ \text{Nuc } A &= \{v \in V; A(v) = o\}. \end{aligned}$$

O primeiro conjunto é chamado de *imagem* da transformação linear e o segundo de *núcleo*.

**Exercício 2.1.1** 1) Prove que o núcleo e a imagem de uma transformação linear são subespaços do domínio e do contradomínio, respectivamente.

2) Demonstre a equivalência:  $A : V \rightarrow W$  é uma transformação linear se, e somente se,  $A(v + \lambda w) = A(v) + \lambda A(w)$  para quaisquer  $v, w \in V$  e  $\lambda \in \mathbb{K}$ .  $\square$

Registremos numa proposição dois fatos simples mas de bastante utilidade.

**Proposição 2.1.1** *Seja  $A : V \rightarrow W$  uma transformação.*

- a)  $A$  é injetiva  $\Leftrightarrow \text{Nuc } A = \{o\}$ .  
 b)  $A$  é sobrejetiva  $\Leftrightarrow \text{Im } A = W$ .

**Demonstração** a) Suponha que  $A$  é injetiva. Como  $A(o) = o$ , somente o vetor nulo, e nenhum outro vetor, pode assumir o valor  $o \in W$ , mostrando que  $\text{Nuc } A = \{o\}$ . Vamos supor que  $\text{Nuc } A = \{o\}$ . Sejam  $v, w \in V$  vetores tais que  $A(v) = A(w)$ . Por linearidade obtemos  $A(v - w) = o$ . Como o núcleo é trivial concluímos que  $v - w = o$ , isto é,  $v = w$ , mostrando a injetividade. A demonstração do item b) é trivial.  $\square$

### Exercícios propostos 2.1.1

- Verifique quais das aplicações são transformações lineares.
  - $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2, A(x, y) = (xy, y)$ .
  - $A : \mathbb{R}^3 \rightarrow \mathbb{R}^2, A(x, y, z) = 3(x - y, x + 2y + z)$ .
  - $A : \mathbb{C}^2 \rightarrow \mathbb{C}^2, A(z, w) = (i - z + w, z - 3w)$ .
  - $A : \mathbb{C}^3 \rightarrow \mathbb{C}^2, A(u, z, w) = (u - 3z, u + 2z - 3w)$ .
- Calcule uma base do núcleo e uma base da imagem para cada uma das transformações lineares encontradas no item anterior.
- Determine escalares  $a, b, c, d$  tais que o operador linear  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2, A(x, y) = (ax + by, cx + dy)$ , tenha como imagem o subespaço  $W_1 = \{(x, y) \in \mathbb{R}^2; x = y\}$  e como núcleo o subespaço  $W_2 = \{(x, y) \in \mathbb{R}^2; 2x = y\}$ .
- Seja  $A : V \rightarrow W$  uma transformação linear. Demonstre as seguintes afirmações.
  - A imagem por  $A$  de um subespaço de  $V$  é um subespaço de  $W$ .
  - Suponha que  $V_1 \subset V$  é o espaço das combinações lineares de  $\varepsilon \subset V$ . Então o conjunto  $A(V_1) = \{A(v); v \in V_1\}$  é o espaço das combinações lineares de  $A(\varepsilon) = \{A(v), v \in \varepsilon\}$ .
  - Se  $W_1 \subset W$  é um subespaço então o conjunto  $A^{-1}(W_1) = \{v \in V; A(v) \in W_1\}$  é um subespaço de  $V$ .
- Assuma que  $V$  é um espaço vetorial sobre  $\mathbb{K}$ . Mostre que toda transformação linear  $A : \mathbb{K} \rightarrow V$  é da forma  $A(x) = xv_0$ , para algum  $v_0 \in V$ .
- Escolhidos dois vetores  $v_1, v_2 \in V$  defina a aplicação  $A : \mathbb{K}^2 \rightarrow V, A(x_1, x_2) = x_1v_1 + x_2v_2$ .

- a) Verifique que  $A$  é uma transformação linear.
- b) Demonstre que  $u$  e  $v$  são linearmente independente  $\Leftrightarrow$  a aplicação é injetiva.
- c) Suponha que  $u$  e  $v$  são linearmente dependentes. Dê uma base para o núcleo.
7. Considere a aplicação  $\Psi : M(2 \times 1, \mathbb{K}) \rightarrow M(3 \times 1, \mathbb{K})$ ,  $\Psi(N) = P_0 N$ , em que  $[P_0]$  é a matriz dada ao lado. Verifique que  $\Psi$  é uma transformação linear.  $\Psi$  é injetiva? É sobrejetiva?
- $$P_0 = \begin{bmatrix} 2 & -2 \\ 1 & 0 \\ 0 & 3 \end{bmatrix}.$$
8. Seja  $A : \mathbb{R} \rightarrow \mathbb{R}$  um operador linear.
- a) Mostre que existe um número  $a_0 \in \mathbb{R}$  tal que  $A(x) = a_0 x$  para todo  $x \in \mathbb{R}$ .
- b) Suponha que  $A(2) = -3$  e calcule  $A(3)$ .
9. Dada uma transformação linear  $A : \mathbb{R}^2 \rightarrow \mathbb{R}$ .
- a) Mostre que existem  $a, b \in \mathbb{R}$  tais que  $A(x_1, x_2) = ax_1 + bx_2$ .
- b) Assuma que  $A(1, -1) = -3$  e  $A(1, 1) = 2$ . Calcule  $A(2, 3)$ .
10. Prove as afirmações abaixo utilizando o Teorema da troca.
- a) Uma transformação linear  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  não pode ser sobrejetiva.
- b) Uma transformação linear  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ , não pode ser injetiva.
11. Dada uma transformação linear  $A : V \rightarrow W$  definimos o *gráfico* de  $A$  como sendo o conjunto  $graf(A) = \{(v, A(v)); v \in V\} \subset V \times W$ .
- a) Prove que  $graf(A)$  é um subespaço de  $V \times W$ .
- b) Calcule a dimensão de  $graf(A)$  quando  $\dim V = n$ .
12. Seja  $A : V \rightarrow V$  um operador linear. Mostre que
- $$V = Im A \oplus Nuc A \quad \Leftrightarrow \quad Im A \cap Nuc A = \{o\}.$$
13. Considere  $V = V_1 \oplus V_2 \oplus \dots \oplus V_k$  uma decomposição em soma direta do espaço vetorial  $V$ . Seja  $\pi_i : V \rightarrow V$  a projeção sobre  $V_i$  ao longo dos outros fatores. Identifique os espaços  $Im \pi_i$  e  $Nuc \pi_i$ .
14. Uma *involução* em  $V$  é um operador linear  $S : V \rightarrow V$  tal que  $S^2 = Id_V$ . Mostre que para esses operadores os conjuntos
- $$W_+ = \{v \in V; S(v) = v\} \quad \text{e} \quad W_- = \{v \in V; S(v) = -v\}$$
- são subespaços e que  $V = W_+ \oplus W_-$ .
15. Seja  $W$  um subespaço de um espaço vetorial  $V$ . Prove que a *projeção quociente*  $\pi : V \rightarrow V/W$ ,  $\pi(v) = \bar{v}$ , é uma transformação linear sobrejetiva.
16. Prove: uma aplicação  $A : \mathbb{R}^m \rightarrow \mathbb{R}^n$  tal que  $A(u+v) = A(u) + A(v)$  para todo  $u, v \in \mathbb{R}^m$  é uma transformação linear. Você precisará do conceito de continuidade.

## 2.2 O espaço das transformações lineares

Tomaremos conhecimento agora de um novo espaço. Denotamos por  $\mathcal{L}(V, W)$  o conjunto de todas as transformações lineares  $A : V \rightarrow W$ . As usuais operações de *soma* duas funções e *multiplicação* uma função por um escalar de  $\mathbb{K}$ , o corpo sobre o qual as estruturas de espaços vetoriais em  $V$  e  $W$  estão definidas, induzem no conjunto  $\mathcal{L}(V, W)$  uma estrutura de espaço vetorial. Com efeito, sejam  $A, B \in \mathcal{L}(V, W)$  e  $v, w \in V$ . Calculemos

$$\begin{aligned} (A + B)(v + w) &\equiv A(v + w) + B(v + w) \\ &= A(v) + A(w) + B(v) + B(w) \\ &= (A + B)(v) + (A + B)(w). \end{aligned}$$

De modo semelhante mostramos que  $(A + B)(\lambda v) = \lambda(A + B)(v)$ . Isto significa que  $A + B \in \mathcal{L}(V, W)$ . Também sem dificuldade alguma é possível mostrar que  $\mu A \in \mathcal{L}(V, W)$  para todo  $\mu \in \mathbb{K}$ . É rotina verificar que essas operações satisfazem aos axiomas da definição de espaço vetorial. O vetor nulo será a transformação linear identicamente nula e  $-A = -(A)$ .

**Exemplo 2.2.1** Dadas as transformações lineares  $A, B : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  definidas por

$$A(x, y, z) = (2y, z - x) \quad \text{e} \quad B(x, y, z) = (x - z, z),$$

calculemos  $A - 2B : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ . Pela definição, obtemos

$$(A - 2B)(x, y, z) = (2y, z - x) - 2(x - z, z) = (2y - 2x + 2z, -z - x).$$

O leitor pode verificar que de fato  $A - 2B$  é uma transformação linear.  $\square$

Uma outra operação envolvendo transformações lineares é a operação de *composição*. Se  $A : V \rightarrow W$  e  $C : W \rightarrow Y$  são duas transformações lineares podemos construir uma outra transformação linear denotada por  $C \circ A : V \rightarrow Y$  e chamada de composta de  $C$  e  $A$ , definindo  $C \circ A(v) = C(A(v))$  para todo  $v \in V$ . Observamos que o contradomínio de  $A$  deve ser o domínio de  $C$  e todos os espaços envolvidos estão sobre o mesmo corpo  $\mathbb{K}$ . De fato, a composta é também uma transformação linear pois se  $u, v \in V$  e  $\lambda \in \mathbb{K}$ , as igualdades abaixo são justificadas pelas definições já apresentadas,

$$\begin{aligned} C \circ A(u + \lambda v) &= C(A(u + \lambda v)) \\ &= C(A(u) + \lambda A(v)) \\ &= C(A(u)) + \lambda C(A(v)) \\ &= C \circ A(u) + \lambda C \circ A(v). \end{aligned}$$

**Exemplo 2.2.2** Considere as transformações lineares  $\mathbb{R}^3 \xrightarrow{A} \mathbb{R}^2 \xrightarrow{C} \mathbb{R}^2$  definidas por  $A(x_1, x_2, x_3) = (x_2, 2x_3)$  e  $C(x_1, x_2) = (x_2, x_1 - x_2)$ . Calculemos a composta,

$$C \circ A(x_1, x_2, x_3) = C(A(x_1, x_2, x_3)) = C(x_2, 2x_3) = (2x_3, x_2 - 2x_3).$$

Observe o domínio e o contradomínio da composta  $C \circ A : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ .  $\square$

**Exercício 2.2.1** Deixaremos como exercícios as demonstrações de várias propriedades algébricas sobre a operação de composição. Prove as seguintes afirmações.

1.  $C \circ (A + B) = C \circ A + C \circ B$  para quaisquer  $A, B \in \mathcal{L}(V, W)$  e  $C \in \mathcal{L}(W, Y)$ .
2.  $(A + B) \circ C = A \circ C + B \circ C$  para quaisquer  $A, B \in \mathcal{L}(W, Y)$  e  $C \in \mathcal{L}(V, W)$ .
3.  $A \circ (B \circ C) = (A \circ B) \circ C$  para quaisquer  $A \in \mathcal{L}(Y, Z)$ ,  $B \in \mathcal{L}(W, Y)$  e  $C \in \mathcal{L}(V, W)$ .
4.  $A \circ Id_V = A$  para qualquer  $A \in \mathcal{L}(V, W)$  em que  $Id_V$  é a aplicação identidade de  $V$ . Da mesma forma  $Id_W \circ B = B$  para qualquer  $B \in \mathcal{L}(W, V)$ .  $\square$

Utilizamos a operação de composição para apresentar uma nova noção. Uma transformação linear  $A : V \rightarrow W$  é *invertível* se existe uma aplicação  $B : W \rightarrow V$  tal que

$$B \circ A = Id_V \in \mathcal{L}(V, V) \quad \text{e} \quad A \circ B = Id_W \in \mathcal{L}(W, W).$$

Quando existe uma tal aplicação diremos  $B$  é a *inversa* de  $A$  e denotaremos esta aplicação inversa por  $A^{-1} : W \rightarrow V$ . Da Teoria de conjuntos sabemos que uma função entre dois conjuntos é invertível se, e somente se, a função é sobrejetiva e injetiva. Em particular, isto é válido para transformações lineares, logo podemos afirmar que

**Proposição 2.2.1** *Uma transformação linear  $A : V \rightarrow W$  é invertível se, e somente se,  $Im A = W$  e  $Nuc A = \{o\}$ .*

**Exemplo 2.2.3** 1) A aplicação identidade de qualquer espaço é invertível e a sua inversa é ela própria.

2) A transformação linear  $A : \mathbb{K}^2 \rightarrow \mathbb{K}^2$ ,  $A(x_1, x_2) = (x_1 - x_2, x_2)$  é invertível e tem como aplicação inversa

$$A^{-1} : \mathbb{K}^2 \rightarrow \mathbb{K}^2, \quad A^{-1}(u_1, u_2) = (u_1 + u_2, u_2).$$

3) Apenas uma das condições exigidas na definição de uma transformação linear invertível não é suficiente para garantir a invertibilidade da transformação.

Compondo as transformações

$$\begin{aligned} A : \mathbb{K}^2 &\rightarrow \mathbb{K}^3, & A(x_1, x_2) &= (x_1, x_2, 0), \\ B : \mathbb{K}^3 &\rightarrow \mathbb{K}^2, & B(u_1, u_2, u_3) &= (u_1, u_2) \end{aligned}$$

obtemos  $B \circ A = Id \in \mathcal{L}(\mathbb{K}^2, \mathbb{K}^2)$ , entretanto  $A \circ B \neq Id \in \mathcal{L}(\mathbb{K}^3, \mathbb{K}^3)$ .  $\square$

**Proposição 2.2.2** *Se  $A : V \rightarrow W$  é invertível, então só existe uma inversa para  $A$  e essa inversa também é uma transformação linear.*

**Demonstração** Iniciemos com a prova da unicidade. Assuma que  $C : W \rightarrow V$  é uma aplicação tal que

$$C \circ A = Id_V \in \mathcal{L}(V, V) \quad \text{e} \quad A \circ C = Id_W \in \mathcal{L}(W, W).$$

Então

$$C = Id_V \circ C = (A^{-1} \circ A) \circ C = A^{-1} \circ (A \circ C) = A^{-1} \circ Id_W = A^{-1}.$$

Vejamus a linearidade da inversa. Dados  $w_1, w_2 \in W$  e  $\lambda \in \mathbb{K}$ , como  $A$  é sobrejetiva é possível determinar dois vetores  $v_1, v_2 \in V$  tais que  $A(v_1) = w_1$  e  $A(v_2) = w_2$ . Sendo assim temos

$$\begin{aligned} A^{-1}(w_1 + \lambda w_2) &= A^{-1}(A(v_1) + \lambda A(v_2)) \\ &= A^{-1}(A(v_1 + \lambda v_2)) \\ &= v_1 + \lambda v_2 \\ &= A^{-1}(w_1) + \lambda A^{-1}(w_2). \end{aligned}$$

Isso termina a demonstração.  $\square$

**Exercício 2.2.2** Se  $A : V \rightarrow W$  e  $C : W \rightarrow Y$  são duas transformações lineares invertíveis, prove que a composta  $C \circ A : V \rightarrow Y$  é uma transformação linear invertível e que  $(C \circ A)^{-1} = A^{-1} \circ C^{-1}$ .  $\square$

**Exercício 2.2.3** Considere  $V = V_1 \oplus V_2 \oplus \dots \oplus V_k$  uma decomposição em soma direta do espaço vetorial  $V$ . Seja  $\pi_i : V \rightarrow V$  a projeção sobre  $V_i$  ao longo das outras parcelas. Verifique as identidades funcionais,

$$\text{a) } \pi_i \circ \pi_i \equiv \pi_i, \quad \text{b) } \pi_i \circ \pi_j \equiv 0 \quad \text{se } i \neq j, \quad \text{c) } Id_V \equiv \pi_1 + \pi_2 + \dots + \pi_k.$$

Reciprocamente, suponha que  $\pi_i : V \rightarrow V$ ,  $i = 1, \dots, k$ , é uma coleção de transformações lineares satisfazendo a), b) e c), prove que existe uma decomposição de  $V$  em soma direta com  $k$  parcelas tais que cada  $\pi_i$  é a projeção sobre a parcela  $i$  ao longo das outras parcelas.  $\square$

Posteriormente, teremos oportunidade de utilizar a seguinte definição envolvendo aplicações entre espaços vetoriais complexos.

**Definição 2.2.1** *Diremos que uma aplicação  $A : V \rightarrow W$  entre espaços vetoriais complexos é uma transformação quase linear quando possuir as propriedades:*

1.  $A(\lambda u) = \bar{\lambda}A(u)$ ,
2.  $A(u + v) = A(u) + A(v)$ , para quaisquer  $u, v \in V$  e  $\lambda \in \mathbb{C}$ .

Estamos indicando o *conjugado complexo* sobrepondo uma barra ao escalar  $\lambda \in \mathbb{C}$ . Um exemplo de transformação quase linear é a aplicação

$$A : \mathbb{C}^3 \rightarrow \mathbb{C}^2, \quad A(z_1, z_2, z_3) = (2\bar{z}_3, \bar{z}_1 - i\bar{z}_2).$$

### Exercícios propostos 2.2.1

1. Calcule, quando possível, a composta  $A \circ B$  e  $B \circ A$  em que  $A$  e  $B$  são as transformações lineares dadas.
  - a)  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad A(x, y) = (2x, y - x),$   
 $B : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad B(x, y) = (x - y, y, -y).$
  - b)  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^3, \quad A(x, y) = (3y, y - 2x, y - x)$   
 $B : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad B(x, y, z) = (x - y, y, -2x).$
  - c)  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad A(x, y) = (2x, y - x)$   
 $B : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad B(x, y) = (x - y, y).$
2. Suponha que uma transformação linear  $A : V \rightarrow W$  é injetiva. Demonstre que se  $\varepsilon = \{v_1, v_2, \dots, v_k\} \subset V$  é um conjunto linearmente independente, então o conjunto  $A(\varepsilon) = \{A(v_1), A(v_2), \dots, A(v_k)\} \subset W$  é linearmente independente.
3. Utilize-se do Teorema da troca e do exercício anterior para mostrar que uma transformação linear  $A : \mathbb{R}^m \rightarrow \mathbb{R}^n$ , com  $m \neq n$ , não pode ser invertível.
4. Suponha que o operador linear  $A : V \rightarrow V$  tenha uma inversa à esquerda, isto é, suponha que exista  $B \in \mathcal{L}(V, V)$  tal que  $B \circ A \equiv Id$ . Prove que  $A$  é invertível.
5. Assuma que  $\beta = \{v_1, v_2\}$  é uma base de  $V$ . Defina a aplicação  $A : \mathbb{R}^2 \rightarrow V$  por  $A(x, y) = xv_1 + yv_2$ . Prove que  $A$  é invertível.
6. Prove que o núcleo e a imagem de uma transformação quase linear são subespaços.
7. Seja  $A$  um operador linear num espaço vetorial  $V$  de dimensão finita. Demonstre as afirmações abaixo.
  - a)  $V \supseteq Im A \supseteq Im A^2 \supseteq Im A^3 \supseteq \dots$
  - b)  $\{o\} \subseteq Nuc A \subseteq Nuc A^2 \subseteq Nuc A^3 \subseteq \dots$

- c) Existe um inteiro  $r \geq 1$  tal que  $\text{Nuc } A^{r+i} = \text{Nuc } A^r$  e  $\text{Im } A^{r+i} = \text{Im } A^r$  para todo inteiro  $i \geq 0$ .
- d)  $V = \text{Im } A^r \oplus \text{Nuc } A^r$

## 2.3 Teorema do núcleo e da imagem

Iniciaremos agora o estudo de transformações lineares entre espaços vetoriais de dimensão finita. Antes de tudo mostraremos a existência (e abundância) destas aplicações. Uma propriedade relevante de uma transformação linear é que ela fica determinada conhecendo-se os valores nos vetores de uma base do seu domínio. A demonstração deste fato é construtiva e a técnica de demonstração será utilizada outras vezes. Essencialmente repetiremos o seguinte procedimento. Suponha que desejemos construir uma transformação linear do  $\mathbb{R}^2$  num espaço vetorial real  $V$ . Para isso, escolhamos dois vetores  $v_1, v_2 \in V$  e definimos  $A(x_1, x_2) = x_1v_1 + x_2v_2$ .

**Proposição 2.3.1** *Sejam  $V$  e  $W$  dois espaços vetoriais sobre o mesmo corpo  $\mathbb{K}$  e  $\beta = \{v_1, v_2, \dots, v_n\}$  uma base ordenada de  $V$ . Dados vetores arbitrários  $w_1, w_2, \dots, w_n \in W$ , existe uma, e somente uma, transformação linear  $A : V \rightarrow W$  tal que  $A(v_i) = w_i$ , para todo  $i = 1, 2, \dots, n$ .*

**Demonstração** Primeiro, demonstremos a existência construindo a transformação. Dado um vetor  $v \in V$ , considere suas coordenadas na base ordenada  $\beta$ , digamos que estas coordenadas sejam  $x_1, x_2, \dots, x_n \in \mathbb{K}$ . Recordamos que estas coordenadas são os únicos escalares que satisfazem a combinação linear  $v = x_1v_1 + x_2v_2 + \dots + x_nv_n$ . Escolhidos vetores  $w_1, w_2, \dots, w_n \in W$ , a aplicação  $A : V \rightarrow W$  definida por

$$A(x_1v_1 + x_2v_2 + \dots + x_nv_n) = x_1w_1 + x_2w_2 + \dots + x_nw_n$$

é uma transformação linear. Com efeito. Se  $u = y_1v_1 + y_2v_2 + \dots + y_nv_n$  e  $\lambda \in \mathbb{K}$  então

$$\begin{aligned} A(v + \lambda u) &= A((x_1 + \lambda y_1)v_1 + (x_2 + \lambda y_2)v_2 + \dots + (x_n + \lambda y_n)v_n) \\ &= (x_1 + \lambda y_1)w_1 + (x_2 + \lambda y_2)w_2 + \dots + (x_n + \lambda y_n)w_n \\ &= (x_1w_1 + x_2w_2 + \dots + x_nw_n) + \lambda(y_1w_1 + y_2w_2 + \dots + y_nw_n) \\ &= A(v) + \lambda A(u). \end{aligned}$$

Isto mostra que  $A$  é uma transformação linear, e evidentemente,  $A(v_i) = w_i$  para todo  $i = 1, \dots, n$ , pois  $v_i = 0v_1 + \dots + 1v_i + \dots + 0v_n$ . Feita a demonstração da existência, examinemos a unicidade. Suponha que  $B : V \rightarrow W$  seja uma transformação linear satisfazendo a condição  $B(v_i) = w_i$  para  $i = 1, \dots, n$ . Pela

definição de  $A$  e linearidade das aplicações temos as seguintes igualdades,

$$\begin{aligned}
 A(v) &= A(x_1v_1 + x_2v_2 + \cdots + x_nv_n) \\
 &= x_1w_1 + x_2w_2 + \cdots + x_nw_n \\
 &= x_1B(v_1) + x_2B(v_2) + \cdots + x_nB(v_n) \\
 &= B(x_1v_1 + x_2v_2 + \cdots + x_nv_n) \\
 &= B(v),
 \end{aligned}$$

para todo  $v \in V$ . Portanto  $A \equiv B$ , concluindo a demonstração da proposição.  $\square$

Para um conhecimento maior de transformações lineares em espaços vetoriais de dimensão finita precisaremos de um resultado básico, conhecido como Teorema do núcleo e da imagem, do qual decorrem várias conclusões importantes.

**Teorema 2.3.1 (Teorema do núcleo e da imagem)** *Seja  $A : V \rightarrow W$  uma transformação linear. Se  $V$  tem dimensão finita então*

$$\dim V = \dim Nuc A + \dim Im A.$$

A dimensão da imagem e a dimensão do núcleo são chamados de posto e nulidade de  $A$ , respectivamente.

**Demonstração** Considere uma base ordenada  $\beta = \{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$  para  $V$  na qual os  $k$  primeiros elementos formam uma base para o núcleo de  $A$ . Iremos mostrar que o conjunto de vetores  $\gamma = \{A(v_{k+1}), A(v_{k+2}), \dots, A(v_n)\}$  é uma base para o subespaço  $Im A$ . Por definição, dado um vetor  $w \in Im A$  existe um vetor  $v \in V$  tal que  $w = A(v)$ . Escrevamos o vetor  $v$  como uma combinação linear dos vetores da base  $\beta$ ,  $v = x_1v_1 + x_2v_2 + \cdots + x_nv_n$ . Levando-se em conta que os  $k$  primeiros vetores de  $\beta$  pertencem ao núcleo de  $A$ , avaliemos,

$$\begin{aligned}
 w &= A(v) \\
 &= A(x_1v_1 + x_2v_2 + \cdots + x_nv_n) \\
 &= x_{k+1}A(v_{k+1}) + x_{k+2}A(v_{k+2}) + \cdots + x_nA(v_n).
 \end{aligned}$$

Daí concluímos que  $\gamma$  é um conjunto de geradores para a imagem de  $A$ . Consideremos agora a combinação linear  $o = y_{k+1}A(v_{k+1}) + y_{k+2}A(v_{k+2}) + \cdots + y_nA(v_n)$ . Por linearidade de  $A$  podemos reescrever esta última equação como

$$o = A(y_{k+1}v_{k+1} + y_{k+2}v_{k+2} + \cdots + y_nv_n).$$

Isto significa que o vetor  $y_{k+1}v_{k+1} + y_{k+2}v_{k+2} + \cdots + y_nv_n \in Nuc A$ . Logo, este último vetor também é uma combinação linear dos  $k$  primeiros vetores da base ordenada  $\beta$ , pois tais vetores formam uma base para o núcleo de  $A$ , isto é,

$$y_{k+1}v_{k+1} + y_{k+2}v_{k+2} + \cdots + y_nv_n = x_1v_1 + x_2v_2 + \cdots + x_kv_k$$

ou equivalentemente,

$$x_1v_1 + x_2v_2 + \cdots + x_kv_k - y_{k+1}v_{k+1} - y_{k+2}v_{k+2} - \cdots - y_nv_n = 0.$$

Como  $\beta$  é linearmente independente, todos os coeficientes desta combinação linear são nulos, em particular,  $y_{k+1} = y_{k+2} = \cdots = y_n = 0$ , mostrando que o conjunto  $\gamma$  é linearmente independente. Sendo assim,

$$\dim V = n = k + (n - k) = \dim Nuc A + \dim Im A. \quad \square$$

**Exercício 2.3.1** Seja  $A : V \rightarrow W$  é uma transformação linear sobrejetiva. Prove que se  $V$  tem dimensão finita então  $W$  tem dimensão finita e  $\dim W \leq \dim V$ .  $\square$

### Exercícios propostos 2.3.1

- Construa uma transformação linear  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^n$  satisfazendo as condições  $A(e_i) = w_i$ , em que  $\alpha = \{e_1, e_2, e_3\}$  é a base canônica do  $\mathbb{R}^3$  e  $\gamma = \{w_1, w_2, w_3\} \subset \mathbb{R}^n$ .
  - $\gamma = \{(1, 1), (1, -1), (2, 1)\} \subset \mathbb{R}^2$ .
  - $\gamma = \{(2, -3, 1), (0, 1, 0), (1, -1, 4)\} \subset \mathbb{R}^3$ .
  - $\gamma = \{(1, 1, 0, 1), (0, 1, 0, 1), (1, 2, 0, 2)\} \subset \mathbb{R}^4$ .
- Encontre uma base para o núcleo e uma base para a imagem para cada uma das transformações lineares do exercício anterior e diga qual o posto e a nulidade.
- Determine uma base para o núcleo e uma base para a imagem das transformações.
  - $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (6x + y - 3z, z - y, 2x - z)$ .
  - $A : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ ,  $A(x, y) = (x + y, 2x - y, y - 2x)$ .
  - $A : \mathbb{C}^3 \rightarrow \mathbb{C}^2$ ,  $A(x, y, z) = (x + y + z, x + y + z)$ .
- Construa um operador linear  $A$  em  $\mathbb{R}^3$  satisfazendo a condição pedida.
  - $Im A$  é gerado por  $\varepsilon = \{(2, -1, 1), (1, 0, 1)\}$ .
  - $Nuc A$  é gerado por  $\varepsilon = \{(2, -1, 1), (1, 0, 1)\}$ .
  - $W = Nuc A + Im A$  é um subespaço próprio de  $\mathbb{R}^3$ .
  - $Im A \subset Nuc A$ .
  - $A$  não é o operador identicamente nulo e  $A^2 \equiv 0$ .
- Seja  $A : V \rightarrow W$  uma transformação linear entre espaços vetoriais de dimensão finita. Prove as afirmações.
  - Se  $\dim V < \dim W$  então  $A$  não é sobrejetiva.
  - Se  $\dim V > \dim W$  então  $A$  não é injetiva.
- Construa um operador linear  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  com a propriedade pedida.

- a)  $A$  reflete cada vetor em relação ao subespaço  $W = \{(x, 2x), x \in \mathbb{R}\}$ .
- b)  $A$  rotaciona cada vetor no sentido anti-horário por um ângulo de  $\frac{\pi}{3}$ .
7. Se um operador linear  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  é não nulo e  $A^2 \equiv 0$ , prove que  $\text{posto}(A) = 1$ .
8. Existe um operador linear  $A : \mathbb{R}^{11} \rightarrow \mathbb{R}^{11}$  tal que  $\text{Im } A = \text{Nuc } A$ ? Justifique sua resposta. E existe um operador linear  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  tal que  $A^2 = -Id$ ?
9. Assuma que  $A : V \rightarrow W$  é uma transformação linear sobrejetiva entre dois espaços vetoriais de dimensão finita e que  $U \subset W$  é um subespaço. Mostre a igualdade  $\dim V - \dim A^{-1}(U) = \dim W - \dim U$ .
10. Seja  $A$  um operador linear no espaço vetorial  $V$  de dimensão finita. Demonstre que  $V = \text{Im } A \oplus \text{Nuc } A \Leftrightarrow \text{Nuc } A = \text{Nuc } A^2 \Leftrightarrow \text{Im } A = \text{Im } A^2$ .
11. Demonstre que para quaisquer dois operadores lineares  $A$  e  $B$  num espaço vetorial  $V$  de dimensão finita valem as relações
- $$\|\text{posto}(A) - \text{posto}(B)\| \leq \text{posto}(A + B) \leq \text{posto}(A) + \text{posto}(B).$$
12. Prove que se dois operadores lineares  $A$  e  $B$  num espaço vetorial  $V$  de dimensão  $n$  satisfazem a condição  $A \circ B \equiv 0$  e  $A \circ B$  é invertível, então  $\text{posto}(A) + \text{posto}(B) = n$ .
13. Sejam  $A : V \rightarrow W$  e  $B : W \rightarrow Z$  transformações lineares entre espaços de dimensão finita. Prove que  $\text{posto}(BoA) \leq \max\{\text{posto}(A), \text{posto}(B)\}$ .
14. Considere a aplicação  $A : M(2, \mathbb{R}^2) \rightarrow M(2, \mathbb{R}^2)$ , definida por  $A(X) = N_0 X - X N_0$ , onde  $N_0$  é a matriz dada ao lado. Verifique que  $A$  é um operador linear e determine uma base para o núcleo e uma base para a imagem do operador.

$$N_0 = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}.$$

## 2.4 Isomorfismos lineares

Uma transformação linear invertível  $A : V \rightarrow W$  é também chamada de *isomorfismo linear*, ou simplesmente *isomorfismo* quando não causar ambigüidades, e nesse caso, quando existe uma tal transformação diremos que  $V$  e  $W$  são *espaços isomorfos*. Podemos facilmente determinar quando uma transformação linear entre dois espaços vetoriais de dimensão finita é um isomorfismo. A seguir daremos dois corolários do Teorema do núcleo e da imagem.

**Corolário 2.4.1** *Dois espaços vetoriais de dimensão finita sobre o mesmo corpo  $\mathbb{K}$  são isomorfos se, e somente se, as dimensões dos espaços são iguais. Em particular, todo espaço vetorial  $V$  de dimensão  $n$  sobre o corpo  $\mathbb{K}$  é isomorfo ao  $\mathbb{K}^n$ .*

**Demonstração** Vamos supor que  $A : V \rightarrow W$  seja um isomorfismo. Como já sabemos,  $A$  é uma transformação linear injetiva e sobrejetiva, ou equivalentemente,

$Nuc A = \{o\}$  e  $Im A = W$ . Pelo Teorema do núcleo e da imagem obtemos as igualdades,

$$\dim V = \dim Nuc A + \dim Im A = \dim W.$$

Provemos a suficiência construindo o isomorfismo. Sejam  $\beta = \{v_1, v_2, \dots, v_n\}$  e  $\gamma = \{w_1, w_2, \dots, w_n\}$  bases ordenadas de  $V$  e  $W$ , respectivamente. Por hipótese temos que  $\#\beta = \#\gamma$ . Um vetor  $v \in V$  é escrito de maneira única como uma combinação linear na forma

$$v = x_1v_1 + x_2v_2 + \dots + x_nv_n, \quad \text{com} \quad x_i \in \mathbb{K}.$$

Considere a transformação linear  $A : V \rightarrow W$ , definida por

$$A(x_1v_1 + x_2v_2 + \dots + x_nv_n) = x_1w_1 + x_2w_2 + \dots + x_nw_n.$$

O núcleo de  $A$  é trivial, pois se  $v \in Nuc A$ , então

$$\begin{aligned} o &= A(v) \\ &= A(x_1v_1 + x_2v_2 + \dots + x_nv_n) \\ &= x_1w_1 + x_2w_2 + \dots + x_nw_n. \end{aligned}$$

Sendo  $\gamma$  uma base de  $W$ , os coeficientes  $x_i$ 's são nulos implicando que o vetor  $v \in V$  é nulo. A transformação também é sobrejetiva, pois um vetor  $w = y_1w_1 + y_2w_2 + \dots + y_nw_n \in W$  é a imagem do vetor  $v = y_1v_1 + y_2v_2 + \dots + y_nv_n \in V$ .  $\square$

**Corolário 2.4.2** *Seja  $A : V \rightarrow W$  uma transformação linear entre espaços vectoriais de dimensão finita. Suponha que  $\dim V = \dim W$ . Então as seguintes afirmações são equivalentes.*

- a)  $A$  é um isomorfismo linear.
- b)  $Nuc A = \{o\}$ .
- c)  $Im A = W$
- d) A imagem por  $A$  de uma base de  $V$  é uma base de  $W$ .

**Demonstração** As implicações a)  $\rightarrow$  b) e b)  $\rightarrow$  c) ficam como exercícios.

c)  $\Rightarrow$  d) Seja  $\beta = \{v_1, v_2, \dots, v_n\}$  uma base ordenada de  $V$ . É fácil mostrar que o conjunto ordenado  $A(\beta) = \{A(v_1), A(v_2), \dots, A(v_n)\}$  é um conjunto ordenado de geradores de  $Im A = W$  com  $n$  vetores e  $n = \dim W$ , pelo Teorema da troca segue que  $A(\beta)$  é uma base de  $W$ .

d)  $\Rightarrow$  a) Vamos supor, por absurdo, que o núcleo de  $A$  é não trivial. Considere uma base ordenada  $\beta = \{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$  de  $V$  na qual os  $k \geq 1$  primeiros

vetores formam uma base para o núcleo de  $A$ . Por hipótese, o conjunto de vetores  $A(\beta) = \{A(v_{k+1}), A(v_{k+2}), \dots, A(v_n)\}$  é uma base de  $W$ , de onde concluímos que  $\dim W = n - k < \dim V$ , uma contradição. Logo  $k = 0$ , significando que o núcleo é trivial, em outras palavras,  $A$  é injetiva. Pelo Teorema do núcleo e da imagem temos que  $n = \dim V = \dim \text{Im } A$ . Como o subespaço imagem de  $A$  tem a mesma dimensão do contradomínio, concluímos que  $\text{Im } A = W$ , isto é,  $A$  é sobrejetiva. Em resumo,  $A$  é injetiva e sobrejetiva, portanto, é um isomorfismo.  $\square$

### Exercícios propostos 2.4.1

1. Mostre que  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $A(x, y) = (2x, y - x)$ , é um isomorfismo e determine a sua inversa.
2. Seja  $\beta = \{v_1, v_2\}$  uma base ordenada do espaço vetorial complexo  $V$ . Defina  $A : \mathbb{C}^2 \rightarrow V$ ,  $A(z, w) = (z - w)v_1 + (z + w)v_2$ . Verifique que  $A$  é um isomorfismo.
3. Prove que a inversa de um isomorfismo linear também é um isomorfismo.
4. Prove que a composição induz uma estrutura de grupo no subconjunto de  $\mathcal{L}(V, V)$  formado pelos operadores invertíveis.
5. Mostre que o corpo dos reais considerado como espaço vetorial sobre o corpo dos racionais não tem dimensão finita.
6. Sejam  $W_1$  e  $W_2$  dois subespaços de um espaço vetorial  $V$  de dimensão finita. Assuma que  $V = W_1 + W_2$ . Demonstre as afirmações.
  - a)  $A : W_1 \times W_2 \rightarrow V$ ,  $A(u, v) = u + v$ , é uma transformação linear.
  - b)  $\dim V = \dim W_1 + \dim W_2 - \dim W_1 \cap W_2$ .
  - c)  $A$  é um isomorfismo  $\Leftrightarrow W_1 \cap W_2 = \{o\}$
7. Sejam  $A$  e  $B$  sejam operadores lineares num espaço vetorial  $V$  de dimensão finita. Demonstre que se  $Id - A \circ B$  é invertível então  $Id - B \circ A$  é também invertível.
8. Demonstre que todo operador linear  $A \in \mathcal{L}(V, V)$  em que  $V$  é um espaço vetorial de dimensão finita é uma soma de operadores invertíveis.
9. Sejam  $W$  e  $U$  subespaços de um espaço vetorial de dimensão finita  $V$ . Prove que os espaços quocientes  $(W + U)/U$  e  $W/(W \cap U)$  são isomorfos.

## 2.5 Matriz de uma transformação linear

Caso o leitor queira consultar alguma notação sobre matrizes reservamos o Capítulo 12 para tratar de tais objetos. Nessa seção descreveremos um processo pelo qual associamos uma matriz a uma transformação linear entre espaços vetoriais de dimensão finita.

Vamos supor que  $A : V \rightarrow W$  seja uma transformação linear entre espaços vetoriais sobre  $\mathbb{K}$  e que  $\beta = \{v_1, v_2, \dots, v_n\}$  e  $\gamma = \{w_1, w_2, \dots, w_m\}$  sejam bases ordenadas de  $V$  e  $W$ , respectivamente. Sendo assim, para cada  $j$ ,  $1 \leq j \leq n$ , existem únicos escalares  $a_{1j}, a_{2j}, \dots, a_{mj} \in \mathbb{K}$  que formam as coordenadas do vetor  $A(v_j) \in W$  na base ordenada  $\gamma \subset W$ , mais precisamente,

$$A(v_j) = a_{1j}w_1 + a_{2j}w_2 + \dots + a_{mj}w_m.$$

A matriz de  $A$  relativa às bases ordenadas  $\beta$  e  $\gamma$  é a matriz  $m \times n$  com entradas em  $\mathbb{K}$  denotada por  $[A]_{\gamma}^{\beta}$  e organizada na forma

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

Ressaltamos que a  $j$ -ésima matriz coluna de  $[A]_{\gamma}^{\beta}$  é precisamente a matriz das coordenadas do vetor  $A(v_j)$  na base ordenada  $\gamma \subset W$ . Uma tal matriz guarda muitas informações sobre a transformação linear  $A$ . Dadas  $\beta$ ,  $\gamma$  e a matriz  $[A]_{\gamma}^{\beta}$  recuperamos a transformação pela identidade

$$A(v_j) = a_{1j}w_1 + a_{2j}w_2 + \dots + a_{mj}w_m.$$

Como sabemos,  $A$  fica determinada conhecendo-se os valores nos vetores de uma base e, essencialmente, estes valores estão registrados nas colunas da matriz. Diremos que a matriz  $[A]_{\gamma}^{\beta}$  é a *representação matricial de  $A$  nas bases ordenadas  $\beta$  e  $\gamma$* . Observamos que a representação depende das bases e das ordens das bases.

**Exemplo 2.5.1** Seja  $A : \mathbb{K}^3 \rightarrow \mathbb{K}^2$ ,

$$A(x, y, z) = (x - 2y, 2x + 3y - z).$$

Determinemos a representação matricial de  $A$  nas bases canônicas  $\alpha_3$  e  $\alpha_2$ , respectivamente. Para isso, necessitamos das avaliações,

$$\begin{aligned} A(1, 0, 0) &= (1, 2) = 1e_1 + 2e_2, \\ A(0, 1, 0) &= (-2, 3) = -2e_1 + 3e_2, \\ A(0, 0, 1) &= (0, -1) = 0e_1 - 1e_2. \end{aligned}$$

Por definição, a representação matricial de  $A$  nas bases canônicas é a matriz

$$[A]_{\alpha_2}^{\alpha_3} = \begin{bmatrix} 1 & -2 & 0 \\ 2 & 3 & -1 \end{bmatrix}.$$

Para enfatizar a diferença, vejamos a representação matricial da mesma transformação linear  $A$ , agora com respeito às bases  $\alpha_3 \subset \mathbb{K}^3$  e  $\gamma = \{(1, 1), (0, 1)\} \subset \mathbb{K}^2$ . Para isto, precisaremos das avaliações,

$$\begin{aligned} A(1, 0, 0) &= (1, 2) = 1(1, 1) + 2(0, 1), \\ A(0, 1, 0) &= (-2, 3) = -2(1, 1) + 5(0, 1), \\ A(0, 0, 1) &= (0, -1) = 0(1, 1) - 1(0, 1). \end{aligned} \quad \implies \quad [A]_{\gamma}^{\alpha_3} = \begin{bmatrix} 1 & -2 & 0 \\ 1 & 5 & -1 \end{bmatrix},$$

que também é uma matriz  $2 \times 3$ , mas com entradas diferentes das entradas da matriz anterior.  $\square$

### Exercícios propostos 2.5.1

- Determine as representações matriciais da base canônica para a base canônica.
  - $A: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (6x + y - 3z, z - y, 2x - z)$ .
  - $A: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ ,  $A(x, y) = (x + y, 2x - y, y - 2x)$ .
  - $A: \mathbb{C}^3 \rightarrow \mathbb{C}^2$ ,  $A(x, y, z) = (x + y + z, x + y + z)$ .
- Calcule a matriz do operador derivação  $\mathcal{D}: \mathbb{K}_n[t] \rightarrow \mathbb{K}_n[t]$  da base  $\beta = \{1, t, \dots, t^n\}$  para a base  $\beta$ .
- Fixada uma permutação  $\sigma: \{1, 2, 3\} \rightarrow \{1, 2, 3\}$ , defina a aplicação  $A: \mathbb{K}^3 \rightarrow \mathbb{K}^3$ ,  $A(x_1, x_2, x_3) = (x_{\sigma(1)}, x_{\sigma(2)}, x_{\sigma(3)})$ .
  - Calcule a representação matricial de  $A$  da base canônica para a base canônica quando  $\sigma(1) = 3$ ,  $\sigma(2) = 1$  e  $\sigma(3) = 2$ .
  - Determine uma base  $\beta$  tal que  $[A]_{\beta}^{\alpha} = [I]$ , onde  $\sigma$  é a permutação acima.

## 2.6 Teorema da representação matricial

Retornemos ao estudo de transformações com o auxílio de matrizes. O leitor perceberá a grande proximidade entre os dois conceitos e os lucros desta associação.

**Teorema 2.6.1 (Teorema da representação matricial)** *Suponha que  $V$  e  $W$  sejam dois espaços vetoriais sobre o corpo  $\mathbb{K}$  e que  $\beta$  e  $\gamma$  sejam bases ordenadas de  $V$  e  $W$ , respectivamente, com  $\#\beta = n$  e  $\#\gamma = m$ . Então a aplicação*

$$\Psi: \mathcal{L}(V, W) \rightarrow M(m \times n, \mathbb{K}), \quad \Psi(A) = [A]_{\gamma}^{\beta},$$

*é um isomorfismo linear. Em particular,  $\dim \mathcal{L}(V, W) = \dim V \cdot \dim W$ .*

**Demonstração** Explicitemos as bases envolvidas,

$$\beta = \{v_1, v_2, \dots, v_n\}, \quad \gamma = \{w_1, w_2, \dots, w_m\}$$

e demonstremos inicialmente a linearidade de  $\Psi$ . Escolhidas quaisquer duas transformações lineares  $A, B \in \mathcal{L}(V, W)$ , escrevamos as combinações lineares

$$\begin{aligned} A(v_j) &= a_{1j}w_1 + a_{2j}w_2 + \dots + a_{mj}w_m, \\ B(v_j) &= b_{1j}w_1 + b_{2j}w_2 + \dots + b_{mj}w_m. \end{aligned}$$

Como sabemos, as representações matriciais são  $[A]_\gamma^\beta = [a_{ij}]$  e  $[B]_\gamma^\beta = [b_{ij}]$ . Dado um escalar  $\lambda \in \mathbb{K}$  é imediato avaliarmos  $A + \lambda B$  no vetor  $v_j \in \beta$ ,

$$(A + \lambda B)(v_j) = (a_{1j} + \lambda b_{1j})w_1 + (a_{2j} + \lambda b_{2j})w_2 + \cdots + (a_{mj} + \lambda b_{mj})w_m.$$

Logo, vale a identidade matricial  $[A + \lambda B]_\gamma^\beta = [a_{ij} + \lambda b_{ij}]$ . Daí seguem as igualdades

$$[A + \lambda B]_\gamma^\beta = [a_{ij} + \lambda b_{ij}] = [a_{ij}] + \lambda [b_{ij}] = [A]_\gamma^\beta + \lambda [B]_\gamma^\beta,$$

isto é,  $\Psi(A + \lambda B) = \Psi(A) + \lambda \Psi(B)$ , mostrando a linearidade de  $\Psi$ .

A injetividade será demonstrada examinando-se o núcleo de  $\Psi$ . Afirmar que  $A \in \text{Nuc } \Psi$ , significa afirmar que todas as entradas da representação matricial  $[A]_\gamma^\beta = [a_{ij}]$  são iguais a zero,  $a_{ij} = 0$ . Isto implica que

$$A(v_j) = a_{1j}w_1 + a_{2j}w_2 + \cdots + a_{mj}w_m = 0, \quad j = 1, \dots, n.$$

Como a aplicação identicamente nula também assume o valor nulo em todos os vetores da base  $\beta$ , pela primeira proposição da seção anterior,  $A : V \rightarrow V$  deve ser a transformação identicamente nula. Logo  $\text{Nuc } \Psi = \{o\}$ , mostrando a injetividade.

Provemos a sobrejetividade de  $\Psi$ . Dada uma matriz  $N = [a_{ij}] \in M(m \times n, \mathbb{K})$ , considere a única transformação linear  $A : V \rightarrow W$  tal que

$$A(v_j) = a_{1j}w_1 + a_{2j}w_2 + \cdots + a_{mj}w_m.$$

Esta construção nos garante que  $\Psi(A) = N$ .

Finalmente, recordamos que isomorfismos lineares preservam dimensões, por isso,  $\dim \mathcal{L}(V, W) = \dim M(m \times n, \mathbb{K})$ , e como a dimensão do espaço das matrizes  $m \times n$  é igual ao produto  $\dim V \cdot \dim W$ , concluímos a demonstração.  $\square$

Note que a representação  $\Psi : \mathcal{L}(V, W) \rightarrow M(m \times n, \mathbb{K})$ ,  $\Psi(A) = [A]_\gamma^\beta$ , depende das bases ordenadas escolhidas. Outro fato importante sobre a matriz  $[A]_\gamma^\beta$  é a possibilidade de relacionarmos as coordenadas de um vetor  $v \in V$  na base ordenada  $\beta \subset V$  com as coordenadas do vetor  $A(v) \in W$  na base ordenada  $\gamma \subset W$  através de um produto matricial. Esse é o significado do

**Lema 2.6.1** *Sejam  $\beta$  e  $\gamma$  duas bases ordenadas dos espaços vetoriais de dimensão finita  $V$  e  $W$ , respectivamente. Se  $A : V \rightarrow W$  é uma transformação linear então  $[A(v)]_\gamma = [A]_\gamma^\beta [v]_\beta$ .*

**Demonstração** Escrevamos  $\beta = \{v_1, v_2, \dots, v_n\}$ ,  $\gamma = \{w_1, w_2, \dots, w_m\}$ , e assumamos que a matriz das coordenadas do vetor  $v$  e a representação matricial de  $A$  sejam

$$[v]_\beta = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad [A]_\gamma^\beta = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Avaliando a transformação no vetor  $v = x_1v_1 + x_2v_2 + \cdots + x_nv_n$  obtemos,

$$\begin{aligned} A(v) &= x_1A(v_1) + x_2A(v_2) + \cdots + x_nA(v_n) \\ &= x_1(a_{11}w_1 + a_{21}w_2 + \cdots + a_{m1}w_m) + \cdots + \\ &\quad x_n(a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{mn}w_m) \\ &= (x_1a_{11} + x_2a_{12} + \cdots + x_na_{1n})w_1 + \cdots + \\ &\quad (x_1a_{m1} + x_2a_{m2} + \cdots + x_na_{mn})w_m. \end{aligned}$$

Pela definição de matriz das coordenadas de um vetor na base  $\gamma$ , podemos escrever,

$$[A(v)]_\gamma = \begin{bmatrix} x_1a_{11} + x_2a_{12} + \cdots + x_na_{1n} \\ x_1a_{21} + x_2a_{22} + \cdots + x_na_{2n} \\ \dots \\ x_1a_{m1} + x_2a_{m2} + \cdots + x_na_{mn} \end{bmatrix}.$$

Sem muito esforço verificamos que a identidade matricial acima pode ser reescrita como  $[A(v)]_\gamma = [A]_\gamma^\beta [v]_\beta$ .  $\square$

**Exemplo 2.6.1** Consideremos a transformação linear

$$A : \mathbb{R}^2 \rightarrow \mathbb{R}^3, \quad A(x, y) = (x - 2y, x, y - x),$$

e as bases ordenadas  $\beta = \{(1, 1), (1, -1)\}$  de  $\mathbb{R}^2$  e  $\gamma = \{(1, 0, 1), (1, 1, 0), (0, 1, 1)\}$  de  $\mathbb{R}^3$ . Calculemos a representação matricial de  $A$  nessas bases,

$$\begin{aligned} A(1, 1) &= (-1, 1, 0) = -1(1, 0, 1) + 0(1, 1, 0) + 1(0, 1, 1), \\ A(1, -1) &= (3, 1, -2) = 2(1, 0, 1) + 1(1, 1, 0) + 0(0, 1, 1). \end{aligned}$$

Sendo assim, vale a representação matricial

$$[A]_\gamma^\beta = \begin{bmatrix} -1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

A matriz das coordenadas do vetor  $v = (2, 0) \in \mathbb{R}^2$  na base  $\beta$  é

$$[v]_\beta = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

portanto, pelo lema anterior, a matriz das coordenadas do vetor  $A(v) \in \mathbb{R}^3$  na base  $\gamma$  é calculado pela fórmula

$$[A(v)]_\gamma = [A]_\gamma^\beta [v]_\beta = \begin{bmatrix} -1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

De onde recuperamos o vetor  $A(v)$  pela combinação linear

$$A(v) = 1(1, 0, 1) + 1(1, 1, 0) + 1(0, 1, 1) = (2, 2, 2).$$

Note que tanto o cálculo das representações matriciais quanto o cálculo das coordenadas de um vetor são feitos resolvendo-se sistemas lineares.  $\square$

Neste momento da teoria é natural indagar qual o algoritmo para acharmos a representação matricial de uma composta de transformações lineares. Veremos no corolário a seguir que existe uma elegante fórmula relacionando uma composta de transformações com um produto de matrizes.

**Corolário 2.6.1** *Sejam  $\beta$ ,  $\gamma$  e  $\eta$  bases ordenadas dos espaços vetoriais de dimensão finita sobre o mesmo corpo,  $V, W$  e  $Y$ , respectivamente. Se  $A \in \mathcal{L}(V, W)$  e  $B \in \mathcal{L}(W, Y)$  então*

$$[B \circ A]_{\eta}^{\beta} = [B]_{\eta}^{\gamma} [A]_{\gamma}^{\beta}.$$

**Demonstração** O lema anterior permite-nos escrever as igualdades matriciais

$$\begin{aligned} [B \circ A(v)]_{\eta} &= [B \circ A]_{\eta}^{\beta} [v]_{\beta}, \\ [B \circ A(v)]_{\eta} &= [B(A(v))]_{\eta} = [B]_{\eta}^{\gamma} [A(v)]_{\gamma} = [B]_{\eta}^{\gamma} [A]_{\gamma}^{\beta} [v]_{\beta}, \end{aligned}$$

para qualquer  $v \in V$ . Destas duas relações segue a identidade

$$\left( [B \circ A]_{\eta}^{\beta} - [B]_{\eta}^{\gamma} [A]_{\gamma}^{\beta} \right) [v]_{\beta} = [0]_{\eta}$$

para todo vetor  $v \in V$ . Isto implica facilmente que  $[B \circ A]_{\eta}^{\beta} = [B]_{\eta}^{\gamma} [A]_{\gamma}^{\beta}$ .  $\square$

### Exercícios propostos 2.6.1

- Calcule a representação matricial nas bases ordenadas indicadas em cada item para a transformação linear  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ ,  $A(x, y) = (x + 2y, y, 2x - 3y)$ .
  - Nas bases canônicas  $\alpha_2$  e  $\alpha_3$ .
  - De  $\alpha_2$  para  $\gamma = \{(1, 0, 2), (2, 1, -3), (0, 0, 1)\} \subset \mathbb{R}^3$ .
  - De  $\beta = \{(1, 1), (1, -1)\} \subset \mathbb{R}^2$  para a base  $\alpha_3$ .
  - De  $\beta = \{(1, 1), (1, -1)\} \subset \mathbb{R}^2$  para  $\gamma = \{(1, 0, 2), (2, 1, -3), (0, 0, 1)\} \subset \mathbb{R}^3$ .
- Seja  $D : \mathbb{K}_3[t] \rightarrow \mathbb{K}_3[t]$  o operador linear obtido por derivação de polinômios. Determine a representação matricial de  $[D]_{\alpha}^{\alpha}$  em que  $\alpha = \{1, t, t^2, t^3\}$ . Qual a representação matricial nessa base para  $D^4$  (a composta de  $D$  quatro vezes)?
- Considere a base ordenada  $\beta = \{v_1, v_2, v_3\}$  do  $\mathbb{R}^3$ , onde  $v_1 = (1, 1, 1)$ ,  $v_2 = (1, 0, -1)$  e  $v_3 = (0, 1, -1)$ . Seja  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  o operador linear  $A(v) = x_1 v_1$ , se  $v = x_1 v_1 + x_2 v_2 + x_3 v_3$ . Calcule as representações matriciais  $[A]_{\alpha_3}^{\beta}$ ,  $[A]_{\beta}^{\beta}$ ,  $[A]_{\beta}^{\alpha_3}$  e  $[A]_{\alpha_3}^{\alpha_3}$ .

4. Sejam  $\beta$  uma base ordenada do espaço vetorial de dimensão finita  $V$  sobre o corpo  $\mathbb{K}$  e  $\lambda_0$  um escalar em  $\mathbb{K}$ . Calcule a representação matricial da homotetia  $A : V \rightarrow V$ ,  $A(v) = \lambda_0 v$  da base  $\beta$  para a base  $\beta$ .
5. Suponha que o operador linear  $A : \mathbb{K}^3 \rightarrow \mathbb{K}^3$  é não nulo e que  $A^2 \equiv 0$ .
- Prove que  $\text{posto}(A) = 1$ .
  - Determine uma base ordenada tal que a representação de  $A$  nessa base seja a matriz indicada ao lado.
 
$$[A]_{\beta}^{\beta} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ a & 0 & 0 \end{bmatrix}.$$
6. Descreva a representação matricial da aplicação identidade de um espaço vetorial de dimensão finita de uma base ordenada para a mesma base com outra ordem. Qual é a representação da identidade quando não trocamos a ordem?
7. Suponha que  $N \in M(m \times n, \mathbb{K})$  é uma matriz tal que  $NP = [0]$  para qualquer matriz coluna  $n \times 1$ ,  $P$ , com entradas no corpo  $\mathbb{K}$ . Demonstre que  $N = [0]$ .

## 2.7 Representação de isomorfismo

Utilizaremos o Teorema da representação matricial para estudar isomorfismos. Como vimos, espaços isomorfos possuem a mesma dimensão, portanto eles são representados por matrizes quadradas. O resultado principal dessa seção garante que a matriz de representação de um isomorfismo é sempre invertível quaisquer que sejam as bases com as quais estamos trabalhando e existe uma relação simples entre a matriz de representação de um isomorfismo e a matriz de representação da inversa do isomorfismo. Antes, deixemos um exercício. Indicaremos por  $I_n \in M(n, \mathbb{K})$  a matriz identidade  $n \times n$ .

**Exercício 2.7.1** Se  $A : V \rightarrow V$  é um operador linear num espaço vetorial de dimensão  $n$ , prove que  $[A]_{\beta}^{\beta} = I_n \Leftrightarrow A = Id_V$ , em que  $\beta$  é uma base de  $V$ .  $\square$

**Corolário 2.7.1** *Sejam  $A : V \rightarrow W$  uma transformação linear entre espaços vetoriais de dimensão finita e  $\beta$  e  $\gamma$  bases ordenadas de  $V$  e  $W$ , respectivamente. Então  $A$  é um isomorfismo se, e somente se, a representação  $[A]_{\gamma}^{\beta}$  é uma matriz invertível. Em particular, se  $A$  é um isomorfismo então*

$$[A^{-1}]_{\beta}^{\gamma} = [A]_{\gamma}^{\beta^{-1}}.$$

**Demonstração**  $\Rightarrow$ ) Assuma que  $n = \dim V$ . Sendo assim, podemos escrever as identidades matriciais

$$I_n = [Id_V]_{\beta}^{\beta}, \quad I_n = [Id_W]_{\gamma}^{\gamma},$$

pois sendo  $A$  um isomorfismo, temos que  $\dim V = \dim W$ . Pelas propriedades da representação de uma composta de transformações temos as igualdades

$$I_n = [Id_V]_\beta^\beta = [A^{-1} \circ A]_\beta^\beta = [A^{-1}]_\beta^\gamma [A]_\gamma^\beta,$$

$$I_n = [Id_W]_\gamma^\gamma = [A \circ A^{-1}]_\gamma^\gamma = [A]_\gamma^\beta [A^{-1}]_\beta^\gamma.$$

Em particular,  $[A^{-1}]_\beta^\gamma$  é invertível e sua inversa é  $([A]_\gamma^\beta)$ .

$\Leftarrow$ ) Vamos assumir que  $[A]_\gamma^\beta$  é uma matriz invertível e que  $P = [b_{ij}]$  é a sua matriz inversa. Como uma matriz invertível é quadrada concluímos que  $\dim V = \dim W$ . Neste caso  $\beta = \{v_1, v_2, \dots, v_n\}$  e  $\gamma = \{w_1, w_2, \dots, w_n\}$  são bases com o mesmo número de elementos. Consideremos a transformação  $B : W \rightarrow V$  cujos valores nos vetores da base  $\gamma$  são

$$B(w_j) = b_{1j}v_1 + b_{2j}v_2 + \dots + b_{nj}v_n.$$

Deste modo, necessariamente temos a representação matricial  $[B]_\beta^\gamma = [P]$ . Recorrendo-se às propriedades da representação de uma composta de transformações lineares podemos escrever

$$[B \circ A]_\beta^\beta = [B]_\beta^\gamma [A]_\gamma^\beta = P [A]_\gamma^\beta = [I]_n = [Id_V]_\beta^\beta,$$

$$[A \circ B]_\gamma^\gamma = [A]_\gamma^\beta [B]_\beta^\gamma = [A]_\gamma^\beta P = I_n = [Id_W]_\gamma^\gamma.$$

A unicidade de representação implica que  $A \circ B = Id_W$  e  $B \circ A = Id_V$ . Como  $A$  é invertível, por definição,  $A$  é um isomorfismo.  $\square$

Recordamos que uma matriz quadrada é invertível se, e somente se, o seu determinante é não nulo. Caso o leitor deseje conhecer a demonstração deste fato, indicamos como referência o capítulo sobre matrizes.

### Exercícios propostos 2.7.1

1. Suponha que  $A : \mathbb{K}^2 \rightarrow \mathbb{K}^2$ ,  $A(x, y) = (ax + cy, bx + dy)$  é um isomorfismo. Prove que  $ad - bc \neq 0$  e que

$$[A^{-1}]_{\alpha_2}^{\alpha_2} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

2. Verifique que  $A : \mathbb{K}^3 \rightarrow \mathbb{K}^3$ , é um isomorfismo e calcule a sua inversa quando

a)  $A(x, y, z) = (x + y + z, 3x + 4y + 3z, 3x + 3y + 4z)$ .

b)  $A(x, y, z) = (2x + y + z, x + 2z, 3x + y + 2z)$ .

3. Seja  $A : V \rightarrow W$  é um isomorfismo linear entre espaços vetoriais de dimensão finita. Prove: se  $V = V_1 \oplus V_2$  então  $W = A(V_1) \oplus A(V_2)$ .
4. Demonstre que qualquer isomorfismo entre espaços vetoriais de dimensão finita pode ser representado pela matriz identidade.
5. Seja  $[A]_{\beta}^{\beta} = [a_{ij}]$  uma representação matricial de um isomorfismo linear  $A$  num espaço vetorial  $V$  de dimensão  $n$ . Prove que se

$$\|a_{ii}\| > \sum_{i \neq j} \|a_{ij}\|,$$

então  $A$  é invertível (*Teorema de Hadamard*).

## 2.8 Matriz mudança de coordenadas

Fixemos duas bases  $\beta$  e  $\gamma$  de um mesmo espaço vetorial  $V$  de dimensão  $n$ . A *matriz mudança de coordenadas* de  $\beta$  para  $\gamma$  é a matriz  $[Id]_{\gamma}^{\beta}$  onde  $Id : V \rightarrow V$  é a aplicação identidade. Esta terminologia é apropriada pois conhecendo-se as coordenadas de um vetor  $v$  na base  $\beta$  e a matriz mudança de coordenadas de  $\beta$  para  $\gamma$  é possível determinar as coordenadas do vetor  $v$  na base  $\gamma$ , para isso, basta utilizar as relações

$$[v]_{\gamma} = [Id(v)]_{\gamma} = [Id]_{\gamma}^{\beta} [v]_{\beta}.$$

Note que  $[Id]_{\gamma}^{\beta}$  é uma matriz quadrada  $n \times n$  em que  $n = \dim V$ .

**Lema 2.8.1** A matriz  $[Id]_{\gamma}^{\beta}$  é invertível e  $[Id]_{\gamma}^{\beta^{-1}} = [Id]_{\beta}^{\gamma}$ .

**Demonstração** A demonstração segue do corolário da seção anterior, pois a aplicação  $Id : V \rightarrow V$  é um isomorfismo e a sua inversa é ela própria.  $\square$

**Exemplo 2.8.1** Os conjuntos

$$\beta = \{(\cos \theta, \operatorname{sen} \theta), (-\operatorname{sen} \theta, \cos \theta)\} \quad \text{e} \quad \alpha = \{(1, 0), (0, 1)\}$$

são bases ordenadas de  $\mathbb{R}^2$ . Calculemos as matrizes mudança de coordenadas  $[Id]_{\alpha}^{\beta}$  e  $[Id]_{\beta}^{\alpha}$ . A primeira delas tem um cálculo bastante simples pois

$$\begin{aligned} (\cos \theta, \operatorname{sen} \theta) &= \cos \theta (1, 0) + \operatorname{sen} \theta (0, 1), \\ (-\operatorname{sen} \theta, \cos \theta) &= -\operatorname{sen} \theta (1, 0) + \cos \theta (0, 1). \end{aligned}$$

Daí obtemos a representação

$$[Id]_{\alpha}^{\beta} = \begin{bmatrix} \cos \theta & -\operatorname{sen} \theta \\ \operatorname{sen} \theta & \cos \theta \end{bmatrix}.$$

Pela relação  $[Id]_{\alpha}^{\beta} = ([Id]_{\beta}^{\alpha})^{-1}$  e com a ajuda da adjunta clássica obtemos

$$[Id]_{\alpha}^{\beta} = \begin{bmatrix} \cos \theta & \operatorname{sen} \theta \\ -\operatorname{sen} \theta & \cos \theta \end{bmatrix}. \quad \square$$

Aa matrizes de mudança de coordenadas permite determinarmos todas as representações matriciais de uma transformação conhecendo-se apenas uma delas.

**Proposição 2.8.1** *Sejam  $A : V \rightarrow W$  uma transformação linear entre espaços vetoriais de dimensão finita e  $\beta_o$  e  $\gamma_o$  bases ordenadas de  $V$  e  $W$  respectivamente. Se  $[A]_{\gamma_o}^{\beta_o}$  é a representação matricial de  $A$  nessas bases então a representação matricial de  $A$  nas bases ordenadas  $\beta_1$  e  $\gamma_1$  de  $V$  e  $W$ , respectivamente, é*

$$[A]_{\gamma_1}^{\beta_1} = [Id_W]_{\gamma_1}^{\gamma_o} [A]_{\gamma_o}^{\beta_o} [Id_V]_{\beta_o}^{\beta_1}.$$

**Demonstração** Por tudo que já foi visto sobre representação matricial de uma composta de transformações lineares, podemos escrever,

$$[A]_{\gamma_1}^{\beta_1} = [Id_W \circ A \circ Id_V]_{\gamma_1}^{\beta_1} = [Id_W]_{\gamma_1}^{\gamma_o} [A \circ Id_V]_{\gamma_o}^{\beta_1} = [Id_W]_{\gamma_1}^{\gamma_o} [A]_{\gamma_o}^{\beta_o} [Id_V]_{\beta_o}^{\beta_1}.$$

Isto termina a demonstração da proposição.  $\square$

Em capítulos posteriores, estudaremos exclusivamente operadores lineares num espaço vetorial de dimensão finita,  $A : V \rightarrow V$ . Com muita frequência utilizaremos representações matriciais dos operadores, que, recordamos, são matrizes quadradas. Neste caso, para aliviar a notação indicaremos por  $[A]_{\beta}$  a representação matricial do operador linear  $A$  da base ordenada  $\beta$  para a mesma base, em lugar de utilizar a notação  $[A]_{\beta}^{\beta}$ . O resultado abaixo sobre *conjugação* de matrizes terá um bom proveito.

Em tempo, diz-se que  $N_0, N_1 \in M(n, \mathbb{K})$  são duas matrizes conjugadas quando existe uma matriz invertível  $R \in M(n, \mathbb{K})$  tal que  $N_1 = R N_0 R^{-1}$ . A demonstração do corolário seguinte fica sob a responsabilidade do leitor.

**Corolário 2.8.1** *Sejam  $A$  um operador linear no espaço vetorial  $V$  de dimensão finita e  $\beta$  e  $\delta$  duas bases ordenadas de  $V$ . Então as representações matriciais  $[A]_{\beta}$  e  $[A]_{\delta}$  são conjugadas, mais precisamente,*

$$[A]_{\delta} = [Id_V]_{\delta}^{\beta} [A]_{\beta} [Id_V]_{\beta}^{\delta}.$$

**Exemplo 2.8.2** O conjunto ordenado  $\beta = \{v_1, v_2, v_3\}$  é uma base ordenada de  $\mathbb{R}^3$ , onde  $v_1 = (1, 1, 1)$ ,  $v_2 = (3, 1, 0)$  e  $v_3 = (1, 0, -1)$ . Calculemos a representação matricial nas bases canônicas de  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,

$$A(v) = x_2 v_1 - (x_1 + x_3) v_2 \quad \text{em que} \quad v = x_1 v_1 + x_2 v_2 + x_3 v_3.$$

Para isso, temos o corolário acima. É fácil calcular a representação na base  $\beta$  pois

$$\begin{aligned} A(v_1) &= 0v_1 - 1v_2 + 0v_3, \\ A(v_2) &= 1v_1 + 0v_2 + 0v_3, \\ A(v_3) &= 0v_1 + 0v_2 + 0v_3. \end{aligned} \quad \implies \quad [A]_\beta = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

A representação matricial da identidade da base  $\beta$  para a base canônica  $\alpha$  também é fácil determinar, pois

$$\begin{aligned} Id(v_1) &= (1, 1, 1) = 1e_1 + 1e_2 + 1e_3, \\ Id(v_2) &= (3, 1, 0) = 3e_1 + 1e_2 + 0e_3, \\ Id(v_3) &= (1, 0, -1) = 1e_1 + 0e_2 - 1e_3, \end{aligned} \quad \implies \quad [Id]_\alpha^\beta = \begin{bmatrix} 1 & 3 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix}.$$

Sabemos também que  $[Id]_\beta^\alpha = ([Id]_\alpha^\beta)^{-1}$ . Com a ajuda da adjunta clássica encontramos a representação

$$[Id]_\beta^\alpha = \begin{bmatrix} -1 & 3 & -1 \\ 1 & -2 & 1 \\ 1 & 3 & -2 \end{bmatrix}.$$

Retornando à relação  $[A]_\alpha = [Id_V]_\alpha^\beta [A]_\beta [Id_V]_\beta^\alpha$  e efetuando o produto das matrizes obtemos a representação desejada,

$$[A]_\alpha = \begin{bmatrix} 7 & -20 & 10 \\ 3 & -8 & 4 \\ 1 & 2 & 1 \end{bmatrix}.$$

Note que a representação matricial do operador na base canônica,

$$[A]_\alpha \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 7x - 20y + 10z \\ 3x - 8y + 4z \\ x + 2y + z \end{bmatrix},$$

nos diz imediatamente que o operador linear  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  é expresso como

$$A(x, y, z) = (7x - 20y + 10z, 3x - 8y + 4z, x + 2y + z). \quad \square$$

### Exercícios propostos 2.8.1

- Dada a base ordenada  $\beta = \{v_1, v_2, v_3\}$  de  $\mathbb{K}^3$ , em que  $v_1 = (1, 1, 1)$ ,  $v_2 = (1, 0, -1)$  e  $v_3 = (0, 1, -1)$ . Dê a representação matricial do operador na base canônica.
  - $A : \mathbb{K}^3 \rightarrow \mathbb{K}^3$ ,  $A(v) = x_1v_1$ , em que  $v = x_1v_1 + x_2v_2 + x_3v_3$ .
  - $A : \mathbb{K}^3 \rightarrow \mathbb{K}^3$ ,  $A(v) = x_3v_1 + (x_2 - x_1)v_3$ , em que  $v = x_1v_1 + x_2v_2 + x_3v_3$ .
  - $A \circ B : \mathbb{K}^3 \rightarrow \mathbb{K}^3$ , em que  $A$  e  $B$  são os operadores dos itens anteriores.
  - $B \circ A : \mathbb{K}^3 \rightarrow \mathbb{K}^3$ , em que  $A$  e  $B$  são os operadores dos itens a) e b).
- Fixado um subespaço  $U \subset V$ , verifique que o subconjunto a seguir é um subespaço e calcule sua dimensão,  $\mathcal{N}(U) = \{A \in \mathcal{L}(V, W); \text{ tal que } A(u) = 0 \text{ para todo } u \in U\}$ .

## 2.9 Subespaços invariantes

Diz-se que um subespaço  $W \subset V$  é um *subespaço invariante* por um operador linear  $A : V \rightarrow V$  quando  $A(W) \subset W$ .

