# Notas de Aula

# Álgebra Linear

**Rodney Josué Biezuner** [1]
Departamento de Matemática
Instituto de Ciências Exatas (ICEx)
Universidade Federal de Minas Gerais (UFMG)

Notas de aula da disciplina *Álgebra Linear II*
dos Cursos de Bacharelado em Matemática e Matemática Computacional,
lecionada pelo autor durante o primeiro semestre de 2006.

26 de junho de 2006

[1] E-mail: rodney@mat.ufmg.br; homepage: http://www.mat.ufmg.br/∼rodney.

# Sumário

**1 Espaços Vetoriais** **3**
1.1 Introdução . . . 3
1.2 Bases e Dimensão . . . 5
1.3 Subespaços Vetoriais . . . 7
1.4 Coordenadas . . . 8
1.5 Somas de Subespaços Vetoriais . . . 9

**2 Aplicações Lineares entre Espaços Vetoriais** **12**
2.1 Definição . . . 12
2.2 Isomorfismos . . . 13
2.3 Espaço Quociente . . . 14
2.4 Teorema do Núcleo e da Imagem . . . 15
2.5 A Álgebra de Aplicações Lineares . . . 17
2.6 Matrizes . . . 18
2.7 Representações de Vetores através de Matrizes Colunas e Mudança de Base . . . 20
2.8 Representações de Aplicações Lineares através de Matrizes . . . 21
2.9 Funcionais Lineares e o Espaço Dual . . . 24
2.10 Transposta de uma Aplicação Linear . . . 26

**3 Determinantes** **27**
3.1 Definição . . . 27
3.2 Existência . . . 28
3.3 Unicidade . . . 29
3.3.1 Permutações . . . 29
3.3.2 Demonstração da Unicidade da Função Determinante . . . 31
3.3.3 Fórmula do Determinante através de Permutações . . . 33
3.4 Propriedades . . . 34
3.5 Regra de Cramer . . . 36

**4 Autovalores e Autovetores** **38**
4.1 Polinômios . . . 38
4.2 Autovalores, Autovetores e Autoespaços . . . 42
4.3 Operadores Diagonalizáveis . . . 43
4.4 Ideais de Polinômios . . . 47
4.5 Polinômio Mínimo e o Teorema de Cayley-Hamilton . . . 47
4.6 Subespaços Invariantes e Operadores Triangularizáveis . . . 50
4.7 Exercícios . . . 55
4.8 Projeções e Decomposição em Soma Direta . . . 56
4.9 Exercícios . . . 60
4.10 Fatoração de Polinômios . . . 61

4.11 Teorema da Decomposição Primária e Teorema Espectral . . . 62
4.12 Exercícios . . . 65
4.13 Exercícios Computacionais . . . 66

**5 Forma Canônica de Jordan** **67**
5.1 Complexificação de um Espaço Vetorial . . . 67
5.2 Forma de Jordan . . . 69
5.3 Cálculo da Forma de Jordan . . . 75
5.4 Base de Jordan . . . 81
5.5 Forma de Jordan Real . . . 87
5.6 Exercícios . . . 88

**6 Espaços com Produto Interno** **89**
6.1 Produto Interno . . . 89
6.2 Norma . . . 92
6.3 Bases Ortonormais e Projeções Ortogonais . . . 97
6.4 Operador Adjunto . . . 100
6.5 Operadores Unitários e Isometrias . . . 103
6.6 Operadores Normais . . . 105
6.7 Teoria Espectral para Operadores Auto-adjuntos . . . 110
6.8 Métodos Variacionais . . . 112
6.9 Formas Bilineares e Sesquilineares . . . 115

# Capítulo 1

# Espaços Vetoriais

## 1.1 Introdução

**Definição 0.** Um **corpo** $\mathbb{F}$ é um conjunto munido de duas operações binárias, *adição* e *multiplicação* que satisfazem as seguintes propriedades:

Adição:

1. Comutatividade: para todos $\alpha, \beta \in \mathbb{F}$
$$\alpha + \beta = \beta + \alpha.$$
2. Associatividade: para todos $\alpha, \beta, \gamma \in \mathbb{F}$
$$\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma.$$
3. Existência de Elemento Neutro: existe um elemento $0 \in \mathbb{F}$ tal que para todo $\alpha \in \mathbb{F}$ temos
$$\alpha + 0 = 0 + \alpha = \alpha.$$
4. Existência de Inverso: para todo $\alpha \in \mathbb{F}$ existe $-\alpha \in \mathbb{F}$ tal que
$$\alpha + (-\alpha) = 0.$$

Multiplicação:

1. Comutatividade: para todos $\alpha, \beta \in \mathbb{F}$
$$\alpha\beta = \beta\alpha.$$
2. Associatividade: para todos $\alpha, \beta, \gamma \in \mathbb{F}$
$$\alpha(\beta\gamma) = (\alpha\beta)\gamma.$$
3. Existência de Elemento Neutro: existe um elemento $1 \in \mathbb{F}$ tal que para todo $\alpha \in \mathbb{F}$ temos
$$\alpha 1 = 1\alpha = \alpha.$$
4. Existência de Inverso: para todo $\alpha \in \mathbb{F}$ existe $\alpha^{-1} \in \mathbb{F}$ tal que
$$\alpha\alpha^{-1} = 1.$$

Distributividade:

1. Para todos $\alpha, \beta, \gamma \in \mathbb{F}$
$$\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma.$$

**Definição 1.** Um **espaço vetorial sobre um corpo** $\mathbb{F}$ é um conjunto $V$ munido de duas operações, *adição de vetores* $V \times V \to V$ e *multiplicação por escalar* $\mathbb{F} \times V \to V$ que satisfazem as seguintes propriedades:

Adição de Vetores:

1. Comutatividade: para todos $x, y \in V$
$$x + y = y + x.$$
2. Associatividade: para todos $x, y, z \in V$
$$x + (y + z) = (x + y) + z.$$
3. Existência de Elemento Neutro: existe um elemento $0 \in V$ tal que para todo $x \in V$ temos
$$x + 0 = 0 + x = x.$$
4. Existência de Inverso: para todo $x \in V$ existe $-x \in V$ tal que
$$x + (-x) = 0.$$

Multiplicação por Escalar:

1. Para todo $x \in V$ vale
$$1x = x.$$
2. Associatividade: para todos $\alpha, \beta \in \mathbb{F}$
$$\alpha(\beta x) = (\alpha\beta) x.$$

Distributividade:

1. Para todos $x, y \in V$ e para todo $\alpha \in \mathbb{F}$
$$\alpha(x + y) = \alpha x + \alpha y.$$
2. Para todo $x \in V$ e para todos $\alpha, \beta \in \mathbb{F}$
$$(\alpha + \beta) x = \alpha x + \beta x.$$

Os elementos de $V$ são chamados *vetores*, e os elementos de $\mathbb{F}$ são chamados *escalares*.

**Proposição 0.** *As seguintes afirmativas são válidas.*

1. $\alpha 0 = 0$.
$$\alpha 0 = \alpha(0 + 0) = \alpha 0 + \alpha 0 \Rightarrow \alpha 0 + (-\alpha 0) = \alpha 0 + \alpha 0 + (-\alpha 0) \Rightarrow 0 = \alpha 0.$$
2. $0v = 0$.
$$\begin{aligned} 0v &= (0 + 0) v = 0v + 0v \\ &\Rightarrow 0v + (-0v) = 0v + 0v + (-0v) \\ &\Rightarrow 0 = 0v. \end{aligned}$$

3. $\alpha x \neq 0$ se $\alpha \neq 0$ e $x \neq 0$.
   Suponha que exista $\alpha \in \mathbb{F}$, $\alpha \neq 0$, tal que $\alpha x = 0$ para algum $x \in V$, $x \neq 0$. Então
   $$\alpha^{-1}(\alpha x) = \alpha^{-1} 0 = 0 \Rightarrow \left(\alpha^{-1}\alpha\right) x = 0 \Rightarrow 1x = 0 \Rightarrow x = 0,$$
   um absurdo.
4. $(-1)\,x = -x$.
   $$\begin{aligned} 0 = 0x = [1 + (-1)]\,x = 1x + (-1)\,x = x + (-1)\,x \\ \Rightarrow -x + x + (-1)\,x = -x + 0 \\ \Rightarrow (-1)\,x = -x. \end{aligned}$$
5. *Unicidade do vetor nulo.*
   $$0' = 0' + 0 = 0.$$

## 1.2 Bases e Dimensão

**Definição 2.** Seja $S \subset V$ um subconjunto qualquer de um espaço vetorial $V$. Uma **combinação linear** de elementos de $S$ é uma soma finita
$$\alpha_1 x_1 + \ldots + \alpha_k x_k$$
com $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ e $x_1, \ldots, x_k \in S$.

**Definição 3.** Dizemos que um conjunto $S \subset V$ é **linearmente dependente** se existir um número finito de elementos $x_1, \ldots, x_k \in S$ e escalares $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ não todos nulos tais que
$$\alpha_1 x_1 + \ldots + \alpha_k x_k = 0.$$
Ou seja, o vetor nulo pode ser escrito como uma combinação linear não-trivial de elementos de $S$. Caso contrário, dizemos que $S$ é **linearmente independente**.

Um subconjunto linearmente independente não pode conter o vetor nulo.

**Lema 1.** *Um subconjunto $S \subset V$ é linearmente dependente se e somente se algum elemento de $S$ puder ser escrito como combinação linear de outros elementos de $S$.*

**Prova.** Se $S$ é linearmente dependente, então existem vetores $x_1, \ldots, x_k \in S$ e escalares $\alpha_1, \ldots, \alpha_k$ não todos nulos tais que
$$\alpha_1 x_1 + \ldots + \alpha_k x_k = 0.$$
Suponha que $\alpha_i \neq 0$. Então podemos escrever
$$x_i = \frac{\alpha_1}{\alpha_i} x_1 + \ldots + \frac{\alpha_{i-1}}{\alpha_i} x_{i-1} + \frac{\alpha_{i+1}}{\alpha_i} x_{i+1} + \ldots + \frac{\alpha_k}{\alpha_i} x_k,$$
isto é, $x_i$ é combinação linear de outros elementos de $S$.

Reciprocamente, se $x_0, x_1, \ldots, x_k \in S$ e $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ são tais que
$$x_0 = \alpha_1 x_1 + \ldots + \alpha_k x_k,$$
então
$$x_0 - \alpha_1 x_1 - \ldots - \alpha_k x_k = 0$$
é uma combinação linear não-trivial de elementos de $S$. ■

**Definição 4.** Dizemos que um conjunto $S \subset V$ **gera** o espaço $V$ se para todo $x \in X$ existirem $x_1, \ldots, x_k \in S$ e escalares $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ tais que

$$x = \alpha_1 x_1 + \ldots + \alpha_k x_k.$$

**Definição 5.** Dizemos que um conjunto $\mathcal{B} \subset V$ é uma **base** para o espaço $V$ se:

- $\mathcal{B}$ gera $V$.
- $\mathcal{B}$ é linearmente independente.

**Definição 6.** Dizemos que $V$ é um **espaço vetorial de dimensão finita** se $V$ possui uma base com um número finito de elementos ou se $V = \{0\}$.

O número de elementos nas bases de um espaço vetorial é um *invariante* do espaço vetorial:

**Teorema 1.** *Todas as bases de um espaço vetorial de dimensão finita possuem o mesmo número de elementos.*

**Prova.** A demonstração deste resultado segue do seguinte lema:

**Lema 2.** *Suponha que $S = \{x_1, \ldots, x_k\}$ gera o espaço vetorial $V$ e que $S' = \{y_1, \ldots, y_l\}$ é um subconjunto linearmente independente de $V$. Então*

$$l \leqslant k.$$

**Prova.** Suponha por absurdo que $l > k$. Como $S$ gera $V$ e $S'$ é linearmente independente (em particular, não contém o vetor nulo) temos

$$y_1 = \alpha_1 x_1 + \ldots + \alpha_k x_k$$

para alguns escalares $\alpha_1, \ldots, \alpha_k$ não todos nulos. Podemos supor $\alpha_1 \neq 0$, reordenando os índices, se necessário. Afirmamos que podemos então substituir $x_1$ por $y_1$, isto é, que o conjunto $S_1 = \{y_1, x_2, \ldots, x_k\}$ gera $V$. De fato, podemos escrever

$$x_1 = \frac{1}{\alpha_1} y_1 - \frac{\alpha_2}{\alpha_1} x_2 - \ldots - \frac{\alpha_k}{\alpha_1} x_k,$$

de modo que se $x = \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k$, então

$$x = \frac{\beta_1}{\alpha_1} y_1 + \left( \beta_2 - \frac{\beta_1 \alpha_2}{\alpha_1} \right) x_2 + \ldots + \left( \beta_k - \frac{\beta_1 \alpha_k}{\alpha_1} \right) x_k.$$

Agora, como $S_1$ gera $V$ e $S'$ é linearmente independente, temos

$$y_2 = \alpha_1 y_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k$$

para alguns escalares $\alpha_1, \ldots, \alpha_k$, com $\alpha_2, \ldots, \alpha_k$ não todos nulos (caso contrário, $y_2$ seria um múltiplo escalar de $y_1$). Supondo $\alpha_2 \neq 0$, reordenando os índices se necessário, usamos o mesmo argumento acima para concluir que podemos substituir $x_2$ por $y_2$, de modo que o conjunto $S_2 = \{y_1, y_2, x_3, \ldots, x_k\}$ gera $V$. Repetindo este procedimento sucessivamente, concluímos que podemos substituir todos os $x_i$ por um número equivalente de $y_i$ (já que, por hipótese de absurdo, $l > k$), e obter que o subconjunto próprio

$$S_k = \{y_1, \ldots, y_k\}$$

de $S'$ gera $V$. Mas então, por definição de conjunto gerador, existem escalares $\alpha_1, \ldots, \alpha_k$ tais que

$$y_{k+1} = \alpha_1 y_1 + \ldots + \alpha_k y_k$$

contrariando o fato que $S'$ é linearmente independente (pelo Lema 1). ■

Sejam $\mathcal{B}_1 = \{x_1, \ldots, x_k\}$ e $\mathcal{B}_2 = \{y_1, \ldots, y_l\}$ duas bases do espaço vetorial de dimensão finita $V$. Aplicando o Lema 2 ao conjunto gerador $\mathcal{B}_1$ e ao conjunto linearmente independente $\mathcal{B}_2$ concluímos que $l \leqslant k$; aplicando o Lema 2 ao conjunto gerador $\mathcal{B}_2$ e ao conjunto linearmente independente $\mathcal{B}_1$ concluímos que $k \leqslant l$. Portanto, $k = l$. ■

**Definição 7.** O número de elementos de uma base qualquer de um **espaço vetorial de dimensão finita** $V$ é chamada a dimensão do espaço e denotada $\dim V$. Se $V = \{0\}$, então definimos $\dim V = 0$.

**Teorema 2.** *Todo espaço vetorial não-nulo gerado por um subconjunto finito possui uma base finita.*

**Prova.** Suponha que $S$ seja um subconjunto finito que gera o subespaço vetorial não-nulo $V$. Se $S$ for linearmente independente, então $S$ é a base procurada e não precisamos fazer nada. Caso contrário, se $S$ é linearmente dependente, podemos retirar um elemento de $S$ e o conjunto resultante ainda gerará $V$ (retire um elemento que seja combinação linear dos demais). Se o conjunto restante for linearmente independente, então ele será uma base finita para $V$. Caso contrário, repetimos o procedimento, até obter um conjunto linearmente independente. ■

**Lema 3.** *Se* $\dim V = n$, *então todo subconjunto de* $V$ *com mais de* $n$ *vetores é linearmente dependente.*

**Prova.** Segue imediatamente do Lema 2. ■

**Teorema 3.** *Todo subconjunto linearmente independente de um espaço vetorial de dimensão finita pode ser completado até uma base do espaço.*

**Prova.** Suponha que $S = \{x_1, \ldots, x_k\}$ seja um subconjunto linearmente independente de $V$. Se $S$ não é uma base para $V$, ou seja, se $k < n$, então existe um vetor $x_{k+1} \in V$ tal que $x_{k+1}$ não é uma combinação linear de elementos de $S$. Segue que o conjunto $S_1 = \{x_1, \ldots, x_k, x_{k+1}\}$ é linearmente independente. Se $k + 1 < n$, repetimos o processo. Se $\dim V = n$, repetimos este processo $n - k$ vezes até encontrar um subconjunto $S_{n-k} = \{x_1, \ldots, x_k, x_{k+1}, \ldots, x_n\}$ que é uma base para $V$. ■

## 1.3 Subespaços Vetoriais

**Definição 8.** Seja $V$ um espaço vetorial sobre um corpo $\mathbb{F}$. Um **subespaço** de $V$ é um subconjunto $W \subset V$ que é ele próprio um espaço vetorial sobre $\mathbb{F}$ com as operações de soma e multiplicação escalar induzidas de $V$.

**Proposição 1.** *Um subconjunto não-vazio* $W \subset V$ *é um subespaço de* $V$ *se e somente se ele é fechado com relação às operações de soma e multiplicação por escalar. Em outras palavras,* $W \subset V$ *é um subespaço de* $V$ *se e somente se* $\alpha x + \beta y \in W$ *para todos* $x, y \in V$ *e para todos* $\alpha, \beta \in \mathbb{F}$.

**Prova.** Suponha que $W \subset V$, $W \neq \varnothing$, é fechado com relação às operações de soma e multiplicação por escalar. Como as operações de soma e multiplicação por escalar definidas em $W$ são herdadas das operações definidas em $V$, comutatividade, associatividade e distributividade são imediatamente válidas e $1x = x$ para todo $x \in W$. Basta apenas verificar as duas propriedades seguintes:

- $0 \in W$: pois se $x \in W$ é qualquer vetor (lembre-se que $W \neq \varnothing$), então $0x = 0 \in W$.
- Se $x \in W$, então $-x \in W$: pois $-x = (-1)\, x \in W$.

A recíproca é óbvia, pois $V$ é um espaço vetorial sobre $\mathbb{F}$ e $W \subset V$ é um espaço vetorial sobre $\mathbb{F}$ com as operações de soma e multiplicação escalar induzidas do próprio $V$. ■

**Teorema 4.** *A interseção de qualquer família de subespaços de um espaço vetorial* $V$ *é um subespaço de* $V$.

**Prova.** Seja $\{W_\lambda\}_{\lambda\in\Lambda}$ uma coleção de subespaços de $V$ e $W = \bigcap_{\lambda\in\Lambda} W_\lambda$ sua interseção. Como cada $W_\lambda$ contém o vetor nulo, segue que $W$ também contém o vetor nulo e podemos usar a Proposição 1 para provar que $W$ é um subespaço. Dados quaisquer $x, y \in W$, temos que $x, y \in W_\lambda$ para cada índice $\lambda \in \Lambda$ (por definição de interseção de conjuntos), logo $\alpha x + \beta y \in W_\lambda$ para todos $\alpha, \beta \in \mathbb{F}$ (pela Proposição 1, pois cada $W_\lambda$ é um subespaço de $V$), portanto $\alpha x + \beta y \in W$ para todos $\alpha, \beta \in \mathbb{F}$ (novamente, pela definição de interseção de conjuntos). Segue da Proposição 1 que $W$ é um subespaço. ■
Segue deste resultado que dado um subconjunto de um espaço vetorial, existe um menor subespaço que o contém:

**Definição 9.** Seja $V$ um espaço vetorial sobre um corpo $\mathbb{F}$ e $S$ um subconjunto de $V$. O **subespaço gerado** por $S$ é a interseção de todos os subespaços de $V$ que contém $S$.

**Proposição 2.** *O subespaço gerado por um subconjunto não-vazio $S$ de um espaço vetorial $V$ é o conjunto de todas as combinações lineares de elementos de $S$.*

**Prova.** Denote por $W$ o subespaço gerado por $S$ e por $W'$ o conjunto de todas as combinações lineares de elementos de $S$. Pela Proposição 1, como $S \subset W$, segue que $W' \subset W$. Por outro lado, também pela Proposição 1 o conjunto $W'$ é um subespaço de $V$, logo $W \subset W'$ por definição (pois $W'$ é um subespaço de $V$ que contém $S$). Assim $W' = W$. ■

**Lema 4.** *Seja $S$ um subconjunto linearmente independente de um espaço vetorial $V$. Suponha que $y$ é um vetor de $V$ que não pertence ao subespaço gerado por $S$. Então $S \cup \{y\}$ é linearmente independente.*

**Prova.** Suponha que $x_1, \ldots, x_k \in S$ e existem escalares $\alpha_1, \ldots, \alpha_k, \beta$ tais que

$$\alpha_1 x_1 + \ldots + \alpha_k x_k + \beta y = 0.$$

Então $\beta = 0$, caso contrário

$$y = \frac{\alpha_1}{\beta} x_1 + \ldots + \frac{\alpha_k}{\beta} x_k,$$

o que implicaria que $y$ pertence ao subespaço gerado por $V$, contrariando a hipótese. Mas então

$$\alpha_1 x_1 + \ldots + \alpha_k x_k = 0,$$

e como $S$ é L.I., segue que $\alpha_1 + \ldots + \alpha_k = 0$. ■

**Teorema 5.** *Se $W$ é um subespaço próprio de um espaço vetorial de dimensão finita $V$, então $W$ também tem dimensão finita e* $\dim W < \dim V$.

**Prova.** O resultado é óbvio se $W$ é o subespaço nulo. Se $W$ não é o subespaço nulo, seja $x \in W$, $x \neq 0$. Pelo Teorema 3, existe uma base de $W$ contendo $x$ (como não sabemos apriori que $W$ tem dimensão finita, não sabemos que $W$ possui uma base, pois não queremos utilizar o resultado que ainda não demonstramos que todo espaço vetorial possui uma base). Como esta base de $W$ é em particular um subconjunto linearmente independente de $V$, ela não pode conter mais que $\dim V$ elementos. Portanto, $\dim W \leqslant \dim V$. Por outro lado, como $W$ é um subespaço próprio de $V$, existe um vetor $y \in V$ tal que $y \notin W$. Adicionando $y$ a qualquer base de $W$, obtemos pelo lema anterior um conjunto linearmente independente de $V$. Isso implica que $\dim W < \dim V$. ■

## 1.4 Coordenadas

**Proposição 3.** *Seja $V$ um espaço vetorial de dimensão finita e $\mathcal{B} = \{x_1, \ldots, x_n\}$ uma base para $V$. Então todo vetor em $V$ se escreve de maneira única como uma combinação linear de vetores de $\mathcal{B}$.*

**Prova.** Suponha que $v \in V$ pode ser representado por duas combinações lineares de vetores de $\mathcal{B}$:

$$v = \alpha_1 x_1 + \ldots + \alpha_n x_n,$$
$$v = \alpha_1' x_1 + \ldots + \alpha_n' x_n.$$

Então

$$(\alpha_1 - \alpha_1') x_1 + \ldots + (\alpha_n - \alpha_n') x_n = 0,$$

donde $\alpha_1 = \alpha_1', \ldots, \alpha_n = \alpha_n'$. ■

**Definição 10.** Seja $V$ um espaço vetorial de dimensão finita e $\mathcal{B} = \{x_1, \ldots, x_n\}$ uma base *ordenada* para $V$. Dado $v \in V$, se

$$v = \alpha_1 x_1 + \ldots + \alpha_n x_n,$$

a $n$-upla $(\alpha_1, \ldots, \alpha_n)$ são chamadas as **coordenadas** de $v$ *com relação à base* $\mathcal{B}$.