Exemplos de subespaços invariantes que o leitor verifica facilmente são  $Nuc A$  e  $Im A$ . Diremos  $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$  é uma *soma direta invariante* pelo operador  $A$ , quando cada parcela da decomposição é um subespaço invariante. Um dos objetivos principais deste texto é construir somas diretas invariantes por um operador linear em um espaço vetorial de dimensão finita. Isto será feito a partir do Capítulo 4, no momento estamos interessados em saber como a propriedade de ser um subespaço invariante é constatada em linguagem matricial. Para isto, consideramos uma base ordenada  $\beta = \{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$  do espaço vetorial  $V$  na qual  $\beta_1 = \{v_1, v_2, \dots, v_r\}$  é uma base ordenada para  $W$ , o subespaço invariante pelo operador  $A : V \rightarrow V$ . Garantimos pela invariância de  $W$  que cada  $v_j \in \beta_1$  tem uma imagem  $A(v_j) \in W$ , portanto

$$A(v_j) = a_{1j}v_1 + a_{2j}v_2 + \cdots + a_{rj}v_r, \quad 1 \leq j \leq r.$$

Sendo assim, o operador linear é representado esquematicamente como

$$[A]_\beta = \begin{bmatrix} [A_1]_{\beta_1} & N_1 \\ [0] & N_2 \end{bmatrix},$$

onde  $[A_1]_{\beta_1}$  é a representação matricial do operador  $A_1 : W \rightarrow W$  induzido por restrição de  $A$ , enquanto  $N_1$  e  $N_2$  são matrizes retangulares. O símbolo  $[0]$  indica uma matriz retangular com todas as entradas nulas. Reciprocamente, quando uma representação matricial de um operador  $A : V \rightarrow V$  numa base ordenada  $\beta = \{v_1, v_2, \dots, v_n\}$  tem a forma

$$[A]_\beta = \begin{bmatrix} R & N_1 \\ [0] & N_2 \end{bmatrix},$$

em que  $R$  é uma matriz quadrada  $r \times r$ , está claramente registrado que

$$A(v_j) = a_{1j}v_1 + a_{2j}v_2 + \cdots + a_{rj}v_r, \quad 1 \leq j \leq r,$$

implicando imediatamente que o subespaço  $W \subset V$  gerado por  $\beta_1 = \{v_1, v_2, \dots, v_r\} \subset \beta$  é um subespaço invariante por  $A$  e vale a identidade matricial  $[A_1]_{\beta_1} = [R]$ , onde  $A_1 : W \rightarrow W$  é o operador linear induzido por restrição de  $A$ .

Examinemos a situação na qual  $V$  admite uma soma direta invariante,

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k, \quad A(V_i) \subset V_i.$$

Selecionando-se para cada  $i$ ,  $1 \leq i \leq k$ , uma base ordenada  $\beta_i$  do subespaço  $V_i$ , construímos para  $V$  a base ordenada  $\beta = \vec{\cup} \beta_i$  e pelos mesmos argumentos apresentados anteriormente podemos representar o operador linear  $A : V \rightarrow V$  na

base ordenada  $\beta$  esquematicamente como

$$[A]_{\beta} = \begin{bmatrix} [A_1]_{\beta_1} & [0] & \cdot & [0] \\ [0] & [A_2]_{\beta_2} & \cdot & [0] \\ \cdot & \cdot & \cdot & \cdot \\ [0] & [0] & \cdot & [A_k]_{\beta_k} \end{bmatrix}.$$

Por economia de espaço, algumas vezes escrevemos esta última matriz em uma forma compacta, tal como

$$[A]_{\beta} = \text{diag}\{[A_1]_{\beta_1}, [A_2]_{\beta_2}, \dots, [A_k]_{\beta_k}\},$$

onde para cada  $i$  a matriz  $[A_i]_{\beta_i}$  é a representação matricial na base ordenada  $\beta_i$  do operador linear  $A_i : V_i \rightarrow V_i$  induzido por  $A$ . As recíprocas destes fatos são triviais e não as faremos. Para finalizar, ressaltamos que o determinante daquela última matriz  $[A]_{\beta}$  é calculado por uma regra bastante simples, a saber,

$$\det[A]_{\beta} = \det[A_1]_{\beta_1} \det[A_2]_{\beta_2} \cdots \det[A_k]_{\beta_k}.$$

### Exercícios propostos 2.9.1

Em todos os exercícios  $V$  denotará um espaço vetorial de dimensão finita.

- Determine uma base para cada espaço invariante (não trivial) pelo operador linear  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  quando
  - $A(x, y) = (x - y, 2x - 2y)$ ;
  - $A(x, y) = (x + y, y)$ ;
  - $A(x, y) = (2x + y, x + y)$ .
- Sejam  $A : V \rightarrow V$  um isomorfismo e  $W \subset V$  um subespaço invariante.
  - Prove que o operador obtido por restrição  $A_0 : W \rightarrow W$  é um isomorfismo.
  - Demonstre que  $W$  é também invariante por  $A^{-1}$ .
- Prove: se dois operadores lineares,  $A, B : V \rightarrow V$ , preservam um subespaço  $W \subset V$  então a soma  $A + B$  e a composta  $A \circ B$  também preservam este mesmo subespaço.
- Demonstre que se dois operadores  $A, B : V \rightarrow V$  comutam e  $A$  preserva um subespaço  $W \subset V$ , então  $B(W)$  é também invariante por  $A$ .
- Se um operador linear  $A : V \rightarrow V$  preserva os subespaços  $W_1$  e  $W_2$ , então  $A$  preserva os subespaços  $W_1 + W_2$  e  $W_1 \cap W_2$ . Prove essa afirmação.
- Fixado um subespaço  $W \subset V$ . Mostre que  $\mathcal{I}(W) = \{A \in \mathcal{L}(V, V); A(W) \subset W\}$  é um subespaço de  $\mathcal{L}(V, V)$  e calcule a sua dimensão em função de  $\dim V$  e  $\dim W$ .
- Dada uma soma direta  $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$ . Qual a dimensão do subespaço de  $\mathcal{L}(V, V)$  formado pelos operadores que preservam a soma direta?
- Suponha que o operador linear  $A : V \rightarrow V$  preserva o subespaço  $W \subset V$ . Mostre que a aplicação no espaço quociente definida por  $\bar{A} : V/W \rightarrow V/W$ ,  $\bar{A}(\bar{v}) = \overline{A(v)}$  está bem definida e é um operador linear.

## Capítulo 3

# Espaço dual

No capítulo anterior contruímos vários espaços vetoriais associados a um espaço vetorial  $V$ , por exemplo,  $\mathcal{L}(V, W)$  onde  $W$  é um outro espaço. Apresentaremos a seguir o espaço dual associado à  $V$ , cujo objetivo maior é demonstrar o Teorema do posto de uma matriz.

### 3.1 Funcionais lineares

Seja  $V$  um espaço vetorial sobre  $\mathbb{K}$ . Diz-se que uma aplicação  $f : V \rightarrow \mathbb{K}$  é chamada de *funcional linear* em  $V$  se possui as propriedades:

$$\text{fl1} \quad f(v + w) = f(v) + f(w),$$

$$\text{fl2} \quad f(\lambda v) = \lambda f(v),$$

para quaisquer  $v, w \in V$  e  $\lambda \in \mathbb{K}$ .

Formalmente não podemos afirmar que um funcional linear é uma transformação linear pois o contradomínio é um corpo e não um espaço vetorial. Entretanto, considerando-se o corpo  $\mathbb{K}$  como um espaço unidimensional sobre si mesmo é natural aquela identificação. Na nossa notação,  $\mathbb{K}^1 \simeq \mathbb{K}$ . Com esta observação podemos transpor sem dificuldade várias informações já demonstradas no capítulo anterior para este contexto, observando que  $\alpha_1 = \{(1)\}$  é a base canônica de  $\mathbb{K}$ .

O *espaço dual* de um espaço vetorial  $V$  sobre o corpo  $\mathbb{K}$  é o conjunto de todos os funcionais lineares  $f : V \rightarrow \mathbb{K}$ , conjunto esse que será denotado por  $V^*$ . As usuais operações de adição de funções e multiplicação de uma função por um escalar de  $\mathbb{K}$ , induzem uma estrutura de espaço vetorial em  $V^*$  e com a identificação acima mostramos imediatamente que  $V^*$  é isomorfo ao espaço  $\mathcal{L}(V, \mathbb{K})$ . Logo, se  $V$  é um espaço vetorial de dimensão finita e  $\beta = \{v_1, v_2, \dots, v_n\} \subset V$  é uma base ordenada, a aplicação

$$\Phi : V^* \rightarrow M(1 \times n, \mathbb{K}), \quad \Phi(f) = [f]_{\alpha_1}^{\beta}$$

é um isomorfismo linear. Portanto,  $\dim V^* = n$ . Recordamos que

$$[f]_{\alpha_1}^{\beta} = [ f(v_1) \quad f(v_2) \quad \cdots \quad f(v_n) ],$$

pois a  $j$ -ésima coluna de  $[f]_{\alpha_1}^{\beta}$  é formada pela coordenada do vetor  $f(v_j)$  na base  $\alpha_1 = \{(1)\}$ . Por convenção, a base  $\alpha_1$  não é explicitada na representação matricial do funcional linear, simplificando a notação,  $[f]_{\alpha_1}^{\beta} = [f]^{\beta}$ . É fácil verificar que

$$f(v) = [f(v)]_{\alpha_1} = [f]^{\beta}[v]_{\beta}.$$

Note que estamos identificando o escalar no primeiro membro da igualdade com as matrizes  $1 \times 1$  nos outros membros.

**Exemplo 3.1.1** 1) A aplicação  $f : \mathbb{K}^2 \rightarrow \mathbb{K}$ ,  $f(x, y) = 2x - 3y$  é um funcional linear em  $\mathbb{K}^2$ . Como

$$f(x, y) = xf(e_1) + yf(e_2) = 2x - 3y,$$

podemos descrever o funcional na forma matricial

$$f(x, y) = [f]^{\alpha_2} [(x, y)]_{\alpha_2} = [ 2 \quad -3 ] \begin{bmatrix} x \\ y \end{bmatrix}.$$

2) O exemplo acima é facilmente generalizado. Fixados  $n$  escalares  $a_1, a_2, \dots, a_n \in \mathbb{K}$  é possível contruir um funcional linear em  $\mathbb{K}^n$  definindo

$$f : \mathbb{K}^n \rightarrow \mathbb{K}, \quad f(x_1, x_2, \dots, x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n.$$

Nesse caso a representação matricial de  $f$  na base canônica é a matriz  $1 \times n$ ,

$$[f]^{\alpha_n} = [ a_1 \quad a_2 \quad \cdots \quad a_n ]. \quad \square$$

**Exercício 3.1.1** Prove as afirmações.

1. Uma aplicação  $f : V \rightarrow \mathbb{K}$  é um funcional linear se, e somente se,  $f(v + \lambda w) = f(v) + \lambda f(w)$  para quaisquer  $v, w \in V$  e  $\lambda \in \mathbb{K}$ .
2. Um funcional linear  $f : V \rightarrow \mathbb{K}$  ou é a aplicação identicamente nula ou é uma aplicação sobrejetora. □

**Exemplo 3.1.2** Reconstruiremos um exemplo importante e clássico de um funcional linear. O *traço* de uma matriz quadrada  $N = [a_{ij}] \in M(n, \mathbb{K})$  é o escalar

$$\text{tr}(N) = a_{11} + a_{22} + \cdots + a_{nn}.$$

É imediato verificarmos que o traço define um funcional linear  $\text{tr} : M(n, \mathbb{K}) \rightarrow \mathbb{K}$  cuja principal propriedade é transcrita pela fórmula  $\text{tr}(N \cdot P) = \text{tr}(P \cdot N)$ , para

quaisquer matrizes  $N = [a_{ij}]$ ,  $P = [b_{ij}] \in M(n, \mathbb{K})$ . A demonstração deste fato é uma manipulação de índices. Vejamos,

$$\operatorname{tr}(N \cdot P) = \sum_{i=1}^n (N \cdot P)_{ii} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_{ji} = \sum_{j=1}^n \sum_{i=1}^n b_{ji} a_{ij} = \operatorname{tr}(P \cdot N).$$

Daí concluímos que o traço é um funcional linear invariante por conjugação, explicando melhor,  $\operatorname{tr}(R \cdot N \cdot R^{-1}) = \operatorname{tr}(N)$ , para qualquer matriz invertível  $R \in M(n, \mathbb{K})$ . Esta propriedade permite definir o traço de um operador linear num espaço vetorial  $V$  de dimensão finita sobre o corpo  $\mathbb{K}$ . Para isso, fixamos uma base ordenada  $\beta \subset V$  e consideramos a aplicação

$$\operatorname{tr} : \mathcal{L}(V, V) \rightarrow \mathbb{K}, \quad \operatorname{tr}(A) = \operatorname{tr}([A]_{\beta}).$$

De fato, a aplicação é um funcional linear e não depende da base ordenada escolhida pois qualquer outra representação é conjugada a essa representação. Relembramos a relação de conjugação envolvida,

$$[A]_{\gamma} = [Id_V]_{\gamma}^{\beta} [A]_{\beta} [Id_V]_{\beta}^{\gamma}, \quad \text{em que} \quad [Id_V]_{\gamma}^{\beta} = [Id_V]_{\beta}^{\gamma^{-1}}. \quad \square$$

**Exercício 3.1.2** Sejam  $V$  e  $W$  dois espaços vetoriais sobre  $\mathbb{K}$  e  $\gamma = \{w_1, \dots, w_m\} \subset W$  uma base ordenada.

1. Prove que se  $\{f_i : V \rightarrow \mathbb{K}\}_{i=1}^m$  é uma coleção de  $m$  funcionais lineares então a aplicação  $A : V \rightarrow W$ , definida por  $A(v) = f_1(v)w_1 + f_2(v)w_2 + \dots + f_m(v)w_m$  é uma transformação linear.
2. Enuncie e demonstre a recíproca da afirmação anterior. □

### Exercícios propostos 3.1.1

1. Determine a representação matricial do funcional linear  $f : \mathbb{K}^3 \rightarrow \mathbb{K}$ ,  $f(x, y, z) = x + 2y - 3z$ , na base ordenada indicada.
  - a) Base canônica de  $\mathbb{K}^3$ .
  - b)  $\beta = \{(3, 0, 0), (2, 2, 0), (1, 1, 1)\} \subset \mathbb{K}^3$ .
2. Construa um funcional linear  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  cujo núcleo é gerado pelo conjunto de vetores  $\varepsilon = \{(2, 2, -1), (0, 3, 2)\}$ .
3. Qual a dimensão do núcleo de um funcional linear  $f : V \rightarrow \mathbb{K}$  quando  $\dim V = n$ ?
4. Sejam  $V$  um espaço vetorial de dimensão finita sobre  $\mathbb{K}$  e  $f : V \rightarrow \mathbb{K}$  um funcional linear. Mostre que existe uma base ordenada  $\beta \subset V$  na qual a representação de  $f$  é da forma  $[f]_{\beta} = \begin{bmatrix} a & 0 & \dots & 0 \end{bmatrix}$ .
5. Fixada uma matriz  $N_0 \in M(n, \mathbb{K})$ , considere o operador linear

$$A : M(n, \mathbb{K}) \rightarrow M(n, \mathbb{K}), \quad A(P) = N_0 P - P N_0.$$

- a) Demonstre que  $A$  não é sobrejivo e que  
 b) se  $A(P) = P$ , então  $P$  não é invertível.
6. Seja  $V$  um espaço vetorial de dimensão finita sobre  $\mathbb{K}$ . Prove que  $V^*$  separa pontos. Isto significa que dados dois vetores distintos  $v, w \in V$  existe um funcional linear  $f : V \rightarrow \mathbb{K}$  tal que  $f(v) \neq f(w)$ .
7. Denote por  $\mathcal{C}^0([a, b], \mathbb{R})$  o espaço vetorial das funções contínuas com valores em  $\mathbb{R}$  definidas no intervalo  $[a, b]$ . Verifique quais das aplicações são funcionais lineares.
- a)  $\mathfrak{S} : \mathcal{C}^0([a, b], \mathbb{R}) \rightarrow \mathbb{R}$ ,  $\mathfrak{S}(f) = \int_a^b f(t)dt$  (integral de Riemann);  
 b)  $\mathcal{L} : \mathcal{C}^0([a, b], \mathbb{R}) \rightarrow \mathbb{R}$ ,  $\mathcal{L}(f) = f(a)$  (avaliação em  $a$ ).
8. Seja  $N \in M(m \times n, \mathbb{R})$ . Prove que  $\text{tr}(NN^t) \geq 0$  e que  $\text{tr}(NN^t) = 0 \Leftrightarrow N = [0]$ .
9. Seja  $N \in M(m \times n, \mathbb{C})$ . Prove que  $\text{tr}(N\overline{[N]^t}) \geq 0$  e que  $\text{tr}(N\overline{[N]^t}) = 0 \Leftrightarrow N = [0]$ .
10. Verifique a seguinte relação entre duas matrizes  $N, P \in M(2, \mathbb{K})$ ,
- $$NP + PN = N \text{tr}(P) + P \text{tr}(N) + (\text{tr}(NP) - \text{tr}(N)\text{tr}(P))I.$$
11. Fixado uma decomposição em soma direta de um espaço vetorial de dimensão finita,  $V = W_1 \oplus W_2 \oplus \dots \oplus W_k$ , prove que qualquer operador linear  $A$  em  $V$  satisfazendo a condição  $A(W_i) \subset W_1 \oplus \dots \oplus \hat{W}_i \oplus \dots \oplus W_k$  tem traço nulo ( $\hat{W}_i$  indica que essa parcela foi omitida).

## 3.2 Base dual

Vamos assumir que  $\beta = \{v_1, v_2, \dots, v_n\}$  seja uma base ordenada de um espaço vetorial  $V$  sobre o corpo  $\mathbb{K}$ . Por um lado, o isomorfismo entre  $\mathcal{L}(V, \mathbb{K})$  e  $V^*$ , garante que  $\dim V = n = \dim V^*$  e por outro lado, a Proposição 3.1 do Capítulo 2 afirma que para cada  $i$ ,  $1 \leq i \leq n$ , existe um, e somente um, funcional linear  $f_i : V \rightarrow \mathbb{K}$  satisfazendo as condições  $f_i(v_j) = \delta_{ij}$ , onde  $\delta_{ij}$  é o delta de Kronecker,

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

Estes dois fatos justificam o

**Lema 3.2.1** *O conjunto ordenado  $\beta^* = \{f_1, f_2, \dots, f_n\} \subset V^*$  é uma base ordenada, chamada de base dual de  $\beta$ .*

**Demonstração** Vamos supor que existam escalares  $a_1, a_2, \dots, a_n \in \mathbb{K}$  tais que a combinação linear dos elementos da base dual  $\beta^*$  com esses coeficientes seja o funcional linear identicamente nulo,

$$0 \equiv a_1 f_1 + a_2 f_2 + \dots + a_n f_n.$$

Avaliando esta combinação linear no vetor  $v_j \in \beta$  obtemos

$$\begin{aligned} 0 &= a_1 f_1(v_j) + a_2 f_2(v_j) + \cdots + a_n f_n(v_j) \\ &= a_1 \delta_{1j} + a_2 \delta_{2j} + \cdots + a_n \delta_{nj} \\ &= a_j, \end{aligned}$$

mostrando que  $\beta^*$  é um subconjunto com  $n$  vetores linearmente independente num espaço de dimensão  $n$ . Portanto,  $\beta^*$  é uma base.  $\square$

**Exercício 3.2.1** Demonstre que qualquer funcional linear  $f \in V^*$  é uma combinação linear dos elementos de  $\beta^*$  cujos coeficientes são os escalares  $\{f(v_i)\}_{i=1}^n \subset \mathbb{K}$ , isto é,

$$f \equiv f(v_1)f_1 + f(v_2)f_2 + \cdots + f(v_n)f_n.$$

Compare os coeficientes com as entradas da matriz de  $f$  na base ordenada  $\beta$ .  $\square$

Resumiremos os comentários acima na

**Proposição 3.2.1** *Se  $V$  é um espaço vetorial de dimensão finita, então o espaço dual  $V^*$  tem dimensão finita e  $\dim V = \dim V^*$ .*

### Exercícios propostos 3.2.1

- Determine a base dual das seguintes bases ordenadas.
  - $\alpha = \{(1, 0), (0, 1)\} \subset \mathbb{C}^2$ .
  - $\gamma = \{(1, 0, 0), (1, 1, 0), (1, 1, 1)\} \subset \mathbb{R}^3$ .
  - $\beta = \{(2, 1), (1, 1)\} \subset \mathbb{R}^2$ .
  - $\delta = \{(1, 1), (-1, 1)\} \subset \mathbb{R}^2$ .
- Calcule as coordenadas do funcional  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ ,  $f(x, y, z) = 3x - 2y - z$ ,
  - na base dual da base canônica;
  - na base dual da base ordenada  $\gamma = \{(1, 0, 0), (1, 1, 0), (1, 1, 1)\} \subset \mathbb{R}^3$ .
- Seja  $V$  um espaço vetorial de dimensão finita. Qual a dimensão de  $(V^*)^*$ ?
- Mostre que todo subespaço  $W$  de dimensão  $n - 1$  de um espaço vetorial  $V$  de dimensão  $n$  é um núcleo de um funcional linear  $f : V \rightarrow \mathbb{K}$ .
- Um subespaço  $W$  de um espaço vetorial  $V$  com  $\dim W = r < n = \dim V$  é a interseção de  $n - r$  subespaços de dimensão  $n - 1$ . Prove essa afirmação.
- Suponha que  $f$  e  $g$  são dois funcionais lineares num espaço vetorial  $V$  de dimensão finita sobre  $\mathbb{K}$  satisfazendo a condição  $\text{Nuc } f \subset \text{Nuc } g$ . Prove que  $g(v) = \lambda f(v)$ , para algum escalar  $\lambda \in \mathbb{K}$  e para todo  $v \in V$ .
- Dado um espaço vetorial de dimensão finita  $V$ , considere a aplicação  $\mathcal{A} : V \rightarrow (V^*)^*$ , onde  $\mathcal{A}[v] : V^* \rightarrow \mathbb{K}$  é a aplicação avaliação em  $v$ , mais precisamente,  $\mathcal{A}[v](f) = f(v)$ . Prove que  $\mathcal{A}$  é um isomorfismo linear entre  $V$  e  $(V^*)^*$ .

### 3.3 A adjunta

Uma transformação linear  $A : V \rightarrow W$  entre espaços vetoriais sobre o mesmo corpo  $\mathbb{K}$ , induz uma aplicação entre os espaços duais  $A^* : W^* \rightarrow V^*$ , chamada de *adjunta* de  $A$  cuja definição utiliza uma composição de aplicações da seguinte forma: um funcional linear  $f \in W^*$  é associado ao funcional linear

$$A^*[f] : V \rightarrow \mathbb{K}, \quad A^*[f](v) = f(A(v)).$$

Esquemáticamente,  $A^*[f]$  é a composta  $V \xrightarrow{A} W \xrightarrow{f} \mathbb{K}$ . O leitor verificará facilmente que  $A^*[f]$  é de fato um funcional linear. Observamos que o conceito de adjunta apresentado aqui não tem relação alguma com o conceito de adjunta clássica utilizado no algoritmo para inverter matrizes.

**Lema 3.3.1** *A adjunta  $A^* : W^* \rightarrow V^*$  é uma transformação linear.*

**Demonstração** Com efeito. Sejam  $f, g \in W^*$  e  $\lambda \in \mathbb{K}$ . Para qualquer vetor  $v \in V$  temos que

$$\begin{aligned} A^*[f + \lambda g](v) &= (f + \lambda g)(A(v)) \\ &= f(A(v)) + \lambda g(A(v)) \\ &= A^*[f](v) + \lambda A^*[g](v). \end{aligned}$$

Isto significa que  $A^*[f + \lambda g] = A^*[f] + \lambda A^*[g]$ , como queríamos demonstrar.  $\square$

**Proposição 3.3.1** *Sejam  $A : V \rightarrow W$  uma transformação linear entre espaços vetoriais de dimensão finita,  $\beta$  e  $\gamma$  bases ordenadas de  $V$  e  $W$ , respectivamente. Então a representação matricial  $[A^*]_{\beta^*}^{\gamma^*}$  da adjunta  $A^* : W^* \rightarrow V^*$  é a matriz transposta de  $[A]_{\gamma}^{\beta}$ . Além disso, cada transformação linear  $A : V \rightarrow W$  define uma única adjunta.*

**Demonstração** Antes de tudo, explicitemos as bases ordenadas envolvidas

$$\left\{ \begin{array}{l} \beta = \{v_1, v_2, \dots, v_n\} \\ \beta^* = \{f_1, f_2, \dots, f_n\} \end{array} \right\}, \quad \left\{ \begin{array}{l} \gamma = \{u_1, u_2, \dots, u_m\} \\ \gamma^* = \{g_1, g_2, \dots, g_m\} \end{array} \right\}$$

, e as representações matriciais  $[A]_{\gamma}^{\beta} = [a_{ij}]$  e  $[A^*]_{\beta^*}^{\gamma^*} = [b_{ij}]$ . Por definição de representação matricial sabemos que valem as igualdades

$$\begin{aligned} A(v_i) &= a_{1i}u_1 + a_{2i}u_2 + \dots + a_{mi}u_m, \\ A^*[g_j] &= b_{1j}f_1 + b_{2j}f_2 + \dots + b_{nj}f_n. \end{aligned}$$

Por um lado, a definição de adjunta e de base dual permite escrever

$$A^*[g_j](v_i) = g_j(A(v_i)) = a_{1i}g_j(u_1) + a_{2i}g_j(u_2) + \cdots + a_{mi}g_j(u_m) = a_{ji}.$$

Por outro lado, temos que

$$A^*[g_j](v_i) = b_{1j}f_1(v_i) + b_{2j}f_2(v_i) + \cdots + b_{nj}f_n(v_i) = b_{ij}.$$

Isto mostra que  $b_{ij} = a_{ji}$ , em outras palavras,  $[A^*]_{\beta^*}^{\gamma^*} = [A]_{\gamma}^{\beta^t}$ . Finalmente, para mostrar a unicidade da adjunta consideremos o isomorfismo linear de representação

$$\Psi : \mathcal{L}(W^*, V^*) \rightarrow M(n \times m, \mathbb{K}), \quad \Psi(B) = [B]_{\beta^*}^{\gamma^*}.$$

A adjunta  $A^*$  é o único elemento de  $\mathcal{L}(W^*, V^*)$  representado pela matriz  $[A]_{\gamma}^{\beta^t}$ .  $\square$

### Exercícios propostos 3.3.1

1. Sejam  $A$  e  $B$  transformações lineares de  $V$  em  $W$  e  $\lambda$  um escalar. Mostre que  $(A + \lambda B)^* = A^* + \lambda B^*$ .
2. Mostre que o traço está no núcleo da adjunta do operador linear  $A : M(n, \mathbb{K}) \rightarrow M(n, \mathbb{K})$ ,  $A(X) = N_0 X - X N_0$ , em que  $N_0 \in M(n, \mathbb{K})$  é uma matriz fixada.
3. Considere a aplicação  $f : \mathbb{R}_3[t] \rightarrow \mathbb{R}$ ,

$$f(p(t)) = \int_0^1 p(t) dt.$$

- a) Verifique que  $f$  é um funcional linear e calcule a representação na base dual de  $\beta = \{1, t, t^2, t^3\}$ .
- b) Mostre que  $A : \mathbb{R}_3[t] \rightarrow \mathbb{R}_3[t]$ ,  $A(p(t)) = p'(t) - p'(t-1) + p'(t+1)$  é um operador linear e calcule  $A^*[f](p(t))$ .

## 3.4 Perpendicular de um conjunto

Se  $\varepsilon$  é um subconjunto não vazio de um espaço vetorial  $V$  sobre o corpo  $\mathbb{K}$ , o *perpendicular* de  $\varepsilon$  é o subconjunto do espaço dual  $V^*$  definido por

$$\varepsilon^{per} = \{f \in V^*; f(v) = 0 \text{ para todo } v \in \varepsilon\}.$$

Uma verificação de rotina mostra que o perpendicular de um conjunto não vazio é um subespaço do espaço dual. Mais interessantes e importantes são as relações dimensionais que surgem quando consideramos espaços vetoriais de dimensão finita.

**Proposição 3.4.1** *Seja  $W$  o subespaço gerado por um conjunto não vazio  $\varepsilon$  de um espaço vetorial de dimensão finita  $V$ . Então*

$$a) \varepsilon^{per} = W^{per};$$

$$b) \dim W^{per} = \dim V - \dim W.$$

**Demonstração** a) Para cada vetor  $w \in W$  existem vetores  $w_1, w_2, \dots, w_k \in \varepsilon$  e escalares  $x_1, x_2, \dots, x_k \in \mathbb{K}$  tais que  $w = x_1 w_1 + x_2 w_2 + \dots + x_k w_k$ , desde que  $\varepsilon$  é um conjunto de geradores para  $W$ . Assim, dado um funcional linear  $f \in \varepsilon^{per}$  seguem por linearidade e por  $f(w_i) = 0$  para todo  $i$ ,  $1 \leq i \leq k$ , as igualdades

$$f(w) = x_1 f(w_1) + x_2 f(w_2) + \dots + x_k f(w_k) = 0,$$

significando que  $f \in W^{per}$ , logo vale a inclusão  $\varepsilon^{per} \subset W^{per}$ . A inclusão oposta é trivial, pois um funcional linear anulando-se nos vetores de  $W$  necessariamente anula-se nos vetores de  $\varepsilon$ .

b) Considere uma base ordenada  $\beta = \{v_1, v_2, \dots, v_n\}$  do espaço  $V$ , na qual os  $k$  primeiros vetores formam uma base de  $W$  e considere a base dual correspondente  $\beta^* = \{f_1, f_2, \dots, f_n\}$ . Note que os elementos de  $\gamma^* = \{f_{k+1}, f_{k+2}, \dots, f_n\}$  são funcionais anulando-se nos  $k$  primeiros vetores de  $\beta$ , de onde segue imediatamente que  $\gamma^* \subset W^{per}$ . Para demonstrar que este subconjunto é uma base só precisaremos verificar que todo funcional linear de  $W^{per}$  é uma combinação linear dos vetores de  $\gamma^*$ . Mas isto é simples de ser verificado, pois dado  $f \in W^{per}$  temos que

$$\begin{aligned} f &\equiv f(v_1)f_1 + \dots + f(v_k)f_k + f(v_{k+1})f_{k+1} + \dots + f(v_n)f_n \\ &\equiv f(v_{k+1})f_{k+1} + \dots + f(v_n)f_n. \end{aligned}$$

Finalmente, contemos as dimensões,  $\dim W^{per} = \#\gamma = n - k = \dim V - \dim W$ .  $\square$

### Exercícios propostos 3.4.1

1. Mostre que o perpendicular de um conjunto  $\varepsilon \subset V$  é um subespaço do dual  $V^*$ .
2. Calcule uma base para  $\varepsilon^{per}$  quando
  - a)  $\varepsilon = \{(2, 1, -1)\} \subset \mathbb{R}^3$ .
  - b)  $\varepsilon = \{(2, 1, -1), (0, 1, 2)\} \subset \mathbb{R}^3$ .
  - c)  $\varepsilon = \{(i, i, i), (1, 0, 1)\} \subset \mathbb{C}^3$ .
  - d)  $\varepsilon = \{(1, 1, 1), (1, 0, 1)\} \subset \mathbb{R}^3$ .
3. Determine uma base para o  $perp$  do subespaço  $W = \{(x, y, z); x = y = z\} \subset \mathbb{R}^3$ .
4. Se  $U \subset W$  são subespaços de  $V$ , mostre que  $W^{per} \subset U^{per}$ .
5. Sejam  $U$  e  $W$  subespaços do espaço vetorial de dimensão finita  $V$ . Quais dos seguintes subspa cos de  $V^*$  são iguais?

$$(U + W)^{per}, \quad U^{per} + W^{per}, \quad (U \cap W)^{per} \quad \text{e} \quad U^{per} \cap W^{per}.$$

## 3.5 Teorema do posto

Essa seção é toda ela dedicada à demonstração do Teorema do posto de uma matriz. Antes disso, devemos estabelecer a relação entre a imagem e o núcleo de

uma transformação linear  $A : V \rightarrow W$  com o núcleo e a imagem (nessa ordem) de sua adjunta  $A^* : W^* \rightarrow V^*$ . Essa relação é estabelecida precisamente utilizando-se o conceito de perpendicular de um conjunto. Lembramos que o posto de uma transformação linear é a dimensão da sua imagem.

**Proposição 3.5.1** *Seja  $A : V \rightarrow W$  uma transformação linear entre dois espaços vetoriais de dimensão finita. Então*

- a)  $Nuc A^* = (Im A)^{per}$ ;
- b)  $posto(A) = posto(A^*)$ ;
- c)  $Im A^* = (Nuc A)^{per}$ .

**Demonstração** a) O primeiro ítem da proposição segue diretamente da definição de adjunta e de perpendicular pois

$$f \in Nuc A^* \Leftrightarrow A^*[f](v) = f(A(v)) = 0 \text{ para todo } v \in V \Leftrightarrow f \in (Im A)^{per}.$$

b) O Teorema do núcleo e da imagem, o ítem a), a Proposição 4.1 e o isomorfismo entre um espaço e seu dual justificam a sequência de igualdades,

$$\begin{aligned} \dim Im A^* &= \dim W^* - \dim Nuc A^* \\ &= \dim W^* - \dim (Im A)^{per} \\ &= \dim W - (\dim W - \dim Im A) \\ &= \dim Im A. \end{aligned}$$

Portanto, o posto de  $A$  é igual ao posto de sua adjunta.

c) Afirmar que  $f \in Im A^*$  é equivalente a afirmar que existe um funcional linear  $g \in W^*$  tal que  $f(v) = g(A(v))$  para qualquer vetor  $v$  pertencente ao espaço  $V$ . Portanto, se  $v$  é um elemento do núcleo de  $A$  então  $f(v) = 0$ , em outras palavras,  $f \in (Nuc A)^{per}$ , mostrando a inclusão  $Im A^* \subset (Nuc A)^{per}$ . Para mostrar a igualdade dos dois espaços basta verificar que as dimensões são iguais. Com efeito, pelo Teorema do núcleo e da imagem e o ítem a) temos

$$\dim(Nuc A)^{per} = \dim V - \dim Nuc A = \dim Im A = \dim A^*. \quad \square$$

Sugerimos ao leitor rever a definição de posto das linhas e das colunas de uma matriz no Capítulo 12. Observe que o próximo resultado é uma versão matricial do ítem b) da proposição acima.

**Teorema 3.5.1 (Teorema do posto de uma matriz)** *O posto das colunas de uma matriz  $N \in M(m \times n, \mathbb{K})$  é igual ao posto de suas linhas.*

**Demonstração** Sejam  $\alpha_n = \{e_1, e_2, \dots, e_n\}$  e  $\alpha_m = \{d_1, d_2, \dots, d_m\}$  as bases canônicas de  $\mathbb{K}^n$  e  $\mathbb{K}^m$ , respectivamente. Considere a única transformação linear  $A : \mathbb{K}^n \rightarrow \mathbb{K}^m$  cuja representação matricial  $[A]_{\alpha_m}^{\alpha_n}$  é a matriz dada  $N = [a_{ij}] \in M(m \times n, \mathbb{K})$ . Isto significa que temos as combinações lineares

$$\begin{aligned} A(e_1) &= a_{11}d_1 + a_{21}d_2 + \dots + a_{m1}d_m, \\ A(e_2) &= a_{12}d_1 + a_{22}d_2 + \dots + a_{m2}d_m, \\ &\dots \\ A(e_n) &= a_{1n}d_1 + a_{2n}d_2 + \dots + a_{mn}d_m. \end{aligned}$$

Note que o posto de  $A$  é o número máximo de vetores linearmente independentes no conjunto  $A(\alpha_n) = \{A(e_1), A(e_2), \dots, A(e_n)\} \subset \text{Im } A$ . Note também que o  $j$ -ésimo elemento de  $A(\alpha_n)$  é uma combinação linear dos vetores  $A(e_{i_1}), A(e_{i_2}), \dots, A(e_{i_r})$  se, e somente se, a  $j$ -ésima coluna  $N^j$  é uma combinação linear das colunas  $N^{i_1}, N^{i_2}, \dots, N^{i_r}$ . Logo, o posto da transformação linear  $A$  é igual ao posto das colunas de sua representação matricial  $N$ !

Consideremos a adjunta  $A^* : \mathbb{K}^{m*} \rightarrow \mathbb{K}^{n*}$  e as respectivas bases duais  $\alpha_m^*$  e  $\alpha_n^*$ . Como sabemos,  $[A^*]_{\alpha_n^*}^{\alpha_m^*} = N^t$ . Pelos mesmos motivos anteriores o posto de  $A^*$  é o posto das colunas da matriz  $N^t$ . Como o posto de  $A$  e o posto de  $A^*$  são iguais seguem as igualdades

$$\begin{aligned} \text{posto das colunas de } N &= \text{posto}(A) \\ &= \text{posto}(A^*) \\ &= \text{posto das colunas de } N^t \\ &= \text{posto das linhas de } N. \end{aligned}$$

Isto termina a demonstração do teorema. □

### Exercícios propostos 3.5.1

1. Calcule o posto das seguintes matrizes com entradas reais.

$$a) N = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 2 & 1 & 2 \end{bmatrix}. \quad b) N = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & -6 & -2 & -1 \\ 3 & -2 & 0 & 1 \end{bmatrix}.$$

2. Qual o posto de uma matriz  $n \times n$  invertível?
3. Prove que o posto de uma matriz quadrada é invariante por conjugação.
4. Mostre que a nulidade de  $A : V \rightarrow W$  é igual a nulidade de sua adjunta se, e somente se,  $\dim V = \dim W$ .
5. Seja  $A : V \rightarrow W$  uma transformação linear. Prove as afirmações.
  - a)  $A$  é sobrejetora  $\Leftrightarrow A^t$  é injetora.
  - b)  $A$  é injetora  $\Leftrightarrow A^t$  é sobrejetora.

## Capítulo 4

# Operadores e polinômios

Os próximos sete capítulos serão dedicados ao estudo de operadores lineares num espaço vetorial de dimensão finita. O objetivo do presente capítulo é construir os dois principais polinômios associados a um operador linear, polinômios minimal e característico, e estabelecer as relações entre eles. Posteriormente faremos largo uso das informações obtidas.

### 4.1 Polinômio minimal

Na seqüência,  $A$  denotará um operador linear num espaço vetorial  $V$  de dimensão finita sobre o corpo  $\mathbb{K}$ . Iniciaremos apresentando uma construção envolvendo polinômios e operadores lineares. Para cada polinômio

$$p(t) = a_r t^r + \cdots + a_1 t + a_0, \quad p(t) \in \mathbb{K}[t],$$

podemos construir um novo operador linear definindo

$$p(A) : V \rightarrow V, \quad p(A) = a_r A^r + \cdots + a_1 A + a_0 Id.$$

Obviamente  $A^i$  indica a composta do operador  $A$  um número  $i$  de vezes e  $Id$  indica o operador identidade em  $V$ . A mesma construção pode ser repetida para matrizes quadradas. Para  $N \in M(n, \mathbb{K})$  definimos

$$p(N) = a_n N^n + \cdots + a_1 N + a_0 I.$$

Nesse caso,  $N^i$  indica a  $i$ -ésima potência de  $N$  e  $I$  indica a matriz identidade  $n \times n$ . A relação entre as duas construções é estabelecida através do Teorema da representação matricial.

**Lema 4.1.1** *Sejam  $\beta$  uma base ordenada do espaço vetorial  $V$  de dimensão  $n$  e  $A : V \rightarrow V$  um operador linear. Seja  $R \in M(n, \mathbb{K})$  uma matriz invertível. Então para cada polinômio  $p(t) \in \mathbb{K}[t]$  valem as relações matriciais:*

- a)  $[p(A)]_\beta = p([A]_\beta)$ ;  
 b)  $p(R[A]_\beta R^{-1}) = R p([A]_\beta) R^{-1}$ .

**Demonstração** a) Assumindo que  $p(t) = a_r t^r + \dots + a_1 t + a_0$ , avaliando  $p(t)$  no operador linear e calculando a representação matricial da avaliação, obtemos

$$\begin{aligned} [p(A)]_\beta &= [a_r A^r + \dots + a_1 A + a_0 Id]_\beta \\ &= a_r [A^r]_\beta + \dots + a_1 [A]_\beta + a_0 [Id]_\beta \\ &= a_r [A]_\beta^r + \dots + a_1 [A]_\beta + a_0 [I] \\ &= p([A]_\beta). \end{aligned}$$

b) Deixaremos como exercício sugerindo a utilização da identidade matricial  $(RNR^{-1})^r = RN^r R^{-1}$  para todo inteiro  $r \geq 0$ .  $\square$

O restante da seção é dedicado ao estudo de um conjunto de polinômios associado a um operador linear.

**Definição 4.1.1** *Sejam  $V$  um espaço vetorial sobre o corpo  $\mathbb{K}$  e  $A : V \rightarrow V$  um operador linear. O ideal anulador de  $A$  é o conjunto de polinômios*

$$\mathfrak{S}_A = \{p(t) \in \mathbb{K}[t]; p(A) \equiv 0\}.$$

*O ideal anulador de uma matriz  $N \in M(n, \mathbb{K})$  é o conjunto de polinômios*

$$\mathfrak{S}_N = \{p(t) \in \mathbb{K}[t]; p(N) \equiv [0]\}.$$

Dois conseqüências do lema anterior são imediatas. Se  $[A]_\beta$  é a representação matricial do operador linear na base  $\beta \subset V$  então  $\mathfrak{S}_A = \mathfrak{S}_{[A]_\beta}$  pois

$$p(A) \equiv 0 \Leftrightarrow [p(A)]_\beta \equiv [0] \Leftrightarrow p([A]_\beta) \equiv [0].$$

O item b) do mesmo lema implica que o ideal anulador de uma matriz é invariante por conjugação, isto é,  $\mathfrak{S}_N = \mathfrak{S}_{RNR^{-1}}$  para qualquer matriz invertível  $R \in M(n, \mathbb{K})$ . É claro que o ideal anulador de  $A$  não é vazio pois contém o polinômio identicamente nulo. Podemos afirmar mais,

**Proposição 4.1.1** *O ideal anulador  $\mathfrak{S}_A$  é não trivial.*

**Demonstração** Vamos supor que  $\dim V = n$ . O conjunto  $\varepsilon = \{Id, A, A^2, \dots, A^{n^2}\}$  é um conjunto linearmente dependente de  $\mathcal{L}(V, V)$  pois esse espaço tem dimensão  $n^2$ . Portanto, existem escalares  $a_0, a_1, \dots, a_{n^2} \in \mathbb{K}$  não todos nulos tais que

$$a_{n^2} A^{n^2} + \dots + a_1 A + a_0 Id \equiv 0.$$

Logo, o polinômio  $p(t) = a_{n2}t^{n^2} + \dots + a_1t + a_0$  não é o polinômio identicamente nulo e pertence ao ideal anulador  $\mathfrak{S}_A$ .  $\square$

Devemos justificar o nome *ideal* utilizado para designar o conjunto  $\mathfrak{S}_A$ .

**Proposição 4.1.2**  $\mathfrak{S}_A$  é um ideal de  $\mathbb{K}[t]$ , isto é,

- a)  $\mathfrak{S}_A$  é um grupo comutativo com a operação de soma de polinômios;
- b) Se  $p(t) \in \mathfrak{S}_A$  e  $q(t) \in \mathbb{K}[t]$ , então o produto de polinômios  $p(t)q(t) \in \mathfrak{S}_A$ .

**Demonstração** Embora simples, faremos a demonstração do item b) para observar um detalhe na notação utilizada. Dados  $p(t) \in \mathfrak{S}_A$  e  $q(t) \in \mathbb{K}[t]$ , o produto  $p(t)q(t)$  quando avaliado no operador  $A : V \rightarrow V$  torna-se uma composta de operadores lineares, a saber,

$$p(A) \circ q(A) : V \rightarrow V.$$

Como por hipótese,  $p(A) \equiv 0$  segue que  $p(t)q(t) \in \mathfrak{S}_A$ .  $\square$

Apresentaremos a seguir um resultado básico para o restante da Teoria. Antes recordamos que um polinômio  $p(t) \in \mathbb{K}[t]$  é dito *mônico* quando o coeficiente do termo de maior grau tem o valor 1.

**Teorema 4.1.1 (Teorema do polinômio minimal)** *Seja  $\mathfrak{S}_A$  o ideal anulador de um operador linear  $A$  no espaço vetorial  $V$  de dimensão finita sobre  $\mathbb{K}$ . Então existe um único polinômio mônico  $m_A(t) \in \mathfrak{S}_A$ , chamado de polinômio minimal de  $A$ , tal que  $\mathfrak{S}_A = m_A(t)\mathbb{K}[t]$ .*

**Demonstração** Note que o único polinômio constante pertencente ao ideal anulador é o polinômio identicamente nulo. Seja  $m_A(t) \in \mathfrak{S}_A$  um polinômio mônico não nulo de menor grau dentre todos do ideal anulador. Pelo comentário acima garantimos que  $\text{grau } m_A(t) \geq 1$ . Como  $\mathfrak{S}_A$  é um ideal, justificamos a inclusão  $m_A(t)\mathbb{K}[t] \subset \mathfrak{S}_A$ , restando mostrar a inclusão oposta. Considere qualquer polinômio  $q(t) \in \mathfrak{S}_A$ , lembramos que por escolha  $\text{grau } q(t) \geq \text{grau } m_A(t)$ . Pelo algoritmo da divisão de Euclides é possível escrever

$$q(t) = m_A(t)p(t) + r(t) \quad \text{em que } p(t), r(t) \in \mathbb{K}[t]$$

com  $\text{grau } r(t) < \text{grau } m_A(t)$ . Novamente, como  $\mathfrak{S}_A$  é um ideal, o polinômio  $r(t) = q(t) - m_A(t)p(t)$  pertence ao ideal. Pela minimalidade do grau de  $m_A(t)$  concluímos que  $r(t) \equiv 0$ , implicando que  $q(t) = m_A(t)p(t) \in m_A(t)\mathbb{K}[t]$ , como desejávamos demonstrar. A demonstração da unicidade do polinômio minimal é trivial.  $\square$

Esse teorema é um resultado de álgebra de polinômios. Para ressaltar esse fato deixamos um exercício que será utilizado outras vezes.

**Exercício 4.1.1** Se  $\mathfrak{S} \subset \mathbb{K}[t]$  é um ideal não trivial, prove que existe um único polinômio mônico  $m(t) \in \mathfrak{S}$  tal que  $\mathfrak{S} = m(t)\mathbb{K}[t]$ . Esse polinômio será chamado de polinômio minimal (ou gerador) do ideal  $\mathfrak{S}$ .  $\square$

### Exercícios propostos 4.1.1

1. Dados o operador linear  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $A(x, y) = (x + y, x - 3y)$  e o polinômio  $p(t) = t^3 - 2t + 1$ , avalie  $p(A)$ .
2. Verifique que o polinômio  $p(t) = t^2 + 1$  pertence ao ideal anulador de  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $A(x, y) = (y, -x)$ . Conclua que  $p(t)$  é de fato o polinômio minimal de  $A$ .
3. Mostre que o gerador do ideal anulador de  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $A(x, y) = (2x + 4y, -x - 2y)$  é o polinômio  $m_A(t) = t^2$ .
4. Suponha que  $\mathfrak{S}_1$  e  $\mathfrak{S}_2$  são dois ideais de  $\mathbb{K}[t]$  tais que  $\mathfrak{S}_1 \subset \mathfrak{S}_2$ . Prove que o polinômio minimal do segundo ideal divide o polinômio minimal do primeiro.
5. Seja  $A : V \rightarrow V$  um operador linear. Prove as afirmações.
  - a)  $A$  comuta com  $p(A)$ , para qualquer  $p(t) \in \mathbb{K}[t]$ .
  - b)  $W \subset V$  é um subespaço invariante por  $A \Leftrightarrow W$  é invariante por  $p(A)$ , para qualquer  $p(t) \in \mathbb{K}[t]$ . Em particular mostre que  $\text{Nuc } p(A)$  é invariante por  $A$ .
  - c) O conjunto dos operadores lineares em  $V$  que comutam com  $A$  formam um subespaço de  $\mathcal{L}(V, V)$  cuja dimensão é maior ou igual ao grau do polinômio minimal de  $A$ .
6. Qual o polinômio minimal do operador identicamente nulo em  $V$ ?

## 4.2 Polinômio característico

O Teorema do polinômio minimal não fornece algoritmo algum para o cálculo de  $m_A(t)$ . Essa dificuldade é parcialmente contornada construindo-se explicitamente um polinômio  $p_A(t) \in \mathfrak{S}_A$ , chamado de *polinômio característico*, que além de ser dividido por  $m_A(t)$ , contém todos os fatores primos do polinômio minimal, facilitando os cálculos em várias situações.

O polinômio característico de uma matriz  $N \in M(n, \mathbb{K})$  é o polinômio mônico de grau  $n$  definido por  $p_N(t) = \det(tI - N)$ , onde  $\det$  está indicando o determinante e  $I$  a matriz identidade  $n \times n$ . Recordamos que se  $R \in M(n, \mathbb{K})$  é uma matriz invertível então  $\det R \neq 0$  e  $\det R^{-1} = (\det R)^{-1}$ . Daí segue que matrizes conjugadas têm os polinômios característicos iguais, pois

$$\det(tI - RNR^{-1}) = \det(R(tI - N)R^{-1}) = \det(tI - N).$$

Essa propriedade permite estender para um operador linear  $A : V \rightarrow V$ , onde  $V$  é um espaço vetorial de dimensão finita sobre  $\mathbb{K}$ , o conceito de polinômio característico. Definimos que  $p_A(t) = p_{[A]_\beta}(t)$ , onde  $[A]_\beta$  é a representação matricial de  $A$  numa base  $\beta$  de  $V$ . Como uma outra representação de  $A$  difere dessa por uma conjugação, o polinômio característico está bem definido. Os comentários acima justificam a notação  $p_A(t) = \det(tI - A)$ .

**Exemplo 4.2.1** A representação matricial na base canônica do operador linear  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $A(x, y) = (x - 2y, x + 3y)$  é a matriz

$$[A]_\alpha = \begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}.$$

Calculemos o polinômio característico de  $A$ ,

$$p_A(t) = p_{[A]_\alpha}(t) = \det(tI - [A]_\alpha) = \det \begin{bmatrix} t-1 & 2 \\ -1 & t-3 \end{bmatrix} = t^2 - 4t + 5.$$

Avaliemos o polinômio na representação de  $A$ ,

$$\begin{aligned} p([A]_\alpha) &= \begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}^2 - 4 \begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix} + 5 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 & -8 \\ 4 & 7 \end{bmatrix} + \begin{bmatrix} -4 & 8 \\ -4 & -12 \end{bmatrix} + \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

Os cálculos acima mostram que o polinômio característico  $p_A(t)$  pertence ao ideal anulador  $\mathfrak{S}_A$ , ilustrando um fato geral que demonstraremos na próxima seção.  $\square$

### Exercícios propostos 4.2.1

1. Calcule o polinômio característico dos seguintes operadores.
  - a)  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (4x + 2y, -x + y, y + 2z)$ .
  - b)  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (x, -x + 2y, 2x - y + 2z)$ .
  - c)  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $A(x, y) = (4x + 8y, -2x - 4y)$ .
2. Identifique as raízes do polinômio característico de uma matriz triangular superior.
3. Mostre que  $p_N(t) = t^2 - (\text{tr } N)t + \det N$  é o polinômio característico da matriz  $2 \times 2$  descrita ao lado.  $N = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ .
4. Calcule o polinômio característico do operador derivação no espaço  $\mathbb{K}_n[t]$ .

5. Qual o polinômio característico de uma involução num espaço de dimensão finita?
6. Mostre que  $p_N(t) = t^3 - ct^2 - bt - a$  é o polinômio característico da matriz descrita ao lado. Defina um operador  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  tal que  $[A]_\beta = N$  onde  $\beta = \{e_2, A(e_2), A^2(e_2)\}$ .
- $$N = \begin{bmatrix} 0 & 0 & a \\ 1 & 0 & b \\ 0 & 1 & c \end{bmatrix}.$$
7. Sejam  $A$  e  $B$  dois operadores lineares num espaço vetorial de dimensão finita  $V$ . Demonstre que se  $A$  é invertível então  $p_{A \circ B}(t) = p_{B \circ A}(t)$ . Esse fato é verdadeiro quando  $A$  e  $B$  são não invertíveis?
8. As aplicações  $\sigma_i : M(n, \mathbb{K}) \rightarrow \mathbb{K}$  definidas pela identidade polinomial
- $$\det(tI - N) = \sigma_n(N)t^n + \sigma_{n-1}(N)t^{n-1} + \cdots + \sigma_1(N)t + \sigma_0(N),$$
- são chamadas de *funções simétricas*.
- Demonstre que  $\sigma_i$ s são invariantes por conjugação e que
  - $\sigma_i(NP) = \sigma_i(PN)$ .
  - Identifique  $\sigma_{n-1}$  e  $\sigma_0$ .
  - Descreva todas as funções simétricas  $\sigma_i : M(3, \mathbb{K}) \rightarrow \mathbb{K}$ .
9. Sejam  $A$  e  $B$  dois operadores lineares no espaço vetorial  $V$  de dimensão dois. Mostre que  $(A \circ B - B \circ A)^2$  comuta com todo elemento de  $\mathcal{L}(V, V)$ .
10. Prove por indução que o polinômio característico de um operador num espaço vetorial de dimensão  $n$  é um polinômio mônico de grau  $n$ . Sugestão: utilize o desenvolvimento de Laplace por linhas para o cálculo de determinantes.

### 4.3 Teorema de Cayley-Hamilton

Para mostrar que o polinômio característico de um operador linear pertence ao seu ideal anulador utilizaremos o conceito de *matriz polinomial*. Uma matriz polinomial é uma matriz do tipo

$$N(t) = [a_{ij}(t)] \quad \text{em que} \quad a_{ij}(t) \in \mathbb{K}[t].$$

Isso significa que as entradas da matriz polinomial  $N(t)$  são polinômios e podem ser reescritas na forma

$$N(t) = N_r t^r + \cdots + N_1 t + N_0,$$

onde  $N_i$  são matrizes quadradas com entradas no corpo  $\mathbb{K}$ . O grau de uma matriz polinomial é o maior grau dentre todos graus das entradas.

**Exemplo 4.3.1** 1) A matriz polinomial  $2 \times 2$  com entradas em  $\mathbb{K}[t]$ ,

$$N(t) = \begin{bmatrix} -t + 1 & t^2 + t - 2 \\ t + 2 & t^2 - t + 3 \end{bmatrix},$$

pode ser apresentada também na forma

$$N(t) = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} t^2 + \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} t + \begin{bmatrix} 1 & -2 \\ 2 & 3 \end{bmatrix}.$$

Nesse caso,  $\text{grau } N(t) = 2$ .

2) Seja  $I$  a matriz identidade  $n \times n$ . Dado o polinômio  $q(t) = a_r t^r + \dots + a_1 t + a_0$  em  $\mathbb{K}[t]$ . Reescrevemos a matriz polinomial  $q(t)I$ , cujo grau é  $r$ ,

$$q(t)I = \begin{bmatrix} q(t) & 0 & 0 \\ 0 & q(t) & 0 \\ 0 & 0 & q(t) \end{bmatrix} = a_r I t^r + \dots + a_1 I t + a_0 I.$$

Observe que o polinômio  $\det(q(t)I) = [q(t)]^n$  tem grau  $rn$ .

3) Já tivemos a oportunidade de utilizar matrizes polinomiais quando definimos o polinômio característico de uma matriz com entradas no corpo  $\mathbb{K}$ . Por exemplo, para calcular o polinômio característico de

$$N_0 = \begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix},$$

consideramos a matriz polinomial  $N(t) = tI - N_0$  e definimos o polinômio característico como sendo  $p_{N_0}(t) = \det N(t) = t^2 - 4t + 5$ .

4) Mais geralmente. Dada a matriz  $N_0 \in M(n, \mathbb{K})$ , consideramos a matriz polinomial  $n \times n$  de grau 1, definida por  $N(t) = tI - N_0$ . A adjunta clássica de  $N(t)$ , aqui denotada por  $\text{adj}(N(t))$ , também é uma matriz polinomial  $n \times n$  e satisfaz a identidade matricial  $N(t) \text{adj}(N(t)) = \det N(t) I$ , ou equivalentemente,  $(tI - N_0) \text{adj}(tI - N_0) = p_{N_0}(t)I$ .  $\square$

Vejamos um lema preparatório para o Teorema de Cayley-Hamilton cuja prova está baseada na apresentação de Birkhoff e MacLane [02].

**Lema 4.3.1** *Seja  $N_0 \in M(n, \mathbb{K})$ . Para cada polinômio  $q(t) \in \mathbb{K}[t]$  existe uma matriz polinomial  $Q(t) = [q_{ij}(t)]$ , satisfazendo a identidade matricial*

$$q(t)I - q(N_0) = (tI - N_0)Q(t).$$

**Demonstração** Se  $q(t) = a_r t^r + \dots + a_1 t + a_0$ , então valem as igualdades

$$\begin{aligned} q(t)I &= a_r t^r I + \dots + a_1 t I + a_0 I, \\ q(N_0) &= a_r N_0^r + \dots + a_1 N_0 + a_0 I. \end{aligned}$$

Por subtração obtemos,

$$q(t)I - q(N_0) = a_r (t^r I - N_0^r) + \dots + a_1 (tI - N_0).$$

Desde que as matrizes  $I$  e  $N_0$  comutam, a matriz polinomial  $tI - N_0$  fatora cada  $tI - N_0^i$ , com  $1 \leq i \leq r$ , do seguinte modo,

$$t^i I - N_0^i = (tI - N_0) (t^{i-1} I + t^{i-2} N_0 + \cdots + t N_0^{i-2} + N_0^{i-1}).$$

Daí podemos facilmente construir a matriz polinomial  $Q(t)$ .  $\square$

**Teorema 4.3.1 (Teorema de Cayley-Hamilton)** *O polinômio característico de uma matriz  $N_0 \in M(n, \mathbb{K})$  pertence ao ideal anulador  $\mathfrak{S}_{N_0}$ .*

**Demonstração** Com efeito, denote por  $p_{N_0}(t)$  o polinômio característico de  $N_0$ . Pelo lema anterior existe uma matriz polinomial  $Q(t) = [q_{ij}(t)]$  tal que

$$p_{N_0}(t)I - p_{N_0}(N_0) = (tI - N_0) Q(t).$$

Por outro lado, considerando a adjunta clássica de  $tI - N_0$  temos a igualdade

$$p_{N_0}(t)I = (tI - N_0) \text{adj}(tI - N_0).$$

Dessas duas equações concluímos por substituição que

$$p_{N_0}(N_0) = (tI - N_0) (\text{adj}(tI - N_0) - Q(t)).$$

Utilizaremos o fato que a matriz  $p_{N_0}(N_0)$  é uma matriz numérica para provar que  $\text{adj}(tI - N_0) - Q(t) \equiv [0]$ , de onde seguirá imediatamente o teorema. Suponha por absurdo que  $\text{adj}(tI - N_0) - Q(t)$  é uma matriz polinomial não nula, digamos

$$\text{adj}(tI - N_0) - Q(t) = P_r t^r + \cdots + P_1 t + P_0,$$

em que  $P_i \in M(n, \mathbb{K})$ ,  $r \geq 0$  e  $P_r$  não é identicamente nula. Sendo assim,

$$(tI - N_0) [\text{adj}(tI - N_0) - Q(t)] = P_r t^{r+1} + \sum (\text{termo com grau} \leq r).$$

Logo,  $p_{N_0}(N_0)$  é uma matriz polinomial com grau  $\geq 1$ , uma contradição.  $\square$

Por tudo que já foi apresentado sobre as relações entre o ideal anulador de um operador linear e o ideal anulador de uma de suas representações matriciais, a prova da seguinte versão do Teorema de Cayley-Hamilton é imediata.

**Corolário 4.3.1** *O polinômio característico de um operador linear num espaço vetorial de dimensão finita pertence ao ideal anulador do operador.*

### Exercícios propostos 4.3.1

1. Calcule o polinômio minimal do operador derivação no espaço dos polinômios  $\mathbb{K}_n[t]$ .
2. Calcule o polinômio minimal de uma involução num espaço de dimensão finita.
3. Demonstre que as matrizes  $N$  e  $P$  não são conjugadas, onde

$$N = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad e \quad P = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}.$$

4. Fixado uma matriz  $N \in M(2, \mathbb{K})$ , considere o operador linear  $A : M(2, \mathbb{K}) \rightarrow M(2, \mathbb{K})$ ,  $A(X) = XN$ . Calcule os polinômios minimal e característico de  $A$  em função de  $m_N(t)$  e  $p_N(t)$ .
5. Mostre que os polinômios mínimos e os polinômios característicos das matrizes quadradas  $N$  e  $N^t$  são iguais.

## 4.4 Sobre a fatoração do polinômio minimal

Para descrever com mais precisão a relação entre os polinômios minimal e característico necessitaremos de alguns conceitos de Álgebra de polinômios. Recordamos que os únicos corpos considerados são os corpos dos números reais e dos complexos.

Um polinômio  $p(t) \in \mathbb{K}[t]$  é *redutível* sobre  $\mathbb{K}$  quando admite uma fatoração

$$p(t) = q(t)r(t)$$

na qual  $q(t), r(t) \in \mathbb{K}[t]$ , *grau*  $q(t) \geq 1$  e *grau*  $r(t) \geq 1$ . Caso contrário  $p(t)$  é dito ser um polinômio *primo*.

Dados os polinômios  $p_1(t), p_2(t), \dots, p_k(t) \in \mathbb{K}[t]$ , o conjunto

$$\mathfrak{S} = p_1(t)\mathbb{K}[t] + p_2(t)\mathbb{K}[t] + \dots + p_k(t)\mathbb{K}[t]$$

é um ideal e seu polinômio minimal é chamado de *máximo divisor comum* de  $p_1(t), p_2(t), \dots, p_k(t)$ . Essa família de polinômios é *relativamente prima* quando o gerador de  $\mathfrak{S}$  é o polinômio constante  $m(t) \equiv 1$ , em outras palavras, quando  $\mathfrak{S} = \mathbb{K}[t]$ .

Uma *fatoração primária*, ou *prima*, de um polinômio  $p(t) \in \mathbb{K}[t]$  é uma fatoração da forma

$$p(t) = a_0 p_1(t)^{r_1} p_2(t)^{r_2} \dots p_k(t)^{r_k},$$

onde  $a_0$  é um escalar em  $\mathbb{K}$ ,  $r_i \geq 1$  são inteiros e  $\{p_1(t), p_2(t), \dots, p_k(t)\}$  é uma coleção de polinômios mônicos, primos e relativamente primos. A menos da ordem dos fatores a fatoração primária é única.

O Teorema Fundamental da Álgebra afirma que todo polinômio em  $\mathbb{C}[t]$  admite uma raiz complexa, de onde segue o fato de que é possível fatorar qualquer polinômio em  $\mathbb{C}[t]$  num produto de polinômios lineares (grau 1). Por outro lado, qualquer polinômio em  $\mathbb{R}[t]$  admite uma fatoração em um produto de polinômios lineares e/ou de grau 2.

Feita essa revisão de Álgebra de polinômios, voltemos ao estudo de operadores lineares num espaço vetorial  $V$  de dimensão finita. Todos os fatores primos do polinômio característico são fatores do polinômio minimal. Registremos essa afirmação na versão matricial.

**Proposição 4.4.1** Se  $p_N(t) = p_1(t)^{s_1} p_2(t)^{s_2} \cdots p_k(t)^{s_k}$  é a fatoração primária do polinômio característico da matriz  $N \in M(n, \mathbb{K})$ , então a fatoração primária do polinômio minimal de  $N$  é da forma

$$m_N(t) = p_1(t)^{r_1} p_2(t)^{r_2} \cdots p_k(t)^{r_k} \quad \text{com } 1 \leq r_i \leq s_i.$$

**Demonstração** Como sabemos existe uma matriz polinomial  $Q(t)$  satisfazendo

$$m_N(t)I - m_N(N) = (tI - N)Q(t).$$

Levando-se em consideração que  $m_N(N) \equiv [0]$ , o cálculo do determinante em ambos os membros da equação nos dá

$$[m_N(t)]^n = p_N(t) \det Q(t).$$

Daí segue que  $p_i(t)$  divide  $[m_N(t)]^n$  para  $i = 1, 2, \dots, k$ , e por conseguinte  $p_i(t)$  divide  $m_N(t)$ . Logo a fatoração primária do polinômio minimal contém uma potência positiva de cada  $p_i(t)$ . Como o Teorema de Cayley-Hamilton garante que  $m_N(t)$  divide o polinômio característico  $p_N(t)$ , podemos afirmar que

$$m_N(t) = p_1(t)^{r_1} p_2(t)^{r_2} \cdots p_k(t)^{r_k} \quad \text{com } 1 \leq r_i \leq s_i. \quad \square$$

**Exemplo 4.4.1** O operador linear  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  cuja representação na base canônica é a matriz

$$[A]_\alpha = \begin{bmatrix} -8 & -4 & -8 \\ 47 & 18 & 39 \\ -8 & -2 & -5 \end{bmatrix}$$

tem polinômio característico  $p_A(t) = (t-1)(t-2)^2$ . Logo, as únicas possibilidades para polinômio minimal de  $A$  são  $m_1(t) = (t-1)(t-2)$  e  $m_2(t) = (t-1)(t-2)^2$ . Verifica-se que  $m_1([A]_\alpha) = [0]$ . Como  $\text{grau } m_1(t) < \text{grau } m_2(t)$  podemos garantir que  $m_A(t) = (t-1)(t-2)$ .  $\square$

#### Exercícios propostos 4.4.1

1. Calcule o polinômio minimal do operador  $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$  quando
  - a)  $A(x, y) = (9x - 16y, 4x - 7y)$ .
  - b)  $A(x, y) = (3y, -x + 4y)$ .
  - c)  $A(x, y) = (4x + 8y, -2x - 4y)$ .
  - d)  $A(x, y, z) = (3x + y - 2z, 2x + 4y - 4z, 2x + y - z)$ .
  - e)  $A(x, y, z) = (4x + 2y, -x + y, y + 2z)$ .
  - f)  $A(x, y, z) = (x, -x + 2y, 2x - y + 2z)$ .
2. Demonstre que se  $z_0 = \lambda + i\tau \in \mathbb{C}$  é uma raiz do polinômio  $p(t)$  com coeficientes reais, então o conjugado  $\bar{z}_0 = \lambda - i\tau$  também é uma raiz de  $p(t)$ .

3. Prove que se  $\lambda_1$  e  $\lambda_2$  são dois escalares distintos de  $\mathbb{K}$ , então  $p(t) = t - \lambda_1$  e  $q(t) = t - \lambda_2$  são polinômios relativamente primos.
4. Suponha que  $p(t) \in \mathbb{R}[t]$  é um polinômio primo e que  $z_0 = \lambda + i\tau$  é uma raiz complexa de  $p(t)$ . Mostre que  $p(t) = (t - \lambda)^2 + \tau^2$ .
5. Calcule o gerador do ideal anulador das seguintes matrizes com entradas reais.

$$a) N = \begin{bmatrix} 4 & -2 \\ 8 & -4 \end{bmatrix}; \quad b) N = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}; \quad c) = \begin{bmatrix} 2 & 0 & 0 & 0 \\ -1 & 1 & 0 & -1 \\ -1 & 0 & 2 & 0 \\ 1 & 1 & 0 & 3 \end{bmatrix}.$$

6. Determine o polinômio característico da matriz com entradas reais descrita ao lado e diga para quais valores de  $\theta$  o polinômio é redutível sobre  $\mathbb{R}$ .  $R_\theta = \begin{bmatrix} \cos \theta & -\operatorname{sen} \theta \\ \operatorname{sen} \theta & \cos \theta \end{bmatrix}$
7. Sejam  $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$  e  $\pi_i : V \rightarrow V$  a projeção sobre  $V_i$  ao longo das outras parcelas. Verifique que  $m_{\pi_i}(t) = t(1 - t)$ .
8. Dados os polinômios  $p_1(t), p_2(t), \dots, p_k(t) \in \mathbb{K}[t]$ , prove  $\mathfrak{S}$  é um ideal, onde 
$$\mathfrak{S} = p_1(t)\mathbb{K}[t] + p_2(t)\mathbb{K}[t] + \cdots + p_k(t)\mathbb{K}[t].$$
9. Um operador linear  $\pi : V \rightarrow V$  é *idempotente* se  $\pi^2 = \pi$ . Assuma que  $V$  tem dimensão finita e que  $\pi$  não é a identidade nem o operador identicamente nulo.
  - a) Calcule os polinômios minimal e característico de um operador idempotente.
  - b) Mostre que  $V = \operatorname{Nuc} \pi \oplus \operatorname{Im} \pi$  e que  $\operatorname{tr} \pi = \dim \operatorname{Im} \pi$ .
  - c) Verifique que  $\pi$  é uma projeção sobre a imagem ao longo do núcleo.
10. Prove que toda projeção ao longo de parcelas é um operador idempotente.

## 4.5 Polinômio minimal de um vetor

O ideal anulador com respeito ao operador linear  $A : V \rightarrow V$  de um vetor  $v \in V$  é o conjunto de polinômios

$$\mathfrak{S}_v = \{p(t) \in \mathbb{K}[t], p(A)(v) = 0\}.$$

De fato, verifica-se facilmente que esse conjunto é um ideal e, portanto, existe um polinômio mônico  $m_v(t) \in \mathfrak{S}_v$  tal que  $\mathfrak{S}_v = m_v(t)\mathbb{K}[t]$ , chamado de *polinômio minimal do vetor*  $v$ . É claro que ocorre a inclusão de ideais  $\mathfrak{S}_A \subset \mathfrak{S}_v$ , implicando que  $m_v(t)$  divide  $m_A(t)$ .

Da mesma forma definimos o ideal anulador de um subespaço  $W$  de  $V$ ,

$$\mathfrak{S}_W = \{p(t) \in \mathbb{K}[t], p(A)(w) = 0, \text{ para todo } w \in W\}.$$

O gerador desse ideal é chamado de polinômio minimal do subespaço  $W$ .

### Exercícios propostos 4.5.1

1. Com a notação dada acima, mostre que o polinômio minimal de um vetor relativo a um operador linear divide o polinômio minimal de um subespaço que o contém.
2. Suponha que um operador linear  $A$  num espaço vetorial de dimensão finita  $V$  preserva um subespaço  $W \subset V$ . Se  $A_0$  é o operador em  $W$  obtido por restrição de  $A$ , mostre que  $p_{A_0}(t) \mid p_A(t)$  ( $\mid$  significa que o primeiro polinômio divide o segundo).
3. Seja  $A$  um operador linear não invertível num espaço vetorial de dimensão finita  $V$  sobre  $\mathbb{K}$ . Existe algum polinômio não constante  $p(t) \in \mathbb{K}[t]$  para o qual o operador  $p(A) : V \rightarrow V$  é invertível?
4. Suponha que  $m_A(t) = p_1(t)^{r_1} p_2(t)^{r_2} \cdots p_k(t)^{r_k}$  é a fatoração primária do polinômio mínimo do operador linear  $A : V \rightarrow V$  num espaço vetorial de dimensão finita.
  - a) Se o núcleo de  $A$  é não trivial demonstre que  $p(t) = t$  divide o polinômio minimal  $m_A(t)$ .
  - b) Assuma que  $q(t)$  e  $m_A(t)$  sejam polinômios relativamente primos. Prove que o operador  $q(A) : V \rightarrow V$  é um isomorfismo.
  - c) Mostre que o polinômio minimal de qualquer vetor em  $\text{Nuc } p_i(A)^{r_i}$  é uma potência do tipo  $p_i(t)^{s_i}$ , com  $s_i \leq r_i$ .
  - d) Conclua que se  $i \neq j$  então  $\text{Nuc } p_i(A)^{r_i} \cap \text{Nuc } p_j(A)^{r_j} = \{0\}$ .
5. Qual o polinômio minimal do vetor nulo relativo a um operador  $A : V \rightarrow V$ ?
6. Assuma que  $V = W_1 \oplus W_2$  é uma decomposição em soma direta de um espaço vetorial de dimensão finita  $V$ . Seja  $A : V \rightarrow V$  a aplicação  $A(v) = v_1 - v_2$ , onde  $v = v_1 + v_2$  é a decomposição correspondente do vetor  $v$ . Calcule o polinômio minimal de um vetor qualquer  $v \in V$ .
7. Seja  $A$  um operador linear num espaço vetorial  $V$  de dimensão finita. Dado uma base ordenada  $\beta = \{v_1, v_2, \dots, v_n\}$  de  $V$ , denote por  $m_i(t)$  o polinômio minimal de  $v_i$ . Prove que o minimal múltiplo comum de  $m_1(t), m_2(t), \dots, m_n(t)$  é precisamente o polinômio minimal de  $A$ .