Base canônica de $\mathbb{F}^n$.

## 1.5 Somas de Subespaços Vetoriais

**Definição 11.** Sejam $S_1, \ldots, S_k$ subconjuntos de um espaço vetorial $V$. Definimos a soma dos subconjuntos $S_1, \ldots, S_k$ como sendo o conjunto

$$\sum_{i=1}^{k} S_i = S_1 + \ldots + S_k = \{x_1 + \ldots + x_k : x_i \in S_i \text{ para } i = 1, \ldots k\}.$$

**Proposição 4.** *Se $W_1, \ldots, W_k$ são subespaços de um espaço vetorial $V$, então a sua soma $W_1 + \ldots + W_k$ é também um subespaço vetorial de $V$ e contém cada um dos subespaços $W_i$, $i = 1, \ldots k$.*

**Prova.** Pois se

$$x = w_1 + \ldots + w_k,$$
$$y = w_1' + \ldots + w_k',$$

são dois vetores quaisquer de $W_1 + \ldots + W_k$ com $w_i, w_i' \in W_i$ para cada $i$ e $\alpha, \beta$ são escalares quaisquer, segue que

$$\alpha x + \beta y = (\alpha w_1 + \beta w_1') + \ldots + (\alpha w_k + \beta w_k') \in W_1 + \ldots + W_k.$$

A última afirmação é óbvia, pois o vetor nulo esta em cada um dos subespaços. ■

**Teorema 6.** *Se $W_1, W_2$ são dois subespaços de dimensão finita de um espaço vetorial $V$, então $W_1 + W_2$ também tem dimensão finita e*

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

**Prova.** Pelo Teorema 5, $W_1 \cap W_2$ tem uma base finita

$$\{x_1, \ldots, x_n\}.$$

Pelo Teorema 3, esta é parte de uma base

$$\mathcal{B}_1 = \{x_1, \ldots, x_n, y_1, \ldots, y_k\}$$

para $W_1$ e parte de uma base

$$\mathcal{B}_2 = \{x_1, \ldots, x_n, z_1, \ldots, z_l\}$$

para $W_2$. O subespaço $W_1 + W_2$ é gerado pelo conjunto

$$\mathcal{B} = \{x_1, \ldots, x_n, y_1, \ldots, y_k, z_1, \ldots, z_l\}.$$

Basta provar que $\mathcal{B}$ é L.I. para terminar a demonstração, pois então $\mathcal{B}$ será uma base para $W_1 + W_2$ e portanto

$$\begin{aligned}\dim W_1 + \dim W_2 &= (n+k) + (n+l) = (n+k+l) + n \\ &= \dim (W_1 + W_2) + \dim (W_1 \cap W_2).\end{aligned}$$

E, de fato, suponha que

$$\sum_{i=1}^{n} \alpha_i x_i + \sum_{i=1}^{k} \beta_i y_i + \sum_{i=1}^{l} \gamma_i z_i = 0.$$

Escrevendo

$$z := \sum_{i=1}^{l} \gamma_i z_i = -\sum_{i=1}^{n} \alpha_i x_i - \sum_{i=1}^{k} \beta_i y_i,$$

vemos que $z \in W_1$ também (certamente $z \in W_2$). Em particular, existem escalares $\delta_1, \ldots, \delta_n$ tais que

$$z = \sum_{i=1}^{n} \delta_i x_i.$$

Logo, subtraindo as duas expressões para $z$, obtemos

$$\sum_{i=1}^{n} \delta_i x_i - \sum_{i=1}^{l} \gamma_i z_i = 0,$$

e como $\{x_1, \ldots, x_n, z_1, \ldots, z_l\}$ é L.I., concluímos que

$$\delta_1 = \ldots = \delta_n = \gamma_1 = \ldots = \gamma_l = 0.$$

Mas então

$$\sum_{i=1}^{n} \alpha_i x_i + \sum_{i=1}^{k} \beta_i y_i = 0$$

e como $\{x_1, \ldots, x_n, y_1, \ldots, y_k\}$ é L.I., segue que

$$\alpha_1 = \ldots = \alpha_n = \beta_1 = \ldots = \beta_k = 0.$$

■

**Definição 12.** Sejam $W_1, W_2$ dois subespaços de um espaço vetorial $V$. Dizemos que o subespaço $W = W_1 + W_2$ é a **soma direta** dos subespaços $W_1$ e $W_2$ se $W_1 \cap W_2 = \{0\}$. Neste caso, denotamos

$$W = W_1 \oplus W_2.$$

**Proposição 5.** *Se* $W = W_1 \oplus W_2$, *então*

$$\dim W = \dim W_1 + \dim W_2.$$

**Prova.** Segue imediatamente do Teorema 6 quando se observa que $W_1 \cap W_2 = \{0\}$. ■

**Proposição 6.** *$W = W_1 \oplus W_2$ se e somente se todo elemento $w \in W$ se escrever de maneira única como uma soma $w = w_1 + w_2$, $w_1 \in W_1$ e $w_2 \in W_2$ (em outras palavras, $W_1$ e $W_2$ são linearmente independentes).*

**Prova.** Assuma que $W_1 \cap W_2 = \{0\}$. Seja $w \in W$ e suponha que

$$\begin{aligned} w &= w_1 + w_2 \\ w &= w_1' + w_2' \end{aligned}$$

$w_1, w_1' \in W_1$ e $w_2, w_2' \in W_2$. Então

$$(w_1 - w_1') + (w_2 - w_2') = 0,$$

donde

$$(w_1 - w_1') = -(w_2 - w_2').$$

Mas então, $(w_1 - w_1') \in W_1 \cap W_2$ e $(w_2 - w_2') \in W_1 \cap W_2$, logo $w_1 - w_1' = 0$ e $w_2 - w_2' = 0$, ou seja, $w_1 = w_1'$ e $w_2 = w_2'$.

Reciprocamente, assuma que todo elemento $w \in W$ se escrever de maneira única como uma soma $w = w_1 + w_2$, $w_1 \in W_1$ e $w_2 \in W_2$, e suponha por absurdo que exista um vetor $x \in W_1 \cap W_2$ tal que $x \neq 0$. Então $0 = x + (-x)$ é um vetor de $W$ que se escreve de pelo menos duas maneiras como a soma de vetores de $W_1$ e $W_2$:

$$\begin{aligned} 0 &= x + (-x), \\ 0 &= 0 + 0. \end{aligned}$$

■

**Teorema 7.** *Seja $V$ um espaço vetorial de dimensão finita. Então todo subespaço $W \subset V$ possui um complemento em $V$, isto é, existe um subespaço $Z \subset V$ tal que*

$$V = W \oplus Z.$$

**Prova.** Se $W = \{0\}$ ou $W = V$, tome $Z = V$ ou $Z = \{0\}$. Caso contrário, seja $\{x_1, \ldots, x_k\}$ uma base para $W$. Complete esta base até uma base para $V$: $\{x_1, \ldots, x_k, y_1, \ldots, y_l\}$. Então tomamos como $Z$ o subespaço gerado pelos vetores $y_1, \ldots, y_l$. De fato, se $W \cap Z \neq \{0\}$, tomando um vetor não-nulo $z \in W \cap Z$, teríamos escalares não todos nulos $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_l$ tais que

$$\begin{aligned} z &= \sum_{i=1}^{k} \alpha_i x_i, \\ z &= \sum_{i=1}^{l} \beta_i y_i, \end{aligned}$$

donde

$$\sum_{i=1}^{k} \alpha_i x_i - \sum_{i=1}^{l} \beta_i y_i = 0$$

seria uma combinação linear não trivial produzindo o vetor nulo, contradizendo o fato que $\{x_1, \ldots, x_k, y_1, \ldots, y_l\}$ é L.I. ■

# Capítulo 2

# Aplicações Lineares entre Espaços Vetoriais

## 2.1 Definição

Definida uma estrutura matemática sobre conjuntos, é importante o estudo das funções entre estes conjuntos que preservam a estrutura.

**Definição 1.** Sejam $V$ e $W$ dois espaços vetoriais sobre um mesmo corpo $\mathbb{F}$. Uma função $T : V \longrightarrow W$ é chamada uma **aplicação linear** se

$$T(\alpha v + \beta w) = \alpha T(v) + \beta T(w)$$

para todos $v \in V, w \in W$ e $\alpha, \beta \in \mathbb{F}$.

Aplicações lineares preservam as operações que definem um espaço vetorial, soma e multiplicação por escalar. Em outras palavras, elas preservam combinações lineares.

**Proposição 1.** *Seja $T : V \longrightarrow W$ uma aplicação linear entre dois espaços vetoriais. Então $T(\mathbf{0}_V) = \mathbf{0}_W$.*

**Prova:** Observe que estamos usando notações diferentes para os vetores nulos de cada espaço por motivos de clareza. Temos

$$T(\mathbf{0}_V) = T(0\mathbf{0}_V) = 0T(\mathbf{0}_V) = \mathbf{0}_W.$$

■

**Teorema 1.** *Uma aplicação linear do espaços vetorial de dimensão finita $V$ para o espaço vetorial $W$ é completamente determinada pelos valores que ela toma em uma base qualquer de $V$.*

**Prova:** Sejam $T : V \longrightarrow W$ uma aplicação linear e $\mathcal{B} = \{x_1, \ldots, x_n\}$ uma base para $V$. Dado um vetor $v \in V$, ele se escreve como uma combinação linear

$$v = \alpha_1 x_1 + \ldots + \alpha_n x_n.$$

Então

$$Tv = T(\alpha_1 x_1 + \ldots + \alpha_n x_n) = \alpha_1 T(x_1) + \ldots + \alpha_n T(x_n).$$

■

Podemos dizer ainda mais: para definir uma aplicação linear $T : V \longrightarrow W$ basta estipular os seus valores em uma base:

**Teorema 2.** *Sejam $V$ um espaço vetorial de dimensão finita, $B = \{x_1, \ldots, x_n\}$ uma base para $V$ e $y_1, \ldots, y_n$ vetores quaisquer de um espaço vetorial $W$. Então existe uma única aplicação linear $T : V \longrightarrow W$ tal que*

$$Tx_i = y_i$$

*para $i = 1, \ldots, n$.*

**Prova:** Como todo vetor $v \in V$ se escreve como uma combinação linear de maneira única

$$v = \alpha_1 x_1 + \ldots + \alpha_n x_n,$$

basta definir

$$Tv = \alpha_1 y_1 + \ldots + \alpha_n y_n.$$

De fato, se

$$\begin{aligned} x &= \alpha_1 x_1 + \ldots + \alpha_n x_n, \\ y &= \beta_1 x_1 + \ldots + \beta_n x_n, \end{aligned}$$

temos, para todos $x, y \in V$ e para todos escalares $\alpha, \beta \in \mathbb{F}$,

$$\begin{aligned} T(\alpha x + \beta y) &= T\left(\alpha \sum_{i=1}^{n} \alpha_i x_i + \beta \sum_{i=1}^{n} \beta_i x_i\right) = T\left(\sum_{i=1}^{n} (\alpha\alpha_i + \beta\beta_i) x_i\right) \\ &= \sum_{i=1}^{n} (\alpha\alpha_i + \beta\beta_i) y_i = \alpha \sum_{i=1}^{n} \alpha_i y_i + \beta \sum_{i=1}^{n} \beta_i y_i \\ &= \alpha T(x) + \beta T(y). \end{aligned}$$

A unicidade de $T$ decorre do teorema anterior. ■

## 2.2 Isomorfismos

**Definição 2.** Um **isomorfismo** entre dois espaços vetoriais $V$ e $W$ sobre um mesmo corpo $\mathbb{F}$ é uma aplicação linear bijetiva $T : V \longrightarrow W$ cuja inversa é linear. Quando existir, dizemos que $V$ e $W$ são **isomorfos**.

**Proposição 2.** *Seja $T : V \longrightarrow W$ uma aplicação linear injetiva entre dois espaços vetoriais. Seja $Z = T(V)$. Então a aplicação inversa $T^{-1} : Z \longrightarrow V$ também é linear.*

**Prova:** Dados $z_1, z_2 \in Z$, sejam $v_1, v_2 \in V$ tais que $T(v_1) = z_1, T(v_2) = z_2$. Dados $\alpha, \beta \in \mathbb{F}$, segue que

$$T(\alpha v_1 + \beta v_2) = \alpha T(v_1) + \beta T(v_2) = \alpha z_1 + \beta z_2.$$

Portanto,

$$T^{-1}(\alpha z_1 + \beta z_2) = \alpha v_1 + \beta v_2 = \alpha T^{-1}(z_1) + \beta T^{-1}(z_2).$$

■
Do ponto de vista da álgebra linear, espaços isomorfos são indistinguíveis.

**Lema 1.** *Uma aplicação linear $T : V \longrightarrow W$ é injetiva se e somente se $T^{-1}(\mathbf{0}_W) = \mathbf{0}_V$.*

**Prova:** Assuma $T^{-1}(\mathbf{0}_W) = \mathbf{0}_V$. Então, se $T(x) = T(y)$, por linearidade segue que $T(x - y) = \mathbf{0}_W$, logo $x - y = \mathbf{0}_V$ e portanto $x = y$, ou seja, $T$ é injetiva. Reciprocamente, assuma $T : V \longrightarrow W$ é injetiva. Como , Se $T(x) = T(y)$, por linearidade $T(x - y) = \mathbf{0}_W$, e como $T(\mathbf{0}_V) = \mathbf{0}_W$, segue da injetividade de $T$ que $x - y = \mathbf{0}_V$, logo $x = y$. ■

**Teorema 3.** *Todo espaço vetorial sobre* $\mathbb{F}$ *de dimensão finita* $n$ *é isomorfo a* $\mathbb{F}^n$.

**Prova:** Seja $\mathcal{B} = \{x_1, \ldots, x_n\}$ uma base para um espaço vetorial $V$ sobre $\mathbb{F}$ de dimensão $n$. Definimos um isomorfismo $T : V \longrightarrow \mathbb{F}^n$ por

$$T(x_i) = e_i,$$

e declarando $T$ linear, ou seja,

$$T(\alpha_1 x_1 + \ldots + \alpha_n x_n) = \alpha_1 e_1 + \ldots + \alpha_n e_n.$$

É fácil ver que $T$ é linear, injetiva e sobrejetiva. ■

**Teorema 4.** *Se* $n \neq m$*, então* $\mathbb{F}^n$ *não é isomorfo a* $\mathbb{F}^m$.

**Prova:** Suponha $n > m$ e assuma por absurdo que $T : \mathbb{F}^n \longrightarrow \mathbb{F}^m$ é um isomorfismo. Mostraremos que $\{Te_1, \ldots, Te_n\}$ é um conjunto L.I. no espaço $\mathbb{F}^m$ de dimensão $m < n$, uma contradição. De fato, se

$$0 = \alpha_1 T(e_1) + \ldots + \alpha_n T(e_n) = T(\alpha_1 e_1 + \ldots + \alpha_n e_n),$$

então $\alpha_1 e_1 + \ldots + \alpha_n e_n = 0$ (pois $T$ é injetiva), o que implica $\alpha_1 + \ldots + \alpha_n = 0$. ■

**Corolário 1.** *Sejam* $V$ *e* $W$ *espaços vetoriais isomorfos. Então* $\dim V = \dim W$.

**Prova:** Segue dos Teoremas 3 e 4 e do fato de que a composta de um isomorfismo é um isomorfismo (a composta de aplicações lineares é uma aplicação linear). ■

## 2.3 Espaço Quociente

**Definição 3.** Seja $W$ um subespaço do espaço vetorial $V$. Se $x, y \in V$, dizemos que $x$ é congruente a $y$ módulo $W$ se $x - y \in W$. Denotamos isso por

$$x \equiv y \operatorname{mod} W.$$

A relação de congruência módulo subespaço é uma relação de equivalência (definir e exemplificar). Denotaremos a classe de equivalência do vetor $v$ por $[v]$, isto é,

$$[v] = \{x \in V : x \sim v\}.$$

Observe que a classe de equivalência do vetor nulo é exatamente o subespaço $W$.

**Definição 4.** Seja $W$ um subespaço do espaço vetorial $V$. As operações de soma e multiplicação por escalar do espaço vetorial $V$ induzem de forma natural operações de soma e multiplicação por escalar no conjunto das classes de equivalência módulo $W$ por

$$[v] + [w] = [v + w],$$
$$\alpha [v] = [\alpha v].$$

Com estas operações, o conjunto das classes de equivalência módulo $W$ torna-se um subespaço vetorial, chamado o **espaço quociente** de $V$ por $W$ e denotado por

$$V/W.$$

Mostre que as operações estão bem definidas e que $V/W$ satisfaz as propriedades de um espaço vetorial.

**Teorema 5.** *Seja*

$$V = W \oplus Z.$$

*Então a aplicação quociente*

$$\begin{aligned} Z &\longrightarrow V/W \\ z &\mapsto [z] \end{aligned}$$

*é um isomorfismo canônico (isto é, independe da escolha de bases).*

*Em particular, se tem $V$ dimensão finita, então o espaço quociente $V/W$ tem dimensão finita e*

$$\dim (V/W) = \dim V - \dim W.$$

**Prova:** Denote a aplicação quociente por $Q$. Então

$$Q(\alpha x + \beta y) = [\alpha x + \beta y] = [\alpha x] + [\beta y] = \alpha [x] + \beta [y] = \alpha Q(x) + \beta Q(y),$$

logo $Q$ é linear. Para ver que $Q$ é injetiva, note que se $Q(x) = 0$ então por definição da aplicação quociente temos que $[x] = 0$, o que significa que $x \in W$, por definição da relação de congruência módulo $W$. Então $x \in Z \cap W = \{0\}$, logo $x = 0$. Para ver que $Q$ é sobrejetiva, seja $[y] \in V$. Temos $y = w + z$, com $w \in W$ e $z \in Z$. Daí, $y - z = w$, ou seja, $[y] = [z]$, donde $[y] = T(z)$.

Como $V/W$ é isomorfo a $Z$ e $V = W \oplus Z$, temos respectivamente

$$\begin{aligned} \dim (V/W) &= \dim Z, \\ \dim V &= \dim Z + \dim W, \end{aligned}$$

donde segue a última afirmativa. ■

Outra maneira de definir isomorfismo canônico é que ele é única aplicação que faz o diagrama seguinte virar comutativo:

$$\begin{tikzcd} Z \arrow[r, "\text{inclusão}"] \arrow[dr] & V \arrow[d, "\text{projeção}"] \\ & V/W \end{tikzcd}$$

## 2.4 Teorema do Núcleo e da Imagem

**Proposição 3.** *Seja $T : V \longrightarrow W$ uma aplicação linear entre dois espaços vetoriais. Se $U$ é um subespaço de $V$, então $T(U)$ é um subespaço de $W$. Reciprocamente, se $Z$ é um subespaço de $W$, então $T^{-1}(Z)$ é um subespaço de $V$.*

**Prova:** Seja $U$ é um subespaço de $V$. Como $0 \in U$, temos $T(U) \neq \varnothing$. Sejam $w_1, w_2 \in T(U)$. Então existem $u_1, u_2 \in U$ tais que $T(u_1) = w_1, T(u_2) = w_2$. Para todos $\alpha, \beta \in \mathbb{F}$ segue que

$$\alpha w_1 + \beta w_2 = \alpha T(u_1) + \beta T(u_2) = T(\alpha u_1 + \beta u_2),$$

e como $\alpha u_1 + \beta u_2 \in U$, concluímos que $\alpha w_1 + \beta w_2 \in T(U)$.

Reciprocamente, seja $Z$ é um subespaço de $W$. Sejam $v_1, v_2 \in T^{-1}(Z)$. Então $T(v_1) =: z_1, T(v_2) =: z_2 \in Z$. Para todos $\alpha, \beta \in \mathbb{F}$ segue que

$$T(\alpha v_1 + \beta v_2) = \alpha T(v_1) + \beta T(v_2) = \alpha z_1 + \beta z_2 \in Z,$$

logo concluímos que $\alpha v_1 + \beta v_2 \in T^{-1}(Z)$. ■

Segue deste resultado que o conjunto imagem $\operatorname{im} T$ de uma aplicação linear $T : V \longrightarrow W$ entre espaços vetoriais é um subespaço de $W$ e que o conjunto $T^{-1}(0)$ é um subespaço de $V$; este último é chamado o **núcleo** da aplicação linear $T$ e denotado $\ker T$.

**Teorema 6.** *Seja $T : V \longrightarrow W$ uma aplicação linear entre dois espaços vetoriais. Então*

$$V/\ker T \quad \text{e} \quad \operatorname{im} T$$

*são isomorfos.*

*Em particular, se $V$ tem dimensão finita,*

$$\dim V = \dim (\ker T\ ) + \dim (\operatorname{im} T).$$

**Prova:** Embora a segunda afirmativa decorre da primeira e do Teorema 4, já que

$$\dim (\operatorname{im} T) = \dim (V/\ker T) = \dim V - \dim (\ker T\ ),$$

vamos dar-lhe uma demonstração independente, já que ela é a mais freqüentemente usada nas aplicações e não necessita da introdução do conceito de espaço quociente. Seja $\{x_1, \ldots, x_k\}$ uma base para $\ker T$ e complete este conjunto L.I. até uma base $\{x_1, \ldots, x_k, x_{k+1}, \ldots, x_n\}$ para $V$. Afirmamos que

$$\mathcal{B} = \{Tx_{k+1}, \ldots, Tx_n\}$$

é uma base para $\operatorname{im} T$. De fato, dado $v = \sum_{i=1}^{n} \alpha_i x_i$, temos $Tv = \sum_{i=k+1}^{n} \alpha_i T x_i$, já que $Tx_1 = \ldots = Tx_k = 0$, portanto $\mathcal{B}$ gera $\operatorname{im} T$. Para provar que $\mathcal{B}$ é L.I., suponha que

$$\sum_{i=k+1}^{n} \alpha_i T x_i = 0.$$

Então,

$$T\left(\sum_{i=k+1}^{n} \alpha_i x_i\right) = 0,$$

o que implica que $\sum_{i=k+1}^{n} \alpha_i x_i \in \ker T$ (a interseção dos subespaços $\langle x_1, \ldots, x_k\rangle$ e $\langle x_{k+1}, \ldots, x_n\rangle$ é o vetor nulo), logo $\sum_{i=k+1}^{n} \alpha_i x_i = 0$ e portanto $\alpha_{k+1} = \ldots = \alpha_n = 0$.