## Capítulo 5

# Decomposição primária

Nesse capítulo é demonstrado o Teorema da decomposição primária para operadores que é o ponto de referência em torno do qual o texto será desenvolvido. Construiremos para cada operador linear  $A$  num espaço vetorial de dimensão finita  $V$  uma decomposição de  $V$  em soma direta invariante por  $A$  associada ao polinômio minimal, decomposição essa que simplifica em muito o estudo qualitativo do operador. A demonstração do Teorema da decomposição primária segue, em essência, a idéia apresentada em [?].

### 5.1 Teorema da decomposição primária

**Teorema 5.1.1 (Teorema da decomposição primária)** *Seja  $A$  um operador linear no espaço vetorial de dimensão finita  $V$  sobre o corpo  $\mathbb{K}$ . Se  $m_A(t) = p_1(t)^{r_1} p_2(t)^{r_2} \cdots p_k(t)^{r_k}$  é a fatoração primária do polinômio minimal do operador linear, então  $V$  decompõe-se numa soma direta invariante por  $A$  na forma*

$$V = Nuc p_1(A)^{r_1} \oplus Nuc p_2(A)^{r_2} \oplus \cdots \oplus Nuc p_k(A)^{r_k}.$$

**Demonstração** Para evitar trivialidades assumiremos que na fatoração primária do polinômio minimal, o número de fatores relativamente primos é  $k \geq 2$ . Feita essa hipótese, para cada  $i$ ,  $1 \leq i \leq k$ , defina o polinômio

$$m_{\hat{i}}(t) = p_1(t)^{r_1} \cdots \widehat{p_i(t)^{r_i}} \cdots p_k(t)^{r_k}$$

(o sinal  $\widehat{\phantom{x}}$  sobre o  $i$ -ésimo fator indica que ele foi suprimido). Os polinômios  $m_{\hat{1}}(t)$ ,  $m_{\hat{2}}(t), \dots, m_{\hat{k}}(t)$  são relativamente primos, portanto, a soma dos ideais gerados pela família de polinômios  $\{m_{\hat{i}}(t)\}_{i=1}^k$  igual ao anel  $\mathbb{K}[t]$ , em outras palavras,

$$\mathbb{K}[t] = m_{\hat{1}}(t)\mathbb{K}[t] + m_{\hat{2}}(t)\mathbb{K}[t] + \cdots + m_{\hat{k}}(t)\mathbb{K}[t].$$

Sendo assim, existem polinômios  $q_1(t), q_2(t), \dots, q_k(t) \in \mathbb{K}[t]$  satisfazendo a identidade polinomial

$$1 \equiv m_{\hat{\gamma}_1}(t)q_1(t) + m_{\hat{\gamma}_2}(t)q_2(t) + \dots + m_{\hat{\gamma}_k}(t)q_k(t).$$

Avaliando a identidade em  $A$  e calculando a avaliação no vetor  $v \in V$  obtemos

$$v = Id(v) = m_{\hat{\gamma}_1}(A) \circ q_1(A)(v) + m_{\hat{\gamma}_2}(A) \circ q_2(A)(v) + \dots + m_{\hat{\gamma}_k}(A) \circ q_k(A)(v).$$

Portanto, se  $\pi_i : V \rightarrow V$  é o operador linear  $\pi_i = m_{\hat{\gamma}_i}(A) \circ q_i(A)$ , podemos afirmar que qualquer vetor  $v \in V$  é uma soma de vetores do tipo

$$v = v_1 + v_2 + \dots + v_k \quad \text{em que} \quad v_i = \pi_i(v) \in Im \pi_i,$$

Isso mostra que  $V$  é a soma das imagens dos operadores  $\pi_i : V \rightarrow V$ ,

$$V = Im \pi_1 + Im \pi_2 + \dots + Im \pi_k.$$

Para provar que  $V$  é a soma dos núcleos dos operadores  $p_i(A)^{r_i} : V \rightarrow V$ , isto é,

$$V = Nuc p_1(A)^{r_1} + Nuc p_2(A)^{r_2} + \dots + Nuc p_k(A)^{r_k},$$

é suficiente provar que valem as inclusões de subespaços  $Im \pi_i \subset Nuc p_i(A)^{r_i}$ . Com efeito, dado um vetor  $w \in Im \pi_i$ , escolhamos um vetor  $v \in V$  tal que

$$w = \pi_i(v) = m_{\hat{\gamma}_i}(A) \circ q_i(A)(v).$$

Lembrando-se da identidade polinomial  $m_A(t) = p_i(t)^{r_i} m_{\hat{\gamma}_i}(t)$  e de que o polinômio minimal de  $A$  pertencer ao ideal anulador de qualquer vetor façamos os cálculos,

$$\begin{aligned} p_i(A)^{r_i}(w) &= p_i(A)^{r_i} \circ m_{\hat{\gamma}_i}(A) \circ q_i(A)(v) \\ &= q_i(A) \circ \underbrace{m_A(A)}_{=o}(v) = o. \end{aligned}$$

Portanto, fica verificado que  $w \in Nuc p_i(A)^{r_i}$  como pretendíamos. Agora, para provar que a soma dos núcleos é uma soma direta é suficiente verificar a condição

$$Nuc p_i(A)^{r_i} \cap \{Nuc p_1(A)^{r_1} + \dots + Nuc p_{i-1}(A)^{r_{i-1}}\} = \{o\} \quad \text{para} \quad i \geq 2.$$

Seja  $v$  um vetor nessa interseção. Isto significa que

$$\begin{cases} v \in Nuc p_i(A)^{r_i} & e \\ v = v_1 + v_2 + \dots + v_{i-1} & \text{com} \quad v_j \in Nuc p_j(A)^{r_j}, \quad 1 \leq j \leq i-1. \end{cases}$$

Levando em conta que polinômios em  $A$  comutam e que  $p_j(A)^{r_j}(v_j) = o$ , para todo  $j$  com  $1 \leq j \leq i-1$ , avaliemos o polinômio  $q(t) = p_1(t)^{r_1} \dots p_{i-1}(t)^{r_{i-1}}$  no vetor  $v$ ,

$$q(A)(v) = \sum_{j=1}^{i-1} (p_1(A)^{r_1} \circ \dots \circ \hat{p}_j(A)^{r_j} \circ \dots \circ p_{i-1}(A)^{r_{i-1}}) \circ \underbrace{p_j(A)^{r_j}(v_j)}_{=0} = 0.$$

Por outro lado,  $p_i(A)^{r_i}(v) = 0$  pois  $v \in \text{Nuc } p_i(A)^{r_i}$ , implicando que os dois polinômios relativamente primos  $p_i(t)^{r_i}$  e  $q(t)$  pertencem ao ideal anulador  $\mathfrak{S}_v$ . Consequentemente, o gerador do ideal é o polinômio  $m_v(t) \equiv 1$ . Sendo assim,  $o = m_v(A)(v) = v$ , concluindo que  $V$  é uma soma direta dos núcleos. Para encerrar a demonstração observamos que  $A$  comuta com qualquer polinômio em  $A$ , daí segue que cada parcela  $\text{Nuc } p_i(A)^{r_i}$  é invariante pelo operador.  $\square$

A soma direta demonstrada no teorema acima (a menos da ordem das parcelas) será chamada de *decomposição primária de  $V$*  determinada por  $A$ . Escolhendo uma base ordenada  $\beta_i$  de  $\text{Nuc } p_i(A)^{r_i}$  obtemos uma representação matricial relativamente simples para o operador na base ordenada  $\beta = \bigcup \beta_i$ ,

$$[A]_\beta = \text{diag} \{ [A_1]_{\beta_1}, [A_2]_{\beta_2}, \dots, [A_k]_{\beta_k} \},$$

em que  $A_i : \text{Nuc } p_i(A)^{r_i} \rightarrow \text{Nuc } p_i(A)^{r_i}$  é o operador induzido por restrição de  $A$ .

**Exemplo 5.1.1** Vamos assumir que a representação matricial na base canônica  $\alpha$  do operador  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  é a matriz

$$[A]_\alpha = \begin{bmatrix} 1 & 4 & -3 \\ 1 & 10 & -7 \\ 1 & 13 & -9 \end{bmatrix}.$$

Calculando o polinômio característico de  $A$  obtemos o polinômio

$$p_A(t) = \det(tI - [A]_\alpha) = t(t-1)^2.$$

Como a avaliação do polinômio  $p(t) = t(t-1)$  em  $[A]_\alpha$  não é nulo pois

$$p([A]_\alpha) = [A]_\alpha([A]_\alpha - I) = \begin{bmatrix} 1 & 1 & -1 \\ 3 & 3 & -3 \\ 4 & 4 & -4 \end{bmatrix},$$

necessariamente o polinômio minimal de  $A$  é  $m_A(t) = t(t-1)^2$ . Sendo assim,  $A$  induz a decomposição primária  $V = \text{Nuc } A \oplus \text{Nuc } (A - Id)^2$ . Calculamos uma base para cada parcela da decomposição primária resolvendo os sistemas lineares,

$$[A]_\alpha \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 4 & -3 \\ 1 & 10 & -7 \\ 1 & 13 & -9 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

$$[A - Id]_\alpha^2 \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & -3 & 2 \\ 2 & -6 & 4 \\ 3 & -9 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Uma base para a primeira parcela da decomposição primária pode ser  $\beta_1 = \{(1, 2, 3)\}$  e uma base para a outra pode ser  $\beta_2 = \{(2, 0, -1), (3, 1, 0)\}$ .  $\square$

**Exemplo 5.1.2** Vamos supor que a representação matricial de  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  na base canônica é a matriz

$$[A]_{\alpha} = \begin{bmatrix} 3 & 4 & -4 \\ 8 & 7 & -8 \\ 10 & 10 & -11 \end{bmatrix}.$$

Calculando o polinômio característico de  $A$  obtemos o polinômio

$$p_A(t) = \det(tI - [A]_{\alpha}) = (t + 1)^2(t - 1).$$

Necessariamente o polinômio minimal é  $m_A(t) = (t + 1)(t - 1)$  pois um cálculo matricial mostra que  $[A]_{\alpha}^2 - I \equiv [0]$ . Portanto, a decomposição primária de  $V$  fica sendo  $V = Nuc(A + Id) \oplus Nuc(A - Id)$ . Determinemos uma base para cada parcela resolvendo os sistemas lineares

$$[A + Id]_{\alpha} [(x, y, z)]_{\alpha} = [(0, 0, 0)]_{\alpha},$$

$$[A - Id]_{\alpha} [(x, y, z)]_{\alpha} = [(0, 0, 0)]_{\alpha}.$$

Daí obtemos as bases  $\beta_1 = \{(1, -1, 0), (1, 0, 1)\}$  e  $\beta_2 = \{(-1, 0, -1)\}$ . Observe a relação entre a dimensão de cada parcela e o grau do fator correspondente na decomposição primária do polinômio característico, relação que será estabelecida posteriormente.  $\square$

### Exercícios propostos 5.1.1

- Determine a decomposição primária induzida por cada operador e encontre uma base ordenada na qual a representação matricial é uma diagonal de matrizes.
  - $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (5x - y + 3z, -6x + 4y - 6z, -6x + 2y - 4z)$ .
  - $A : \mathbb{C}^3 \rightarrow \mathbb{C}^3$ ,  $A(x, y, z) = (2y + z, -y, 2x - y - z)$ .
  - $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (x, x + 2y, x + y - 3z)$ .
  - $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (3z, 3x, 3y)$ .
  - $A : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ ,  $A(x, y, z, w) = (y, z, w, x)$ .
- Sejam  $A$  e  $B$  dois operadores no espaço vetorial  $V$  de dimensão finita. Mostre que se os operadores comutam então  $B$  preserva cada parcela da decomposição primária determinada por  $A$ .
- Assuma que o operador  $A$  num espaço vetorial de dimensão finita  $V$  é uma  $n$ -ésima raiz do operador  $B$ , isto é,  $A^n = B$  para  $n \geq 1$ . Mostre que  $A$  preserva cada parcela da decomposição primária determinada por  $B$ .
- Suponha que o operador linear  $A : V \rightarrow V$  é tal que  $A^3 \equiv A$ .
  - Determine as possíveis decomposições primárias determinada por  $A$ .

- b) Represente  $A$  numa base ordenada que é uma união ordenada de bases ordenadas das parcelas da decomposição.
5. Dados dois polinômios relativamente primos,  $p(t)$  e  $q(t)$ , cujo produto  $p(t)q(t)$  é o polinômio minimal de um operador linear  $A$  de um espaço vetorial de dimensão finita  $V$ . Demonstre que  $V = Nuc p(A) \oplus Nuc q(A)$ .
6. Considere o operador linear  $A : \mathbb{R}_n[t] \rightarrow \mathbb{R}_n[t]$ ,  $A(p(t)) = p(1-t)$ .
- Mostre que  $A$  é uma involução.
  - Dê a decomposição primária determinada pelo operador.
  - Calcule uma base para cada parcela da decomposição primária.

## 5.2 Corolários

Apresentaremos algumas conseqüências do Teorema da decomposição primária. Por hipótese, em todos corolários dessa seção  $A : V \rightarrow V$  denotará um operador linear num espaço vetorial de dimensão finita  $V$  sobre o corpo  $\mathbb{K}$ ,

$$m_A(t) = p_1(t)^{r_1} p_2(t)^{r_2} \cdots p_k(t)^{r_k}$$

indicará a fatoração primária do seu polinômio minimal, enquanto a decomposição primária induzida por  $A$  será registrada como

$$V = Nuc p_1(A)^{r_1} \oplus Nuc p_2(A)^{r_2} \oplus \cdots \oplus Nuc p_k(A)^{r_k}.$$

**Corolário 5.2.1** *O polinômio minimal de  $A_i : Nuc p_i(A)^{r_i} \rightarrow Nuc p_i(A)^{r_i}$ , o operador induzido por restrição de  $A$ , é o polinômio  $m_{A_i}(t) = p_i(t)^{r_i}$ .*

**Demonstração** Para qualquer vetor  $v$  em  $Nuc p_i(A)^{r_i}$  valem as igualdades

$$p_i(A_i)^{r_i}(v) = p_i(A)^{r_i}(v) = o,$$

de onde concluímos que o polinômio  $p_i(t)^{r_i}$  pertence ao ideal anulador  $\mathfrak{S}_{A_i} = m_{A_i}(t)\mathbb{K}[t]$ . Sendo  $p_i(t)^{r_i}$  um polinômio mônico e primo, necessariamente teremos que  $m_{A_i}(t) = p_i(t)^{s_i}$ , com  $1 \leq s_i \leq r_i$ . Por absurdo, suponhamos que para algum  $i$  ocorra a desigualdade  $s_i < r_i$ . Sendo assim, considere o polinômio

$$m(t) = p_1(t)^{s_1} p_2(t)^{s_2} \cdots p_k(t)^{s_k}$$

cujos graus são estritamente menores que o grau do polinômio minimal de  $A$ . Pelo Teorema da decomposição primária um vetor  $v \in V$  é expresso como

$$v = v_1 + v_2 + \cdots + v_k \quad \text{com} \quad v_i \in Nuc p_i(A)^{r_i}.$$

Como  $p_i(A)^{s_i}(v_i) = o$  e polinômios em  $A$  comutam, segue as igualdades,

$$m(A)(v) = \sum_{i=1}^k m(A)(v_i)$$

$$\begin{aligned}
&= \sum_{i=1}^k (p_1(A)^{s_1} \circ \cdots \circ \widehat{p_i(A)^{s_i}} \circ \cdots \circ p_k(A)^{s_k}) \circ \underbrace{p_i(A)^{s_i}(v_i)}_{=o} \\
&= o.
\end{aligned}$$

Logo, o polinômio  $m(t)$  anula o operador linear  $A$  e  $\text{grau } m(t) < \text{grau } m_A(t)$ . Uma contradição. Portanto,  $m_{A_i}(t) = p_i(t)^{r_i}$ , como queríamos demonstrar.  $\square$

O polinômio característico do operador determina a dimensão de cada parcela da decomposição primária. Registremos este fato num corolário.

**Corolário 5.2.2** *Se  $p_A(t) = p_1(t)^{s_1} p_2(t)^{s_2} \cdots p_k(t)^{s_k}$  é a fatoração primária do polinômio característico de  $A$ , então  $\text{grau } p_i(t)^{s_i} = \dim \text{Nuc } p_i(A)^{r_i}$ .*

**Demonstração** Se  $\beta = \overrightarrow{\cup} \beta_i$  é uma base ordenada construída após escolha de uma base ordenada  $\beta_i$  do subespaço  $\text{Nuc } p_i(t)^{r_i}$ , a representação matricial de  $A$  é uma diagonal de matrizes, a saber,

$$[A]_{\beta} = \text{diag} \{ [A_1]_{\beta_1}, [A_2]_{\beta_2}, \dots, [A_k]_{\beta_k} \},$$

em que  $A_i : \text{Nuc } p_i(A)^{r_i} \rightarrow \text{Nuc } p_i(A)^{r_i}$  é o operador obtido por restrição de  $A$ . Calculemos o polinômio característico de  $A$  em função dos polinômios característicos das restrições. Seja  $I_i$  a matriz identidade de mesma dimensão que  $[A_i]_{\beta_i}$ . Então

$$\begin{aligned}
p_A(t) &= \det (tI - [A]_{\beta}) \\
&= \det (tI_1 - [A_1]_{\beta_1}) \cdots \det (tI_k - [A_k]_{\beta_k}) \\
&= p_{A_1}(t) \cdots p_{A_k}(t).
\end{aligned}$$

Pelo corolário acima, sabemos que  $m_{A_i}(t) = p_i(t)^{r_i}$ . Como o polinômio minimal  $m_{A_i}(t)$  divide o polinômio característico  $p_{A_i}(t)$  podemos afirmar que  $p_{A_i}(t) = p_i(t)^{z_i}$  para algum inteiro  $z_i \geq r_i$ . Portanto, uma fatoração primária para  $p_A(t)$  é

$$p_A(t) = p_1(t)^{z_1} p_2(t)^{z_2} \cdots p_k(t)^{z_k}.$$

Entretanto, a fatoração primária é única e por hipótese

$$p_A(t) = p_1(t)^{s_1} p_2(t)^{s_2} \cdots p_k(t)^{s_k}.$$

Logo,  $z_i = s_i$ , isto é,  $p_{A_i}(t) = p_i(t)^{s_i}$ . Recordamos que o grau de um polinômio característico é igual à dimensão do espaço, no caso aqui considerado temos,

$$\dim \text{Nuc } p_i(A)^{r_i} = \text{grau } p_i(t)^{s_i}. \quad \square$$

Na demonstração do Teorema da decomposição primária está subjacente um fato que deve ser ressaltado pela sua utilidade.

**Corolário 5.2.3** *A projeção  $\pi_i : V \rightarrow V$  sobre  $Nuc p_i(A)^{r_i}$  ao longo das outras parcelas é um polinômio em  $A$ .*

**Demonstração** Revisemos a demonstração do Teorema da decomposição primária. Alí, definimos a família de polinômios relativamente primos  $\{m_{\hat{\pi}_i}(t)\}_{i=1}^k$ , onde

$$m_{\hat{\pi}_i}(t) = p_1(t)^{r_1} \cdots \hat{p}_i(t)^{r_i} \cdots p_k(t)^{r_k}.$$

Como

$$\mathbb{K}[t] = m_{\hat{\pi}_1}(t)\mathbb{K}[t] + m_{\hat{\pi}_2}(t)\mathbb{K}[t] + \cdots + m_{\hat{\pi}_k}(t)\mathbb{K}[t],$$

escolhemos polinômios  $q_1(t), q_2(t), \dots, q_k(t) \in \mathbb{K}[t]$  satisfazendo a identidade

$$1 \equiv m_{\hat{\pi}_1}(t)q_1(t) + m_{\hat{\pi}_2}(t)q_2(t) + \cdots + m_{\hat{\pi}_k}(t)q_k(t).$$

Definimos o operador  $\pi_i : V \rightarrow V$  por  $\pi_i = m_{\hat{\pi}_i}(A) \circ q_i(A)$  e concluímos que

$$\begin{cases} V = Im \pi_1 + Im \pi_2 + \cdots + Im \pi_k, \\ Id = \pi_1 + \pi_2 + \cdots + \pi_k. \end{cases}$$

Observe que  $\pi_i$  é um polinômio em  $A$ ! Além disso, provamos a inclusão de subespaços  $Im \pi_i \subset Nuc p_i(A)^{r_i}$ . Na verdade, nessa inclusão vale a igualdade pois, por absurdo, suponha que para algum  $i$  ocorra a desigualdade dimensional  $\dim Im \pi_i < \dim Nuc p_i(A)^{r_i}$ , então

$$\dim V \leq \sum_{i=1}^k \dim Im \pi_i < \sum_{i=1}^k \dim Nuc p_i(A)^{r_i} = \dim V,$$

evidentemente uma contradição. Logo, a decomposição primária de  $V$  é também expressa por

$$V = Im \pi_1 \oplus Im \pi_2 \oplus \cdots \oplus Im \pi_k.$$

Para provar que  $\pi_i$  é a projeção de  $V$  sobre  $Nuc p_i(A)^{r_i} = Im \pi_i$  ao longo das outras parcelas basta mostrar que  $\pi_i \circ \pi_j \equiv 0$  se  $i \neq j$  e que  $\pi_i \circ \pi_i \equiv \pi_i$ . Com efeito. Se  $i \neq j$  o polinômio que define  $\pi_i \circ \pi_j$  é divisível pelo polinômio minimal de  $A$ , daí a nulidade da composta. A segunda afirmação é obtida compondo  $\pi_i$  com a identidade

$$Id = \pi_1 + \pi_2 + \cdots + \pi_k.$$

Isto termina a demonstração do corolário.  $\square$

O Teorema da decomposição primária simplifica bastante o estudo dos subespaços invariantes pois ele é uma soma direta de subespaços invariantes contidos nas parcelas da decomposição primária.

**Corolário 5.2.4** *Se  $W \subset V$  é um subespaço invariante pelo operador  $A$  então*

$$W = (W \cap Nuc p_1(A)^{r_1}) \oplus (W \cap Nuc p_2(A)^{r_2}) \oplus \cdots \oplus (W \cap Nuc p_k(A)^{r_k}).$$

**Demonstração** A inclusão

$$W \supset (W \cap Nuc p_1(A)^{r_1}) + (W \cap Nuc p_2(A)^{r_2}) + \cdots + (W \cap Nuc p_k(A)^{r_k})$$

é independente de  $W$  ser invariante por  $A$  ou não. Verifiquemos a inclusão oposta. Qualquer vetor  $w \in W$  é decomposto de modo único como

$$w = \pi_1(w) + \pi_2(w) + \cdots + \pi_k(w),$$

onde  $\pi_i$  é a projeção de  $V$  sobre a  $i$ -ésima parcela da decomposição primária ao longo das outras parcelas. Como vimos,  $\pi_i$  é um polinômio em  $A$ . Desde que  $W$  é invariante por  $A$ , é claro que  $W$  é invariante por qualquer polinômio em  $A$ , em particular por  $\pi_i$ . Logo  $\pi_i(w) \in W \cap Nuc p_i(A)^{r_i}$ , mostrando a inclusão desejada. Portanto

$$W = (W \cap Nuc p_1(A)^{r_1}) + (W \cap Nuc p_2(A)^{r_2}) + \cdots + (W \cap Nuc p_k(A)^{r_k}).$$

Como a decomposição primária é uma soma direta, concluímos imediatamente que essa última soma é também uma soma direta.  $\square$

Terminaremos a seção mostrando a existência de um vetor especial. O próximo resultado será utilizado repetidas vezes.

**Corolário 5.2.5** *Existe um vetor  $v \in V$  cujo polinômio minimal é igual ao polinômio minimal do operador linear  $A$ .*

**Demonstração** A demonstração será feita em três afirmações.

*Afirmção 1.* Existe um vetor  $v_i \in Nuc p_i(A)^{r_i}$  cujo polinômio minimal  $m_{v_i}(t)$  é precisamente  $p_i(t)^{r_i}$ .

O polinômio minimal de qualquer vetor em  $Nuc p_i(A)^{r_i}$  é uma potência de  $p_i(t)$ . Por absurdo, suponha que todo vetor  $v \in Nuc p_i(A)^{r_i}$  possui um polinômio minimal  $m_v(t) = p_i(t)^{r_v}$  satisfazendo  $r_v < r_i$ . Considere o polinômio  $p(t)^r$  onde  $r$  é o máximo dos inteiros  $r_v$ . Dessa forma,  $p_i(A)^r(v) = 0$  para todo  $v \in Nuc p_i(A)^{r_i}$ , contradizendo o primeiro corolário desta seção.

*Afirmção 2.* Se  $i \neq j$  o operador linear  $p_i(A)^n : Nuc p_j(A)^{r_j} \rightarrow Nuc p_j(A)^{r_j}$  é um isomorfismo linear para todo inteiro  $n \geq 0$ .

Se  $v$  é um vetor no núcleo do operador  $p_i(A)^n : Nuc p_j(A)^{r_j} \rightarrow Nuc p_j(A)^{r_j}$ , então os polinômios  $p_i(t)^n$  e  $p_j(t)^{r_j}$  pertencem ao ideal anulador  $\mathfrak{S}_v$  de onde concluímos que seu polinômio minimal  $m_v(t)$  divide dois polinômios relativamente primos. Logo,  $m_v(t) \equiv 1$ , implicando que  $v = o$ .

*Afirmção 3.* O vetor  $v = v_1 + v_2 + \cdots + v_k$  tem polinômio mínimo igual ao polinômio minimal de  $A$ , onde  $v_i$  é o vetor obtido na Afirmção 1.

Deixaremos como exercício a prova desta afirmação.  $\square$

**Exercício 5.2.1** Utilize as afirmações 1 e 2 do corolário acima para provar que a imagem do operador linear  $p_i(A)^{r_i} : V \rightarrow V$  é o subespaço

$$\text{Im } p_i(A)^{r_i} = \text{Nuc } p_1(A)^{r_1} \oplus \cdots \oplus \widehat{\text{Nuc } p_i(A)^{r_i}} \oplus \cdots \oplus \text{Nuc } p_k(A)^{r_k}. \quad \square$$

### Exercícios propostos 5.2.1

1. Considere o operador linear  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (-y, 2x + 3y, -x + 2y)$ .
  - a) Calcule  $m_A(t)$ .
  - b) Examine a identidade polinomial  $(t - 2)^2 + (t - 1)(3 - t) = 1$  e explicita as projeções  $\pi_i$ 's sobre as parcelas da decomposição primária.
2. Encontre um vetor cujo polinômio minimal é igual ao polinômio minimal do operador.
  - a)  $A : \mathbb{C}^3 \rightarrow \mathbb{C}^3$ ,  $A(x, y, z) = (0, 2x, x + 3y)$ .
  - b)  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (2x + y + z, -2x + y + 3z, 3x + y - z)$
3. Seja  $\pi : V \rightarrow V$  um operador linear num espaço vetorial  $V$  de dimensão finita sobre  $\mathbb{K}$  satisfazendo a condição  $\pi^2 \equiv \pi$ .
  - a) Qual o polinômio minimal de  $\pi$ ?
  - b) Se  $\pi$  não é a identidade e nem o operador nulo, prove que  $V = \text{Im } \pi \oplus \text{Nuc } \pi$ , com cada parcela não trivial.
4. Denote por  $V_1$  e  $V_2$  os subespaços unidimensionais de  $\mathbb{R}^2$  gerados pelos vetores  $v_1 = (1, 2)$  e  $v_2 = (1, -1)$ , respectivamente.
  - a) Mostre que  $\mathbb{R}^2 = V_1 \oplus V_2$ .
  - b) Se  $\pi_1$  é a projeção de  $\mathbb{R}^2$  sobre  $V_1$  ao longo da outra parcela, dê a fatoração primária de  $m_{\pi_1}(t)$ .
  - c) Qual a decomposição primária induzida por  $\pi_1$ ?
  - d) Calcule a representação matricial de  $\pi_1$  na base canônica.
5. Suponha que o polinômio minimal do operador linear não identicamente nulo  $A : V \rightarrow V$  seja um produto de fatores lineares distintos. Demonstre que existe um operador linear  $B : V \rightarrow V$  tal que  $P = A \circ B$  é idempotente, isto é,  $P^2 = P$ .
6. Determine o polinômio característico do operador induzido numa parcela da decomposição primária definida pelo operador.
7. Assuma que  $A : V \rightarrow V$  é um operador linear sobre um espaço vetorial  $V$  de dimensão finita sobre  $\mathbb{K}$  preservando a soma direta  $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$ .
  - a) Se  $A_i : V_i \rightarrow V_i$  é o operador induzido por restrição de  $A$ , prove que o polinômio minimal  $m_{A_i}(t)$  pertence ao ideal anulador  $\mathfrak{S}_{A_i}$ , para todo  $i$ .
  - b) O polinômio minimal de  $A$  é o produto dos polinômios mínimos  $m_{A_i}(t)$ ?

8. Seja  $A$  um operador linear num espaço vetorial de dimensão finita  $V$ . Prove que se  $u$  e  $v$  são vetores de parcelas diferentes da decomposição primária, então o polinômio minimal da soma  $u + v$  é o produto dos polinômios mínimo de  $u$  e de  $v$ .
9. Descreva todos os subespaços invariantes pelo operador linear.
  - a)  $A : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ ,  $A(x, y, z, w) = (0, x - w, y + 4w, z + w)$ .
  - b)  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (x - 6y, x + 2z, x + 2z)$ .

### 5.3 Subespaços cíclicos

Quando o polinômio minimal de um operador linear  $A : V \rightarrow V$  é uma potência de um único polinômio primo,  $m_A(t) = p(t)^r$ , o Teorema da decomposição primária praticamente nada acrescenta ao nosso conhecimento sobre o operador. Um exemplo típico dessa situação é o operador induzido por restrição a uma das parcelas primárias. Como sabemos, o induzido tem um polinômio minimal da forma  $p(t)^r$ . Para contornar essa e outras dificuldades consideramos o subespaço invariante de menor dimensão que contém um dado vetor e procuramos decompor o espaço  $V$  em uma soma direta de tais subespaços. Precisemos os conceitos.

Seja  $A$  um operador linear num espaço vetorial de dimensão finita  $V$  sobre o corpo  $\mathbb{K}$ . Um espaço invariante por  $A$  que contém um vetor  $v \in V$  deve conter todos os iterados de  $v$  pelo operador  $A$ , isto é, deve conter o conjunto

$$\varepsilon = \{v, A(v), A^2(v), \dots, A^n(v), \dots\}.$$

Como sabemos, o menor subespaço que contém o conjunto  $\varepsilon$  é aquele formado pelas combinações lineares de seus elementos, o qual será chamado de subespaço *A-cíclico gerado por  $v$*  e denotado por  $\mathcal{C}_A(v)$  ou, quando não causar ambigüidades, por  $\mathcal{C}(v)$ . Verifica-se que tal subespaço é invariante por  $A$ . Um modo conciso de definir uma combinação linear de iterados de  $v$  por  $A$  é considerar um polinômio

$$p(t) = a_n t^n + \dots + a_1 t + a_0$$

em  $\mathbb{K}[t]$  e fazer a avaliação

$$p(A)(v) = a_n A^n(v) + \dots + a_1 A(v) + a_0 Id(v).$$

Desse modo, temos a seguinte descrição do subespaço *A-cíclico gerado por  $v$* ,

$$\mathcal{C}_A(v) = \{p(A)(v); p(t) \in \mathbb{K}[t]\}.$$

Relacionaremos algumas propriedades sobre subespaços cíclicos que serão utilizadas posteriormente sem nenhuma referência. Observamos que o conhecimento do polinômio minimal do vetor é a principal fonte de informações sobre o subespaço cíclico gerado por ele, como veremos na proposição a seguir.

**Proposição 5.3.1** *Seja  $A$  um operador linear num espaço vetorial  $V$  de dimensão finita sobre o corpo  $\mathbb{K}$  e seja  $v \in V$  um vetor não nulo. Então*

- a) *o conjunto  $\beta = \{v, A(v), \dots, A^{k_0-1}(v)\}$  é uma base para  $\mathcal{C}(v)$ , onde  $k_0$  é o grau do polinômio minimal de  $v$ . Em particular,  $\dim \mathcal{C}(v) = \text{grau } m_v(t)$ ;*
- b) *vale a igualdade dos ideais anuladores:  $\mathfrak{S}_v = \mathfrak{S}_{\mathcal{C}(v)}$ ;*
- c) *se  $A_0 : \mathcal{C}(v) \rightarrow \mathcal{C}(v)$  é o operador linear induzido por restrição de  $A$ , então o polinômio minimal de  $A_0$  é igual ao polinômio minimal de  $v$ .*

**Demonstração** a) Escolha um vetor  $w \in \mathcal{C}(v)$ . Por definição de subespaço cíclico, existe um polinômio

$$p(t) = a_n t^n + \dots + a_1 t + a_0, \quad \text{tal que} \quad w = p(A)(v).$$

Pelo algoritmo da divisão de Euclides temos que

$$p(t) = q(t)m_v(t) + r(t) \quad \text{com} \quad \text{grau } r(t) < \text{grau } m_v(t) = k_0.$$

Daí segue que  $w = r(A)(v)$ , em outras palavras,  $w$  é uma combinação linear de vetores de  $\beta = \{v, A(v), \dots, A^{k_0-1}(v)\}$ , mostrando que  $\beta$  é um conjunto de geradores para  $\mathcal{C}(v)$ . Os vetores de  $\beta$  são linearmente independentes pois uma combinação linear

$$a_{k_0-1}A^{k_0-1}(v) + \dots + a_1A(v) + a_0v = o$$

define um polinômio  $p(t) = a_{k_0-1}t^{k_0-1} + \dots + a_1t + a_0$  no ideal anulador  $\mathfrak{S}_v$  satisfazendo a condição  $\text{grau } p(t) < \text{grau } m_v(t)$ . Logo,  $p(t) \equiv 0$ , ou equivalentemente,  $a_{k_0-1} = \dots = a_1 = a_0 = 0$ , como desejávamos demonstrar.

b) Provenos apenas a inclusão  $\mathfrak{S}_v \subset \mathfrak{S}_{\mathcal{C}(v)}$ . Seja  $p(t)$  um polinômio no ideal anulador  $\mathfrak{S}_v$ . Dado um vetor  $w \in \mathcal{C}(v)$ , por definição de subespaço cíclico, existe um polinômio  $q(t) \in \mathbb{K}[t]$  tal que  $w = q(A)(v)$ . Observe que

$$p(A)(w) = p(A) \circ q(A)(v) = q(A) \circ \underbrace{p(A)(v)}_{=o} = o.$$

Portanto,  $p(A)$  é um operador que anula qualquer vetor em  $\mathcal{C}(v)$ , significando que  $p(t) \in \mathfrak{S}_{\mathcal{C}(v)}$ .

c) Novamente, demonstraremos apenas a inclusão  $\mathfrak{S}_{A_0} \subset \mathfrak{S}_v$ , a inclusão oposta ficará como exercício. Afirmar que  $p(t)$  pertence ao ideal anulador  $\mathfrak{S}_{A_0}$  significa afirmar que  $p(A_0)(w) = o$  para qualquer vetor  $w \in \mathcal{C}(v)$ . Em particular, como  $v \in \mathcal{C}(v)$  temos que  $p(A_0)(v) = o$ . Desde que  $p(A)(v) = p(A_0)(v) = o$  concluímos que  $p(t)$  pertence ao ideal anulador  $\mathfrak{S}_v$ , mostrando a inclusão  $\mathfrak{S}_{A_0} \subset \mathfrak{S}_v$  como pretendíamos.  $\square$

**Exercícios propostos 5.3.1**

1. Quantos subespaços são invariantes por  $A : \mathbb{R}^4 \rightarrow \mathbb{R}^4$  se o polinômio minimal do operador é  $m_A(t) = (t^2 + 1)(t^2 + t + 1)$ ?
2. Sejam  $A$  um operador linear em  $V$ . Prove que  $\mathcal{C}(u) = \mathcal{C}(v) \Leftrightarrow$  existe um polinômio  $g(t) \in \mathbb{K}[t]$  relativamente primo com  $m_u(t)$  tal que  $g(A)(u) = v$ .
3. Dado  $A$  um operador linear no espaço vetorial  $V$  de dimensão finita sobre  $\mathbb{K}$ , seja  $v \in V$  um vetor cujo polinômio minimal é  $m_v(t) = t^k + a_{k-1}t^{k-1} + \dots + a_1t + a_0$ . Denote por  $A_0$  o operador em  $\mathcal{C}(v)$  induzido por restrição de  $A$ . Calcule a representação de  $A_0$  na base ordenada  $\beta = \{v, A(v), \dots, A^{k-1}(v)\}$ . A matriz obtida é chamada de *matriz companheira*.

**5.4 Espaços cíclicos**

Seja  $A$  um operador linear num espaço vetorial  $V$  de dimensão finita sobre  $\mathbb{K}$ . Diremos que o espaço é *A-cíclico* quando existe um vetor  $v \in V$  tal que  $V = \mathcal{C}(v)$ . Espaços *A-cíclicos* admitem uma caracterização bastante simples através dos dois principais polinômios associados ao operador. Com as hipóteses acima, temos a

**Proposição 5.4.1** *O espaço  $V$  é A-cíclico se, e somente se, o polinômio característico e o polinômio mínimo de  $A$  são iguais.*

**Demonstração**  $\Rightarrow$ ) Suponha que exista um vetor  $v \in V$  tal que  $V = \mathcal{C}(v)$ . Então valem as igualdades

$$\text{grau } p_A(t) = \dim V = \dim \mathcal{C}(v) = \text{grau } m_v(t).$$

Como  $m_v(t)$  divide  $m_A(t)$  e  $m_A(t)$  divide  $p_A(t)$ , como os polinômios são mônicos e como  $\text{grau } m_v(t) = \text{grau } p_A(t)$  concluímos que  $m_v(t) = m_A(t) = p_A(t)$ .

$\Leftarrow$ ) Vamos supor a igualdade de polinômios  $m_A(t) = p_A(t)$ . Escolhamos um vetor  $v \in V$  cujo polinômio minimal é igual ao polinômio minimal de  $A$ . Então

$$\dim V = \text{grau } p_A(t) = \text{grau } m_A(t) = \text{grau } m_v(t) = \dim \mathcal{C}(v).$$

A igualdade dimensional implica imediatamente que  $V = \mathcal{C}(v)$ .  $\square$

**Exemplo 5.4.1** Consideremos o operador  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (z, x, y)$ . Se  $\alpha = \{e_1, e_2, e_3\}$  é a base canônica do  $\mathbb{R}^3$  é fácil verificar que  $A(e_1) = e_2$ ,  $A(e_2) = e_3$  e  $A(e_3) = e_1$ . Como os iterados do vetor  $e_1$  contém a base canônica,  $\mathbb{R}^3$  é um espaço *A-cíclico*, mais precisamente,  $\mathbb{R}^3 = \mathcal{C}(e_1)$ . Examinemos com mais detalhes o operador. O Teorema da decomposição primária afirma que  $\mathbb{R}^3$  decompõe-se na soma direta *A*-invariante,

$$\mathbb{R}^3 = \text{Nuc}(A + Id) \oplus \text{Nuc}(A^2 + A + Id),$$

pois o polinômio minimal é  $m_A(t) = (t+1)(t^2+t+1)$ . Com um pouco de cálculo determinamos as seguintes decomposições cíclicas para as parcelas primárias,

$$\begin{cases} \text{Nuc}(A + Id) = \mathcal{C}(v_1) & \text{com } v_1 = (1, 1, 1), \\ \text{Nuc}(A^2 + A + Id) = \mathcal{C}(v_2) & \text{com } v_2 = (1, -1, 0). \end{cases}$$

Logo, também é possível decompor o espaço  $\mathbb{R}^3$  em duas parcelas  $A$ -cíclicas, a saber,  $\mathbb{R}^3 = \mathcal{C}(v_1) \oplus \mathcal{C}(v_2)$ . Na próxima seção descreveremos melhor as várias possibilidades de decomposições em soma direta por subespaços  $A$ -cíclicos.  $\square$

### Exercícios propostos 5.4.1

1. Se o polinômio característico do operador  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  não tem raízes reais, prove que  $\mathbb{R}^2$  é um espaço  $A$ -cíclico.
2. Verifique que  $\mathbb{R}^3$  é um espaço  $A$ -cíclico e determine um vetor  $v$  tal que  $\mathbb{R}^3 = \mathcal{C}(v)$  quando  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  é o operador
  - a)  $A(x, y, z) = (x + 4y - 3z, x + 10y - 7z, x + 13y - 9z)$ .
  - b)  $A(x, y, z) = (x + 2z, -x + z, x + y + 2z)$ .
3. Seja  $A : V \rightarrow V$  um operador linear num espaço vetorial de dimensão finita. Suponha que  $V$  é  $A$ -cíclico. Mostre que um operador  $B : V \rightarrow V$  que comuta com  $A$  é um polinômio em  $A$ .
4. Prove que se um espaço vetorial de dimensão finita  $V$  é  $A^n$ -cíclico para algum inteiro  $n \geq 2$ , então  $V$  é  $A$ -cíclico.
5. Demonstre que um subespaço de um espaço  $A$ -cíclico é também  $A$ -cíclico.
6. Seja  $A$  um operador linear num espaço vetorial  $V$  de dimensão finita. Suponha que  $V = \mathcal{C}(v)$  e que  $m_v(t) = p(t)^r$  é a fatoração primária do polinômio minimal de  $v$ . Mostre que se  $s \leq r$  então  $p(A)^s(V)$  é um subespaço  $A$ -cíclico gerado por  $p(A)^s(v)$  com dimensão  $\text{grau } p(t)^r - \text{grau } p(t)^s$ .

## 5.5 Sobre a decomposição cíclica

Um dos nossos objetivos é construir uma decomposição "canônica" do espaço em uma soma direta de subespaços  $A$ -cíclicos,

$$V = \mathcal{C}(v_1) \oplus \mathcal{C}(v_2) \oplus \cdots \oplus \mathcal{C}(v_m).$$

Esse será o tópico tratado nos próximos dois capítulos. Como vimos no último exemplo, sem impor alguma restrição é possível construir várias decomposições

$A$ -cíclicas de  $V$  não isomorfas em geral. Para deixar mais claro essa idéia, apresentaremos a seguir um teorema que garante um tipo de unicidade de decomposição cíclica impondo um mínimo de restrições sobre as parcelas da decomposição. Embora seja um procedimento matematicamente exdrúxulo apresentar um teorema de unicidade sem antes saber se o objeto existe, estamos convencidos da sua conveniência. Na próxima definição utilizaremos a terminologia usual. Seja

$$V = Nuc p_1(A)^{r_1} \oplus Nuc p_2(A)^{r_2} \oplus \cdots \oplus Nuc p_k(A)^{r_k}$$

a decomposição primária determinada pelo operador linear  $A : V \rightarrow V$ .

**Definição 5.5.1** *Uma decomposição  $A$ -cíclica de  $V$  é uma decomposição em soma direta de subespaços  $A$ -cíclicos na qual cada parcela da decomposição primária está decomposta na forma,*

$$Nuc p_i(A)^{r_i} = \mathcal{C}(v_1) \oplus \mathcal{C}(v_2) \oplus \cdots \oplus \mathcal{C}(v_{m_i}),$$

$$grau p_i(t)^{r_i} = \dim \mathcal{C}(v_1) \geq \dim \mathcal{C}(v_2) \geq \cdots \geq \dim \mathcal{C}(v_{m_i}).$$

Em tempo, uma representação de Jordan de um operador é obtida a partir de uma tal decomposição. Já dissemos que a existência de uma decomposição  $A$ -cíclica será demonstrada posteriormente. Aqui nos ocuparemos da unicidade dimensional da decomposição  $A$ -cíclica explicada no teorema abaixo. O estudo fica reduzido a operadores para os quais o polinômio minimal é potência de um único polinômio primo  $p(t)$  pois, em geral, só precisaremos examinar o operador  $A_i : Nuc p_i(A)^{r_i} \rightarrow Nuc p_i(A)^{r_i}$ , induzido por restrição de  $A$ , cujo polinômio minimal é  $p_i(t)^{r_i}$ .

**Teorema 5.5.1 (Teorema da unicidade dimensional cíclica)** *Seja  $A$  um operador linear no espaço vetorial  $V$  de dimensão finita sobre o corpo  $\mathbb{K}$  cujo polinômio minimal é potência de um único polinômio primo,  $m_A(t) = p(t)^r$ . Se existe uma decomposição  $A$ -cíclica*

$$V = \mathcal{C}(v_1) \oplus \mathcal{C}(v_2) \oplus \cdots \oplus \mathcal{C}(v_k),$$

$$grau p(t)^r = \dim \mathcal{C}(v_1) \geq \dim \mathcal{C}(v_2) \geq \cdots \geq \dim \mathcal{C}(v_k),$$

*então esta decomposição é dimensionalmente única no seguinte sentido. Para qualquer outra decomposição  $A$ -cíclica*

$$V = \mathcal{C}(u_1) \oplus \mathcal{C}(u_2) \oplus \cdots \oplus \mathcal{C}(u_n),$$

$$grau p(t)^r = \dim \mathcal{C}(u_1) \geq \dim \mathcal{C}(u_2) \geq \cdots \geq \dim \mathcal{C}(u_n),$$

*o número de parcelas são iguais e  $\dim \mathcal{C}(v_i) = \dim \mathcal{C}(u_i)$*

A demonstração do teorema é um laborioso cálculo dimensional e, certamente, quebrará o ritmo de leitura. Se por um lado a compreensão do enunciado é importante, por outro lado a técnica de demonstração não mais se repetirá, não ocorrendo perda alguma caso a leitura da demonstração seja omitida. No momento, o mais importante é compreender e ilustrar o resultado solucionando alguns problemas propostos. Em última análise a aspereza da demonstração é consequência de estarmos demonstrando a unicidade da representação de Jordan para um operador. De qualquer modo, a prova está na próxima seção. Finalmente, observamos que o Teorema da unicidade dimensional cíclica está mostrando o isomorfismo entre todas as decomposições cíclicas definidas por um operador.

**Exemplo 5.5.1** Suponha que os polinômios mínimo e característico de um operador linear  $A : \mathbb{R}^5 \rightarrow \mathbb{R}^5$  sejam, respectivamente,

$$m_A(t) = (t - 1)^2(t + 3) \quad \text{e} \quad p_A(t) = (t - 1)^3(t + 3)^2.$$

Pelo Teorema da decomposição primária e corolários sabemos que

$$\mathbb{R}^5 = \text{Nuc}(A - Id)^2 \oplus \text{Nuc}(A + 3Id),$$

com  $\dim \text{Nuc}(A - Id)^2 = 3$  e  $\dim \text{Nuc}(A + 3Id) = 2$ . Se existe uma decomposição  $A$ -cíclica de  $\mathbb{R}^5$  ela deve satisfazer as seguintes condições dimensionais,

$$\left\{ \begin{array}{l} \text{Nuc}(A - Id)^2 = \mathcal{C}(v_1) \oplus \mathcal{C}(v_2), \\ \text{grau}(t - 1)^2 = \underbrace{\dim \mathcal{C}(v_1)}_{=2} > \underbrace{\dim \mathcal{C}(v_2)}_{=1} \end{array} \right\} \left\{ \begin{array}{l} \text{Nuc}(A + 3Id) = \mathcal{C}(v_3) \oplus \mathcal{C}(v_4), \\ \text{grau}(t + 3) = \underbrace{\dim \mathcal{C}(v_3)}_{=1} = \underbrace{\dim \mathcal{C}(v_4)}_{=1}. \end{array} \right.$$

A representação matricial do operador na base ordenada  $\beta = \{v_1, A(v_1)\} \cup \{v_2\} \cup \{v_3\} \cup \{v_4\}$ , fica sendo a matriz ao lado. A matriz será a mesma se escolhermos qualquer outra decomposição cíclica respeitando-se, é claro, a ordem das parcelas da decomposição primária.  $\square$

$$[A]_{\beta} = \begin{bmatrix} 0 & -1 & & & \\ 1 & 2 & & & \\ & & 1 & & \\ & & & -3 & \\ & & & & -3 \end{bmatrix}$$

### Exercícios propostos 5.5.1

1. Descreva todas as decomposições  $A$ -cíclicas que podem ocorrer quando  $A$  é um operador linear num dos seguintes espaços.

- a)  $V = \mathbb{R}^2$ .   b)  $V = \mathbb{R}^3$ .   c)  $V = \mathbb{C}^2$ .   d)  $V = \mathbb{C}^3$ .

2. A representação matricial na base ordenada  $\beta = \{v_1, v_2, v_3, v_4, v_5\}$  do operador linear  $B : \mathbb{R}^5 \rightarrow \mathbb{R}^5$  é a matriz ao lado.
- a) Mostre que  $\mathbb{R}^5 = \mathcal{C}_B(v_1) \oplus \mathcal{C}_B(v_2)$  e calcule a dimensão de cada parcela.
- b) Determine uma decomposição  $A$ -cíclica quando  $A \equiv Id + B$ .
3. Descreva as possíveis decomposições cíclicas determinadas pelo operador  $A : \mathbb{R}^6 \rightarrow \mathbb{R}^6$  supondo que os polinômios mínimo e característico são os indicados.
- a)  $m_A(t) = (t^2 + 1)(t - 3)$  e  $p_A(t) = (t^2 + 1)^2(t - 3)^2$ .
- b)  $m_A(t) = (t^2 + 1)(t - 3)^2$  e  $p_A(t) = (t^2 + 1)^2(t - 3)^2$ .
- c)  $m_A(t) = (t^2 - 1)(t - 3)^2$  e  $p_A(t) = (t^2 - 1)^2(t - 3)^2$ .
4. Seja  $A$  um operador linear no espaço vetorial  $V$  de dimensão finita sobre  $\mathbb{K}$ . Prove as seguintes afirmações sobre espaços cíclicos.
- a) Se  $w \in \mathcal{C}(v)$  então  $\mathcal{C}(w) \subset \mathcal{C}(v)$ .
- b)  $q(A)(\mathcal{C}(v)) = \mathcal{C}(q(A)(v))$  para qualquer polinômio  $q(t) \in \mathbb{K}[t]$ .
- c) Se existe uma decomposição  $V = \mathcal{C}(v_1) \oplus \mathcal{C}(v_2) \oplus \cdots \oplus \mathcal{C}(v_m)$ , então
- $$p(A)(V) = \mathcal{C}(p(A)(v_1)) \oplus \mathcal{C}(p(A)(v_2)) \oplus \cdots \oplus \mathcal{C}(p(A)(v_m)),$$
- para qualquer polinômio  $p(t) \in \mathbb{K}[t]$ . (**Sugestão:** para o item 2 compare os conjuntos

$$\begin{cases} q(A)(\mathcal{C}(v)) = \{q(A) \circ p(A)(v) \in \mathbb{K}[t]\}, \\ \mathcal{C}(q(A)(v)) = \{p(A) \circ q(A)(v) \in \mathbb{K}[t]\}. \end{cases}$$

No item 3 utilize os anteriores e um critério para identificar uma soma direta.)

## 5.6 Apêndice

Esse apêndice é dedicado à prova do Teorema da unicidade dimensional da decomposição cíclica. Antes de iniciá-la examine o Exercício Proposto 4 da seção anterior.

**Demonstração do Teorema da unicidade dimensional** Recordamos que por hipótese a fatoração primária do polinômio mínimo de  $A$  é  $m_A(t) = p(t)^r$ , portanto qualquer vetor tem polinômio minimal do tipo  $m_v(t) = p(t)^z$  e por conseguinte temos que  $\dim \mathcal{C}(v) = \text{grau } p(t)^z$ .

Dada uma decomposição de  $V$  em subespaços  $A$ -cíclicos é conveniente apre-

sentá-la de outra forma,

$$V = \begin{cases} \mathcal{C}(v_{11}) \oplus \mathcal{C}(v_{12}) \oplus \cdots \oplus \mathcal{C}(v_{1k_1}) \oplus & \leftarrow \dim \mathcal{C}(v_{ij}) = \text{grau } p(t)^{r_1} \\ \mathcal{C}(v_{21}) \oplus \mathcal{C}(v_{22}) \oplus \cdots \oplus \mathcal{C}(v_{2k_2}) \oplus & \leftarrow \dim \mathcal{C}(v_{ij}) = \text{grau } p(t)^{r_2} \\ \vdots & \\ \mathcal{C}(v_{m1}) \oplus \mathcal{C}(v_{m2}) \oplus \cdots \oplus \mathcal{C}(v_{mk_m}), & \leftarrow \dim \mathcal{C}(v_{ij}) = \text{grau } p(t)^{r_m} \end{cases}$$

onde estão relacionados por linha os subespaços de dimensão  $\text{grau } p(t)^{r_1}$ , indicada ao lado, cujos expoentes estão ordenados como  $r = r_1 > r_2 > \cdots > r_m \geq 1$ .

Consideremos uma segunda decomposição  $A$ -cíclica,

$$V = \begin{cases} \mathcal{C}(u_{11}) \oplus \mathcal{C}(u_{12}) \oplus \cdots \oplus \mathcal{C}(u_{1n_1}) \oplus & \leftarrow \dim \mathcal{C}(u_{ij}) = \text{grau } p(t)^{s_1} \\ \mathcal{C}(u_{21}) \oplus \mathcal{C}(u_{22}) \oplus \cdots \oplus \mathcal{C}(u_{2n_2}) \oplus & \leftarrow \dim \mathcal{C}(u_{ij}) = \text{grau } p(t)^{s_2} \\ \vdots & \\ \mathcal{C}(u_{d1}) \oplus \mathcal{C}(u_{d2}) \oplus \cdots \oplus \mathcal{C}(u_{dn_d}), & \leftarrow \dim \mathcal{C}(u_{ij}) = \text{grau } p(t)^{s_d} \end{cases}$$

com  $r = s_1 > \cdots > s_d \geq 1$ . Comparemos inicialmente os comprimentos das primeiras linhas avaliando o operador  $p(A)^{s_1-1}$  em ambas decomposições, lembrando que por definição de decomposição cíclica temos as igualdades  $r = r_1 = s_1$ ,

$$p(A)^{s_1-1}(V) = \mathcal{C}(p(A)^{s_1-1}(v_{11})) \oplus \cdots \oplus \mathcal{C}(p(A)^{s_1-1}(v_{1k_1})),$$

$$p(A)^{s_1-1}(V) = \mathcal{C}(p(A)^{s_1-1}(u_{11})) \oplus \cdots \oplus \mathcal{C}(p(A)^{s_1-1}(u_{1n_1})),$$

como

$$\dim \mathcal{C}(p(A)^{s_1-1}(v_{1j})) = \text{grau } p(t) = \dim \mathcal{C}(p(A)^{s_1-1}(u_{1j})),$$

podemos contar as dimensões das parcelas e obter

$$\dim p(A)^{s_1-1}(V) = k_1 \text{grau } p(t) = n_1 \text{grau } p(t).$$

Portanto,  $k_1 = n_1$ , como queríamos provar.

A demonstração seguirá por indução. Vamos supor que já tenhamos mostrado que as duas decomposições satisfazem a unicidade dimensional até a  $j$ -ésima linha, isto é, se  $1 \leq i \leq j$ , as  $i$ -ésimas linhas têm o mesmo número de parcelas e as parcelas têm as mesmas dimensões, em outras palavras,

$$k_i = n_i \quad \text{e} \quad \dim \mathcal{C}(v_{pi}) = \dim \mathcal{C}(u_{pi}) \quad \text{para} \quad 1 \leq p \leq k_{im} \quad \text{e} \quad 1 \leq i \leq j.$$

Note que se numa decomposição existe a  $(j+1)$ -ésima linha na outra decomposição necessariamente existe uma  $(j+1)$ -linha, caso contrário, facilmente chegaríamos a uma contradição sobre a dimensão do espaço  $V$  utilizando a hipótese

de indução. Vamos supor por absurdo que vale uma desigualdade entre os  $(j+1)$ -ésimos expoentes, por exemplo,  $r_{j+1} > s_{j+1}$  (para  $r_{j+1} < s_{j+1}$  os argumentos são semelhantes). Sendo assim, avaliando o operador  $p(A)^{s_{j+1}-1}$  nas duas decomposições podemos afirmar que

1. Na primeira decomposição obtemos o espaço  $p(A)^{s_{j+1}-1}(V)$  com  $j+1$  linhas, cada linha com  $k_i$  parcelas e cada parcela na  $i$ -ésima linha tem dimensão igual ao  $\text{grau } p(t)^{r_i-(s_{j+1}-1)}$ ;
2. Na segunda decomposição obtemos o mesmo espaço  $p(A)^{s_{j+1}-1}(V)$  com  $j$  linhas, cada linha com  $n_i$  parcelas e cada parcela da  $i$ -ésima linha tem dimensão  $\text{grau } p(t)^{s_i-(s_{j+1}-1)}$ .

Isso é uma contradição sobre a dimensão de  $p(A)^{s_{j+1}-1}(V)$ . Com efeito. Por hipótese de indução sabemos que  $r_i = s_i$  e  $k_i = n_i$  se  $1 \leq i \leq j$ , e na primeira avaliação obtemos uma dimensão para  $p(A)^{s_{j+1}-1}(V)$  maior que aquela fornecida pela segunda avaliação pois esta última possui uma linha a menos. Isto termina a demonstração do Teorema da unicidade dimensional da decomposição cíclica.  $\square$

## Capítulo 6

# Representação canônica (I)

O Teorema da decomposição primária fornece uma macro decomposição de um operador. O objetivo desse (e do capítulo Representação canônica (II)) é mostrar como cada parcela da decomposição primária pode ser decomposta em subespaços cíclicos. Feito isso teremos condições de construir a representação de Jordan, o retrato de um operador linear com seu melhor perfil. A técnica utilizada é estudar operadores diagonalizáveis e operadores nilpotentes. Para deixar claro a estrutura desse capítulo faremos um resumo da linha de desenvolvimento. Novamente, recordamos que os únicos corpos considerados são  $\mathbb{R}$  ou  $\mathbb{C}$ .

Examinando a fatoração primária do polinômio minimal de um operador  $A : V \rightarrow V$ ,  $m_A(t) = p_1(t)^{r_1} p_2(t)^{r_2} \cdots p_k(t)^{r_k}$ , iremos construir uma base especial para cada parcela da decomposição primária correspondente a um dos quatro tipos de fatores:

$$\begin{array}{ll} \text{I} & p(t) = (t - \lambda), \\ \text{II} & p(t) = (t - \lambda)^r, \\ \text{III} & p(t) = (t - \lambda)^2 + \tau^2, \\ \text{IV} & p(t) = [(t - \lambda)^2 + \tau^2]^r, \end{array}$$

onde  $r \in \mathbb{N}$ ,  $r > 1$ ,  $\tau \in \mathbb{R}$  e  $\tau > 0$ .

É claro que um polinômio do tipo  $p(t) = (t - \lambda)^2 + \tau^2$ ,  $\tau > 0$  só ocorre como fator primo de um polinômio minimal quando o espaço vetorial é real. Nesse capítulo estudaremos o caso I e II deixando os outros para um apêndice, pois envolve uma técnica complementar. Para simplificar os enunciados das proposições sobre a existência de uma decomposição cíclica iremos assumir que a decomposição primária possui uma única parcela e nos casos com mais de uma parcela adaptamos o resultado fazendo a restrição do operador a cada uma das parcelas da decomposição primária.

## 6.1 Autovalores e autovetores

Na sequência  $A$  denotará um operador linear no espaço vetorial  $V$  de dimensão finita sobre o corpo  $\mathbb{K}$ . Nessa seção estudaremos as raízes do polinômio característico. Um escalar  $\lambda \in \mathbb{K}$  é um *autovalor* de  $A$  se, e somente se, existe um vetor não nulo  $v \in V$  tal que  $A(v) = \lambda v$ . Um vetor não nulo satisfazendo a condição  $A(v) = \lambda v$  é chamado de *autovetor associado* ao autovalor  $\lambda$ . O *autoespaço associado* ao autovalor  $\lambda$  é o conjunto definido por

$$V_\lambda = \{v \in V, A(v) = \lambda v\}.$$

Verifica-se sem esforço algum que  $V_\lambda$  é um subespaço invariante pelo operador  $A$ . Note que somente o vetor nulo em  $V_\lambda$  não é um autovetor associado ao autovalor  $\lambda$ . Em alguns textos autovalor é chamado de valor próprio ou valor característico e, acompanhando a terminologia, autovetor é chamado de vetor próprio ou vetor característico. Caracterizemos um autovalor de um operador  $A : V \rightarrow V$ .

**Proposição 6.1.1** *São equivalentes as seguintes afirmações.*

- a)  $\lambda \in \mathbb{K}$  é autovalor de  $A$ .
- b)  $\lambda Id - A : V \rightarrow V$  é um operador linear não invertível.
- c)  $\lambda \in \mathbb{K}$  é uma raiz do polinômio característico  $p_A(t)$ .

**Demonstração**  $a) \Rightarrow b)$  Se  $\lambda$  é autovalor de  $A$ , por definição existe um vetor não nulo tal que  $A(v) = \lambda v$ . Logo  $v$  pertence ao núcleo do operador  $\lambda Id - A : V \rightarrow V$ , significando que esse operador é não invertível.

$b) \Rightarrow a)$  Se  $\lambda Id - A$  não é invertível, pelo Teorema do núcleo e da imagem concluímos que  $Nuc(\lambda Id - A)$  não é trivial. Logo existe um vetor  $v$  não nulo tal que  $(\lambda Id - A)(v) = 0$ . É imediato concluir que  $\lambda$  é autovalor e  $v$  é autovetor associado, mostrando  $a)$ .

$b) \Leftrightarrow c)$  Basta utilizar um critério já demonstrado anteriormente. O operador  $\lambda Id - A$  não é invertível  $\Leftrightarrow 0 = \det(\lambda Id - A) = p_A(\lambda)$ .  $\square$

Quando  $V$  é um espaço vetorial real, alguns operadores podem não ter autovalores, para que isso ocorra é suficiente que o polinômio característico não tenha raízes reais. Em tal caso, o espaço  $V$  deve ter dimensão par. Se  $m_A(t) = (t - \lambda)p_2(t)^{r_2} \cdots p_k(t)^{r_k}$  é a decomposição primária do polinômio característico de  $A$ , o leitor pode mostrar que a parcela  $Nuc(A - \lambda Id)$  da decomposição primária de  $V$  é o autoespaço  $V_\lambda$ .

A condição  $b)$  da proposição acima permite transferir o conceito de autovalor para matrizes quadradas  $N \in M(n, \mathbb{K})$ . Um escalar  $\lambda \in \mathbb{K}$  é *autovalor da matriz*

$N$  se, e somente se, a matriz  $\lambda I - N$  não é invertível. Da mesma forma, um autovalor de uma matriz é uma raiz do seu polinômio característico. Portanto, os autovalores de um operador  $A : V \rightarrow V$  são os autovalores de uma representação matricial  $[A]_\beta$ .

**Exercício 6.1.1** Qual a dimensão do subespaço cíclico gerado por um autovetor de um operador linear?  $\square$

**Exemplo 6.1.1** Calculando o polinômio característico do operador linear

$$A : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad A(x, y) = (x + y, x + y),$$

obtemos o polinômio  $p_A(t) = t(t - 2)$ , logo os seus autovalores são escalares  $\lambda_1 = 0$  e  $\lambda_2 = 2$ . O Teorema da decomposição primária garante a decomposição  $V = \text{Nuc } A \oplus \text{Nuc}(A - 2Id)$ . É imediato concluir que cada parcela da decomposição é um autoespaço associado a um autovalor, fato que permite-nos reescrever a decomposição como soma direta dos autoespaços  $V = V_{\lambda_1} \oplus V_{\lambda_2}$ . Com um cálculo simples encontramos que  $u = (1, -1)$  e  $v = (1, 1)$  são autovetores associados aos autovalores  $\lambda_1 = 0$  e  $\lambda_2 = 2$ , respectivamente. Na base ordenada  $\beta = \{u, v\}$  a representação de  $A$  é a matriz diagonal

$$[A]_\beta = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}. \quad \square$$

Embora trivial, deixaremos registrado o primeiro resultado sobre a existência de uma decomposição cíclica.

**Teorema 6.1.1 (Teorema da decomposição cíclica I)** *Seja  $A$  um operador linear num espaço vetorial  $V$  de dimensão  $n$  sobre  $\mathbb{K}$  com polinômio minimal  $m_A(t) = t - \lambda$ . Então existem vetores  $v_1, v_2, \dots, v_n \in V$  tais que*

$$\begin{cases} V = Z(v_1) \oplus Z(v_2) \oplus \dots \oplus Z(v_n), \\ \text{grau}(t - \lambda) = \dim Z(v_1) = \dim Z(v_2) = \dots = \dim Z(v_n). \end{cases}$$

*Essa decomposição  $A$ -cíclica é dimensionalmente única.*

**Demonstração** É claro que  $V = V_\lambda = \text{Nuc}(A - \lambda Id)$ . Portanto, se escolhermos uma base ordenada qualquer  $\beta = \{v_1, v_2, \dots, v_n\}$ , o espaço decompõe-se numa soma direta dos espaços  $A$ -cíclicos  $\{\mathcal{C}(v_i)\}_{i=1}^n$ . Como a base é formada de autovetores, os subespaços cíclicos têm dimensão 1.  $\square$

Nas condições da proposição acima, a base  $\beta$  construída é chamada de *base de Jordan* e a *representação Jordan* de  $A$  é a matriz diagonal  $n \times n$ ,  $[A]_\beta = \{\lambda, \lambda, \dots, \lambda\}$ .

$$[A]_\beta = \begin{bmatrix} \lambda & & & \\ & \lambda & & \\ & & \ddots & \\ & & & \lambda \end{bmatrix}.$$

### Exercícios propostos 6.1.1

Todos os espaços considerado a seguir têm dimensão finita.

1. Qualquer operador num espaço vetorial complexo ou num espaço vetorial real de dimensão ímpar tem pelo menos um autovalor. Prove a afirmação.
2. Se  $m_A(t) = (t-\lambda)p_2(t)^{r_2} \cdots p_k(t)^{r_k}$  é a decomposição primária do polinômio minimal de  $A : V \rightarrow V$ , prove que a parcela  $\text{Nuc}(A - \lambda Id)$  da decomposição primária correspondente é o autoespaço  $V_\lambda$ .
3. Sejam  $v_1, v_2, \dots, v_k$  autovetores do operador  $A : V \rightarrow V$  associados aos autovalores  $\lambda_1, \lambda_2, \dots, \lambda_k$ , respectivamente. Se os autovalores são dois a dois distintos, mostre que os autovetores são linearmente independentes.
4. Se todo vetor não nulo  $v \in V$  é autovetor associado a um autovalor  $\lambda$  de um operador linear  $A : V \rightarrow V$ , então  $A = \lambda Id$ .
5. Seja  $A$  um operador linear num espaço vetorial  $V$  de dimensão finita sobre o corpo  $\mathbb{K}$ . Se  $p(t) \in \mathbb{K}[t]$  é um polinômio tal que  $p(A) \equiv 0$ , então toda raiz de  $p(t)$  é um autovalor de  $A$ ?
6. Demonstre as seguintes afirmações sobre um operador linear  $A : V \rightarrow V$ .
  - a) Se  $\lambda$  é um autovalor de  $A$  então  $p(\lambda)$  é um autovalor de  $p(A)$ , para qualquer polinômio  $p(t) \in \mathbb{K}[t]$ .
  - b)  $A$  é invertível  $\Leftrightarrow$  todo autovalor de  $A$  é diferente de zero.
  - c) Quando  $A$  é invertível, o operador  $A^{-1}$  é um polinômio em  $A$ .
  - d) Se  $A$  é invertível então o conjunto de autovetores de  $A$  e  $A^{-1}$  coincidem.
7. Utilizando um argumento indutivo sobre a dimensão demonstre que todo operador linear  $A$  num espaço vetorial complexo  $V$  de dimensão finita admite uma representação por uma matriz triangular superior  $[A]_\beta$ .

## 6.2 Operadores diagonalizáveis

Um operador linear  $A : V \rightarrow V$  é *diagonalizável* se existe uma base  $\beta \subset V$  formada por autovetores. O termo diagonalizável é sugestivo, pois numa base

$\beta = \{v_1, v_2, \dots, v_n\}$  formada por autovetores a representação do operador é uma matriz diagonal,

$$[A]_\beta = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\},$$

desde que  $A(v_i) = \lambda_i v_i$ . Observamos que os autovalores não são necessariamente distintos. Para saber sob quais condições podemos diagonalizar um operador, basta examinar a fatoração primária do seu polinômio minimal e verificar se ela é da forma  $m_A(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_k)$  onde  $\lambda_i \neq \lambda_j$  se  $i \neq j$ .

**Proposição 6.2.1** *Seja  $A$  um operador linear num espaço vetorial de dimensão finita  $V$ . Então  $A$  é diagonalizável  $\Leftrightarrow$  o polinômio minimal de  $A$  fatora-se em um produto de polinômios lineares distintos.*

**Demonstração**  $\Rightarrow$ ) Escolha uma base ordenada  $\gamma = \{v_1, v_2, \dots, v_n\}$  constituída de autovetores do operador. Reindexando  $\gamma$  podemos construir uma nova base ordenada

$$\beta = \{v_{11}, \dots, v_{1s_1}\} \vec{\cup} \cdots \vec{\cup} \{v_{k1}, \dots, v_{ks_k}\}$$

na qual todos os elementos de um mesmo subconjunto indicado são autovetores de um único autovalor,  $A(v_{ij}) = \lambda_i v_{ij}$ , para todo  $j$ ,  $1 \leq j \leq k_i$ . Feito isso, representamos o operador  $A$  como uma matriz diagonal,

$$[A]_\beta = \text{diag}\{\lambda_1 I_{s_1}, \lambda_2 I_{s_2}, \dots, \lambda_k I_{s_k}\},$$

em que  $I_{s_i}$  é a matriz identidade  $s_i \times s_i$ . Dessa representação matricial segue que podemos decompor  $V$  em uma soma direta de autoespaços

$$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k}.$$

Logo, dado um vetor  $v \in V$  podemos decompô-lo em  $k$  parcelas,

$$v = v_1 + v_2 + \cdots + v_k \text{ com } v_i \in V_{\lambda_i}.$$

Também pela representação matricial é possível concluir que o polinômio característico do operador é da forma

$$p_A(t) = (t - \lambda_1)^{s_1} (t - \lambda_2)^{s_2} \cdots (t - \lambda_k)^{s_k}.$$

Consideremos o polinômio

$$m(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_k),$$

avaliemos  $m(t)$  no operador e calculemos a avaliação num vetor  $v = v_1 + v_2 + \cdots + v_k$ . Como os fatores de  $m(A)$  comutam, desde que são polinômios em  $A$ , e  $(A - \lambda_i Id)(v_i) = 0$  temos que

$$m(A)(v) = \sum_{i=1}^k m(A)(v_i)$$

$$\begin{aligned}
&= \sum_{i=1}^k \left( (A - \lambda_1 Id) \circ \cdots \circ (A - \widehat{\lambda_i Id}) \circ \cdots \circ (A - \lambda_k Id) \right) \circ \underbrace{(A - \lambda_i Id)(v_i)}_{=o} \\
&= o.
\end{aligned}$$

Isto é suficiente para demonstrar que  $m(t) = m_A(t)$ , pois ele é aquele polinômio de menor grau no ideal anulador de  $A$  cuja fatoração contém todos os fatores primos do polinômio característico.

$\Leftrightarrow$ ) Vamos assumir que

$$m_A(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_k), \quad \text{com } \lambda_i \neq \lambda_j \text{ se } i \neq j.$$

Sendo assim, cada parcela da decomposição primária

$$V = Nuc(A - \lambda_1 Id) \oplus Nuc(A - \lambda_2 Id) \oplus \cdots \oplus Nuc(A - \lambda_k Id)$$

é um autoespaço. Escolhida uma base ordenada  $\beta_i$  de  $V_{\lambda_i} = Nuc(A - \lambda_i Id)$ , a união ordenada  $\beta = \vec{\cup} \beta_i$  é uma base ordenada de  $V$  formada por autovetores. Por definição,  $A$  é diagonalizável.  $\square$

Diremos que uma matriz quadrada é *diagonalizável* quando ela é conjugada a uma matriz diagonal.

### Exercícios propostos 6.2.1

- Verifique se o operador é diagonalizável e, em caso positivo, determine uma base ordenada que diagonaliza o operador e dê a representação matricial nesta base.
  - $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (x, x + 2y, x + y - 3z)$ .
  - $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (x - y, 2x + 2y + 2z, x + y - 2z)$ .
  - $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (x - 2y, 3y - 4z, -y + 3z)$ .
  - $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (x, x + 2y, x + y - 3z)$ .
  - $A : \mathbb{C}^3 \rightarrow \mathbb{C}^3$ ,  $A(x, y, z) = (x, -2x + y + 2z, -2x + 2y + 3z)$ .
  - $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $A(x, y) = (5x - y, x + 3y)$ .
  - $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (y, z, -x - y)$ .