Para obter o isomorfismo desejado, considere o diagrama

$$\begin{array}{lcl} V & \xrightarrow{T} & \operatorname{im} T \\ \downarrow Q & \nearrow_{U} & \\ V/\ker T & & \end{array}$$

onde $Q : V \longrightarrow V/\ker T$ é a projeção canônica $Q(v) = [v]$. Vamos definir o isomorfismo $U$ de tal forma que o diagrama seja comutativo, isto é, $T = U \circ Q$:

$$U([v]) = Tv.$$

Observe que $U$ está bem definida, porque se $[v] = [w]$, então $v - w \in \ker T$, isto é, $T(v - w) = 0$, donde $Tv = Tw$. Além disso, $U$ é linear porque

$$U(\alpha [v] + \beta [w]) = U([\alpha v + \beta w]) = T(\alpha v + \beta w) = \alpha T(v) + \beta T(w) = \alpha U([v]) + \beta U([w]).$$

$U$ é injetiva porque se $U([v]) = Tv = 0$, então $v \in \ker T$, logo $[v] = 0$ em $V/\ker T$. Finalmente, $U$ é sobrejetiva, porque dado $w \in \operatorname{im} T$, temos $w = Tv$ para algum $v \in V$, logo $w = U([v])$. ■

**Corolário 2.** *Sejam $V$ e $W$ espaços vetoriais com a mesma dimensão. Então uma aplicação linear $T : V \longrightarrow W$ é injetiva se e somente se ela é sobrejetiva.*

**Prova:** Pois

$$\dim W = \dim V = \dim (\ker T\ ) + \dim (\operatorname{im} T),$$

logo $\dim (\ker T\ ) = 0$ se e somente se $\dim (\operatorname{im} T) = \dim W$. ■

## 2.5 A Álgebra de Aplicações Lineares

**Proposição 4.** *A composta de aplicações lineares é uma aplicação linear.*

**Prova:** Sejam $V, W, Z$ espaços vetoriais e $S : V \longrightarrow W, T : W \longrightarrow Z$ aplicações lineares. Então $T \circ S : V \longrightarrow Z$ satisfaz

$$(T \circ S)(\alpha x + \beta y) = T[S(\alpha x + \beta y)] = T[\alpha S(x) + \beta S(y)] = \alpha T[S(x)] + \beta T[S(y)] = \alpha (T \circ S)(x) + \beta (T \circ S)(y).$$

■

**Corolário 3.** *Isomorfismo é uma relação de equivalência no conjunto dos espaços vetoriais.*

**Definição 5.** Sejam $V$ e $W$ espaços vetoriais sobre um corpo $\mathbb{F}$. Denote o conjunto das aplicações lineares de $V$ em $W$ por $\mathcal{L}(V, W)$. Definimos as operações de soma e multiplicação por escalar de elementos de $\mathcal{L}(V, W)$ por

$$\begin{aligned} (T + S)(v) &:= T(v) + S(v), \\ (\alpha T)(v) &:= \alpha T(v), \end{aligned}$$

para todo $v \in V$.

Se $V = W$, denotamos $\mathcal{L}(V, W)$ simplesmente por $\mathcal{L}(V)$.

**Proposição 5.** *$\mathcal{L}(V, W)$ é um espaço vetorial sobre $\mathbb{F}$.*

**Proposição 6.** *Se $V$ tem dimensão $n$ e $W$ tem dimensão $m$, então $\mathcal{L}(V, W)$ tem dimensão $nm$.*

**Prova:** Sejam $\mathcal{B}_V = \{x_1, \ldots, x_n\}$ e $\mathcal{B}_W = \{y_1, \ldots, y_m\}$ bases ordenadas de $V$ e $W$ respectivamente. Para cada par de índices $ij$, $i = 1, \ldots, n$, $j = 1, \ldots, m$, defina $E^{ij} : V \longrightarrow W$ como sendo a única aplicação linear que satisfaz

$$E^{ij}(x_k) = \delta_{ik} y_j, \quad k = 1, \ldots, n,$$

onde $\delta_{ik}$ é o delta de Kronecker. Em outras palavras,

$$E^{ij}(x_k) = \begin{cases} y_j & \text{se } i = k, \\ 0 & \text{se } i \neq k. \end{cases}$$

Afirmamos que $\mathcal{B} = \left\{E^{ij}\right\}_{i=1,\ldots,n, j=1,\ldots,m}$ formam uma base para $\mathcal{L}(V, W)$. Como este conjunto tem $nm$ elementos, isto provará o resultado. E, de fato, se $T : V \longrightarrow W$ é uma aplicação linear, para cada $i = 1, \ldots, n$ temos

$$T(x_k) = \sum_{j=1}^{m} \alpha_{kj} y_j.$$

para alguns escalares $\alpha_{ij} \in \mathbb{F}$. Mostraremos que

$$T = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{ij} E^{ij},$$

o que provará que $\mathcal{B}$ gera $\mathcal{L}(V, W)$. Com efeito, para cada $k = 1, \ldots, n$ temos

$$\begin{aligned} \left( \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{ij} E^{ij} \right)(x_k) &= \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{ij} E^{ij}(x_k) = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{ij} (\delta_{ik} y_j) \\ &= \sum_{j=1}^{m} \left( \sum_{i=1}^{n} \delta_{ik} \alpha_{ij} \right) y_j = \sum_{j=1}^{m} \alpha_{kj} y_j \\ &= T(x_k) \end{aligned}$$

e o resultado segue por unicidade. Para provar que $\mathcal{B}$ é L.I., suponha que

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\alpha_{ij}E^{ij}=0.$$

Então

$$\left(\sum_{i=1}^{n}\sum_{j=1}^{m}\alpha_{ij}E^{ij}\right)(x_k)=0$$

para todo $k=1,\dots,n$, ou seja, esta é aplicação nula, e seguindo o argumento anterior, podemos escrever para cada $k=1,\dots,n$

$$0=0\,(x_k)=\sum_{j=1}^{m}\alpha_{kj}y_j,$$

donde $\alpha_{k1}=\ldots=\alpha_{km}=0$. ■

**Definição 6.** Em $\mathcal{L}(V)$, o produto de dois operadores lineares é definido por

$$TS:=T\circ S.$$

Uma **álgebra linear** sobre $\mathbb{F}$ é um espaço vetorial $X$ sobre $\mathbb{F}$ em que uma operação adicional de **multiplicação** de vetores é definida, satisfazendo as seguintes propriedades:

(i) (Associatividade) Para todos $v,w,z\in X$ temos

$$v\,(wz)=(vw)\,z,$$

(ii) (Distributividade com relação à Soma) Para todos $v,w,z\in X$ temos

$$v\,(w+z)=vw+vz,$$
$$(v+w)\,z=vz+wz;$$

(iii) (Distributividade com relação à Multiplicação por Escalar) Para todo $\alpha\in\mathbb{F}$ temos

$$\alpha\,(vw)=(\alpha v)\,w=v\,(\alpha w)\,.$$

Se existir um elemento $1\in X$ tal que $1v=v1=v$ para todo $v\in X$, dizemos que $X$ é uma **álgebra linear com identidade** sobre $\mathbb{F}$. Se a operação de multiplicação for comutativa, dizemos que é uma **álgebra comutativa**.

**Proposição 7.** *$\mathcal{L}(V)$ é uma álgebra linear com identidade. Ela não é comutativa se* $\dim V\geqslant 2$.

## 2.6 Matrizes

**Definição 7.** Uma matriz $m\times n$ sobre o corpo $\mathbb{F}$ é uma função $A:\{1,\dots,m\}\times\{1,\dots,n\}\longrightarrow\mathbb{F}$.

As **entradas** da matriz $A$ são os escalares $A\,(i,j)$ representados por $A_{ij}$ ou por $a_{ij}$; a matriz $A$ também pode ser denotada por $A=(a_{ij})$.

Uma matriz $m\times 1$ é chamada uma **matriz coluna** e uma matriz $1\times n$ é chamada uma **matriz linha**.

**Definição 8.** Sobre o conjunto $M_{m\times n}(\mathbb{F})$ das matrizes $m\times n$ sobre o corpo $\mathbb{F}$ definimos as operações de soma e multiplicação por escalar por

$$(A+B)_{ij}:=A_{ij}+B_{ij},$$
$$(\alpha A)_{ij}:=\alpha A_{ij}.$$

**Proposição 8.** *$M_{m\times n}(\mathbb{F})$ é um espaço vetorial sobre $\mathbb{F}$.*

**Definição 9.** Dadas duas matrizes sobre o corpo $\mathbb{F}$, $A_{m\times p}$ e $B_{p\times n}$, a matriz produto $AB$ é a matriz $m \times p$ definida por

$$(AB)_{ij} = \sum_{r=1}^{p} a_{ir} b_{rj}.$$

**Proposição 9.** *O produto de matrizes satisfaz as seguintes propriedades*

*(i) (Associatividade) Para todas matrizes $A \in M_{m\times p}(\mathbb{F})$, $B \in M_{p\times q}(\mathbb{F})$ e $C \in M_{q\times n}(\mathbb{F})$ vale*

$$A(BC) = (AB)C.$$

*(ii) (Distributividade com relação à soma) Para todas matrizes $A, B, C \in M_{m\times n}(\mathbb{F})$ vale*

$$A(B + C) = AB + AC,$$
$$(A + B)C = AC + BC.$$

*(iii) (Distributividade com relação à multiplicação por escalar) Para toda matriz $A \in M_{m\times n}(\mathbb{F})$ e para todo escalar $\alpha \in \mathbb{F}$ vale*

$$\alpha(AB) = (\alpha A)B = A(\alpha B).$$

*(iv) (Existência de elemento neutro) Defina $I_n$ como sendo a matriz*

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

*Então, para toda matriz $A \in M_{m\times n}(\mathbb{F})$ vale*

$$AI_n = I_m A = A.$$

*Em particular, $M_n(\mathbb{F})$ é uma álgebra linear com identidade.*

**Prova:** (i) De fato, se $A = (a_{ij})_{m\times p}, B = (b_{ij})_{p\times q}$ e $C = (c_{ij})_{q\times n}$, então os produtos estão todos definidos e nós temos

$$\begin{aligned} [A(BC)]_{ij} &= \sum_{r=1}^{p} a_{ir}[BC]_{rj} = \sum_{r=1}^{p} a_{ir}\left(\sum_{s=1}^{q} b_{rs}c_{sj}\right) = \sum_{r=1}^{p}\sum_{s=1}^{q} a_{ir}\left(b_{rs}c_{sj}\right) = \sum_{r=1}^{p}\sum_{s=1}^{q}\left(a_{ir}b_{rs}\right)c_{sj} \\ &= \sum_{s=1}^{q}\sum_{r=1}^{p}\left(a_{ir}b_{rs}\right)c_{sj} = \sum_{s=1}^{q}\left(\sum_{r=1}^{p} a_{ir}b_{rs}\right)c_{sj} = \sum_{s=1}^{q}[AB]_{is}c_{sj} = [(AB)C]_{ij}. \end{aligned}$$

■

**Definição 10.** Uma matriz quadrada $A$, $n \times n$, é **invertível** (ou **não-singular**) se existe uma matriz $B$, $n \times n$, tal que

$$AB = BA = I.$$

$B$ é chamada a **inversa** de $A$.

Se $A$ não possui inversa, dizemos que $A$ é **singular** (esta terminologia se explica pelo fato que as matrizes que não possuem inversa serem uma minoria entre todas as matrizes, minoria em um sentido matematicamente preciso além do alcance deste curso), ou não-invertível.

**Proposição 10.** *Valem os seguintes fatos:*

*(i) Se uma matriz possui uma inversa, então esta inversa é única.*

*(ii) Se $A$ é invertível, então $A^{-1}$ também é e $(A^{-1})^{-1} = A$.*

*(iii) Se $A, B$ são invertíveis, então $AB$ também é e*

$$(AB)^{-1} = B^{-1}A^{-1}.$$

**Prova:** (i) Suponha que

$$AB_1 = B_1A = I.$$
$$AB_2 = B_2A = I.$$

Tomando a equação $B_1A = I$, por exemplo, e multiplicando ambos os lados desta equação à direita por $B_2$, obtemos

$$(B_1A)B_2 = IB_2 \Rightarrow B_1(AB_2) = B_2 \Rightarrow B_1I = B_2 \Rightarrow B_1 = B_2.$$

(iii) Para verificar isso, temos que mostrar que

$$(AB)B^{-1}A^{-1} = I,$$
$$B^{-1}A^{-1}(AB) = I.$$

Provaremos a primeira identidade, já que a demonstração da segunda é análoga. De fato,

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

■

**Questão:** Se $A$ e $B$ são matrizes $n \times n$ tais que o produto $AB$ é invertível, então $A$ e $B$ também são necessariamente invertíveis? E se $A$ e $B$ são matrizes tais que o produto $AB$ é invertível, então $A$ e $B$ também são necessariamente invertíveis?

## 2.7 Representações de Vetores através de Matrizes Colunas e Mudança de Base

**Definição 11.** Dado um espaço vetorial $V$ de dimensão finita, e uma base $\mathcal{B} = \{x_1, \ldots, x_n\}$ para $V$, representamos um vetor $v$ de $V$ com relação à base $\mathcal{B}$ através de uma matriz coluna

$$v = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

onde as entradas $\alpha_1, \ldots, \alpha_n$ são as coordenadas de $v$ com relação à base $\mathcal{B}$. Esta representação de $v$ também será denotada por

$$[v]_{\mathcal{B}}\,.$$

**Teorema 7.** *Sejam $\mathcal{B} = \{x_1, \ldots, x_n\}$ e $\mathcal{B}' = \{x'_1, \ldots, x'_n\}$ duas bases para o espaço vetorial $V$. Então existe uma única matriz invertível $P$ tal que*

$$[v]_{\mathcal{B}'} = P\,[v]_{\mathcal{B}}\,,$$
$$[v]_{\mathcal{B}} = P^{-1}\,[v]_{\mathcal{B}'}\,,$$

*para todo vetor $v \in V$, chamada a* ***matriz de mudança de base*** *de $\mathcal{B}$ para $\mathcal{B}'$. As colunas de $P$ são dadas pelas coordenadas dos vetores da base $\mathcal{B}$ com relação à base $\mathcal{B}'$, ou seja,*

$$P_i = [x_i]_{\mathcal{B}'}, \quad i = 1, \ldots, n.$$

**Prova:** Por hipótese, denotando $P = (p_{ij})$, os vetores da base $\mathcal{B}$ se escrevem em relação à base $\mathcal{B}'$ na forma

$$x_j = \sum_{i=1}^{n} p_{ij} x_i', \quad j = 1, \ldots, n.$$

Dado um vetor $v \in V$, suas coordenadas em relação às bases $\mathcal{B}$ e $\mathcal{B}'$ são, respectivamente,

$$v = \sum_{j=1}^{n} \alpha_j x_j,$$
$$v = \sum_{i=1}^{n} \alpha_i' x_i'.$$

Como

$$v = \sum_{j=1}^{n} \alpha_j x_j = \sum_{j=1}^{n} \alpha_j \left( \sum_{i=1}^{n} p_{ij} x_i' \right) = \sum_{i=1}^{n} \left( \sum_{j=1}^{n} p_{ij} \alpha_j \right) x_i',$$

segue que (unicidade das coordenadas)

$$\alpha_i' = \sum_{j=1}^{n} p_{ij} \alpha_j.$$

Portanto, provamos que existem matrizes $n \times n$ $P$ e $Q$ tais que

$$[v]_{\mathcal{B}'} = P [v]_{\mathcal{B}},$$
$$[v]_{\mathcal{B}} = Q [v]_{\mathcal{B}'}.$$

Em particular

$$[v]_{\mathcal{B}'} = P [v]_{\mathcal{B}} = PQ [v]_{\mathcal{B}'},$$
$$[v]_{\mathcal{B}} = Q [v]_{\mathcal{B}'} = QP [v]_{\mathcal{B}},$$

para todo $v \in V$. Isso implica (pela unicidade da matriz identidade) que

$$PQ = QP = I,$$

ou seja, $Q = P^{-1}$. ■

## 2.8 Representações de Aplicações Lineares através de Matrizes

Seja $T : V \longrightarrow W$ uma aplicação linear entre espaços vetoriais de dimensão finita. Escolha bases $\mathcal{B}_V = \{x_1, \ldots, x_n\}$ para $V$ e $\mathcal{B}_W = \{y_1, \ldots, y_m\}$ para $W$. Então, cada vetor $x \in V$ se escreve com relação à base $\mathcal{B}_V$ na forma

$$x = \alpha_1 x_1 + \ldots + \alpha_n x_n = \sum_{j=1}^{n} \alpha_j x_j,$$

para alguns escalares $\alpha_1, \ldots, \alpha_n$. Daí, segue que

$$Tx = T\left(\sum_{j=1}^{n} \alpha_j x_j\right) = \sum_{j=1}^{n} \alpha_j T(x_j).$$

Escreva os vetores $Tx_1, \ldots, Tx_n$ em relação à base $\mathcal{B}_W$ na forma

$$T(x_j) = \sum_{i=1}^{m} a_{ij} y_i,$$

isto é, na forma de matriz coluna:

$$T(x_j) = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

A matriz $A = (a_{ij})_{m \times n}$ é chamada a **representação matricial** da aplicação linear $T$ com relação às bases $\mathcal{B}_V$ e $\mathcal{B}_W$. Esta representação de $T$ também será denotada por

$$[T]_{\mathcal{B}_V, \mathcal{B}_W}.$$

**Teorema 8.** *Sejam $\mathcal{B}_V = \{x_1, \ldots, x_n\}$ e $\mathcal{B}_W = \{y_1, \ldots, y_m\}$ bases para os espaços vetoriais $V$ e $W$, respectivamente. Então a aplicação*

$$\Phi : \mathcal{L}(V, W) \longrightarrow M_{m \times n}(\mathbb{F})$$

*definida por*

$$T \mapsto [T]_{\mathcal{B}_V, \mathcal{B}_W}$$

*é um isomofismo entre espaços vetoriais. Mais que isso, se $V = W$, então ela é um isomorfismo entre álgebras lineares com identidade.*

**Prova:** Verificar que $\Phi$ é um isomorfismo entre espaços vetoriais é um exercício. Para ver que $\Phi$ é um isomorfismo entre álgebras lineares com identidade, basta verificar que o produto é preservado, já que é fácil ver que $\Phi$ leva a aplicação identidade na matriz identidade (exercício). A preservação do produto é um corolário do resultado mais geral enunciado no teorema a seguir. ■

**Teorema 9.** *Sejam $V$, $W$ e $Z$ espaços vetoriais de dimensão finita com bases $\mathcal{B}_V = \{x_1, \ldots, x_n\}$, $\mathcal{B}_W = \{y_1, \ldots, y_m\}$ e $\mathcal{B}_Z = \{z_1, \ldots, z_p\}$, respectivamente. Sejam $T : V \longrightarrow W$ e $S : W \longrightarrow Z$ aplicações lineares. Então*

$$[S \circ T]_{\mathcal{B}_V, \mathcal{B}_Z} = [S]_{\mathcal{B}_W, \mathcal{B}_Z} [T]_{\mathcal{B}_V, \mathcal{B}_W}.$$

**Prova:** Sejam

$$\begin{aligned} [T]_{\mathcal{B}_V, \mathcal{B}_W} &= A = (a_{ij})_{m \times n}, \\ [S]_{\mathcal{B}_W, \mathcal{B}_Z} &= B = (b_{ij})_{p \times m}. \end{aligned}$$

Temos

$$\begin{aligned} (S \circ T)(x_j) = S[T(x_j)] &= S\left(\sum_{i=1}^{m} a_{ij} y_i\right) = \sum_{i=1}^{m} a_{ij} S(y_i) \\ &= \sum_{i=1}^{m} a_{ij} \sum_{k=1}^{p} b_{ki} z_k = \sum_{k=1}^{p} \left(\sum_{i=1}^{m} b_{ki} a_{ij}\right) z_k \\ &= \sum_{k=1}^{p} (BA)_{kj} z_k. \end{aligned}$$

. ■

**Teorema 10.** *Sejam* $\mathcal{B} = \{x_1, \ldots, x_n\}$ *e* $\mathcal{B}' = \{x'_1, \ldots, x'_n\}$ *duas bases para o espaço vetorial* $V$. *Seja* $T : V \longrightarrow W$ *uma aplicação linear. Então existe uma única matriz invertível* $P$ *tal que*

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

*As colunas de* $P$ *são dadas pelas coordenadas dos vetores da base* $\mathcal{B}$ *com relação à base* $\mathcal{B}'$, *ou seja,*

$$P_i = [x_i]_{\mathcal{B}'}\,, \quad i = 1, \ldots, n.$$

**Prova:** Pelo Teorema 7,

$$[v]_{\mathcal{B}} = P^{-1}\,[v]_{\mathcal{B}'}\,,$$
$$[Tv]_{\mathcal{B}} = P^{-1}\,[Tv]_{\mathcal{B}'}\,,$$

para todo vetor $v \in V$. Por definição, também temos

$$[Tv]_{\mathcal{B}} = [T]_{\mathcal{B}}\,[v]_{\mathcal{B}}\,.$$

Logo,

$$P^{-1}\,[Tv]_{\mathcal{B}'} = [T]_{\mathcal{B}}\,P^{-1}\,[v]_{\mathcal{B}'}\,,$$

donde

$$[Tv]_{\mathcal{B}'} = P\,[T]_{\mathcal{B}}\,P^{-1}\,[v]_{\mathcal{B}'}\,.$$

Mas, por definição,

$$[Tv]_{\mathcal{B}'} = [T]_{\mathcal{B}'}\,[v]_{\mathcal{B}'}\,,$$

logo

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

o que significa que

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

■

Observe que a matriz $P$ nada mais é que a matriz que representa a aplicação linear $U$ que leva a base $\mathcal{B}'$ na base $\mathcal{B}$ em relação à base $\mathcal{B}'$:

$$Ux'_i = x_i,$$

isto é,

$$P = [U]_{\mathcal{B}'}\,.$$

Observe também que

$$P^{-1} = [U]_{\mathcal{B}}^{-1}\,.$$

Assim, o resultado do teorema anterior também pode ser expresso na forma

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

**Definição 12.** Sejam $A, B \in M_n(\mathbb{F})$ duas matrizes quadradas. Dizemos que $A$ e $B$ são **semelhantes** se existe uma matriz invertível $P \in M_n(\mathbb{F})$ tal que $B = P^{-1}AP$.

Segue da recíproca do Teorema 10 que duas matrizes são semelhantes se em cada espaço vetorial sobre $\mathbb{F}$ elas representam a mesma aplicação linear em relação a duas bases (possivelmente) distintas. Observe que similaridade é uma relação de equivalência em $M_n(\mathbb{F})$.

## 2.9 Funcionais Lineares e o Espaço Dual

**Definição 13.** Seja $V$ um espaço vetorial sobre um corpo $\mathbb{F}$. Uma aplicação linear $f : V \longrightarrow \mathbb{F}$ é chamada um **funcional linear**.

**Definição 14.** O espaço dos funcionais lineares $\mathcal{L}(V, \mathbb{F})$ é denotado por $V^*$ e chamado o **espaço dual** de $V$.

Observe que $\dim V^* = \dim V$.

**Definição 15.** Sejam $V$ um espaço vetorial sobre um corpo $\mathbb{F}$ e $\mathcal{B} = \{x_1, \ldots, x_n\}$ uma base para $V$. A base dual de $\mathcal{B}$ é a base $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ para $V^*$ definida por

$$f_i(x_j) = \delta_{ij}.$$

$\mathcal{B}^*$ é de fato um conjunto linearmente independente pois, se $\sum_{i=1}^{n} \alpha_i f_i = 0$, então

$$\alpha_j = \sum_{i=1}^{n} \alpha_i f_i(x_j) = 0(x_j) = 0$$

para todo $j$. O resultado a seguir mostra que $\mathcal{B}^*$ também gera $V^*$.

**Teorema 11.** *Sejam $V$ um espaço vetorial de dimensão finita sobre um corpo $\mathbb{F}$, $\mathcal{B} = \{x_1, \ldots, x_n\}$ uma base para $V$ e $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ a correspondente base dual para $V^*$. Então para todo funcional linear $f \in V^*$ nós temos*

$$f = \sum_{i=1}^{n} f(x_i) f_i$$

*e para todo vetor $x \in V$ nós temos*

$$x = \sum_{i=1}^{n} f_i(x) x_i.$$

**Prova:** Seja $f \in V^*$. Então

$$f(x_j) = \left[ \sum_{i=1}^{n} f(x_i) f_i \right] (x_j)$$

e por unicidade da aplicação linear definida em uma base, segue a primeira equação. A segunda equação segue do fato de que se $x = \sum_{j=1}^{n} \alpha_j x_j$, então

$$f_i(x) = \sum_{j=1}^{n} \alpha_j f_i(x_j) = \alpha_i.$$

■

Segue do Teorema 11 que as funções $f_i$ nada mais são que as funções coordenadas.

**Definição 16.** O espaço dos funcionais lineares $\mathcal{L}(V^*, \mathbb{F})$ definidos no dual de $V$ é denotado por $V^{**}$ e chamado o **duplo dual** de $V$.

**Teorema 12.** *Seja $V$ um espaço vetorial de dimensão finita sobre um corpo $\mathbb{F}$. Então $V$ e $V^{**}$ são canonicamente isomorfos. O isomorfismo canônico $\Phi : V \longrightarrow V^{**}$ é definido por*

$$\Phi(x) = L_x,$$

*onde $L_x \in V^{**}$ é definido por*

$$L_x(f) = f(x).$$

**Prova:** O funcional $L_x$ é de fato linear, pois

$$L_x(\alpha f + \beta g) = (\alpha f + \beta g)(x) = \alpha f(x) + \beta g(x) = \alpha L_x(f) + \beta L_x(g).$$

A aplicação $\Phi$ é linear pois

$$\Phi(\alpha x + \beta y) = L_{\alpha x + \beta y} = \alpha L_x + \beta L_y$$

porque para todo $f \in V^*$ temos

$$L_{\alpha x + \beta y}(f) = f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) = \alpha L_x(f) + \beta L_y(f).$$

$\Phi$ é injetiva porque $L_x = 0$ se e somente se $f(x) = 0$ para todo $f$, isto é, se e somente se $x$ é o vetor nulo (ou seja, se $x$ não é o vetor nulo existe um funcional que leva $x$ em um escalar não-nulo; um exemplo é um funcional coordenada apropriado). A sobrejetividade de $\Phi$ decorre do Corolário 2 quando observamos que

$$\dim V = \dim V^* = \dim V^{**}.$$

■

**Corolário 4.** *Seja $V$ um espaço vetorial de dimensão finita sobre um corpo $\mathbb{F}$. Se $L \in V^{**}$ então existe um único vetor $x \in V$ tal que*

$$L(f) = f(x)$$

*para todo $f \in V^*$.*

**Corolário 5.** *Seja $V$ um espaço vetorial de dimensão finita sobre um corpo $\mathbb{F}$. Toda base para $V^*$ é o dual de uma base para $V$.*

**Prova:** Seja $\mathcal{B}^* = \{f_1, \dots, f_n\}$ uma base qualquer para $V^*$. Seja $\mathcal{B}^{**} = \{L_1, \dots, L_n\}$ sua base dual em $V^{**}$, ou seja,

$$L_i(f_j) = \delta_{ij}.$$

Usando o corolário anterior, sejam $x_1, \dots, x_n \in V$ os únicos vetores tais que

$$L_i(f) = f(x_i)$$

para todo $f \in V^*$, para todo $i$. Então $\mathcal{B} = \{x_1, \dots, x_n\}$ é uma base para $V$ pelo Teorema 12 (isomorfismos levam bases em bases) e

$$f_i(x_j) = L_j(f_i) = \delta_{ij},$$

de modo que $\mathcal{B}^*$ é a base dual de $\mathcal{B}$. ■

## 2.10 Transposta de uma Aplicação Linear

**Teorema 13.** *Sejam $V, W$ espaços vetoriais sobre o corpo $\mathbb{F}$. Então, para cada aplicação linear $T : V \longrightarrow W$ existe uma única aplicação linear $T^t : W^* \longrightarrow V^*$ tal que*

$$\left(T^t g\right)(v) = g\,(Tv)$$

*para todo $v \in V$ e para todo $g \in W^*$.*

**Prova:** De fato, se $g$ é um funcional linear em $W$, se definirmos

$$f\,(v) = g\,(Tv)\,,$$

então $f$ é um funcional linear em $V$ ($f$ é a composta de duas aplicações lineares). Além disso, $T^t$ é linear porque

$$\begin{aligned}\left[T^t\left(\alpha g + \beta h\right)\right](v) &= \left(\alpha g + \beta h\right)(Tv) = \alpha g\,(Tv) + \beta h\,(Tv) = \alpha\left(T^t g\right)(v) + \beta\left(T^t h\right)(v) \\ &= \left[\alpha\left(T^t g\right) + \beta\left(T^t h\right)\right](v)\end{aligned}$$

para todo $v \in V$. ■

**Definição 17.** Seja $T : V \longrightarrow W$ uma aplicação linear entre dois espaços vetoriais. Então a **transposta** de $T$ é aplicação linear $T^t : W^* \longrightarrow V^*$ definida pelo teorema acima.

**Teorema 14.** *Sejam $V$, $W$ espaços vetoriais de dimensão finita sobre o corpo $\mathbb{F}$. Sejam $\mathcal{B}_V$ e $\mathcal{B}_W$ bases para $V$ e $W$, respectivamente e $\mathcal{B}_V^*$ e $\mathcal{B}_W^*$ suas respectivas bases duais. Seja $T : V \longrightarrow W$ uma aplicação linear. Se*

$$A = (a_{ij}) = [T]_{\mathcal{B}_V, \mathcal{B}_W}$$

*é a matriz de $T$ com respeito às bases $\mathcal{B}_V$ e $\mathcal{B}_W$, então*

$$A^t = (a_{ji}) = \left[T^t\right]_{\mathcal{B}_V^*, \mathcal{B}_W^*}$$

*é a matriz da aplicação transposta $T^t$ com respeito às bases duais $\mathcal{B}_V^*$ e $\mathcal{B}_W^*$.*

**Prova:** Sejam

$$\begin{aligned}\mathcal{B}_V &= \{x_1, \ldots, x_n\}\,, \quad \mathcal{B}_W = \{y_1, \ldots, y_m\}\,, \\ \mathcal{B}_V^* &= \{f_1, \ldots, f_n\}\,, \quad \mathcal{B}_W^* = \{g_1, \ldots, g_m\}\,.\end{aligned}$$

Denote por $B = (b_{ij})$ a matriz da aplicação transposta. Então, por definição,

$$\begin{aligned}Tx_j &= \sum_{i=1}^m a_{ij} y_i, \quad j = 1, \ldots, n, \\ T^t g_j &= \sum_{i=1}^n b_{ij} f_i, \quad j = 1, \ldots, m.\end{aligned}$$

Por outro lado,

$$\left(T^t g_j\right)(x_i) = g_j\,(Tx_i) = g_j\left(\sum_{k=1}^m a_{ki} y_k\right) = \sum_{k=1}^m a_{ki}\, g_j\,(y_k) = \sum_{k=1}^m \delta_{jk} a_{ki} = a_{ji}.$$

Como também

$$\left(T^t g_j\right)(x_i) = \left(\sum_{k=1}^n b_{kj} f_k\right)(x_i) = \sum_{k=1}^n b_{kj} f_k\,(x_i) = \sum_{k=1}^n \delta_{ik} b_{kj} = b_{ij},$$

concluímos que

$$b_{ij} = a_{ji}.$$

■

# Capítulo 3

# Determinantes

## 3.1 Definição

Definiremos a função determinante a partir das propriedades que queremos que ela satisfaça. Provaremos depois que de fato existe uma função que satisfaz estas propriedades.

**Definição 1.** Identificaremos o espaço das matrizes quadradas $M_n(\mathbb{F})$ com o espaço $\mathbb{F}^n \times \ldots \times \mathbb{F}^n$, identificando colunas de uma matriz com vetores de $\mathbb{F}^n$. Uma função **determinante** é uma função

$$D : \mathbb{F}^n \times \ldots \times \mathbb{F}^n \to \mathbb{F}$$

satisfazendo as seguintes propriedades:

(D1) $D$ é uma função *multilinear*:

$$D(c_1, \ldots, \alpha c_i + \beta c_i', \ldots, c_n) = \alpha D(c_1, \ldots, c_i, \ldots, c_n) + \beta D(c_1, \ldots, c_i', \ldots, c_n).$$

(D2) $D$ é uma função *alternada*:

$$D(c_1, \ldots, c_i, \ldots, c_j, \ldots, c_n) = -D(c_1, \ldots, c_j, \ldots, c_i, \ldots, c_n).$$

(D3) $D(I) = 1$.

Além de escrever $D(A)$ ou $D(c_1, \ldots, c_n)$ para a matriz $A$ cujas colunas são $c_1, \ldots, c_n$, também usaremos por exemplo a notação

$$\begin{aligned} D(\alpha c_i + \beta c_i') &= \alpha D(c_i) + \beta D(c_i'), \\ D(c_i, c_j) &= -D(c_j, c_i), \end{aligned}$$

quando estiver claro no contexto que as outras colunas são mantidas fixadas (neste caso, a função determinante é uma função apenas da(s) coluna(s) variáveis). Denotaremos por $\mathbb{K}$ qualquer subcorpo dos complexos (inclusive o próprio corpo dos complexos).

**Proposição 1.** *Seja $D : \mathbb{K}^n \times \ldots \times \mathbb{K}^n \to \mathbb{K}$ uma função multilinear. Então as afirmativas seguintes são equivalentes:*

(i) *$D$ é alternada.*

(ii) *Se $c_i = c_j$ para $i \neq j$, então $D(c_1, \ldots, c_n) = 0$.*

(iii) *Se $c_i = c_{i+1}$ para algum $i$, então $D(c_1, \ldots, c_n) = 0$.*

**Prova:** (i) $\Rightarrow$ (ii) Pois $D(c_i, c_i) = -D(c_i, c_i)$ implica $D(c_i, c_i) = 0$ em um subcorpo dos complexos. (ii) $\Rightarrow$ (i) Usando primeiro multilinearidade e depois (i), segue que

$$0 = D(c_i + c_j, c_i + c_j) = D(c_i, c_i) + D(c_i, c_j) + D(c_j, c_i) + D(c_j, c_j) = D(c_j, c_i) + D(c_i, c_j),$$

logo

$$D(c_i, c_j) = -D(c_j, c_i).$$

(ii) $\Rightarrow$ (iii) Imediato. (iii) $\Rightarrow$ (ii) Por indução, suponha que provamos que sempre que $c_i = c_{i+j}$ temos $D(c_1, \ldots, c_k) = 0$ para todo $j = 1, \ldots, k$. O caso $k = 1$ é exatamente (iii). Vamos provar que isso implica que se $c_i = c_{i+k+1}$ temos $D(c_1, \ldots, c_n) = 0$. De fato,

$$D(c_i, \ldots, c_{i+k}, c_{i+k+1}) = D(c_i, \ldots, c_{i+k}, c_{i+k} + c_{i+k+1}),$$

porque $D(c_i, \ldots, c_{i+k}, c_{i+k}) = 0$ por (iii), e

$$D(c_i, \ldots, c_{i+k}, c_{i+k} + c_{i+k+1}) = D(c_i, \ldots, c_{i+k} + c_{i+k+1}, c_{i+k} + c_{i+k+1})$$

pois, por hipótese de indução, $D(c_i, \ldots, c_{i+k+1}, c_{i+k} + c_{i+k+1}) = 0$ e

$$D(c_i, \ldots, c_{i+k} + c_{i+k+1}, c_{i+k} + c_{i+k+1}) = 0,$$

novamente por causa de (iii). ■

Para matrizes sobre um corpo $\mathbb{F}$ arbitrário, a propriedade (D2) na definição de determinante é trocada pela condição (ii) ou (iii) (elas são equivalentes para quaisquer corpos), e obtemos a mesma teoria de determinantes para corpos arbitrários.

## 3.2 Existência

**Definição 2.** Seja $A \in M_n(\mathbb{F})$. Denotamos por $A(i|j)$ a matriz em $M_{n-1}(\mathbb{F})$ obtida ao se eliminar a $i$-ésima linha e a $j$-ésima coluna de $A$.

**Teorema 1.** (Existência da Função Determinante) *Existe pelo menos uma função determinante.*

**Prova:** A função determinante é construída indutivamente. Em $M_1(\mathbb{F}) = \mathbb{F}$ definimos simplesmente $\det A = \det(a) = a$. Em $M_1(\mathbb{F})$, definimos

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

É fácil verificar que estas funções satisfazem as propriedades (D1)-(D3). Em geral, tendo definido uma função determinante em $M_1(\mathbb{F}), \ldots, M_{n-1}(\mathbb{F})$, definimos uma função determinante em $M_n(\mathbb{F})$ através da fórmula

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

fixando algum $i$ (por exemplo, $i = 1$). Esta é a chamada *fórmula do determinante através da expansão em cofatores segundo a i-ésima linha de A*. Vamos verificar que a função assim definida satisfaz as propriedades (D1)-(D3):

(D1) Sejam

$$\begin{aligned} A &= (a_{ij}) = (c_1, \ldots, \alpha c_k + \beta c_k', \ldots, c_n), \\ B &= (b_{ij}) = (c_1, \ldots, c_k, \ldots, c_n), \\ C &= (c_{ij}) = (c_1, \ldots, c_k', \ldots, c_n). \end{aligned}$$

Então

$$\det A = \sum_{\substack{j=1 \\ j \neq k}}^{n} (-1)^{i+j} a_{ij} \det A(i|j) + (-1)^{i+k} (\alpha b_{ik} + \beta c_{ik}) \det A(i|k).$$

Mas

$$\begin{aligned}
\sum_{\substack{j=1 \\ j \neq k}}^{n} (-1)^{i+j} a_{ij} \det A(i|j) &= \sum_{\substack{j=1 \\ j \neq k}}^{n} (-1)^{i+j} a_{ij} \left[\alpha \det B(i|j) + \beta \det C(i|j)\right] \\
&= \alpha \sum_{\substack{j=1 \\ j \neq k}}^{n} (-1)^{i+j} a_{ij} \det B(i|j) + \beta \sum_{\substack{j=1 \\ j \neq k}}^{n} (-1)^{i+j} a_{ij} \det C(i|j),
\end{aligned}$$

enquanto que

$$(-1)^{i+k} (\alpha b_{ik} + \beta c_{ik}) \det A(i|k) = \alpha (-1)^{i+k} a_{ik} \det B(i|k) + \beta (-1)^{i+k} c_{ik} \det C(i|k).$$

Portanto,

$$\begin{aligned}
\det A &= \alpha \sum_{j=1}^{n} (-1)^{i+j} b_{ij} \det B(i|j) + \beta \sum_{j=1}^{n} (-1)^{i+j} c_{ij} \det C(i|j) \\
&= \alpha \det B + \beta \det C.
\end{aligned}$$

(D2) Em vista da Proposição 1, basta provar que se $A$ tem duas colunas adjacentes iguais então $\det A = 0$. Seja $A = (c_1, \ldots, c_n)$ e suponha que $c_k = c_{k+1}$. Se $j \neq k$ e $j \neq k+1$, então a matriz $A(i|j)$ tem duas colunas iguais, logo $\det A(i|j) = 0$ e

$$\det A = (-1)^{i+k} a_{i,k} \det A(i|k) + (-1)^{i+k+1} a_{i,k+1} \det A(i|k+1).$$

Como $c_k = c_{k+1}$, temos $a_{i,k} = a_{i,k+1}$ e $A(i|k) = A(i|k+1)$, portanto $\det A = 0$.
(D3) Se $I_n$ é a matriz identidade $n \times n$, então $I_n(i|i) = I_{n-1}$ é a matriz identidade $(n-1) \times (n-1)$. Logo,

$$\det I_n = \sum_{j=1}^{n} (-1)^{i+j} \delta_{ij} \det I_n(i|j) = (-1)^{2i} \det I_{n-1} = 1.$$

■

## 3.3 Unicidade

### 3.3.1 Permutações

**Definição 3.** Seja $I = \{1, \ldots, n\}$. Uma **permutação** de grau $n$ é uma bijeção $p : I \longrightarrow I$. O conjunto das permutações de grau $n$ será denotado por $S_n$.

Uma permutação $p$ pode ser representada pela matriz

$$p = \begin{pmatrix} 1 & 2 & \ldots & n \\ p_1 & p_2 & \ldots & p_n \end{pmatrix}$$

ou pela matriz $A = (a_{ij})$, onde

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

ou seja, na coluna $j$ o valor 1 é colocado na entrada da linha $p_j$, os demais valores da coluna sendo iguais a 0. Em outras palavras, a matriz $A$ é a representação matricial de um operador linear $T : \mathbb{F}^n \longrightarrow \mathbb{F}^n$ tal que

$$Te_j = e_{p_j} \quad j = 1, \ldots, n.$$

Chamaremos a matriz $A$ de **matriz da permutação** $p$. (Apesar da primeira também ser uma matriz, a segunda representação matricial é mais importante porque a composta de permutações equivale à multiplicação destas matrizes, como veremos a seguir.) O conjunto das matrizes de permutação será denotado por $\mathcal{A}_n$.

**Exemplo 1.** A permutação de grau 5

$$p = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 5 & 3 & 1 \end{pmatrix}$$

tem como matriz

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

**Definição 4.** Um grupo é um conjunto $G$ munido de uma operação binária

$$\begin{aligned} G \times G &\longrightarrow G \\ (x, y) &\mapsto xy \end{aligned}$$

que satisfaz as seguintes propriedades:

(i) (Associatividade) Para todos $x, y, z \in G$ vale

$$x\,(yz) = (xy)\,z.$$

(ii) (Existência de Elemento Neutro) Existe um elemento $e \in G$ tal que para todo $x \in G$ vale

$$ex = xe = x.$$

(iii) (Existência de Inverso) Para todo elemento $x \in G$ existe um elemento $x^{-1} \in G$ tal que

$$xx^{-1} = x^{-1}x = e.$$

Qualquer espaço vetorial é um grupo com relação à operação de soma de vetores. O conjunto das matrizes invertíveis com relação à operação de multiplicação de matrizes também é um grupo. Dado um conjunto $S$, conjunto de todas as bijeções de $S$ em $S$ sob a operação de composição de funções é um grupo. Em particular,

**Proposição 2.** *O conjunto $S_n$ das permutações de grau $n$ sob a operação de composição de permutações é um grupo.*

**Exemplo 2.** Se

$$p = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 5 & 3 & 1 \end{pmatrix} \text{ e } q = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 5 & 4 & 3 \end{pmatrix},$$

então

$$qp = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 3 & 5 & 2 \end{pmatrix} \text{ e } p^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 2 & 4 & 1 & 3 \end{pmatrix}$$

**Definição 5.** Dados dois grupos $G$ e $H$, um **homomorfismo** entre eles é uma aplicação $\phi : G \longrightarrow H$ que preserva a operação de grupo, isto é,

$$\phi(xy) = \phi(x)\,\phi(y)\,.$$

O homomorfismo é um **isomorfismo** entre grupos se $\phi$ for uma bijeção.

Denote por $\mathcal{A}_n$ o grupo das matrizes cujas entradas em todas as linhas e colunas são 0 exceto em uma posição em cada linha e cada coluna, em que o valor é 1. Este é exatamente o conjunto das matrizes que representam permutações. Este conjunto torna-se um grupo sob a operação de multiplicação de matrizes.

**Proposição 3.** *O grupo $S_n$ é isomorfo a $\mathcal{A}_n$.*

**Prova:** Ambos os grupos são isomorfos ao grupo das aplicações lineares $T : \mathbb{F}^n \longrightarrow \mathbb{F}^n$ tais que

$$Te_j = e_{p_j} \quad j = 1, \ldots, n,$$

para alguma permutação $p$. ■
Em outras palavras, a composta de duas permutações é representada pelo produto das matrizes de suas permutações.

**Definição 6.** Uma **transposição** é uma permutação $\tau : I \longrightarrow I$ que satisfaz

$$\begin{aligned} \tau_i &= j, \\ \tau_j &= i, \\ \tau_k &= k \quad \text{para todo } k \neq i, j. \end{aligned}$$

**Exemplo 3.** A permutação de grau 7

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 2 & 6 & 4 & 5 & 3 & 7 \end{pmatrix}$$

cuja matriz é

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

é uma transposição.

### 3.3.2 Demonstração da Unicidade da Função Determinante

**Lema 1.** *Se $D_1, D_2 : \mathbb{F}^n \times \ldots \times \mathbb{F}^n \rightarrow \mathbb{F}$ são duas funções multilineares alternadas (i.e., satisfazem as propriedades* (D1) *e* (D2) *da Definição 1) tais que $D_1(I) = D_2(I)$, então $D_1(A) = D_2(A)$ para toda matriz de permutação $A$.*

**Prova:** Seja $A$ uma matriz de permutação. Um número finito de trocas de colunas (no máximo $n-1$) transforma a matriz $A$ na matriz identidade: transponha o vetor $e_1$ para a primeira coluna (se ele já não for a primeira coluna), obtendo uma matriz $A_1$; depois transponha o vetor $e_2$ para a segunda coluna (se ele já não for a segunda coluna), obtendo a matriz $A_2$ e assim sucessivamente.