- Mostre que a matriz  $N \in M(3, \mathbb{R})$ , descrita abaixo, é diagonalizável.

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

- Mostre que a matriz  $N \in M(n, \mathbb{K})$ , descrita abaixo é conjugada a uma matriz diagonal.

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

- Demonstre que todo operador  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  do tipo  $A(x, y) = (ax + by, bx + cy)$  é diagonalizável.

5. Mostre que toda involução e todo operador idempotente num espaço vetorial de dimensão finita é diagonalizável.
6. A restrição de um operador diagonalizável a um subespaço invariante é um operador diagonalizável? Justifique sua resposta.
7. Suponha que  $A : V \rightarrow V$  é um operador linear diagonalizável num espaço vetorial de dimensão  $n$ . Prove que  $V$  é  $A$ -cíclico  $\Leftrightarrow A$  tem  $n$  autovalores distintos.
8. Se  $A : V \rightarrow V$  é um operador diagonalizável e  $W_1 \subset V$  é um subespaço invariante, então existe outro subespaço invariante  $W_2 \subset V$  tal que  $V = W_1 \oplus W_2$ .
9. Se  $p_A(t) = (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_n)$  com  $\lambda_i \neq \lambda_j$  se  $i \neq j$  é o polinômio característico do operador linear  $A : V \rightarrow V$ , demonstre que  $\text{tr } A^r = \sum_{k=1}^n \lambda_k^r$ .
10. Suponha que  $A : V \rightarrow V$  é um operador diagonalizável tal que  $p_A(t) = m_A(t)$ . Quantos subespaços invariantes existem?

11. Encontre uma condição sobre os valores de  $a, b, c, d, f$  necessária e suficiente para que a matriz  $N$ , descrita ao lado, seja conjugada a uma matriz diagonal.

$$N = \begin{bmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ b & c & 3 & 0 \\ d & e & f & 3 \end{bmatrix}.$$

12. Prove que um operador linear é diagonalizável se, e somente se, qualquer representação matricial do operador é uma matriz diagonalizável.
13. Um operador que comuta com qualquer operador diagonalizável é uma homotetia. Prove essa afirmação.
14. Determine  $N \in M(3, \mathbb{R})$  tal que
15. Calcule  $N^{17}$  quando  $N$  é a matriz descrita abaixo.

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

$$N = \begin{bmatrix} 6 & -2 \\ -2 & 9 \end{bmatrix}.$$

16. Dado um operador linear  $A$  num espaço vetorial  $V$  de dimensão dois prove que  $V = \mathcal{C}(v)$  ( $A$ -cíclico) ou  $A$  é uma homotetia. Os dois fatos não ocorrem simultaneamente.
17. Suponha que um operador linear  $A$  num espaço vetorial real  $V$  de dimensão  $n$  tenha  $n$  autovalores distintos. Demonstre que nem sempre existe um operador linear  $B$  tal que  $A = B^2$ .
18. Fixado uma matriz  $N \in M(2, \mathbb{K})$ , considere o operador linear  $A : M(2, \mathbb{K}) \rightarrow M(2, \mathbb{K})$ ,  $A(X) = XN$ . Mostre que  $A$  não é diagonalizável.
19. Dada a matriz  $A \in M(2, \mathbb{K})$ , defina um operador linear  $A : M(2, \mathbb{K}) \rightarrow M(2, \mathbb{K})$  por  $A(X) = NX - XN$ . Se  $N$  é diagonalizável então  $A$  é diagonalizável.
20. Sejam  $D$  um operador diagonalizável num espaço vetorial  $V$  de dimensão  $n$  sobre o corpo  $\mathbb{K}$  com autovalores  $\{\lambda_j\}_{j=1}^n$  (possivelmente com repetições) e  $p(t) \in \mathbb{K}[t]$ . Prove que  $p(D)$  é diagonalizável e que  $\{p(\lambda_j)\}_{j=1}^n$  são os autovalores de  $p(D)$ .

### 6.3 Operadores nilpotentes

No Teorema da decomposição primária surge um operador linear com propriedades especiais. As restrições  $p(A) : Nuc p_i(A)^{r_i} \rightarrow Nuc p_i(A)^{r_i}$  satisfazem a identidade  $p_i(A)^{r_i} \equiv 0$ . Tais tipos de operadores recebem um nome particular.

Diz-se que um operador  $B$  num espaço vetorial de dimensão finita  $V$  é *nilpotente* se seu polinômio minimal é do tipo  $m_B(t) = t^r$ . O inteiro  $r$  é chamado de *nilpotência* de  $B$  e satisfaz a desigualdade  $r \leq \dim V$  pois

$$r = \text{grau } m_B(t) \leq \text{grau } p_B(t) = \dim V.$$

**Exercício 6.3.1** Mostre os três fatos sobre um operador nilpotente  $B : V \rightarrow V$ .

a) O operador nilpotente define uma sequência de núcleos do tipo

$$V = Nuc B^r \supset Nuc B^{r-1} \supset \dots \supset Nuc B \supset Nuc B^0 = \{o\}.$$

b)  $B$  induz por restrição um operador  $B : Nuc B^i \rightarrow Nuc B^{i-1}$ .

c) Um vetor  $v$  que pertence ao subespaço  $Nuc B^i$  mas não pertence ao subespaço  $Nuc B^{i-1}$  tem polinômio minimal  $m_v(t) = t^i$ .

De modo análogo, diremos que uma matriz  $N \in M(n, \mathbb{K})$  é nilpotente com nilpotência  $r$  se seu polinômio minimal é o polinômio  $m_N(t) = t^r$ . A relação entre as duas definições é óbvia, afirmar que um operador  $B$  é nilpotente é equivalente a afirmar que qualquer representação matricial  $[B]_\beta$  é uma matriz nilpotente.

**Exercício 6.3.2** Verifique que a matriz  $N \in M(2, \mathbb{K})$  é nilpotente e calcule o traço e o determinante,

$$N = \begin{bmatrix} 10 & -4 \\ 25 & -10 \end{bmatrix}. \quad \square$$

**Exemplo 6.3.1** Seja  $\mathbb{K}_n[t]$  o anel dos polinômios com grau menor ou igual à  $n$ . A derivação  $D : \mathbb{K}_n[t] \rightarrow \mathbb{K}_n[t]$ ,  $D(p(t)) = p'(t)$ , é nilpotente e  $m_D(t) = t^{n+1}$ .  $\square$

**Exemplo 6.3.2** O operador linear

$$B : \mathbb{K}^n \rightarrow \mathbb{K}^n, \quad B(x_1, x_2, \dots, x_n) = (0, x_1, x_2, \dots, x_{n-1}),$$

é nilpotente com polinômio minimal  $m_B(t) = t^n = p_B(t)$ . Portanto, o espaço é  $B$ -cíclico. Nesse caso, é simples verificar que  $\mathbb{R}^n = \mathcal{C}(e_1)$  em que  $e_1$  é o primeiro

elemento da base canônica pois  $\alpha \in \mathcal{C}(e_1)$ . Logo, sua representação na base  $\beta = \{e_1, B(e_2), \dots, B^{r-1}(e_1)\} = \alpha$  é a matriz  $n \times n$  descrita ao lado. Um dos nossos objetivos é mostrar que qualquer operador nilpotente pode ser representado por uma diagonal de matrizes como essa.  $\square$

$$[B]_{\beta} = \begin{bmatrix} 0 & 0 & & & \\ 1 & 0 & & & \\ & 1 & \ddots & & \\ & & \ddots & 0 & 0 \\ & & & 1 & 0 \end{bmatrix}.$$

**Exemplo 6.3.3** Vamos supor que  $A : V \rightarrow V$  é um operador linear tal que  $m_A(t) = (t - \lambda)^n$  no qual  $n = \dim V$ . Considere o operador linear

$$B : V \rightarrow V, \quad B(v) = (A - \lambda Id)(v).$$

Sendo assim, é imediato concluir que o polinômio minimal de  $B$  é o polinômio  $m_B(t) = t^n$ . Por definição, o operador  $B$  é nilpotente com nilpotência  $n$ , de onde segue que o espaço vetorial é  $B$ -cíclico. Escolhendo um vetor  $v \in V$  tal que  $V = \mathcal{C}(v)$  e considerando a base cíclica  $\beta = \{v, B(v), \dots, B^{n-1}(v)\}$  temos para representação matricial de  $B$  na base  $\beta$  a matriz  $n \times n$  semelhante à matriz do exemplo anterior. Como  $[B]_{\beta} = [A - \lambda Id]_{\beta}$ , podemos escrever

$$[B]_{\beta} = \begin{bmatrix} 0 & 0 & & & \\ 1 & 0 & & & \\ & 1 & \ddots & & \\ & & \ddots & 0 & 0 \\ & & & 1 & 0 \end{bmatrix}, \quad [A]_{\beta} = \begin{bmatrix} \lambda & 0 & & & \\ 1 & \lambda & & & \\ & 1 & \ddots & & \\ & & \ddots & \lambda & 0 \\ & & & 1 & \lambda \end{bmatrix}. \quad \square$$

Na demonstração do próximo teorema de decomposição necessitaremos de algumas informações.

**Lema 6.3.1** *A sequência de núcleos de um operador nilpotente  $B$  com nilpotência  $r$  satisfaz as seguintes propriedades.*

a) *Todas as inclusões são próprias,*

$$V = \text{Nuc } B^r \supsetneq \text{Nuc } B^{r-1} \supsetneq \dots \supsetneq \text{Nuc } B \supsetneq \text{Nuc } B^0 = \{o\}.$$

b) *Se  $W$  é qualquer subespaço complementar à  $\text{Nuc } B^j$  em  $\text{Nuc } B^{j+1}$ , em outras palavras, se  $\text{Nuc } B^{j+1} = W \oplus \text{Nuc } B^j$ , então  $B(W) \cap \text{Nuc } B^{j-1} = \{o\}$ .*

**Demonstração** a) Seja  $v \in V = \text{Nuc } B^r$  um vetor cujo polinômio minimal é igual ao polinômio minimal de  $B$ , isto é,  $m_v(t) = t^r$ . É claro que o polinômio minimal de  $v_i = B^i(v)$  é  $m_i(t) = t^{r-i}$ , portanto  $v_i \in \text{Nuc } B^{r-i}$ . Suponha por

absurdo que  $Nuc B^i = Nuc B^{i-1}$ , para algum  $i$ , então  $B^{(i-1)}(v_i) = o$ , significando que  $B^{i-1}(B^{r-i}(v)) = o$ . Logo o polinômio  $m(t) = t^{r-1}$  está no ideal anulador de  $v$  e tem grau menor que o grau do seu polinômio minimal, uma contradição. Isso mostra que todas as inclusões são próprias.

b) A argumentação é a mesma. Se existisse um vetor não nulo na interseção  $B(W) \cap Nuc B^{j-1}$  esse vetor teria polinômio mínimo  $m_v(t) = t^j$  e o polinômio  $p(t) = t^{j-1}$  pertenceria ao ideal anulador  $\mathfrak{S}_v$ , evidentemente uma contradição.  $\square$

### Exercícios propostos 6.3.1

1. Prove que toda matriz conjugada a uma matriz nilpotente é nilpotente.
2. Descreva uma matriz  $n \times n$  que é simultaneamente diagonalizável e nilpotente.
3. O produto de duas matrizes  $n \times n$  que comutam com uma delas nilpotente, implica que a outra é uma matriz nilpotente. Verifique essa afirmação.
4. Uma matriz  $2 \times 2$ ,  $N$ , é nilpotente  $\Leftrightarrow tr N = 0$  e  $det N = 0$ .
5. O operador obtido por restrição de um operador linear nilpotente a um subespaço invariante é um operador nilpotente?
6. Demonstre que um operador nilpotente  $B$  com nilpotência  $r$  num espaço vetorial de dimensão finita  $V$  determina a sequência de subespaço

$$V \supseteq Im B \supseteq Im B^2 \supseteq \dots \supseteq Im B^{r-1} \supseteq Im B^r = \{o\}.$$

7. Seja  $A$  um operador linear no espaço vetorial de dimensão finita  $V$ . Prove a afirmação: se  $A^s \equiv 0$  para algum inteiro  $s > dim V$  então  $A^r \equiv 0$  para algum inteiro  $r \leq dim V$ .

8. Mostre que o operador linear  $B : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ ,  $B(x_1, x_2, x_3, x_4) = (x_2, x_3, x_4, 0)$ , é nilpotente e encontre uma base ordenada  $\beta \subset \mathbb{R}^4$  tal que a representação matricial de  $B$  é uma matriz como descrita ao lado.

$$[B]_{\beta} = \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & 1 & 0 & \\ & & 1 & 0 \end{bmatrix}.$$

9. Verifique que as matrizes ao lado são nilpotente calculando diretamente as potências  $J, J^2, J^3$ , etc.

$$J_1 = \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & 1 & 0 & \\ & & 1 & 0 \end{bmatrix}, \quad J_2 = \begin{bmatrix} 0 & 0 & 0 \\ a & 0 & 0 \\ c & b & 0 \end{bmatrix}.$$

10. Uma matriz triangular inferior (respect. superior) com todos os elementos da diagonal principal iguais a zero é nilpotente?
11. Seja  $N \in M(n, \mathbb{K})$ . Se  $tr N^j = 0$  para todo  $j \geq 1$ , então  $N$  é nilpotente.
12. Se  $N$  e  $P \in M(n, \mathbb{K})$  são matrizes tais que  $NP - PN$  comuta com  $N$  então  $NP - PN$  é nilpotente (*Lema de Jacobson*).

## 6.4 Decomposição cíclica II

Nessa seção construiremos uma decomposição cíclica para um operador linear cujo polinômio minimal é potência positiva do polinômio  $p(t) = (t-1)$ . A demonstração é construtiva e deve ser repetida nos exemplos. Antes de enunciá-lo, colocaremos uma observação como exercício.

**Exercício 6.4.1** Dados um operador linear  $A : V \rightarrow V$  e um escalar  $\lambda \in \mathbb{K}$ , considere o operador linear  $B : V \rightarrow V$ ,  $B(v) = (A - \lambda Id)(v)$ . Mostre que para todo  $v \in V$  vale a igualdade de subespaços cíclicos  $\mathcal{C}_A(v) = \mathcal{C}_B(v)$ . Sugestão: utilize o desenvolvimento de Taylor de um polinômio em torno de  $t = 1$ .  $\square$

**Teorema 6.4.1 (Teorema da decomposição cíclica II)** *Seja  $A$  um operador num espaço vetorial  $V$  de dimensão finita sobre o corpo  $\mathbb{K}$  com polinômio minimal  $m_A(t) = (t - \lambda)^r$ . Então existem vetores  $v_1, v_2, \dots, v_n \in V$  tais que*

$$\begin{cases} V = \mathcal{C}(v_1) \oplus \mathcal{C}(v_2) \oplus \dots \oplus \mathcal{C}(v_n), \\ \text{grau } m_A(t) = \dim \mathcal{C}(v_1) \geq \dim \mathcal{C}(v_2) \geq \dots \geq \dim \mathcal{C}(v_n). \end{cases}$$

*Essa decomposição é dimensionalmente única.*

**Demonstração** Pelo exercício acima é suficiente construir uma decomposição cíclica para o operador  $B : V \rightarrow V$ ,  $B(v) = (A - Id)(v)$  que é nilpotente com  $m_B(t) = t^r$ . Uma decomposição cíclica para  $B$  é também uma decomposição cíclica para  $A$ . Consideremos a sequência de núcleos

$$V = Nuc B^r \supsetneq Nuc B^{r-1} \supsetneq \dots \supsetneq Nuc B \supsetneq Nuc B^0 = \{o\}.$$

Sendo as inclusões próprias, é possível escolher um subespaço  $W_0$  complementar ao subespaço  $Nuc B^{r-1}$  em  $V = Nuc B^r$ , e uma base ordenada  $\beta_0$  de  $W_0$ . Guardemos esses dados,

$$V = W_0 \oplus Ker B^{r-1} \quad \text{e} \quad \beta_0 = \{v_{01}, v_{02}, \dots, v_{0k_0}\} \subset W_0.$$

Observemos que cada vetor na base  $\beta_0$  tem polinômio minimal  $m_{v_{0i}}(t) = m_B(t) = t^r$ . Isso implica que o subespaço cíclico gerado por um vetor  $v_{0i} \in \beta_0$  tem dimensão  $r$ . Mostremos que o conjunto de vetores

$$B(\beta_0) = \{B(v_{01}), B(v_{02}), \dots, B(v_{0k_0})\}$$

é linearmente independente no espaço  $Nuc B^{r-1}$ . Com efeito, considere a combinação linear

$$a_1 B(v_{01}) + a_2 B(v_{02}) + \dots + a_{k_0} B(v_{0k_0}) = o.$$

Avaliando o operador  $B^{r-2}$  na combinação linear temos as implicações

$$a_1 B^{r-1}(v_{01}) + a_2 B^{r-1}(v_{02}) + \cdots + a_{k_0} B^{r-1}(v_{0k_0}) = o \Rightarrow$$

$$a_1 v_{01} + a_2 v_{02} + \cdots + a_{k_0} v_{0k_0} \in \text{Ker} B^{r-1} \Rightarrow$$

$$a_1 v_{01} + a_2 v_{02} + \cdots + a_{k_0} v_{0k_0} \in \text{Ker} B^{r-1} \cap W_0 = \{o\} \Rightarrow$$

$$a_1 = a_2 = \cdots = a_{k_0} = 0.$$

Pelo que foi demonstrado no lema anterior é possível escolher um subespaço  $W_1 \subset \text{Nuc} B^{r-1}$  satisfazendo as condições

$$\text{Nuc} B^{r-1} = W_1 \oplus \text{Nuc} B^{r-2} \quad \text{e} \quad B(W_0) \subset W_1.$$

Escolhemos para base ordenada de  $W_1$  um conjunto da forma

$$\beta_1 = B(\beta_0) \vec{\cup} \{v_{11}, v_{12}, \dots, v_{1k_1}\}.$$

Tal construção decompõe  $V$  em soma direta,

$$V = W_0 \oplus W_1 \oplus \text{Nuc} B^{r-2}$$

e fornece uma base ordenada  $\beta_1$  de  $W_1$  na qual todo elemento possui polinômio minimal  $t^{r-1}$ , implicando que o espaço cíclico gerado por um elemento da base  $\beta_1$  tem dimensão  $r-1$ . Repetindo esse processo um número  $r$  de vezes, sempre recorrendo-se ao lema, construímos uma soma direta  $V = W_0 \oplus W_1 \oplus \cdots \oplus W_{r-1}$  na qual cada parcela satisfaz a condição

$$W_i \subset \text{Nuc} B^{r-i} \quad \text{e} \quad W_i \cap \text{Nuc} B^{r-(i+1)} = \{0\}.$$

Além disso, construímos uma base ordenada para  $W_i$  pela regra indutiva

$$\beta_i = B(\beta_{i-1}) \vec{\cup} \{v_{i1}, v_{i2}, \dots, v_{ik_i}\}$$

onde cada elemento  $v_{ij}$  desta base é um vetor com polinômio mínimo  $t^{r-i}$ . A tabela abaixo apresenta esquematicamente a base ordenada de  $V$  obtida no final do processo,

$W_0$	$v_{o1}$	..	$v_{ok_0}$					
$W_1$	$B(v_{o1})$	..	$B(v_{ok_0})$	$v_{11}$	..	$v_{1k_1}$		
..	..		..	..		..		
$W_{r-1}$	$B^{r-1}(v_{o1})$	..	$B^{r-1}(v_{ok_0})$	$B^{r-2}(v_{11})$	..	$B^{r-2}(v_{1k_1})$	$v_{r-1,1}$	.. $v_{r-1,k_{r-1}}$
	$\mathcal{C}(v_{o1})$	..	$\mathcal{C}(v_{ok_0})$	$\mathcal{C}(v_{11})$	..	$\mathcal{C}(v_{1k_1})$	$\mathcal{C}(v_{r-1,1})$	.. $\mathcal{C}(v_{r-1,k_{r-1}})$

A  $i$ -ésima linha descreve a base ordenada para o espaço  $W_i$  e a  $j$ -ésima coluna contém a base ordenada para o espaço  $B$ -cíclico gerado pelo vetor no alto da mesma coluna. Como vale a soma direta  $V = W_0 \oplus W_1 \oplus \cdots \oplus W_{r-1}$ , a dimensão de  $V$  é a soma dos número de vetores em cada linha. Por outro lado, é evidente que obtemos o mesmo total ao somarmos o número de vetores em cada coluna, isto é, a dimensão de  $V$  é a soma das dimensões dos espaços  $B$ -cíclicos indicados na base da tabela. Portanto,  $V$  é a soma direta desses espaços  $B$ -cíclicos.

Finalmente, o vetor  $v_{ij} \in W_i$  no alto da tabela tem, por construção, polinômio minimal  $t^{r-i}$ , logo as dimensões dos espaços  $B$ -cíclicos indicados na base da tabela formam uma sequência decrescente na ordem de apresentação, da esquerda para a direita, com o primeiro elemento satisfazendo a condição  $\dim \mathcal{C}(v_{01}) = \text{grau } m_B(t)$ . Isto mostra a existência da decomposição  $A$ -cíclica. A unicidade dimensional já foi mostrada no capítulo anterior.  $\square$

Com as hipóteses e notação do Teorema da decomposição cíclica II, temos que

**Corolário 6.4.1** *A dimensão do núcleo de  $B = A - \lambda Id$  é igual ao número de parcelas da decomposição cíclica.*

**Demonstração** Examinando a tabela construída na demonstração, verificamos que  $W_{r-1} = \text{Nuc } B$ . Como a dimensão desse núcleo é igual ao número de vetores na última linha, o número de vetores nessa linha é igual ao número de colunas da tabela e cada coluna corresponde a uma base de uma das parcelas da decomposição cíclica, temos mostrado o corolário.  $\square$

### Exercícios propostos 6.4.1

- Se  $A : V \rightarrow V$  é nilpotente com nilpotência  $r = \dim V - 1$  então  $A$  admite dois autovetores linearmente independentes.
- Considere o operador linear  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $A(x, y) = (10x + 25y, -4x - 10y)$ .
  - Prove que  $\mathbb{R}^2 = \mathcal{C}(v)$  para algum  $v \in \mathbb{R}^2$ .
  - Determine uma base  $\beta$  tal que a representação matricial seja a matriz dada ao lado.
 
$$[A]_\beta = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$
- Demonstre que toda matriz nilpotente,  $2 \times 2$  e não nula é conjugada à matriz descrita ao lado.
 
$$N = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$
- Considere o operador  $B : \mathbb{R}^5 \rightarrow \mathbb{R}^5$ ,  $B(t, x, y, z, w) = (0, t, x, 0, y)$ .
  - Mostre que  $B$  é nilpotente com nilpotência 3.
  - Verifique que  $\mathbb{R}^5 = \mathcal{C}(e_1) \oplus \mathcal{C}(e_3)$ .
  - Calcule  $[B]_\beta$  onde  $\beta = \{e_1, B(e_1), B^2(e_1)\} \cup \{e_3, B(e_3)\}$ .

- d) Represente nessa mesma base ordenada o operador  $A : \mathbb{R}^5 \rightarrow \mathbb{R}^5$ ,  
 $A(t, x, y, z, w) = (2t, t + 2x, x + 2y, 2z, y + 2w)$ .

5. Suponha que  $N \in M(n, \mathbb{K})$  é nilpotente. Prove as afirmações.

- a)  $\det(N + I) = 1$ .  
 b) Se  $P \in M(n, \mathbb{K})$  é uma matriz que comuta com  $N$  então  $\det(N + P) = \det P$ .

## 6.5 Representação de Jordan

Suponha que

$$\begin{cases} V = \mathcal{C}(v_1) \oplus \mathcal{C}(v_2) \oplus \cdots \oplus \mathcal{C}(v_n) \\ \text{grau } m_A(t) = \dim \mathcal{C}(v_1) \geq \dim \mathcal{C}(v_2) \geq \cdots \geq \dim \mathcal{C}(v_n) \end{cases},$$

é uma decomposição cíclica determinada pelo operador  $A : V \rightarrow V$  com polinômio minimal  $m_A(t) = (t - 1)^r$ . A base ordenada  $\beta = \vec{\cup} \beta_i$ ,  $1 \leq i \leq n$ , em que

$$\beta_i = \{v_i, (A - \lambda Id)(v_i), \dots, (A - \lambda Id)^{k_i}(v_i)\} \subset \mathcal{C}(v_i)$$

é chamada de *base de Jordan*. Um *bloco de Jordan* de comprimento  $r$  e autovalor  $\lambda \in \mathbb{K}$  é uma matriz denotada por  $J_r(\lambda) = [\eta_{ij}] \in M(r, \mathbb{K})$  n qual as entradas são definidas por

$$\eta_{ij} = \begin{cases} \lambda & \text{se } i - j = 0 \\ 1 & \text{se } i - j = 1 \\ 0 & \text{se } i - j \notin \{0, 1\} \end{cases}.$$

Graficamente, temos a forma descrita ao lado. Note que um bloco de Jordan com comprimento  $r$  e autovalor  $\lambda = 0$  é uma matriz nilpotente com nilpotência  $r$ . Identificando-se o espaço das matrizes  $1 \times 1$  com o corpo das entradas, justifica-se a notação  $J_1(\lambda) = \lambda$ .

$$J_r(\lambda) = \begin{bmatrix} \lambda & & & & & \\ 1 & \lambda & & & & \\ & 1 & \lambda & & & \\ & & \ddots & \ddots & & \\ & & & 1 & \lambda & \\ & & & & 1 & \lambda \end{bmatrix}.$$

A representação do operador  $A$  na base de Jordan será chamada de *representação de Jordan*. Como consequência do Teorema da decomposição cíclica II temos as representações de Jordan para os operadores  $B = A - \lambda Id$  e  $A$ , quando  $m_A(t) = (t - 1)^r$  descritas por

$$\begin{cases} [B]_\beta = \text{diag}\{J_{r_1}(0), J_{r_2}(0), \dots, J_{r_n}(0)\} \\ \text{grau } t^r = r_1 \geq r_2 \geq \cdots \geq r_n \end{cases},$$

e

$$\begin{cases} [A]_{\beta} = \text{diag}\{J_{r_1}(\lambda), J_{r_2}(\lambda), \dots, J_{r_n}(\lambda)\} \\ \text{grau}(t - \lambda)^r = r_1 \geq r_2 \geq \dots \geq r_n \end{cases} .$$

Em outras palavras, a representação matricial do operador induzido por restrição à parcela  $\mathcal{C}(v_i)$  dá origem a um bloco de Jordan  $J_{r_i}(\lambda)$  com comprimento  $r_i = \#\beta_i$ . Note que pelo corolário do mesmo teorema, o número de blocos de Jordan e a dimensão de  $\text{Nuc } B$  são iguais.

### Exercícios propostos 6.5.1

1. Determine uma base ordenada  $\beta \subset \mathbb{R}^2$  para a qual a representação do operador  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $A(x, y) = (y, 0)$ , é um bloco de Jordan.
2. Mostre que um bloco de Jordan  $J_r(0)$ , não admite uma raiz quadrada, isto é, não existe uma matriz  $N$  tal que  $N^2 = J_r(0)$ .
3. Prove que para todo inteiro  $n > 0$  existe uma matriz  $P_n \in M(3, \mathbb{K})$  tal que  $P_n^n = I + J_3(0)$ .

## 6.6 Exemplos

Ilustraremos com alguns exemplos as decomposições cíclicas para ilustrar, bem como a representação de Jordan obtida da decomposição.

**Exemplo 6.6.1** Vamos assumir que o operador linear  $A : \mathbb{R}^4 \rightarrow \mathbb{R}^4$  seja representado na base canônica pela matriz ao lado.

O polinômio característico é o polinômio  $p_A(t) = (t - 2)^4$ . Um simples cálculo matricial mostra que  $([A]_{\alpha} - 2I)^2 = [0]$ , significando que  $m_A(t) = (t - 2)^2$  é o polinômio minimal de  $A$ . Sigamos o roteiro dado na demonstração do Teorema da

$$[A]_{\alpha} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ -1 & 1 & 0 & -1 \\ -1 & 0 & 2 & 0 \\ 1 & 1 & 0 & 3 \end{bmatrix} .$$

decomposição cíclica II para construirmos uma decomposição  $A$ -cíclica de  $\mathbb{R}^4$ . Primeiro, consideramos o operador nilpotente com nilpotência dois,  $B : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ ,  $B(v) = (A - 2Id)v$  e sua sequência de núcleos

$$\mathbb{R}^4 = \text{Nuc } B^2 \supsetneq \text{Nuc } B \supsetneq \{0\} .$$

Construímos uma decomposição  $\mathbb{R}^4 = W_0 \oplus W_1$  onde a última parcela é  $W_1 = \text{Nuc } B$  enquanto  $W_0 \subset \text{Nuc } B^2$ . Calculemos uma base para o núcleo  $\text{Nuc } B$  resol-

vendo o sistema linear  $[B]_{\alpha} [v]_{\alpha} = [0]$ . Vejamos, do sistema

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

obtemos que o núcleo é gerado pelos vetores  $v_3 = (0, 0, 1, 0)$  e  $v_4 = (0, -1, 0, 1)$ . Escolhidos dois vetores linearmente independentes, por exemplo,  $v_1 = (1, 0, 0, 0)$  e  $v_2 = (0, 0, 0, 1)$  para geradores de um espaço  $W_0$ , obtemos a decomposição  $B$ -cíclica (e  $A$ -cíclica)

$$\begin{cases} \mathbb{R}^4 = \mathcal{C}(v_1) \oplus \mathcal{C}(v_2), \\ \text{grau}(t-2)^2 = \dim \mathcal{C}(v_1) = \dim \mathcal{C}(v_2). \end{cases}$$

As representações de  $B$  e  $A$  na base de Jordan  $\beta = \{v_1, B(v_1)\} \cup \{v_2, B(v_2)\}$  são as matrizes

$$[B]_{\beta} = \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & & 0 & \\ & & 1 & 0 \end{bmatrix} \quad \text{e} \quad [A]_{\beta} = \begin{bmatrix} 2 & & & \\ 1 & 2 & & \\ & & 2 & \\ & & 1 & 2 \end{bmatrix}.$$

Ressaltamos que as representações são as mesmas independentes dos vetores  $v_1$  e  $v_2$  escolhidos para gerar um subespaço complementar ao núcleo de  $B$ .  $\square$

**Exemplo 6.6.2** Dado o operador  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (-y, 2x+3y, -x+2z)$ , é claro que a sua representação na base canônica é a matriz ao lado. O polinômio característico de  $A$  tem uma fatoração primária que é um produto de fatores lineares,  $p_A(t) = (t-1)(t-2)^2$  e o polinômio minimal é igual ao polinômio característico. Logo, a decomposição primária determinada pelo operador é da forma

$$\mathbb{R}^3 = V_1 \oplus \text{Nuc}(A - 2Id)^2$$

em que a primeira parcela  $V_1$  é o autoespaço associado ao autovalor  $\lambda = 1$ . Obstruções dimensionais implicam que para cada parcela só existe uma única decomposição cíclica possível, a menos de isomorfismo, a saber,

$$\begin{cases} V_1 = \mathcal{C}(v_1) \\ \text{grau}(t-1) = \dim \mathcal{C}(v_1) \end{cases} \quad \text{e} \quad \begin{cases} \text{Nuc}(A - 2Id)^2 = \mathcal{C}(v_2) \\ \text{grau}(t-2)^2 = \dim \mathcal{C}(v_2) \end{cases}.$$

Sendo assim, na base de Jordan  $\beta = \{v_1\} \cup \{v_2, (A - 2Id)(v_2)\}$  obtemos a representação de Jordan para o operador. Determinemos explicitamente uma base para a qual a representação de  $A$  é a representação de Jordan acima. O autovetor associado ao autovalor  $\lambda = 1$  é determinado pelas soluções da equação  $A(x, y, z) = 1(x, y, z)$ , em termos matriciais, temos

$$[A]_{\beta} = \begin{bmatrix} 1 & & \\ & 2 & 0 \\ & 1 & 2 \end{bmatrix}.$$

$$\begin{bmatrix} 1 & 1 & 0 \\ -2 & -2 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Resolvendo o sistema, escolhemos  $v_1 = (1, -1, 1)$  para autovetor associado a  $\lambda = 1$ . Passemos ao cálculo do vetor  $v_2$ . Pela construção feita no Teorema da decomposição cíclica II devemos escolher um vetor em  $Nuc(A - 2Id)^2$  que não esteja em  $Nuc(A - 2Id)$ . Resolvendo o sistema  $[A - 2Id]_{\alpha}[v]_{\alpha} = [0]$ , isto é,

$$\begin{bmatrix} 2 & 1 & 0 \\ -2 & -1 & 0 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

concluimos que  $Nuc(A - 2Id)^2$  é gerado pelos vetores  $u_1 = (0, 0, 1)$  e  $u_2 = (1, -2, 0)$ . Como o primeiro deles está em  $Nuc(A - 2Id)$  devemos escolher o outro vetor  $u_2$  para ser o gerador da parcela  $\mathcal{C}(v_2)$  da decomposição cíclica.  $\square$

### Exercícios propostos 6.6.1

1. Descreva a representação de Jordan do operador conhecendo-se os polinômios característico e mínimo. Pode ocorrer que existam várias possibilidades. Sendo assim, só podemos determinar exatamente qual delas representa o operador se conhecermos explicitamente o operador linear.

a)  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $p_A(t) = m_A(t) = (t - 1)(t - 2)^2$ .

b)  $A : \mathbb{R}^5 \rightarrow \mathbb{R}^5$ ,  $p_A(t) = (t - 1)^3(t - 2)^2$  e  $m_A(t) = (t - 1)(t - 2)^2$ .

c)  $A : \mathbb{R}^6 \rightarrow \mathbb{R}^6$ ,  $p_A(t) = (t + 3)^4(t + 2)^2$  e  $m_A(t) = (t + 3)^2(t + 2)^2$ .

d)  $A : \mathbb{R}^7 \rightarrow \mathbb{R}^7$ ,  $p_A(t) = t^4(t + 3)^3$  e  $m_A(t) = t^2(t + 3)^2$ .

2. Encontre a representação e a base de Jordan para os seguintes operadores lineares.

a)  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (x + y, y, -2x - 2y + 2z)$ .

b)  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (x, x + 2y - z, 2x + 4y - 2z)$ .

c)  $A : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ ,  $A(x, y, z, w) = (0, x + w, x + w, -y + z)$ .

d)  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (7x + y + 4z, -x - 7y - 4z, -6x + 6y)$ .

3. Seja  $A$  um operador linear no espaço vetorial de dimensão finita  $V$  com polinômio minimal  $m_A(t) = (t - \lambda)^r$ . Treine sua redação em Matemática mostrando que duas representações de Jordan de  $A$  são sempre iguais.
4. Uma matriz de Jordan é uma matriz do tipo

$$\begin{cases} J = \text{diag}\{J_{r_1}(\lambda), J_{r_2}(\lambda), \dots, J_{r_n}(\lambda)\} \\ r_1 \geq r_2 \geq \dots \geq r_n \end{cases}.$$

- a) Mostre que as seguintes matrizes são conjugadas a uma matriz de Jordan.

$$i) N = \begin{bmatrix} 1 & 1 & 1 \\ 4 & 10 & 13 \\ -3 & -7 & -9 \end{bmatrix}. \quad ii) P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -16 & 32 & 24 & 8 \end{bmatrix}.$$

- b) Mostre que toda matriz  $3 \times 3$  não nula e nilpotente é conjugada a uma das duas matrizes,

$$J = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{ou} \quad J^2.$$

## 6.7 Aplicação: decomposição $D + N$

Já vimos que um operador linear num espaço vetorial de dimensão finita é diagonalizável quando o seu polinômio minimal é um produto de fatores lineares sem repetições. De certa forma esse resultado admite uma generalização. Mostraremos que se o polinômio mínimo do operador é um produto de fatores lineares com repetições então ele é uma soma de um operador diagonalizável com um nilpotente que comutam entre si, resultado bastante utilizado nas aplicações de Álgebra Linear. Para isso necessitaremos do

**Lema 6.7.1** *Sejam  $D_1$  e  $D_2$  dois operadores diagonalizáveis num espaço vetorial de dimensão finita  $V$ . Então existe uma base que diagonaliza simultaneamente os dois operadores se, e somente se, os operadores comutam.*

**Demonstração**  $\Rightarrow$ ) Antes de tudo, recordamos que quaisquer duas matrizes diagonais comutam, portanto, fixado uma base ordenada  $\beta$  de  $V$  que diagonaliza simultaneamente os operadores, pelas propriedades de representação matricial de uma composta temos as implicações,

$$[D_1]_\beta [D_2]_\beta = [D_2]_\beta [D_1]_\beta \Leftrightarrow [D_1 \circ D_2]_\beta = [D_2 \circ D_1]_\beta.$$

Pelo Teorema da representação matricial concluímos que  $D_1 \circ D_2 = D_2 \circ D_1$ .

$\Leftarrow$ ) Consideremos a fatoração primária do polinômio mínimo de  $D_1$ ,

$$m_{D_1}(t) = (t - \lambda_1)(t - \lambda_2) \dots (t - \lambda_k), \quad \lambda_i \neq \lambda_j \text{ se } i \neq j.$$

Como sabemos, as parcelas da decomposição primária de  $V$  determinada por  $D_1$  são os autoespaços,  $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$ . A hipótese de comutatividade implica que o operador  $D_2$  preserva cada um dos autoespaços. De fato, se  $v \in V_{\lambda_i}$ , então

$$D_1(D_2(v)) = D_2(D_1(v)) = D_2(\lambda_i v) = \lambda_i D_2(v).$$

Isso significa que  $D_2(v)$  é um autovetor associado ao mesmo autovalor  $\lambda_i$ . Do mesmo modo, a fatoração primária do polinômio minimal de  $D_2$  fatora-se em um produto de polinômios lineares sem repetições,

$$m_{D_2}(t) = (t - \rho_1)(t - \rho_2) \cdots (t - \rho_k), \quad \rho_i \neq \rho_j \text{ se } i \neq j,$$

e  $D_1$  preserva cada parcela da decomposição primária determinada por  $D_2$ ,  $V = V_{\rho_1} \oplus V_{\rho_2} \oplus \cdots \oplus V_{\rho_l}$ . Então, por um dos corolários do Teorema da decomposição primária vale a decomposição

$$V_{\rho_j} = (V_{\rho_j} \cap V_{\lambda_1}) \oplus (V_{\rho_j} \cap V_{\lambda_2}) \oplus \cdots \oplus (V_{\rho_j} \cap V_{\lambda_k})$$

para todo  $j$ . Logo, escolhendo uma base ordenada  $\beta_{ji}$  para  $V_{\rho_j} \cap V_{\lambda_i}$ , a união ordenada  $\beta = \vec{\cup} \left\{ \vec{\cup} \beta_{ji} \right\}$  é uma base ordenada de  $V$  formada por autovetores de  $D_1$  e  $D_2$ . Portanto, os operadores são simultaneamente diagonalizáveis.  $\square$

Operadores diagonalizáveis e operadores nilpotentes descrevem todos os operadores num espaço vetorial complexo e muitos outros em espaços vetoriais reais.

**Teorema 6.7.1 (Teorema da decomposição  $D + N$ )** *Seja  $A : V \rightarrow V$  um operador linear num espaço vetorial de dimensão finita sobre o corpo  $\mathbb{K}$ . Se  $m_A(t) = (t - \lambda_1)^{r_1}(t - \lambda_2)^{r_2} \cdots (t - \lambda_k)^{r_k}$  é a fatoração primária do polinômio minimal então existem um único operador nilpotente  $B$  e um único diagonalizável  $D$  satisfazendo as condições  $A = B + D$  e  $B \circ D = D \circ B$ .*

**Demonstração** Vamos assumir por um momento que  $m_A(t) = (t - \lambda)^r$ . Defina

$$D = \lambda Id \quad \text{e} \quad B = A - \lambda Id.$$

É claro que  $A = B + D$  e que  $B \circ D = D \circ B$ . Também é claro que  $D$  é diagonalizável e que  $B$  é nilpotente com nilpotência  $r$ . Verifiquemos a unicidade. Suponha que  $B_1$  é nilpotente e que  $D_1$  é diagonalizável satisfazendo as condições  $A = B_1 + D_1$  e  $B_1 \circ D_1 = D_1 \circ B_1$ . É imediato concluirmos que

$$B_1 \circ A = A \circ B_1 \quad \text{e} \quad D_1 \circ A = A \circ D_1.$$

Logo, esses dois operadores comutam com  $B$  e  $D$  pois os dois últimos são polinômios em  $A$ . Recordamos que a nilpotência de qualquer operador nilpotente é menor que  $n = \dim V$ . Levando-se em conta que os operadores comutam calculemos pelo

binômio de Newton

$$(D - D_1)^{2n} = (B_1 - B)^{2n} = \sum_{j=0}^{2n} \frac{(-1)^j 2n!}{j! (2n-j)!} B_1^j \circ B^{2n-j} = 0,$$

pois ou  $j$  ou  $2n - j$  são maiores que  $n$ . Como o único operador diagonalizável e nilpotente é o operador identicamente nulo, segue que  $D_1 = D$  e por conseguinte  $B_1 = B$ .

No caso do polinômio minimal ter mais de um fator consideramos as projeções sobre cada parcela da decomposição primária ao longo das outras,  $\pi_i : V \rightarrow V$ , e definimos

$$D = \lambda_1 \pi_1 + \lambda_2 \pi_2 + \cdots + \lambda_k \pi_k,$$

$$B = (A - \lambda_1 Id) \circ \pi_1 + (A - \lambda_2 Id) \circ \pi_2 + \cdots + (A - \lambda_k Id) \circ \pi_k.$$

A demonstração segue análoga desde que polinômios em  $A$  comutam com qualquer operador que comuta com  $A$ .  $\square$

**Exemplo 6.7.1** O polinômio  $p_A(t) = (t - 1)^2$  é o polinômio característico do operador linear  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $A(x, y) = (3x - y, 4x - y)$  e seu polinômio minimal é  $m_A(t) = p_A(t)$ . Logo, pela construção feita na proposição acima devemos definir os operadores  $D = Id$  e  $B = A - Id$ . De fato, verificamos que  $B$  é nilpotente com nilpotência dois pois,

$$[A - Id]_{\alpha} = \begin{bmatrix} 2 & -1 \\ 4 & -2 \end{bmatrix}. \quad \square$$

### Exercícios propostos 6.7.1

1. Determine a decomposição  $D + N$  dos operadores.

a)  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (x, x + y, x + y + 2z)$ .

b)  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $A(x, y) = (5x + 9y, -x - y)$ .

c)  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (8x - y - 5z, -2x + 3y + z, 4x - y - z)$ .

2. Prove que se  $B$  e  $C$  são dois operadores lineares num espaço vetorial  $V$  de dimensão finita que comutam e  $B$  é nilpotente, então os operadores  $C$  e  $A \equiv B + C$  têm os mesmos autovalores.

3. A representação matricial do operador  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (x, x+y, x+y+2z)$  na base canônica é da forma

$$[A]_{\alpha} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}}_{[D]_{\alpha}} + \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix}}_{[B]_{\alpha}}$$

O operador diagonal  $D$  e o operador nilpotente  $B$  definidos pelas representações indicadas correspondem à decomposição  $D + N$  para o operador  $A$ ! Essa afirmação é falsa ou verdadeira?

4. Mostre que toda matriz quadrada com entradas complexas é a soma de uma matriz diagonal com uma matriz nilpotente que comutam. Esse fato é verdadeiro para matrizes quadradas com entradas reais?

## Capítulo 7

# Espaços Euclidianos

Um produto interno num espaço vetorial real é uma aplicação pela qual podemos realizar as noções de comprimento e ângulo da Geometria Euclidiana. Em espaços equipados com um produto interno existem operadores chamados de operadores normais que são as origens de vários grupos matriciais clássicos. Os dois próximos capítulos estão direcionados para o estudo de operadores normais em um espaço vetorial real de dimensão finita equipado com um produto interno. Nesse capítulo recapitularemos a terminologia e os fatos básicos necessários para estudá-los. A maior parte dos tópicos aqui estudados é normalmente apresentado nos cursos introdutórios de Álgebra Linear.

### 7.1 Produto interno

Um *produto interno* num espaço vetorial real  $V$  é uma aplicação  $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$  possuindo três propriedades, a saber. Para quaisquer  $u, v, w \in V$  e  $\lambda \in \mathbb{R}$ , temos que

1.  $\langle u, v \rangle = \langle v, u \rangle;$  (simétrica)
2.  $\langle u + \lambda v, w \rangle = \langle u, w \rangle + \lambda \langle v, w \rangle;$  (bilinear)
3.  $\langle v, v \rangle > 0 \Leftrightarrow v \neq 0.$  (positiva definida)

Por simplicidade, indicamos um produto interno em  $V$  pelo símbolo  $\langle \cdot, \cdot \rangle$  omitindo o domínio e o contra domínio da aplicação. Para outras propriedades elementares indicamos o primeiro exercício proposto dessa seção.

**Exemplo 7.1.1** 1) Chamaremos de *produto interno canônico* do  $\mathbb{R}^n$  a aplicação que a cada par de vetores  $u = (x_1, x_2, \dots, x_n), v = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$  associa o número real

$$\langle u, v \rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

A menos que seja dito explicitamente o contrário a partir desse momento o espaço  $\mathbb{R}^n$  estará equipado com o produto interno canônico.

2) Qualquer espaço vetorial real  $V$  de dimensão finita admite um produto interno. Para construí-lo, fixamos uma base ordenada  $\beta = \{v_1, v_2, \dots, v_n\} \subset V$  e definimos

$$\langle u, v \rangle_\beta = x_1y_1 + x_2y_2 + \cdots + x_ny_n,$$

em que  $u = x_1v_1 + x_2v_2 + \cdots + x_nv_n$  e  $v = y_1v_1 + y_2v_2 + \cdots + y_nv_n$ . Essa construção acrescenta um pouco mais de informação: todo espaço vetorial de dimensão finita admite vários produtos internos. Posteriormente, essa questão será estudada comparando dois produtos internos no mesmo espaço.

3) Um produto interno num espaço vetorial  $V$  induz, por restrição, um produto interno em qualquer subespaço  $W \subset V$ .

4) Podemos definir um produto interno no espaço vetorial das funções contínuas do intervalo  $[0, 1]$  com valores na reta,  $C^0([0, 1], \mathbb{R})$ , do seguinte modo,

$$\langle f, g \rangle = \int_0^1 f(t)g(t) dt.$$

A integral considerada é a integral de Riemann. □

### Exercícios propostos 7.1.1

1. Deduza diretamente da definição de produto interno num espaço vetorial real  $V$  que para quaisquer  $u, v \in V$  e  $\lambda \in \mathbb{R}$  valem as propriedades
  - a)  $\langle u, w + \lambda v \rangle = \langle u, w \rangle + \lambda \langle u, v \rangle$ ;
  - b)  $\langle v, v \rangle = 0 \Leftrightarrow v = 0$ .
2. Verifique que  $\langle (x_1, x_2), (y_1, y_2) \rangle = 2x_1y_1 + 3x_2y_2$  é um produto interno em  $\mathbb{R}^2$ .
3. Assuma que  $A$  é um isomorfismo linear num espaço vetorial real  $V$  de dimensão finita equipado com um produto interno  $\langle \cdot, \cdot \rangle$ . Verifique que  $\langle \langle v, w \rangle \rangle = \langle A(v), A(w) \rangle$  é outro produto interno no mesmo espaço.
4. Mostre que a aplicação  $\langle N, P \rangle = \text{tr}(NP^t)$  é um produto interno no espaço das matrizes  $M(m \times n, \mathbb{R})$ .

## 7.2 Norma

Uma *norma* num espaço vetorial real  $V$  é uma aplicação com valores não negativos,  $\| \cdot \| : V \rightarrow [0, \infty)$ , possuindo as seguintes propriedades para quaisquer  $u, v \in V$  e  $\lambda \in \mathbb{R}$ ,

1.  $\|v\| = 0 \Leftrightarrow v = 0$ ;
2.  $\|\lambda v\| = |\lambda| \|v\|$ ;
3.  $\|u + v\| \leq \|u\| + \|v\|$ . (primeira desigualdade triangular)

A notação  $|\lambda|$  indica o *valor absoluto* do escalar real. Se  $\mathbb{R}$  é considerado como um espaço vetorial sobre si mesmo, o valor absoluto é de fato uma norma. O valor  $\|v\|$  é interpretado, geometricamente, como o comprimento do vetor  $v \in V$ , por isso diremos que um vetor  $u \in V$  é *unitário* quando  $\|u\| = 1$ . Um produto interno  $\langle \cdot, \cdot \rangle$  num espaço vetorial real  $V$  sempre induz uma norma em  $V$ , para isso, basta definir

$$\|\cdot\| : V \rightarrow [0, \infty), \quad \|v\| = \sqrt{\langle v, v \rangle}.$$

Essa será sempre a norma considerada num espaço com produto interno. Entretanto para verificar que, de fato, a aplicação assim definida é uma norma necessitamos da *Desigualdade de Schwarz*.

**Teorema 7.2.1 (Desigualdade de Schwarz)** *Seja  $V$  um espaço vetorial real com um produto interno  $\langle \cdot, \cdot \rangle$ . Então a aplicação  $\|\cdot\| : V \rightarrow [0, \infty)$ ,  $\|v\| = \sqrt{\langle v, v \rangle}$  satisfaz a desigualdade*

$$|\langle u, v \rangle| \leq \|u\| \|v\|.$$

para quaisquer  $u, v \in V$ . Em consequência, a aplicação  $\|\cdot\|$  é uma norma.

**Demonstração** Por definição de produto interno, para qualquer escalar  $t \in \mathbb{R}$  temos a desigualdade,

$$0 \leq \langle tu - v, tu - v \rangle = \|u\|^2 t^2 - 2\langle u, v \rangle t + \|v\|^2.$$

Logo, o discriminante  $\Delta$  do polinômio em  $t$  é não positivo,

$$\Delta = 4\langle u, v \rangle^2 - 4\|u\|^2 \|v\|^2 \leq 0.$$

Daí segue a desigualdade de Schwarz. Para demonstrar que a aplicação é uma norma devemos verificar os três itens da definição. Entretanto faremos apenas a demonstração da desigualdade triangular que é uma consequência da desigualdade de Schwarz, deixando os outros itens como exercícios. Pela desigualdade mostrada temos

$$\langle u + v, u + v \rangle = \|u\|^2 + \|v\|^2 + 2\langle u, v \rangle \leq \|u\|^2 + \|v\|^2 + 2\|u\| \|v\|.$$

Portanto,  $\|u + v\|^2 \leq (\|u\| + \|v\|)^2$ , implicando que  $\|u + v\| \leq \|u\| + \|v\|$ . □

**Exercício 7.2.1** Seja  $\|\cdot\|$  uma norma definida pelo produto interno  $\langle \cdot, \cdot \rangle$  num espaço vetorial real  $V$ .

1. Mostre a *identidade de polarização*,

$$\langle u, v \rangle = \frac{1}{4} \|u + v\|^2 - \frac{1}{4} \|u - v\|^2.$$

Sugestão: desenvolva o segundo membro da identidade.

2. Prove e guarde as caracterizações pois elas serão utilizadas inúmeras vezes.

a)  $\langle u, v \rangle = 0$  para todo  $u \in V \Leftrightarrow v = 0$ ;

b)  $\langle u, v \rangle = \langle u, w \rangle$  para todo  $u \in V \Leftrightarrow v = w$ . □

As normas aqui consideradas sempre serão definidas por um produto interno no espaço vetorial, embora muitas vezes esse fato não esteja explicitado. Para deixar claro que nem toda norma em  $V$  é dessa forma, remetemos o leitor para o primeiro exercício proposto dessa seção, ali está dado um critério para saber quando uma norma é induzida ou não por um produto interno.

### Exercícios propostos 7.2.1

1. Demonstre que uma norma  $\| \cdot \|$  num espaço vetorial real  $V$  induzida por um produto interno  $\langle \cdot, \cdot \rangle$  em  $V$  satisfaz a *lei do paralelogramo*, isto é,

$$\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2).$$

2. Se  $v = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ , defina a aplicação

$$\| \cdot \| : \mathbb{R}^n \rightarrow [0, \infty), \quad \|v\| = \max\{|x_i|, 1 \leq i \leq n\}.$$

Prove que  $\| \cdot \|$  é uma norma mas que não é induzida por um produto interno.

3. Prove que para quaisquer duas funções contínuas  $f, g : [0, 1] \rightarrow \mathbb{R}$  vale a desigualdade

$$\left| \int_0^1 f(t)g(t) dt \right| \leq \left| \int_0^1 f(t)^2 dt \right|^{1/2} \left| \int_0^1 g(t)^2 dt \right|^{1/2}.$$

4. Dados dois vetores não nulos e distintos,  $u_0$  e  $u_1$ , num espaço vetorial real  $V$  equipado com um produto interno, mostre que se um vetor  $v \in V$  satisfaz a igualdade

$$\|v - u_0\| + \|v - u_1\| = \|u_0 - u_1\|,$$

então existem  $\lambda_0, \lambda_1 \in [0, 1]$  tais que  $v = \lambda_0 u_0 + \lambda_1 u_1$  e  $\lambda_0 + \lambda_1 = 1$ .

5. Prove que para qualquer função contínua positiva  $f : \mathbb{R} \rightarrow \mathbb{R}$  vale a desigualdade

$$\left( \int_0^1 f(t) dt \right) \left( \int_0^1 \frac{1}{f(t)} dt \right) \geq 1.$$

6. Seja  $\| \cdot \|$  uma norma em  $V$ . Prove que para quaisquer  $u, v, w \in V$  vale a desigualdade

$$\|u - v\| \|w\| \leq \|u - w\| \|v\| + \|w - v\| \|u\|.$$

7. Demonstre que uma norma  $\| \cdot \|$  num espaço vetorial real de dimensão finita  $V$  satisfazendo a Lei do paralelogramo é induzida por um produto interno (você irá precisar do conceito de continuidade).

### 7.3 Ortogonalidade

A desigualdade de Schwarz permite definir o ângulo entre dois vetores não nulos  $u$  e  $v$  num espaço vetorial real  $V$  equipado com um produto interno  $\langle \cdot, \cdot \rangle$ . Como

$$-1 \leq \frac{\langle u, v \rangle}{\|u\| \|v\|} \leq 1,$$

definimos que o *ângulo entre os dois vetores* não nulos é o único  $\theta \in [0, \pi]$  tal que

$$\langle u, v \rangle = \|u\| \|v\| \cos \theta.$$

Essa definição estabelece a relação fundamental entre os conceitos de produto interno, norma (comprimento de vetor) e ângulo entre dois vetores não nulos que será explorada ao longo dos próximos capítulos. Por exemplo, diremos que dois vetores são *ortogonais* quando  $\langle u, v \rangle = 0$ . Para vetores não nulos serem ortogonais significa que o ângulo entre eles é  $\theta = \frac{\pi}{2}$ .

**Exercício 7.3.1** Se  $W$  é um subespaço do espaço vetorial real  $V$  equipado com um produto interno  $\langle \cdot, \cdot \rangle$ , demonstre que o complemento ortogonal  $W^\perp$  é também um subespaço e que  $W \cap W^\perp = \{o\}$ . Por definição, o complemento ortogonal de  $W$  é o conjunto

$$W^\perp = \{v \in V; \langle v, w \rangle = 0 \text{ para todo } w \in W\}. \quad \square$$

#### Exercícios propostos 7.3.1

1. Considere o produto interno canônico em  $\mathbb{R}^2$ .
  - a) Mostre que para cada vetor unitário  $u \in \mathbb{R}^2$  existe um único  $\theta \in [0, 2\pi)$  tal que  $u = (\cos \theta, \sin \theta)$ .
  - b) Dados os vetores unitários  $u = (\cos \theta, \sin \theta)$  e  $v = (\cos \alpha, \sin \alpha)$ , verifique a identidade,  $\langle u, v \rangle = \cos(\theta - \alpha)$ .
  - c) Mostre que para o ângulo  $\theta$  entre dois vetores não nulos  $u = (x_1, y_1)$  e  $v = (x_2, y_2)$  vale
 
$$\cos \theta = \frac{x_1 x_2 + y_1 y_2}{\|u\| \|v\|} \quad \text{e} \quad \sin \theta = \frac{x_1 y_2 - x_2 y_1}{\|u\| \|v\|}.$$
2. Verifique que a base canônica de  $M(m \times n, \mathbb{R})$  é uma base ortonormal em relação ao produto interno  $\langle N, Q \rangle = \text{tr}(NQ^t)$ .
3. Seja  $\| \cdot \|$  a norma definida por um produto interno num espaço vetorial real  $V$ .
  - a) Prove o *Teorema de Pitágoras*:  $u$  e  $v$  são ortogonais  $\Leftrightarrow \|u\|^2 + \|v\|^2 = \|u+v\|^2$ .
  - b) Quais as condições sobre os vetores  $u$  e  $v$  para que  $|\langle u, v \rangle| = \|u\| \|v\|$ ?
  - c) Demonstre a *segunda desigualdade triangular*,  $|\|u\| - \|v\|| \leq \|u - v\|$ .

4. Sejam  $u$  e  $v$  vetores não nulos num espaço com produto interno  $V$ . Demonstre:
- os vetores  $\|v\|u + \|u\|v$  e  $\|v\|u - \|u\|v$  são ortogonais;
  - o ângulo entre o vetor  $\|v\|u - \|u\|v$  e  $u$  e o ângulo entre o vetor  $\|v\|u + \|u\|v$  e  $v$  são iguais.

## 7.4 Espaços Euclidianos

Um *espaço Euclidiano* é um par  $(V, \langle \cdot, \cdot \rangle)$  no qual  $V$  é um espaço vetorial real de dimensão finita e  $\langle \cdot, \cdot \rangle$  é um produto interno em  $V$ .

Com a finalidade de ser menos repetitivo, reservamos a notação  $(V, \langle \cdot, \cdot \rangle)$  para designar única e exclusivamente um espaço Euclidiano. Avisamos que em alguns textos não é exigido a condição de dimensão finita na definição, fato que não ocorrerá aqui. Em tais espaços existem bases especiais que destacaremos com uma definição. Uma base ordenada  $\beta = \{v_1, v_2, \dots, v_n\} \subset V$  é uma *base ortogonal* de  $(V, \langle \cdot, \cdot \rangle)$  quando  $\langle v_i, v_j \rangle = 0$  se  $i \neq j$ . Diremos que a base ordenada  $\beta$  é *ortonormal* quando  $\langle v_i, v_j \rangle = \delta_{ij}$  (delta de Kronecker).

**Exemplo 7.4.1** 1) A base canônica do  $\mathbb{R}^n$  é uma base ortonormal com respeito ao produto interno canônico.

2) Como vimos qualquer espaço vetorial real  $V$  de dimensão finita admite uma estrutura de espaço Euclidiano. Para construí-la, fixamos uma base ordenada  $\beta = \{v_1, v_2, \dots, v_n\} \subset V$  e consideramos

$$\langle u, v \rangle_\beta = x_1y_1 + x_2y_2 + \dots + x_ny_n,$$

em que  $u = x_1v_1 + x_2v_2 + \dots + x_nv_n$  e  $v = y_1v_1 + y_2v_2 + \dots + y_nv_n$ . Nessa construção  $\beta$  torna-se uma base ortonormal.  $\square$

**Exercício 7.4.1** Suponha que  $\delta = \{v_1, v_2, \dots, v_k\}$  é um conjunto de vetores não nulo e dois a dois ortogonais no espaço Euclidiano  $V$ . Prove que  $\delta$  é um conjunto linearmente independente. Em particular, se  $\#\delta = \dim V$  então  $\delta$  é uma base ortogonal.  $\square$

Bases ortonormais em espaços Euclidianos existem e a demonstração desse fato é construtiva. O método de construção utilizado para mostrar a existência é chamado de *processo de ortogonalização de Gram-Schmidt* que é tão relevante quanto o resultado em si.

**Proposição 7.4.1** *Todo subespaço não trivial  $W$  de um espaço Euclidiano  $V$  possui uma base ortogonal.*

**Demonstração** Escolha  $\gamma = \{w_1, w_2, \dots, w_k\}$  uma base ordenada qualquer de  $W$ . Denote por  $W_i$  o subespaço de dimensão  $i$  gerado pelos  $i$ -ésimos primeiros vetores dessa base,  $\gamma_i = \{w_1, w_2, \dots, w_i\}$ . Sendo assim, valem as inclusões próprias de subespaços

$$W_0 = \{0\} \subsetneq W_1 \subsetneq W_2 \subsetneq \dots \subsetneq W_k = W.$$

Feitos essas preliminares iniciemos a construção indutiva de uma base ortogonal pelo processo de ortogonalização de Gram-Schmidt. A base ortogonal de  $W_1$  será  $\beta_1 = \{v_1\}$  em que  $v_1 = w_1$ . Para construir uma base ortogonal para  $W_2$  consideramos o conjunto ordenado  $\beta_2 = \beta_1 \vec{\cup} \{v_2\}$  onde

$$v_2 = w_2 - \frac{\langle w_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1.$$

O vetor  $v_2$  está bem definido pois  $v_1$  não sendo nulo temos que  $\langle v_1, v_1 \rangle > 0$ . Note que também o vetor  $v_2$  não é nulo, caso contrário concluimos que  $w_1$  e  $w_2$  são vetores linearmente dependentes contrariando o fato de  $\gamma$  ser uma base de  $W$ . Por outro lado verificamos facilmente que  $\langle v_1, v_2 \rangle = 0$  de onde segue que  $\beta_2 \subset W_2$  é um conjunto linearmente independente num espaço vetorial de dimensão dois, implicando que  $\beta_2 = \beta_1 \vec{\cup} \{v_2\}$  é uma base ortogonal de  $W_2$ . Por hipótese de indução, vamos assumir que já construímos uma base ortogonal  $\beta_i = \{v_1, v_2, \dots, v_i\}$  para o subespaço  $W_i$ . Seja

$$\beta_{i+1} = \beta_i \vec{\cup} \{v_{i+1}\},$$

em que

$$v_{i+1} = w_{i+1} - \frac{\langle w_{i+1}, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle w_{i+1}, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 - \dots - \frac{\langle w_{i+1}, v_i \rangle}{\langle v_i, v_i \rangle} v_i.$$

Novamente,  $v_{i+1}$  está bem definido e é um vetor em  $W_{i+1}$ . O vetor  $v_{i+1}$  não é nulo, caso contrário teremos  $w_{i+1} \in W_i$  contrariando a hipótese de  $\gamma$  ser linearmente independente, desde que cada  $v_i$  é combinação linear de  $\gamma_i$ . Uma simples verificação mostra que  $\beta_{i+1}$  é um conjunto de vetores não nulos dois a dois ortogonais no subespaço  $W_{i+1}$  cuja dimensão é  $i+1$ . Segue que  $\beta_{i+1}$  é uma base ortogonal desse espaço. Continuando o processo um número de vezes igual à  $\dim W$ , obtemos uma base ortogonal de  $W$ .  $\square$

**Corolário 7.4.1** *Todo subespaço não trivial de um espaço Euclidiano  $V$ , possui uma base ortonormal. Em particular,  $V$  possui uma base ortonormal.*

**Demonstração** Pelo processo de Gram-Schmidt podemos construir uma base ortogonal  $\gamma = \{v_1, v_2, \dots, v_k\}$  de  $W$ . O conjunto ordenado  $\beta = \{u_1, u_2, \dots, u_n\}$ ,

onde  $u_i = \frac{1}{\|v_i\|}v_i$ , é formado por vetores unitários dois a dois ortogonais, logo  $\beta$  é uma base ortonormal de  $W$ .  $\square$

Quando um espaço Euclidiano  $V$  é uma soma direta de subespaços mutuamente ortogonais, isto é,  $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$  e

$$\langle v_i, v_j \rangle = 0 \quad \text{quando} \quad v_i \in V_i, \quad v_j \in V_j \quad \text{para} \quad i \neq j,$$

chamamos a decomposição de *soma direta ortogonal* e esse fato será registrado graficamente por

$$V = V_1 \boxplus V_2 \boxplus \cdots \boxplus V_k.$$

O símbolo  $\boxplus$  é o mesmo símbolo usualmente utilizado na Geometria Euclidiana para indicar o ângulo reto entre duas retas.

**Exercício 7.4.2** Prove as afirmação sobre um subespaço  $W$  de  $(V, \langle \cdot, \cdot \rangle)$ .

1. Um vetor  $v \in V$  é ortogonal a todo elemento  $w \in W$  se, e somente se,  $v$  é ortogonal aos vetores de uma base de  $W$ .
2. Se  $W$  é um subespaço próprio então o complemento ortogonal  $W^\perp$  é não vazio.  $\square$

A existência de uma base ortogonal permite uma fácil demonstração para a

**Proposição 7.4.2** Se  $W$  é um subespaço do espaço Euclidiano  $V$  então

$$V = W \boxplus W^\perp.$$

**Demonstração** Já sabemos que  $W \cap W^\perp = \{0\}$ . Vamos supor por absurdo que a soma direta ortogonal  $U = W \boxplus W^\perp$  é um subespaço próprio de  $V$ . Valendo essa hipótese, seja  $v \in V$  um vetor não nulo ortogonal à  $U$ . Então  $v$  é ortogonal a todo vetor de  $W$ , logo  $v \in W^\perp$ . Mas  $v$  também é ortogonal à  $W^\perp$ , portanto  $\langle v, v \rangle = 0$ , implicando que  $v = 0$ , uma contradição. Isso mostra que  $V = W \boxplus W^\perp$ .  $\square$

**Exercício 7.4.3** Seja  $\beta = \{v_1, v_2, \dots, v_n\}$  uma base ortonormal de  $(V, \langle \cdot, \cdot \rangle)$ .

1. Mostre que os coeficientes da combinação linear  $v = x_1v_1 + x_2v_2 + \cdots + x_nv_n$  são os escalares  $x_i = \langle v, v_i \rangle$ .
2. Se  $w = y_1v_1 + y_2v_2 + \cdots + y_nv_n$  então  $\langle v, w \rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n$ .
3. Conclua que  $\|v\| = (x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2}$ .  $\square$

A utilização de bases ortonormais é uma ferramenta básica para o estudo de operadores em espaços Euclidianos. Ilustraremos a facilidade operacional quando trabalhamos com tais tipos de bases. Por exemplo, calculemos a representação  $[A]_\beta = [a_{ij}]$  de um operador  $A$  num espaço Euclidiano  $V$  quando a base  $\beta = \{v_1, v_2, \dots, v_n\}$  é ortonormal. Por definição de representação matricial sabemos que

$$A(v_j) = a_{1j}v_1 + a_{2j}v_2 + \dots + a_{nj}v_n.$$

Avaliando o produto interno em ambos membros da equação com o vetor  $v_i$  obtemos que (observe a ordem dos índices)

$$a_{ij} = \langle A(v_j), v_i \rangle.$$

### Exercícios propostos 7.4.1

1. Ortogonalize pelo processo de Gram-Schmidt as bases ordenadas.
  - a)  $\beta = \{(1, 2), (1, 3)\} \subset \mathbb{R}^2$ .
  - b)  $\beta = \{(1, 1, 1), (1, -1, 0), (0, 1, 2)\} \subset \mathbb{R}^3$ .
  - c)  $\beta = \{(1, 2, 1), (2, 0, -2), (-4, -4, 4)\} \subset \mathbb{R}^3$ .
2. Seja  $W$  o subespaço do  $\mathbb{R}^4$  gerado pelos vetores  $v_1 = (1, 1, 0, 0)$  e  $v_2 = (0, 1, 1, 1)$ .
  - a) Determine uma base para  $W^\perp$ .
  - b) Estenda a base obtida para uma base ortogonal do  $\mathbb{R}^4$ .
3. Mostre que qualquer conjunto  $\delta = \{v_1, v_2, \dots, v_k\}$  de vetores unitários e dois a dois ortogonais num espaço Euclidiano  $V$  de dimensão  $n$  pode ser estendido a uma base ortonormal  $\beta = \{v_1, v_2, \dots, v_n\}$ .
4. Fixado um vetor  $v_0$  não nulo de  $(V, \langle \cdot, \cdot \rangle)$  defina a aplicação
 
$$\pi : V \rightarrow V, \quad \pi(w) = w - \frac{\langle w, v_0 \rangle}{\langle v_0, v_0 \rangle} v_0.$$
  - a) Verifique que  $\pi$  é uma projeção e identifique o núcleo e a imagem.
  - b) Mostre que  $\pi$  é uma *projeção ortogonal*, isto é,  $V = \text{Im } \pi \boxplus \text{Nuc } \pi$ .
5. Ortogonalize a base ordenada  $\beta = \{1, t, t^2, t^3\}$  de  $\mathbb{R}_3[t]$  quando o produto interno considerado é

$$\langle p(t), q(t) \rangle = \int_0^1 p(t) q(t) dt.$$

Determine uma base para o subespaço ortogonal ao subespaço formado pelos polinômios de grau no máximo 1.

6. Se  $W_1$  e  $W_2$  são dois subespaços do espaço Euclidiano  $V$ , prove as identidades.
  - a)  $(W_1^\perp)^\perp = W_1$ .
  - b)  $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$ .

$$c) W_1^\perp \cap W_2^\perp = W_1^\perp + W_2^\perp.$$

7. Considere o conjunto ordenado  $\beta = \{v_1, v_2, \dots, v_n\}$  do espaço Euclidiano  $(V, \langle \cdot, \cdot \rangle)$ .

(a) Se  $\beta$  é uma base ortonormal, prove a *desigualdade de Bessel*,

$$\sum_{i=1}^n \langle w, v_i \rangle^2 \leq \|w\|^2.$$

(b) Reciprocamente, se vale a desigualdade de Bessel para todo  $w \in V$ , então o conjunto  $\beta$  é uma base.

8. Considere no espaço  $M(n, \mathbb{R})$  o produto interno  $\langle N, Q \rangle = \text{tr}(NQ^t)$ . Demonstre que  $M(n, \mathbb{R}) = \mathcal{S} \boxplus \mathcal{A}$  onde a primeira parcela é o subespaço das matrizes simétricas e a segunda é o subespaço das matrizes anti-simétricas.

9. Seja  $A$  um operador linear no espaço Euclidiano  $V$  de dimensão  $n$ . Mostre que se  $\delta = \{u_1, u_2, \dots, u_k\}$  é uma base ortonormal da imagem de  $A$ , então

$$A(v) = \langle A(v), u_1 \rangle u_1 + \langle A(v), u_2 \rangle u_2 + \dots + \langle A(v), u_k \rangle u_k.$$

10. Seja  $V$  um espaço Euclidiano e  $W \subset V$  um subespaço próprio. Demonstre:

a) dado um vetor  $v \in V$  existe um único vetor  $w_0 \in W$  tal que  $v - w_0$  é ortogonal à  $W$  e

b) para qualquer outro vetor  $w \in W$  com  $w \neq w_0$  vale a desigualdade  $\|v - w_0\| \leq \|v - w\|$ .

11. Determine a aplicação  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  que reflete ortogonalmente os vetores em relação ao subespaço  $W = \{(x, y, z) \in \mathbb{R}^3; x + y + z = 0\}$ .

12. Uma projeção  $\pi$  num espaço Euclidiano é *ortogonal* quando  $V = \text{Im } \pi \boxplus \text{Nuc } \pi$ . Prove que uma projeção  $\pi$  é ortogonal  $\Leftrightarrow \|\pi(v)\| \leq \|v\|$  para qualquer  $v \in V$ .

## 7.5 Representação de um funcional linear

Faremos outra aplicação da existência de uma base ortonormal cujo resultado, o Teorema da representação de um funcional linear, é fundamental para o restante do texto. O espaço dual  $V^*$  de um espaço Euclidiano  $V$  admite uma descrição muito simples quando utilizamos o produto interno. Fixado um vetor  $v_0 \in V$ , segue diretamente da definição de produto interno que a aplicação

$$f_0 : V \rightarrow \mathbb{R}, \quad f_0(w) = \langle w, v_0 \rangle,$$

é um funcional linear. Diremos que o vetor  $v_0 \in V$  *representa o funcional linear*  $f_0$ . A recíproca também é verdadeira, todo funcional linear num espaço Euclidiano é representado por um vetor, registremos este fato no

**Teorema 7.5.1 (Teorema da representação de um funcional linear)** *Sejam  $(V, \langle \cdot, \cdot \rangle)$  um espaço Euclidiano e  $\Lambda : V \rightarrow V^*$  a aplicação que a cada vetor  $v \in V$  associa ao funcional linear*

$$\Lambda[v] : V \rightarrow \mathbb{R}, \quad \Lambda[v](w) = \langle w, v \rangle.$$

*Então  $\Lambda : V \rightarrow V^*$  é um isomorfismo linear. Em particular, cada funcional linear é representado por um único vetor.*

**Demonstração** É claro que a aplicação  $\Lambda : V \rightarrow V^*$ , está bem definida, isto é,  $\Lambda[v] : V \rightarrow \mathbb{R}$  é um funcional linear. Para mostrar que  $\Lambda$  é um isomorfismo linear basta mostrar que  $\Lambda$  é linear e sobrejetiva pois  $\dim V = \dim V^*$ .