Observe que trocar duas colunas de $A$ equivale efetivamente a multiplicar a permutação $p$ por uma transposição. De fato, se no primeiro passo

$$\begin{aligned}p_1 &= j \neq 1, \\ p_k &= 1,\end{aligned}$$

ou seja, em termos da matriz de permutação $A$, o vetor $e_1$ ocupa a coluna $k$, consideramos a transposição $\tau_1 = k$, de modo que

$$(p\tau)_1 = p(\tau_1) = p_k = 1,$$

isto é, a matriz da permutação $p\tau$ possui o vetor $e_1$ na primeira coluna. Em geral, se através de um certo número de permutações obtivemos uma permutação cuja matriz possui os vetores $e_1, \ldots, e_{m-1}$ nas colunas $1, \ldots, m-1$, respectivamente, se a coluna $m$ está ocupada pelo vetor $e_j \neq e_m$ e o vetor $e_m$ ocupa a coluna $k$, isto é, a permutação representada por esta matriz satisfaz

$$\begin{aligned}q_m &= j \neq m, \\ q_k &= m,\end{aligned}$$

consideramos a transposição $\tau_m = k$, de modo que

$$(q\tau)_m = q(\tau_m) = q_k = m,$$

e a matriz desta permutação possui o vetor $e_m$ na coluna $m$. Portanto, provamos que dada uma permutação $p$, existem transposições $\tau^1, \ldots, \tau^k$ tais que

$$p\tau^1 \ldots \tau^k = \text{id}.$$

Em particular, como a inversa de uma transposição é ela própria, segue que

$$p = \tau^k \ldots \tau^1.$$

Se forem necessárias $k$ transposições para transformar a matriz $A$ na matriz identidade, definindo

$$\begin{aligned}A_0 &= \text{ matriz de permutação de } p = A, \\ A_1 &= \text{ matriz de permutação de } p\tau^1, \\ A_2 &= \text{ matriz de permutação de } p\tau^1\tau^2, \\ &\vdots \\ A_{k-1} &= \text{ matriz de permutação de } p\tau^1\tau^2 \ldots \tau^{k-1}, \\ A_k &= \text{ matriz de permutação de } p\tau^1\tau^2 \ldots \tau^{k-1}\tau^k = \text{id } = I,\end{aligned}$$

concluímos que, se $D$ é uma função multilinear alternada,

$$D(A) = (-1) D(A_1) = (-1)^2 D(A_2) = \ldots = (-1)^{k-1} D(A_{k-1}) = (-1)^k D(I),$$

Assim, o valor de uma função multilinear alternada de qualquer matriz de permutação é caracterizado pelo valor que ela assume na matriz identidade. Em particular,

$$\begin{aligned}D_1(A) &= (-1)^k D_1(I), \\ D_2(A) &= (-1)^k D_2(I),\end{aligned}$$

donde segue o resultado. ■

**Teorema 2.** (Unicidade da Função Determinante) *Se* $D_1, D_2 : \mathbb{F}^n \times \ldots \times \mathbb{F}^n \to \mathbb{F}$ *são duas funções multilineares alternadas (i.e., satisfazem as propriedades* (D1) *e* (D2) *da Definição 1) tais que* $D_1(I) = D_2(I)$, *então* $D_1 = D_2$.

*Em particular, existe uma única função determinante.*

**Prova:** Sejam $c_1, \ldots, c_n \in \mathbb{F}^n$ vetores arbitrários. Escrevendo estes vetores em termos da base canônica de $\mathbb{F}^n$:

$$c_j = \sum_{i=1}^{n} a_{ij} e_i,$$

e utilizando a multilinearidade das funções $D_1$ e $D_2$, segue que

$$D_1(c_1, \ldots, c_n) = \sum_{i_1, \ldots, i_n = 1}^{n} a_{i_1 1} \ldots a_{i_n n} D_1(e_{i_1} \ldots e_{i_n}),$$
$$D_2(c_1, \ldots, c_n) = \sum_{p_1, \ldots, p_n = 1}^{n} a_{i_1 1} \ldots a_{i_n n} D_2(e_{i_1} \ldots e_{i_n}).$$

Como as funções são alternadas, temos que $D_1(e_{i_1} \ldots e_{i_n}) = D_2(e_{i_1} \ldots e_{i_n}) = 0$ sempre que a função $i : I \longrightarrow I$ não for uma permutação, isto é, sempre que $i$ for tal que $i_k = i_l$ para algum para $k \neq l$. Logo,

$$D_1(c_1, \ldots, c_n) = \sum_{p \in S_n} a_{p_1 1} \ldots a_{p_n n} D_1(e_{p_1} \ldots e_{p_n}),$$
$$D_2(c_1, \ldots, c_n) = \sum_{p \in S_n} a_{p_1 1} \ldots a_{p_n n} D_2(e_{p_1} \ldots e_{p_n}),$$

e o resultado segue do Lema 1. ■

**Corolário 1.** *O cálculo do determinante de uma matriz pode ser feito através da expansão em cofatores a partir de qualquer linha da matriz.*

**Corolário 2.** (Decomposição das Permutações) *Toda permutação é um produto de transposições.*

*Além disso, se* $p = \tau_1 \ldots \tau_k$ *é uma decomposição de* $p$ *como um produto de transposições, então*

$$\det(c_{p_1}, \ldots, c_{p_n}) = (-1)^k \det(c_1, \ldots, c_n).$$

*Em outras palavras, quando as posições das colunas de uma matriz são trocadas através de uma permutação, o sinal do determinante não se altera se esta permutação é um número par de transposições e o sinal muda se ela é um número ímpar de transposições.*

**Prova:** Segue da demonstração do Lema 1. ■

### 3.3.3 Fórmula do Determinante através de Permutações

**Definição 7.** O **sinal** de uma permutação $p$ é definido por

$$\operatorname{sign} p = \det A$$

onde $A$ é a matriz de $p$.

**Proposição 4.** *Se* $p = \tau_1 \ldots \tau_k$ *é uma decomposição de* $p$ *como um produto de transposições, então* $\operatorname{sign} p = (-1)^k$.

**Prova:** Pois se $A = (e_{p_1}, \ldots, e_{p_n})$, segue do Corolário 1 que

$$\operatorname{sign} p = \det (e_{p_1}, \ldots, e_{p_n}) = (-1)^k \det (e_1, \ldots, e_n) = (-1)^k \det I = (-1)^k .$$

■

**Corolário 3.** (Fórmula do Determinante através de Permutações) *Temos*

$$\det (c_1, \ldots, c_n) = \sum_{p \in S_n} (\operatorname{sign} p)\, a_{p_1 1} \ldots a_{p_n n}$$

**Prova:** Pelo Teorema 2, temos

$$\det (c_1, \ldots, c_n) = \sum_{p \in S_n} a_{p_1 1} \ldots a_{p_n n} \det (e_{p_1} \ldots e_{p_n}) .$$

O resultado segue da Proposição 4. ■

**Proposição 5.** *O sinal de uma permutação satisfaz as seguintes propriedades:*

(i) *Se $id$ é a permutação identidade, então* $\operatorname{sign} id = 1$.

(ii) *Se $\tau$ é uma transposição, então* $\operatorname{sign} \tau = -1$.

(iii) *Se $p, q$ são permutações, então* $\operatorname{sign} (pq) = \operatorname{sign} p \operatorname{sign} q$.

**Prova:** Se $p = \tau_1 \ldots \tau_k$ e $q = \sigma_1 \ldots \sigma_l$, então $pq = \tau_1 \ldots \tau_k \sigma_1 \ldots \sigma_l$, portanto

$$\operatorname{sign} (pq) = (-1)^{k+l} = (-1)^k (-1)^l = \operatorname{sign} p \operatorname{sign} q.$$

■

## 3.4 Propriedades

**Proposição 6.** (Determinante da Transposta)

$$\det \left(A^t\right) = \det A.$$

**Prova:** Temos

$$\det A = \sum_{p \in S_n} (\operatorname{sign} p)\, a_{p_1 1} \ldots a_{p_n n}.$$

Agora, observe que se $p_i = j$, então $i = p_j^{-1}$, logo $a_{p_i i} = a_{j p_j^{-1}}$. Como $\operatorname{sign} p = \operatorname{sign} p^{-1}$, segue que

$$\begin{aligned}
\det A &= \sum_{p^{-1} \in S_n} \left(\operatorname{sign} p^{-1}\right) a_{1 p_1^{-1}} \ldots a_{n p_n^{-1}} = \sum_{q \in S_n} (\operatorname{sign} q)\, a_{1 q_1} \ldots a_{n q_n} \\
&= \sum_{q \in S_n} (\operatorname{sign} q) \left(A^t\right)_{q_1 1} \ldots \left(A^t\right)_{q_n n} \\
&= \det \left(A^t\right) .
\end{aligned}$$

■

**Corolário 4.** *O cálculo do determinante de uma matriz pode ser feito através da expansão em cofatores a partir de qualquer coluna da matriz.*

**Proposição 7.** (Determinante do Produto)

$$\det(AB) = \det A \det B.$$

**Prova:** Denote as colunas de $A$, $B$ e $AB$ respectivamente por $A_j$, $B_j$ e $(AB)_j$. Observe que

$$\begin{aligned}
(AB)_{1j} &= \sum_{r=1}^{n} a_{1r} b_{rj}, \\
(AB)_{2j} &= \sum_{r=1}^{n} a_{2r} b_{rj}, \\
&\vdots \\
(AB)_{nj} &= \sum_{r=1}^{n} a_{nr} b_{rj},
\end{aligned}$$

de modo que podemos escrever

$$(AB)_j = \sum_{r=1}^{n} b_{rj} A_r.$$

Portanto,

$$\det(AB) = \det\left(\sum_{r=1}^{n} b_{r1} A_r, \ldots, \sum_{r=1}^{n} b_{rn} A_r\right).$$

Expandindo esta expressão (isto é, usando a multilinearidade e alternalidade do determinante da mesma forma que fizemos no Teorema 2), obtemos

$$\begin{aligned}
\det(AB) &= \sum_{p \in S_n} b_{p_1 1} \ldots b_{p_n n} \det(A_{p_1}, \ldots, A_{p_n}) = \sum_{p \in S_n} (\operatorname{sign} p)\, b_{p_1 1} \ldots b_{p_n n} \det(A_1, \ldots, A_n) \\
&= \sum_{p \in S_n} (\operatorname{sign} p)\, b_{p_1 1} \ldots b_{p_n n} \det A = \det A \left[\sum_{p \in S_n} (\operatorname{sign} p)\, b_{p_1 1} \ldots b_{p_n n}\right] \\
&= \det A \det B.
\end{aligned}$$

■

**Corolário 5.** (Determinante da Inversa) *Se* $A$ *for invertível, então* $\det A \neq 0$ *e*

$$\det\left(A^{-1}\right) = \frac{1}{\det A}.$$

**Prova:** Pois

$$\det\left(AA^{-1}\right) = \det A \det\left(A^{-1}\right)$$

e

$$\det\left(AA^{-1}\right) = \det I = 1.$$

■

**Corolário 6.** *Matrizes semelhantes possuem o mesmo determinante.*

**Prova:** Pois, se $B = P^{-1}AP$, então

$$\det B = \det\left(P^{-1}AP\right) = \det P^{-1} \det A \det P = \frac{1}{\det P} \det A \det P = \det A.$$

■

## 3.5 Regra de Cramer

Seja $A$ uma matriz $n \times n$ e $b \in \mathbb{F}^n$ um vetor. Considere a equação (que corresponde a um sistema linear)

$$Ax = b.$$

Suponha que $x = \sum_{j=1}^{n} x_j e_j$ seja uma solução para esta equação. Se $A_j$ denota a $j$-ésima coluna da matriz $A$, temos

$$Ax = A\left(\sum_{j=1}^{n} x_j e_j\right) = \sum_{j=1}^{n} x_j A e_j = \sum_{j=1}^{n} x_j A_j = b.$$

Logo,

$$b_i = \sum_{j=1}^{n} x_j a_{ij}.$$

Denote por $\widetilde{A}_k$ a matriz obtida de $A$ através da substituição da $k$-ésima coluna de $A$ pelo vetor $b$. Então

$$\widetilde{A}_k = (A_1, \ldots, A_{k-1}, b, A_{k+1}, \ldots, A_n) = \left(A_1, \ldots, A_{k-1}, \sum_{j=1}^{n} x_j A_j, A_{k+1}, \ldots, A_n\right).$$

Daí,

$$\begin{aligned}
\det \widetilde{A}_k &= \det \left(A_1, \ldots, A_{k-1}, \sum_{j=1}^{n} x_j A_j, A_{k+1}, \ldots, A_n\right) \\
&= \sum_{j=1}^{n} x_j \det (A_1, \ldots, A_{k-1}, A_j, A_{k+1}, \ldots, A_n) \\
&= x_k \det (A_1, \ldots, A_{k-1}, A_k, A_{k+1}, \ldots, A_n) \\
&= x_k \det A.
\end{aligned}$$

Portanto, se $\det A \neq 0$ e existir uma solução $x$ para o sistema $Ax = b$, então esta solução é única e é dada por

$$x_k = \frac{\det \widetilde{A}_k}{\det A}. \tag{3.1}$$

Podemos dizer mais: se $\det A \neq 0$, então a expressão acima fornece a única solução para o sistema $Ax = b$ (veja Teorema 3 a seguir). Esta é a chamada **regra de Cramer**.

**Definição 8.** A **adjunta clássica** da matriz $A$ é definida como sendo a matriz transposta da matriz de cofatores da matriz $A$, isto é,

$$(\text{adj } A)_{ij} = (-1)^{i+j} \det A\,(j|i)\,.$$

**Teorema 3.** *Temos*

$$(\text{adj } A)\, A = A\, (\text{adj } A) = (\det A)\, I.$$

*Em particular, se* $\det A \neq 0$*, então* $A$ *é invertível e*

$$A^{-1} = \frac{\text{adj } A}{\det A}.$$

**Prova:** Temos

$$\begin{aligned}
\left[(\text{adj } A)\, A\right]_{ij} &= \sum_{r=1}^{n} (\text{adj } A)_{ir}\, a_{rj} = \sum_{r=1}^{n} (-1)^{i+r} \det A\,(r|i)\, a_{rj} \\
&= \sum_{r=1}^{n} (-1)^{i+r}\, a_{rj} \det A\,(r|i) \\
&= \sum_{r=1}^{n} (-1)^{i+r}\, b_{ri} \det B\,(r|i) \\
&= \det B.
\end{aligned}$$

onde a matriz $B$ é obtida a partir da matriz $A$ quando substituímos a $i$-ésima coluna de $A$ pela sua $j$-ésima coluna. Assim, se $i \neq j$, temos que a matriz $B$ possui duas colunas iguais, logo $\det B = 0$ e concluímos que

$$\left[(\text{adj } A)\, A\right]_{ij} = 0 \quad \text{se } i \neq j.$$

Se $i = j$, então

$$\sum_{r=1}^{n} (-1)^{i+r}\, a_{rj} \det A\,(r|i) = \sum_{r=1}^{n} (-1)^{i+r}\, a_{ri} \det A\,(r|i) = \det A\,.$$

Em outras palavras,

$$\left[(\text{adj } A)\, A\right]_{ij} = (\det A)\, \delta_{ij}.$$

Portanto, $(\text{adj } A)\, A = (\det A)\, I$.

Para provar que $A\,(\text{adj } A) = (\det A)\, I$, observe que $A^t\,(i|j) = A\,(j|i)^t$, logo

$$(-1)^{i+j} \det A^t\,(j|i) = (-1)^{j+i} \det A\,(j|i)^t = (-1)^{j+i} \det A\,(j|i)\,,$$

o que significa que o cofator $i, j$ da transposta $A^t$ é o cofator $j, i$ de $A$, ou seja,

$$\text{adj } \left(A^t\right) = (\text{adj } A)^t\,,$$

ou seja, a adjunta clássica da transposta de $A$ é a transposta da adjunta clássica de $A$. Já sabemos que

$$\left(\text{adj } A^t\right) A^t = \left(\det A^t\right) I,$$

donde

$$(\text{adj } A)^t\, A^t = (\det A)\, I.$$

Tomando a transposta de ambos os lados, obtemos o resultado desejado. ■

**Corolário 7.** *Uma matriz é invertível se e somente se o seu determinante é diferente de zero.*

A regra de Cramer pode ser obtida a partir da adjunta clássica. De fato, se $Ax = b$, então

$$(\text{adj } A)\, Ax = (\text{adj } A)\, b,$$

donde

$$(\det A)\, x = (\text{adj } A)\, b.$$

Se $\det A \neq 0$, temos que

$$x = \frac{\text{adj } A}{\det A} b,$$

ou seja,

$$x_j = \frac{1}{\det A} \sum_{i=1}^{n} (\text{adj } A)_{ji}\, b_i = \frac{1}{\det A} \sum_{i=1}^{n} (-1)^{i+j}\, b_i \det A\,(i|j) = \frac{1}{\det A} \det \widetilde{A}_j.$$

# Capítulo 4

# Autovalores e Autovetores

## 4.1 Polinômios

Nesta seção introduziremos alguns fatos básicos sobre polinômios.

**Definição 01.** Seja $\mathbb{F}$ um corpo. Denotaremos o espaço vetorial das funções (seqüências) $f : \mathbb{N} \longrightarrow \mathbb{F}$ por $\mathbb{F}^\infty$.

Observe que os vetores em $\mathbb{F}^\infty$ nada mais são que seqüências de escalares em $\mathbb{F}$: $f = (f_0, f_1, f_2, \ldots)$. Definimos um produto em $\mathbb{F}^\infty$ associando a cada par de vetores $f, g$ o vetor $fg$ definido por

$$(fg)_n = \sum_{i=0}^{n} f_i g_{n-i} \quad n = 0, 1, 2, \ldots$$

Deste modo $\mathbb{F}^\infty$ torna-se uma álgebra linear comutativa com identidade sobre $\mathbb{F}$ (o vetor $1 := (1, 0, 0, \ldots)$ é a identidade). O vetor $(0, 1, 0, \ldots)$ desempenha um papel fundamental e é denotado por $x$. Observe que

$$\begin{aligned} x^2 &= (0, 0, 1, 0, \ldots), \\ x^3 &= (0, 0, 0, 1, 0, \ldots), \\ &\vdots \\ x^n &= \Big(0, 0, 0, 0, \ldots, 0, \underset{n}{1}, 0, \ldots\Big). \end{aligned}$$

Também denotamos $x^0 := 1$. A álgebra linear $\mathbb{F}^\infty$ é as vezes chamada a **álgebra das séries formais** sobre $\mathbb{F}$, o elemento $f = (f_0, f_1, f_2, \ldots)$ sendo freqüentemente denotado na forma

$$f = \sum_{n=0}^{\infty} f_n x^n.$$

O conjunto $\{1, x, x^2, \ldots\}$ é um conjunto linearmente independente mas não é uma base para $\mathbb{F}^\infty$.

**Definição 02.** Denotamos por $\mathbb{F}[x]$ o subespaço de $\mathbb{F}^\infty$ gerado por $1, x, x^2, \ldots$. Um elemento de $\mathbb{F}[x]$ é chamado um **polinômio** sobre $\mathbb{F}$.

O **grau** de um polinômio $f \neq 0$ é o inteiro $n$ tal que $f_n \neq 0$ e $f_i = 0$ para todo $i > n$.

Em vista das definições acima, um polinômio $f \in \mathbb{F}[x]$ pode então ser denotado na forma

$$f = f_0 x^0 + f_1 x + f_2 x^2 + \ldots + f_n x^n.$$

Os escalares $f_0, f_1, f_2, \ldots, f_n$ são chamados os **coeficientes** do polinômio $f$. Um polinômio $f$ de grau $n$ é chamado um **polinômio mônico** se $f_n = 1$.

**Proposição 01.** *Sejam $f, g$ polinômios não nulos sobre $\mathbb{F}$. Então*

(i) *$fg$ é um polinômio não-nulo.*

(ii) grau $(fg) =$ grau $f+$ grau $g$.

(iii) *Se $f, g$ são mônicos, então $fg$ é mônico.*

(iv) *$fg$ é um polinômio escalar se e somente se $f, g$ são polinômios escalares.*

(v) Se $f + g \neq 0$, então grau $(f + g) \leqslant \max($ grau $f,$ grau $g)$.

**Prova:** Suponha que $f$ e $g$ tem graus $n$ e $m$, respectivamente, ou seja,

$$f = \sum_{i=0}^{n} f_i x^i \quad \text{e} \quad g = \sum_{i=0}^{m} g_i x^i,$$

com

$$f_n \neq 0 \quad \text{e} \quad g_m \neq 0.$$

Por definição, se $k$ é um natural, temos

$$(fg)_{n+m+k} = \sum_{i=0}^{n+m+k} f_i g_{n+m+k-i}.$$

Claramente, para que tenhamos $f_i g_{m+n+k+i} \neq 0$, é necessário que $i \leqslant n$ (pois $f_i = 0$ se $i > n$) e que $n+m+k-i \leqslant m$ (pois $g_{n+m+k-i} = 0$ se $n+m+k-i > m$), ou seja, $i \geqslant n+k$. Assim, se $f_i g_{m+n+k+i} \neq 0$, então $n + k \leqslant i \leqslant n$, o que implica $k = 0$ e $i = n$. Portanto,

$$(fg)_{n+m} = f_n g_m \tag{4.1}$$

e

$$(fg)_{n+m+k} = 0 \text{ se } k > 0. \tag{4.2}$$

As afirmações (i), (ii) e (iii) seguem destes dois fatos. A afirmação (iv) é uma conseqüência de (ii) e a afirmação (v) é óbvia. ■

**Corolário 01.** $\mathbb{F}[x]$ *é uma álgebra linear comutativa com identidade sobre* $\mathbb{F}$.

**Corolário 02.** *Sejam $f, g, h$ polinômios sobre $\mathbb{F}$ tais que $f \neq 0$ e $fg = fh$. Então $g = h$.*

**Prova:** O resultado segue imediatamente da Proposição 01 (i) quando escrevemos $f(g - h) = 0$. ■

**Corolário 03.** *Sejam*

$$f = \sum_{i=0}^{n} f_i x^i \quad e \quad g = \sum_{j=0}^{m} g_j x^j$$

*Então*

$$fg = \sum_{j=0}^{n+m} \left( \sum_{i=0}^{j} f_i g_{j-i} \right) x^j = \sum_{i=0}^{n} \sum_{j=0}^{m} f_i g_j x^{i+j}.$$

**Definição 03.** Seja $\mathcal{A}$ uma álgebra linear comutativa com identidade sobre $\mathbb{F}$ e para cada elemento $a \in \mathcal{A}$ adote a convenção $a^0 = 1$, onde 1 é a identidade de $\mathcal{A}$. A cada polinômio $f = \sum_{i=0}^{n} f_i x^i$ sobre $\mathbb{F}$ associamos um elemento $f(a) \in \mathcal{A}$ pela regra

$$f(a) = \sum_{i=0}^{n} f_i a^i.$$

**Proposição 02.** *Seja $\mathcal{A}$ uma álgebra linear comutativa com identidade sobre o corpo $\mathbb{F}$. Sejam $f, g$ polinômios sobre $\mathbb{F}$, $a \in \mathcal{A}$ e $\alpha, \beta \in \mathbb{F}$. Então*

(i) $(\alpha f + \beta g)(a) = \alpha f(a) + \beta g(a)$.

(ii) $(fg)(a) = f(a)\, g(a)$.

**Prova:** Provaremos apenas (ii). Sejam

$$f = \sum_{i=0}^{n} f_i x^i \quad \text{e} \quad g = \sum_{j=0}^{m} g_j x^j.$$

Então, pelo Corolário 03,

$$fg = \sum_{i,j=0}^{n,m} f_i g_j x^{i+j},$$

de modo que (usando (i)),

$$(fg)(a) = \sum_{i,j=0}^{n,m} f_i g_j a^{i+j} = \left( \sum_{i=0}^{n} f_i a^i \right) \left( \sum_{j=0}^{m} g_j a^j \right) = f(a)\, g(a).$$

■

**Corolário 04.** *Se $T : V \longrightarrow V$ é um operador linear definido sobre o espaço vetorial $V$ sobre o corpo $\mathbb{F}$ e $p$ é um polinômio sobre $\mathbb{F}$ tal que $p = (x - r_1) \dots (x - r_n)$, então*

$$p(T) = (T - r_1 I) \dots (T - r_n I).$$

**Lema 01.** *Sejam $f, d$ polinômios não nulos sobre $\mathbb{F}$ tais que* grau $d \leqslant$ grau $f$. *Então existe um polinômio $g$ sobre $\mathbb{F}$ tal que ou*

$$f - dg = 0,$$

*ou*

$$\text{grau } (f - dg) < \text{ grau } f.$$

**Prova:** Escreva

$$f = f_n x^n + \sum_{i=0}^{n-1} f_i x^i \quad \text{e} \quad d = d_m x^m + \sum_{i=0}^{m-1} d_i x^i,$$

com

$$f_n \neq 0 \quad \text{e} \quad d_m \neq 0.$$

Então $m \leqslant n$ e ou

$$f - \left( \frac{f_n}{d_m} \right) x^{n-m} d = 0,$$

ou

$$\text{grau } \left[ f - \left( \frac{f_n}{d_m} \right) x^{n-m} d \right] < \text{ grau } f.$$

Tomamos $g = f - \left( \dfrac{f_n}{d_m} \right) x^{n-m} d$. ■

**Teorema 01.** (Divisão de Polinômios) *Se $f, d$ são polinômios sobre $\mathbb{F}$ com $d \neq 0$, Então existem polinômios únicos $q, r$ sobre $\mathbb{F}$ tais que*

$$f = dq + r$$

*com ou $r = 0$ ou*

$$\text{grau } r < \text{ grau } d.$$

**Prova:** Se $f = 0$ ou grau $f <$ grau $d$, podemos tomar $q = 0$ e $r = d$.

No caso em que $f \neq 0$ e grau $d \leqslant$ grau $f$, existe um polinômio $g_1$ tal que $f - dg_1 = 0$ ou grau $(f - dg_1) <$ grau $f$. Se grau $(f - dg_1) <$ grau $d$, tomamos $q = g_1$ e $r = f - dg_1$. Caso contrário, usamos novamente o lema anterior e encontramos um polinômio $g_2$ tal que ou $(f - dg_1) - dg_2 = 0$ ou grau $[(f - dg_1) - dg_2] <$ grau $(f - dg_1)$. Este processo pode ser continuado até obter o resultado deste teorema.

Para provar a unicidade dos polinômios $q$ e $r$, suponha que também existam outros polinômios $q_1, r_1$ tais que

$$f = dq_1 + r_1$$

com $r_1 = 0$ ou grau $r_1 <$ grau $d$. Então $dq + r = dq_1 + r_1$, donde

$$d\,(q - q_1) = r_1 - r.$$

Se $q \neq q_1$, então

$$\text{grau } d + \text{grau } \ (q - q_1) = \text{grau } \ (r_1 - r)\,,$$

mas isso contradiz grau $(r_1 - r) <$ grau $d$. Portanto $q = q_1$, o que implica $r = r_1$. ■

**Definição 04.** Dados $f, d$ polinômios sobre $\mathbb{F}$ com $d \neq 0$, se existe um polinômio $q$ sobre $\mathbb{F}$ tal que $f = dq$, dizemos que $d$ **divide** $f$ (ou, alternativamente, que $f$ é **divisível** por $d$, ou ainda que $f$ é um **múltiplo** de $d$) e chamamos $q$ o **quociente** de $f$ por $d$ e denotamos $q = f/d$. Se $f = dq + r$ com $r \neq 0$, dizemos que $r$ é o **resto** da divisão de $f$ por $q$.

**Corolário 05.** *Seja $f$ um polinômio sobre $\mathbb{F}$ e $\alpha \in \mathbb{F}$. Então $f$ é divisível por $x - \alpha$ se e somente se $f(\alpha) = 0$.*

**Prova:** Pelo teorema, $f = (x - \alpha)\,q + r$, onde $r$ é um polinômio escalar (isto é, ou $r = 0$ ou grau $r <$ grau $(x - \alpha) = 1$, isto é, grau $r = 0$). Segue da Proposição 02 que

$$f(\alpha) = (\alpha - \alpha)\,q(\alpha) + r(\alpha) = r(\alpha) = r.$$

Portanto, $r = 0$ se e somente se $f(\alpha) = 0$. ■

**Definição 05.** Dado um polinômio $f$ sobre $\mathbb{F}$, dizemos que $\alpha \in \mathbb{F}$ é uma **raiz** de $f$ se $f(\alpha) = 0$. Se $\alpha$ é uma raiz de $f$, a **multiplicidade** de $\alpha$ como uma raiz de $f$ é o maior inteiro positivo $r$ tal que $(x - \alpha)^r$ divide $f$.

**Corolário 06.** *Um polinômio de grau $n$ sobre $\mathbb{F}$ tem no máximo $n$ raízes.*

**Prova:** Por indução em $n$. O resultado é obviamente verdadeiro para polinômios de grau 0 e de grau 1. Assuma o resultado verdadeiro para polinômios de grau $n - 1$. Se $f$ possui grau $n$ e $\alpha$ é uma raiz de $f$ então $f = (x - \alpha)\,q$ e $q$ é um polinômio de grau $n - 1$, que tem no máximo $n - 1$ raízes, pela hipótese de indução. Como $f(\beta) = (\beta - \alpha)\,q(\beta) = 0$ se e somente se $\alpha = \beta$ ou se $\beta$ for uma raiz de $q$, segue o resultado. ■

**Teorema 02.** (Teorema Fundamental da Álgebra) *Todo polinômio sobre o corpo $\mathbb{C}$ dos complexos possui pelo menos uma raiz.*

**Corolário 07.** *Todo polinômio complexo $f$ pode ser escrito, a menos de uma constante complexa, como o produto de polinômios de grau 1, ou seja,*

$$f = c\,(x - r_1) \ldots (x - r_n)\,.$$

**Prova:** Pelo Teorema Fundamental da Álgebra, se $f$ tem grau $n$, temos $f = (x - r_1)\, q_1$ e $q_1$ tem grau $n-1$. Podemos aplicar novamente o Teorema Fundamental da Álgebra ao polinômio $q$. Procedendo desta forma, chegamos a $f = (x - r_1) \ldots (x - r_n)\, q_n$, com grau $q_n = 0$, isto é, $q_n = c \in \mathbb{C}$. ■

## 4.2 Autovalores, Autovetores e Autoespaços

**Definição 1.** Seja $V$ um espaço vetorial sobre o corpo $\mathbb{F}$ e $T : V \longrightarrow V$ um um operador linear. Dizemos que $\lambda \in \mathbb{F}$ é um **autovalor** de $T$ se existe um vetor $v \in V$, $v \neq 0$, tal que

$$Tv = \lambda v.$$

Se $\lambda$ é um autovalor de $T$ e $v$ é qualquer vetor (mesmo nulo) tal que $Tv = \lambda v$, dizemos que $v$ é um **autovetor** de $T$ associado a $\lambda$.

Observe que o conjunto $V_\lambda = \{v \in V : Tv = \lambda v\}$ dos autovetores de $T$ associados ao autovalor $\lambda$ é um subespaço de $V$, pois é o núcleo do operador $T - \lambda I$; ele é chamado o **autoespaço** de $T$ associado a $\lambda$. Como $\lambda$ é um autovalor de $T$ se e somente se o operador $T - \lambda I$ não é injetivo, segue que $\lambda$ é um autovalor de $T$ se e somente se $\det(T - \lambda I) = 0$. Portanto, é importante estudar o polinômio $\det(xI - T)$.

**Definição 2.** O polinômio $p_c(x) = \det(xI - T)$ é chamado o **polinômio característico** de $T$.

O polinômio característico de $T$ é um polinômio mônico de grau $n$, como pode-se ver da fórmula para o determinante em termos de cofatores (exercício) e os autovalores de $T$ são exatamente as raízes do seu polinômio característico (o nome *polinômio característico* vem do fato de que autovalores são também chamados *valores característicos*). Este último fato também implica que a análise de um operador linear depende muito do corpo sobre o qual o espaço vetorial está definido, pois se um polinômio possui raízes em um corpo $\mathbb{F}$, estas raízes podem não estar presentes em um subcorpo $\mathbb{F}'$; assim, o mesmo operador $T : V \longrightarrow V$ pode não possuir autovalores quando $V$ é considerado sobre $\mathbb{F}'$, ao invés de ser considerado sobre $\mathbb{F}$. Em particular, ele pode ser diagonalizável quando considerado sobre $\mathbb{F}$, mas não quando considerado sobre $\mathbb{F}'$.

**Proposição 1.** *Um operador linear sobre um espaço de dimensão $n$ possui no máximo $n$ autovalores distintos.*

**Prova:** Pois um polinômio em $\mathbb{F}[x]$ possui no máximo $n$ raízes. ■

**Proposição 2.** *Um operador linear sobre um espaço vetorial complexo possui pelo menos um autovalor.*

**Prova:** Pois todo polinômio em $\mathbb{C}[x]$ possui pelo menos uma raiz.

Vamos dar uma segunda demonstração deste resultado sem usar o polinômio característico. Seja $T : V \longrightarrow V$ um um operador linear sobre um espaço vetorial complexo $V$ de dimensão $n$. Então, dado $v \neq 0$, os vetores $v, Tv, \ldots, T^n v$ são linearmente dependentes em $V$ (pois constituem um conjunto com $n+1$ vetores), logo existem escalares $a_0, a_1, \ldots, a_n$ não todos nulos tais que

$$a_0 v + a_1 T v + \ldots + a_n T^n v = 0.$$

Considere o polinômio em $\mathbb{C}[z]$ com estes coeficientes, isto é,

$$a_0 + a_1 z + \ldots + a_n z^n.$$

Em $\mathbb{C}[z]$ este polinômio pode ser fatorado em um produto de termos lineares:

$$a_0 + a_1 z + \ldots + a_n z^n = c(z - \lambda_1)\ldots(z - \lambda_m)$$

(se $a_n \neq 0$, teríamos exatamente $n$ termos e $c = a_n$). Segue que

$$0 = (a_0 I + a_1 T + \ldots + a_n T^n) v = a_n (T - \lambda_1 I)\ldots(T - \lambda_m I) v.$$

Em particular, necessariamente temos que pelo menos algum operador $T - \lambda_j I$ não é injetivo (pois a composta de bijeções é uma bijeção), e neste caso $\lambda_j$ é um autovalor para $T$. ■

**Definição 3.** O conjunto dos autovalores de $T$ é chamado o **espectro** de $T$ e será denotado por $\sigma(T)$. A **multiplicidade algébrica** de um autovalor de $T$ é a sua multiplicidade como raiz do polinômio característico, isto é, $d$ é a multiplicidade algébrica do autovalor $\lambda$ se o polinômio característico de $T$ se escreve na forma $p_c(x) = (x - \lambda)^d q(x)$ e $q(x)$ não possui $\lambda$ como raiz.

De maneira análoga definimos os autovalores e o polinômio característico de uma matriz. É claro que os autovalores e o polinômio característico de um operador são os autovalores e o polinômio característico de qualquer uma de suas representações matriciais. Uma demonstração independente deste resultado é dada a seguir:

**Proposição 3.** *Matrizes semelhantes possuem os mesmos autovalores.*

**Prova:** Pois se $B = P^{-1}AP$, então

$$\begin{aligned} \det(xI - B) &= \det\left(xI - P^{-1}AP\right) = \det\left[P^{-1}(xI - A)P\right] \\ &= \det P^{-1} \det(xI - A) \det P = \det(xI - A). \end{aligned}$$

■

**Exemplo 1.** A matriz

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

possui o polinômio característico

$$\det(xI - A) = \det \begin{bmatrix} x & 1 \\ -1 & x \end{bmatrix} = x^2 + 1.$$

Se $A$ é considerada uma matriz sobre $\mathbb{C}$, então $A$ possui dois autovalores distintos, $\pm i$, enquanto que sobre $\mathbb{R}$ $A$ não possui autovalores. □

## 4.3 Operadores Diagonalizáveis

A importância do estudo de autovalores está contida na próxima definição:

**Definição 4.** Dizemos que um operador linear $T : V \longrightarrow V$ é **diagonalizável** se existir uma base para $V$ formada por autovetores de $T$.

Se $\mathcal{B} = \{v_1, \ldots, v_n\}$ é uma base para $V$ constituída de autovetores de $T$, isto é, se

$$Tv_i = \lambda_i v_i \quad \text{para } i = 1, \ldots, n,$$

então a matriz de $T$ com relação a esta base é uma matriz diagonal, com os autovalores de $T$ ocupando a diagonal principal da matriz:

$$[T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}.$$

Observe que os autovalores de $T$ não precisam ser distintos para isso acontecer (de fato, eles podem ser todos iguais, caso em que o operador $T$ é um múltiplo escalar da identidade, o múltiplo escalar sendo precisamente o único autovalor de $T$).

**Proposição 4.** *Um conjunto de autovetores não-nulos correspondentes a autovalores dois a dois distintos é linearmente independente.*

**Prova:** A demonstração é por indução sobre o número de autovetores. Suponha o resultado provado para um conjunto de $k-1$ autovetores. Sejam $\lambda_1, \dots, \lambda_k$ um conjunto de autovalores de $T$, dois a dois distintos. Sejam $v_1, \dots, v_k$ autovetores não-nulos respectivamente correspondentes a estes autovalores. Suponha que

$$\alpha_1 v_1 + \dots + \alpha_k v_k = 0.$$

Aplicando $T$ a esta equação obtemos

$$\alpha_1 T v_1 + \dots + \alpha_k T v_k = 0,$$

ou seja,

$$\alpha_1 \lambda_1 v_1 + \dots + \alpha_k \lambda_k v_k = 0. \tag{4.3}$$

Por outro lado, se ao invés de aplicar $T$ à equação como fizemos, nós a multiplicarmos pelo autovalor $\lambda_k$ obtemos

$$\alpha_1 \lambda_k v_1 + \dots + \alpha_k \lambda_k v_k = 0. \tag{4.4}$$

Subtraindo as duas equações, segue que

$$\alpha_1 (\lambda_1 - \lambda_k) v_1 + \dots + \alpha_{k-1} (\lambda_{k-1} - \lambda_k) v_{k-1} = 0. \tag{4.5}$$

Pela hipótese de indução, sabemos que $v_1, \dots, v_{k-1}$ são linearmente independentes, logo

$$\alpha_1 (\lambda_1 - \lambda_k) = \dots = \alpha_{k-1} (\lambda_{k-1} - \lambda_k) = 0.$$

Como $\lambda_k - \lambda_i \neq 0$ para todo $i \neq k$, segue que

$$\alpha_1 = \dots = \alpha_{k-1} = 0.$$

Em particular, de $\alpha_1 v_1 + \dots + \alpha_{k-1} v_{k-1} + \alpha_k v_k = 0$ segue que

$$\alpha_k v_k = 0,$$

logo $\alpha_k = 0$ também. ■

**Teorema 1.** *Seja $V$ um espaço vetorial de dimensão $n$. Se o operador linear $T : V \longrightarrow V$ possui $n$ autovalores distintos, então ele é diagonalizável.*

**Prova:** Sejam $\lambda_1, \dots, \lambda_n$ os autovalores de $T$ e $v_1, \dots, v_n$ autovetores respectivamente correspondentes. Segue do resultado anterior que $\{v_1, \dots, v_n\}$ é um subconjunto linearmente independente de $V$, logo é uma base para $V$. ■

**Exemplo 2.** A matriz

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

possui o polinômio característico

$$\det(xI - A) = \det \begin{bmatrix} x & 1 & 0 & 0 \\ -1 & x & 0 & 0 \\ 0 & 0 & x & 1 \\ 0 & 0 & -1 & x \end{bmatrix} = \left(x^2+1\right)^2.$$

Se $A$ é considerada uma matriz sobre $\mathbb{C}$, então $A$ possui dois autovalores distintos, $\pm i$, enquanto que sobre $\mathbb{R}$ $A$ não possui autovalores. Apesar disso, $A$ é diagonalizável sobre $\mathbb{C}$, possuindo 4 autovetores distintos:

$$\begin{bmatrix} 0 \\ 0 \\ -i \\ 1 \end{bmatrix}, \begin{bmatrix} -i \\ 1 \\ 0 \\ 0 \end{bmatrix} \text{ associados ao autovalor } -i,$$

e

$$\begin{bmatrix} 0 \\ 0 \\ i \\ 1 \end{bmatrix}, \begin{bmatrix} i \\ 1 \\ 0 \\ 0 \end{bmatrix} \text{ associados ao autovalor } i$$

□

**Exemplo 3.** A matriz

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

possui o polinômio característico

$$\det(xI - A) = \det \begin{bmatrix} x & -1 \\ 0 & x \end{bmatrix} = x^2.$$

Tanto faz se $A$ for considerada uma matriz sobre $\mathbb{R}$ ou sobre $\mathbb{C}$, $A$ possui apenas um autovalor e o autoespaço associado a este autovalor tem dimensão 1, com o vetor $(1,0)$ sendo um autovetor associado ao autovalor 0. Portanto, $A$ não é diagonalizável. □

**Lema 1.** *Se $Tv = \lambda v$ e $p$ é um polinômio qualquer, então $p(T)v = p(\lambda)v$.*

**Lema 2.** *Seja $V$ um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$ e $T : V \longrightarrow V$ um um operador linear. Sejam $\lambda_1, \ldots, \lambda_k$ autovalores de $T$ e $W_i$ o autoespaço associado a $\lambda_i$. Se $W = W_1 + \ldots + W_k$, então*

$$W = W_1 \oplus \ldots \oplus W_k.$$

*Equivalentemente, se $\mathcal{B}_i$ é uma base para $W_i$, então $\mathcal{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_k\}$ é uma base para $W$ e*

$$\dim W = \dim W_1 + \ldots + \dim W_k.$$

**Prova:** Para provar que os autoespaços de $T$ são linearmente independentes, precisamos mostrar que dados $x_i \in W_i$ tais que $x_1 + \ldots + x_k = 0$, então $x_i = 0$ para todo $i$ (isso provará que se $\mathcal{B}_i$ é uma base para $W_i$, então $\mathcal{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_k\}$ é uma base para $W$, pois qualquer combinação linear de vetores de $\mathcal{B}$ se escreve desta forma). Se $p$ é um polinômio qualquer, então

$$0 = p(T)\,0 = p(T)(x_1 + \ldots + x_k) = p(T)\,x_1 + \ldots + p(T)\,x_k = p(\lambda_1)\,x_1 + \ldots + p(\lambda_k)\,x_k.$$

Escolha polinômios $p_1, \ldots, p_k$ tais que

$$p_i(\lambda_j) = \delta_{ij}.$$

Por exemplo,

$$p_i = \prod_{j \neq i} \frac{x - \lambda_j}{\lambda_i - \lambda_j}.$$

Então

$$0 = p_i(T)\,0 = \sum_{j=1}^{n} p_i(\lambda_j)\,x_j = \sum_{j=1}^{n} \delta_{ij} x_j = x_i.$$

■

**Teorema 2.** *Seja $V$ um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$ e $T : V \longrightarrow V$ um um operador linear. Sejam $\lambda_1, \ldots, \lambda_k$ os autovalores distintos de $T$ e $W_i$ o autoespaço associado a $\lambda_i$. As seguintes afirmações são equivalentes:*

(i) *$T$ é diagonalizável.*

(ii) *O polinômio característico de $T$ é*

$$f = (x - \lambda_1)^{d_1} \ldots (x - \lambda_k)^{d_k}$$

*e $\dim W_i = d_i$ para todo $i$.*

(iii) *Temos*

$$\dim V = \dim W_1 + \ldots + \dim W_k.$$

**Prova:** (i) ⇒ (ii) Se $T$ é diagonalizável, então $T$ possui uma representação matricial em blocos na forma

$$[T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 I_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 I_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_k I_k \end{bmatrix}$$

em relação a alguma base $\mathcal{B}$ de $V$, onde cada bloco identidade $I_i$ tem tamanho $d_i$. Segue imediatamente que o polinômio característico de $T$ é

$$\det(xI - T) = (x - \lambda_1)^{d_1} \ldots (x - \lambda_k)^{d_k}.$$

Como a matriz $[T - \lambda_i I]_{\mathcal{B}}$ tem exatamente $d_i$ zeros em sua diagonal principal, segue que $\dim W_i = d_i$.
(ii) ⇒ (iii) Se $f = (x - \lambda_1)^{d_1} \ldots (x - \lambda_k)^{d_k}$, então

$$\dim V = \text{ grau } f = d_1 + \ldots + d_k.$$

(iii) ⇒ (i) Segue do lema que $V = W_1 \oplus \ldots \oplus W_k$, logo $V$ possui uma base formada por autovetores de $T$. ■

**Exemplo 4.** A matriz

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

possui o polinômio característico $\det(xI - A) = (x-2)^2 (x-1)$, com autoespaços $W_1 = \langle (3, -1, 3) \rangle$ e $W_2 = \langle (2,1,0), (2,0,1) \rangle$. Portanto, $A$ é diagonalizável. □

**Exemplo 5.** Determine se a matriz

$$B = \begin{bmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{bmatrix}$$

é diagonalizável sobre $\mathbb{R}$ e sobre $\mathbb{C}$. Encontre bases para os autoespaços de $B$. $\square$

## 4.4 Ideais de Polinômios

**Definição 06.** Seja $\mathbb{F}$ um corpo. Um **ideal** em $\mathbb{F}[x]$ é um subespaço $M$ de $\mathbb{F}[x]$ tal que $fg \in M$ sempre que $f \in \mathbb{F}[x]$ e $g \in M$.

Em outras palavras, um ideal de polinômios é um subespaço de $\mathbb{F}[x]$ que é fechado com relação à multplicação, não só de seus próprios elementos, mas também por outros elementos de $\mathbb{F}[x]$. Ele deixa de ser uma subálgebra de $\mathbb{F}[x]$ porque não precisa conter a identidade.

**Exemplo 01.** Dado $d \in \mathbb{F}[x]$, o conjunto $M = d\mathbb{F}[x]$ de todos os múltiplos polinomiais de $d$ é um ideal. De fato, $M$ é não-vazio pois contém $d$ e se $f, g \in \mathbb{F}[x]$ e $\alpha, \beta \in \mathbb{F}$, então

$$\begin{aligned} \alpha(df) + \beta(dg) &= d(\alpha f + \beta g), \\ f(dg) &= d(fg). \end{aligned}$$

$M$ é chamado o **ideal principal gerado por** $d$.

**Teorema 03.** *Se $M$ é um ideal de $\mathbb{F}[x]$, então existe um único polinômio mônico $d$ em $\mathbb{F}[x]$ tal que $M$ é o ideal principal gerado por $d$.*

**Prova:** Por hipótese, $M$ contém um polinômio não-nulo. Entre todos os polinômios não-nulos de $M$ existe um polinômio $d$ de grau mínimo, que podemos assumir mônico, multiplicando $d$ pelo (polinômio) escalar apropriado, se necessário. Agora, se $f \in M$, então

$$f = dq + r$$

com $r = 0$ ou grau $r <$ grau $d$. Como $d \in M$, $dq \in M$ e $f - dq = r \in M$ também. Por escolha de $d$ (não existe polinômio não-nulo em $M$ de grau menor que $d$), concluímos que $r = 0$. Portanto $M = d\mathbb{F}[x]$.