A linearidade: dados os vetores  $u, v \in V$  e o escalar  $\lambda \in \mathbb{R}$ , para qualquer  $w \in V$  temos que

$$\Lambda[u + \lambda v](w) = \langle u + \lambda v, w \rangle = \langle u, w \rangle + \lambda \langle v, w \rangle = (\Lambda[u] + \lambda \Lambda[v])(w).$$

Logo,  $\Lambda[u + \lambda v] = \Lambda[u] + \lambda \Lambda[v]$  mostrando a linearidade de  $\Lambda : V \rightarrow V^*$ .

A sobrejetividade: escolhida uma base ortonormal  $\beta = \{v_1, v_2, \dots, v_n\}$  de  $V$  sabemos que um vetor  $w \in V$  é escrito de maneira única como uma combinação linear dos elementos dessa base,  $w = x_1 v_1 + x_2 v_2 + \dots + x_n v_n$ . Fixado um funcional linear  $f : V \rightarrow \mathbb{R}$ , e avaliando esse funcional no vetor  $w$  obtemos que

$$f(w) = x_1 f(v_1) + x_2 f(v_2) + \dots + x_n f(v_n).$$

Desse modo, o vetor

$$v_f = f(v_1)v_1 + f(v_2)v_2 + \dots + f(v_n)v_n$$

representa o funcional linear. Com efeito. Sendo a base  $\beta$  ortonormal calculemos,

$$\Lambda[v_f](w) = \langle w, v_f \rangle = x_1 f(v_1) + x_2 f(v_2) + \dots + x_n f(v_n) = f(w).$$

Isso significa que  $\Lambda[v_f] = f$ , concluindo a demonstração.  $\square$

### Exercícios propostos 7.5.1

1. Considere o funcional linear  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $f(x, y) = 3x - 2y$ . Represente o funcional utilizando os seguintes produtos internos em  $\mathbb{R}^2$ .
  - a) Produto interno canônico.
  - b)  $\langle (x, y), (z, w) \rangle = 2xz + 3yw$ .
  - c)  $\langle (x, y), (z, w) \rangle = 2xz - xw - yz + 2yw$ .
2. Considere o seguinte produto interno no espaço vetorial dos polinômios  $\mathbb{R}_2[t] = \{p(t) \in \mathbb{R}[t]; \text{ tal que grau } p(t) \leq 2\}$ ,

$$\langle f, g \rangle = \int_0^1 f(t)g(t) dt,$$

onde a integral considerada é a integral de Riemann. Represente o funcional linear avaliação em  $t = 0$ , ou seja,  $\mathcal{A} : \mathbb{R}_2[t] \rightarrow \mathbb{R}$ ,  $\mathcal{A}(p(t)) = p(0)$ .

## 7.6 Operador transposto

Um operador linear  $A$  num espaço vetorial  $V$  sobre o corpo  $K$  induz um operador linear no espaço dual  $A^* : V^* \rightarrow V^*$ , chamado de adjunta de  $A$  e definido por

$$A^*[f] : V \rightarrow K, \quad A^*[f](v) = f(A(v)).$$

Nessa seção iremos estudar a adjunta de um operador num espaço Euclidiano. Com o auxílio do Teorema da representação de um funcional linear transferimos o estudo da adjunta para o estudo de um operador no espaço Euclidiano  $V$ .

Se  $A$  é um operador linear em  $(V, \langle \cdot, \cdot \rangle)$ , pelo Teorema da representação existe um único vetor  $v \in V$  representando o funcional linear  $f : V \rightarrow \mathbb{R}$  e existe um único vetor  $v^* \in V^*$  representando o funcional linear  $A^*[f] : V \rightarrow \mathbb{R}$ . Pela definição de adjunta temos que

$$A^*[f](w) = f(A(w)) = \langle A(w), v \rangle,$$

por outro lado, pelo Teorema da representação de um funcional podemos escrever

$$A^*[f](w) = \langle w, v^* \rangle$$

para todo  $w \in V$ . Esses comentários permitem-nos definir uma aplicação

$$A^t : V \rightarrow V, \quad v \mapsto v^*,$$

satisfazendo a identidade  $\langle w, A^t(v) \rangle = \langle A(w), v \rangle$  para quaisquer  $v, w \in V$ .

**Proposição 7.6.1** *Seja  $A$  um operador linear no espaço Euclidiano  $V$ . Então existe um único operador linear  $A^t : V \rightarrow V$ , chamado de transposto de  $A$ , tal que  $\langle w, A^t(v) \rangle = \langle A(w), v \rangle$  para quaisquer  $v, w \in V$ .*

**Demonstração** A existência de uma aplicação  $A^t : V \rightarrow V$  satisfazendo a condição  $\langle w, A^t(v) \rangle = \langle A(w), v \rangle$  para quaisquer  $v, w \in V$  está feita nos comentários anteriores. Mostremos a linearidade. Dados  $u, v \in V$  e  $\lambda \in \mathbb{R}$ , para qualquer  $w \in V$  temos as igualdades

$$\begin{aligned} \langle w, A^t(u + \lambda v) \rangle &= \langle A(w), u + \lambda v \rangle \\ &= \langle A(w), u \rangle + \lambda \langle A(w), v \rangle \\ &= \langle w, A^t(u) \rangle + \langle w, \lambda A^t(v) \rangle \\ &= \langle w, A^t(u) + \lambda A^t(v) \rangle. \end{aligned}$$

Sendo a igualdade  $\langle w, A^t(u + \lambda v) \rangle = \langle w, A^t(u) + \lambda A^t(v) \rangle$  verdadeira para todo  $w \in V$ , concluímos que  $A^t(u + \lambda v) = A^t(u) + \lambda A^t(v)$ , como queríamos demonstrar.

A unicidade do operador transposto tem uma demonstração que, em essência, será utilizada outras vezes. Se  $B : V \rightarrow V$  é um operador linear tal que  $\langle A(v), w \rangle = \langle w, B(v) \rangle$  para quaisquer  $w, v \in V$ , então  $\langle w, A^t(v) - B(v) \rangle = 0$  para todo  $w \in V$ , implicando que  $A^t(v) = B(v)$  para todo  $v \in V$ .  $\square$

Como já vimos, a adjunta  $A^*$  depende apenas do operador  $A$ , o mesmo não ocorrendo com o transposto que depende do operador e do produto interno que define a estrutura de espaço Euclidiano com o qual estamos trabalhando. Cada produto interno no espaço vetorial determina um representante diferente para o mesmo funcional linear e, portanto, um operador transposto diferente. Recordamos que temos uma definição de matriz transposta. Estabeleceremos a relação entre esses conceitos utilizando representações em bases ortonormais.

**Proposição 7.6.2** *Sejam  $A$  um operador linear no espaço Euclidiano  $V$  e  $\beta$  uma base ortonormal de  $V$ . Então a representação matricial do operador transposto de  $A$  é a transposta da representação matricial de  $A$ . De outra maneira,  $[A^t]_\beta = [A]_\beta^t$ .*

**Demonstração** Consideremos as representações matriciais  $[A]_\beta = [a_{ij}]$  e  $[A^t]_\beta = [b_{ij}]$ , onde  $\beta = \{v_1, v_2, \dots, v_n\}$  é uma base ortonormal de  $V$ . Como sabemos as entradas das representações matriciais são obtidas pelas fórmulas

$$a_{ij} = \langle A(v_j), v_i \rangle, \quad b_{ij} = \langle A^t(v_j), v_i \rangle.$$

Portanto, pela definição de operador transposto segue que

$$b_{ij} = \langle A^t(v_j), v_i \rangle = \langle v_j, A(v_i) \rangle = \langle A(v_i), v_j \rangle = a_{ji},$$

como pretendíamos provar.  $\square$

Quando  $V$  é um espaço Euclidiano podemos definir a aplicação

$$\Psi : \mathcal{L}(V, V) \rightarrow \mathcal{L}(V, V), \quad \Psi(A) = A^t.$$

Entre as muitas propriedades de  $\Psi$  a proposição abaixo mostra que ela é uma transformação linear.

**Proposição 7.6.3** *Valem as seguintes propriedades para quaisquer operadores lineares  $A$  e  $B$  em  $(V, \langle \cdot, \cdot \rangle)$  e qualquer escalar  $\lambda \in \mathbb{R}$ .*

a)  $(A + \lambda B)^t = A^t + \lambda B^t$ .

b)  $(A \circ B)^t = B^t \circ A^t$ .

c)  $(A^t)^t = A$ .

**Demonstração 1.** Por um lado, temos as igualdades,

$$\begin{aligned}\langle (A + \lambda B)(u), v \rangle &= \langle A(u), v \rangle + \lambda \langle B(u), v \rangle \\ &= \langle u, A^t(v) + \lambda B^t(v) \rangle \\ &= \langle u, (A^t + \lambda B^t)(v) \rangle,\end{aligned}$$

por outro lado,  $\langle (A + \lambda B)(u), v \rangle = \langle u, (A + \lambda B)^t(v) \rangle$ . A unicidade do transposto implica que  $(A + \lambda B)^t = A^t + \lambda B^t$ .

2. A demonstração é novamente uma manipulação elementar de propriedades,

$$\langle (A \circ B)(u), v \rangle = \langle B(u), A^t(v) \rangle = \langle u, (B^t \circ A^t)(v) \rangle.$$

Também sabemos da identidade  $\langle (A \circ B)(u), v \rangle = \langle u, (A \circ B)^t(v) \rangle$ . Pela unicidade do transposto concluímos que  $(A \circ B)^t = B^t \circ A^t$ .

3. Deixaremos a demonstração aos cuidados do leitor.  $\square$

Merece ser enfatizado por um lema duas propriedades sobre o transposto de um operador linear, não pela dificuldade de demonstração que é nenhuma, mas pelo seu frequente uso.

**Lema 7.6.1** *Um operador linear  $A$  num espaço Euclidiano  $V$ , possui as seguintes propriedades.*

- a) Se  $A$  preseva um subespaço  $W$ , então  $A^t$  preserva  $W^\perp$ .
- b) O operador  $A$  induz uma decomposição em soma direta ortogonal

$$V = \text{Im } A \boxplus \text{Nuc } A^t.$$

**Demonstração 1)** Se  $v \in W^\perp$ , pela invariância de  $W$  temos que  $\langle v, A(w) \rangle = 0$  para qualquer vetor  $w \in W$ , e por definição de transposto vale a igualdade  $\langle A^t(v), w \rangle = 0$ . Portanto  $A^t(v) \in W^\perp$ , significando que  $W^\perp$  é invariante por  $A^t$ .

2. Considere a decomposição em soma direta ortogonal  $V = \text{Im } A \boxplus (\text{Im } A)^\perp$ . Verifiquemos a inclusão  $(\text{Im } A)^\perp \subset \text{Nuc } A^t$ . Se  $w$  é um vetor em  $(\text{Im } A)^\perp$ , então para todo  $v \in V$  temos a igualdade  $\langle A(v), w \rangle = 0$ , pois  $A(v)$  pertence ao subespaço imagem de  $A$ . Logo,  $\langle v, A^t(w) \rangle = 0$  para qualquer  $v \in V$ . Isso significa que  $A^t(w) = 0$ , mostrando a inclusão desejada. A inclusão oposta admite uma demonstração semelhante.  $\square$

Recomendamos ao leitor a leitura do Exercício proposto número 2, ali estão registrados alguns fatos que serão utilizados no próximo capítulo.

### Exercícios propostos 7.6.1

1. Calcule o operador transposto de  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $A(x, y) = (x - 2y, 3x + y)$ .
2. Seja  $A$  um operador em  $(V, \langle \cdot, \cdot \rangle)$ .
  - a) Prove que  $B : V \rightarrow V$  é o operador transposto de  $A \Leftrightarrow [B]_\beta = [A]_\beta^t$ , em que  $\beta$  é uma base ortonormal de  $V$ .
  - b) Demonstre que se  $A$  é invertível então  $A^t$  é invertível e  $(A^{-1})^t = (A^t)^{-1}$ .
  - c) Mostre que  $A \circ A^t \equiv 0 \Leftrightarrow A \equiv 0$ .
  - d) Verifique que  $p(A)^t = p(A^t)$  para qualquer polinômio  $p(t) \in \mathbb{R}[t]$ .
  - e) Os ideais anuladores de  $A$  e  $A^t$  são iguais, ou equivalentemente, os polinômios mínimo são iguais. Prove essa afirmação.

3. Fixado um vetor não nulo  $v_0$  em um espaço Euclidiano  $V$ , calcule o operador transposto de

$$\pi : V \rightarrow V, \quad \pi(w) = w - \frac{\langle w, v_0 \rangle}{\langle v_0, v_0 \rangle} v_0.$$

4. Mostre que a aplicação  $\langle N, P \rangle = \text{tr}(NP^t)$  é um produto interno no espaço das matrizes  $M(n, \mathbb{R})$ . Responda as perguntas abaixo considerando esse produto interno.
  - a) Calcule o complemento ortogonal do subespaço das matrizes diagonais.
  - b) Assuma que  $P_0$  é uma matriz invertível em  $M(n, \mathbb{R})$ . Determine o transposto do operador  $A : M(n, \mathbb{R}) \rightarrow M(n, \mathbb{R})$ , quando
    - i)  $A(N) = P_0 N$ ,
    - ii)  $A(N) = N P_0$ ,
    - iii)  $A(N) = P_0^{-1} N P_0$ .

5. Dado o produto interno no espaço  $\mathbb{R}_3[t] = \{p(t) \in \mathbb{R}[t]; \text{grau } p(t) \leq 3\}$  definido por

$$\langle p(t), q(t) \rangle = \int_0^1 p(t) q(t) dt$$

Descreva o transposto da derivação  $D : \mathbb{R}_3[t] \rightarrow \mathbb{R}_3[t]$ ,  $p(t) \mapsto p'(t)$ .

6. Fixados vetores  $u_0$  e  $v_0$  de um espaço Euclidiano  $V$ , calcule o transposto do operador linear  $A : V \rightarrow V$ ,  $A(v) = \langle v, u_0 \rangle v_0$ .
7. Calcule uma base ortonormal para os  $\text{Nuc } A$ ,  $\text{Im } A$ ,  $\text{Nuc } A^t$  e  $\text{Im } A^t$  quando o operador é  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (x + y, x + y, x + y)$ .
8. Mostre a seguinte fórmula para a transposta de uma matriz quadrada constituída de quatro submatrizes,

$$\begin{bmatrix} P & Q \\ R & S \end{bmatrix}^t = \begin{bmatrix} P^t & R^t \\ Q^t & S^t \end{bmatrix}.$$

## Capítulo 8

# Operadores normais

Iniciaremos o estudo da classe de operadores lineares num espaço Euclidiano que comutam com o seu transposto, chamados de operadores normais. As representações matriciais canônicas desses operadores serão apresentadas tendo em vista que alguns tipos de matrizes que surgem dessas representações formam subgrupos importantes no espaço das matrizes quadradas. A linha de desenvolvimento dessa seção é simples. Primeiro estudaremos a decomposição primária de um operador normal e depois a decomposição cíclica de cada parcela da decomposição primária.

### 8.1 Operadores normais

Diz-se que um operador  $A$  num espaço Euclidiano  $V$  é *normal* se  $A \circ A^t = A^t \circ A$ .

A terminologia utilizada é apropriada pois o conceito de normalidade estará sempre associado ao conceito de ortonormalidade e as boas representações desses operadores serão obtidas em bases ortonormais dando origem às matrizes normais. Apresentemos alguns exemplos de operadores normais em  $(V, \langle \cdot, \cdot \rangle)$ .

**Exemplo 8.1.1** 1) Conhecendo-se um único operador normal  $A : V \rightarrow V$  é possível contruir vários outros. Por exemplo, o operador transposto  $A^t$  bem como qualquer potência  $A^n$  com  $n \geq 0$  são normais. Além desse, podemos provar que os operadores  $B = A - A^t$  e  $C = \lambda A$ , com  $\lambda \in \mathbb{R}$ , são também operadores normais. Mostra-se por indução que  $p(A)$  é normal para qualquer polinômio  $p(t) \in \mathbb{R}[t]$ .

2) Se  $A : V \rightarrow V$  é qualquer operador, o operador  $P = A^t \circ A$  é chamado de *operador de Gram* associado ao operador  $A$ . Ele é um operador normal.

3) A composta e a soma de dois operadores normais num espaço Euclidiano, em geral, não são operadores normais, entretanto, uma condição suficiente para serem normais é que os dois operadores comutem.  $\square$

**Exercício 8.1.1** Demonstre o seguinte critério de normalidade num espaço Euclidiano. Um operador  $A : V \rightarrow V$  é normal  $\Leftrightarrow$  a representação  $[A]_\beta$  é uma matriz normal para qualquer base ortonormal  $\beta$  de  $V$ .  $\square$

Apresentaremos as primeiras propriedades de um operador normal estabelecendo as relações entre o núcleo e a imagem de  $A$  e de seu transposto.

**Proposição 8.1.1** Se  $A$  é um operador normal num espaço Euclidiano  $(V, \langle \cdot, \cdot \rangle)$ , então  $\|A(v)\| = \|A^t(v)\|$  para qualquer vetor  $v \in V$ . Em consequência,

- a) os núcleos de  $A$  e  $A^t$  são iguais;
- b)  $V$  decompõe-se numa soma direta ortogonal,  $V = \text{Im } A \oplus \text{Nuc } A$ ;
- c) as imagens de  $A$  e  $A^t$  são iguais.

**Demonstração** A comutatividade de  $A$  e de seu transposto  $A^t$  justifica as igualdades,

$$\langle A(v), A(v) \rangle = \langle v, A^t \circ A(v) \rangle = \langle v, A \circ A^t(v) \rangle = \langle A^t(v), A^t(v) \rangle,$$

para qualquer vetor  $v \in V$ . Portanto,  $\|A(v)\|^2 = \|A^t(v)\|^2$ , de onde segue a afirmação.

As consequências são imediatas. a) Da equivalência  $\|A(v)\| = 0$  se, e somente se,  $\|A^t(v)\| = 0$  segue a igualdade entre os núcleos de  $A$  e  $A^t$ . b) Sabemos que  $V = \text{Im } A \oplus \text{Nuc } A^t$  para qualquer operador  $A : V \rightarrow V$ . Pelo primeiro item temos a decomposição desejada. O item c) é deixado para exercício.  $\square$

### Exercícios propostos 8.1.1

1. Verifique que o operador de Gram associado a um operador  $A$  num espaço Euclidiano  $V$  é invertível se, e somente se,  $A$  é invertível.
2. Dê um exemplo de um operador  $A$  num espaço Euclidiano  $V$  tal que  $A \circ A^t \neq A^t \circ A$ .
3. Demonstre que qualquer matriz normal  $N \in M(2, \mathbb{R})$  é da forma

$$N = r \begin{bmatrix} \cos \theta & -\text{sen } \theta \\ \text{sen } \theta & \cos \theta \end{bmatrix} \quad \text{ou} \quad N = \begin{bmatrix} a & b \\ b & c \end{bmatrix},$$

onde  $r, a, b, c \in \mathbb{R}$  e  $\theta \in [0, 2\pi)$ .

4. Assuma que um operador normal  $A : V \rightarrow V$  preserva o subespaço  $W \subset V$ .
  - a) Prove que o operador  $A_0 : W \rightarrow W$  obtido por restrição de  $A$  é normal.
  - b) Demonstre que  $V = W \oplus W^\perp$  é uma soma direta ortogonal invariante por  $A$ .
  - c) Conclua que essa soma direta também é invariante por  $A^t$ .

## 8.2 Decomposição normal

Operadores normais recebem uma versão particular do Teorema da decomposição primária. Novamente, veremos que toda informação está concentrada no polinômio minimal do operador.

**Teorema 8.2.1 (Teo. da decomposição normal)** *Seja  $A$  um operador normal num espaço Euclidiano  $V$ . Então*

- a) a fatoração primária do polinômio minimal de  $A$  é um produto de polinômios mônicos primos sem repetições,  $m_A(t) = p_1(t)p_2(t) \cdots p_k(t)$ ;
- b) a decomposição primária de  $V$  é uma soma direta ortogonal,
- $$V = \text{Nuc } p_1(A) \boxplus \text{Nuc } p_2(A) \boxplus \cdots \boxplus \text{Nuc } p_k(A).$$

**Demonstração** a) Por absurdo, vamos assumir que  $m_A(t) = p_i(t)^2 q(t)$  para algum índice  $i$ . Essa hipótese implica que para qualquer  $v \in V$  o vetor  $w = p_i(A) \circ q(A)(v)$  pertence à interseção  $\text{Im } p_i(A) \cap \text{Nuc } p_i(A)$ . Entretanto, essa interseção é trivial pois  $V = \text{Im } p_i(A) \boxplus \text{Nuc } p_i(A)$  desde que  $p_i(A)$  é um operador normal. Isso significa que  $p_i(A) \circ q(A)(v) = 0$  para todo  $v \in V$ . Sendo assim, o polinômio  $p_i(t)q(t)$  pertence ao ideal anulador  $\mathfrak{S}_A$  e tem grau menor que grau do gerador do ideal, evidentemente uma contradição. Isso mostra que a fatoração primária do polinômio minimal não tem fatores repetidos.

b) Como sabemos, a projeção  $\pi_i : V \rightarrow V$  sobre a  $i$ -ésima parcela da decomposição primária ao longo das outras parcelas é um polinômio em  $A$ , portanto,  $\pi_i$  é um operador normal e por conseguinte  $V = \text{Im } \pi_i \boxplus \text{Nuc } \pi_i$ . Como para todo  $i$  temos que

$$\begin{cases} \text{Im } \pi_i = \text{Nuc } p_i(A), \\ \text{Nuc } \pi_i = \text{Nuc } p_1(A) \oplus \cdots \oplus \widehat{\text{Nuc } p_i(A)} \oplus \cdots \oplus \text{Nuc } p_k(A), \end{cases}$$

segue que as parcelas da decomposição primária são duas a duas ortogonais.  $\square$

A decomposição primária induzida por um operador normal será chamada de *decomposição normal*. Comparemos as decomposições normais de um operador e do seu transposto num espaço Euclidiano  $V$ .

**Corolário 8.2.1** *Um operador normal e seu transposto induzem a mesma decomposição normal.*

**Demonstração** Recordamos que os polinômios minimais de um operador  $A$  e de seu transposto  $A^t$  são iguais, fato geral que independe da normalidade. Portanto, a decomposição normal definida por  $A^t$  fica sendo

$$V = \text{Nuc } p_1(A^t) \boxplus \text{Nuc } p_2(A^t) \boxplus \cdots \boxplus \text{Nuc } p_k(A^t).$$

Como  $p_i(A^t) = p_i(A)^t$  e  $p_i(A)$  é também normal, temos as igualdades dos núcleos,  $\text{Nuc } p_i(A^t) = \text{Nuc } p_i(A)^t = \text{Nuc } p_i(A)$ .  $\square$

### Exercícios propostos 8.2.1

1. Seja  $A$  um operador normal no espaço Euclidiano  $V$ .
  - a) Prove que autovetores de  $A$  associados a autovalores distintos são ortogonais.
  - b) Mostre que se  $A^k(v) = 0$  para algum  $k \geq 2$ , então  $A(v) = 0$ .
2. Descreva todos operadores num espaço Euclidiano que são nilpotentes e normais.
3. Assuma que  $A$  é um operador nilpotente num espaço Euclidiano  $V$  e que o espaço é  $A$ -cíclico,  $V = \mathcal{C}(v)$ .
  - a) Determine a decomposição normal do operador de Gram  $P = A^t \circ A$
  - b) e conclua que  $P$  é uma projeção ortogonal.
  - c) Calcule a representação matricial de  $P$  na base cíclica gerada por  $v$ .
4. Com as mesmas hipóteses do item anterior, exceto que  $V$  seja  $A$ -cíclico, podemos concluir que  $P$  é uma projeção ortogonal?
5. Prove que se  $A$  e  $B$  são operadores normais no espaço Euclidiano  $V$  que comutam então  $A$  comuta com  $B^t$ .

## 8.3 Operadores simétricos

Iniciamos o estudo de tipos particulares de operadores normais.

Um operador  $A$  num espaço Euclidiano  $V$  é *simétrico* quando  $A = A^t$ .

É trivial verificar que um operador simétrico é normal. O principal teorema sobre tais operadores que iremos demonstrar é o Teorema espectral, garantindo que um operador simétrico é diagonalizável numa base ortonormal de autovetores. Sem muito esforço esse resultado pode ser demonstrado para espaços Euclidianos de dimensão dois e depois generalizado para espaços com dimensões maiores utilizando o Teorema da decomposição normal. Deixaremos uma série de informações sobre operadores simétricos em um espaço Euclidiano  $V$  como exercícios.

**Exercício 8.3.1** 1) Um operador de Gram é simétrico.

2) Se  $A : V \rightarrow V$  é simétrico então  $p(A)$  é simétrico para todo  $p(t) \in \mathbb{R}[t]$ .

3)  $A : V \rightarrow V$  é simétrico  $\Leftrightarrow$  a representação  $[A]_\beta$  em qualquer base ortonormal  $\beta$  de  $V$  é uma matriz simétrica.  $\square$

Demonstremos um lema auxiliar que contém quase todas as informações necessárias para a demonstração do Teorema espectral.

**Lema 8.3.1** *Seja  $A$  um operador simétrico em  $(V, \langle \cdot, \cdot \rangle)$ . Se  $\dim V = 2$  então valem as afirmações:*

- a) *o polinômio característico fatora-se num produto de dois polinômios lineares,  $p_A(t) = (t - \lambda_1)(t - \lambda_2)$ , em que  $\{\lambda_1, \lambda_2\} \subset \mathbb{R}$  são autovalores não necessariamente distintos;*
- b) *o operador é diagonalizável por uma base ortonormal de autovetores.*

**Demonstração** a). Fixado uma base ortonormal  $\gamma = \{v_1, v_2\} \subset V$ , a representação de  $A$  nessa base é uma matriz simétrica, digamos que seja

$$[A]_\gamma = \begin{bmatrix} a & b \\ b & c \end{bmatrix}.$$

O discriminante  $\Delta$  do polinômio característico  $p_A(t) = t^2 - (a + c)t + (ac - b^2)$  é não negativo,  $\Delta = (a - c)^2 + b^2 \geq 0$ . Portanto,  $p_A(t)$  pode ser fatorado em polinômios lineares,  $p_A(t) = (t - \lambda_1)(t - \lambda_2)$ , em que  $\{\lambda_1, \lambda_2\} \subset \mathbb{R}$  são escalares não necessariamente distintos.

- b) Para demonstrar esse item é conveniente examinar dois casos.

*Primeiro caso* Assumamos que o discriminante seja nulo. Nesse caso, as entradas da representação  $[A]_\beta$  tem os valores  $a = c$  e  $b = 0$ . Portanto, o operador  $A$  tem um autovalor  $\lambda = a$ , todo vetor de  $V$  é um autovetor e

$$[A]_\beta = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}.$$

Isso mostra que a base ortonormal  $\gamma = \{v_1, v_2\}$ , ou qualquer outra base ortonormal, é formada de autovetores. Observamos que o polinômio minimal de  $A$  é o polinômio linear  $m_A(t) = (t - \lambda)$  e que  $A$  é uma homotetia por  $\lambda$ .

*Segundo caso* Quando o discriminante é positivo as raízes do polinômio característico  $\{\lambda_1, \lambda_2\} \subset \mathbb{R}$  são distintas, implicando que o polinômio minimal é igual ao polinômio característico,  $m_A(t) = (t - \lambda_1)(t - \lambda_2)$ , e cada parcela da decomposição normal

$$V = \text{Nuc}(A - \lambda_1 Id) \boxplus \text{Nuc}(A - \lambda_2 Id)$$

é um autoespaço unidimensional  $V_{\lambda_i}$ . Escolhendo um vetor unitário  $u_i \in V_{\lambda_i}$ , a base  $\beta = \{u_1, u_2\}$  é ortonormal e formada de autovetores.  $\square$

**Teorema 8.3.1 (Teorema espectral)** *Seja  $A$  um operador linear num espaço Euclidiano  $V$ . Se  $A$  é simétrico, então*

- a) a fatoração primária do polinômio minimal de  $A$  é um produto de polinômios lineares distintos,  $m_A(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_k)$ ;
- b) a decomposição normal determinada por  $A$  é uma soma direta ortogonal de autoespaços,  $V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k}$ .

Em particular,  $A$  é diagonalizável por uma base ortonormal de autovetores.

**Demonstração** Pelo Teorema da decomposição normal sabemos que o polinômio minimal de  $A$  fatora-se em um produto de polinômios mônicos, primos e dois a dois distintos,  $m_A(t) = p_1(t)p_2(t) \cdots p_k(t)$ . Por absurdo, assumamos que algum fator  $p_i(t)$  seja um polinômio de grau dois. Considere o espaço  $A$ -cíclico  $\mathcal{C}(v) \subset \text{Nuc } p_i(A)$  gerado por um vetor não nulo  $v \in \text{Nuc } p_i(A)$ . Como  $m_v(t)$  é um polinômio não constante que divide o polinômio primo  $p_i(t)$ , necessariamente temos que  $m_v(t) = p_i(t)$  e por conseguinte,

$$\dim \mathcal{C}(v) = \text{grau } m_v(t) = 2.$$

O operador linear induz por restrição um operador simétrico em  $(\mathcal{C}(v), \langle \cdot, \cdot \rangle)$ , pois esse espaço é invariante por  $A$  e  $A^t = A$ . Aqui surge a contradição. Por um lado, o polinômio minimal da restrição é o polinômio primo de grau dois  $p_i(t)$ , por outro lado, o lema anterior afirma que esse polinômio fatora-se em um produto de polinômios lineares! Logo, cada fator da decomposição primária é um polinômio linear.

A demonstração de 2 é imediata. Em particular, para diagonalizar o operador escolhamos uma base ortonormal  $\beta_i \subset \text{Nuc } p_i(A)$  e construímos a base ortonormal  $\beta = \bigcup \beta_i$  de  $V$ .  $\square$

**Corolário 8.3.1** *Um operador num espaço Euclidiano é simétrico se, e somente se, o operador é diagonalizável por uma base ortonormal de autovetores.*

Não roubaremos do leitor o prazer de demonstrar esse corolário.

**Exercício 8.3.2** Prove que um operador normal  $A$  em  $(V, \langle \cdot, \cdot \rangle)$  é simétrico se, e somente se, todas as raízes do polinômio característico são reais.  $\square$

### Exercícios propostos 8.3.1

1. Verifique que os operadores são simétricos e diagonalize-os.

a)  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $A(x, y) = (4x + 12y, 12x - 3y)$ .

- b)  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (2z, -y, 2x)$ .
- c)  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (x + z, -y, x + z)$ .
2. Fixado um vetor não nulo  $v_0$  no espaço Euclidiano  $V$ , defina a aplicação  $\pi : V \rightarrow V$ ,  $\pi(w) = \langle w, v_0 \rangle v_0$ .
- a) Mostre que  $\pi$  é um operador linear simétrico.
- b) Identifique as parcelas da decomposição normal determinada por  $\pi$ .
3. Suponha que  $A$  é um operador linear diagonalizável no espaço real de dimensão finita  $V$ . Mostre que podemos definir um produto interno em  $V$  que torna  $A$  um operador simétrico.
4. Sejam  $V = W_1 \oplus W_2$  uma decomposição em soma direta do espaço Euclidiano  $V$  e  $\pi_i : V \rightarrow V$  a projeção sobre a  $i$ -ésima parcela ao longo da outra. Prove as equivalências.
- a)  $\pi_i$  é um operador normal.
- b)  $\pi_i$  é um operador simétrico.
- c) A soma direta é ortogonal.
5. Suponha que um operador linear simétrico  $A$  em  $(V, \langle \cdot, \cdot \rangle)$  é tal que  $A^n \equiv Id$  para algum inteiro  $n \geq 2$ . Demonstre que  $A^2 \equiv Id$ .
6. Prove que dois operadores simétricos em  $(V, \langle \cdot, \cdot \rangle)$  comutam se, e somente se, eles são simultaneamente diagonalizáveis.
7. Uma involução  $\mathcal{S}$  num espaço Euclidiano  $V$  é normal se, e somente se,  $\mathcal{S}$  é simétrica.
8. Prove que o traço de um operador simétrico num espaço Euclidiano é a soma das raízes do polinômio característico (contando as multiplicidades).

## 8.4 Decomposição cíclica normal

Para cada operador  $A$  num espaço Euclidiano  $V$  iremos construir uma decomposição  $A$ -cíclica cujas parcelas são duas a duas ortogonais. Isso permitirá fazer uma representação matricial do operador, utilizada para identificar os exemplos clássicos de operadores normais. Por clareza, os teoremas sobre decomposições cíclicas estão direcionados para o estudo da restrição do operador a uma das parcelas da decomposição normal. Iniciemos com uma versão dos teoremas das decomposições cíclicas para operadores normais.

**Teorema 8.4.1 (Teorema da decomposição cíclica normal)** *Seja  $A$  um operador normal no espaço Euclidiano  $V$ . Se  $m_A(t) = p(t)$  é a fatoração primária do polinômio minimal de  $A$  então existem vetores  $v_1, v_2, \dots, v_n \in V$  tais que*

$$\begin{cases} V = \mathcal{C}(v_1) \boxplus \mathcal{C}(v_2) \boxplus \cdots \boxplus \mathcal{C}(v_n), \\ \text{grau } m_A(t) = \dim \mathcal{C}(v_1) = \dim \mathcal{C}(v_2) = \cdots = \dim \mathcal{C}(v_n). \end{cases}$$

Essa decomposição  $A$ -cíclica é dimensionalmente única.

**Demonstração** Dividiremos a demonstração em dois casos.

*Primeiro caso.* Assumamos que o polinômio minimal de  $A$  é o polinômio linear  $m_A(t) = (t - \lambda)$ . Sendo assim, todo vetor é um autovetor associado ao autovalor  $\lambda$  e escolhendo-se uma base ortonormal  $\beta = \{v_1, v_2, \dots, v_n\}$  de  $V = V_\lambda$  obtemos a decomposição  $A$ -cíclica

$$\begin{cases} V = \mathcal{C}(v_1) \boxplus \mathcal{C}(v_2) \boxplus \cdots \boxplus \mathcal{C}(v_n), \\ \text{grau } m_A(t) = 1 = \dim \mathcal{C}(v_1) = \dim \mathcal{C}(v_2) = \cdots = \dim \mathcal{C}(v_n). \end{cases}$$

*Segundo caso* Vamos supor que o polinômio minimal é um polinômio primo de grau dois,  $m_A(t) = (t - \lambda)^2 + \tau^2$ , com  $\tau > 0$ . Nesse caso considere o operador de Gram  $P : V \rightarrow V$ ,  $P(v) = A^t \circ A(v)$ . Sendo  $P$  um operador simétrico, pelo Teorema espectral podemos escolher um autovetor  $v \in V$  associado a um autovalor  $\mu \in \mathbb{R}$ , mais precisamente,  $P(v) = \mu v$ . Demonstraremos que a soma direta ortogonal

$$V = \mathcal{C}_A(v) \boxplus \mathcal{C}_A(v)^\perp$$

é invariante por  $A$  e  $A^t$ . Como  $\mathcal{C}_A(v)$  é invariante por  $A$ , então o complemento ortogonal  $\mathcal{C}_A(v)^\perp$  é invariante por  $A^t$ . Logo, é suficiente mostrar que a primeira parcela é invariante também por  $A^t$  pois isso implicará que a segunda será invariante por  $A$ . Reduzindo ainda mais a questão. É suficiente demonstrar que  $\gamma = \{v, A(v)\}$  é uma base ordenada de  $\mathcal{C}_A(v)$  tal que  $A^t(\gamma) \subset \mathcal{C}_A(v)$ . De fato  $\gamma$  é uma base pois todo espaço  $A$ -cíclico não trivial tem dimensão dois desde que o polinômio minimal de um vetor não nulo é igual ao polinômio minimal de  $A$ . Verifiquemos a condição  $A^t(\gamma) \subset \mathcal{C}_A(v)$ . Dado o elemento  $A(v) \in \gamma$  temos que

$$A^t(A(v)) = P(v) = \mu v,$$

mostrando que  $A^t(A(v)) \in \mathcal{C}_A(v)$ . Consideremos o outro elemento  $v \in \gamma$ . Para mostrar que  $A^t(v) \in \mathcal{C}_A(v)$  devemos levar em conta a comutatividade dos operadores  $A$ ,  $A^t$  e  $P$  como também o desenvolvimento  $m_A(t) = t^2 - 2\lambda t + (\lambda^2 + \tau^2)$ . Observe que  $o = A^t(o) = A^t(m_A(A)(v))$ , portanto,

$$\begin{aligned} o &= A^t(A^2(v) - 2\lambda A(v) + (\lambda^2 + \tau^2)v) \\ &= P(A(v)) - 2\lambda P(v) + (\lambda^2 + \tau^2)A^t(v) \\ &= \mu A(v) - 2\lambda \mu v + (\lambda^2 + \tau^2)A^t(v). \end{aligned}$$

Dessas igualdades é possível concluir que  $A^t(v)$  é uma combinação linear dos elementos da base  $\gamma$ , mais precisamente,

$$A^t(v) = \frac{-1}{\lambda^2 + \tau^2} (2\mu\lambda v - \mu A(v)).$$

Isso demonstra que  $A^t(\gamma) \subset \mathcal{C}_A(v)$ , como desejávamos.

A continuação da demonstração do segundo caso é feita por indução sobre a dimensão de  $V$ . Devemos mostrar que se  $A$  é um operador simétrico no espaço Euclidiano  $V$  com  $\dim V = 2n$  e polinômio mínimo  $m_A(t) = (t - \lambda)^2 + \tau^2$ , então existem vetores  $v_1, v_2, \dots, v_n \in V$  tais que

$$\begin{cases} V = \mathcal{C}(v_1) \boxplus \mathcal{C}(v_2) \boxplus \dots \boxplus \mathcal{C}(v_n), \\ \text{grau } m_A(t) = 2 = \dim \mathcal{C}(v_1) = \dim \mathcal{C}(v_2) = \dots = \dim \mathcal{C}(v_n). \end{cases}$$

Note que para  $n = 1$  a verificação é trivial pois  $V = \mathcal{C}(v)$  para qualquer vetor não nulo  $v \in V$ . A construção acima é o passo de indução que reduz a dimensão em 2, pois  $A$  e  $A^t$  preservam o subespaço cíclico  $\mathcal{C}(v_1)^\perp$ . Deixaremos para o leitor completar esse final de demonstração.  $\square$

Uma decomposição  $A$ -cíclica determinada por um operador normal será chamada de *decomposição  $A$ -cíclica normal*.

**Corolário 8.4.1** *Uma decomposição  $A$ -cíclica normal é também uma decomposição  $A^t$ -cíclica normal.*

**Demonstração** Vamos supor que a fatoração primária do polinômio minimal do operador normal  $A$  seja  $m_A(t) = p(t)$ . Desse modo, considere uma decomposição  $A$ -cíclica normal

$$\begin{cases} V = \mathcal{C}(v_1) \boxplus \mathcal{C}(v_2) \boxplus \dots \boxplus \mathcal{C}(v_n), \\ \text{grau } m_A(t) = \dim \mathcal{C}(v_1) = \dim \mathcal{C}(v_2) = \dots = \dim \mathcal{C}(v_n). \end{cases}$$

Como o operador  $A$  preserva a soma direta ortogonal

$$W_i = \mathcal{C}(v_1) \boxplus \dots \boxplus \widehat{\mathcal{C}(v_i)} \boxplus \dots \boxplus \mathcal{C}(v_n),$$

então  $A^t$  preserva o complemento ortogonal de  $W_i$  que é  $W_i^\perp = \mathcal{C}(v_i)$ . A igualdade dos polinômios mínimos  $m_A(t)$  e  $m_{A^t}(t)$ , garante que cada parcela  $A$ -cíclica é uma parcela  $A^t$ -cíclica.  $\square$

#### Exercícios propostos 8.4.1

1. Suponha que a representação do operador  $A : \mathbb{R}^4 \rightarrow \mathbb{R}^4$  na base canônica seja a matriz ao lado.

a) Mostre que  $A$  é normal (em relação ao produto interno canônico).

$$[A]_{\alpha_4} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

b) Dê uma decomposição  $A$ -cíclica normal explicitando os vetores que geram as parcelas.

2. Mostre que  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (z, x, y)$  é normal e determine uma decomposição  $A$ -cíclica.

3. Considere o operador  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (y - 3z, -x - 2z, 2x + 3y)$ .

a) Verifique que  $A$  é um operador normal.

b) Determine uma decomposição cíclica normal explicitando os vetores que geram as parcelas cíclicas.

4. Demonstre que toda matriz normal  $N \in M(3, \mathbb{R})$  é conjugada a uma matriz da forma

$$Q = \begin{bmatrix} \mu & 0 \\ 0 & P \end{bmatrix} \quad \text{onde} \quad P = r \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \quad \text{ou} \quad P = \begin{bmatrix} a & b \\ b & d \end{bmatrix}.$$

## 8.5 Representação de operadores normais

O Teorema da decomposição cíclica permite-nos construir uma representação bastante simples de um operador normal pois sabemos decompor o espaço  $V$  em subespaços unidimensionais ou bidimensionais invariantes pelo operador. Para isso, recordemos uma notação matricial. Denotamos por  $R(\lambda, \tau)$  a matriz  $2 \times 2$

$$R(\lambda, \tau) = \begin{bmatrix} \lambda & -\tau \\ \tau & \lambda \end{bmatrix}, \quad \text{em que} \quad \lambda, \tau \in \mathbb{R}.$$

Note que  $p(t) = (t - \lambda)^2 + \tau^2$  é o polinômio característico de  $R(\lambda, \tau)$ .

**Proposição 8.5.1** *Se  $A$  é um operador normal no espaço Euclidiano  $V$  de dimensão dois com polinômio minimal  $m_A(t) = (t - \lambda)^2 + \tau^2$ ,  $\tau > 0$ , então existe uma base ortonormal  $\beta$  tal que*

$$[A]_{\beta} = \begin{bmatrix} \lambda & -\tau \\ \tau & \lambda \end{bmatrix}.$$

**Demonstração** Seja  $\beta = \{u_1, u_2\}$  uma base ortonormal de  $V$ . Digamos que a representação matricial de  $A$  nesta base seja a matriz

$$[A]_{\beta} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

A irreduzibilidade do polinômio minimal de  $A$  garante que  $c \neq 0$ , na verdade podemos assumir que  $c > 0$ , caso contrário consideramos a base ordenada  $\beta = \{u_1, -u_2\}$  e essa condição ficará satisfeita. Para demonstrar a proposição é suficiente mostrar que  $b = -c$  e  $a = d$ . As hipóteses de  $A$  ser um operador normal e a base ser ortonormal implicam que  $[A]_\beta [A]_\beta^t = [A]_\beta^t [A]_\beta$ . Efetuando-se o produto matricial obtemos quatro equações das quais destacaremos duas,

$$\begin{cases} b^2 = c^2 \\ ac + bd = ab + cd \end{cases} .$$

Como o polinômio minimal tem grau dois e a dimensão de  $V$  também é dois, necessariamente o polinômio característico de  $A$  é igual ao polinômio mínimo,  $p_A(t) = t^2 - (a + d)t + (ad - bc)$ . A condição de  $p_A(t)$  ser um polinômio primo, é equivalente a ter discriminante negativo,  $\Delta = (a - d)^2 + 4bc < 0$ . Como  $b^2 = c^2$  e  $c > 0$  necessariamente temos que  $b = -c$ . Por outro lado, substituindo esse valor de  $b$  em  $ac + bd = ab + cd$  determinamos a equação  $2c(a - d) = 0$ , de onde segue que  $a = d$ . Em resumo, a representação de  $A$  na base ortonormal  $\beta$  é a matriz

$$[A]_\beta = \begin{bmatrix} a & -c \\ c & a \end{bmatrix}$$

com  $c > 0$ . Calculando novamente o polinômio minimal obtemos  $m_A(t) = (t - a)^2 + c^2$ . Sendo assim,  $\lambda = a$  e  $\tau = c$ , como desejávamos demonstrar.  $\square$

Na próxima proposição descreveremos uma representação canônica para um operador normal. Recordamos que os enunciados dizem respeito às restrições do operador às parcelas da decomposição normal.

**Proposição 8.5.2** *Se  $m_A(t) = p(t)$  é a fatoração primária do polinômio minimal do operador normal  $A$  num espaço Euclidiano  $V$ , então existe uma base ortonormal  $\beta$  de  $V$  tal que*

- a)  $[A]_\beta = \text{diag}\{\lambda, \lambda, \dots, \lambda\}$  se  $m_A(t) = (t - \lambda)$ , ou  
 b)  $[A]_\beta = \text{diag}\{R(\lambda, \tau), R(\lambda, \tau), \dots, R(\lambda, \tau)\}$ , se  $m_A(t) = (t - \lambda)^2 + \tau^2$ ,  $\tau > 0$ .

**Demonstração** Escolha qualquer decomposição  $A$ -cíclica normal,

$$\begin{cases} V = \mathcal{C}(v_1) \boxplus \mathcal{C}(v_2) \boxplus \dots \boxplus \mathcal{C}(v_n), \\ \text{grau } m_A(t) = \dim \mathcal{C}(v_1) = \dim \mathcal{C}(v_2) = \dots = \dim \mathcal{C}(v_n). \end{cases}$$

a) Se o polinômio minimal é linear,  $m_A(t) = (t - \lambda)$ , cada parcela cíclica é unidimensional e cada gerador  $v_i$  é um autovetor (que pode ser assumido unitário) associado ao autovalor  $\lambda$ . Logo, para diagonalizar o operador escolhemos como base ortonormal o conjunto  $\beta = \{v_1, v_2, \dots, v_n\}$ .

b) Se o polinômio minimal é primo de grau dois,  $m_A(t) = (t - \lambda)^2 + \tau^2$ ,  $\tau > 0$ , cada parcela cíclica tem dimensão dois e é invariante por  $A$  e  $A^t$ . Portanto  $A$  induz por restrição um operador normal em cada parcela  $\mathcal{C}(v_i)$  com o mesmo polinômio mínimo. Pela proposição anterior, existe uma base ortonormal  $\beta_i$  de  $\mathcal{C}(v_i)$  cuja representação da restrição  $A_i$  é a matriz

$$[A_i]_{\beta_i} = \begin{bmatrix} \lambda & -\tau \\ \tau & \lambda \end{bmatrix}.$$

Finalizando a demonstração, escolhamos a base ordenada  $\beta = \vec{\cup} \beta_i$  para obter a representação desejada.  $\square$

O conjunto de matrizes

$$\mathbb{C}_M = \{R(\lambda, \tau); \lambda, \tau \in \mathbb{R}\}$$

equipado com as operações de soma e multiplicação de matrizes tem uma estrutura de corpo. Para deixar claro a relação entre os números complexos e várias representações matriciais aqui descritas, construiremos um isomorfismo de corpo entre os números complexos e esse subconjunto especial de matrizes  $2 \times 2$ . O isomorfismo é a aplicação

$$\Phi: \mathbb{C} \rightarrow \mathbb{C}_M, \quad \Phi(\lambda + i\tau) = \begin{bmatrix} \lambda & -\tau \\ \tau & \lambda \end{bmatrix}.$$

Dado um número complexo não nulo  $z = \lambda + i\tau$  existe um único  $\theta$ ,  $0 \leq \theta < 2\pi$ , chamado de *ângulo polar*, tal que  $z = \|z\|(\cos \theta + i \operatorname{sen} \theta)$ , onde  $\|z\| = \sqrt{\lambda^2 + \tau^2}$ . Essa expressão é chamada de *decomposição polar* de  $z$  e o ângulo  $\theta$  satisfaz as identidades,

$$\cos \theta = \frac{\lambda}{\sqrt{\lambda^2 + \tau^2}} \quad \text{e} \quad \operatorname{sen} \theta = \frac{\tau}{\sqrt{\lambda^2 + \tau^2}}.$$

Transportando a decomposição polar via isomorfismo  $\Phi$  para  $\mathbb{C}_M$  temos a decomposição polar da matriz correspondente,

$$R(\lambda, \tau) = \sqrt{\lambda^2 + \tau^2} \begin{bmatrix} \cos \theta & -\operatorname{sen} \theta \\ \operatorname{sen} \theta & \cos \theta \end{bmatrix}.$$

Definimos a matriz rotação por  $\theta$  como sendo a matriz ortogonal  $2 \times 2$

$$R_\theta = \begin{bmatrix} \cos \theta & -\operatorname{sen} \theta \\ \operatorname{sen} \theta & \cos \theta \end{bmatrix}.$$

No final do capítulo descreveremos uma generalização dessa idéia mostrando uma decomposição polar para uma matriz qualquer. Finalmente, observamos a forma especial do produto matricial

$$R(\lambda, \tau)R(\lambda, \tau)^t = (\lambda^2 + \tau^2)I_2.$$

### Exercícios propostos 8.5.1

1. Suponha que  $m_A(t) = (t^2 + a^2)(t^2 + b^2)$  é o polinômio minimal de um operador normal  $A : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ . Dê a representação matricial canônica para  $A$ .
2. Suponha que  $p_A(t) = (t - 2)^2(t^2 + 1)^2 [(t - 1)^2 + 1]$  seja o polinômio característico de um operador normal  $A : \mathbb{R}^8 \rightarrow \mathbb{R}^8$ . Determine o polinômio minimal e uma representação matricial canônica para  $A$ .
3. Um operador  $A$  no espaço Euclidiano  $V$  é normal  $\Leftrightarrow \|A(v)\| = \|A^t(v)\|$  para qualquer  $v \in V$ ?

## 8.6 Operadores anti-simétricos

Um operador normal  $A$  num espaço Euclidiano  $V$  é *anti-simétrico* se  $A^t = -A$ .

Matricialmente um operador anti-simétrico é identificado pela representação numa base ortonormal, ela é uma matriz antisimétrica. A recíproca dessa afirmação também é verdadeira. Examinemos a decomposição normal de um operador anti-simétrico,

$$V = Nuc p_1(A) \boxplus Nuc p_2(A) \boxplus \cdots \boxplus Nuc p_k(A).$$

Cada fator  $p_i(t)$  do polinômio minimal do operador ou é linear,  $p(t) = t - \lambda$ , ou é um polinômio primo,  $p(t) = (t - \lambda)^2 + \tau^2$ , em que sempre assumimos que  $\tau > 0$ .

Suponha que algum fator  $p_i(t)$  seja linear, então a parcela correspondente da decomposição normal é o autoespaço  $V_{\lambda_i}$  e numa base ortonormal  $\beta_i \subset Nuc p_i(A)$  o operador  $A_i$  obtido por restrição de  $A$  é diagonalizado. Como a base é ortonormal, a matriz  $[A_i]_{\beta_i}$  é antisimétrica e diagonal, logo  $[A_i]_{\beta_i} = [0]$ , significando que a restrição é nula. Portanto, o fator considerado é  $p_i(t) = t$  e a parcela correspondente na decomposição normal é precisamente o núcleo do operador anti-simétrico  $A$ . Isso mostra que só é possível existir um fator linear na decomposição primária do polinômio minimal e esse fator linear é  $p(t) = t$ .

Assuma que  $p_i(t) = (t - \lambda_i)^2 + \tau_i^2$ . Pelo Teorema da decomposição cíclica normal existem  $v_1, v_2, \dots, v_n \in Nuc p_i(A)$  tais que

$$\begin{cases} Nuc p_i(A) = \mathcal{C}(v_1) \boxplus \mathcal{C}(v_2) \boxplus \cdots \boxplus \mathcal{C}(v_n), \\ grau p_i(t) = 2 = \dim \mathcal{C}(v_1) = \dim \mathcal{C}(v_2) = \cdots = \dim \mathcal{C}(v_n). \end{cases}$$

É possível escolher uma base ortonormal  $\beta_i \subset \mathcal{C}(v_i)$ , tal que a restrição do operador tem como representação a matriz

$$[A_i]_{\beta_i} = \begin{bmatrix} \lambda_i & -\tau_i \\ \tau_i & \lambda_i \end{bmatrix}.$$

Entretanto essa matriz é antisimétrica, implicando que  $\lambda_i = 0$ . Logo o fator é da forma  $p_i(t) = t^2 + \tau_i^2$  e a representação na base  $\beta_i$  do operador restrito ao

subespaço  $\mathcal{C}(v_i)$  é a matriz

$$R(0, \tau_i) = \tau_i \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Observamos que as raízes complexas desse polinômio são imaginárias.

**Exercício 8.6.1** Prove que se todas as raízes do polinômio característico de um operador normal num espaço Euclidiano são imaginárias, então o operador é anti-simétrico.  $\square$

### Exercícios propostos 8.6.1

1. Mostre que toda matriz antisimétrica  $N \in M(2, \mathbb{R})$  é da forma

$$N = r \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{onde } r \in \mathbb{R}.$$

2. Fixado um vetor unitário  $v_0 \in \mathbb{R}^3$ , defina a aplicação  $\mathcal{V} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $\mathcal{V}(v) = v_0 \wedge v$ , onde  $\wedge$  indica o produto vetorial dos vetores  $v_0$  e  $v$ .
  - a) Mostre que  $\mathcal{V}$  é um operador anti-simétrico.
  - b) Encontre a matriz de  $\mathcal{V}$  na base canônica.
  - c) Explícite a transposta de  $\mathcal{V}$  em termos de  $v_0$ .
3. Qual o traço de um operador anti-simétrico?
4. Prove: um operador  $A$  em  $(V, \langle \cdot, \cdot \rangle)$  é anti-simétrico  $\Leftrightarrow \langle v, A(v) \rangle = 0$  para todo  $v \in V$ .

## 8.7 Operadores ortogonais

Um operador  $U$  num espaço Euclidiano  $V$  é dito *ortogonal* se  $U \circ U^t = Id = U^t \circ U$ .

Da definição segue que operadores ortogonais são invertíveis. Por isso, para verificar quando um operador é ortogonal só precisamos verificar uma das duas condições,  $U \circ U^t = Id$  ou  $U^t \circ U = Id$ .

**Exercício 8.7.1** É possível identificar matricialmente um operador ortogonal. Um operador  $U$  num espaço Euclidiano  $V$  é ortogonal  $\Leftrightarrow$  a representação  $[U]_\beta$  é uma matriz ortogonal para qualquer base ortonormal  $\beta \subset V$ .  $\square$

Examinemos em poucas linhas algumas propriedades de matrizes ortogonais. Um dos exemplos importantes para o cálculo matricial é a matriz mudança de coordenadas entre bases ortonormais.

**Proposição 8.7.1** *Se  $\beta$  e  $\gamma$  são duas bases ortonormais no espaço Euclidiano  $V$ , então a matriz mudança de coordenadas  $[Id]_{\gamma}^{\beta}$  é uma matriz ortogonal.*

**Demonstração** Digamos que  $\beta = \{v_1, v_2, \dots, v_n\}$  e  $\gamma = \{w_1, w_2, \dots, w_n\}$ . Para aliviar a notação denotaremos por  $Q$  a matriz mudança de coordenadas  $[Id]_{\gamma}^{\beta}$ . Por definição de representação matricial, a  $i$ -ésima coluna de  $Q$  é formada pelos coeficientes da combinação linear  $v_i = a_{1i}w_1 + a_{2i}w_2 + \dots + a_{ni}w_n$  enquanto a  $j$ -ésima linha de  $Q^t$  é formada pelos coeficientes da combinação linear  $v_j = a_{1j}w_1 + a_{2j}w_2 + \dots + a_{nj}w_n$ . Reescrevamos o produto interno  $\delta_{ij} = \langle v_i, v_j \rangle$  levando em conta a ortonormalidade das bases,

$$\delta_{ij} = \sum_{k,m=1}^n a_{ki}a_{mj} \langle w_i, w_j \rangle = \sum_{k=1}^n a_{ki}a_{kj} = (Q^t Q)_{ii}$$

(recordamos que  $\delta_{ij}$  é o delta de Kronecker). Portanto,  $Q^t Q = I_n$ . Apenas essa condição é suficiente para mostrar a ortogonalidade da matriz mudança de coordenadas  $Q$ .  $\square$

As linhas de uma matriz ortogonal  $n \times n$ ,  $Q = [a_{ij}]$ , admitem uma leitura bastante geométrica. Se a  $i$ -ésima linha de  $Q$  é identificada com as coordenadas de um vetor  $v_i \in \mathbb{R}^n$  na base canônica, mais precisamente,  $Q_i = [v_i]_{\alpha}$ , então a condição  $Q^t Q = I_n$  implica que o conjunto ordenado  $\beta = \{v_1, v_2, \dots, v_n\}$  é uma base ortonormal do  $\mathbb{R}^n$  relativa ao produto interno canônico. O mesmo ocorre quando a identificação é com a  $j$ -ésima coluna de  $Q$ .

Feitas essas considerações matriciais, voltemos ao estudo de operadores ortogonais construindo uma representação matricial apropriada. Recordamos que  $R_{\theta}$  é a matriz de rotação por  $\theta$ .

**Proposição 8.7.2** *Se  $m_A(t) = (t - \lambda)^2 + \tau^2$ , com  $\tau > 0$ , é a fatoração primária do polinômio minimal de um operador normal  $A$  num espaço Euclidiano  $V$ . Então*

- a) o operador é ortogonal  $\Leftrightarrow$
- b)  $m_A(t) = (t - \lambda)^2 + \tau^2$  com  $\lambda^2 + \tau^2 = 1 \Leftrightarrow$
- c) existe uma base ortonormal  $\beta$  na qual a representação matricial é da forma  $[A]_{\beta} = \text{diag}\{R_{\theta}, R_{\theta}, \dots, R_{\theta}\}$ .

**Demonstração** Pelo Teorema da decomposição cíclica normal existem  $v_1, v_2, \dots, v_n \in V = \text{Nuc } m_A(A)$  tais que

$$\begin{cases} \text{Nuc } m_A(A) = \mathcal{C}(v_1) \boxplus \mathcal{C}(v_2) \boxplus \cdots \boxplus \mathcal{C}(v_n), \\ \text{grau } m_A(t) = 2 = \dim \mathcal{C}(v_1) = \dim \mathcal{C}(v_2) = \cdots = \dim \mathcal{C}(v_n). \end{cases}$$

Sendo  $A$  um operador normal, é possível escolher uma base ortonormal  $\beta_i \subset \mathcal{C}(v_i)$  na qual a restrição do operador a cada parcela cíclica tem como representação

$$[A_i]_{\beta_i} = \begin{bmatrix} \lambda & -\tau \\ \tau & \lambda \end{bmatrix}.$$

Portanto, se considerarmos a base ortonormal  $\beta = \vec{\cup} \beta_i$ , obtemos a representação

$$[A]_{\beta} = \text{diag}\{[A_1]_{\beta_1}, [A_2]_{\beta_2}, \dots, [A_n]_{\beta_n}\}.$$

Por tudo que já foi visto sobre representações em bases ortonormais podemos escrever

$$[A \circ A^t]_{\beta} = \text{diag}\{[A_1]_{\beta_1}[A_1]_{\beta_1}^t, [A_2]_{\beta_2}[A_2]_{\beta_2}^t, \dots, [A_n]_{\beta_n}[A_n]_{\beta_n}^t\}.$$

$a) \Leftrightarrow b)$  A condição  $A \circ A^t = Id$  é equivalente à condição  $[A_i]_{\beta_i}[A_i]_{\beta_i}^t = I_2$  para todo  $i$  que por sua vez é equivalente à condição  $\lambda^2 + \tau^2 = 1$ .

$b) \Leftrightarrow c)$  Deixaremos como exercício. □

**Exercício 8.7.2** Se  $m_A(t) = (t - \lambda)$  é a fatoração primária do polinômio minimal de um operador normal  $A$  num espaço Euclidiano  $V$  então  $A$  é ortogonal  $\Leftrightarrow |\lambda| = 1$ . Prove essa afirmação. □

### Exercícios propostos 8.7.1

1. Demonstre que a multiplicação de matrizes induz uma estrutura de grupo no conjunto das matrizes ortogonais  $O(n, \mathbb{R})$ .
2. Descreva todas as matrizes ortogonais  $Q \in M(2, \mathbb{R})$ .
3. Mostre que se um operador ortogonal em  $\mathbb{R}^2$  fixa um vetor não nulo então o operador é uma involução ortogonal.
4. Verifique que o operador linear  $A$  é normal e determine a decomposição normal quando  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = \frac{1}{3}(2x + 2y + z, -2x + y + 2z, x - 2y + 2z)$ . Dê uma representação matricial canônica.
5. No espaço das matrizes  $M(n, \mathbb{R})$ , considere o produto interno dado por  $\langle N, P \rangle = \text{tr}(NP^t)$ . Fixado  $Q_0$ , uma matriz  $n \times n$  com entradas reais, prove que o operador linear  $U : M(n, \mathbb{R}) \rightarrow M(n, \mathbb{R})$ ,  $U(N) = Q_0N$  é um operador ortogonal  $\Leftrightarrow Q_0$  é uma matriz ortogonal.
6. Suponha que seja dada uma decomposição em soma direta  $V = W \boxplus W^\perp$  de um espaço Euclidiano  $V$ . Seja  $v = u + w$  a decomposição correspondente do vetor. Demonstre que a aplicação  $U : V \rightarrow V$ ,  $U(v) = u - w$ , é ortogonal. Dê uma representação matricial canônica para  $U$ .

7. Determine os operadores lineares de um espaço Euclidiano que comutam com todos operadores ortogonais.
8. Mostre que as matrizes  $N$  e  $P$  são conjugadas, onde
- $$N = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & 1 \\ -1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \quad \text{e} \quad P = \begin{bmatrix} \cos \frac{\pi}{4} & -\operatorname{sen} \frac{\pi}{4} & & \\ \operatorname{sen} \frac{\pi}{4} & \cos \frac{\pi}{4} & & \\ & & \cos \frac{\pi}{4} & -\operatorname{sen} \frac{\pi}{4} \\ & & \operatorname{sen} \frac{\pi}{4} & \cos \frac{\pi}{4} \end{bmatrix}.$$
9. Denote por  $W$  um subespaço unidimensional gerado pelo vetor unitário  $u$  de um espaço Euclidiano  $V$ . Considere a decomposição em soma direta  $V = W^\perp \boxplus W$  e a aplicação  $A(v) = v - 2\langle v, u \rangle u$ . Demonstre as afirmações.
- $A$  é uma involução.
  - $A$  é uma reflexão ortogonal em torno do subespaço  $W^\perp$ .
  - $A$  é ortogonal.
10. Sejam  $A$  e  $B$  dois operadores lineares em  $(V, \langle \cdot, \cdot \rangle)$  que comutam. Assuma que  $A$  é simétrico e que  $B$  é anti-simétrico. Prove as afirmações.
- $\langle A(v), B(v) \rangle = 0$  para todo  $v \in V$ .
  - $\|(A(v) + B(v))\| = \|(A(v) - B(v))\|$  para todo  $v \in V$ .
  - $C = (A + B) \circ (A - B)$  é um operador ortogonal.
11. Considere o operador linear  $A(x, y) = (ax + by, cx + dy)$  em  $\mathbb{R}^2$ . Determine condições necessárias e suficientes sobre  $a, b, c$  e  $d$  para que  $A$  seja um operador ortogonal em relação ao produto interno  $\langle (x_1, y_1), (x_2, y_2) \rangle = 2x_1x_2 + 3y_1y_2$ .

## 8.8 Espectro

O *espectro*  $\sigma(A)$  de um operador  $A$  num espaço vetorial  $V$  de dimensão finita sobre  $K$  ( $\mathbb{R}$  ou  $\mathbb{C}$ ) é o conjunto das raízes complexas do polinômio característico de  $A$ . Contando as repetições, o polinômio característico possui um número de raízes igual à  $\dim V$ .

O objetivo dessa seção é fazer um resumo do estudo feito sobre operadores normais clássicos examinando a distribuição do espectro no plano complexo. Na sequência,  $A$  denotará um operador normal no espaço Euclidiano  $V$  e  $m_A(t) = p_1(t)p_2(t) \cdots p_k(t)$  denotará a fatoração primária do seu polinômio mínimo, recordamos que os fatores são primos e sem repetições. O Teorema da decomposição normal garante que

$$V = \operatorname{Nuc} p_1(A) \boxplus \operatorname{Nuc} p_2(A) \boxplus \cdots \boxplus \operatorname{Nuc} p_k(A).$$

e os teoremas de decomposições cíclicas descreveram processos para construirmos uma base ortonormal especial  $\beta_i$  para a  $i$ -ésima parcela da decomposição normal. A saber. O operador  $A_i$  induzido por restrição de  $A$  tem uma das duas representações,

$$[A_i]_{\beta_i} = \underbrace{\text{diag}\{\lambda_i, \lambda_i, \dots, \lambda_i\}}_{n_i \text{ vezes}} \quad \text{se} \quad p_i(t) = (t - \lambda_i),$$

ou

$$[A_i]_{\beta_i} = \underbrace{\text{diag}\{R(\lambda_i, \tau_i), R(\lambda_i, \tau_i), \dots, R(\lambda_i, \tau_i)\}}_{n_i \text{ vezes}} \quad \text{se} \quad p_i(t) = (t - \lambda_i)^2 + \tau_i.$$

A representação canônica do operador normal  $A$  é a representação na base ortonormal  $\beta = \vec{\cup} \beta_i$ ,

$$[A]_{\beta} = \text{diag}\{[A_1]_{\beta_1}, [A_2]_{\beta_2}, \dots, [A_k]_{\beta_k}\}.$$

Essa representação permite afirmar que o espectro de  $A$  é a união dos espectros dos induzidos  $A_i$ . Com efeito. Segue da representação matricial construída que

$$p_A(t) = p_{A_1}(t)p_{A_2}(t) \cdots p_{A_k}(t).$$

em que cada fator é um polinômio linear  $p_{A_i}(t) = (t - \lambda_i)$  ou um polinômio primo de grau dois  $p_{A_i}(t) = (t - \lambda_i)^2 + \tau_i^2$ . O estudo espectral feito anteriormente fica resumido na

**Proposição 8.8.1** *Seja  $A$  um operador normal num espaço Euclidiano  $V$ . Então*

- a)  $A$  é simétrico  $\Leftrightarrow$  o espectro  $\sigma(A)$  está contido no eixo real;
- b)  $A$  é anti-simétrico  $\Leftrightarrow$  o espectro  $\sigma(A)$  está contido no eixo imaginário;
- c)  $A$  é ortogonal  $\Leftrightarrow$  o espectro  $\sigma(A)$  está contido no círculo unitário.

### Exercícios propostos 8.8.1

1. Se  $\lambda$  não pertence ao espectro de um operador linear  $A : V \rightarrow V$  onde  $V$  é um espaço vetorial de dimensão finita, então  $A + \lambda Id$  é invertível.
2. Se  $A$  é um operador normal num espaço Euclidiano  $V$  satisfazendo a identidade  $A^n \equiv Id$  para algum inteiro  $n \geq 1$ , então  $A$  é um operador ortogonal.
3. Dê um exemplo de um operador linear  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  tal que  $A^3 \equiv Id$ .
4. Dê a representação matricial canônica para um operador normal  $A : V \rightarrow V$  se o seu polinômio minimal é  $m_A(t) = t^4 - 1$ .

5. Se  $\lambda$  pertence ao espectro de um operador ortogonal então  $\frac{1}{\lambda}$  também pertence ao espectro. Examine essa afirmação.
6. Suponha que  $-1$  não está no espectro do operador  $A : V \rightarrow V$ . Prove as afirmações
  - a)  $Id + A$  é um operador invertível.
  - b) Se  $A$  é ortogonal então  $B = (Id - A) \circ (Id + A)^{-1}$  é antisimétrica.
  - c) Se  $A$  é antisimétrica então  $B = (Id - A) \circ (Id + A)^{-1}$  é ortogonal.

## 8.9 Operadores que preservam a estrutura

Diz-se que um operador  $U$  num espaço Euclidiano  $(V, \langle \cdot, \cdot \rangle)$  *preserva o produto interno* quando  $\langle U(v), U(w) \rangle = \langle v, w \rangle$  para quaisquer  $v, w \in V$ .

Tais aplicações também são chamadas de *isometrias lineares*. Na seção final desse capítulo examinaremos o conceito de isometria com mais vagar.

**Lema 8.9.1** *Toda aplicação num espaço Euclidiano que preserva o produto interno é um operador linear invertível.*

**Demonstração** Seja  $U$  uma aplicação em  $(V, \langle \cdot, \cdot \rangle)$  que preserva o produto interno. Fixemos uma base ortonormal qualquer  $\beta = \{v_1, v_2, \dots, v_n\} \subset V$ . Segue diretamente da hipótese as igualdades,

$$\delta_{ij} = \langle U(v_i), U(v_j) \rangle = \langle v_i, v_j \rangle.$$

Logo, o conjunto  $U(\beta) = \{U(v_1), U(v_2), \dots, U(v_n)\}$  é formado por  $n$  vetores unitários dois a dois ortogonais. Portanto  $U(\beta)$  é uma base ortonormal de  $V$ . Sendo assim, a aplicação  $U : V \rightarrow V$  pode ser escrita como

$$U(v) = \langle U(v), U(v_1) \rangle U(v_1) + \langle U(v), U(v_2) \rangle U(v_2) + \dots + \langle U(v), U(v_n) \rangle U(v_n).$$

Note que cada aplicação  $a_i : V \rightarrow \mathbb{R}$  definida por  $a_i(v) = \langle U(v), U(v_i) \rangle$  é um funcional linear pois, como  $U$  preserva o produto interno concluímos que  $a_i(v) = \langle v, v_i \rangle$ . Agora é imediato verificar que  $U$  é um operador linear.

Agora, se  $v \in Nuc U$  então

$$0 = \langle U(v), U(v) \rangle = \langle v, v \rangle = \|v\|^2.$$

Isso mostra que o núcleo é trivial, portanto,  $U$  é um operador linear invertível.  $\square$

Existem várias propriedades que caracterizam um operador que preserva a estrutura Euclidiana, entre as quais, a ortogonalidade é uma delas. Listemos algumas numa única proposição.