Se $d'$ fosse outro polinômio mônico em $\mathbb{F}[x]$ tal que $M = d'\mathbb{F}[x]$, então existem polinômios não-nulos $f, g$ tais que $d = d'f$ e $d' = dg$. Logo, $d = dgf$ e daí

$$\text{grau } d = \text{ grau } d + \text{ grau } f + \text{ grau } g,$$

donde grau $f =$ grau $g = 0$. Como $f, g$ são mônicos, segue que $f = g = 1$ e portanto $d = d'$. $\blacksquare$

## 4.5 Polinômio Mínimo e o Teorema de Cayley-Hamilton

Seja $V$ um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$ e $T : V \longrightarrow V$ um um operador linear. Considere as primeiras $n^2 + 1$ potências de $T$:

$$I, T, T^2, \ldots, T^{n^2}.$$

Como o espaço $\mathcal{L}(V)$ dos operadores lineares sobre $V$ tem dimensão $n^2$, estes vetores são linearmente dependentes, isto é, existem escalares $a_0, a_1, \ldots, a_{n^2}$ tais que

$$a_0 I + a_1 T + \ldots + a_{n^2} T^{n^2} = 0.$$

Em outras palavras, o polinômio $p = a_0 + a_1 x + \ldots + a_{n^2} x^{n^2}$ *anula* o operador $T$, isto é,

$$p(T) = 0.$$

**Proposição 5.** *O conjunto dos polinômios que anulam um operador é um ideal em* $\mathbb{F}[x]$.

**Prova:** Pois, se $f, g$ anulam $T$ e $\alpha, \beta \in \mathbb{F}$, então

$$(\alpha f + \beta g)(T) = \alpha f(T) + \beta g(T) = 0$$

e se $f \in \mathbb{F}[x]$ e $g$ anula $T$,

$$(fg)(T) = f(T)\, g(T) = 0.$$

■

**Definição 5.** Seja $V$ um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$ e $T : V \longrightarrow V$ um um operador linear. O **polinômio mínimo** para $T$ é o gerador mônico do ideal dos polinômios anuladores de $T$.

Assim, se um polinômio anula $T$, ele é um múltiplo polinomial do polinômio mínimo.

**Teorema 3.** *Os polinômios mínimo e característico de um operador linear possuem as mesmas raízes, exceto por multiplicidades.*

**Prova:** Seja $p$ o polinômio mínimo de um operador linear $T$. Para provar o teorema, basta provar que $p(\lambda) = 0$ se e somente se $\lambda$ é um autovalor de $T$.

Suponha em primeiro lugar que $p(\lambda) = 0$ para algum $\lambda \in \mathbb{R}$. Então

$$p = (x - \lambda)\, q$$

e daí

$$0 = p(T) = (x - \lambda)(T)\, q(T) = (T - \lambda I)\, q(T)$$

Como grau $q <$ grau $p$, pela definição de polinômio mínimo não podemos ter $q(T) = 0$. Seja $v$ um vetor tal que $q(T)\, v = w \neq 0$. Então,

$$0 = (T - \lambda I)\, q(T)\, v = (T - \lambda I)\, w$$

e portanto $\lambda$ é um autovalor de $T$.

Reciprocamente, seja $\lambda$ é um autovalor de $T$. Então existe um vetor $v \neq 0$ tal que $Tv = \lambda v$. Pelo Lema 1,

$$0 = p(T)\, v = p(\lambda)\, v,$$

donde $p(\lambda) = 0$. ■

**Corolário 1.** *Se* $T : V \longrightarrow V$ *é um operador diagonalizável e* $\lambda_1, \ldots, \lambda_k$ *são os seus autovalores distintos, então seu polinômio mínimo é o polinômio*

$$p = (x - \lambda_1) \ldots (x - \lambda_k)$$

**Prova:** Por definição de operador diagonalizável (existe uma base para $V$ consistindo apenas de autovetores de $T$), qualquer vetor $v$ escreve-se na forma $v = v_1 + \ldots + v_k$ com $v_j$ um autovetor associado a $\lambda_j$. Como

$$(T - \lambda_1 I) \ldots (T - \lambda_k I)\, v_j = 0$$

para todo $i$, porque $(T - \lambda_j I)\, v_j = 0$ e polinômios em $T$ comutam, segue que

$$p(T)\, v = p(T) \sum_{j=1}^{k} v_j = \sum_{j=1}^{k} p(T)\, v_j = \sum_{j=1}^{k} (T - \lambda_1 I) \ldots (T - \lambda_k I)\, v_j = 0 \quad \text{para todo } v \in V.$$

■

Veremos na próxima seção que o polinômio mínimo de um operador linear ser um produto de fatores distintos é de fato equivalente a ele ser diagonalizável.

**Exemplo 6.** Encontre o polinômio mínimo para a matriz

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

Vimos no Exemplo 4 que $A$ possui o polinômio $(x-2)^2(x-1)$ como polinômio característico e que $A$ é diagonalizável. Logo seu polinômio mínimo é $p = (x-2)(x-1) = x^2 - 3x + 2$. Em particular, $A^2 - 3A + 2I = 0$. $\square$

**Exemplo 7.** Encontre o polinômio mínimo sobre $\mathbb{R}$ para a matriz

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

Vimos no Exemplo 1 que $B$ possui o polinômio $x^2 + 1$ como polinômio característico e que $B$ não é diagonalizável sobre $\mathbb{R}$, pois seu polinômio característico não possui raízes reais. No entanto, sobre $\mathbb{C}$, o polinômio característico se fatora $x^2 + 1 = (x+i)(x-i)$ e $B$ possui dois autovalores distintos, e portanto é diagonalizável. Assim, o polinômio mínimo sobre $\mathbb{C}$ para esta matriz é $x^2 + 1$. Como este polinômio possui coeficientes reais, ele também é o polinômio mínimo para $B$ sobre $\mathbb{R}$. $\square$

**Teorema 4.** (Teorema de Cayley-Hamilton) *O polinômio característico de um operador linear anula este operador.*

*Em particular, segue que o polinômio característico de um operador é divisível pelo polinômio mínimo deste operador.*

**Prova:** Seja $T : V \longrightarrow V$ um operador linear. A demonstração será por indução sobre $n = \dim V$. Se $n = 1$, o resultado é óbvio. Suponha o resultado válido para qualquer operador linear definido sobre qualquer espaço vetorial de dimensão $n-1$. Seja $T : V \longrightarrow V$ um operador linear definido sobre um espaço vetorial de dimensão $n$.

Seja $T : V \longrightarrow V$ um operador linear e $\mathcal{B} = \{x_1, \ldots, x_n\}$ uma base para $V$. Seja $A = [T]_{\mathcal{B}}$, ou seja,

$$Tx_j = \sum_{i=1}^{n} a_{ij} x_i \quad \text{para } j = 1, \ldots, n.$$

Estas equações podem ser reescritas na forma equivalente

$$\sum_{i=1}^{n} (\delta_{ij} T - a_{ij} I) x_i = 0 \quad \text{para } j = 1, \ldots, n.$$

Considere matrizes sobre a álgebra comutativa com identidade $\mathcal{A}$ dos polinômios em $T$. Em outras palavras, matrizes sobre $\mathcal{A}$ possuem polinômios em $T$ em suas entradas. Considere em particular a matriz $B$ definida por

$$b_{ij} = \delta_{ij} T - a_{ji} I.$$

Afirmamos que

$$\det B = f(T)$$

onde $f$ é o polinômio característico de $T$. Por exemplo, quando $n = 2$, temos

$$B = \begin{bmatrix} T - a_{11} I & -a_{21} I \\ -a_{12} I & T - a_{22} I \end{bmatrix}$$

e

$$\det B = (T - a_{11}I)(T - a_{22}I) - a_{12}a_{21}I$$
$$= T^2 - (a_{11} + a_{22})T + (a_{11}a_{22} - a_{12}a_{21})I$$
$$= f(T)$$

onde $f = x^2 - (a_{11} + a_{22})x + (a_{11}a_{22} - a_{12}a_{21})$ é o polinômio característico de $T$. No caso geral isso também é claro, porque o polinômio característico de $T$ é o determinante da matriz $xI - A$, cujas entradas são polinômios da forma

$$(xI - A)_{ij} = \delta_{ij}x - a_{ij}I,$$

e o determinante não se altera se considerarmos a transposta desta matriz.

Logo, para mostrar que $f(T) = 0$, basta mostrar que

$$(\det B)\, x_k = 0 \quad \text{para } k = 1, \ldots, n.$$

Por definição de $B$, os vetores $x_1, \ldots, x_n$ satisfazem as equações

$$\sum_{i=1}^{n} b_{ji}x_i = 0 \quad \text{para } j = 1, \ldots, n.$$

Seja $\widetilde{B} = \operatorname{adj} B$ a adjunta clássica de $B$, isto é, $\widetilde{B}B = (\det B)\, I$. Da equação acima, para cada par $k, j$ temos

$$\sum_{i=1}^{n} \widetilde{b}_{kj}b_{ji}x_i = \widetilde{b}_{kj}\sum_{i=1}^{n} b_{ji}x_i = 0.$$

Logo, somando em $j$, temos

$$\sum_{j=1}^{n}\sum_{i=1}^{n} \widetilde{b}_{kj}b_{ji}x_i = 0.$$

Daí,

$$0 = \sum_{j=1}^{n}\sum_{i=1}^{n} \widetilde{b}_{kj}b_{ji}x_i = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} \widetilde{b}_{kj}b_{ji}\right)x_i = \sum_{i=1}^{n}\left(\widetilde{B}B\right)_{ki}x_i = \sum_{i=1}^{n}\delta_{ki}(\det B)\, x_i = (\det B)\, x_k.$$

Mais tarde daremos uma demonstração independente deste teorema independente de determinantes. ■

## 4.6 Subespaços Invariantes e Operadores Triangularizáveis

**Definição 6.** Seja $T : V \longrightarrow V$ um operador linear. Dizemos que um subespaço $W \subset V$ é um **subespaço invariante** sob $T$ se $T(W) \subset W$.

**Exemplo 8.** O subespaço nulo e o espaço todo são invariantes sob qualquer operador linear $T$. O núcleo de $T$ e a imagem de $T$ também são invariantes sob $T$. □

**Exemplo 9.** Considere o espaço vetorial $\mathbb{F}[x]$ e o operador linear derivada $D$. Então o subespaço dos polinômios de grau menor que ou igual a $n$, onde $n$ é um inteiro não-negativo qualquer, é invariante sob $D$. □

**Exemplo 10.** Qualquer autoespaço de $T$ é invariante sob $T$. □

**Exemplo 11.** O operador linear em $\mathbb{R}^2$ representado na base canônica pela matriz

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

não possui outros subespaços invariantes além dos triviais, isto é, além do subespaço nulo e de $\mathbb{R}^2$. $\square$

Quando um subespaço $W \subset V$ é invariante sob $T$, $T$ *induz* um operador linear sobre $W$, o operador restrição $T|_W : W \longrightarrow W$. Denotaremos o operador restrição por $T_W$. Seja $\mathcal{B}_W = \{x_1, \ldots, x_m\}$ uma base para $W$ e $\mathcal{B} = \{x_1, \ldots, x_m, x_{m+1}, \ldots, x_n\}$ um completamento desta base até uma base para $V$. Seja $A = [T]_{\mathcal{B}}$ a representação matricial do operador $T$ com relação a $\mathcal{B}$, de modo que

$$Tx_j = \sum_{i=1}^{n} a_{ij} x_i \quad \text{para } j = 1, \ldots, n.$$

Como $W$ é invariante sob $T$, temos

$$Tx_j = \sum_{i=1}^{m} a_{ij} x_i \quad \text{para } j = 1, \ldots, m,$$

isto é, $a_{ij} = 0$ para $i > m$, se $j \leqslant m$. Em outras palavras, a matriz para $T$ tem a forma em blocos

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix},$$

onde $B$ é uma matriz $m \times m$, $C$ é uma matriz $m \times (n - m)$ e $D$ é uma matriz $(n - m) \times (n - m)$. Além disso, o bloco $B$ é a representação matricial do operador restrição $T_W$ com relação à base $\mathcal{B}_W$ de $W$, isto é,

$$B = [T_W]_{\mathcal{B}_W}.$$

**Proposição 6.** *Seja $W$ um espaço invariante de $V$ sob $T$. Então o polinômio característico para o operador restrição $T_W$ divide o polinômio característico para $T$ e o polinômio mínimo para o operador restrição $T_W$ divide o polinômio mínimo para $T$.*

**Prova:** Escrevendo como na discussão acima

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix},$$

onde $A = [T]_{\mathcal{B}}$ e $B = [T_W]_{\mathcal{B}_W}$, segue imediatamente que

$$\det (xI_n - A) = \det (xI_m - B) \det (xI_{n-m} - D)$$

o que prova a afirmação sobre polinômios característicos. Para provar a afirmação sobre polinômios mínimos, note que

$$A^k = \begin{bmatrix} B^k & C_k \\ 0 & D^k \end{bmatrix},$$

para todo $k$, onde $C_k$ é alguma matriz $m \times (n - m)$, donde, se $p$ é um polinômio qualquer, temos

$$p(A) = \begin{bmatrix} p(B) & \widetilde{C}_k \\ 0 & p(D) \end{bmatrix}$$

para alguma matriz $\widetilde{C}_k$, $m \times (n - m)$. Portanto, qualquer polinômio que anula $A$ também anula $B$ (e $D$). $\blacksquare$

**Definição 7.** Seja $W$ um espaço invariante de $V$ sob $T$ e $v \in V$. O conjunto dos polinômios

$$S(v; W) = \{f \in \mathbb{F}[x] : f(T)\, v \in W\}$$

é chamado o **$T$-condutor de $v$ para $W$**.

**Proposição 7.** *Seja $W$ um espaço invariante de $V$ sob $T$. Então $W$ é invariante sob qualquer polinômio em $T$.*

*Em particular, para qualquer $v \in V$, o $T$-condutor de $v$ para $W$ é um ideal de polinômios.*

**Prova:** Se $w \in W$, então $Tw \in W$ e por indução $T^k w \in W$ para todo $k$; tomando combinações lineares (porque $W$ é um subespaço vetorial), concluímos que $f(T)\, w \in W$ para todo polinômio $f$.

$S(v; W)$ é um subespaço vetorial, pois se $f, g \in S(v; W)$ então $(\alpha f + \beta g)(T)\, v = \alpha [f(T)\, v] + \beta [g(T)\, v] \in W$. Se $f \in \mathbb{F}[x]$ e $g \in S(v; W)$, então

$$[(fg)(T)]\, v = [f(T)\, g(T)]\, v = f(T)\, [g(T)\, v] \in W$$

porque $g(T)\, v \in W$ e $W$ é invariante sob qualquer polinômio em $T$. ■
O único gerador mônico do ideal $S(v; W)$ também é chamado o **$T$-condutor de $v$ para $W$**.

**Corolário 2.** *Todo $T$-condutor divide o polinômio mínimo para $T$.*

**Prova:** Pois todo $T$-condutor contém o polinômio mínimo por definição. ■

**Lema 3.** *Sejam $V$ um espaço vetorial de dimensão finita e $T : V \longrightarrow V$ um operador linear tal que o polinômio mínimo para $T$ é um produto de fatores lineares*

$$p = (x - \lambda_1)^{r_1} \dots (x - \lambda_k)^{r_k}.$$

*Seja $W$ um subespaço próprio de $V$ invariante sob $T$. Então existe um vetor $v \in V \backslash W$ tal que $(T - \lambda I)\, v \in W$ para algum autovalor $\lambda$ de $T$.*

*Em outras palavras, existe um vetor $v \in V \backslash W$ tal que o seu $T$-condutor para $W$ é um polinômio linear.*

**Prova:** Seja $z \in V \backslash W$ um vetor qualquer. Seja $f$ o $T$-condutor de $z$ para $W$. Então $f$ divide o polinômio mínimo de $T$. Como $z \notin W$, $f$ não é um polinômio escalar. Portanto,

$$f = (x - \lambda_1)^{s_1} \dots (x - \lambda_k)^{s_k}$$

onde $s_j < r_j$ e pelo menos algum $s_j \neq 0$. Escolha um índice $j$ tal que $s_j \neq 0$ e escreva

$$f = (x - \lambda_j)\, g.$$

Por definição de $f$, o vetor $v = g(T)\, z \notin W$, mas

$$(T - \lambda_j I)\, v = (T - \lambda_j I)\, g(T)\, z = f(T)\, z \in W.$$

■

**Definição 8.** Uma matriz $A = (a_{ij})$ é uma **matriz triangular** se $a_{ij} = 0$ sempre que $i < j$ (**triangular superior**) ou se $a_{ij} = 0$ sempre que $i > j$ (**triangular inferior**).

**Definição 9.** Seja $T : V \longrightarrow V$ um operador linear. Dizemos que $T$ é **triangularizável** se existe uma base de $V$ tal que a matriz de $T$ em relação a esta base é uma matriz triangular.

**Teorema 5.** *Seja $T : V \longrightarrow V$ um operador linear sobre um espaço vetorial de dimensão finita sobre $\mathbb{F}$. Então $T$ é triangularizável se e somente se o seu polinômio mínimo é um produto de fatores lineares sobre $\mathbb{F}$.*

**Prova:** Se $T$ é triangularizável, então existe uma base $\mathcal{B} = \{x_1, \ldots, x_n\}$ para $V$ tal que

$$[T]_{\mathcal{B}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ 0 & a_{22} & a_{23} & \ldots & a_{2n} \\ 0 & 0 & a_{33} & \ldots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & a_{nn} \end{bmatrix}.$$

Segue que o polinômio característico de $T$ é o produto de fatores lineares

$$\det (xI - T) = (x - a_{11}) \ldots (x - a_{nn}).$$

Como o polinômio mínimo divide o polinômio característico, o mesmo vale para ele.

Reciprocamente, se o polinômio mínimo para $T$ se escreve na forma

$$p = (x - \lambda_1)^{r_1} \ldots (x - \lambda_k)^{r_k},$$

aplicamos o Lema 3 repetidamente da seguinte forma para encontrar uma base $\mathcal{B} = \{x_1, \ldots, x_n\}$ para $V$ em relação à qual a matriz de $T$ é triangular. Aplique o lema ao subespaço invariante $W = \{0\}$ para obter o vetor $x_1$. Como $(T - \lambda_1 I) x_1 = 0$ para algum autovalor $\lambda_1$, segue que o subespaço $W_1 = \langle x_1 \rangle$ é invariante sob $T$. Podemos então aplicar novamente o lema ao subespaço $W_2$ para obter o vetor $x_2 \in V \backslash W_1$. Em particular, $\{x_1, x_2\}$ é L.I., e como $(T - \lambda_2 I) x_2 \in W_1$ para algum autovalor $\lambda_2$, segue que o subespaço $W_2 = \langle x_1, x_2 \rangle$ é invariante sob $T$; observe que $Tx_2 = a_{12} x_1 + \lambda_2 x_2$ para algum escalar $a_{12}$. Continuando desta forma, em cada passo $j$ encontramos vetores $x_1, \ldots, x_j$ linearmente independentes tais que $Tx_j = a_{1j} x_1 + \ldots + a_{j-1,j} x_{j-1} + \lambda_j x_j$ e o subespaço $W_j = \langle x_1, \ldots, x_j \rangle$ é portanto invariante sob $T$. ■

**Corolário 3.** *Todo operador complexo é triangularizável.*

Obteremos agora uma terceira caracterização para operadores diagonalizáveis em termos do polinômio mínimo:

**Teorema 6.** *Seja $T : V \longrightarrow V$ um operador linear sobre um espaço vetorial de dimensão finita sobre $\mathbb{F}$. Então $T$ é diagonalizável se e somente se o seu polinômio mínimo é um produto de fatores lineares distintos sobre $\mathbb{F}$.*

**Prova:** Já vimos no Corolário 1 que se $T$ é diagonalizável, então seu polinômio mínimo é um produto de fatores lineares distintos. Reciprocamente, suponha que o polinômio mínimo de um operador $T$ é o produto

$$p = (x - \lambda_1) \ldots (x - \lambda_k)$$

de fatores lineares distintos e suponha por absurdo que $T$ não é diagonalizável. Então, por definição, se $W$ é o subespaço gerado pelos autovetores de $T$, segue que $W \neq V$. Mas $W$ é um subespaço invariante sob $T$, logo pelo Lema 3 existe um vetor $v \in V \backslash W$ e um autovalor $\lambda_j$ tal que

$$w = (T - \lambda_j I) v \in W.$$

Como $w \in W$, existem vetores $w_i \in W_i$, onde $W_i$ é o autoespaço associado a $\lambda_i$ tal que

$$w = w_1 + \ldots + w_k.$$

Portanto, para qualquer polinômio $f$, temos

$$f(T) w = f(\lambda_1) w_1 + \ldots + f(\lambda_k) w_k \in W.$$

Agora, escrevendo $p = (x - \lambda_j) q$ para algum polinômio $q$ que não possui $\lambda_j$ como raiz e

$$q - q(\lambda_j) = (x - \lambda_j) g$$

para algum polinômio $g$ (pois $\lambda_j$ é uma raiz do polinômio $q - q(\lambda_j)$), temos

$$q(T) v - q(\lambda_j) v = g(T)(T - \lambda_j I) v = g(T) w \in W.$$

Como

$$0 = p(T) v = (T - \lambda_j I) q(T) v,$$

concluímos que $q(T) v$ é um autovetor de $T$ associado a $\lambda_j$ e, em particular, $q(T) v \in W$ também. Logo,

$$q(\lambda_j) v = q(T) v - g(T) w \in W$$

e como $v \notin W$, necessariamente $q(\lambda_j) = 0$, contradição. ■

Assim, para determinar se $T$ é diagonalizável, basta encontrar os autovalores distintos $\lambda_1, \ldots, \lambda_k$ de $T$ e determinar se o operador $(T - \lambda_1 I) \ldots (T - \lambda_k I)$ é o operador nulo ou não.