**Proposição 8.9.1** *As seguintes afirmações sobre um operador linear  $U$  num espaço Euclidiano  $V$  são equivalentes.*

- a) *O operador preserva o produto interno;*
- b) *O operador preserva a norma;*
- c) *O operador é ortogonal;*
- d) *O operador transforma base ortonormal em base ortonormal.*

**Demonstração** a)  $\Rightarrow$  b) Se  $U$  preserva o produto interno temos que

$$\|U(v)\|^2 = \langle U(v), U(v) \rangle = \langle v, v \rangle = \|v\|^2.$$

Isso mostra que  $U$  preserva a norma induzida pelo produto interno.

b)  $\Rightarrow$  c) Se  $u, v \in V$  são dois vetores quaisquer, calculemos o produto interno  $\langle u, U^t \circ U(v) \rangle$  pela identidade de polarização,

$$\langle u, U^t \circ U(v) \rangle = \langle U(u), U(v) \rangle = \frac{1}{4} \|U(u) + U(v)\|^2 - \frac{1}{4} \|U(u) - U(v)\|^2.$$

Por hipótese,  $U$  é um operador linear que preserva a norma, então

$$\begin{aligned} \langle u, U^t \circ U(v) \rangle &= \frac{1}{4} \|U(u+v)\|^2 - \frac{1}{4} \|U(u-v)\|^2 \\ &= \frac{1}{4} \|u+v\|^2 - \frac{1}{4} \|u-v\|^2 \\ &= \langle u, v \rangle. \end{aligned}$$

Como  $\langle u, U^t \circ U(v) \rangle = \langle u, v \rangle$  para quaisquer  $u, v \in V$  então  $U^t \circ U = Id$ , significando que  $U$  é ortogonal.

c)  $\Rightarrow$  d) Suponha que  $U$  é um operador ortogonal e que  $\beta = \{v_1, v_2, \dots, v_n\}$  é uma base ortonormal do espaço Euclidiano  $V$ . Calculemos,

$$\langle U(v_i), U(v_j) \rangle = \langle v_i, U^t \circ U(v_j) \rangle = \langle v_i, v_j \rangle = \delta_{ij}.$$

Esse cálculo mostra que  $U(\beta) = \{U(v_1), U(v_2), \dots, U(v_n)\}$  é um conjunto de  $n$  vetores unitários dois a dois ortogonais num espaço de dimensão  $n$ . Como sabemos, sob essas condições  $U(\beta)$  é uma base ortonormal de  $V$ .

d)  $\Rightarrow$  a) Suponhamos que  $U$  aplica base ortonormal em base ortonormal. Sejam  $\beta = \{v_1, v_2, \dots, v_n\}$  uma base ortonormal de  $V$ ,  $v = a_1v_1 + a_2v_2 + \dots + a_nv_n$  e  $w = b_1v_1 + b_2v_2 + \dots + b_nv_n$ . Considerando a ortonormalidade de  $\beta$  e  $U(\beta)$  obtemos, por um lado, a igualdade

$$\langle U(v), U(w) \rangle = \sum_{i,j=1}^n a_i b_j \langle U(v_i), U(v_j) \rangle = \sum_{i=1}^n a_i b_i,$$

e, por outro lado,

$$\langle v, w \rangle = \sum_{i,j=1}^n a_i b_j \langle v_i, v_j \rangle = \sum_{i,j=1}^n a_i b_i.$$

Desses cálculos concluímos que  $\langle U(v), U(w) \rangle = \langle v, w \rangle$ .  $\square$

### Exercícios propostos 8.9.1

1. Demonstre que um operador linear num espaço Euclidiano que transforma vetores unitários em vetores unitários é ortogonal.
2. Uma aplicação  $A$  num espaço Euclidiano  $V$  que preserva a norma é um operador linear, portanto é ortogonal. Prove essa afirmação.
3. Um operador linear  $A$  em  $(V, \langle \cdot, \cdot \rangle)$  é dito *coforme* se preserva o ângulo entre dois vetores. Mostre que um operador conforme é da forma  $A \equiv rU$  em que  $U$  é ortogonal e  $r \neq 0$ .
4. Um operador  $A$  em  $(V, \langle \cdot, \cdot \rangle)$  é *isogonal* se preserva ortogonalidade, isto é, se  $\langle v, w \rangle = 0 \Rightarrow \langle A(v), A(w) \rangle = 0$ . Prove que todo operador isogonal fatora-se num composição de uma homotetia com um operador ortogonal,  $A \equiv rU$  com  $r \neq 0$ .
5. Demonstre as recíprocas das afirmações feitas nos itens acima
6. Uma aplicação  $A$  num espaço Euclidiano  $V$  que transforma vetores unitários em vetores unitários é um operador ortogonal?
7. Dado um operador invertível  $A$  num espaço Euclidiano  $(V, \langle \cdot, \cdot \rangle)$  mostre que existe uma base ortonormal  $\beta = \{v_1, v_2, \dots, v_n\}$  tal que  $A(\beta) = \{A(v_1), A(v_2), \dots, A(v_n)\}$  é uma base ortogonal.

## 8.10 Comparando estruturas Euclidianas

Surpreendentemente, existe uma estreita relação entre dois produtos internos no mesmo espaço vetorial  $V$ . Conhecendo-se um deles podemos descrever facilmente todos os outros. Para explicar com precisão essa afirmação necessitamos de um novo conceito.

Diz-se que  $P : V \rightarrow V$  é um *operador positivo* num espaço Euclidiano  $V$  se

$p_1$   $P$  é simétrico;

$p_2$   $\langle v, P(v) \rangle > 0$  para qualquer vetor não nulo  $v \in V$ .

Estamos interessados na recíproca do seguinte resultado.

**Exercício 8.10.1** Assuma que  $P$  é um operador positivo em  $(V, \langle \cdot, \cdot \rangle)$ . Então  $P$  é invertível e a aplicação  $\langle \langle \cdot, \cdot \rangle \rangle : V \times V \rightarrow V$  definida por  $\langle \langle v, w \rangle \rangle = \langle v, P(w) \rangle$  é um outro produto interno em  $V$ .  $\square$

Antes de demonstrar que dois produtos internos num mesmo espaço vetorial estão relacionados por um operador positivo como indicado no exercício acima, vejamos, entre outros fatos, que operadores positivos existem e que é simples construí-los. Relembramos que o Teorema espectral garante que todo autovalor de um operador positivo é real.

**Proposição 8.10.1** *As seguintes afirmações sobre um operador  $P$  num espaço Euclidiano  $V$  são equivalentes.*

- a) *O operador é positivo.*
- b) *O operador é simétrico e os autovalores são positivos.*
- c)  *$P$  é um operador de Gram associado a um operador invertível  $A : V \rightarrow V$ .*

**Demonstração** a)  $\Rightarrow$  b) Seja  $\lambda \in \mathbb{R}$  um autovalor do operador positivo  $P : V \rightarrow V$ . Escolhamos um autovetor  $v$  associado a esse autovalor e calculemos,

$$0 < \langle v, P(v) \rangle = \langle v, \lambda v \rangle = \lambda \|v\|^2.$$

Como  $\|v\| > 0$ , concluímos que o autovalor é positivo.

b)  $\Rightarrow$  c) Assumamos que todos os autovalores  $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$  (contando as repetições) do operador simétrico  $P$  são positivos. Pelo Teorema espectral podemos escolher uma base ortonormal de autovetores  $\beta = \{v_1, v_1, \dots, v_n\}$  que diagonaliza o operador. Seja  $A : V \rightarrow V$  o único operador linear cuja representação na base  $\beta$  é a matriz

$$[A]_\beta = \text{diag}\{\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n}\}.$$

O operador  $A$  é invertível pois seu determinante é positivo e  $[A^t]_\beta = [A]_\beta^t$  desde que a base é ortonormal. Portanto,

$$[A^t \circ A]_\beta = [A^t]_\beta [A]_\beta = [A]_\beta^t [A]_\beta = [P]_\beta.$$

Pela unicidade da representação matricial obtemos que  $P = A^t \circ A$ .

c)  $\Rightarrow$  a) Suponhamos que  $P = A^t \circ A$  em que  $A$  é um operador invertível. É claro que  $P$  é simétrico. Para mostrar que  $P$  é positivo escolhamos qualquer vetor não nulo  $v \in V$  e calculemos

$$\langle v, P(v) \rangle = \langle v, A^t \circ A(v) \rangle = \langle A(v), A(v) \rangle = \|A(v)\|^2.$$

Como  $A$  é invertível então  $\|A(v)\| > 0$ . Isso mostra que  $\langle v, P(v) \rangle > 0$  para qualquer vetor não nulo  $v \in V$ .  $\square$

**Proposição 8.10.2** *Seja  $(V, \langle \cdot, \cdot \rangle)$  um espaço Euclidiano. Suponha que  $\langle \cdot, \cdot \rangle$  seja um outro produto interno em  $V$ . Então existe um único operador positivo  $P : V \rightarrow V$  tal que  $\langle \langle v, w \rangle \rangle = \langle v, P(w) \rangle$ .*

**Demonstração** Para definir uma aplicação  $P : V \rightarrow V$  utilizaremos o Teorema da representação de um funcional linear. Dado um vetor  $w \in V$  considere o funcional linear  $f : V \rightarrow \mathbb{R}$ ,  $f_w(v) = \langle \langle v, w \rangle \rangle$ . Considere o único vetor  $w' \in V$  representando esse funcional linear com relação ao produto interno  $\langle \cdot, \cdot \rangle$ , em outras palavras,

$$f_w(v) = \langle \langle v, w \rangle \rangle = \langle v, w' \rangle.$$

Defina a aplicação  $P : V \rightarrow V$  por  $P(w) = w'$ . Observe que  $P$  é uma aplicação "simétrica", pois

$$\langle v, P(w) \rangle = \langle \langle v, w \rangle \rangle = \langle \langle w, v \rangle \rangle = \langle w, P(v) \rangle = \langle P(v), w \rangle.$$

Portanto, se  $P$  for um operador linear, ele é simétrico. Mostremos a linearidade de  $P$ . Dados  $u, v, w \in V$  e  $\lambda \in \mathbb{R}$ , calculemos,

$$\langle u, P(v + \lambda w) \rangle = \langle P(u), v + \lambda w \rangle = \langle P(u), v \rangle + \lambda \langle P(u), w \rangle.$$

Novamente, pela simetria da aplicação  $P$  seguem as igualdades,

$$\langle u, P(v + \lambda w) \rangle = \langle u, P(v) \rangle + \lambda \langle u, P(w) \rangle = \langle u, P(v) + \lambda P(w) \rangle.$$

Como vale a igualdade  $\langle u, P(v + \lambda w) \rangle = \langle u, P(v) + \lambda P(w) \rangle$  para qualquer  $u \in V$  então  $P(v + \lambda w) = P(v) + \lambda P(w)$ . Feito isso, resta mostrar a segunda condição de positividade. Mas isso também é simples, pois se  $v \in V$  é um vetor não nulo, segue da definição de produto interno que  $\langle v, P(v) \rangle = \langle \langle v, v \rangle \rangle > 0$ . Isso termina a demonstração da proposição.  $\square$

Uma matriz  $N \in M(n, \mathbb{R})$  é *positiva* quando  $N = [a_{ij}]$  é simétrica e

$$\sum_{i,j=1}^n a_{ij} x_i x_j > 0,$$

para toda sequência de reais  $(x_1, x_2, \dots, x_n)$  não todos nulos. Em outras palavras,  $N$  é positiva se for uma representação matricial na base canônica de um operador positivo  $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$  onde estamos considerando em  $\mathbb{R}^n$  o produto interno canônico.

### Exercícios propostos 8.10.1

1. Seja  $A$  um operador num espaço Euclidiano  $V$ . Demonstre as afirmações.
  - a)  $Nuc P = Nuc A$ , onde  $P$  é o operador de Gram associado à  $A$ .
  - b) Se  $A$  é simétrico, existe um escalar  $c > 0$  tal que  $A + cId$  é positivo.
2. Dado um operador positivo  $P$  no espaço Euclidiano  $V$  mostre que existem infinitos operadores  $A : V \rightarrow V$  tais que  $P = A^t \circ A$ .
3. O conjunto  $\mathcal{P} \subset \mathcal{L}(V, V)$  formado por todos operadores lineares positivos no espaço Euclidiano  $V$  é convexo, isto é, dados dois operadores  $P_1, P_2 \in \mathcal{P}$  o segmento  $sP_1 + tP_2$  está contido em  $\mathcal{P}$ , quando  $s + t = 1$ ,  $s \geq 1$  e  $t \geq 1$ .
4. Prove as seguintes afirmações sobre um operador positivo num espaço Euclidiano  $(V, \langle \cdot, \cdot \rangle)$ . As bases consideradas são ortonormais.
  - a) Cada entrada da diagonal da representação  $[P]_\beta$  é positiva.
  - b) A matriz  $[P]_\beta$  é simétrica com todas as entradas positivas e  $\det [P]_\beta > 0$ .
  - c) Existe um operador  $A : V \rightarrow V$  tal que  $P = A^2$ .
5. (Problema de autovalores generalizados) Sejam  $A$  um operador simétrico e  $P$  um operador positivo no espaço Euclidiano  $V$ . Demonstre que existe uma base ordenada  $\beta = \{v_1, v_2, \dots, v_n\} \subset V$  tal que para todo  $v_i \in \beta$  existe  $\lambda \in \mathbb{R}$  com  $A(v_i) = \lambda P(v_i)$ .

## 8.11 Decomposição polar

Como uma bela aplicação da teoria já apresentada demonstraremos a existência de uma decomposição para operadores lineares equivalente à decomposição polar para números complexos. Para prová-la ampliaremos as idéias apresentadas na seção anterior.

Um operador  $P$  num espaço Euclidiano  $V$  é *não negativo* se  $P$  é simétrico e  $\langle v, P(v) \rangle \geq 0$  para todo vetor não nulo  $v \in V$ . Registraremos uma proposição cuja demonstração em nada difere da sua similar para operadores positivos.

**Proposição 8.11.1** *As seguintes afirmações sobre um operador  $P$  num espaço Euclidiano  $V$  são equivalentes.*

- a) *O operador é não negativo.*
- b) *O operador é simétrico e os autovalores são não negativos.*
- c)  *$P$  é um operador de Gram associado a um operador  $A : V \rightarrow V$ .*

**Exercício 8.11.1** Prove que para cada operador  $A$  em um espaço Euclidiano  $V$  existe um único operador não negativo  $P$  tal que  $P^2 = A^t \circ A$ . Denotaremos esse operador de Gram por  $P = \sqrt{A^t \circ A}$ .  $\square$

**Teorema 8.11.1 (Teorema da decomposição polar)** *Para cada operador  $A$  num espaço Euclidiano  $V$  existe um único operador não negativo  $P$  e um operador ortogonal  $U$  tal que  $A = U \circ P$ . Além disso, se  $A$  é invertível então  $P$  é positivo e  $U$  é unicamente determinado.*

**Demonstração** Seja  $P$  o único operador não negativo tal que  $P^2 = A^t \circ A$ . Antes de tudo, observamos que da simetria de  $P$  segue a soma direta ortogonal  $V = \text{Im } P \boxplus \text{Nuc } P$ . Outro fato que deve ser notado é a coincidência dos núcleos  $\text{Nuc } P = \text{Nuc } A$  cuja demonstração é consequência das seguintes igualdades para todo  $v \in V$ ,

$$\|P(v)\|^2 = \langle P(v), P(v) \rangle = \langle P^2(v), v \rangle = \langle A^t \circ A(v), v \rangle = \langle A(v), A(v) \rangle = \|A(v)\|^2.$$

Então, pelo Teorema do núcleo e da imagem concluímos que  $\dim \text{Im } P = \dim \text{Im } A$  e, por conseguinte,  $\dim \text{Nuc } P = \dim \text{Im } A^\perp$ .

Defina uma aplicação  $U : \text{Im } P \boxplus \text{Nuc } P \rightarrow \text{Im } A \boxplus \text{Im } A^\perp$  do modo seguinte,

$$U(P(v) + u) = A(v) + U_0(u),$$

onde  $U_0 : \text{Nuc } P \rightarrow \text{Im } A^\perp$  é qualquer operador linear entre esses subespaços de dimensões iguais que preserve a norma induzida. A aplicação  $U$  está bem definida, o seu valor só depende de  $P(v)$  e não do vetor  $v$ . Com efeito,

$$P(v) = P(w) \Leftrightarrow v - w \in \text{Nuc } P = \text{Nuc } A \Leftrightarrow A(v) = A(w).$$

Portanto,  $U(P(v) + u) = U(P(w) + u)$ . Com um procedimento de rotina verifica-se que  $U$  é linear. Mais ainda,  $U$  é ortonormal pois preserva a norma. De fato. Recordando-se das ortogonalidades dos subespaços envolvidos e da definição de  $U_0$  temos que

$$\|U(P(v) + u)\|^2 = \|A(v) + U_0(u)\|^2 = \|A(v)\|^2 + \|U_0(u)\|^2.$$

Por outro lado,

$$\|P(v) + u\|^2 = \|P(v)\|^2 + \|u\|^2 = \|A(v)\|^2 + \|U_0(u)\|^2.$$

Pela construção é claro que se  $v \in V$  então  $U(P(v)) = U(P(v) + 0) = A(v)$ , como pretendíamos. Provemos a unicidade do operador não negativo. Suponha que tenhamos a decomposição  $U_0 \circ P_0 = A$ , onde  $U_0$  é ortogonal e  $P_0$  é não negativo. Então

$$A^t \circ A = P_0^t \circ U_0^t \circ U_0 \circ P_0 = P_0 \circ P_0 = P_0^2$$

Portanto, o operador simétrico na decomposição polar é o único operador não negativo satisfazendo a condição  $A^t \circ A = P^2$  como tínhamos escolhido inicialmente.

Se  $A$  é invertível então  $\text{Nuc } P = \text{Nuc } A = \{0\}$ . Logo,  $P$  é não negativo e invertível, condições equivalentes a ser positivo. Sendo assim, o operador ortonormal

$U$  é único pois necessariamente ele de ser  $U = A \circ P^{-1}$ .  $\square$

### Exercícios propostos 8.11.1

- Determine a decomposição polar dos operadores quando o produto interno considerado é o produto interno canônico.
  - $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $A(x, y) = (x - y, x + y)$ .
  - $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $A(x, y, z) = (y, z, 0)$ .
  - $A : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ ,  $A(x, y, z, w) = (x - w, y + z, -y + z, x + w)$ .
- Enuncie uma versão matricial para a decomposição polar de uma matriz quadrada.
- Dê a decomposição polar das matrizes.

$$\text{a) } N = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

$$\text{b) } N = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

- Se  $V = W \boxplus W^\perp$  é uma soma direta de um espaço Euclidiano  $V$  e  $\pi : V \rightarrow V$  é a projeção ao longo da primeira parcela sobre a segunda, prove as afirmações.
  - $\pi$  é não negativo.
  - Existe um operador  $A : V \rightarrow V$  tal que  $A^2 = Id + \pi$ .
- Mostre que um operador linear  $A$  num espaço Euclidiano  $V$  é normal  $\Leftrightarrow$  as parcelas da decomposição polar  $A = PU$  satisfazem a condição de comutatividade  $PU = UP$ .
- Um operador  $A$  num espaço Euclidiano  $V$  é *não expansivo* quando  $\|A(v)\| \leq \|v\|$  para qualquer  $v \in V$ . Prove que todo operador em  $V$  é uma composta de um não expansivo com um simétrico.

## 8.12 Isometrias

Num espaço Euclidiano podemos definir uma distância utilizando a norma induzida pelo produto interno. Estudaremos os movimentos rígidos (transformações do espaço) que preservam a distância entre pontos, também chamadas de isometrias. Veremos que tais aplicações admitem uma descrição bastante simples. Para apresentar a questão de modo preciso coloquemos uma definição.

Um espaço métrico é um par  $(\mathcal{S}, d)$  no qual  $\mathcal{S}$  é um conjunto e  $d : \mathcal{S} \times \mathcal{S} \rightarrow [0, +\infty)$  é uma função não negativa chamada de *distância*, ou *métrica*, satisfazendo três condições. Para quaisquer  $x, y, z \in \mathcal{S}$  valem:

$$\text{d.1 } d(x, y) = 0 \Leftrightarrow x = y; \quad (\text{não degenerada})$$

$$\text{d.2 } d(x, y) = d(y, x); \quad (\text{simétrica})$$

$$d.3 \quad d(x, y) \leq d(x, z) + d(z, y). \quad (\text{desigualdade triangular})$$

Nesse texto não estamos interessados em estudar tais espaços em toda sua generalidade, caso o leitor tenha interesse maior deve procurar uma apresentação sistemática em livros de Topologia Geral que tratam de espaços métricos. Estamos interessados apenas numa métrica particular de um espaço Euclidiano.

Num espaço Euclidiano  $(V, \langle \cdot, \cdot \rangle)$  a norma define uma métrica, a saber,

$$d : V \times V \rightarrow [0, +\infty), \quad d(u, v) = \|u - v\|.$$

É simples verificar que essa aplicação é uma métrica e somente essa métrica será considerada num dado espaço Euclidiano.

Uma *isometria*,  $F$ , ou *movimento rígido*, em  $V$  é uma aplicação que preserva a métrica, isto significa que  $F : V \rightarrow V$  é uma aplicação tal que  $d(F(u), F(v)) = d(u, v)$  para quaisquer vetores  $u, v \in V$ .

**Exemplo 8.12.1** 1) Um operador ortogonal  $U$  num espaço Euclidiano é uma isometria pois preserva a norma,

$$d(U(u), U(v)) = \|U(u) - U(v)\| = \|U(u - v)\| = \|u - v\| = d(u, v).$$

2) Uma translação por  $v_0 \in V$  é a aplicação  $T : V \rightarrow V$ ,  $T(u) = u + v_0$ . Translações são isometrias pois para quaisquer  $u, v \in V$  temos que

$$d(T(u), T(v)) = \|T(u) - T(v)\| = \|u + v_0 - v - v_0\| = \|u - v\| = d(u, v).$$

O nosso objetivo é demonstrar que qualquer isometria é uma composta de uma translação com um operador ortogonal.  $\square$

**Exercício 8.12.1** Prove as afirmações sobre translações em espaços Euclidianos.

1. Dados  $u_1, v_1 \in V$  existe uma única translação  $T$  tal que  $T(u_1) = v_1$ .
2. A composta de duas translações  $S$  e  $T$  é uma translação e  $S \circ T = T \circ S$ .
3. Se  $T$  é uma translação por  $v_0$  então  $T$  é invertível e  $T^{-1}$  é uma translação por  $-v_0$ .  $\square$

**Lema 8.12.1** *Seja  $U$  uma isometria num espaço Euclidiano  $V$ . Então  $U$  fixa o vetor nulo se, e somente se,  $U$  é um operador ortogonal.*

**Demonstração** Uma isometria  $U$  de  $(V, \langle \cdot, \cdot \rangle)$  que fixa o vetor nulo preserva a norma pois

$$\|v\| = d(v, 0) = d(U(v), U(0)) = \|U(v) - U(0)\| = \|U(v)\|.$$

Como sabemos,  $U$  deve ser um operador ortogonal. A recíproca já foi comentada anteriormente.  $\square$

Provemos o principal resultado dessa seção.

**Teorema 8.12.1 (Teorema da classificação das isometrias)** *Para cada isometria  $F$  num espaço Euclidiano  $V$  existem um único operador ortogonal  $U$  e uma única translação  $T$  tal que  $F = T \circ U$ .*

**Demonstração** Se  $T$  é a translação por  $F(0)$  então  $T^{-1}$  é uma translação por  $-F(0)$ . Logo a aplicação  $U = T^{-1} \circ F$  é uma isometria pois é uma composta de isometrias e preserva a origem. Com efeito,

$$U(0) = T^{-1} \circ F(0) = F(0) - F(0) = 0.$$

Portanto,  $U$  é um operador ortogonal e  $F = T \circ U$ . Provemos a unicidade dessa decomposição. Assuma que  $F = T_0 \circ U_0$  em que  $T_0$  é uma translação e  $U_0$  é um operador ortogonal. Então  $T \circ U = T_0 \circ U_0$  implicando que  $U = T^{-1} \circ T_0 \circ U_0$ . Desde que  $U$  e  $U_0$  são operadores lineares, eles fixam o vetor nulo, e avaliando essa composta  $u = 0$  obtemos que  $0 = T^{-1} \circ T_0(0)$ . Entretanto, a única translação que fixa pontos é a identidade. Sendo assim,  $Id = T^{-1} \circ T_0$ , de onde seguem as igualdades  $T \equiv T_0$  e  $U \equiv U_0$ , como queríamos provar.  $\square$

### Exercícios propostos 8.12.1

- Determine a decomposição polar quando a aplicação for uma isometria.
  - $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $F(x, y, z) = (x + 1, y - 3, z)$ .
  - $G : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $G(x, y, z) = (2x + 1, y - 3, -5z)$ .
  - $H : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $H(x, y, z) = -(x, y, z)$ .
- Uma isometria num espaço Euclidiano  $V$  admite uma decomposição do tipo  $U_0 \circ T_0$ ?
- Prove que o conjunto das isometrias de um espaço Euclidiano é um grupo com a operação de composição de funções. Esse grupo é chamado de *grupo Euclidiano*.
- Um operador linear  $A$  num espaço Euclidiano  $V$  é uma *semelhança* quando
 
$$\|A(u) - A(v)\| = r\|u - v\|, \quad \text{para algum real } r > 0.$$
 Prove que uma semelhança é da forma  $A(v) = rU(v) + v_0$  onde  $U$  é ortogonal.
- Dados um vetor não nulo  $v_0$  em  $(V, \langle \cdot, \cdot \rangle)$  e um escalar não nulo  $\lambda \in \mathbb{R}$ , determine uma condição necessária e suficiente para que  $A : V \rightarrow V$ ,  $A(v) = v + \lambda\langle v, v_0 \rangle v_0$  seja uma isometria.

## Capítulo 9

# Espaços unitários

Um espaço unitário é um espaço vetorial complexo de dimensão finita equipado com um produto interno. A maioria dos resultados sobre espaços Euclidianos apresentados admite uma versão equivalente em espaços unitários e as demonstrações podem ser transladadas utilizando-se os mesmos argumentos, por isso, na ausência de uma demonstração, se achar necessário, o leitor pode refazê-la com a leitura da proposição equivalente para o caso Euclidiano. Veremos que o estudo de operadores normais em espaços unitários é bastante simplificado pelo fato do corpo dos números complexos ser algebricamente fechado.

### 9.1 Espaços unitários

Seja  $V$  um espaço vetorial complexo. Um *produto interno* em  $V$  é uma aplicação  $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$  satisfazendo as seguintes condições para quaisquer  $u, v, w \in V$  e  $\lambda \in \mathbb{C}$ :

1.  $\langle u, v \rangle = \overline{\langle v, u \rangle}$ ;
2.  $\langle u + \lambda v, w \rangle = \langle u, w \rangle + \lambda \langle v, w \rangle$ ;
3.  $\langle v, v \rangle > 0 \Leftrightarrow v \neq o$ .

Um produto interno num espaço vetorial complexo também é chamado de *produto Hermitiano positivo definido*. Seguem da definição as propriedades:

- a)  $\langle u, v + \lambda w \rangle = \langle u, v \rangle + \bar{\lambda} \langle u, w \rangle$ ;
- b)  $\langle v, v \rangle = 0 \Leftrightarrow v = o$ .

**Exemplo 9.1.1** 1) Entenderemos por *produto interno canônico* do  $\mathbb{C}^n$  a aplicação que cada par de vetores  $u = (x_1, x_2, \dots, x_n)$ ,  $v = (y_1, y_2, \dots, y_n) \in \mathbb{C}^n$  é associado ao número complexo

$$\langle u, v \rangle = x_1 \overline{y_1} + x_2 \overline{y_2} + \cdots + x_n \overline{y_n}.$$

A menos que seja dito explicitamente o contrário,  $\mathbb{C}^n$  estará sempre equipado com o produto interno canônico.

2) Todo espaço vetorial complexo de dimensão finita admite um produto interno. A construção utilizando as coordenadas de um vetor numa base ordenada do espaço é semelhante a aquela feita para espaços Euclidianos.

3) A aplicação  $\langle N, P \rangle = \text{tr}(NP^*)$  é um produto interno no espaço das matrizes  $M(m \times n, \mathbb{C})$ , onde a notação  $P^*$  está indicando a matriz adjunta de  $P = [a_{ij}]$ , isto é,  $P^* = [b_{ij}]$  onde  $b_{ij} = \overline{a_{ji}}$ .  $\square$

Da mesma forma, um produto interno  $\langle \cdot, \cdot \rangle$  em um espaço vetorial complexo  $V$  induz uma norma  $\| \cdot \| : V \rightarrow [0, +\infty)$ , ao definirmos  $\|v\| = \sqrt{\langle v, v \rangle}$ . Para demonstrar que a aplicação é de fato uma norma precisaremos da desigualdade de Schwarz. Com as hipóteses e notações acima temos a

**Proposição 9.1.1** *Para quaisquer  $u, v \in V$  vale a desigualdade  $\| \langle u, v \rangle \| \leq \|u\| \|v\|$ . Em consequência, a aplicação  $\| \cdot \|$  é uma norma, isto é,*

$$a) \|v\| = 0 \Leftrightarrow v = 0;$$

$$b) \|\lambda v\| = \|\lambda\| \|v\|;$$

$$c) \|u + v\| \leq \|u\| + \|v\|. \quad (\text{primeira desigualdade triangular})$$

**Demonstração** Provaremos apenas a desigualdade de Schwarz, ressaltando que a demonstração apresentada pode ser integralmente adaptada para espaços vetoriais reais. Dados dois escalares  $s, t \in \mathbb{C}$ , por definição de produto interno temos que

$$0 \leq \langle su - tv, su - tv \rangle = s\overline{s}\|u\|^2 - \overline{st}\langle u, v \rangle - s\overline{t}\langle u, v \rangle + t\overline{t}\|v\|^2.$$

Para os valores  $s = \|v\|^2$  e  $t = \langle u, v \rangle$  obtemos a desigualdade

$$0 \leq \|v\|^4 \|u\|^2 - \|v\|^2 \langle u, v \rangle \overline{\langle u, v \rangle} - \|v\|^2 \langle u, v \rangle \overline{\langle u, v \rangle} + \langle u, v \rangle \overline{\langle u, v \rangle} \|v\|^2.$$

Simplificando a expressão obtemos a desigualdade

$$0 \leq \|v\|^2 \left( \|v\|^2 \|u\|^2 - \langle u, v \rangle \overline{\langle u, v \rangle} \right).$$

Como  $\langle u, v \rangle \overline{\langle u, v \rangle} = |\langle u, v \rangle|^2$  podemos concluir a demonstração.  $\square$

**Exercício 9.1.1** Demonstre a *identidade de polarização* que relaciona o produto interno  $\langle \cdot, \cdot \rangle$  com a sua norma  $\| \cdot \|$  num espaço vetorial complexo  $V$ ,

$$\langle u, v \rangle = \frac{1}{4} \|u + v\|^2 - \frac{1}{4} \|u - v\|^2 + \frac{i}{4} \|u + iv\|^2 - \frac{i}{4} \|u - iv\|^2.$$

Sugestão: desenvolva o segundo membro da equação.  $\square$

As definições de ortogonalidade de vetores, complemento ortogonal de subespaços e bases ortonormais são semelhantes às definições correspondentes num espaço Euclidiano.

Um *espaço unitário* é um par  $(V, \langle \cdot, \cdot \rangle)$  no qual  $V$  é um espaço vetorial complexo de dimensão finita e  $\langle \cdot, \cdot \rangle$  é um produto interno de  $V$ . A existência de bases ortonormais num espaço unitário é provada aplicando o processo de ortogonalização de Gram-Schmidt que em nada difere do seu similar real. Utilizaremos também a notação

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$$

para indicar uma soma direta de subespaços dois a dois ortogonais. Finalmente, recordamos que as entradas da representação  $[A]_\beta = [a_{ij}]$  de um operador linear  $A : V \rightarrow V$  numa base ortonormal  $\beta = \{v_1, v_2, \dots, v_n\}$  é obtida pela mesma fórmula descrita no caso Euclidiano, ou seja,  $a_{ij} = \langle A(v_j), v_i \rangle$ .

### Exercícios propostos 9.1.1

1. Prove as afirmações sobre a norma  $\|\cdot\|$  de um espaço unitário  $V$ .
  - a)  $\|u + v\| \leq \|u\| + \|v\|$ . (primeira desigualdade triangular)
  - b)  $|\|u\| - \|v\|| \leq \|u - v\|$ . (segunda desigualdade triangular)
  - c)  $\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)$ . (lei do paralelogramo)
2. Mostre que  $\langle N, N^* \rangle = \text{tr}(NN^*)$  é um produto interno em  $M(n, \mathbb{C})$ . Dê uma base ortonormal para esse espaço unitário.

## 9.2 O operador adjunto

Para construir o operador adjunto precisaremos do

**Teorema 9.2.1 (Teorema da representação de um funcional linear)** *Seja  $(V, \langle \cdot, \cdot \rangle)$  um espaço unitário. A aplicação  $\Lambda : V \rightarrow V^*$ , que a cada vetor  $v \in V$  associa o funcional linear*

$$\Lambda[v] : V \rightarrow \mathbb{C} \quad \Lambda[v](w) = \langle w, v \rangle,$$

*é um isomorfismo quase linear. Em particular, cada funcional linear é representado por um vetor.*

**Demonstração** A quase linearidade de  $\Lambda : V \rightarrow V^*$  é verificada pelas igualdades a seguir. Fixados  $u, v \in V$  e  $\lambda \in \mathbb{C}$ , para qualquer vetor  $w \in V$  temos que

$$\Lambda[u + \lambda v](w) = \langle w, u + \lambda v \rangle = \langle w, u \rangle + \bar{\lambda} \langle w, v \rangle = (\Lambda[u] + \bar{\lambda} \Lambda[v])(w).$$

Provemos a sobrejetividade. Escolhida uma base ortonormal  $\beta = \{v_1, v_2, \dots, v_n\}$  do espaço unitário, para cada vetor  $w \in V$  existe uma única combinação linear tal que  $w = x_1 v_1 + x_2 v_2 + \dots + x_n v_n$ . Um funcional linear  $f : V \rightarrow \mathbb{C}$  avaliado em  $w$  assume o valor

$$f(w) = x_1 f(v_1) + x_2 f(v_2) + \dots + x_n f(v_n).$$

Daí podemos encontrar o vetor que representa  $f$ , o leitor pode verificar que  $v_f = \overline{f(v_1)}v_1 + \overline{f(v_2)}v_2 + \dots + \overline{f(v_n)}v_n$  é o representante procurado. Como  $\dim V = \dim V^*$ , concluímos que  $\Lambda : V \rightarrow V^*$  é um isomorfismo quase linear.  $\square$

Para dirimir confusões notacionais, vejamos novamente a definição de adjunta de um operador  $A : V \rightarrow V$ . A adjunta de  $A$  é a aplicação  $A^* : V^* \rightarrow V^*$ , onde  $A^*[f] : V \rightarrow \mathbb{C}$  é o funcional linear  $A^*[f](v) = f(A(v))$ . Como visto no Capítulo 3, fixada uma base  $\beta$  de  $V$  as representações de  $A$  e  $A^*$  estão relacionadas pela fórmula  $[A^*]_{\beta^*} = [A]_{\beta}^t$ , em que  $\beta^*$  é a base dual de  $\beta$ .

Passaremos a definir o *operador adjunto* de  $A : V \rightarrow V$ . Dado um funcional linear  $f : V \rightarrow \mathbb{C}$ , existe um único vetor  $v \in V$  que o representa, como também existe um único vetor  $v^* \in V$  que representa o funcional  $A^*[f]$ . Então para qualquer  $w \in V$  valem as relações,

$$A^*[f](w) = f(A(w)) = \langle A(w), v \rangle \quad \text{e} \quad A^*[f](w) = \langle w, v^* \rangle.$$

O *operador adjunto* de  $A : V \rightarrow V$  é a aplicação  $A^* : V \rightarrow V$  satisfazendo a identidade  $\langle A(w), v \rangle = \langle w, A^*(v) \rangle$  para quaisquer  $v, w \in V$ . Embora exista a coincidência na notação, nas próximas seções  $A^*$  significará sempre o operador adjunto. Observamos que com os mesmos procedimentos em espaços Euclidianos contruímos o operador transposto.

**Proposição 9.2.1** *Para cada operador  $A$  num espaço unitário  $V$  existe um único operador linear  $A^* : V \rightarrow V$ , chamado de operador adjunto de  $A$ , tal que*

$$\langle A(w), v \rangle = \langle w, A^*(v) \rangle,$$

*para quaisquer vetores  $v, w \in V$ . Mais ainda, a representação numa base ortonormal  $\beta$  de  $A^*$  é a matriz adjunta da representação de  $A$ , isto é,  $[A^*]_{\beta} = [A]_{\beta}^*$ .*

**Demonstração** Provemos apenas a última conclusão. As entradas das representações matriciais  $[A]_{\beta} = [a_{ij}]$  e  $[A^*]_{\beta} = [b_{ij}]$  na base ortonormal  $\beta = \{v_1, \dots, v_n\}$  são calculadas pelas fórmulas  $a_{ij} = \langle A(v_j), v_i \rangle$  e  $b_{ij} = \langle A^*(v_j), v_i \rangle$ . Então

$$b_{ij} = \langle A^*(v_j), v_i \rangle = \overline{\langle v_i, A^*(v_j) \rangle} = \overline{\langle A(v_i), v_j \rangle} = \overline{a_{ji}}. \quad \square$$

**Exercício 9.2.1** Prove as afirmações sobre representações matriciais de oper-

adores numa base ortonormal  $\beta$  de  $(V, \langle \cdot, \cdot \rangle)$ .

1. Se a representação de  $B$  é a matriz adjunta da representação de  $A$  então  $B$  é o operador adjunto de  $A$ .
2. Se  $A$  é invertível então o operador  $A^*$  também é invertível,  $(A^{-1})^* = (A^*)^{-1}$  e  $[A^{-1}]_{\beta}^* = [A^*]_{\beta}^{-1}$ .  $\square$

Para um espaço unitário  $V$ , a aplicação  $\Psi : \mathcal{L}(V, V) \rightarrow \mathcal{L}(V, V)$ ,  $\Psi(A) = A^*$ , é uma aplicação quase linear. Registremos esse e outros fatos sobre  $\Psi$ .

**Proposição 9.2.2** *Valem as seguintes propriedades para quaisquer operadores  $A$  e  $B$  no espaço unitário  $V$  e qualquer escalar  $\lambda \in \mathbb{C}$ .*

- a)  $(A + \lambda B)^* = A^* + \bar{\lambda}B^*$ .
- b)  $(A \circ B)^* = B^* \circ A^*$ .
- c)  $(A^*)^* = A$ .

**Demonstração** a) Se  $v, w \in V$  e  $\lambda \in \mathbb{C}$ , calculemos

$$\langle (A + \lambda B)(v), w \rangle = \langle A(v), w \rangle + \lambda \langle B(v), w \rangle = \langle v, A^*(w) \rangle + \lambda \langle v, B^*(w) \rangle.$$

Como o produto interno é quase linear na segunda variável, podemos escrever

$$\langle (A + \lambda B)(v), w \rangle = \langle v, (A^* + \bar{\lambda}B^*)(w) \rangle.$$

Entretanto,  $\langle (A + \lambda B)(v), w \rangle = \langle v, (A + \lambda B)^*(w) \rangle$ . Pela unicidade do operador adjunto concluímos a igualdade  $(A + \lambda B)^* = A^* + \bar{\lambda}B^*$ . Faça os itens b) e c).  $\square$

**Exercício 9.2.2** *Mostre as afirmações sobre um operador  $A$  num espaço  $(V, \langle \cdot, \cdot \rangle)$ .*

1. Se  $A$  preserva um subespaço  $W$  então o operador adjunto  $A^*$  preserva  $W^{\perp}$ .
2.  $V = \text{Im } A \oplus \text{Nuc } A^*$ .  $\square$

Para estabelecer uma relação entre o ideal anulador de um operador e de seu operador adjunto é conveniente fixar uma notação. Se  $p(t) = a_n t^n + \cdots + a_1 t + a_0$  é um polinômio com coeficientes complexos, denotaremos por  $\bar{p}(t)$  o polinômio conjugado, mais claramente,  $\bar{p}(t) = \bar{a}_n t^n + \cdots + \bar{a}_1 t + \bar{a}_0$ . Note que um escalar  $\lambda \in \mathbb{C}$  é uma raiz de  $p(t)$  se, e somente se,  $\bar{\lambda}$  é uma raiz de  $\bar{p}(t)$ .

**Exercício 9.2.3** *Mostre que para qualquer operador  $A$  no espaço unitário  $V$  o polinômio minimal de  $A^*$  é o conjugado complexo do polinômio minimal de  $A$ , isto é  $m_{A^*}(t) = \bar{m}_A(t)$ .  $\square$*

### 9.3 Operadores normais

*Operadores normais* em espaços unitários são definidos do mesmo modo que operadores normais em espaços Euclidianos. São aqueles que comutam com seu adjunto. Matricialmente identificamo-os através de uma representação. O operador é normal se, e somente se, sua representação numa base ortonormal é uma matriz normal. Existem diversas maneiras de construirmos operadores normais, entre as quais destacamos o operador de Gram  $P$  associado a um operador  $A$ , que por definição é  $P = A^* \circ A$ .

**Proposição 9.3.1** *Se  $A$  é um operador normal num espaço unitário  $(V, \langle \cdot, \cdot \rangle)$  então  $\|A(v)\| = \|A^*(v)\|$  para qualquer vetor  $v \in V$ . Em consequência,*

- a) *os núcleos de  $A$  e  $A^*$  coincidem;*
- b)  *$V$  decompõe-se na soma direta ortogonal,  $V = \text{Im } A \sqcup \text{Nuc } A$ ;*
- c) *as imagens de  $A$  e  $A^*$  são iguais.*

Quase todas demonstrações das proposições dessa seção são iguais às demonstrações apresentadas no caso Euclidianos. A diferença entre os resultados fica por conta do corpo dos números complexos ser algebricamente fechado, fato que surge com plena força na versão do Teorema da decomposição primária para operadores normais em espaços unitários.

**Teorema 9.3.1 (Teorema da decomposição normal)** *Seja  $A$  um operador num espaço unitário  $V$ . Se  $A$  é normal então*

- a) *a fatoração primária do polinômio minimal de  $A$  é um produto de polinômios lineares distintos,  $m_A(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_k)$ ,*
- b) *e a decomposição primária de  $V$  é uma soma direta ortogonal de autoespaços*

$$V = V_{\lambda_1} \sqcup V_{\lambda_2} \sqcup \cdots \sqcup V_{\lambda_k}.$$

*Em particular, um operador normal é diagonalizável numa base ortonormal de autovetores.*

Com as mesmas hipóteses do Teorema da decomposição normal, temos o

**Corolário 9.3.1** *O polinômio minimal de  $A^*$  é o conjugado complexo de  $m_A(t)$  e as parcelas das decomposições normais de  $A$  e  $A^*$  são iguais,  $V_{\lambda_i} = V_{\overline{\lambda_i}}$ .*

**Demonstração** A relação entre os polinômios mínimos é independente da normalidade. Agora, como  $A$  é normal então  $A - \lambda_i Id$  é também normal. Como os núcleos de um operador normal e de seu adjunto coincidem, temos que

$$Nuc(A^* - \bar{\lambda}_i Id) = Nuc(A - \lambda_i Id)^* = Nuc(A - \lambda_i Id)$$

e a demonstração do corolário está completa.  $\square$

**Corolário 9.3.2** *Seja  $A$  um operador linear num espaço unitário  $V$ . Então  $A$  é normal se, e somente se,  $A$  é diagonalizável numa base ortonormal de autovetores.*

**Exercício 9.3.1** Demonstre que um operador  $A$  em  $(V, \langle \cdot, \cdot \rangle)$  é normal  $\Leftrightarrow A^*$  é um polinômio em  $A$ .

## 9.4 Operadores autoadjuntos

Um operador  $A$  num espaço unitário  $V$  é autoadjunto quando  $A^* = A$ , ou equivalentemente, quando  $[A^*]_\beta = [A]_\beta^*$  para uma (e portanto qualquer) base ortonormal  $\beta \subset V$ . É claro que um operador autoadjunto é normal. Note que os autovalores de um operador autoadjunto são reais pois se  $v \in V$  é um autovetor associado ao autovalor  $\lambda \in \mathbb{C}$ , então

$$\lambda \|v\|^2 = \lambda \langle v, v \rangle = \langle A(v), v \rangle = \langle v, A(v) \rangle = \langle v, \lambda v \rangle = \bar{\lambda} \langle v, v \rangle = \bar{\lambda} \|v\|^2.$$

Operadores autoadjuntos admitem uma caracterização que não existe similar para operadores simétricos em espaços Euclidianos.

**Teorema 9.4.1 (Teorema do operador autoadjunto)** *Um operador  $A$  num espaço unitário  $V$  é autoadjunto se, e somente se,  $\langle v, A(v) \rangle$  é real para todo  $v \in V$ .*

**Demonstração** Provemos a necessidade. Como

$$\langle v, A(v) \rangle = \overline{\langle A(v), v \rangle} = \overline{\langle v, A^*(v) \rangle} = \overline{\langle v, A(v) \rangle},$$

necessariamente  $\langle v, A(v) \rangle$  é um escalar real, desde que ele é igual ao seu conjugado. A demonstração da suficiência é simples porém um pouco mais elaborada. Para dois vetores  $u, v \in V$ , examinemos as igualdades,

$$\begin{cases} \langle u + v, A(u + v) \rangle = \langle u, A(u) \rangle + \langle v, A(u) \rangle + \langle u, A(v) \rangle + \langle v, A(v) \rangle, \\ \langle u + iv, A(u + iv) \rangle = \langle u, A(u) \rangle + \langle iv, A(u) \rangle + \langle u, A(iv) \rangle + \langle v, A(iv) \rangle. \end{cases}$$

Note que a hipótese  $\langle w, A(w) \rangle \in \mathbb{R}$  para qualquer  $w \in V$  implica que os fatores  $\langle v, A(u) \rangle + \langle u, A(v) \rangle$  e  $i \langle v, A(u) \rangle - i \langle u, A(v) \rangle$  são reais, ou dito de outro modo, são iguais a seus conjugados complexos,

$$\begin{cases} \langle v, A(u) \rangle + \langle u, A(v) \rangle = \langle A(u), v \rangle + \langle A(v), u \rangle \\ i \langle v, A(u) \rangle - i \langle u, A(v) \rangle = -i \langle A(v), u \rangle + i \langle A(v), u \rangle. \end{cases}$$

Multiplicando a segunda equação pelo escalar imaginário  $i$  e somando as duas equações, a menos do fator 2, obtemos que  $\langle u, A(v) \rangle = \langle A(u), v \rangle$ .  $\square$

Finalmente, um operador  $A$  em  $(V, \langle \cdot, \cdot \rangle)$  é *antiadjunto* quando  $A^* = -A$ .

### Exercícios propostos 9.4.1

1. Seja  $A$  que um operador normal  $A$  num espaço unitário  $V$ . Prove que  $A$  é autoadjunto se, e somente se, o seu espectro está contido no eixo real.
2. Um operador  $A$  num espaço unitário  $V$  é normal se, e somente se,  $A = A_1 + iA_2$  onde  $A_1$  e  $A_2$  são operadores autoadjuntos que comutam.
3. Seja  $A$  um operador normal em  $(V, \langle \cdot, \cdot \rangle)$ . Prove as afirmações.
  - a)  $A$  é antiadjunto  $\Leftrightarrow$  a representação de  $A$  numa base ortonormal é uma matriz anti-adjunta.
  - b)  $A$  é antiadjunto  $\Leftrightarrow$  o seu espectro está contido no eixo imaginário.

## 9.5 Operadores unitários

Diz-se que um operador  $U$  em  $(V, \langle \cdot, \cdot \rangle)$  é *unitário* quando  $U \circ U^* = Id = U^* \circ U$ .

Portanto, operadores unitários são normais, invertíveis e sua inversa é seu operador adjunto. Matricialmente são identificados através de uma representação numa base ortonormal  $\beta \subset V$ . O operador  $U$  é unitário se, e somente se,  $[U]_\beta$  é uma matriz unitária. O espectro de um operador unitário está contido no círculo unitário pois se  $v \in V$  é um autovetor associado ao autovalor  $\lambda \in \mathbb{C}$  temos que

$$\|v\|^2 = \langle v, v \rangle = \langle v, U^* \circ U(v) \rangle = \langle U(v), U(v) \rangle = \lambda \bar{\lambda} \|v\|^2.$$

Logo  $|\lambda|^2 = 1$ .

Toda aplicação num espaço unitário que preserva o produto interno é um operador linear. Vejamos algumas caracterizações de um operador unitário.

**Proposição 9.5.1** *As seguintes afirmações sobre um operador linear  $U$  num espaço unitário  $V$  são equivalentes.*

- a) O operador preserva o produto interno.
- b) O operador preserva a norma.

- c) O operador é unitário.
- d) O operador aplica base ortonormal em base ortonormal.

### Exercícios propostos 9.5.1

1. Mostre que a matriz mudança de coordenadas entre duas bases ortonormais de um espaço unitário é uma matriz unitária.
2. Prove que um operador normal num espaço unitário é um operador unitário se, e somente se, o espectro está contido no círculo unitário.
3. Demonstre que o produto usual de matrizes induz uma estrutura de grupo no espaço das matrizes unitárias,  $U(n, \mathbb{C}) = \{Q \in M(n, \mathbb{C}); QQ^* = I = Q^*Q\}$ .

## 9.6 Decomposição polar

Um operador  $P$  num espaço unitário  $V$  é um *operador positivo* se  $\langle v, P(v) \rangle > 0$  para todo vetor não nulo  $v \in V$ . Pelo Teorema do operador autoadjunto podemos afirmar que  $P$  autoadjunto. Num espaço Euclidiano  $V$  apenas a condição  $\langle v, P(v) \rangle > 0$  para todo vetor não nulo  $v$  não caracteriza um operador positivo.

**Exemplo 9.6.1** O operador  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $A(x, y) = (x - y, x + y)$  satisfaz a condição  $\langle (x, y), A(x, y) \rangle > 0$  em relação ao produto interno canônico para todo vetor não nulo pois  $\langle (x, y), A(x, y) \rangle = x^2 + y^2$  e o operador não é simétrico.  $\square$

Daremos uma seqüência de fatos cujas demonstrações são semelhantes às demonstrações já feitas no caso Euclidiana.

**Exercício 9.6.1** Assuma que  $P$  é um operador positivo no espaço unitário  $(V, \langle \cdot, \cdot \rangle)$ . Mostre que  $P$  é invertível e que a aplicação  $\langle \langle \cdot, \cdot \rangle \rangle : V \times V \rightarrow V$  definida por  $\langle \langle v, w \rangle \rangle = \langle v, P(w) \rangle$  é um outro produto interno em  $V$ .  $\square$

**Proposição 9.6.1** As seguintes afirmações sobre um operador  $P$  num espaço unitário  $V$  são equivalentes.

- a) O operador é positivo.
- b) O operador é autoadjunto e os autovalores são positivos.
- c)  $P$  é um operador de Gram associado a um operador invertível  $A : V \rightarrow V$ .

**Proposição 9.6.2** *Suponha que  $\langle \langle \cdot, \cdot \rangle \rangle$  seja um outro produto interno no espaço unitário  $(V, \langle \cdot, \cdot \rangle)$ . Então existe um único operador positivo  $P : V \rightarrow V$  tal que  $\langle \langle v, w \rangle \rangle = \langle v, P(w) \rangle$ .*

Um operador  $P$  num espaço  $(V, \langle \cdot, \cdot \rangle)$  é não negativo se  $\langle v, P(v) \rangle \geq 0$  para todo vetor não nulo  $v \in V$ .

**Proposição 9.6.3** *As seguintes afirmações sobre um operador  $P$  num espaço unitário  $V$  são equivalentes.*

- a) *O operador é não negativo.*
- b) *O operador é autoadjunto e os autovalores são não negativos.*
- c)  *$P$  é um operador de Gram associado a um operador  $A : V \rightarrow V$ .*

**Exercício 9.6.2** Prove que para cada operador  $A$  em um espaço unitário  $V$  existe um único operador não negativo tal que  $P^2 = A^* \circ A$ . Denotaremos esse operador de Gram por  $P = \sqrt{A^* \circ A}$ .  $\square$

**Teorema 9.6.1 (Teo. da decomposição polar)** *Para cada operador linear  $A$  num espaço unitário  $V$ , existe um único operador não negativo  $P$  e um operador unitário  $U$  tal que  $A = U \circ P$ . Além disso, se  $A$  é invertível então  $P$  é positivo e  $U$  é unicamente determinado.*

### Exercícios propostos 9.6.1

1. Determine a decomposição polar dos operadores quando o produto interno considerado é o produto interno canônico.
  - a)  $A : \mathbb{C}^2 \rightarrow \mathbb{C}^2$   $A(x, y) = (x, x + y)$ .
  - b)  $A : \mathbb{C}^3 \rightarrow \mathbb{C}^3$   $A(x, y, z) = (y, z, x)$ .
2. Prove que um operador  $A$  num espaço unitário  $V$  é normal  $\Leftrightarrow A = PU$  onde  $P$  é não negativo,  $U$  é unitário e  $PU = UP$ .

# Capítulo 10

## Formas bilineares

Um produto interno num espaço Euclidiano é um exemplo de uma forma bilinear, ou dito de outro modo, é um exemplo de uma aplicação com duas variáveis e linear em cada uma delas. O objetivo desse capítulo é estudar tais aplicações bilineares mas agora sem exigir a condição de ser positiva definida.

### 10.1 Formas bilineares

Seja  $V$  um espaço vetorial sobre o corpo  $\mathbb{K}$ . Uma aplicação  $g : V \times V \rightarrow \mathbb{K}$  é uma *forma bilinear* quando para quaisquer vetores  $u, v, w \in V$  e para qualquer escalar  $\lambda \in \mathbb{K}$  valem as igualdades

1.  $g(u + \lambda v, w) = g(u, w) + \lambda g(v, w)$ ,
2.  $g(u, v + \lambda w) = g(u, v) + \lambda g(u, w)$ ,

Essencialmente, uma forma bilinear é uma aplicação com duas variáveis na qual ao fixarmos uma delas obtemos um funcional linear. Um produto interno num espaço Euclidiano é um exemplo, entretanto, o mesmo não ocorre num espaço unitário pois o seu produto interno é quase linear na segunda variável.

**Exemplo 10.1.1** 1) Ao definirmos a aplicação  $g : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$ , por

$$g((z_1, z_2, \dots, z_n), (w_1, w_2, \dots, w_n)) = z_1 w_1 + z_2 w_2 + \dots + z_n w_n,$$

obtemos uma forma bilinear em  $\mathbb{C}^n$  que não é um produto interno. O produto interno canônico é obtido pela fórmula  $\langle u, v \rangle = g(u, \bar{v})$ .

2) Dados dois funcionais lineares  $f_1, f_2 : V \rightarrow \mathbb{K}$  num espaço vetorial  $V$  podemos contruir uma forma bilinear chamada de *produto tensorial* de  $f_1$  por  $f_2$ , denotada por  $f_1 \otimes f_2 : V \times V \rightarrow \mathbb{K}$ , e definida por  $f_1 \otimes f_2(u, v) \equiv f_1(u)f_2(v)$ .

3) Dados dois funcionais lineares  $f_1, f_2 : V \rightarrow \mathbb{K}$  num espaço vetorial  $V$  podemos contruir uma forma bilinear chamada de *produto exterior* de  $f_1$  por  $f_2$ , denotada por  $f_1 \wedge f_2 : V \times V \rightarrow \mathbb{K}$ , e definida por  $f_1 \wedge f_2(u, v) \equiv f_1(u)f_2(v) - f_1(v)f_2(u)$ .

4) A aplicação  $g : M(n, \mathbb{K}) \rightarrow \mathbb{K}$ ,  $g(X, Y) = \text{tr}(X^t Q_0 Y)$  é uma forma bilinear, onde  $Q_0 \in M(n, \mathbb{K})$  é uma matriz fixada a priori.  $\square$

**Exercício 10.1.1** Denote por  $\mathcal{B}(V, \mathbb{K})$  o conjunto de todas formas bilineares no espaço vetorial  $V$  sobre o corpo  $\mathbb{K}$ . Mostre que as usuais operações de somar duas funções e mutiplicar uma função por um escalar de  $\mathbb{K}$ , induzem uma estrutura de espaço vetorial no conjunto  $\mathcal{B}(V, \mathbb{K})$ .  $\square$

### Exercícios propostos 10.1.1

- Verifique que as formas abaixo são bilineares e escreva-as como uma soma de produto tensoriais.
  - $g : \mathbb{K}^3 \rightarrow \mathbb{K}$ ,  $g((x_1, y_1, z_1), (x_2, y_2, z_2)) = (x_1 + y_1 + z_1)(x_2 + y_2 + z_2)$ .
  - $g : \mathbb{K}^3 \rightarrow \mathbb{K}$ ,  $g((x_1, y_1, z_1), (x_2, y_2, z_2)) = x_1 y_2 + y_1 z_2 + z_1 x_2$ .
- Sejam  $f_1$  e  $f_2$  dois funcionais lineares num espaço vetorial  $V$ . Prove:
  - Se  $g = f_1 \otimes f_2 + f_2 \otimes f_1$ , então  $g(u, v) = g(v, u)$  para quaisquer  $u, v \in V$ .
  - Se  $g = f_1 \wedge f_2$ , então  $g(u, v) = -g(v, u)$  para quaisquer  $u, v \in V$ .
- Uma forma sesqui-linear num espaço vetorial complexo  $V$  é uma aplicação  $g : V \times V \rightarrow \mathbb{C}$  tal que para quaisquer  $u, v \in V$  e  $\lambda \in \mathbb{C}$  satisfaz
  - Sq1.  $g(u + \lambda v, w) = g(u, w) + \lambda g(v, w)$ ;
  - Sq2.  $g(u, \lambda v + w) = \bar{\lambda} g(u, v) + g(u, w)$ .
  - Verifique que se  $A$  é um operador linear em  $V$  e  $g$  é sesqui-linear então  $h(u, v) = g(u, h(v))$  é sesqui-linear e que
  - se  $A$  é um operador quase linear em  $V$  e  $g$  é uma forma bilinear, então  $h(u, v) = g(u, A(v))$  é sesqui-linear.

## 10.2 Representação matricial

Quando o espaço vetorial  $V$  tem dimensão finita é possível identificar cada forma bilinear  $g : V \times V \rightarrow \mathbb{K}$  a uma matriz quadrada. Para construir a representação fixemos uma base ordenada  $\beta = \{v_1, v_2, \dots, v_n\} \subset V$  e consideremos os vetores

$$u = x_1 v_1 + x_2 v_2 + \dots + x_n v_n \quad \text{e} \quad v = y_1 v_1 + y_2 v_2 + \dots + y_n v_n.$$

Utilizando-se da bilinearidade, a avaliação de  $g$  no par  $(u, v)$  fica sendo

$$g(u, v) = \sum_{i,j}^n g(v_i, v_j) x_i y_j.$$

Note que  $g(v_i, v_j)$  são escalares em  $\mathbb{K}$ . A matriz da forma bilinear  $g$  na base ordenada  $\beta$  é a matriz  $n \times n$  definida por  $[g]_\beta = [g_{ij}]$  em que  $g_{ij} = g(v_i, v_j)$ . Desse modo, obtemos um algoritmo que expressa  $g$  com muita clareza, a saber,

$$g(u, v) = [u]_\beta^t [g]_\beta [v]_\beta,$$

ressaltando que estamos identificando o escalar do membro esquerdo da igualdade com a matriz  $1 \times 1$  do membro direito. Reciprocamente, dada uma matriz quadrada  $Q \in M(n, \mathbb{K})$  e fixada a base ordenada  $\beta \subset V$ , para construímos uma forma bilinear  $g : V \times V \rightarrow \mathbb{K}$  cuja representação é  $Q$  basta definir

$$g(u, v) = [u]_\beta^t Q [v]_\beta.$$

**Exemplo 10.2.1** 1) A representação matricial do produto interno canônico do  $\mathbb{R}^n$  na base canônica é a matriz identidade, pois se  $u = (x_1, x_2, \dots, x_n)$  e  $v = (y_1, y_2, \dots, y_n)$ , então

$$\begin{aligned} \langle u, v \rangle &= x_1 y_1 + x_2 y_2 + \dots + x_n y_n \\ &= \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \end{aligned}$$

pois  $\langle e_i, e_j \rangle = \delta_{ij}$ . De forma compacta,  $\langle u, v \rangle = [u]_\alpha I [v]_\alpha$ .

2) A forma bilinear  $g : \mathbb{K}^3 \times \mathbb{K}^3 \rightarrow \mathbb{K}$ ,

$$g((x_1, x_2, x_3), (y_1, y_2, y_3)) = x_1 y_1 - 2x_1 y_3 + 5x_2 y_2 + x_3 y_1 - x_3 y_2 - 4x_3 y_3,$$

é transcrita matricialmente na base canônica como o produto matricial

$$g((x_1, x_2, x_3), (y_1, y_2, y_3)) = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 5 & 0 \\ 1 & -1 & -4 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}.$$

As entradas da representação de  $g$  são obtidas avaliando-se  $g_{ij} = g(e_i, e_j)$ .  $\square$

Resumiremos os comentários sobre a representação de uma forma bilinear numa proposição cuja demonstração não será apresentada pois nessa altura do texto os argumentos utilizados já são bem conhecidos do leitor.

**Proposição 10.2.1** *Sejam  $V$  um espaço vetorial de dimensão  $n$  sobre o corpo  $\mathbb{K}$  e  $\beta \subset V$  uma base ordenada. Então a aplicação*

$$\Psi : \mathcal{B}(V, \mathbb{K}) \rightarrow M(n, \mathbb{K}), \quad \Psi(g) = [g]_{\beta},$$

é um isomorfismo linear. Em particular,  $\dim \mathcal{B}(V, \mathbb{K}) = n^2$ .

**Exercício 10.2.1** Definimos o núcleo à direita e o núcleo à esquerda de uma forma bilinear  $g$  num espaço vetorial  $V$  como sendo os conjuntos, respectivamente,

$$dNuc\ g = \{u \in V; g(u, w) = 0 \ \forall w \in V\},$$

$$eNuc\ g = \{v \in V; g(w, v) = 0 \ \forall w \in V\}.$$

Demonstre que os dois conjuntos são subespaços de  $V$ . Construa um exemplo no qual os dois núcleos são diferentes.  $\square$

### Exercícios propostos 10.2.1

1. Determine a representação matricial da forma bilinear na base canônica.
  - a)  $g : \mathbb{K}^3 \rightarrow \mathbb{K}^3$ ,  $g((x_1, y_1, z_1), (x_2, y_2, z_2)) = (x_1 + y_1 + z_1)(x_2 + y_2 + z_2)$ .
  - b)  $g : \mathbb{K}^3 \rightarrow \mathbb{K}^3$ ,  $g((x_1, y_1, z_1), (x_2, y_2, z_2)) = x_1y_2 + y_1z_2 + z_1x_2$ .
2. Seja  $V$  um espaço Euclidiano.
  - a) Prove que a representação matricial do produto interno em qualquer base ortonormal é a matriz identidade.
  - b) Se  $A$  é um operador linear em  $V$ , mostre que  $g(u, v) = \langle u, A(v) \rangle$  é uma forma bilinear. Dê as representações matriciais de  $g$  e de  $A$  numa base ortonormal.
3. Seja  $\beta = \{v_1, v_2, \dots, v_n\}$  uma base do espaço vetorial  $V$  e  $\beta^* = \{f_1, f_2, \dots, f_n\}$  sua base dual. Mostre que  $\beta^{\otimes} = \{f_i \otimes f_j\}$  é uma base para o espaço das formas bilineares  $\mathcal{B}(V, \mathbb{K})$  e qualquer forma bilinear  $g$  em  $V$  é escrita como  $g = \sum_{i,j=1}^n g(v_i, v_j) f_i \otimes f_j$ .

## 10.3 Propriedades da representação

Comparemos as representações matriciais de uma mesma forma bilinear em duas bases ordenadas distintas de um espaço vetorial de dimensão finita. Ressaltamos que as duas matrizes construídas não são conjugadas, necessariamente.

**Proposição 10.3.1** *Sejam  $\beta$  e  $\gamma$  duas bases ordenadas do espaço vetorial de dimensão finita  $V$ . Se  $g : V \times V \rightarrow \mathbb{K}$  é uma forma bilinear então  $[g]_{\gamma} = Q^t [g]_{\beta} Q$  em que  $Q = [id]_{\beta}^{\gamma}$ .*

**Demonstração** Como sabemos, as relações entre as coordenadas dos vetores  $u, v \in V$  nas duas bases ordenadas são  $[u]_{\beta} = Q [u]_{\gamma}$  e  $[v]_{\beta} = Q [v]_{\gamma}$  onde  $Q = [id]_{\beta}^{\gamma}$ . Sendo assim, avaliemos  $g(u, v)$ ,

$$g(u, v) = [u]_{\beta}^t [g]_{\beta} [v]_{\beta} = (Q[u]_{\gamma})^t [g]_{\beta} (Q[v]_{\gamma}) = [u]_{\gamma}^t (Q^t [g]_{\beta} Q) [v]_{\gamma}.$$

Pela unicidade da representação da forma bilinear  $g$  na base ordenada  $\gamma$  concluímos que  $[g]_{\gamma} = Q^t [g]_{\beta} Q$ .  $\square$

Repetiremos uma construção já apresentada anteriormente ao estudarmos um produto interno num espaço Euclidiano. Fixado uma forma bilinear  $g$  num espaço vetorial  $V$ , temos uma aplicação  $\Lambda : V \rightarrow V^*$ ,  $v \mapsto \Lambda[v]$ , onde  $\Lambda[v]$  denota o funcional linear

$$\Lambda[v] : V \rightarrow \mathbb{K}, \quad \Lambda[v](w) = g(w, v).$$

É rotina verificar que a aplicação  $\Lambda : V \rightarrow V^*$  é uma transformação linear. Quando a forma bilinear é um produto interno num espaço Euclidiano sabemos que  $\Lambda$  é um isomorfismo linear. Entretanto, como isso não ocorre com qualquer forma bilinear, distinguiremos as duas situações através de uma definição.

Diremos que  $g$  é *não degenerada* (ou *não singular*) se a transformação linear  $\Lambda : V \rightarrow V^*$  é um isomorfismo linear. Caso contrário diremos que  $g$  é *degenerada* (ou *singular*). Essa distinção pode ser estabelecida matricialmente.

**Lema 10.3.1** *Uma forma bilinear  $g$  é degenerada se, e somente se, a matriz de  $g$  numa base ordenada  $\gamma \subset V$  não é invertível.*

**Demonstração** Por definição  $g$  é degenerada se, e somente se, existe um vetor não nulo  $v_0 \in V$  tal que o funcional linear  $f_0 : V \rightarrow \mathbb{K}$ ,  $f_0(w) = g(w, v_0)$  é identicamente nulo, ou equivalentemente, existe um vetor  $v_0 \in V$  tal que  $[w]_{\gamma} [g]_{\gamma} [v_0]_{\gamma} = 0$  para todo  $w \in V$ . Entretanto, isso ocorre se, e somente se,  $[g]_{\gamma} [v_0]_{\gamma} = [0]$ . Portanto,  $g$  é degenerada se, e somente se,  $\lambda = 0$  é um autovalor de  $[g]_{\gamma} \Leftrightarrow [g]_{\gamma}$  é não invertível.  $\square$

Como duas representações de  $g$  são conjugadas, a proposição acima afirma um pouco mais. A forma bilinear  $g$  é degenerada se, e somente se,  $[g]_{\gamma}$  é não invertível para qualquer base ordenada  $\gamma \subset V$ .

Formas bilineares em espaços Euclidianos admitem uma elegante descrição em termos de operadores lineares. Do mesmo modo que descrevemos qualquer produto interno em um dado espaço Euclidiano utilizando um operador positivo, também é possível descrever qualquer forma bilinear a partir do produto interno que define a estrutura Euclidiana.

**Exemplo 10.3.1** Um operador linear  $A$  num espaço Euclidiano  $(V, \langle \cdot, \cdot \rangle)$  dá origem a uma forma bilinear  $g : V \times V \rightarrow \mathbb{K}$ , bastando para isso definir  $g(u, v) = \langle u, A(v) \rangle$ . A recíproca dessa afirmação também é verdadeira. Se não vejamos.

Seja  $g$  uma forma bilinear no espaço Euclidiano  $(V, \langle \cdot, \cdot \rangle)$ . Para cada vetor  $v \in V$  considere o funcional linear  $f_v : V \rightarrow \mathbb{R}$ ,  $f_v(u) = g(u, v)$ . Pelo Teorema da representação de um funcional linear, existe um único vetor  $v' \in V$  tal que  $g(u, v) = \langle u, v' \rangle$ . Sendo assim, defina a aplicação  $A : V \rightarrow V$ ,  $v \mapsto v'$ . Mostra-se que  $A$  é o único operador linear satisfazendo a condição  $g(u, v) = \langle u, A(v) \rangle$  para quaisquer vetores  $u, v \in V$ .  $\square$

**Exercício 10.3.1** Se  $g$  é uma forma bilinear no espaço unitário  $(V, \langle \cdot, \cdot \rangle)$ , então existe um e somente um operador quase linear  $A : V \rightarrow V$  tal que  $g(u, v) = \langle u, A(v) \rangle$  para quaisquer  $u, v \in V$ . Demonstre esse teorema de representação.  $\square$

### Exercícios propostos 10.3.1

- Para cada forma bilinear  $g$  em  $\mathbb{R}^2$  determine o operador linear  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  tal que  $g(u, v) = \langle u, A(v) \rangle$ , onde o produto interno considerado é o canônico.
  - $g((x_1, y_1), (x_2, y_2)) = x_1y_2 - y_1x_2$ .
  - $g((x_1, y_1), (x_2, y_2)) = x_1x_2 + x_1y_2 - x_2y_1$ .
- Seja  $g$  uma forma bilinear não degenerada no espaço vetorial de dimensão finita  $V$ . Dada qualquer outra forma bilinear  $f : V \times V \rightarrow \mathbb{R}$ , existe um único operador linear  $A : V \rightarrow V$  representando  $f$ , isto é,  $f(u, v) = g(u, A(v))$ .
- Mostre que no exercício acima se eliminarmos a hipótese "não degenerada" o resultado deixa de ser verdadeiro.
- Suponha que a forma bilinear  $g$  no espaço Euclidiano  $(V, \langle \cdot, \cdot \rangle)$  é representado pelo operador linear  $A : V \rightarrow V$ ,  $g(u, v) = \langle u, A(v) \rangle$ . Prove as igualdades.
  - $dNuc g = Nuc A^t$ .
  - $eNuc g = Nuc A$ .
- Demonstre que os núcleos à direita e à esquerda de uma forma bilinear num espaço vetorial de dimensão finita têm as mesmas dimensões (utilize o teorema do Posto).
- Prove que uma forma bilinear  $g$  num espaço vetorial de dimensão finita  $V$  é não degenerada  $\Leftrightarrow dNuc g = eNuc g = \{o\}$ .

## 10.4 Formas bilineares simétricas

Uma forma bilinear  $g : V \times V \rightarrow \mathbb{K}$  é *simétrica* quando  $g(u, v) = g(v, u)$  para quaisquer  $u, v \in V$ .

Além do produto interno num espaço vetorial real, já conhecemos várias outras formas bilineares simétricas, por exemplo, a aplicação  $g : M(n, \mathbb{K}) \rightarrow \mathbb{K}$  definida

por  $g(N, P) = \text{tr}(NP)$ . Num espaço vetorial  $V$  de dimensão finita, a identificação de uma forma bilinear simétrica  $g$  é bastante simples. Dada qualquer base ordenada  $\beta = \{v_1, v_2, \dots, v_n\} \subset V$  seja  $g_{ij} = g(v_i, v_j)$  as entradas da representação de  $g$  na base  $\beta$ . Da hipótese de simetria segue que  $g_{ij} = g_{ji}$ . Então, a forma bilinear  $g$  é simétrica se, e somente se,  $[g]_\beta$  é uma matriz simétrica. Observamos que a simetria da representação  $[g]_\beta$  é independente da base considerada.

**Exercício 10.4.1** Seja  $g$  uma forma bilinear simétrica num espaço vetorial  $V$ . Prove que  $dNuc g = eNuc g$ . Por causa dessa propriedade escreveremos simplesmente "o núcleo de  $g$ ", sem mais comentários.  $\square$

Pela proximidade entre os conceitos, podemos transpor várias noções definidas num espaço com produto interno para um espaço equipado com uma forma bilinear simétrica. Por exemplo, diremos que os vetores  $u, v \in V$  são  *$g$ -ortogonais* quando  $g(u, v) = 0$ . Ressaltamos que o conceito de  *$g$ -ortogonalidade* não é exatamente igual ao conceito de ortogonalidade num espaço vetorial real equipado com produto interno pois não estamos assumindo que  $g$  seja positiva definida. Por isso, é possível existir algum vetor não nulo em  $V$  que é  *$g$ -ortogonal* a si mesmo. Por exemplo, o vetor  $v = (1, -1) \in \mathbb{R}^2$  é  *$g$ -ortogonal* a si mesmo quando estamos considerando a forma bilinear simétrica

$$g : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}, \quad g((x_1, x_2), (y_1, y_2)) = x_1y_1 + x_2y_1 + x_1y_2 + x_2y_2.$$

Do mesmo modo pode ocorrer que  $g(u, u) < 0$ , entretanto, somente com a forma bilinear identicamente nula podemos ter cada um dos vetores do espaço  *$g$ -ortogonal* a si mesmo. Para mostrar essa afirmação precisaremos de duas definições.

Uma aplicação  $q : V \rightarrow \mathbb{K}$  é uma *forma quadrática* quando existe uma forma bilinear  $g$  em  $V$  tal que  $q(v) = g(v, v)$ . A forma bilinear utilizada para definir uma forma quadrática pode ser substituída por uma forma bilinear simétrica, a saber,

$$g_0(u, v) = \frac{1}{2}(g(u, v) + g(v, u)).$$

É imediato concluir que também vale a identidade  $q(v) = g_0(v, v)$ . Levando em conta esses comentários, sempre que citarmos uma forma quadrática estaremos supondo que a forma bilinear associada  $g$  é simétrica.

Conhecendo-se a representação matricial da forma bilinear simétrica  $g$  numa base ordenada  $\beta = \{v_1, v_2, \dots, v_n\} \subset V$ , obtemos os valores da forma quadrática associada  $q$  no vetor  $v \in V$  exclusivamente em função das coordenadas do vetor na base considerada,

$$q(v) = [v]_\beta^t [g]_\beta [v]_\beta = \sum_{i,j=1}^n g(v_i, v_j) x_i x_j.$$

Reciprocamente, dada uma forma quadrática (que é um polinômio homogêneo

de grau dois em  $n$  variáveis) em função das coordenadas dos vetores de uma base ordenada  $\beta$ ,

$$q(v) = \sum_{i \leq j=1}^n g(v_i, v_j) x_i x_j,$$

recuperamos a única forma bilinear simétrica que define  $q$  através da representação matricial

$$[g(v_i, v_j)]_{\beta} = \frac{T+T^t}{2}$$

em que  $T$  é a matriz triangular superior

$$T = [b_{ij}] \quad \text{com} \quad b_{ij} = \begin{cases} g_{ij} & \text{se } i \leq j \\ 0 & \text{se } i > j \end{cases}.$$

**Exemplo 10.4.1** 1) A forma quadrática associada ao produto interno canônico do  $\mathbb{R}^n$  é o polinômio homogêneo de grau dois em  $n$  variáveis  $q(x_1, x_2, \dots, x_n) = x_1^2 + x_2^2 + \dots + x_n^2$ .

2) Considere a forma bilinear simétrica  $g : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ ,

$$g((x_1, x_2), (y_1, y_2)) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

A representação matricial indicada está na base canônica. Nesse exemplo, existem vetores não nulos que são  $g$ -ortogonais a todos os outros vetores, é o caso do vetor  $v = (1, -1)$ . A explicação provém da matriz  $[g]_{\alpha}$ . Ela não é invertível, isto é,  $g$  é uma forma bilinear simétrica degenerada.

3) A forma quadrática em  $\mathbb{R}^3$  dada na base canônica por  $q(v) = -2x^2 + 4xy - 6yz$  onde  $v = (x, y, z)$  tem uma representação matricial da forma

$$q(v) = [x, y, z] \begin{bmatrix} -2 & 2 & 0 \\ 2 & 0 & -3 \\ 0 & -3 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}. \quad \square$$

**Exercício 10.4.2** Seja  $g$  uma forma bilinear simétrica em  $V$ . Prove a *identidade de polarização*,

$$g(u, v) = \frac{1}{4}(q(u+v) - q(u-v)).$$

Conclua que duas formas bilineares simétricas,  $g_1$  e  $g_2$ , determinam a mesma forma quadrática se, e somente se,  $g_1 \equiv g_2$ .  $\square$

Voltemos à questão inicial sobre o conceito de  $g$ -ortogonalidade.

**Lema 10.4.1** Se todo vetor  $w \in V$  é  $g$ -ortogonal a si mesmo, então  $g \equiv 0$ .

**Demonstração** Seja  $q$  a forma quadrática associada à  $g$ . Vamos supor que  $g(w, w) = 0$  para todo  $w \in V$ . Pela identidade de polarização temos que

$$g(u, v) = \frac{1}{4}(g(u + v, u + v) - g(u - v, u - v)) = 0, \quad \text{para todo } u, v \in V. \quad \square$$

### Exercícios propostos 10.4.1

1. Considere a forma bilinear em  $\mathbb{R}^3$ ,  $g((x_1, x_2, x_3), (y_1, y_2, y_3)) = (x_1 - 2x_2)(y_1 - 2y_2)$ .
  - a) Mostre que  $g$  é simétrica e determine a representação matricial na base canônica.
  - b) Encontre uma base para o núcleo de  $g$ .
  - c) Descreva a forma quadrática determinada por  $g$ .
2. Se  $g$  é uma forma bilinear em  $V$  e  $v_0$  é um vetor, prove que o conjunto  $W(v_0)^\perp = \{w \in V; g(v_0, w) = 0\}$  é um subespaço.
3. Prove que uma representação matricial de uma forma bilinear simétrica num espaço vetorial de dimensão finita é uma matriz simétrica, independente da base escolhida.
4. Seja  $(V, \langle \cdot, \cdot \rangle)$  um espaço Euclidiano. Mostre que para cada forma bilinear simétrica  $g$  existe um, e somente um, operador simétrico  $A : V \rightarrow V$  tal que  $g(u, v) = \langle u, A(v) \rangle$ .
5. Dados dois funcionais lineares  $f_1$  e  $f_2$  no espaço vetorial  $V$  de dimensão finita, verifique que  $g \equiv f_1 \otimes f_2 + f_2 \otimes f_1$  é uma forma bilinear simétrica. Determine a representação  $[g]_\beta$  conhecendo-se as representações  $[f_1]_\beta$  e  $[f_2]_\beta$ .
6. Demonstre que o conjunto  $\mathcal{S}(V, \mathbb{K}) \subset \mathcal{B}(V, \mathbb{K})$  formado pelas formas bilineares simétricas num espaço vetorial  $V$  é um subespaço. Calcule a dimensão de  $\mathcal{S}(V, \mathbb{K})$  sabendo-se que  $\dim V = n$ .
7. Assuma que a forma bilinear simétrica  $g$  em  $V$  é *semi-positiva definida*, isto é,  $g(u, u) \geq 0$ . Prove a desigualdade de Schwartz,  $g(u, v)^2 \leq g(u, u)g(v, v)$ , onde  $q$  é a forma quadrática associada à  $g$ .
8. Dada a aplicação  $g : M(n, \mathbb{K}) \times M(n, \mathbb{K}) \rightarrow \mathbb{K}$ ,  $g(N, P) = n \operatorname{tr}(NP) - \operatorname{tr}(N) \operatorname{tr}(P)$ .
  - a) Verifique que  $g$  é uma forma bilinear simétrica.
  - b) Calcule os valores  $g(P, I)$  e conclua que  $g$  é degenerada.

## 10.5 Bases g-ortonormais

Por simplicidade de redação introduziremos uma notação. Um par  $(V, g)$  significará um espaço vetorial  $V$  de dimensão finita sobre  $\mathbb{K}$  no qual está definida uma forma bilinear simétrica não identicamente nula  $g$ . Em alguns textos essa notação é reservada apenas para o caso em que a forma bilinear simétrica é não degenerada, mas aqui ela poderá ser ou não degenerada. Para ressaltar esse fato chamaremos

de *índice de nulidade* de  $(V, g)$  ao inteiro  $i_0$  definido por  $i_0 = \dim Nuc g$ . Note que a hipótese de  $g$  ser não identicamente nula implica que  $i_0 < \dim V$ . Por outro lado, se  $g$  for não degenerada, então  $i_0 = 0$ .

Uma base ordenada  $\beta = \{v_1, v_2, \dots, v_n\}$  de  $(V, g)$  é uma *base g-ortogonal* quando  $g(v_i, v_j) = 0$  para  $i \neq j$ . O objetivo dessa seção é mostrar que bases g-ortogonais existem. Na verdade exibiremos um processo para construir uma *base g-ortonormal*, isto é, uma base ordenada  $\beta = \{u_1, u_2, \dots, u_n\}$  tal que

$$g(u_i, u_j) = 0, \quad \text{se } i \neq j \quad \text{e} \quad g(u_i, u_i) \in \{-1, 0, 1\}.$$

Numa tal base, a representação  $[g]_\beta$  é uma matriz diagonal formada por 0 ou  $\pm 1$ .

Seja  $v_0$  um vetor do espaço vetorial  $V$ . Denote por  $W(v_0)$  o subespaço gerado pelo vetor  $v_0$  e por  $W(v_0)^\perp$  o seu subespaço g-ortogonal, em outras palavras,

$$W(v_0)^\perp = \{w \in V; g(v_0, w) = 0\}.$$

Como vimos, existe um vetor  $v_0$  tal que  $g(v_0, v_0) \neq 0$  pois estamos sempre assumindo que a forma bilinear simétrica não é identicamente nula.

**Lema 10.5.1** *Se  $g(v_0, v_0) \neq 0$  então  $V = W(v_0)^\perp \boxplus W(v_0)$  e  $Nuc g \subset W(v_0)^\perp$ .*

**Demonstração** A demonstração é uma adaptação do processo de ortogonalização de Gram-Schmidt. Provaremos inicialmente que  $V = W(v_0)^\perp + W(v_0)$ . Dado um vetor  $u \in V$  considere o vetor  $u^\perp$  definido por

$$u^\perp = u - \frac{g(u, v_0)}{g(v_0, v_0)} v_0.$$

Uma avaliação simples mostra que  $u^\perp \in W(v_0)^\perp$ . Logo  $u = u^\perp + \lambda v_0$ , significando que  $V$  é a soma dos subespaços indicados. Vamos supor que  $u \in W(v_0)^\perp \cap W(v_0)$ . Então, por um lado  $u = \lambda v_0$  para algum  $\lambda \in \mathbb{K}$  e por outro lado,

$$0 = g(v_0, u) = \lambda g(v_0, v_0).$$

Como  $g(v_0, v_0) \neq 0$ , concluímos que  $\lambda = 0$  e por conseguinte  $u = 0$ . Isso mostra que a soma é uma soma direta. A demonstração da inclusão do núcleo de  $g$  no ortogonal de  $W(v_0)$  deixaremos como exercício.  $\square$

**Proposição 10.5.1** *Se  $g$  é uma forma bilinear simétrica num espaço vetorial  $V$  de dimensão  $n$ , então existem vetores não nulos  $v_1, v_2, \dots, v_{n-i_0} \in V$  tais que*

$$V = W(v_1) \boxplus W(v_2) \boxplus \dots \boxplus W(v_{n-i_0}) \boxplus Nuc g,$$

*em que a soma direta é g-ortogonal,  $i_0$  é o índice de nulidade e  $g(v_i, v_i) \neq 0$ . Em consequência,  $V$  admite uma base g-ortogonal.*

**Demonstração** Escolha um vetor não nulo  $v_n \in V$  talque  $g(v_n, v_n) \neq 0$  e considere a decomposição  $V = W(v_n)^\perp \boxplus W(v_n)$ . Note que  $\dim W(v_n)^\perp = n - 1$  e que a restrição de  $g$  ao subespaço  $W(v_n)^\perp$  é uma forma bilinear simétrica com o mesmo índice de nulidade  $i_0$  pois o  $Nuc g \subset W(v_n)^\perp$ . Desse modo, utilizamos uma hipótese indutiva para obter vetores  $v_1, \dots, v_{n-1-i_0} \in W(v_n)^\perp$  tais que

$$W(v_n)^\perp = W(v_1) \boxplus \dots \boxplus W(v_{n-1-i_0}) \boxplus Nuc g.$$

Agora é fácil obter a decomposição desejada. Portanto, se  $\beta_0 = \{v_{n-i_0+1}, \dots, v_n\}$  é uma base ordenada de  $Nuc g$ , caso o núcleo não seja trivial, então a união ordenada  $\beta = \{v_1, \dots, v_{n-i_0}\} \cup \beta_0$  é uma base ortogonal de  $V$ .  $\square$

Na proposição acima construímos uma base  $g$ -ortogonal  $\beta$  de  $(V, g)$  na qual está contida uma base de  $Nuc g$ . Uma base para o núcleo está sempre presente numa base  $g$ -ortogonal, permitindo-nos determinar o índice de nulidade através de qualquer uma delas.

**Corolário 10.5.1** *Para qualquer base  $g$ -ortogonal  $\beta$  de  $(V, g)$  o conjunto de vetores  $\beta_0 = \{v \in \beta, g(v, v) = 0\}$  é uma base de  $Nuc g$ .*

**Demonstração** Seja  $\beta = \{v_1, v_2, \dots, v_n\}$  uma base  $g$ -ortogonal ordenada de tal modo que o conjunto  $\beta_0 = \{v \in \beta, g(v, v) = 0\}$  é formado pelos  $j_0$ -ésimos últimos elementos. Como a base é  $g$ -ortogonal, segue que  $\beta_0$  está contido no núcleo. Resta verificar que esse conjunto gera  $Nuc g$ . Com efeito. Dado um vetor  $v = x_1 v_1 + x_2 v_2 + \dots + x_n v_n$  no núcleo de  $g$ , consideremos qualquer elemento  $v_k \in \beta$  com  $0 \leq k \leq n - j_0$  e façamos a avaliação,

$$0 = g(v, v_k) = x_k g(v_k, v_k).$$

A condição  $g(v_k, v_k) \neq 0$  implica que  $x_k = 0$ , de onde concluímos que qualquer vetor do núcleo é uma combinação linear dos  $j_0$  últimos elementos de  $\beta$ . Logo  $j_0 = i_0$ , como desejávamos demonstrar.  $\square$

**Corolário 10.5.2** *Se  $V$  é um espaço vetorial complexo, então existe uma base  $g$ -ortonormal  $\beta = \{v_1, v_2, \dots, v_n\}$  de  $(V, g)$  tal que*

$$g(v_i, v_j) = 0 \quad \text{se} \quad i \neq j \quad \text{e} \quad g(v_i, v_i) \in \{0, 1\}.$$

**Demonstração** Seja  $\gamma = \{v_1, \dots, v_{n-i_0}, v_{n-i_0+1}, \dots, v_n\}$  uma base  $g$ -ortogonal construída como na proposição anterior. Defina uma base ordenada formada pelos vetores

$$\begin{cases} u_i = v_i, & \text{se } v_i \text{ pertence ao núcleo de } g \text{ e} \\ u_i = \frac{1}{\tau_i} v_i & \text{se } v_i \text{ não pertence ao núcleo de } g, \end{cases}$$

em que  $\tau_i$  é uma das raízes quadradas de  $g(v_i, v_i) \neq 0$ .  $\square$

### Exercícios propostos 10.5.1

1. Prove que uma forma bilinear simétrica  $g$  num espaço vetorial complexo  $V$  de dimensão finita admite uma representação do tipo

$$[g]_{\beta} = \text{diag}\{I_{n-i_0}, [0]\}.$$

2. Para qualquer matriz simétrica  $N \in M(n, \mathbb{C})$  existe uma matriz invertível  $P \in M(n, \mathbb{C})$  tal que  $P^t N P$  é diagonal. Demonstre a afirmação.

## 10.6 Teorema do índice

Nessa seção estudaremos apenas formas bilineares simétricas num espaço vetorial real  $V$  de dimensão finita. Indexaremos um subespaço  $W$  de  $(V, g)$  por  $+$  ou por  $-$  quando a forma bilinear  $g$  restrita ao subespaço é positiva definida ou negativa definida, respectivamente. Mais claramente,

$$\begin{cases} W_+ & \text{quando } g(w, w) > 0 \text{ para todo vetor não nulo } w \in W, \\ W_- & \text{quando } g(w, w) < 0 \text{ para todo vetor não nulo } w \in W. \end{cases}$$

Como o subespaço trivial cumpre as duas exigências, simultaneamente, ele será indexado conforme a conveniência do momento.

**Teorema 10.6.1 (Teorema do índice)** *Seja  $g$  uma forma bilinear no espaço vetorial real  $V$  de dimensão finita. Então existem subespaços  $W_-$  e  $W_+$  tais que*

$$V = W_- \boxplus \text{Nuc } g \boxplus W_+.$$

*Mais ainda, o inteiro  $i_- = \dim W_-$ , chamado de índice de Morse de  $(V, g)$ , depende apenas da forma bilinear no seguinte sentido, para qualquer outra decomposição em soma direta  $g$ -ortogonal  $V = U_- \boxplus \text{Nuc } g \boxplus U_+$  temos que  $i_- = \dim U_-$ .*

**Demonstração** A menos de reordenação, uma base  $g$ -ortonormal  $\beta$  de  $(V, g)$  decompõe-se como  $\beta = \beta_- \vec{\cup} \beta_0 \vec{\cup} \beta_+$  onde  $\beta_0$  é uma base para o núcleo de  $g$ ,

$$\beta_- = \{v \in \beta; g(v, v) < 0\} \quad \text{e} \quad \beta_+ = \{v \in \beta; g(v, v) > 0\}.$$

Se definimos  $W_-$  (respec.  $W_+$ ) como o subespaço gerado por  $\beta_-$  (respec.  $\beta_+$ ) a  $g$ -ortonormalidade da base  $\beta$  garante a decomposição  $V = W_- \boxplus \text{Nuc } g \boxplus W_+$ .

Suponha que tenhamos uma outra decomposição  $g$ -ortogonal para o espaço, digamos que seja  $V = U_- \boxplus \text{Nuc } g \boxplus U_+$ . Vamos supor por absurdo que exista um vetor não nulo  $u_- \in U_- \cap (\text{Nuc } g \boxplus W_+)$ . Sendo assim, podemos escrever

$u_- = w_0 + w_+$  com  $w_0 \in Nuc\ g$  e  $w_+ \in W_+$ . Fazemos a avaliação  $g(u_-, u_-)$  levando em conta que  $u \in U_-$ ,

$$0 > g(u_-, u_-) = g(w_0 + w_+, w_0 + w_+) = g(w_+, w_+) \geq 0.$$

Evidentemente, isso é contraditório. Como aquela interseção é nula podemos considerar o subespaço de  $V$  formado pela soma direta  $U_- \oplus (Nuc\ g \boxplus W_+)$  e concluir utilizando um cálculo dimensional simples que  $\dim U_- \leq \dim W_-$ . Do mesmo modo provamos que  $\dim U_+ \leq \dim W_+$ . Novamente suponha por absurdo que pelo menos uma dessas duas desigualdades é estrita, então

$$\dim V = \dim U_- + \dim Nuc\ g + \dim U_+ < \dim W_- + \dim Nuc\ g + \dim W_+ = \dim V.$$

Uma contradição. Logo, o inteiro  $i_- = \dim W_-$  não depende da decomposição.  $\square$

**Corolário 10.6.1** *Seja  $g$  é uma forma bilinear simétrica no espaço vetorial real  $V$  de dimensão finita. Então existe uma base  $g$ -ortonormal*

$$\beta = \{v_1, \dots, v_{i_-}, u_1, \dots, u_{i_0}, w_1, \dots, w_{i_+}\}$$

tal que  $g(v_i, v_j) = -\delta_{ij}$ ,  $g(w_i, w_j) = \delta_{ij}$  e  $g(u_i, u_j) = 0$ .

Deixaremos para o leitor a responsabilidade pela demonstração. Um par  $(V, g)$  é dito espaço *pseudo Euclidiano* quando  $V$  é um espaço vetorial real de dimensão finita e  $g$  é uma forma bilinear simétrica não degenerada. Nesse caso, temos uma decomposição  $g$ -ortogonal dada pelo Teorema do índice, na forma

$$V = W_- \boxplus W_+,$$

em que cada parcela da decomposição é uma soma direta de espaços unidimensionais dois a dois  $g$ -ortogonais. Um espaço pseudo Euclidiano  $(V, g)$  de dimensão 4 é dito *espaço de Minkowski* se o índice de Morse de  $g$  é  $i_- = 3$ .

### Exercícios propostos 10.6.1

- Seja  $A$  um operador no espaço Euclidiano  $(V, \langle \cdot, \cdot \rangle)$ .
  - Verifique que  $g(u, v) = \langle A(u), A(v) \rangle$  é uma forma bilinear simétrica.
  - Calcule os índices de nulidade e de Morse de  $g$ .
- Todos os índices associados a uma forma bilinear simétrica  $g$  podem ser detectados por uma base  $g$ -ortonormal. Prove que para quaisquer duas bases  $g$ -ortonormais  $\beta$  e  $\gamma$  de  $V$  temos as igualdades  $i_0 = \#\beta_0 = \#\gamma_0$  e  $i_- = \#\beta_- = \#\gamma_-$ .
- Mostre que toda forma bilinear simétrica  $g$  em um espaço vetorial real  $V$  de dimensão  $n$  admite uma representação da forma

$$[g]_\beta = \text{diag}\{-I_{i_-}, [0], I_{i_+}\},$$

em que  $i_+ = n - (i_0 + i_-)$  e  $[0]$  é a matriz nula  $i_0 \times i_0$ .

4. Seja  $\mathcal{H} \subset M(2, \mathbb{C})$  o subespaço formado pelas matrizes Hermitianas (ou autoadjuntas),  $N^* = N$ .
- Verifique que  $\mathcal{H}$  tem uma estrutura natural de um espaço vetorial real de dimensão 4 e que  $\beta = \{I, H_1, H_2, H_3\}$  é uma base ordenada de  $\mathcal{H}$ , onde os elementos de  $\beta$  são as *matrizes de Pauli*,
- $$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad H_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad H_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad H_3 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}.$$
- Prove que  $g(N, P) = \frac{1}{2}(tr(NP) - tr(N)tr(P))$  é uma forma bilinear simétrica não degenerada em  $\mathcal{H}$  (com valores em  $\mathbb{R}$ ).
  - Demonstre as fórmulas  $g(N, N) = -detN$  e  $g(N, I) = -\frac{1}{2}tr(N)$ .
  - Conclua que  $\beta$  é uma base  $g$ -ortonormal.
  - Dê uma decomposição da forma  $\mathcal{H} = W_+ \oplus W_-$  e calcule o índice de  $g$ .
  - Mostre que o espaço  $g$ -ortogonal à identidade é o conjunto das matrizes Hermitianas com traço nulo.
5. Seja  $g$  uma forma bilinear simétrica no espaço vetorial real  $V$  de dimensão  $n$ .
- Determine uma base  $\beta$  de  $V$  para a qual a forma quadrática associada tem a forma diagonal  $q(u) = \sum_{i=1}^n \lambda_i(x_i)^2$ .
  - Demonstre a Lei de inércia de Silvestre: o número de coeficientes positivos em qualquer forma diagonal de  $q$  é o mesmo.

## 10.7 Formas bilineares antisimétricas

Diz-se que uma forma bilinear  $g$  no espaço vetorial  $V$  é *antisimétrica* quando  $g(u, v) = -g(v, u)$ , para quaisquer vetores  $u, v \in V$ .

Num espaço vetorial de dimensão finita uma forma bilinear antisimétrica é reconhecida matricialmente através de uma representação pois em qualquer base ordenada  $\beta$  de  $V$  a matriz  $[g]_\beta$  é antisimétrica. A recíproca dessa afirmação também é verdadeira. Note que o núcleo à direita e o núcleo à esquerda são coincidentes, permitindo-nos definir o índice de nulidade de  $(V, g)$  como sendo o inteiro  $i_0 = \dim Nuc g$ .

Diremos que dois vetores  $u, v \in V$  são  $g$ -ortogonais se  $g(u, v) = 0$ . Por essa definição, todo vetor de  $V$  é  $g$ -ortogonal a si mesmo pois a condição de anti-simetria  $g(v, v) = -g(v, v)$  implica que  $g(v, v) = 0$ . Uma base ordenada  $\beta = \{v_1, v_2, \dots, v_n\}$  será dita  *$g$ -ortonormal* quando  $g(v_i, v_j) \in \{-1, 0, 1\}$ .

Na sequência estaremos interessados em demonstrar que bases  $g$ -ortonormais existem em espaços vetoriais de dimensão finita. Indicaremos por  $W(u_0, v_0)$  o

subespaço de  $V$  gerado por dois vetores  $u_0$  e  $v_0$ . O complemento ortogonal de  $W(u_0, v_0)$  é o conjunto  $W(u_0, v_0)^\perp = \{v \in V; g(v, w) = 0 \forall w \in W(u_0, v_0)\}$ . Observe que sendo  $g$  não identicamente nula deve existir dois vetores satisfazendo  $g(u_0, v_0) \neq 0$  e, sendo assim, o subespaço gerado tem dimensão dois pois os vetores são necessariamente linearmente independentes.

**Lema 10.7.1** *Se  $g$  é antisimétrica e  $g(u_0, v_0) \neq 0$ , então*

$$V = W(u_0, v_0) \boxplus W(u_0, v_0)^\perp \quad e \quad Nuc\,g \subset W(u_0, v_0)^\perp.$$

**Demonstração** Utilizaremos a mesma idéia do processo de ortogonalização de Gram-Schmidt. Dado um vetor  $u$  em  $V$  considere o vetor

$$u^\perp = u - \frac{g(u, v_0)}{g(u_0, v_0)}u_0 + \frac{g(u, u_0)}{g(u_0, v_0)}v_0.$$

Com uma simples avaliação estabelecemos as igualdades

$$g(u^\perp, u_0) = 0 = g(u^\perp, v_0),$$

mostrando que  $V = W(u_0, v_0) + W(u_0, v_0)^\perp$ . Agora consideremos um vetor  $u \in W(u_0, v_0) \cap W(u_0, v_0)^\perp$ . Nesse caso,  $u = \lambda u_0 + \mu v_0$ , com  $\lambda, \mu \in \mathbb{K}$ , e

$$0 = g(u, v_0) = g(\lambda u_0 + \mu v_0, v_0) = \lambda g(u_0, v_0).$$

Daí concluímos que  $\lambda = 0$  desde que, por hipótese,  $g(u_0, v_0) \neq 0$ . Do mesmo modo, fazendo a avaliação  $g(u, u_0)$  obtemos o valor  $\mu = 0$ . Em resumo,

$$W(u_0, v_0) \cap W(u_0, v_0)^\perp = \{o\},$$

provando que a soma dos subespaços é uma soma direta  $g$ -ortogonal. A inclusão deixaremos como exercício.  $\square$

**Proposição 10.7.1** *Seja  $g$  uma forma bilinear antisimétrica num espaço vetorial de dimensão finita  $V$ . Então existem vetores  $u_1, v_1, \dots, u_k, v_k \in V$  tais que*

$$V = W(u_1, v_1) \boxplus W(u_2, v_2) \boxplus \dots \boxplus W(u_k, v_k) \boxplus Nuc\,g$$

*em que a soma direta é  $g$ -ortogonal e  $g(u_i, v_i) = -1$ . Em consequência,  $V$  admite uma base  $g$ -ortonormal.*

**Demonstração** Para evitar trivialidades iremos assumir que  $g$  não é identicamente nula, isso implica que  $\dim V \geq 2$ . A demonstração será feita por indução sobre a dimensão de  $V$ . Se  $\dim V = 2$ , escolhemos quaisquer vetores  $u_1, v_1 \in V$  satisfazendo a condição  $g(u_1, v_1) \neq 0$ . A independência linear dos vetores garante que  $V = W(u_1, v_1)$ . A menos de substituir o vetor  $u_1$  pelo vetor  $\frac{1}{g(u_1, v_1)}u$  e de reordenar o conjunto  $\{u_1, v_1\}$ , essa é a decomposição desejada e  $\{u_1, v_1\}$  é uma base  $g$ -ortonormal. Vamos supor que a afirmação considerada seja verdadeira para

qualquer espaço vetorial de dimensão menor ou igual à  $n$ . Seja  $g$  uma forma bilinear antisimétrica em um espaço  $V$  de dimensão  $n + 1$ . Considere dois vetores não nulos, digamos que sejam  $u_k, v_k \in V$ , tais que  $g(u_k, v_k) \neq 0$ . Pelo lema anterior vale a decomposição  $g$ -ortogonal

$$V = W(u_k, v_k)^\perp \boxplus W(u_k, v_k).$$

Note que o subespaço  $W(u_k, v_k)^\perp$  tem dimensão  $n - 1$ , a restrição de  $g$  é antisimétrica e  $\text{Nuc } g \subset W(u_k, v_k)^\perp$ . Por hipótese de indução, decomponemos esse subespaço e obter a decomposição de  $V$  pretendida. A base ordenada

$$\beta = \{u_1, v_1\} \vec{\cup} \cdots \vec{\cup} \{u_k, v_k\} \vec{\cup} \gamma,$$

em que  $\gamma$  é uma base ordenada do núcleo de  $g$ , caso não seja trivial, é uma base  $g$ -ortonormal.  $\square$

**Exercício 10.7.1** Com as hipótese da proposição acima mostre que a representação de  $g$  na base  $\beta$  é a matriz diagonal de blocos

$$[g]_\beta = \text{diag}\{\underbrace{J, J, \dots, J}_{k \text{ vezes}}, [0]\}, \quad \text{onde} \quad J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

e  $[0]$  é a matriz nula  $i_0 \times i_0$ .  $\square$

Um par  $(V, g)$  é chamado de *espaço simplético* se  $V$  é um espaço vetorial real de dimensão finita e  $g$  é uma forma bilinear antisimétrica não degenerada.

### Exercícios propostos 10.7.1

1. Mostre que toda forma bilinear antisimétrica num espaço vetorial unidimensional é identicamente nula.
2. Prove que a representação matricial de uma forma bilinear num espaço vetorial de dimensão finita é uma matriz antisimétrica se, e somente se, a forma é antisimétrica.
3. Seja  $g$  uma forma bilinear antisimétrica num espaço Euclidiano  $(V, \langle \cdot, \cdot \rangle)$ . Construa um operador linear anti-simétrico  $A : V \rightarrow V$  tal que  $g(u, v) = \langle u, A(v) \rangle$ .
4. Prove que o subconjunto  $\mathcal{A}(V, \mathbb{K}) \in \mathcal{B}(V, \mathbb{K})$  formado pelas formas bilineares anti-simétricas num espaço vetorial  $V$  é um subespaço. Qual a dimensão de  $\mathcal{A}(V, \mathbb{K})$  quando  $\dim V = n$ ?
5. Demonstre que o espaço das formas bilineares  $\mathcal{B}(V, \mathbb{K})$  num espaço vetorial  $V$  decompõe-se na soma direta  $\mathcal{B}(V, \mathbb{K}) = \mathcal{S}(V, \mathbb{K}) \oplus \mathcal{A}(V, \mathbb{K})$  na qual a primeira parcela é o subespaço das formas bilineares simétricas e a outra é o subespaço das anti-simétricas.

6. Denote por  $\mathcal{S}(n, \mathbb{R})$  (respectivamente  $\mathcal{A}(n, \mathbb{R})$ ) o subespaço de  $M(n, \mathbb{R})$  formado pelas matrizes simétricas (respectivamente antisimétricas). Seja  $g$  a forma bilinear simétrica em  $M(n, \mathbb{R})$  definida por  $g(N, P) = \text{tr}(NP)$ . Prove as afirmações.
- $g$  é positiva definida em  $\mathcal{S}(n, \mathbb{R})$ .
  - $g$  é negativa definida em  $\mathcal{A}(n, \mathbb{R})$ .
  - $g(N, P) = 0 \Leftrightarrow N \in \mathcal{S}(n, \mathbb{R})$  e  $P \in \mathcal{A}(n, \mathbb{R})$ .
  - Conclua que  $M(n, \mathbb{R}) = \mathcal{S}(n, \mathbb{R}) \oplus \mathcal{A}(n, \mathbb{R})$ .
  - Determine o índice de  $g$ .
7. Determine o índice da forma bilinear simétrica  $g(N, P) = \text{tr}(NP) - \text{tr}(N)\text{tr}(P)$  em  $M(n, \mathbb{R})$ .
8. Demonstre que um espaço vetorial real  $V$  de dimensão  $n$  admite uma estrutura de espaço simplético  $\Leftrightarrow n = 2m$ .
9. Fixados os vetores  $v_0, w_0 \in \mathbb{R}^3$ , prove que  $q : \mathbb{R}^3 \rightarrow \mathbb{R}$ ,  $q(u) = \langle u \wedge v_0, u \wedge w_0 \rangle$  (onde  $\wedge$  indica o produto vetorial) é uma forma quadrática. Determine a matriz de  $q$  na base canônica.
10. Seja  $g$  uma forma bilinear no espaço vetorial  $V$ . Um operador linear  $A : V \rightarrow V$  preserva  $g$  se  $g(A(u), A(v)) = g(u, v)$ .
- Mostre que a composta de dois operadores que preservam  $g$  é um operador que preserva  $g$ .
  - Assuma que  $g$  é não degenerada e demonstre que o conjunto  $\mathcal{O}(g)$  formado por todos os operadores lineares que preservam  $g$  é um grupo com a operação de composição de operadores e  $\det A \in \{-1, 1\}$ .
11. Um operador  $A$  preserva a forma quadrática  $q$  associada a uma forma bilinear simétrica  $g$  se  $q(A(v)) = q(v)$ . Prove que  $A$  preserva a forma bilinear simétrica  $g \Leftrightarrow A$  preserva a forma quadrática associada  $q$ .
12. Considere a forma bilinear em  $\mathbb{R}^2$  definida por  $g((x_1, y_1), (x_2, y_2)) = x_1y_1 - x_2y_2$ .
- Verifique que  $g$  é não degenerada.
  - Calcule o grupo  $\mathcal{O}(g)$  formado pelos operadores que preservam  $g$ .
13. A mesma questão anterior para a forma  $g((x_1, y_1), (x_2, y_2)) = x_1y_2$ .

## Capítulo 11

# Representação canônica (II)

Numa primeira leitura esse capítulo pode ser dispensado sem prejuízo para a compreensão posterior do texto.

Dando continuidade ao estudo de decomposições  $A$ -cíclicas de um operador, examinaremos os casos nos quais o espaço vetorial  $V$  é real e o polinômio minimal é uma potência de um polinômio primo de grau dois em  $\mathbb{R}[t]$ ,

$$\begin{array}{ll} \text{III} & p(t) = (t - \lambda)^2 + \tau^2, \\ \text{IV} & p(t) = [(t - \lambda)^2 + \tau^2]^r, \end{array}$$

em que  $\tau > 0$  e  $r > 1$ . Para isso necessitaremos de um novo conceito.

### 11.1 Complexificação

O processo de complexificação de um espaço vetorial real em tudo é semelhante ao processo para construir os números complexos a partir dos números reais. Ilustremos a construção com um modelo.

**Exemplo 11.1.1** O polinômio primo  $m_A(t) = (t - 1)^2 + 1 \in \mathbb{R}[t]$  é o polinômio minimal do operador linear  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $A(x, y) = (x - y, x + y)$ . Na construção que faremos, o operador complexificado de  $A$  será naturalmente identificado com o operador linear  $A_c : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ ,  $A(z, w) = (z - w, z + w)$ . Note que o polinômio minimal de  $A_c$  é igual ao polinômio minimal de  $A$ , mas agora podemos fatorá-lo em um produto de polinômios lineares distintos, condição suficiente para diagonalizá-lo. Pretendemos utilizar essa propriedade de diagonalização para construir uma decomposição  $A$ -cíclica para o espaço.  $\square$

Como conjunto, o complexificado  $V_c$  de um espaço vetorial real  $V$  é o produto cartesiano  $V_c = V \times V$ . Um elemento do complexificado é um par ordenado  $(u, v)$

mas que será escrito como  $u + iv$ . Para equipar o conjunto  $V_c$  com uma estrutura de espaço vetorial complexo, definimos a adição de vetores e a multiplicação de um vetor por um escalar complexo pelas regras, respectivamente: se  $v = v_1 + iv_2$  e  $w = w_1 + iw_2$  são elementos de  $V_c$  e  $\lambda + i\tau \in \mathbb{C}$  definimos

$$\begin{aligned} v_1 + iv_2 + w_1 + iw_2 &= (v_1 + w_1) + i(v_2 + w_2), \\ (\lambda + i\tau)(v_1 + iv_2) &= (\lambda v_1 - \tau v_2) + i(\tau v_1 + \lambda v_2). \end{aligned}$$

O espaço vetorial complexo  $V_c$  será chamado de *complexificado* de  $V$ . Nessa construção  $V$  fica canonicamente identificado com o subconjunto

$$V = \{v_1 + iv_2 \in V_c; v_2 = 0\},$$

que não é um subespaço complexo de  $V_c$  mas é fechado em relação à adição de vetores. Uma base ordenada  $\beta = \{u_1, u_2, \dots, u_n\}$  de  $V$  é naturalmente identificada com um base ordenada de  $V_c$  pois qualquer vetor  $w \in V_c$  é escrito de maneira única como

$$w = z_1 u_1 + z_2 u_2 + \dots + z_n u_n \quad \text{com } z_i \in \mathbb{C}.$$

Se cada escalar  $z_i$  é real então o vetor  $w$  pertence à  $V$ . Logo, como espaço vetorial complexo as dimensões de  $V$  e  $V_c$  satisfazem a relação

$$\dim_{\mathbb{C}} V_c = \dim_{\mathbb{R}} V.$$

Definimos o *complexificado do operador*  $A : V \rightarrow V$  como sendo o único operador  $A_c : V_c \rightarrow V_c$  que coincide com  $A$  no subconjunto  $V$ . Esse operador existe pois como sabemos um operador é determinado pelos valores nos vetores de uma base do espaço. Portanto, dada uma base ordenada  $\beta = \{u_1, u_2, \dots, u_n\}$  do espaço  $V$  (que é também uma base de  $V_c$ ) consideramos o único operador  $A_c$  tal que

$$[A_c]_{\beta} = [A]_{\beta}.$$

Em outras palavras,

$$A_c(z_1 u_1 + z_2 u_2 + \dots + z_n u_n) = z_1 A(u_1) + z_2 A(u_2) + \dots + z_n A(u_n).$$

Dessas definições segue que

$$A_c(v_1 + iv_2) = A(v_1) + iA(v_2)$$

e que  $V$  é um subconjunto invariante por  $A_c$  pois a restrição ao conjunto  $V$  é o operador linear  $A$ . Como o polinômio minimal de um operador é igual ao polinômio minimal de qualquer representação matricial concluímos que  $m_A(t) = m_{A_c}(t)$ . Conclusões semelhantes valem para os polinômios característicos de  $A$  e  $A_c$ , bem como para os polinômios mínimos de um vetor  $v \in V \subset V_c$  relativos ao operador original e ao operador complexificado. A igualdade dos polinômios é justificada pela inclusão natural  $\mathbb{R}[t] \subset \mathbb{C}[t]$ .

Chamaremos a aplicação

$$\chi : V_c \rightarrow V_c, \quad \chi(v_1 + iv_2) = v_1 - iv_2,$$

de *conjugação complexa* em  $V_c$ . Tal aplicação é quase linear, isto é,

$$\chi(zv) = \bar{z}v, \quad \text{para qualquer } z \in \mathbb{C}$$

e satisfaz a identidade  $\chi^2 \equiv Id_{V_c}$ . Note que o subconjunto  $V \subset V_c$  é caracterizado como sendo o subconjunto dos pontos fixos de  $\chi$ , isto é  $V = \{v \in V_c; \chi(v) = v\}$ . Segue das definições a comutatividade  $\chi \circ A_c = A_c \circ \chi$ .

### Exercícios propostos 11.1.1

1. Seja  $A$  um operador num espaço vetorial real  $V$  de dimensão finita.
  - a) O complexificado de  $A_c : V_c \rightarrow V_c$  é diagonalizável  $\Leftrightarrow$  o polinômio minimal de  $A$  é um produto de polinômios primos dois a dois distintos.
  - b) Se  $W$  é um subespaço de  $V_c$  então  $\chi(W)$  é também um subespaço.
  - c) Se  $V_c = W_1 \oplus W_2$  então  $V_c = \chi(W_1) \oplus \chi(W_2)$ .
  - d)  $\mathcal{C}_{A_c}(\chi(w)) = \chi(\mathcal{C}_{A_c}(w))$ .

## 11.2 Espaço A- simples

Voltemos a um dos objetivos desse capítulo, estudar operadores lineares  $A$  num espaço vetorial  $V$  de dimensão finita sobre  $\mathbb{R}$  cujo polinômio minimal é uma potência do polinômio primo  $p(t) = (t - \lambda)^2 + \tau^2$  com  $\tau > 0$ . Observamos dois fatos. Primeiro. As raízes complexas desse polinômio primo são os escalares  $z = \lambda + i\tau$  e  $z = \lambda - i\tau$ . Segundo. O polinômio  $p(t)$  é o polinômio característico da matriz  $2 \times 2$  com entradas reais

$$R(\lambda, \tau) = \begin{bmatrix} \lambda & -\tau \\ \tau & \lambda \end{bmatrix}.$$

Iniciemos com o estudo do tipo mais simples de tais operadores lineares cujo polinômio minimal é  $m_A(t) = (t - \lambda)^2 + \tau^2$  com  $\tau > 0$ .

**Definição 11.2.1** *Seja  $A$  um operador linear num espaço vetorial real  $V$ . Diremos que  $V$  é  $A$ -simples se os únicos espaços invariantes são os subespaços  $\{0\}$  ou  $V$ .*

**Exercício 11.2.1** Um espaço vetorial real  $V$  de dimensão 1 é  $A$ -simples para qualquer operador  $A : V \rightarrow V$ . Prove essa afirmação.  $\square$

Existe uma forte restrição sobre a dimensão de um espaço vetorial  $A$ -simples.

**Proposição 11.2.1** *Seja  $A$  um operador linear simples num espaço vetorial real  $V$  com  $\dim V \geq 2$ . Se  $V$  é  $A$ -simples então o espaço vetorial tem dimensão dois e seu polinômio minimal é um polinômio primo de grau dois,*

$$m_A(t) = (t - \lambda)^2 + \tau^2 \quad \text{com } \tau > 0.$$

*Em particular, o operador complexificado  $A_c : V_c \rightarrow V_c$  é diagonalizável.*

**Demonstração** Observamos que  $V$  é um espaço cíclico pois se  $v_0 \in V$  é um vetor não nulo então  $\mathcal{C}(v_0)$  é um subespaço não trivial invariante pelo operador, como  $V$  é  $A$ -simples, necessariamente teremos  $V = \mathcal{C}(v_0)$ . Daí concluímos que o polinômio característico do operador é igual ao seu polinômio minimal bem como é igual ao polinômio minimal de qualquer vetor não nulo.

A fatoração primária do seu polinômio característico é potência de um único polinômio primo, isto é,  $p_A(t) = p(t)^r$ , de outro modo a decomposição primária de  $V$  teria mais de uma parcela contrariando a simplicidade. Examinemos o expoente  $r$ . Suponhamos por absurdo que  $r > 1$ . Sendo assim, o vetor  $w = p(A)(v_0)$  é não nulo e o polinômio  $p(t)^{r-1}$  pertence ao ideal  $\mathfrak{S}_w$ , uma contradição. Em resumo,  $p_A(t)$  é um polinômio primo. Note que seu grau não pode ser um pois neste caso teríamos  $A = \lambda Id$  implicando que o operador admitiria infinitos subespaços invariantes contrariando novamente a simplicidade. A conclusão da demonstração agora é direta. Como  $p_A(t) \in \mathbb{R}[t]$  é um polinômio primo e não linear então  $p_A(t) = m_A(t) = (t - \lambda)^2 + \tau^2$ , com  $\tau > 0$ . Logo,  $\dim V = \text{grau } p_A(t) = 2$  como desejávamos demonstrar. Finalmente, o complexificado  $A_c$  é diagonalizável pois  $p_{A_c}(t) = p_A(t)$  e esse polinômio fatora-se em  $\mathbb{C}[t]$  como um produto de dois fatores lineares distintos.  $\square$

**Exercício 11.2.2** Demonstre a recíproca da proposição anterior. Se  $\dim V = 2$  e  $m_A(t) = (t - \lambda)^2 + \tau^2$ , então  $V$  é  $A$ -simples.  $\square$

Como sempre, vejamos uma representação matricial "canônica" para um operador  $A$  quando espaço vetorial real é  $A$ -simples de dimensão dois. A construção da base ordenada feita na demonstração da proposição abaixo deve ser repetida nos exemplos.

**Proposição 11.2.2** *Seja  $A$  um operador linear num espaço vetorial real  $V$  de dimensão dois com polinômio minimal  $m_A(t) = (t - \lambda)^2 + \tau^2$ ,  $\tau > 0$ . Então existe uma base ordenada  $\beta \subset V$  tal que*

$$[A]_\beta = \begin{bmatrix} \lambda & -\tau \\ \tau & \lambda \end{bmatrix}.$$

**Demonstração** Como sabemos, o polinômio minimal do complexificado  $A_c : V \rightarrow V$  é igual ao polinômio minimal de  $A$ . Fatorando-se em  $\mathbb{C}[t]$  obtemos

$$m_{A_c}(t) = (t - \mu)(t - \bar{\mu}) \quad \text{em que } \mu = \lambda + i\tau.$$

Por hipótese  $\tau \neq 0$ , logo o operador  $A_c$  é diagonalizável. Escolhamos uma base ordenada conveniente para os nossos propósitos.

Primeiro, verifiquemos que a conjugação complexa aplica uma parcela primária isomorficamente na outra,

$$\chi : Nuc(A_c - \mu Id_{V_c}) \rightarrow Nuc(A_c - \bar{\mu} Id_{V_c})$$

Seja  $w_1 \in V_c$  um autovetor associado ao autovalor  $\mu$ . Recordando que os operadores  $A_c$  e  $\chi$  comutam, calculemos,

$$A_c(\chi(w_1)) = \chi(A_c(w_1)) = \chi(\mu w_1) = \bar{\mu}\chi(w_1).$$

Portanto,  $\chi(w_1) \in V_{\bar{\mu}} = Nuc(A_c - \bar{\mu} Id_{V_c})$ . Isso implica que  $\gamma = \{w_1, \chi(w_1)\}$  é uma base ordenada de  $V_c$  pois sua dimensão é dois. Vamos assumir que  $w_1 = u_1 + iv_1$ . O ponto principal da demonstração é provar que o conjunto  $\beta = \{v_1, u_1\}$  (observe a ordem) é uma base de  $V$ . Note que um outro modo de descrever esta base é utilizando a conjugação complexa

$$v = \frac{1}{2i}(w - \chi(w)) \quad \text{e} \quad u = \frac{1}{2}(w + \chi(w)).$$

Esses dois vetores pertencem ao conjunto  $V$  pois são pontos fixos da conjugação complexa. Utilizando-se da independência linear de  $w_1$  e  $\chi(w_1)$  em  $V_c$  mostramos facilmente que  $v_1$  e  $u_1$  são linearmente independentes sobre  $\mathbb{R}$ . Resta calcular a representação matricial  $[A]_{\beta}$ . Pelas definições temos que

$$\mu w_1 = A_c(w_1) = A(u_1) + iA(v_1).$$

De onde segue que

$$\begin{cases} A(v_1) = \lambda v_1 + \tau u_1 \\ A(u_1) = -\tau v_1 + \lambda u_1 \end{cases}.$$

Com isso, terminamos a demonstração. □

**Exemplo 11.2.1** O polinômio característico do operador linear  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $A(x, y) = (2x + 5y, -x)$  é o polinômio primo  $p_A(t) = (t - 1)^2 + 2^2$ . Portanto, seu polinômio minimal é esse mesmo polinômio. O complexificado do operador é canonicamente identificado com o operador  $A_c : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ ,  $A_c(z, w) = (2z + 5w, -z)$ . Como o polinômio característico fatora-se em  $\mathbb{C}[t]$  num produto de fatores lineares distintos,

$$p_{A_c}(t) = (t - (1 + 2i))(t - (1 - 2i)),$$

podemos diagonalizá-lo. Calculemos um autovetor do complexificado associado ao

autovalor  $\mu = 1 + 2i$  resolvendo o sistema linear  $A_c(z, w) = (1 + 2i)(z, w)$ . Isto é,

$$\begin{cases} 2z + 5w = (1 + 2i) \\ -z = (1 + 2i)w \end{cases}.$$

Encontramos que o vetor  $(1 + 2i, -1) \in \mathbb{C}^2$  é um autovetor associado à  $\mu$ . Escolhendo para base ordenada de  $V$  o conjunto  $\beta = \{(2, 0), (1, -1)\}$  obtemos a representação construída na proposição. Vejamos,

$$[A]_\beta = [Id]_\beta^\alpha [A]_\alpha [Id]_\alpha^\beta = -\frac{1}{2} \begin{bmatrix} -1 & -1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & 5 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}. \quad \square$$

### 11.3 Decomposição cíclica III

**Teorema 11.3.1 (Teorema da decomposição cíclica III)** *Seja  $A$  um operador linear num espaço vetorial real  $V$  de dimensão finita com polinômio mínimo  $m_A(t) = (t - \lambda)^2 + \tau^2$ ,  $\tau > 0$ . Então existem vetores  $v_1, v_2, \dots, v_n \in V$  tais que*

$$\begin{cases} V = \mathcal{C}(v_1) \oplus \mathcal{C}(v_2) \oplus \dots \oplus \mathcal{C}(v_n), \\ \text{grau}(t - \lambda)^2 + \tau^2 = \dim \mathcal{C}(v_1) = \dim \mathcal{C}(v_2) = \dots = \dim \mathcal{C}(v_n). \end{cases}$$

*Essa decomposição  $A$ -cíclica é dimensionalmente única.*

**Demonstração** Como já vimos, o polinômio minimal do complexificado  $A_c : V_c \rightarrow V_c$  fatora-se em dois polinômios lineares distintos,

$$m_{A_c}(t) = (t - \mu)(t - \bar{\mu}) \quad \text{em que} \quad \mu = \lambda + i\tau.$$

Sendo assim, a decomposição primária de  $V_c$  é uma soma direta de dois autoespaços,  $V_c = V_\mu \oplus V_{\bar{\mu}}$ . Uma simples verificação, mostra que as restrições da conjugação complexa  $\chi : V_\mu \rightarrow V_{\bar{\mu}}$  e  $\chi : V_{\bar{\mu}} \rightarrow V_\mu$  estão bem definidas. Na verdade, essas duas aplicações são isomorfismos quase lineares. Portanto, dada uma base ordenada  $\beta = \{w_1, w_2, \dots, w_n\}$  do autoespaço  $V_\mu$ , o conjunto de vetores  $\chi(\beta) = \{\chi(w_1), \chi(w_2), \dots, \chi(w_n)\}$  é uma base ordenada do autoespaço  $V_{\bar{\mu}}$ . Tendo em mãos essas duas bases, consideremos a seguinte base ordenada de  $V_c$ ,

$$\gamma = \{w_1, \chi(w_1)\} \cup \{w_2, \chi(w_2)\} \cup \dots \cup \{w_n, \chi(w_n)\}.$$

É claro que  $V_c = W_1 \oplus W_2 \oplus \dots \oplus W_n$  onde  $W_i$  é o subespaço bidimensional gerado pelo conjunto  $\gamma_i = \{w_i, \chi(w_i)\}$ . Logo

$$\dim_{\mathbb{C}} V_c = 2n = \dim_{\mathbb{R}} V.$$

Recordamos que cada  $W_i$  é um subespaço invariante por  $\chi$  pois  $\chi(\gamma_i) = \gamma_i$ , como também é invariante por  $A_c$  pois a base  $\gamma_i$  é formada de autovetores do complexificado. Finalmente, provemos que existe a decomposição  $A$ -cíclica de  $V$ .

Iremos provar que  $V = V_1 \oplus V_2 \oplus \cdots \oplus V_n$  em que  $V_i = W_i \cap V$ . Inicialmente observamos que  $V_i$  é um espaço vetorial real pois se  $u, v \in V_i$  e  $\rho \in \mathbb{R}$ , a combinação linear  $u + \rho v$  pertence ao subespaço  $W_i \subset V_c$  como também pertence ao conjunto  $V$  pois ele um espaço vetorial real. Note que os vetores

$$v_i = \frac{1}{2i}(w_i - \chi(w_i)) \quad \text{e} \quad u_i = \frac{1}{2}(w_i + \chi(w_i))$$

são pontos fixos da conjugação complexa  $\chi$  e pertencem ao subespaço  $W_i \subset V_c$ , portanto são vetores em  $V$ . A independência linear do conjunto  $\beta_i = \{v_i, u_i\}$  sobre  $\mathbb{R}$  é consequência da independência linear sobre  $\mathbb{C}$ . Isso implica que  $\dim_{\mathbb{R}} V_i \geq 2$ . Agora, como  $V_c$  é uma soma direta dos  $W_i$ 's, a condição

$$W_i \cap (W_1 \cap W_2 \cap \cdots \cap W_{i-1}) = \{o\}$$

implica que vale a condição

$$(W_i \cap V) \cap ((W_1 \cap V) + (W_2 \cap V) + \cdots + (W_{i-1} \cap V)) = \{o\}.$$

Portanto temos uma soma direta  $V_1 \oplus V_2 \oplus \cdots \oplus V_n \subset V$ . Para demonstrar que nessa inclusão vale a igualdade, calculemos as dimensões lembrando que  $\dim_{\mathbb{R}} V \geq 2$ ,

$$2n = \dim_{\mathbb{C}} V_c = \dim_{\mathbb{R}} V \geq \sum_{i=1}^n \dim_{\mathbb{R}} V_i \geq 2n.$$

Daí seguem as afirmações:  $\dim_{\mathbb{R}} V = 2$ , o conjunto  $\beta_i = \{v_i, u_i\} \subset V_i$  é uma base e  $V = V_1 \oplus V_2 \oplus \cdots \oplus V_n$ . Por outro lado essa soma direta é invariante por  $A$  pois ela é invariante por  $A_c$  e o operador complexificado restrito ao subconjunto  $V \subset V_c$  coincide com o operador  $A$ . Finalmente, a soma direta construída é uma decomposição cíclica pois o polinômio mínimo de  $v_i$  é  $m_{v_i}(t) = (t - \lambda)^2 + \tau^2$ ,  $\tau > 0$  desde que ele divide  $m_A(t)$ . Como  $\mathcal{C}(v_i)$  é um subespaço de dimensão dois de um espaço de mesma dimensão, obtemos que  $V_i = \mathcal{C}(v_i)$ .  $\square$

Examinemos a questão da representação matricial de um operador linear  $A : V \rightarrow V$  em que  $V$  é um espaço vetorial real de dimensão finita e o polinômio minimal do operador é do tipo  $m_A(t) = (t - \lambda)^2 + \tau^2$ , com  $\tau > 0$ . Como vimos no teorema acima, podemos obter vetores  $v_1, v_2, \dots, v_n \in V$  tais que

$$\begin{cases} V = \mathcal{C}(v_1) \oplus \mathcal{C}(v_2) \oplus \cdots \oplus \mathcal{C}(v_n), \\ \text{grau}(t - \lambda)^2 + \tau^2 = \dim \mathcal{C}(v_1) = \dim \mathcal{C}(v_2) = \cdots = \dim \mathcal{C}(v_n). \end{cases}$$

Os geradores dos subespaços  $A$ -cíclicos são elementos de uma base ordenada  $\beta_i = \{v_i, u_i\}$  de  $\mathcal{C}(v_i)$  e foram escolhidos utilizando a conjugação complexa,

$$v_i = \frac{1}{2i}(w_i - \chi(w_i)) \quad \text{e} \quad u_i = \frac{1}{2}(w_i + \chi(w_i)).$$

Para contruir uma representação de  $A$ , o fato principal a ser observado é que as restrições  $A_i : \mathcal{C}(v_i) \rightarrow \mathcal{C}(v_i)$  definem uma estrutura de espaço  $A$ -simples em  $\mathcal{C}(v_i)$ ,

portanto, sem acrescentar outras informações às já utilizadas na construção de uma representação canônica para um operador num espaço  $A$ -simples feita na seção anterior, podemos afirmar que naquela base ordenada  $\beta_i$  temos a representação matricial

$$[A_i]_{\beta_i} = \begin{bmatrix} \lambda & -\tau \\ \tau & \lambda \end{bmatrix}.$$

Logo, com os conhecimentos já disponíveis obtemos a seguinte representação na base ordenada  $\beta = \vec{\cup} \beta_i$  para o operador  $A$ ,

$$[A]_{\beta} = \text{diag}\{\underbrace{R(\lambda, \tau), R(\lambda, \tau), \dots, R(\lambda, \tau)}_{n \text{ vezes}}\},$$

onde, recordamos, estamos utilizando a notação

$$R(\lambda, \tau) = \begin{bmatrix} \lambda & -\tau \\ \tau & \lambda \end{bmatrix}.$$

Nas hipóteses do Teorema da decomposição cíclica III, a base obtida pela construção alí descrita também será chamada de *base de Jordan*.

### Exercícios propostos 11.3.1

1. Para cada item, determine uma base de Jordan para o operador  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  e represente-o nessa base.
  - a)  $A(x, y, z) = (-5x - y, 2x + 3y + 10z, x - y)$ ;
  - b)  $A(x, y, z) = (x + y, -x + y, -4x - 2y + 2z)$ .
2. Seja  $A$  um operador num espaço vetorial real  $V$  de dimensão finita. Diremos que  $V$  é *A-semi-simples* quando  $V = V_1 \oplus V_2 \oplus \dots \oplus V_k$  e cada parcela é  $A$ -simples. Prove as equivalências.
  - a)  $V$  é  $A$ -semi-simples.
  - b) O polinômio minimal de  $A$  é um produto de polinômios primos sem repetições.
  - c) O complexificado  $A_c : V_c \rightarrow V_c$  é diagonalizável.

## 11.4 Decomposição cíclica IV

Finalizando o nosso estudo de decomposições cíclicas apresentaremos o último caso cujas hipótese incluem as hipóteses de teorema anterior quando a potência do polinômio primo considerada foi  $r = 1$ .

**Teorema 11.4.1 (Teorema da decomposição cíclica IV)** *Seja  $A$  um operador linear no espaço vetorial real  $V$  com polinômio minimal  $m_A(t) = [(t - \lambda)^2 + \tau^2]^r$  com  $\tau > 0$ . Então existem vetores  $v_1, v_2, \dots, v_n \in V$  tais que*

$$\begin{cases} V = \mathcal{C}(v_1) \oplus \mathcal{C}(v_2) \oplus \dots \oplus \mathcal{C}(v_n), \\ \text{grau } [(t - \lambda)^2 + \tau^2]^r = \dim \mathcal{C}(v_1) \geq \dim \mathcal{C}(v_2) \geq \dots \geq \dim \mathcal{C}(v_n). \end{cases}$$

*Essa decomposição  $A$ -cíclica é dimensionalmente única.*

**Demonstração** É suficiente demonstrar o teorema para o operador  $B : V \rightarrow V$ ,  $B(v) = (A - \lambda Id)(v)$  cujo polinômio minimal é  $m_B(t) = (t^2 + \tau^2)^r$  desde que uma decomposição  $B$ -cíclica é uma decomposição  $A$ -cíclica.

O polinômio minimal do complexificado  $B_c : V_c \rightarrow V_c$  fatora-se primariamente em  $m_{B_c}(t) = (t - i\tau)^r(t + i\tau)^r$ . Logo,

$$V_c = Nuc(B_c - \tau Id_c)^r \oplus Nuc(B_c + \tau Id_c)^r$$

é a decomposição primária correspondente de  $V_c$ . Pelo Teorema da decomposição cíclica II podemos determinar vetores  $w_1, w_2, \dots, w_n \in Nuc(B_c - \tau Id_c)^r$  tais que

$$\begin{cases} Nuc(B_c - \tau Id_c)^r = \mathcal{C}_{B_c}(w_1) \oplus \mathcal{C}_{B_c}(w_2) \oplus \dots \oplus \mathcal{C}_{B_c}(w_n), \\ r = \dim_{\mathbb{C}} \mathcal{C}_{B_c}(w_1) \geq \dim_{\mathbb{C}} \mathcal{C}_{B_c}(w_2) \geq \dots \geq \dim_{\mathbb{C}} \mathcal{C}_{B_c}(w_n). \end{cases}$$

A próxima etapa é mostrar que podemos transferir essa decomposição para a outra parcela via o isomorfismo quase linear  $\chi$ , e obter a decomposição

$$\begin{cases} Nuc(B_c + \tau Id_c)^r = \mathcal{C}_{B_c}(\chi(w_1)) \oplus \mathcal{C}_{B_c}(\chi(w_2)) \oplus \dots \oplus \mathcal{C}_{B_c}(\chi(w_n)), \\ r = \dim_{\mathbb{C}} \mathcal{C}_{B_c}(\chi(w_1)) \geq \dim_{\mathbb{C}} \mathcal{C}_{B_c}(\chi(w_2)) \geq \dots \geq \dim_{\mathbb{C}} \mathcal{C}_{B_c}(\chi(w_n)), \end{cases}$$

em que

$$\chi : Nuc(B_c - \tau Id_c)^r \rightarrow Nuc(B_c + \tau Id_c)^r$$

é a restrição do operador quase linear de conjugação complexa  $\chi : V_c \rightarrow V_c$ . De fato, a restrição está bem definida. Para demonstrar essa afirmação utilizamos a comutatividade  $B_c \circ \chi = \chi \circ B_c$  com a finalidade de obter indutivamente a identidade funcional

$$\chi \circ (B_c - \tau Id_c)^s = (B_c + \tau Id_c)^s \circ \chi, \quad 0 \leq s \leq r.$$

Feito isso, para cada vetor  $w \in Nuc(B_c - \tau Id_c)^s$  temos as implicações

$$(B_c - \tau Id_c)^s(w) = 0 \Leftrightarrow \chi \circ (B_c - \tau Id_c)^s(w) = 0 \Leftrightarrow (B_c + \tau Id_c)^s \circ \chi(w) = 0,$$

mostrando que  $\chi(w) \in Nuc(B_c + \tau Id_c)^s$ . Como  $\chi^2 \equiv Id_c$  segue que aquela restrição é um isomorfismo quase linear entre as parcelas da decomposição primária de  $V_c$ . Sendo assim,  $\chi$  preserva somas diretas e a decomposição  $B_c$ -cíclica da segunda parcela da decomposição primária está construída, bastando apenas verificar que

$$\chi(\mathcal{C}_{B_c}(w_i)) = \mathcal{C}_{B_c}(\chi(w_i)).$$

Mas, essa igualdade é consequência da comutatividade dos operadores  $B_c$  e  $\chi$ . Para finalizar, resta mostrar que as desigualdades dimensionais são as mesmas. Isso também é simples, pois o polinômio minimal do vetor  $w_i$  é  $m_{w_i}(t) = (t - i\tau)^s$  se, e somente se, o polinômio minimal de  $\chi(w_i)$  é  $m_{\chi(w_i)} = (t + i\tau)^s$ . Essa afirmação segue diretamente da fórmula indutiva vista acima.

Construída a decomposição  $B_c$ -cíclica do complexificado  $V_c$ , passemos à construção da decomposição  $B$ -cíclica do espaço vetorial real  $V$ . Escolhamos o conjunto de vetores

$$\gamma = \{v_1, v_2, \dots, v_n\}, \quad \text{em que} \quad v_i = \frac{1}{2i}(w_i - \chi(w_i)),$$

para ser o conjunto de geradores das parcelas da decomposição  $B$ -cíclica. Com efeito. Note que o polinômio minimal de  $v_i$  é o produto dos polinômios mínimos de  $w_i$  e de seu conjugado complexo  $\chi(w_i)$ ,

$$m_{v_i}(t) = (t - i\tau)^{r_i}(t + i\tau)^{r_i} = (t^2 + \tau^2)^{r_i},$$

pois  $w_i$  e  $\chi(w_i)$  são vetores de parcelas diferentes da decomposição primária de  $V_c$ . Logo,

$$\dim_{\mathbb{R}} \mathcal{C}_B(v_i) = \dim_{\mathbb{C}} \mathcal{C}_{B_c}(w_i) + \dim_{\mathbb{C}} \mathcal{C}_{B_c}(\chi(w_i)).$$

Como  $v_i$  é uma combinação linear de  $w_i$  e  $\chi(w_i)$  e como o operador complexificado  $B_c$  restrito ao conjunto  $V$  coincide com  $B$ , temos a inclusão

$$\mathcal{C}_B(v_i) \subset \mathcal{C}_{B_c}(w_i) \oplus \mathcal{C}_{B_c}(\chi(w_i)).$$

Essa inclusão implica na existência da soma direta

$$V \supset \mathcal{C}_B(v_1) \oplus \mathcal{C}_B(v_2) \oplus \dots \oplus \mathcal{C}_B(v_n).$$

Para mostrar que temos a igualdade é suficiente um cálculo dimensional,

$$\begin{aligned} \dim_{\mathbb{R}} V &\geq \sum \dim_{\mathbb{R}} \mathcal{C}_B(v_i) \\ &= \sum (\dim_{\mathbb{C}} \mathcal{C}_{B_c}(w_i) + \dim_{\mathbb{C}} \mathcal{C}_{B_c}(\chi(w_i))) \\ &= \dim_{\mathbb{C}} V_c. \end{aligned}$$

A igualdade  $\dim_{\mathbb{R}} V = \dim_{\mathbb{C}} V_c$ , implica que a dimensão de  $V$  é igual à soma das dimensões das parcelas  $\mathcal{C}_B(v_i)$ , informação suficiente para concluirmos que

$$V = \mathcal{C}_B(v_1) \oplus \mathcal{C}_B(v_2) \oplus \dots \oplus \mathcal{C}_B(v_n).$$

A ordem decrescente das dimensões é justificada pela igualdade

$$\dim_{\mathbb{R}} \mathcal{C}_B(v_i) = \dim_{\mathbb{C}} \mathcal{C}_{B_c}(w_i) + \dim_{\mathbb{C}} \mathcal{C}_{B_c}(\chi(w_i)),$$

e pelas desigualdades dimensionais na decomposição  $B_c$ -cíclica de  $V_c$ . A unicidade dimensional segue do teorema correspondente. Isso encerra a demonstração do teorema.  $\square$

Construiremos a seguir uma representação matricial "canônica" para um operador  $A : V \rightarrow V$  satisfazendo as hipóteses do Teorema da decomposição cíclica IV. Seguiremos *ipsis litteris* a notação utilizada na demonstração.

Para não sobrecarregar demasiadamente o texto iremos assumir que a decomposição  $A$ -cíclica (e  $B$ -cíclica) obtida tem apenas uma parcela, isto é,

$$\begin{cases} V = \mathcal{C}_B(v), \\ \text{grau } m_B(t) = \dim_{\mathbb{R}} \mathcal{C}_B(v). \end{cases}$$

No caso geral os procedimentos seguidos para uma são os mesmos para todas parcelas. Note que as condições fixadas sobre o número de parcelas pressupõem que o operador complexificado  $B_c : V_c \rightarrow V_c$  determina uma decomposição primária cujas parcelas são  $B_c$ -cíclicas.

$$V_c = \mathcal{C}_{B_c}(w) \oplus \mathcal{C}_{B_c}(\chi(w)).$$

Dito de outra forma,

$$\text{Nuc}(B_c - i\tau Id_c)^r = \mathcal{C}_{B_c}(w),$$

$$\text{Nuc}(B_c + i\tau Id_c)^r = \mathcal{C}_{B_c}(\chi(w)).$$

As restrições dimensionais sobre as parcelas  $B_c$ -cíclicas ficam sendo

$$r = \dim_{\mathbb{C}} \mathcal{C}_{B_c}(w) = \dim_{\mathbb{C}} \mathcal{C}_{B_c}(\chi(w)).$$

Feitas essas hipóteses, considere as bases de Jordan para cada parcela primária,

$$\gamma_c = \{w, (B_c - i\tau)(w), \dots, (B_c - i\tau)^{r-1}(w)\} \subset \mathcal{C}_{B_c}(w),$$

$$\chi(\gamma_c) = \{\chi(w), (B_c + i\tau)(\chi(w)), \dots, (B_c + i\tau)^{r-1}(\chi(w))\} \subset \mathcal{C}_{B_c}(\chi(w)).$$

Para os nossos propósitos é mais conveniente reescrever os dois conjuntos ordenados como

$$\gamma_c = \{w_0, w_1, \dots, w_{r-1}\} \quad \text{e} \quad \chi(\gamma_c) = \{\chi(w_0), \chi(w_1), \dots, \chi(w_{r-1})\}.$$

É claro que uma igualdade entre os conjuntos ordenados significa que o primeiro elemento  $w_0 = w$ , o segundo elemento  $w_1 = (B_c - i\tau Id)(w)$ , etc. A identidade

$$B_c \circ (B_c - i\tau)^j = (B_c - i\tau)^{j+1} + i\tau(B_c - i\tau)^j$$

avaliada no elemento  $w_0$  fornece os seguintes valores nos vetores da base,

$$B_c(w_j) = \begin{cases} i\tau w_j + w_{j+1} & \text{se } j = 1, 2, \dots, r-2 \\ i\tau w_{r-1} & \text{se } j = r-1 \end{cases},$$

e, de modo semelhante, obtemos que

$$B(\chi(w_j)) = \begin{cases} -i\tau\chi(w_j) + \chi(w_{j+1}) & \text{se } j = 1, 2, \dots, r-2 \\ -i\tau\chi(w_{r-1}) & \text{se } j = r-1 \end{cases}.$$

Tendo em mãos essas bases construímos uma base ordenada de  $V$  considerando o conjunto ordenado

$$\beta = \{v_0, u_0\} \vec{\cup} \{v_1, u_1\} \vec{\cup} \cdots \vec{\cup} \{v_{r-1}, u_{r-1}\},$$

onde os vetores são definidos por

$$v_i = \frac{1}{2i}(w_i - \chi(w_i)) \quad \text{e} \quad u_i = \frac{1}{2}(w_i + \chi(w_i)).$$

Os vetores  $v_i$  e  $u_i$  são pontos fixos da conjugação complexa, logo pertencem ao conjunto  $V$ . Por outro lado a independência linear de  $\beta$  sobre  $\mathbb{C}$  garante a independência linear de  $\beta$  sobre  $\mathbb{R}$ . Como

$$\dim_{\mathbb{C}} V_c = 2r = \dim_{\mathbb{R}} V,$$

necessariamente  $\beta$  é uma base de  $V$ . Essa base será chamada de base de Jordan. Recordando que  $B_c$  coincide com  $B$  em  $V$ , calculemos a representação matricial na base ordenada  $\beta$ . Para  $j = 1, \dots, r-2$  temos os valores

$$\begin{aligned} B(v_j) &= \frac{1}{2i}(B_c(w_i) - B_c(\chi(w_i))) \\ &= \frac{1}{2i}(w_{j+i} + i\tau w_j - \chi(w_{j+1}) + i\tau\chi(w_j)) \\ &= v_{j+1} + \tau u_j, \end{aligned}$$

e  $B(v_{r-1}) = \tau u_{r-1}$ . Em resumo,

$$B(v_j) = \begin{cases} \tau u_j + v_{j+1} & \text{se } j = 1, 2, \dots, r-1 \\ \tau u_{r-1} & \text{se } j = r-1 \end{cases}.$$

Do mesmo modo, calculamos que

$$B(u_j) = \begin{cases} -\tau v_j + u_{j+1} & \text{se } j = 1, 2, \dots, r-1 \\ -\tau v_{r-1} & \text{se } j = r-1 \end{cases}.$$

Sendo assim, a representação canônica de Jordan para o operador  $A : V \rightarrow V$  fica sendo

$$[A]_{\beta} = [B]_{\beta} + \lambda I,$$

mais explicitamente,



# Capítulo 12

## Matrizes

O leitor familiarizado com os conceitos envolvendo matrizes, se desejar, pode dispensar a leitura deste capítulo. Aquí foram relacionadas todas as definições e propriedades que são utilizadas no texto. Uma matriz é um objeto matemático cuja existência é independente do conceito de transformação linear, embora possua uma relação estreita com tais aplicações, matrizes também são largamente utilizadas para cálculos em várias áreas Científicas. Aquítambém os únicos corpos considerados são o corpo dos números complexos e o corpo dos números reais.

### 12.1 Matrizes

Uma *matriz*  $m \times n$  com entradas no corpo  $K$  é uma sequência de escalares  $N = (a_{ij})$  com  $a_{ij} \in K$  organizada na forma ao lado. Resumiremos a apresentação de uma matriz em  $N = [a_{ij}]$ . O índice  $i$  de  $a_{ij}$  indica a linha na qual a entrada encontra-se e o índice  $j$  indica a coluna.

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

A notação  $N_i$  indicará a  $i$ -ésima matriz linha de  $N$ , a matriz  $1 \times n$  cujas entradas são as entradas da  $i$ -ésima linha de  $N$ . Da mesma forma,  $N^j$  indicará a  $j$ -ésima matriz coluna de  $N$ , a matriz  $m \times 1$  formada pelas entradas da  $j$ -ésima coluna de  $N$ . Mais claramente,

$$N_i = [ a_{i1} \quad a_{i2} \quad \cdots \quad a_{in} ] \quad \text{e} \quad N^j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}.$$

Induzimos uma estrutura de espaço vetorial no conjunto das matrizes  $m \times n$  com entradas em  $K$ , conjunto este denotado por  $M(m \times n, K)$ , definido a adição

de matrizes e a multiplicação de uma matriz por um escalar, respectivamente por

$$\begin{aligned} N + P &= [a_{ij} + b_{ij}], \\ \lambda N &= [\lambda a_{ij}], \end{aligned}$$

em que  $N = [a_{ij}]$ ,  $P = [b_{ij}] \in M(m \times n, K)$  e  $\lambda \in K$ . O vetor nulo do espaço é a matriz identicamente nula, isto é, a matriz com todas as entradas nulas, e  $-N = [-a_{ij}]$ . Com esta estrutura  $M(m \times n, K)$  torna-se um espaço vetorial de dimensão  $mn$ . Chamaremos de base canônica ao conjunto  $\alpha = \{E_{11}, E_{12}, \dots, E_{mn}\}$ , onde  $E_{ij}$  denota a matriz cuja  $ij$ -ésima entrada vale 1 e todas outras entradas são nulas. A definição da estrutura linear estabelece uma identificação natural entre o corpo  $K$  como espaço vetorial sobre si mesmo e o espaço das matrizes  $1 \times 1$  com entradas em  $K$ .

**Definição 12.1.1** *Seja  $N = [a_{ij}] \in M(m \times n, K)$ .*

- a) *A transposta de  $N$  é a matriz  $N^t = [a_{ij}^t] \in M(n \times m, K)$  em que  $a_{ij}^t = a_{ji}$ .*  
 b) *A adjunta de  $N$  é a matriz  $N^* = [a_{ij}^*] \in M(n \times m, K)$  tal que  $a_{ij}^* = \overline{a_{ji}}$ .*

Note que a matriz transposta e a matriz adjunta de uma matriz  $m \times n$  é uma matriz  $n \times m$  e quando o corpo com o qual estamos trabalhando é o corpo dos números reais é evidente que  $N^t = N^*$ .

**Exemplo 12.1.1** Ilustraremos as últimas definições com alguns cálculos e nenhum comentário. Considere as matrizes  $2 \times 3$  com entradas complexas

$$N = \begin{bmatrix} 0 & 1 & 1+i \\ 1 & i & 3+i \end{bmatrix} \quad \text{e} \quad P = \begin{bmatrix} 2-i & 0 & 4 \\ 3i & 1 & 2 \end{bmatrix}.$$

Seguindo as definições temos

$$\begin{aligned} N + P &= \begin{bmatrix} 2-i & 1 & 5+i \\ 1+3i & 1+i & 5+i \end{bmatrix}, & 2N &= \begin{bmatrix} 0 & 2 & 2+2i \\ 2 & 2i & 6+2i \end{bmatrix}, \\ N^t &= \begin{bmatrix} 0 & 1 \\ 1 & i \\ 1+i & 3+i \end{bmatrix} & \text{e} & N^* &= \begin{bmatrix} 0 & 1 \\ 1 & -i \\ 1-i & 3-i \end{bmatrix}. \end{aligned} \quad \square$$

Dadas duas matrizes

$$N = [a_{ij}] \in M(m \times n, K) \quad \text{e} \quad P = [b_{ij}] \in M(n \times r, K),$$

utilizamos o fato dos comprimentos das linhas de  $N$  ser igual ao comprimento das colunas de  $P$ , para construir uma nova matriz  $NP = [c_{ij}] \in M(m \times r, K)$ , chamada de produto de  $N$  por  $P$ , uma matriz  $m \times r$  cujas entradas são definidas pela regra

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

**Exemplo 12.1.2** Efetuando o produto das seguintes matrizes  $3 \times 2$  e  $2 \times 2$

$$N = \begin{bmatrix} 1 & 2 \\ 0 & -1 \\ 3 & 0 \end{bmatrix} \quad \text{e} \quad P = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$$

obtemos a matriz

$$NP = \begin{bmatrix} 1 \times 1 + 2 \times 2 & 1 \times 1 + 2 \times 0 \\ 0 \times 1 - 1 \times 2 & 0 \times 1 - 1 \times 0 \\ 3 \times 1 + 0 \times 2 & 3 \times 1 + 0 \times 0 \end{bmatrix} = \begin{bmatrix} 5 & 1 \\ -2 & 0 \\ 3 & 3 \end{bmatrix}.$$

Ressaltamos que não podemos efetuar o produto na ordem  $P$  e  $N$ .  $\square$

**Exercício 12.1.1** Quando for possível efetuar as operações, demonstre as seguintes identidades matriciais, onde  $N$ ,  $P$  e  $Q$  são matrizes com entradas no corpo  $K$ .

- $N(P + Q) = NP + NQ$ .
- $(P + Q)N = PN + QN$ .
- $(NP)Q = N(PQ)$ .
- $(\lambda N)P = \lambda(NP) = N(\lambda P)$  para qualquer escalar  $\lambda \in K$ .

**Solução** Ilustrando a técnica de demonstração provaremos a). Assumamos que

$$N = [a_{ij}] \in M(m \times n, K) \quad \text{e que} \quad P = [b_{ij}], Q = [c_{ij}] \in M(n \times r, K).$$

Por definição de soma de matrizes sabemos que

$$N(P + Q) = [d_{ij}] \quad \text{e} \quad d_{ij} = \sum_{k=1}^n a_{ik}(b_{kj} + c_{kj}).$$

Desenvolvendo o último somatório temos as igualdades

$$d_{ij} = \sum_{k=1}^n (a_{ik}b_{kj} + a_{ik}c_{kj}) = \sum_{k=1}^n a_{ik}b_{kj} + \sum_{k=1}^n a_{ik}c_{kj}.$$

Neste último membro da igualdade, o primeiro somatório é a  $ij$ -ésima entrada de  $NP$  e o segundo somatório é a  $ij$ -ésima entrada de  $NQ$ . Portanto temos que  $N(P + Q) = NP + NQ$ .  $\square$

**Exemplo 12.1.3** Demonstremos a identidade matricial  $(NP)^t = P^t N^t$ , em que  $N = [a_{ij}]$  e  $P = [b_{ij}]$  são matrizes  $m \times n$  e  $n \times r$ , respectivamente. Inicialmente escrevamos o produto  $N$  por  $P$ ,

$$NP = [c_{ij}] \quad \text{com} \quad c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

Por definição de transposta, valem as relações

$$(NP)^t = [c_{ij}^t] \text{ onde } c_{ij}^t = c_{ji} \quad \text{e} \quad c_{ji} = a_{j1}b_{1i} + a_{j2}b_{2i} + \cdots + a_{jn}b_{ni}.$$

Por outro lado, calculemos as entradas de  $P^tN^t = [d_{ij}]$ ,

$$\begin{aligned} d_{ij} &= b_{i1}^t a_{1j}^t + b_{i2}^t a_{2j}^t + \cdots + b_{in}^t a_{nj}^t \\ &= b_{1i} a_{j1} + b_{2i} a_{j2} + \cdots + b_{ni} a_{jn} \\ &= a_{j1} b_{1i} + a_{j2} b_{2i} + \cdots + a_{jn} b_{ni}. \end{aligned}$$

Portanto,  $c_{ij}^t = d_{ij}$ , como queríamos demonstrar.  $\square$

Recordamos que estamos indicando por  $N^j$  a  $j$ -ésima matriz coluna e por  $N_i$  a  $i$ -ésima matriz linha de  $N \in M(m \times n, \mathbb{K})$ . O *posto das colunas* de  $N$  é definido como sendo a dimensão do subespaço de  $M(m \times 1, \mathbb{K})$  gerado pelas matrizes  $\{N^1, N^2, \dots, N^n\}$ . Em outras palavras, o posto das colunas de  $N$  é o número máximo de colunas linearmente independentes. Da mesma forma definimos o *posto das linhas* de  $N$  como sendo a dimensão do subespaço de  $M(1 \times n, \mathbb{K})$  gerado pelas matrizes  $\{N_1, N_2, \dots, N_m\}$ . Na última seção do Capítulo 3, existe a demonstração de uma propriedade, de certa forma, surpreendente: o posto das colunas é igual ao posto das linhas.

### Exercícios propostos 12.1.1

1. Verifique que  $(N^t)^t = N$  e  $(N^*)^* = N$  para qualquer matriz  $N \in M(m \times n, \mathbb{K})$ .
2. É possível um produto de matrizes ser igual a matriz nula sem que um dos fatores ser a matriz nula?
3. Prove que a aplicação  $\Phi : M(m \times n, \mathbb{K}) \rightarrow M(n \times m, \mathbb{K})$ ,  $\Phi(N) = N^t$ , é um isomorfismo linear. Qual a sua inversa?
4. Aplicação  $\Phi : M(m \times n, \mathbb{C}) \rightarrow M(n \times m, \mathbb{C})$ ,  $\Phi(N) = N^*$  é um isomorfismo linear?

## 12.2 Matrizes quadradas

Uma *matriz quadrada* é uma matriz cujas linhas têm o mesmo comprimento das colunas. Simplificaremos a notação indicando o espaço vetorial das matrizes  $n \times n$  por  $M(n, \mathbb{K})$ . Estes espaços vetoriais são mais ricos em propriedades algébricas do que os espaços de matrizes não quadradas, a diferença fica por conta do produto de matrizes. Ao efetuarmos um produto entre dois elementos de  $M(n, \mathbb{K})$  obtendo uma outra matriz neste mesmo espaço. Observamos que o produto de matrizes é

não comutativo. Além disso a forma quadrada da matriz permite destacar vários tipos especiais de matrizes, como veremos na sequência.

A *diagonal* de uma matriz  $N \in M(n, \mathbb{K})$  é a subsequência formada pelas *ii*-ésimas entradas,  $(a_{11}, a_{22}, \dots, a_{nn})$ . Diremos que  $N$  é uma *matriz diagonal* quando toda entrada não pertencente à diagonal principal tem o valor zero. Esquemáticamente uma matriz diagonal tem a forma indicada ao lado.

$$N = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}.$$

Algumas vezes indicaremos uma matriz diagonal de modo mais conciso,

$$N = \text{diag}\{a_{11}, a_{22}, \dots, a_{nn}\}.$$

Um caso particular de matriz diagonal é a matriz identidade denotada por  $I$  e definida como

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

Quando desejamos enfatizar o tamanho  $n \times n$  da matriz identidade utilizamos uma indexação do tipo  $I_n$ . Um outro modo prático de indicar a matriz identidade é com o uso do *delta de Kronecker*,

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

O delta de Kronecker permite-nos escrever a matriz identidade como  $I_n = [\delta_{ij}]$ .

**Exemplo 12.2.1** A matriz identidade  $I_n$  possui uma propriedade especial. Para qualquer matriz quadrada  $n \times n$ ,  $N = [a_{ij}]$ , valem as identidades matriciais  $NI_n = N = I_n N$ . A demonstração desta afirmação segue diretamente da definição de produto de matrizes. Provaremos apenas a igualdade  $NI_n = N$ , a segunda igualdade é feita de modo análogo. Escrevamos

$$NI_n = [c_{ij}] \quad \text{em que} \quad c_{ij} = a_{i1}\delta_{1j} + a_{i2}\delta_{2j} + \cdots + a_{in}\delta_{nj}.$$

Observe que única parcela não nula da soma que descreve  $c_{ij}$  é exatamente quando  $i = j$ . Sendo assim, obtemos  $c_{ij} = a_{ij}$ , significando que  $NI_n = N$ .  $\square$

**Definição 12.2.1** Uma matriz  $N \in M(n, \mathbb{K})$  é invertível, ou não singular, se existir uma matriz  $P \in M(n, \mathbb{K})$  tal que  $PN = I_n = NP$ . Quando existe, a matriz  $P$  é chamada de inversa de  $N$  e denotada por  $N^{-1}$ .

**Exercício 12.2.1** Mostre as afirmações sobre matrizes quadradas invertíveis.

- a) Quando uma matriz quadrada tem inversa, a inversa é única.
- b) Se  $N$  e  $P$  são invertíveis, então  $NP$  é invertível e  $(NP)^{-1} = P^{-1}N^{-1}$ .
- c) Sejam  $N$  e  $P$  duas matrizes  $n \times n$ . Prove que  $NP = I_n \Leftrightarrow I_n = PN$ .  $\square$

As afirmações sobre determinantes contidas nesse parágrafo estão demonstradas no capítulo seguinte. Faremos apenas um breve resumo dos fatos necessário para o desenvolvimento do texto, antecipando um algoritmo para inversão de matrizes. Indicamos por  $N_{\widehat{ij}}$  a  $ij$ -ésima matriz reduzida de  $N \in M(n, \mathbb{K})$ , isto significa que a matriz  $N_{\widehat{ij}}$  é a matriz  $(n-1) \times (n-1)$  obtida de  $N$  por supressão da  $i$ -ésima linha e da  $j$ -ésima coluna. A matriz dos cofatores de  $N$  é a matriz  $n \times n$  definida como

$$C_N = [c_{ij}] \quad \text{em que} \quad c_{ij} = (-1)^{i+j} \det N_{\widehat{ij}}.$$

A adjunta clássica de  $N$  a matriz tranposta da matriz dos cofatores,  $\text{adj}(N) = C_N^t$ . Um resultado bem conhecido da Teoria dos determinantes afirma que

$$N \text{adj}(N) = (\det N)I_n = \text{adj}(N)N.$$

Daí segue imediatamente a afirmação

$$N \in M(n, \mathbb{K}) \quad \text{é invertível se, e somente se,} \quad \det N \neq 0.$$

Mais ainda, a inversa de  $N$ , se existe, é

$$N^{-1} = \frac{1}{\det N} \text{adj}(N).$$

**Exercício 12.2.2** Calcule a adjunta clássica da matriz

$$N = \begin{bmatrix} a & b \\ c & d \end{bmatrix}. \quad \square$$

**Definição 12.2.2** Sejam  $N, P \in M(n, \mathbb{K})$ . Diremos que  $N$  é conjugada à  $P$  quando existe uma matriz invertível  $R \in M(n, \mathbb{K})$  tal que  $P = RNR^{-1}$ .

### Exercícios propostos 12.2.1

1. A título de apresentação definimos vários subconjunto de  $M(n, \mathbb{K})$  que são sub-espacos vetoriais, cujas verificações ficarão aos cuidados do leitor. Calcule também as dimensões. *Matrizes de traço nulo; Matrizes triangulares superiormente; Matrizes triangulares inferiormente.*
2. Calcule as potências  $N^i$ ,  $i = 0, 1, 2, \dots$ , das matrizes.

$$a) N = \begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix}. \quad b) N = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}. \quad c) N = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

3. Prove que conjugação é uma relação de equivalência em  $M(n, \mathbb{K})$ .
4. Se  $N$  é conjugada à  $P$ , demonstre que valem as afirmações.
  - a)  $N^i$  é conjugada à  $P^i$ , para todo inteiro  $i \geq 0$ .
  - b) Se  $N$  é invertível então  $P$  é invertível.
  - c) Se  $N$  é invertível então  $N^i$  é conjugada à  $P^i$ , para todo inteiro  $i \in \mathbb{Z}$ .
5. Se  $N \in M(n, \mathbb{K})$  é invertível então  $N^t$  e  $N^*$  são invertíveis. Explícite as inversas.
6. Por que o produto de matrizes não induz uma estrutura de grupo em  $M(n, \mathbb{K})$ ?
7. Considere a aplicação  $A : M(2, \mathbb{K}) \rightarrow M(2, \mathbb{K})$ ,  $A(P) = PN_0$ , onde  $N_0$  é a matriz ao lado. Mostre que  $A$  é um operador linear e calcule uma base para o núcleo e uma para a imagem.
 
$$N = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}.$$
8. Seja  $N_0 \in M(2, \mathbb{K})$  uma matriz invertível. Verifique que  $A : M(2, \mathbb{K}) \rightarrow M(2, \mathbb{K})$ ,  $A(P) = PN_0$ , é um isomorfismo linear e explícite a sua inversa.

### 12.3 Matrizes normais

Embora a apresentação de matrizes normais possa ser feita de forma unificada, sem ressaltar o corpo com o qual estamos trabalhando, faremos a apresentação, primeiro para o corpo dos números reais e depois para o corpo dos números complexos, deixando mais claras as diferenças. Além disso, como a nomenclatura matricial que utilizamos é clássica, temos mais um motivo para separar as definições.

**Matrizes normais com entradas reais.** Diz-se que uma matriz  $N \in M(n, \mathbb{R})$  é chamada de *matriz normal* quando ela comuta com a sua transposta,  $NN^t = N^tN$ . Dentre as matrizes normais com entradas reais destacamos três tipos especiais, a saber:

- a) matriz *simétrica*, se  $N^t = N$ ;
- b) matriz *antisimétrica*, se  $N^t = -N$ ;
- c) matriz *ortogonal*, se  $N^tN = I_n = NN^t$ .

De fato é um exercício elementar verificar que os três tipos de matrizes acima definidos são matrizes normais.

**Matrizes normais com entradas complexas** Uma matriz  $N \in M(n, \mathbb{C})$  é chamada de *matriz normal* quando ela comuta com a sua adjunta,  $NN^* = N^*N$ . Dentre as matrizes normais com entradas complexas destacamos três tipos:

- a) matriz *autoadjunta*, se  $N^* = N$ ;

- b) matriz *antiadjunta*, se  $N^* = -N$ ;
- c) matriz *unitária*, se  $N^*N = I_n = NN^*$ .

**Exercícios propostos 12.3.1**

- Verifique que os seguintes conjuntos são subespaços e calcule as dimensões.
  - Matrizes simétricas:  $\mathcal{S} = \{N \in M(n, \mathbb{R}); N^t = N\}$ ;
  - Matrizes antisimétricas:  $\mathcal{AS} = \{N \in M(n, \mathbb{R}); N^t = -N\}$ ;
  - Matrizes autoadjuntas:  $\mathcal{A}^* = \{N \in M(n, \mathbb{C}); N^* = N\}$ ;
  - Matrizes antiadjuntas:  $\mathcal{AA}^* = \{N \in M(n, \mathbb{C}); N^* = -N\}$ .

- O conjunto das matrizes ortogonais  $n \times n$  é um subespaço de  $M(n, \mathbb{K})$ ?
- Descreva todas as matrizes ortogonais  $Q \in M(2, \mathbb{R})$ .
- Classifique as matrizes normais com entradas complexas.

$$a) N = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \quad b) N = \begin{bmatrix} \cos \theta & -\operatorname{sen} \theta \\ \operatorname{sen} \theta & \cos \theta \end{bmatrix}.$$

$$c) N = \begin{bmatrix} 1 & -3 & 1 \\ -3 & 0 & 5 \\ 1 & 5 & -7 \end{bmatrix}. \quad d) N = \begin{bmatrix} 2 & -3+i & 1 \\ -3-i & 6 & 5-2i \\ 1 & 5+2i & -7 \end{bmatrix}.$$

- Dadas duas matrizes  $N, P \in M(n, \mathbb{R})$ , verifique as afirmações.
  - $NN^t$  é uma matriz simétrica.
  - $NP - P^tN^t$  é uma matriz anti-simétrica.
  - Se  $N$  é normal então todas as potências positivas de  $N$  são normais.
- Prove que o produto de matrizes induz uma estrutura de grupo nos conjuntos.
  - $O(n, \mathbb{R}) = \{Q \in M(n, \mathbb{R}); QQ^t = I_n = Q^tQ\}$ .
  - $U(n, \mathbb{C}) = \{U \in M(n, \mathbb{C}); UU^* = I_n = U^*U\}$ .
- Demonstre que  $M(n, \mathbb{C})$  decompõe-se numa soma direta na qual uma das parcelas é o espaço das matrizes auto-adjuntas e a outra parcela é o espaço das matrizes anti-adjuntas. Enuncie e demonstre uma afirmação análoga para  $M(n, \mathbb{R})$ .

## 12.4 Apêndice: Grupos, anéis e corpos

Para comodidade do leitor, apresentaremos sem detalhes algumas estruturas algébricas utilizadas no texto. Maiores informações podem ser obtidas consultando-se um livro de Álgebra.

**Definição 12.4.1** Um conjunto  $G$  é dito ser um grupo se existe uma operação binária  $G \times G \rightarrow G$ ,  $(x, y) \mapsto x.y$ , chamada de produto, satisfazendo os axiomas:

1. existe um elemento  $e \in G$ , chamado de identidade, tal que  $ex = x = xe$  para qualquer  $x \in G$ ;
2. a operação é associativa:  $(x.y).z = x.(y.z)$  para quaisquer  $x, y, z \in G$ ;
3. para cada  $x \in G$  existe um elemento  $x^{-1} \in G$ , chamado de inverso de  $x$ , tal que  $x.x^{-1} = e = x^{-1}.x$ .

Num grupo só existe uma única identidade, bem como cada elemento só admite um único inverso. Um subconjunto  $H$  de um grupo  $G$  é dito ser um *subgrupo* quando ele satisfaz os axiomas da estrutura de grupo em relação à operação binária induzida por restrição. Diremos que um grupo  $G$  é *comutativo* (ou *Abeliano*) quando  $x.y = y.x$  para quaisquer  $x, y \in G$ . Neste caso, muitas vezes a operação binária é denotada pelo sinal de  $+$ , o elemento identidade pelo dígito 0, sendo chamado de zero, e o elemento inverso de  $x$  é indicado por  $-x$ .

Diremos que uma aplicação entre grupos,  $\Phi : G \rightarrow F$ , é um *homomorfismo* quando  $\Phi(x.y) = \Phi(x).\Phi(y)$ , para quaisquer  $x, y \in G$ . Um homomorfismo injetor e sobrejetor é chamado de *isomorfismo*. O *núcleo* de um homomorfismo de grupos,  $\text{Nuc } \Phi = \{x \in G; \Phi(x) = e\}$ , é um subgrupo de  $G$  e  $\Phi$  é injetor se, e somente se, o núcleo é trivial.

Uma *permutação* do conjunto  $\{1, 2, \dots, n\}$  é uma aplicação injetora (e sobrejetora),  $\sigma : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$ . Denotamos por  $\mathcal{S}_n$  o conjunto de todas as permutações desse conjunto. A composição de funções induz uma estrutura de grupo em  $\mathcal{S}_n$  na qual o elemento identidade é a aplicação identidade e o inverso de  $\sigma$  é a sua função inversa. Uma *transposição* é uma permutação que troca apenas dois elementos de  $\{1, 2, \dots, n\}$  deixando os outros fixos. Um resultado básico da Teoria de grupos nos diz que toda permutação é um produto de transposições não necessariamente de maneira única. Entretanto, uma permutação é um produto de um número par de transposições ou um de número ímpar, exclusivamente. Essa propriedade permite definir um homomorfismo de grupos

$$\text{sin} : \mathcal{S}_n \rightarrow \mathbb{Z}_2, \quad \text{sin}(\sigma) = (-1)^{k(\sigma)},$$

onde  $k(\sigma)$  é um número de transposições utilizadas para descrever  $\sigma$ .

**Definição 12.4.2** Um anel comutativo com identidade é um conjunto não vazio  $\mathbb{K}$  no qual estão definidas duas operações binárias

$$\mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K}, \quad (x, y) \mapsto x + y,$$

$$\mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K}, \quad (x, y) \mapsto x \cdot y,$$

chamadas de adição e multiplicação, respectivamente, satisfazendo os seguintes axiomas para quaisquer  $x, y, z \in \mathbb{K}$ :

1.  $\mathbb{K}$  é um grupo comutativo com relação à operação de adição;
2. a multiplicação satisfaz as propriedades:
  - a) existe um elemento  $1 \in \mathbb{K}$  tal que  $1 \cdot x = x$ ;
  - b) é comutativa:  $x \cdot y = y \cdot x$ ;
  - c) é associativa:  $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ ;
3. a adição e a multiplicação são distributivas:  $x \cdot (y + z) = x \cdot y + x \cdot z$ .

**Definição 12.4.3** Seja  $\mathbb{K}$  um anel comutativo com identidade. Um subconjunto  $\mathfrak{S} \subset \mathbb{K}$  é chamado de ideal quando

1.  $\mathfrak{S}$  é um subgrupo de  $\mathbb{K}$  com relação à operação de adição;
2. se  $x \in \mathfrak{S}$  e  $y \in \mathbb{K}$ , então  $x \cdot y \in \mathfrak{S}$ .

**Definição 12.4.4** Um anel comutativo com identidade  $\mathbb{K}$  é chamado de corpo quando para qualquer  $x \in \mathbb{K}$  distinto de zero existe um elemento  $x^{-1} \in \mathbb{K}$  denominado de inverso multiplicativo de  $x$ , tal que  $x \cdot x^{-1} = 1$ .

Além dos corpos dos números reais e dos números complexos, os únicos anéis considerados nesse texto são os anéis dos polinômios com coeficientes em um daqueles corpos.

**Definição 12.4.5** Sejam  $\mathbb{K}$  e  $\mathbb{L}$  dois corpos. Uma aplicação  $\Phi : \mathbb{K} \rightarrow \mathbb{L}$  é um isomorfismo de corpos se para quaisquer  $x, y \in \mathbb{K}$  a aplicação satisfaz as condições:

1.  $\Phi(x + y) = \Phi(x) + \Phi(y)$ ;
2.  $\Phi(x \cdot y) = \Phi(x) \cdot \Phi(y)$ ;
3.  $\Phi$  é injetora e sobrejetora.

## Capítulo 13

# Determinantes

Esse é um capítulo complementar. Novamente, se o leitor está familiarizado com a teoria pode dispensar a leitura. A partir desse momento, a menos que seja dito o contrário, indicaremos por  $\mathbb{K}$  um anel comutativo com identidade. Esta escolha tem uma explicação. Estamos interessados em apresentar a teoria de determinantes tanto para matrizes cujas entradas são elementos de um corpo quanto para matrizes cujas entradas são polinômios. A demonstração da existência do determinante é indutiva e a apresentação escolhida está baseada num elegante tratamento dado por Artin [1] e detalhado com muita simplicidade por Lang [11].

### 13.1 Determinante

Definiremos determinante como uma aplicação de um espaço de matrizes quadradas com entradas no anel  $\mathbb{K}$  tomando valores neste mesmo anel,  $\det : M(n, \mathbb{K}) \rightarrow \mathbb{K}$ , possuindo algumas propriedades que serão especificadas nos próximos parágrafos. Com o propósito de recordar apresentemos três exemplos.

**Exemplo 13.1.1** 1) Considerando a identificação natural entre as matrizes  $1 \times 1$  e o anel das entradas  $\mathbb{K}$ , definimos o determinante de um escalar como sendo o próprio escalar.

2) Para matrizes  $2 \times 2$  definimos a aplicação  $\det : M(2, \mathbb{K}) \rightarrow \mathbb{K}$  pela regra

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

Verifica-se diretamente da definição as seguintes propriedades, onde  $\lambda \in \mathbb{K}$ .

a)  $\det I_2 = 1$ ;

b)  $\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = 0$ , se as colunas  $N^1$  e  $N^2$  são iguais;

$$c) \det \begin{bmatrix} a_{11} + \lambda b_{11} & a_{12} \\ a_{21} + \lambda b_{21} & a_{22} \end{bmatrix} = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \lambda \det \begin{bmatrix} b_{11} & a_{12} \\ b_{21} & a_{22} \end{bmatrix};$$

$$d) \det \begin{bmatrix} a_{11} & a_{12} + \lambda b_{12} \\ a_{21} & a_{22} + \lambda b_{22} \end{bmatrix} = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \lambda \det \begin{bmatrix} a_{11} & b_{12} \\ a_{21} & b_{22} \end{bmatrix}.$$

As identidades *c)* e *d)* dizem que o determinante como uma aplicação das duas matrizes colunas possui as propriedades de uma transformação linear ao fixarmos uma das colunas. Essa lista de propriedades será utilizada para definirmos a aplicação determinante num espaço vetorial de matrizes quadradas  $n \times n$ .

3) Recordemos a definição de determinante de uma matriz  $3 \times 3$ . A regra que apresentaremos será generalizada para demonstrar indutivamente a existência de um determinante no espaço das matrizes  $n \times n$ . Defina  $\det : M(3, \mathbb{K}) \rightarrow \mathbb{K}$  pela regra conhecida como *desenvolvimento de Laplace por uma linha*, que no nosso caso será o desenvolvimento pela primeira linha,

$$\begin{aligned} \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} &= a_{11} \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} \\ &\quad - a_{12} \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} \\ &\quad + a_{13} \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}. \end{aligned}$$

Verifica-se diretamente que essa aplicação possui as mesmas propriedades: a) o determinante da matriz identidade  $3 \times 3$  tem o valor 1; b) quando duas colunas adjacentes são iguais, o determinante da matriz é nulo; c) e d) o determinante é uma aplicação linear numa coluna quando fixamos as outras colunas.  $\square$

### Exercícios propostos 13.1.1

1. Calcule os determinantes das matrizes com entradas complexas.

$$a) N = \begin{bmatrix} 2 & -2 \\ 4 & 1 \end{bmatrix}. \quad b) N = \begin{bmatrix} i & 0 & 2i \\ -1 & 3 & 1 \\ i & 2 & 0 \end{bmatrix}. \quad c) N = \begin{bmatrix} 1 & 3 & 4 \\ 1 & -1 & 0 \\ 1 & -2 & -1 \end{bmatrix}.$$

2. Calcule os determinantes das matrizes com entradas em  $\mathbb{R}[t]$ .

$$a) N = \begin{bmatrix} 3t - 1 & t - 1 \\ t + 1 & t - 1 \end{bmatrix}. \quad b) N = \begin{bmatrix} 1 & 0 & 0 \\ t & 1 & 0 \\ t & t & 1 \end{bmatrix}.$$

### 13.2 Definição

Recordamos que dada uma matriz  $N = [a_{ij}] \in M(n, \mathbb{K})$  estamos indicando por  $N^j$  a  $j$ -ésima matriz coluna de  $N$ . Ao escrevermos  $N = (N^1, N^2, \dots, N^n)$  ficará subentendido a identificação de  $N$  com um elemento do produto direto

$$\underbrace{M(n \times 1, \mathbb{K}) \times M(n \times 1, \mathbb{K}) \times \cdots \times M(n \times 1, \mathbb{K})}_{n \text{ vezes}}$$

via o isomorfismo linear natural existente entre o espaço  $M(n, \mathbb{K})$  e esse produto direto. Quando for conveniente trocaremos de notação sem mais explicações.

**Definição 13.2.1** *Diz-se que uma aplicação  $\mathcal{D} : M(n, \mathbb{K}) \rightarrow \mathbb{K}$  é um determinante se possuir as seguintes propriedades:*

1.  $\mathcal{D}(I_n) = 1$ ;
2.  $\mathcal{D}(N^1, \dots, N^j, N^{j+1}, \dots, N^n) = 0$  se  $N^j = N^{j+1}$  para algum  $j$ .
3.  $\mathcal{D}(N^1, \dots, N^j + \lambda P, \dots, N^n) = \mathcal{D}(N^1, \dots, N^{j-1}, N^j, N^{j+1}, \dots, N^n) + \lambda \mathcal{D}(N^1, \dots, N^{j-1}, P, N^{j+1}, \dots, N^n)$ , para qualquer matriz coluna  $P \in M(n \times 1, \mathbb{K})$  e qualquer  $\lambda \in \mathbb{K}$ .

Posteriormente demonstraremos que para cada espaço  $M(n, \mathbb{K})$  existe uma única aplicação satisfazendo essas três condições. Antes vejamos algumas propriedades conhecidas sobre o cálculo de determinantes que são consequências da definição.

**Proposição 13.2.1** *Suponha que exista uma aplicação  $\mathcal{D} : M(n, \mathbb{K}) \rightarrow \mathbb{K}$  seja um determinante. Seja  $N \in M(n, \mathbb{K})$ . Então valem as afirmações.*

1. Se uma coluna de  $N$  é nula então  $\mathcal{D}(N) = 0$ .
2. O determinante troca de sinal quando duas colunas adjacentes são permutadas.
3. Quando duas colunas de  $N$  são iguais, o seu determinante é nulo.
4. Se  $P$  é uma matriz obtida somando-se a uma coluna de  $N$  uma combinação linear de outras colunas, então  $\mathcal{D}(P) = \mathcal{D}(N)$ .
5. Se uma coluna de  $N$  é uma combinação linear de outras colunas então  $\mathcal{D}(N) = 0$ .

**Demonstração 1.** Suponha que  $N = (N^1, \dots, N^{j-1}, [0], N^{j+1}, \dots, N^n)$ . Seguem da definição— ao as igualdades,

$$\mathcal{D}(N) = \mathcal{D}(N^1, \dots, N^{j-1}, [0] + [0], N^{j+1}, \dots, N^n) = 2\mathcal{D}(N).$$

De onde obtemos que  $\mathcal{D}(N) = 0$ .

2. Observamos que  $\mathcal{D}(N^1, \dots, N^j + N^{j+1}, N^j + N^{j+1}, \dots, N^n) = 0$  pois duas colunas são iguais. Ao desenvolvermos esse determinante obtemos apenas duas parcelas, quais sejam

$$0 = \mathcal{D}(N^1, \dots, N^j, N^{j+1}, \dots, N^n) + \mathcal{D}(N^1, \dots, N^{j+1}, N^j, \dots, N^n),$$

demonstrando a afirmação 2.

3. Suponha que as colunas  $N^{j_1}$  e  $N^{j_2}$  da matriz  $N$  são iguais. Com uma sequência de transposição de colunas é possível colocar essas duas colunas iguais em posições adjacentes. Pelo item anterior, o determinante da matriz  $P$  obtida no final das transposições difere do determinante de  $N$  possivelmente por sinal. Como  $\mathcal{D}(P) = 0$ , então  $\mathcal{D}(N) = 0$ .

4. Fixemos a coluna  $N^{j_0}$  da matriz  $N = (N^1, N^2, \dots, N^n)$  e escolhamos  $r$  escalares  $a_{j_1}, a_{j_2}, \dots, a_{j_r} \in \mathbb{K}$  indexados por inteiros satisfazendo as condições  $1 \leq j_s \leq n$  e  $j_s \neq j_0$ . Consideremos a matriz

$$P = (N^1, \dots, N^{j_0-1}, N^{j_0} + \sum_{s=1}^r a_{j_s} N^{j_s}, N^{j_0+1}, \dots, N^n).$$

Avaliando-se o determinante de  $P$ , levando-se em consideração a propriedade d3, temos que

$$\mathcal{D}(P) = \mathcal{D}(N) + \sum_{s=1}^r a_{j_s} \mathcal{D}(N^1, \dots, N^{j_0-1}, N^{j_s}, N^{j_0+1}, \dots, N^n).$$

Examinemos as parcelas do somatório. Como a matriz coluna  $N^{j_s}$  em cada parcela é igual a uma outra matriz coluna, pelo item anterior, concluímos que o determinante de cada parcela é nulo, portanto  $\mathcal{D}(P) = \mathcal{D}(N)$ .

5. Vamos assumir que a coluna  $N^j$  da matriz  $N = (N^1, N^2, \dots, N^n)$  é uma combinação linear de outras colunas. Para anular a coluna  $N^j$  basta somar uma combinação linear conveniente das outras colunas, obtendo uma matriz na forma  $P = (N^1, \dots, [0], \dots, N^n)$ . Por itens anteriores segue que  $\mathcal{D}(N) = \mathcal{D}(P) = 0$ .  $\square$

### Exercícios propostos 13.2.1

1. Justifique o valor zero para o determinante das matrizes.

$$a) N = \begin{bmatrix} 1 & 3 & 1 \\ 2 & 0 & 2 \\ -1 & 2 & -1 \end{bmatrix}. \quad b) N = \begin{bmatrix} 2 & 1 & 3 \\ -1 & 3 & 2 \\ 1 & 3 & 4 \end{bmatrix}. \quad c) N = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 0 & 1 \\ -1 & 0 & 2 \end{bmatrix}.$$

### 13.3 Existência

Para construir um determinante  $\det : M(n, \mathbb{K}) \rightarrow \mathbb{K}$  necessitaremos de algumas notações. Denotaremos por  $N_{\widehat{ij}}$  a matriz  $(n-1) \times (n-1)$  obtida de  $N = [a_{ij}] \in M(n, \mathbb{K})$  por supressão da  $i$ -ésima linha e da  $j$ -ésima coluna. Chamaremos  $N_{\widehat{ij}}$  de  $ij$ -ésima matriz reduzida de  $N$ . Recordando a definição de determinante de uma matriz  $N = [a_{ij}] \in M(3, \mathbb{K})$ , é possível reescrever aquela fórmula utilizando-se as matrizes reduzidas,

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11} \det N_{\widehat{11}} - a_{12} \det N_{\widehat{12}} + a_{13} \det N_{\widehat{13}}.$$

**Proposição 13.3.1** *Para cada inteiro  $n > 0$  existe um determinante no espaço das matrizes  $n \times n$ .*

**Demonstração** Como foi comentado anteriormente a demonstração da existência de um determinante para cada inteiro positivo  $n$  é indutiva. Já sabemos que existem determinantes em  $M(1, \mathbb{K})$ ,  $M(2, \mathbb{K})$  e  $M(3, \mathbb{K})$ . Vamos supor por indução que já tenhamos mostrado a existência de um determinante em cada espaço vetorial  $M(r, \mathbb{K})$ ,  $1 \leq r \leq n-1$ . Defina uma aplicação  $\det : M(n, \mathbb{K}) \rightarrow \mathbb{K}$  pela regra conhecida como desenvolvimento de Laplace pela primeira linha,

$$\det [N] = (-1)^{1+1} a_{11} \det N_{\widehat{11}} + (-1)^{1+2} a_{12} \det N_{\widehat{12}} + \cdots + (-1)^{1+n} a_{1n} \det N_{\widehat{1n}}.$$

Mostremos que esta aplicação é um determinante.

1. Se  $I_n = [\delta_{ij}]$  é a matriz identidade  $n \times n$ , é evidente que  $I_{\widehat{11}}$  é a matriz identidade  $(n-1) \times (n-1)$ . Portanto,

$$\begin{aligned} \det I_n &= (-1)^{1+1} \delta_{11} \det I_{\widehat{11}} + (-1)^{1+2} \delta_{12} \det I_{\widehat{12}} + \cdots + (-1)^{1+n} \delta_{1n} \det I_{\widehat{1n}} \\ &= \det I_{\widehat{11}} \\ &= \det I_{n-1}. \end{aligned}$$

Logo,  $\det I_n = 1$ , desde que por hipótese de indução temos que  $\det I_{n-1} = 1$ .

2. Suponha que as colunas  $N^{j_0}$  e  $N^{j_0+1}$  da matriz  $N = [a_{ij}] \in M(n, \mathbb{K})$  sejam iguais. Sendo assim, se  $\mathbb{K} \neq j_0$  as  $1k$ -ésimas matrizes reduzidas de  $N$  possuem duas colunas iguais, implicando por hipótese de indução que

$$\begin{aligned} \det N &= (-1)^{1+1} a_{11} \det N_{\widehat{11}} + (-1)^{1+2} a_{12} \det N_{\widehat{12}} + \cdots + (-1)^{1+n} a_{1n} \det N_{\widehat{1n}} \\ &= (-1)^{1+j_0} a_{1j_0} \det N_{\widehat{1j_0}} + (-1)^{1+j_0+1} a_{1j_0+1} \det N_{\widehat{1j_0+1}}. \end{aligned}$$

A igualdade entre as colunas  $N^{j_0}$  e  $N^{j_0+1}$  implica na igualdade entre as entradas  $a_{1j_0}$  e  $a_{1j_0+1}$  e na igualdade das matrizes reduzidas  $N_{\widehat{1j_0}}$  e  $N_{\widehat{1j_0+1}}$ . Examinado-se os expoentes  $(-1)^{1+j_0}$  e  $(-1)^{1+j_0+1}$  na última equação concluímos que  $\det N = 0$ .

3. Mostremos a multilinearidade da aplicação  $\det : M(n, \mathbb{K}) \rightarrow \mathbb{K}$ . Dados um escalar  $\lambda \in \mathbb{K}$ , uma matriz  $N = [a_{ij}] \in M(n, \mathbb{K})$  e a matriz coluna  $P^{j_0} = [b_{ij_0}] \in M(n \times 1, \mathbb{K})$ , calculemos o determinante de

$$Q = [c_{ij}] = (N^1, \dots, N^{j_0} + \lambda P^{j_0}, \dots, N^n).$$

Para isto, identifiquemos as parcelas das matrizes reduzidas de  $Q$  utilizando a notação  $P = ([0], \dots, [0], P^{j_0}, [0], \dots, [0])$ . Sendo assim, temos as matrizes

$$\begin{aligned} Q_{\widehat{1k}} &= N_{\widehat{1k}} + \lambda P_{\widehat{1k}} \quad \text{se } k \neq j_0, \\ Q_{\widehat{1j_0}} &= N_{\widehat{1j_0}}. \end{aligned}$$

A hipótese de indução sobre a multilinearidade do determinante para matrizes  $(n-1) \times (n-1)$  justifica as igualdades

$$\begin{aligned} \det Q &= \sum_{k=1}^n (-1)^{1+k} c_{1k} \det Q_{\widehat{1k}} \\ &= \left\{ \sum_{\substack{k \neq j_0 \\ k=1, n}} (-1)^{1+k} a_{1k} \det (N_{\widehat{1k}} + \lambda P_{\widehat{1k}}) \right\} + (-1)^{1+j_0} (a_{1j_0} + \lambda b_{1j_0}) \det N_{\widehat{1j_0}} \\ &= \left\{ \sum_{k=1}^n (-1)^{1+k} a_{1k} \det N_{\widehat{1k}} \right\} + \lambda \det (N^1, \dots, P^{j_0}, \dots, N^n) \\ &= \det N + \lambda \det (N^1, \dots, P^{j_0}, \dots, N^n), \end{aligned}$$

como queríamos demonstrar.  $\square$

### Exercícios propostos 13.3.1

1. Calcule o determinante das matrizes.

$$a) N = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 2 & 1 & -1 & 0 \\ 0 & -2 & 0 & 4 \\ 1 & 0 & 2 & 0 \end{bmatrix}. \quad b) N = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

2. Mostre a identidade

$$\det \begin{bmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{bmatrix} = (x-y)(y-z)(z-x).$$

## 13.4 Unicidade

Para mostrar a unicidade do determinante nas matrizes quadradas  $n \times n$  utilizaremos o grupo das permutações de um conjunto com  $n$  elementos. Caso o leitor deseje rever as propriedades deste grupo indicamos o Apêndice no final do texto.

**Proposição 13.4.1** *Um determinante  $\mathcal{D} : M(n, \mathbb{K}) \rightarrow \mathbb{K}$  pode ser descrito por*

$$\mathcal{D}(N^1, N^2, \dots, N^n) = \sum \sin(\sigma) a_{\sigma(1),1} a_{\sigma(2),2} \dots a_{\sigma(n),n},$$

em que  $N = [a_{ij}]$ , o somatório é feito sobre todas as permutações  $\sigma$  dos inteiros  $\{1, 2, \dots, n\}$  e  $\sin(\sigma)$  é o sinal da permutação. Essa expressão só depende das propriedades listadas na Definição de determinante. Em consequência, existe um, e somente um, determinante definido em cada espaço  $M(n, \mathbb{K})$ .

**Demonstração** Fixemos a seguinte notação matricial,

$$E^j = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow \{ \text{i-ésima linha.} \}$$

Note que as colunas da matriz  $N$  são combinações lineares dessas matrizes,

$$\begin{aligned} N^1 &= a_{11}E^1 + a_{21}E^2 + \dots + a_{n1}E^n, \\ N^2 &= a_{12}E^1 + a_{22}E^2 + \dots + a_{n2}E^n, \\ &\dots \\ N^n &= a_{1n}E^1 + a_{2n}E^2 + \dots + a_{nn}E^n. \end{aligned}$$

Utilizando as propriedades listadas na definição de determinante, o valor de  $\mathcal{D}(N)$  pode ser escrito como

$$\begin{aligned} \mathcal{D}(N^1, N^2, \dots, N^n) &= \sum_{\sigma} \mathcal{D}(a_{\sigma(1),1}E^{\sigma(1)}, a_{\sigma(2),2}E^{\sigma(2)}, \dots, a_{\sigma(n),n}E^{\sigma(n)}) \\ &= \sum_{\sigma} a_{\sigma(1),1} a_{\sigma(2),2} \dots a_{\sigma(n),n} \mathcal{D}(E^{\sigma(1)}, E^{\sigma(2)}, \dots, E^{\sigma(n)}) \end{aligned}$$

Como a matriz  $E^{\sigma} = (E^{\sigma(1)}, E^{\sigma(2)}, \dots, E^{\sigma(n)})$  é obtida da matriz identidade,  $I_n = (E^1, E^2, \dots, E^n)$ , por uma sequência de transposições de duas colunas adjacentes e cada transposição de duas colunas adjacentes troca apenas o sinal do determinante mas não o seu valor absoluto, no final do processo se tivermos realizado  $k(\sigma)$  transposições teremos o valor

$$\begin{aligned} \mathcal{D}(E^{\sigma(1)}, E^{\sigma(2)}, \dots, E^{\sigma(n)}) &= (-1)^{k(\sigma)} \mathcal{D}(E^1, E^2, \dots, E^n) \\ &= (-1)^{k(\sigma)} \mathcal{D}(I_n) \\ &= \sin(\sigma). \end{aligned}$$

Logo, um determinante  $\mathcal{D} : M(n, \mathbb{K}) \rightarrow \mathbb{K}$  é descrito pela expressão

$$\mathcal{D}(N^1, N^2, \dots, N^n) = \sum_{\sigma} \sin(\sigma) a_{\sigma(1),1} a_{\sigma(2),2} \dots a_{\sigma(n),n}.$$

Como essa expressão só depende das propriedades de determinante, fica estabelecida a unicidade, pois a existência já foi exibida na seção anterior.  $\square$

### Exercícios propostos 13.4.1

1. Demonstre utilizando indução sobre  $n$  que o determinante de uma matriz triangular superiormente é o produto das entradas da diagonal.
2. Defina indutivamente a aplicação  $\widetilde{\det} : M(n, \mathbb{K}) \rightarrow \mathbb{K}$  por
 
$$\widetilde{\det} [N] = (-1)^{i+1} a_{i1} \widetilde{\det} N_{\widehat{i1}} + (-1)^{i+2} a_{i2} \widetilde{\det} N_{\widehat{i2}} + \dots + (-1)^{i+n} a_{in} \widetilde{\det} N_{\widehat{in}}.$$
 (desenvolvimento de Laplace pela  $i$ -ésima linha). Prove que essa aplicação é um determinante e conclua que  $\det N = \widetilde{\det} N$ .
3. Uma aplicação  $\mathcal{A} : M(n, \mathbb{K}) \rightarrow \mathbb{K}$  é chamada de  $n$ -linear alternada se possui as propriedades,

$$la.1 \quad \mathcal{A}(N^1, \dots, N^j, N^{j+1}, \dots, N^n) = 0, \text{ se } N^j = N^{j+1} \text{ para algum } j.$$

$$la.2 \quad \mathcal{A}(N^1, \dots, N^j + \lambda P, \dots, N^n) = \mathcal{A}(N^1, \dots, N^{j-1}, N^j, N^{j+1}, \dots, N^n) \\ + \lambda \mathcal{A}(N^1, \dots, N^{j-1}, P, N^{j+1}, \dots, N^n),$$

para qualquer matriz coluna  $P \in M(n \times 1, \mathbb{K})$  e qualquer  $\lambda \in \mathbb{K}$ .

Mostre que  $\mathcal{A}(N) = \mathcal{A}(I_n) \det(N)$ .

## 13.5 Propriedades

Veremos agora algumas propriedades sobre determinantes que são muito conhecidas. Iniciaremos mostrando que o determinante da transposta de uma matriz  $N = [a_{ij}] \in M(n, \mathbb{K})$  é igual ao determinante de  $N$ .

**Proposição 13.5.1** *Vale a igualdade  $\det N = \det N^t$ .*

**Demonstração** Como sabemos, o determinante da matriz  $N = [a_{ij}]$  é calculado pela somatório

$$\det N = \sum_{\sigma} \sin(\sigma) a_{\sigma(1),1} a_{\sigma(2),2} \dots a_{\sigma(n),n},$$

somatório é feito sobre todas as permutações do conjunto  $\{1, 2, \dots, n\}$ . Se  $\sigma^{-1}$  é a permutação inversa de  $\sigma$  e  $\sigma(i) = j$  então  $a_{\sigma(i),i} = a_{j,\sigma^{-1}(j)}$ . Portanto, valem as igualdades dos produtos

$$a_{\sigma(1),1} a_{\sigma(2),2} \dots a_{\sigma(n),n} = a_{1,\sigma^{-1}(1)} a_{2,\sigma^{-1}(2)} \dots a_{n,\sigma^{-1}(n)}.$$

Desde que o sinal de uma permutação é igual ao sinal de sua permutação inversa podemos escrever

$$\begin{aligned}\det N &= \sum_{\sigma^{-1}} \sin(\sigma^{-1}) a_{1,\sigma^{-1}(1)} a_{2,\sigma^{-1}(2)} \cdots a_{n,\sigma^{-1}(n)} \\ &= \sum_{\sigma} \sin(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)} \\ &= \det N^t,\end{aligned}$$

como desejávamos provar.  $\square$

O "determinante do produto é o produto dos determinantes". Sejam  $N, P \in M(n, \mathbb{K})$ .

**Proposição 13.5.2**  $\det NP = \det N \det P$ .

**Demonstração** Utilizaremos as notações  $N = [a_{ij}]$ ,  $P = [b_{ij}]$  e  $Q = NP = [c_{ij}]$ . A  $j$ -ésima matriz coluna da matriz produto  $Q$  é uma combinação linear das colunas de  $N$  cujos coeficientes são as entradas de  $P$ , mais precisamente

$$Q^j = b_{1j}N^1 + b_{2j}N^2 + \cdots + b_{nj}N^n.$$

Avaliemos o determinante do produto levando em consideração essa combinação linear,

$$\begin{aligned}\det NP &= \det (Q^1, Q^2, \dots, Q^n) \\ &= \det \left( \sum_{k=1}^n b_{k1}N^k, \dots, \sum_{k=1}^n b_{kn}N^k \right) \\ &= \sum_{\sigma} \det \left( b_{\sigma(1),1}N^{\sigma(1)}, b_{\sigma(2),2}N^{\sigma(2)}, \dots, b_{\sigma(n),n}N^{\sigma(n)} \right) \\ &= \sum_{\sigma} b_{\sigma(1),1} b_{\sigma(2),2} \cdots b_{\sigma(n),n} \det \left( N^{\sigma(1)}, N^{\sigma(2)}, \dots, N^{\sigma(n)} \right).\end{aligned}$$

A matriz  $N = (N^{\sigma(1)}, N^{\sigma(2)}, \dots, N^{\sigma(n)})$  é obtida da matriz  $N = (N^1, N^2, \dots, N^n)$  por uma seqüência de transposições de duas colunas adjacentes. Como cada transposição de duas colunas adjacentes troca o sinal do determinante, no final do processo se tivermos realizado  $k(\sigma)$  transposições obteremos o resultado

$$\det (N^{\sigma(1)}, N^{\sigma(2)}, \dots, N^{\sigma(n)}) = (-1)^{k(\sigma)} \det (N^1, N^2, \dots, N^n).$$

Finalmente, como  $\sin(\sigma) = (-1)^{k(\sigma)}$ , obtemos

$$\det NP = \sum_{\sigma} b_{\sigma(1),1} b_{\sigma(2),2} \cdots b_{\sigma(n),n} \det (N^1, N^2, \dots, N^n) = \det N \det P. \quad \square$$

**Exercícios propostos 13.5.1**

1. Mostre que o determinante de uma matriz  $3 \times 3$ ,  $N = [a_{ij}]$ , tem a forma:  
 $\det N = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{12}a_{23}a_{31} - a_{13}a_{21}a_{32} - a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31}$ .
2. Demonstre as seguintes afirmações sobre uma matriz  $N \in M(n, \mathbb{K})$ .
  - a) Se duas linhas de  $N$  são iguais então  $\det N = 0$ .
  - b) Se  $P$  é obtida de  $N$  adicionando-se a uma linha de  $N$  uma combinação linear de outras linhas então  $\det P = \det N$ .
3. Demonstre que o determinante de uma matriz  $N = [a_{ij}] \in M(n, \mathbb{K})$  também pode ser calculado com o desenvolvimento de Laplace pela  $j$ -ésima coluna,  
 $\det [N] = (-1)^{1+j}a_{1j} \det N_{\widehat{1j}} + (-1)^{2+j}a_{2j} \det N_{\widehat{2j}} + \cdots + (-1)^{n+j}a_{nj} \det N_{\widehat{nj}}$ .
4. Sejam  $\mathbb{K}$  um corpo e  $A : \mathbb{K}^n \rightarrow \mathbb{K}^n$  um operador linear. Se  $\{v_1, v_2, \dots, v_n\}$  é uma coleção de  $n$  vetores em  $\mathbb{K}^n$  demonstre a fórmula a seguir, em que  $\beta$  é uma base ordenada de  $\mathbb{K}^n$ ,  
 $\det ([A(v_1)]_\beta, [A(v_2)]_\beta, \dots, [A(v_n)]_\beta) = \det [A]_\beta ([v_1]_\beta, [v_2]_\beta, \dots, [v_n]_\beta)$ .

**13.6 Adjunta clássica**

Apresentaremos um algoritmo envolvendo as  $ij$ -ésimas matrizes reduzidas de uma matriz  $N = [a_{ij}] \in M(n, \mathbb{K})$  cuja consequência é uma fórmula que explicita quando uma matriz quadrada com entradas num anel  $\mathbb{K}$  admite uma inversa e deixa claro como devemos construir a inversa de uma matriz, caso exista.

O  $ij$ -ésimo cofator da matriz  $N = [a_{ij}]$  é o elemento de  $\mathbb{K}$  definido como

$$c_{ij} = (-1)^{i+j} \det N_{\widehat{ij}},$$

e a *adjunta clássica* de  $N$  é a matriz transposta da matriz dos cofatores,

$$\text{adj}(N) = [c_{ij}]^t.$$

**Exemplo 13.6.1** Calculemos a adjunta clássica da matriz  $N$  com entradas em  $\mathbb{R}[t]$ ,

$$N = \begin{bmatrix} 1 & t & 0 \\ t-1 & t^2 & t+1 \\ -1 & 0 & 2 \end{bmatrix}.$$

Para isso, calculemos os cofatores e formemos a matriz

$$\text{adj}(N) = \begin{bmatrix} \det N_{\widehat{11}} & -\det N_{\widehat{21}} & \det N_{\widehat{31}} \\ -\det N_{\widehat{12}} & \det N_{\widehat{22}} & -\det N_{\widehat{32}} \\ \det N_{\widehat{13}} & -\det N_{\widehat{23}} & \det N_{\widehat{33}} \end{bmatrix} = \begin{bmatrix} 2t^2 & 2t & t^2+t \\ 1-3t & 2 & -t-1 \\ -t^2 & t & t \end{bmatrix}. \square$$

**Exercício 13.6.1** Verifique que  $\text{adj}(N^t) = \text{adj}(N)^t$ . (Mostre  $N_{\widehat{ij}} = N_{\widehat{ji}}^t$ ).  $\square$

A proposição a seguir descreve relações matriciais envolvendo uma matriz, sua adjunta clássica e seu determinante, com várias e importantes conseqüências.

**Proposição 13.6.1**  $\text{adj}(N)N = (\det N)I_n = N\text{adj}(N)$ .

**Demonstração** Escrevamos a  $jk$ -ésima entrada do produto  $\text{adj}(N)N = [d_{ij}]$  onde  $N = [a_{ij}]$  e  $\text{adj}(N) = [c_{ij}^t]$ ,

$$\begin{aligned} d_{jk} &= c_{j1}^t a_{1k} + c_{j2}^t a_{2k} + \cdots + c_{jn}^t a_{nk}. \\ &= (-1)^{1+j} a_{1k} c_{1j} + (-1)^{2+j} a_{2k} c_{2j} + \cdots + (-1)^{n+j} a_{nk} c_{nj} \\ &= (-1)^{1+j} a_{1k} N_{\widehat{1j}} + (-1)^{2+j} a_{2k} N_{\widehat{2j}} + \cdots + (-1)^{n+j} a_{nk} N_{\widehat{nj}}. \end{aligned}$$

Daí concluímos que se  $k = j$ , temos o valor  $d_{jj} = \det N$  pois a última linha é o desenvolvimento de Laplace do determinante de  $N$  pela  $j$ -ésima coluna.

Denote por  $P = [b_{ij}]$  a matriz obtida de  $N$  substituindo a coluna  $N^{j_0}$  pela coluna  $N^k$ , para  $j_0 \neq k$ . Sendo assim valem as igualdades  $\det P = 0$  e  $P_{\widehat{j_0}} = N_{\widehat{j_0}}$ . Calculemos o determinante de  $P$  com o desenvolvimento pela  $j_0$ -ésima coluna,

$$\begin{aligned} 0 &= \det P \\ &= (-1)^{1+j_0} b_{1j_0} \det P_{\widehat{1j_0}} + (-1)^{2+j_0} b_{2j_0} \det P_{\widehat{2j_0}} + \cdots + (-1)^{n+j_0} b_{nj_0} \det P_{\widehat{nj_0}} \\ &= (-1)^{1+j_0} a_{1k} \det N_{\widehat{1j_0}} + (-1)^{2+j_0} a_{2k} \det N_{\widehat{2j_0}} + \cdots + (-1)^{n+j_0} a_{nk} \det N_{\widehat{nj_0}} \\ &= (-1)^{1+j_0} a_{1k} c_{1j_0} + (-1)^{2+j_0} a_{2k} c_{2j_0} + \cdots + (-1)^{n+j_0} a_{nk} c_{nj_0} \\ &= c_{j_0 1}^t a_{1k} + c_{j_0 2}^t a_{2k} + \cdots + c_{j_0 n}^t a_{nk} \\ &= d_{j_0 k}, \end{aligned}$$

mostrando que o produto da adjunta clássica de  $N$  com  $N$  é da forma

$$\text{adj}(N)N = \begin{bmatrix} \det N & & \\ & \ddots & \\ & & \det N \end{bmatrix} = (\det N)I_n.$$

Para finalizar, utilizaremos a identidade  $\text{adj}(N^t) = \text{adj}(N)^t$ ,

$$N\text{adj}(N) = (\text{adj}(N)^t N^t)^t = (\text{adj}(N^t) N^t)^t = ((\det N^t)I_n)^t = (\det N)I_n. \quad \square$$

**Corolário 13.6.1** Assuma que  $\mathbb{K}$  é um corpo. Uma matriz  $N \in M(n, \mathbb{K})$  é invertível  $\Leftrightarrow \det N \neq 0$ . Mais ainda, se  $N$  é invertível então

$$N^{-1} = \frac{1}{\det N} \text{adj}(N) \quad e \quad \det N^{-1} = (\det N)^{-1}.$$

**Exercícios propostos 13.6.1**

1. Se  $ad - bc \neq 0$  mostre que

$$N = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ é invertível e } N^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

2. Calcule a inversa, se existir, da matriz  $N$  com entradas reais.

$$a) N = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}. \quad b) N = \begin{bmatrix} 2 & 10 & 3 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{bmatrix}. \quad c) N = \begin{bmatrix} 1 & -1 & 3 \\ 4 & 2 & 3 \\ 5 & 1 & 6 \end{bmatrix}.$$

3. Calcule  $\det N$ ,  $\text{adj}(N)$  e  $N^{-1}$  quando

$$N = \begin{bmatrix} 2 & 0 & 5 \\ 1 & 2 & 1 \\ -3 & 1 & 3 \end{bmatrix}.$$

4. Dada a matriz por blocos abaixo, onde  $P$  e  $R$  são matrizes quadradas, demonstre que  $\det N = (\det P)(\det R)$ .

$$N = \begin{bmatrix} P & Q \\ 0 & R \end{bmatrix}.$$

3. Prove que o determinante é invariante por conjugação de matrizes.
4. Sejam  $u = (a_1, a_2, a_3)$ ,  $v = (b_1, b_2, b_3)$ ,  $w = (c_1, c_2, c_3) \in \mathbb{R}^3$ . Demonstre que  $\beta = \{u, v, w\}$  é linearmente dependente  $\Leftrightarrow$  o determinante da matriz  $3 \times 3$  cujas colunas são os vetores  $u$ ,  $v$  e  $w$  é igual a zero.
5. Enuncie e demonstre um resultado equivalente para  $n$  vetores em  $\mathbb{R}^n$ .
6. Mostre que  $\text{adj}(PQ) = \text{adj}(Q)\text{adj}(P)$  para quaisquer matrizes quadradas  $P$  e  $Q$ .

**13.7 Regra de Cramer**

Coloquemos o seguinte problema: determinar  $n$  escalares  $x_1, x_2, \dots, x_n$  em um corpo  $\mathbb{K}$  satisfazendo as condições,

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases},$$

onde os coeficientes  $a_{ij}$  e os valores  $b_k$  são escalares conhecidos. Uma expressão como esta é chamada de sistema linear com  $m$  equações e  $n$  incógnitas. Existe uma apresentação matricial muito cômoda para um sistema linear, a saber,

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Em resumo, escrevemos  $NX = P$ . Dessa apresentação matricial de sistemas obtemos uma outra mais apropriada para os nossos objetivos,

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

ou mais compactamente,  $x_1 N^1 + x_2 N^2 + \cdots + x_n N^n = P$ .

No que segue, iremos nos restringir ao estudo de um sistema linear de  $n$  equações com  $n$  incógnitas. Neste caso, a matriz  $N$  é uma matriz quadrada  $n \times n$ , possibilitando utilizar a Teoria de determinantes a fim encontrar os valores das incógnitas, através de um método chamado de Regra de Cramer. Para isso, calculemos o determinante da matriz obtida por substituição da  $i_0$ -ésima coluna de  $N$  pela matriz coluna  $P$ ,

$$\begin{aligned} \det(N^1, \dots, N^{i_0-1}, P, N^{i_0+1}, \dots, N^n) &= \det\left(N^1, \dots, N^{i_0-1}, \sum_j x_j N^j, N^{i_0+1}, \dots, N^n\right) \\ &= \sum_j x_j \det(N^1, \dots, N^{i_0-1}, N^j, N^{i_0+1}, \dots, N^n). \end{aligned}$$

Quando  $j \neq i_0$  o valor do determinante  $\det(N^1, \dots, N^{i_0-1}, N^j, N^{i_0+1}, \dots, N^n)$  é nulo pois duas colunas são iguais. Por outro lado, quando  $j = i_0$  temos a igualdade,

$$\det(N^1, \dots, N^{i_0-1}, N^{i_0}, N^{i_0+1}, \dots, N^n) = \det N.$$

Retornando com estes dados obtemos

$$\det(N^1, \dots, N^{i_0-1}, P, N^{i_0+1}, \dots, N^n) = x_{i_0} \det N.$$

Os comentários acima mostram a Regra de Cramer, registrada na seguinte proposição.

**Proposição 13.7.1** *Suponha que  $NX = P$  é um sistema linear  $n \times n$  no corpo  $\mathbb{K}$ . Se  $\det N \neq 0$  então*

$$x_i = \frac{\det Q}{\det N}, \quad \text{para } i = 1, 2, \dots, n,$$

em que  $Q$  é a matriz obtida pela substituição da  $i$ -ésima coluna de  $N$  pela matriz coluna  $P$ .

**Exercício 13.7.1** Resolva os seguintes sistemas em  $\mathbb{R}$ .

$$a) \begin{cases} 2x + 2y + z = 5 \\ 6x - y + 2z = 1 \\ 2x - 4y = 0 \end{cases} \quad b) \begin{cases} 2x - 6y - z = 0 \\ 4x + 5y + 3z = 11 \\ 3x + 4y + 2z = 0 \end{cases} \quad \square$$

# Bibliography

- [01] Artin, Emil; *Galois theory*; Notre Dame Math. Lectures n 2, (1953).
- [02] Birkhoff, G. & Mac Lane, S.: *A Survey of Modern Algebra*; The Macmillan Comp., New York (1958).
- [03] Boldrini, Costa, Figueiredo & Wetzler; *Álgebra Linear*; (3ª Ed.), Editora Habra.
- [04] Cullen, C.; *Matrices and Linear Transformations*; Add-Wesley Comp. (1966).
- [05] dos Santos, N. M.; *Vetores e Matrizes*; Ao Livro Técnico S/A (1982)
- [06] Garcia, A. & Lequain, Y.; *Álgebra um curso de introdução*; Projeto Euclides.
- [08] Halmos; *Espaços vetoriais de dimensão finita*; Trad. laPenha; Editora Campos, RJ.
- [09] Hoffman, K. & Kunze, R.; *Linear Algebra*; Prentice-Hall (1967).
- [12] Lang, S.; *Linear Algebra*; Addison-Wesley Publish. Comp. (1966).
- [15] Lima, Elon Lages; *Álgebra Linear* (2ª Ed.) IMPA; 1996.
- [16] Serfati, M.; *Exercices de Mathématiques, vol 1*; Ed. Belin (1980).
- [17] Shilov, Georgi E.; *Linear Álgebra*; 1ª Prentice-Hall, Inc. (1971) New Jersey.
- [18] Steinbruch, A.& Winterle,P.; *Álgebra Linear*; 2ª Ed. São Paulo, McGraw-Hill (1987).

# Índice analítico

<b>A</b>	
Adjunta	
clássica	197,211
de uma transformação linear	53
Anel, definição	200
Ângulo	
entre dois vetores	113
polar	135
Aplicação	
identidade	22
$n$ -linear alternada	209
significado	22
Autoespaço	22
Autovalor	
de uma matriz	89
de um operador	89
Autovetor	89
<b>B</b>	
Base	
canônica de $\mathbb{K}^n$	9
de Jordan	91
definição	9
dual	51
$g$ -ortogonal	171
ordenada	16
ortogonal	114
ortonormal	114
<b>C</b>	
Cofator de uma matriz	211
Combinação linear	5
Complemento ortogonal	111
Complexificado	
de um espaço vetorial real	179
de um operador	180
ortogonal de um subespaço	113
Composição de transformações lineares	26
Conjugação complexa	181
Conjunto	
de geradores	7
linearmente dependente	7
linearmente independente	8
ordenado	10
Coordenadas de um vetor numa base	16
Corpo, definição	201
<b>D</b>	
Decomposição	
cíclica	82
cíclica normal	130
normal	126
polar	
de um complexo	135
de um operador	148
primária	72
Delta de Kronecker	51
Desenvolvimento de Laplace	
pela primeira linha	206
Desigualdade	
de Bessel	118
de Schwarz	111,153
triangular	
primeira	111,153
segunda	114,154
Determinante	
definição	204
de uma matriz $2 \times 2$	202
de uma matriz $3 \times 3$	203
Dimensão de um espaço vetorial	13
Distância	149

## ÍNDICE ANALÍTICO

217

### E

Espaço	
A-cíclico	81
A-simples	181
das transformações lineares	26
de dimensão finita	13
de Minkowski	174
dual	48
Euclidiano	114
indecomponível	191
irredutível	191
pseudo Euclidiano	174
quociente	7
simplético	177
trivial	2
unitário	154
vetorial	
conjugado	4
definição	1
Espaços isomorfos	33
Espectro de um operador	140

### F

Fatoração primária de um polinômio	66
Forma bilinear	
antisimétrica	175
definição	162
degenerada	166
não degenerada	166
não singular	166
semi-positiva definida	170
sesqui-linear	163
simétrica	167
singular	166
Forma quadrática	168
Funções	
pares	21
ímpares	21
simétricas	63
Funcional linear	48

### G

Geradores	
de um espaço vetorial	7
de um ideal	61

Gráfico de uma transf. linear	25
Grupo	
Abeliano, comutativo	200
das permutações	200
definição	200
Euclidiano	151

### H

Homotetia	23
-----------	----

### I

Ideal	
definição	201
anulador	
de um operador	59
de um subespaço	68
de um vetor	68
de uma matriz	59
Identidade	
de polarização	111,153
Imagem de uma transformação linear	23
Índice	
de nulidade de uma forma	171

Involução	25
Isometria	150
Isomorfismo	
de grupos	200
linear	33

### L

Lei	
de inércia de Sylvester	175
do paralelogramo	112,154
Lema de Jacobson	97
Linearmente	
dependente	7
independente	8

### M

Matriz	
adjunta	193
antiadjunta	199
antisimétrica	198
autoadjunta	198
coluna	192

de Pauli	175	isogonal	144
de uma transformação linear	35	não negativo	147
diagonalizável	93	nilpotente	95
dos cofatores	211	normal	124,157
linha	192	ortogonal	137
mudança de coordenadas	43	positivo	144
nilpotente	95	que preserva o produto interno	142
normal	198	simétrico	127
ortogonal	198	transposto	120
polinomial	63	unitário	159
positiva	146		
quadrada	195	<b>P</b>	
reduzida	197	Permutação	200
simétrica	198	Perpendicular de um conjunto	54
transposta	193	Polinômio	
unitária	199	característico	61
Matrizes conjugadas	44	conjugado	156
Máximo divisor comum	66	de uma matriz	61
Métrica	149	de uma transformação	61
Movimento rígido	150	minimal	
		de um ideal	61
<b>N</b>		de um operador	62
$n$ -ésima raiz de um operador	73	de um subespaço	68
Nilpotência		de um vetor	68
de uma matriz	95	mônico	60
definição	95	primo	66
Norma	110	reduzível	66
Núcleo		Polinômios relativamente primos	66
à direita de uma forma bilinear	165	Posto de uma transformação linear	31
à esquerda de uma forma bilinear	165	Processo ortog. de Gram-Schmidt	114,154
de uma transformação linear	23	Projeção	
Nulidade de uma transformação linear	31	ortogonal	117
		quociente	25
<b>O</b>		sobre um subespaço	23
Operador linear		Produto	
adjunto	155	Cartesiano	3
antiadjunto	159	de matrizes	193
antisimétrico	136	direto	3
autoadjunto	158	exterior	163
conforme	144	Hermitiano	152
de Gram	124	interno	
de semelhança	151	em espaço vetorial complexo	152
definição	22	em espaço vetorial real	109
diagonalizável	91	canônico do $\mathbb{C}^n$	152
idempotente	68	canônico do $\mathbb{R}^n$	109

tensorial	162	do núcleo e da imagem	31
<b>R</b>		do operador autoadjunto	158
Representação		do posto de uma matriz	56
matricial		espectral	128
de uma forma bilinear	163	Traço de uma matriz	49
de Jordan	91,101,190	Transformação linear	
de uma transformação linear	36	definição	22
vetorial		identicamente nula	23
de um funcional linear	119,154	invertível	27
<b>S</b>		Transformação quase linear	29
Soma		Translação	150
de subespaços	18	Transposição	200
de transformações lineares	26	<b>U</b>	
direta	18	União ordenada de conjuntos	10
direta ortogonal	116	Vetor	
Subespaço		nulo	2
A-cíclico	79	unitário	111
definição	4	Vetores	
invariante	46	g-ortogonais	168
Subgrupo	200	ortogonais	113
<b>T</b>		<b>Símbolos</b>	
Teorema		$[v]_{\beta}^{\gamma}$	17
da classificação das isometrias	151	$\mathbb{R}$	2
da decomposição		$\mathbb{C}$	2
cíclica I	90	$\mathbb{K}^n$	2
cíclica II	98	$\mathbb{K}[t]$	3
cíclica III	184	$M(m \times n, \mathbb{K})$	3
cíclica IV	187	$\mathcal{S}(\mathbb{Z}, \mathbb{R})$	4
cíclica normal	130	$\ell^1(\mathbb{R})$	4
D + N	106	$\mathbb{K}_n[t]$	5
normal	126,157	$\#$	11
polar	148,161	$Im A$	23
primária	70	$Nuc$	23
da representação		$\circ$	26
de um funcional linear	119,154	$\mathcal{L}(V, W)$	26
matricial	37	$A^{-1}$	27
da troca de Steinitz	11	$[A]_{\gamma}^{\beta}$	36
da unicidade dimensional cíclica	83	$[A]_{\gamma}^{\beta}$	36
de Cayley-Hamilton	63	$diag\{N - 1, \dots, N_r\}$	47
de Hadamard	43	$p(t)$	58
de Pitágoras	113	$\mathfrak{S}_A$	59
do índice de Morse	173	$\mathfrak{S}_N$	59
		$m_A(t)$	60

$p_A(t)$	62	$\psi, \Psi$	psi
$\mathfrak{S}_v$	68	$\omega, \Omega$	ômega
$\mathfrak{S}_W$	68		
$\langle , \rangle$	109		
$\  \ $	110		
$   $	111		
$W^\perp$	113		
$\square$	116		
$R(\lambda, \tau)$	133		
$R_\theta$	135		
$\otimes$	162		
$\wedge$	163		
$\mathcal{B}(V, \mathbb{K})$	163		
$\mathcal{A}(V, \mathbb{K})$	177		
$N^t$	193		
$N^*$	193		
$M(m \times n, K)$	193		
$M(n, K)$	196		
$I_n$	196		
$\text{adj}(N)$	197		
$N_{ij}^{\widehat{\phantom{ij}}}$	197		
$N_{ij}^{\widehat{\phantom{ij}}}$	206		
<b>Letras gregas</b>			
$\alpha$	alfa		
$\beta$	beta		
$\gamma, \Gamma$	gama		
$\delta, \Delta$	delta		
$\epsilon, \varepsilon$	epsilon		
$\zeta$	zeta		
$\eta$	eta		
$\theta, \Theta, \vartheta$	teta		
$\iota$	iota		
$\kappa$	capa		
$\lambda, \Lambda$	lambda		
$\mu$	mu		
$\nu$	ni		
$\xi, \Xi$	qui		
$\emptyset$	o		
$\pi, \Pi, \varpi$	pi		
$\rho, \varrho$	rô		
$\sigma, \Sigma, \varsigma$	sigma		
$\tau$	tau		
$\upsilon, \Upsilon$	upsilon		
$\phi, \varphi, \Phi$	fi		