**Demonstração alternativa do Teorema de Cayley-Hamilton.** É fácil ver que o teorema de Cayley-Hamilton vale para matrizes triangulares. De fato, se $\mathcal{B} = \{x_1, \ldots, x_n\}$ é uma base para $V$ em relação à qual $T$ é representada por uma matriz triangular $A = (a_{ij})$, então o polinômio característico para $A$ é

$$f = (x - a_{11}) \ldots (x - a_{nn}).$$

Para provar que $f$ anula $T$, isto é, que

$$(T - a_{11} I) \ldots (T - a_{nn} I)$$

é operador nulo sobre $V$, podemos proceder por indução na dimensão de $V$. O resultado é claramente válido para $n = 1$. Assuma o resultado válido para $n - 1$ e escreva a matriz $A$ em blocos

$$A = \begin{bmatrix} B & C \\ 0 & a_{nn} \end{bmatrix}$$

onde $B$ é uma matriz $(n-1) \times (n-1)$ e $C$ é uma matriz coluna $(n-1) \times 1$. Observe que

$$(T - a_{nn} I) V \subset \langle x_1, \ldots, x_{n-1} \rangle =: W$$

$W$ é um subespaço invariante por $T$ de dimensão $n - 1$, com

$$[T_W]_{\{x_1, \ldots, x_{n-1}\}} = B,$$

cujo polinômio característico é exatamente

$$f = (x - a_{11}) \ldots (x - a_{n-1,n-1}).$$

Logo, por hipótese de indução,

$$(T_W - a_{11} I_{n-1}) \ldots (T_W - a_{n-1,n-1} I_{n-1})$$

é o operador nulo sobre $W$. Portanto a composta

$$(T - a_{11} I) \ldots (T - a_{n-1,n-1} I)(T - a_{nn} I) = (T_W - a_{11} I_{n-1}) \ldots (T_W - a_{n-1,n-1} I_{n-1})(T - a_{nn} I)$$

é o operador nulo sobre $V$ (observe que os vetores de $W$ têm as últimas coordenadas nulas, logo faz sentido aplicar o lado direito a vetores de $V$, apesar das matrizes identidade possuírem tamanho $n-1$). Do Teorema 5 segue que todo operador linear definido em um espaço vetorial sobre um corpo algebricamente fechado é triangularizável (lembre-se que dizemos que um corpo é algebricamente fechado se todo polinômio sobre este corpo possui uma raiz ou, equivalentemente, todo polinômio sobre o corpo se escreve como um produto de fatores lineares). [Observe que para provar esta afirmação do Teorema 5, que se o polinômio mínimo de um operador se fatora como um produto de fatores lineares então o operador é triangularizável, não usamos o teorema de Cayley-Hamilton. Ele foi usado no Teorema 5 para provar a recíproca desta afirmação, que não entra no presente argumento.] Qualquer corpo é um subcorpo de um corpo algebricamente fechado. ■

## 4.7 Exercícios

1. Seja $T$ o operador linear em $\mathbb{R}^4$ representado na base canônica pela matriz
$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \end{bmatrix}.$$
Sob que condições em $a$, $b$ e $c$ o operador $T$ é diagonalizável?

2. Sejam $A, B \in M_n(\mathbb{F})$. Prove que se $I - AB$ é invertível, então $I - BA$ também é invertível e
$$(I - BA)^{-1} = I + B(I - AB)^{-1} A.$$
Use este resultado para provar que se $A, B \in M_n(\mathbb{F})$ são duas matrizes quaisquer, então $AB$ e $BA$ possuem os mesmos autovalores. Será que $AB$ e $BA$ possuem o mesmo polinômio característico? Será que $AB$ e $BA$ possuem o mesmo polinômio mínimo?

3. Seja $A \in M_n(\mathbb{F})$ uma matriz diagonal com polinômio característico
$$f = (x - \lambda_1)^{d_1} \dots (x - \lambda_k)^{d_k}$$
onde $\lambda_1, \dots, \lambda_k$ são distintos. Mostre que subconjunto em $M_n(\mathbb{F})$ das matrizes $B$ tais que $AB = BA$ é um subespaço vetorial de dimensão
$$d_1^2 + \dots + d_k^2.$$

4. Seja $A \in M_n(\mathbb{F})$ e considere o operador linear $T : M_n(\mathbb{F}) \longrightarrow M_n(\mathbb{F})$ definido por $T(B) = AB$. Será verdade que $A$ e $T$ possuem os mesmos autovalores? Será que $A$ e $T$ possuem o mesmo polinômio característico? Será que $A$ e $T$ possuem o mesmo polinômio mínimo?

5. Seja $\mathbb{F}$ um corpo arbitrário e $a, b, c \in \mathbb{F}$. Considere a matriz
$$A = \begin{bmatrix} 0 & 0 & c \\ 1 & 0 & b \\ 0 & 1 & a \end{bmatrix}.$$
Mostre que o polinômio característico para $A$ é $p = x^3 - ax^2 - bx - c$ e que este também é o polinômio mínimo para $A$.

6. Seja $A$ a matriz real
$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ -2 & -2 & 2 & 1 \\ 1 & 1 & -1 & 0 \end{bmatrix}$$
Mostre que o seu polinômio característico é $p = x^2(x-1)^2$ e que este também é o seu polinômio mínimo. Se $A$ for considerada uma matriz complexa, $A$ é diagonalizável?

7. Seja $T$ o operador linear sobre $\mathbb{R}^2$ cuja matriz na base canônica é
$$A = \begin{bmatrix} 1 & -1 \\ 2 & 2 \end{bmatrix}.$$
Prove que os únicos subespaços de $\mathbb{R}^2$ invariantes sob $T$ são os triviais. Se $U$ é um operador sobre $\mathbb{C}^2$ cuja matriz na base canônica é $A$, mostre que $U$ possui um subespaço invariante unidimensional.

8. Seja
$$A = \begin{bmatrix} 0 & 1 & 0 \\ 2 & -2 & 2 \\ 2 & -3 & 2 \end{bmatrix}.$$
$A$ é semelhante sobre o corpo dos números reais a uma matriz triangular? Se for, encontre uma tal matriz triangular.

9. Prove que toda matriz $A$ tal que $A^2 = A$ é semelhante a uma matriz diagonal.

10. Seja $T : V \longrightarrow V$ um operador linear tal que todo subespaço de $V$ é invariante sob $T$. Mostre que $T$ é um múltiplo escalar do operador identidade.

11. Seja $A$ uma matriz real $3 \times 3$. Mostre que se $A$ não é semelhante sobre $\mathbb{R}$ a uma matriz triangular, então $A$ é semelhante sobre $\mathbb{C}$ a uma matriz diagonal.

12. A seguinte afirmação é falsa ou verdadeira? *Se uma matriz triangular é semelhante a uma matriz diagonal, então ela já é diagonal.*

13. Seja $T$ um operador linear sobre um espaço vetorial $V$ sobre $\mathbb{F}$ de dimensão finita e $f \in \mathbb{F}[x]$. Prove que $\Lambda$ é um autovalor de $f(T)$ se e somente se $\Lambda = f(\lambda)$, onde $\lambda$ é um autovalor de $T$ (Compare com o Lema 1).

## 4.8 Projeções e Decomposição em Soma Direta

Recordamos a seguinte definição:

**Definição 10.** Dizemos que os subespaços $W_1, \ldots, W_k$ de um espaço vetorial $V$ são **linearmente independentes** se
$$w_1 + \ldots + w_k = 0, \quad \text{com } w_i \in W_i \text{ para cada } i,$$
implicar que $w_i = 0$ para todo $i$. Neste caso, a soma $W = W_1 + \ldots + W_k$ é chamada uma **soma direta** e é denotada por
$$W = W_1 \oplus \ldots \oplus W_k.$$

**Lema 4.** *Sejam $V$ um espaço vetorial de dimensão finita e $W_1, \ldots, W_k$ subespaços de $V$. As seguintes afirmações são equivalentes:*

(i) *$W_1, \ldots, W_k$ são linearmente independentes.*

(ii) *Para cada $2 \leqslant j \leqslant k$ nós temos*
$$W_j \cap (W_1 + \ldots + W_{j-1}) = \{0\}.$$

(iii) *Se $\mathcal{B}_i$ é uma base para $W_i$ então $\mathcal{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_k\}$ é uma base para $W_1 + \ldots + W_k$.*

**Prova:** (i) $\Rightarrow$ (ii) Seja $w \in W_j \cap (W_1 + \ldots + W_{j-1})$. Então $w \in W_j$ e $w = w_1 + \ldots + w_{j-1}$ para alguns vetores $w_i \in W_i$. Como
$$w_1 + \ldots + w_{j-1} - w + 0 + \ldots + 0 = 0$$
concluímos que $w_1 = \ldots = w_{j-1} = w = 0$.
(ii) $\Rightarrow$ (i) Suponha que
$$w_1 + \ldots + w_k = 0$$
com $w_i \in W_i$ para cada $i$. Se existe algum $w_j$ não nulo, seja $j$ o maior inteiro tal que $w_j \neq 0$. Então $w_1 + \ldots + w_j = 0$ e
$$w_j = -w_1 - \ldots - w_{j-1}$$
contradizendo $W_j \cap (W_1 + \ldots + W_{j-1}) = \{0\}$.
(i) $\Leftrightarrow$ (iii) Óbvio. ■

**Definição 11.** Seja $V$ um espaço vetorial. Uma projeção de $V$ é um operador linear $E : V \longrightarrow V$ tal que $E^2 = E$.

**Proposição 8.** *Seja $E : V \longrightarrow V$ uma projeção. Então*

$$V = \operatorname{im} E \oplus \ker E,$$

*com*

$$x = Ex + (I - E)\, x$$

*sendo esta a única decomposição de um vetor $x \in V$ como uma soma de elementos em $\operatorname{im} E$ e $\ker E$. Em particular, $x \in \operatorname{im} E$ se e somente se*

$$Ex = x.$$

*Além disso, se $V = W \oplus Z$, existe uma única projeção $E : V \longrightarrow V$ tal que $W = \operatorname{im} E$ e $Z = \ker E$.*

**Prova:** Suponha que $x \in \operatorname{im} E \cap \ker E$. Então, por definição, $Ex = 0$ e $x = Ey$ para algum $y \in V$. Mas então $0 = Ex = E^2 y = Ey = x$. Segue do Lema 4 (ii) que $\ker E$ e $\operatorname{im} E$ são linearmente independentes.

Observe que $E\,[(I - E)\, x] = Ex - E^2 x = 0$, portanto $(I - E)\, x \in \ker E$. A unicidade da decomposição segue do fato de $\ker E$ e $\operatorname{im} E$ serem linearmente independentes.

Se $V = W \oplus Z$, então cada vetor $v \in V$ se escreve de maneira única na forma $v = w + z$ para alguns vetores $w \in W$ e $z \in Z$, logo podemos definir um operador linear $E : W \oplus Z \longrightarrow V$ por $E(w + z) = w$. Então $E$ é um operador linear bem definido e temos $E^2 v = E\,(Ev) = Ew = w = Ev$ para todo vetor $v \in V$.
■

Na notação da proposição anterior, se $W = \operatorname{im} E$ e $Z = \ker E$, dizemos que $E$ é a **projeção sobre $W$ ao longo de $Z$**.

**Proposição 9.** *Seja $E : V \longrightarrow V$ uma projeção sobre um espaço de dimensão finita. Então $E$ é diagonalizável.*

**Prova:** Se $\mathcal{B}' = \{x_1, \ldots, x_m\}$ é uma base para $\operatorname{im} E$ e $\mathcal{B}'' = \{x_{m+1}, \ldots, x_n\}$ é uma base para $\ker E$, então $\mathcal{B} = \{x_1, \ldots, x_n\}$ é uma base que diagonaliza $E$, pois

$$[E]_{\mathcal{B}} = \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix}.$$

■

O próximo resultado generaliza a Proposição 8 (veja os Exercícios 18 e 19).

**Teorema 7.** (Decomposição em Soma Direta) *Se $V = W_1 \oplus \ldots \oplus W_k$, então existem $k$ operadores lineares $E_1, \ldots, E_k : V \longrightarrow V$ tais que*

(i) *Cada $E_i$ é uma projeção.*

(ii) *$E_i E_j = 0$ se $i \neq j$.*

(iii) *$E_1 + \ldots + E_k = I$.*

(iv) *$\operatorname{im} E_i = W_i$.*

*Reciprocamente, se existem $k$ operadores lineares $E_1, \ldots, E_k : V \longrightarrow V$ que satisfazem as condições* (ii)-(iv)*, então vale também* (i) *e $V = W_1 \oplus \ldots \oplus W_k$.*

**Prova:** Suponha que $V = W_1 \oplus \ldots \oplus W_k$. Dado $v = v_1 + \ldots + v_k$ com $v_i \in W_i$, defina

$$E_j v = v_j.$$

Então $E_j$ é um operador linear bem definido e $E_j^2 = E_j$, isto é, $E_j$ é uma projeção. Além disso, para todo $j$ vale

$$\operatorname{im} E_j = W_j \quad \text{e} \quad \ker E_j = W_1 + \ldots + W_{j-1} + W_{j+1} + \ldots + W_k,$$

de modo que $E_i E_j = 0$ se $i \neq j$. Como

$$v = v_1 + \ldots + v_k = E_1 v + \ldots + E_k v,$$

temos $I = E_1 + \ldots + E_k$.

Reciprocamente, suponha que existem $k$ operadores lineares $E_1, \ldots, E_k : V \longrightarrow V$ que satisfazem as condições (ii)-(iv). Obtemos (ii) multiplicando (iii) por cada $E_i$. Como, por (iii),

$$v = E_1 v + \ldots + E_k v$$

temos que

$$V = W_1 + \ldots + W_k.$$

Além disso, esta expressão para $v$ é única. De fato, se $v = v_1 + \ldots + v_k$ com $v_i \in W_i$, temos, usando (i) e (ii),

$$E_j v = \sum_{i=1}^{k} E_j v_i = \sum_{i=1}^{k} E_j E_i v_i = E_j^2 v_j = E_j v_j = v_j.$$

■

As decomposições em soma direta $V = W_1 \oplus \ldots \oplus W_k$ mais úteis de um espaço vetorial $V$ são aquelas em que cada um dos subespaços $W_i$ são invariantes sob algum operador linear $T$. Neste caso, o operador $T$ induz um operador $T_i = T|_{W_i}$ sobre cada um dos subespaços invariantes. Se

$$v = v_1 + \ldots + v_k$$

é a expressão única de $v$ como a soma de vetores nos subespaços invariantes, temos

$$Tv = Tv_1 + \ldots + Tv_k = T_1 v_1 + \ldots + T_k v_k.$$

Dizemos que $T$ é a **soma direta** dos operadores $T_1, \ldots, T_k$. [Observe porém que a expressão $T_1 + \ldots + T_k$ propriamente dita não faz sentido, já que cada operador $T_i$ tem domínio diferente dos demais.] Se $\mathcal{B}_i$ é uma base para $W_i$, de modo que $\mathcal{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_k\}$ é uma base para $V$, temos

$$[T]_{\mathcal{B}} = \begin{bmatrix} [T_1]_{\mathcal{B}_1} & 0 & \ldots & 0 \\ 0 & [T_2]_{\mathcal{B}_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & [T_k]_{\mathcal{B}_k} \end{bmatrix}.$$

O objetivo é encontrar uma decomposição do espaço $V$ em soma direta de subespaço invariantes por $T$ de tal forma que os operadores $T_i$ tenham uma forma simples.

**Lema 5.** *Sejam $T : V \longrightarrow V$ um operador linear e $E : V \longrightarrow V$ uma projeção. O núcleo e a imagem de $E$ são invariantes sob $T$ se e somente se $ET = TE$, isto é, se e somente se $T$ comuta com $E$.*

**Prova:** Suponha que $T$ comuta com $E$. Dado $w \in \operatorname{im} E$, então $Ew = w$ e

$$Tw = T(Ew) = E(Tw),$$

de modo que $Tw \in \operatorname{im} E$ e portanto $\operatorname{im} E$ é invariante sob $T$. Dado $z \in \ker E$, temos $Ez = 0$ e

$$E(Tz) = T(Ez) = T0 = 0,$$

de modo que $Tz \in \ker E$ e portanto $\ker E$ é invariante sob $T$.

Reciprocamente, suponha que $W = \operatorname{im} E$ e $Z = \ker E$ são invariantes sob $T$. Seja $v \in V$ um vetor qualquer. Então,

$$v = Ev + (I - E)v,$$

onde $Ev \in \operatorname{im} E$ e $(I - E)v \in \ker E$. Temos

$$Tv = TEv + T(I - E)v.$$

Pela invariância de $\operatorname{im} E$ e $\ker E$ sob $T$, temos

$$\begin{aligned} TEv &= Ew, \\ T(I - E)v &= z \end{aligned}$$

para alguns vetores $w \in \operatorname{im} E$ e $z \in \ker E$. Logo,

$$ETv = E(TEv + T(I - E)v) = E(Ew + z) = E^2 w = Ew = TEv.$$

■

**Teorema 8.** (Decomposição em Soma Direta de Subespaços Invariantes) *Seja $T : V \longrightarrow V$ um operador linear sobre um espaço vetorial de dimensão finita e $V = W_1 \oplus \ldots \oplus W_k$ uma decomposição em soma direta de subespaços invariantes. Então existem $k$ operadores lineares $E_1, \ldots, E_k : V \longrightarrow V$ tais que*

(i) *Cada $E_i$ é uma projeção.*

(ii) $E_i E_j = 0$ *se* $i \neq j$.

(iii) $E_1 + \ldots + E_k = I$.

(iv) $\operatorname{im} E_i = W_i$.

(v) $TE_i = E_i T$ *para todo* $i$.

*Reciprocamente, se existem $k$ operadores lineares $E_1, \ldots, E_k : V \longrightarrow V$ que satisfazem as condições* (ii)-(v)*, então vale também* (i) *e $V = W_1 \oplus \ldots \oplus W_k$.*

**Prova:** Em vista do Teorema 7, basta provar que obtemos uma decomposição em soma direta de subespaços invariantes se e somente se $T$ comuta com cada $E_i$.

Suponha que $T$ comuta com $E_i$. Segue imediatamente do lema anterior que $\operatorname{im} E_i$ é invariante sob $T$.

Reciprocamente, suponha que cada $W_i$ é invariante sob $T$. Observando que

$$\begin{aligned} \operatorname{im} E_i &= W_i, \\ \ker E_i &= W_1 + \ldots + W_{i-1} + W_{i+1} + \ldots + W_k \end{aligned}$$

são invariantes sob $T$, segue do lema anterior que $T$ comuta com $E_i$. ■

No próximo resultado, descreveremos um operador diagonalizável na linguagem de decomposição em soma direta de subespaços invariantes, o que ajudará a entender outros teoremas mais profundos de decomposição em soma direta que veremos ao longo do curso.

**Teorema 9.** (Decomposição em Soma Direta de Subespaços Invariantes para Operadores Diagonalizáveis) *Seja $T : V \longrightarrow V$ um operador linear sobre um espaço de dimensão finita $V$.*

*Se $T$ é diagonalizável, $\lambda_1, \ldots, \lambda_k$ são os autovalores distintos de $T$ e $W_1, \ldots, W_k$ são os autoespaços respectivamente correspondentes a estes autovalores, então existem $k$ operadores lineares $E_1, \ldots, E_k : V \longrightarrow V$ tais que*

(i) *Cada $E_i$ é uma projeção.*

(ii) $E_i E_j = 0$ *se* $i \neq j$.

(iii) $E_1 + \ldots + E_k = I$.

(iv) $\operatorname{im} E_i = W_i$.

(v) $T = \lambda_1 E_1 + \ldots + \lambda_k E_k$.

*Reciprocamente, se existem $k$ operadores lineares não-nulos $E_1, \ldots, E_k : V \longrightarrow V$ e $k$ escalares distintos $\lambda_1, \ldots, \lambda_k$ que satisfazem as condições* (ii), (iii) *e* (v)*, então $T$ é diagonalizável, $\lambda_1, \ldots, \lambda_k$ são os autovalores distintos de $T$, $W_1, \ldots, W_k$ são os autoespaços respectivamente correspondentes a estes autovalores, e as condições* (i) *e* (iv) *também são satisfeitas.*

**Prova:** Suponha $T$ diagonalizável. Então já vimos que $V = W_1 \oplus \ldots \oplus W_k$. Então (i)-(iv) seguem do Teorema 7. Para verificar (v), observe que para cada vetor $v \in V$ temos

$$v = E_1 v + \ldots + E_k v,$$

logo

$$Tv = TE_1 v + \ldots + TE_k v = \lambda_1 E_1 v + \ldots + \lambda_k E_k v.$$

Reciprocamente, suponha que existam operadores lineares $E_1, \ldots, E_k : V \longrightarrow V$ e escalares distintos $\lambda_1, \ldots, \lambda_k$ que satisfazem as condições (ii), (iii) e (v). Como $E_i E_j = 0$ se $i \neq j$, se multiplicarmos $E_1 + \ldots + E_k = I$ por $E_i$ imediatamente obtemos $E_i^2 = E_i$, isto é, cada $E_i$ é uma projeção.

Multiplicando $T = \lambda_1 E_1 + \ldots + \lambda_k E_k$ por $E_i$, obtemos $TE_i = \lambda_i E_i$, o que mostra que $\operatorname{im} E_i = W_i \subset \ker (T - \lambda_i I)$. Como $E_i \neq 0$ por hipótese, isso prova que $W_i \neq \{0\}$, ou seja, $\lambda_i$ é um autovalor de $T$. E em particular, como os subespaços $W_i$ estão contidos em autoespaços associados a autovalores distintos, eles são linearmente independentes. Portanto, segue de (iii) que $V = W_1 \oplus \ldots \oplus W_k$. Concluímos que $T$ é diagonalizável, pois possui $V$ uma base de autovetores de $T$. Não existem outros autovalores de $T$ além de $\lambda_1, \ldots, \lambda_k$, pois se $\lambda$ é um autovalor de $T$, então

$$T - \lambda I = (\lambda_1 - \lambda) E_1 + \ldots + (\lambda_k - \lambda) E_k,$$

de modo que se $(T - \lambda I) v = 0$, devemos ter $(\lambda_i - \lambda) E_i v = 0$ para todo $i$. Se $v \neq 0$, então $E_j v \neq 0$ para algum $j$, logo $\lambda = \lambda_j$.

Só falta mostra que $W_i = \ker (T - \lambda_i I)$ para todo $i$. Já provamos que $W_i \subset \ker (T - \lambda_i I)$. Se $v \in \ker (T - \lambda_i I)$, isto é, se $Tv = \lambda_j v$, então

$$0 = (T - \lambda_j I) v = (\lambda_1 - \lambda_j) E_1 v + \ldots + (\lambda_k - \lambda_j) E_k v,$$

donde $(\lambda_i - \lambda_j) E_i v = 0$ para todo $i$, logo $E_i v = 0$ para todo $i \neq j$. Daí, segue que $v = E_1 v + \ldots + E_k v = E_j v \in W_j$. ■

## 4.9 Exercícios

14. Seja $T : V \longrightarrow V$ um operador linear e $E$ uma projeção. Prove que a imagem de $E$ é invariante sob $T$ se e somente se $ETE = TE$.

15. Seja $T : V \longrightarrow V$ um operador linear que comuta com toda projeção de $V$. O que você pode dizer sobre $T$?

16. Se $E$ é uma projeção e $f$ um polinômio, então $f(E) = aI + bE$. Encontre uma expressão para $a$ e $b$ em termos dos coeficientes de $f$.

17. Mostre que se $E$ é a projeção sobre $W$ ao longo de $Z$, então $I - E$ é a projeção sobre $Z$ ao longo de $W$.

18. Sejam $E_1, \ldots, E_k : V \longrightarrow V$ operadores lineares tais que $E_1 + \ldots + E_k = I$.

    1. Prove que se $E_i E_j = 0$ sempre que $i \neq j$, então cada $E_i$ é uma projeção.

    2. Prove a recíproca no caso $k = 2$, isto é, se $E_1, E_2$ são projeções tais que $E_1 + E_2 = I$, então $E_1 E_2 = 0$.

    3. Prove a recíproca no caso geral, isto é, se $E_1, \ldots, E_k$ são projeções tais que $E_1 + \ldots + E_k = I$, então $E_i E_j = 0$ sempre que $i \neq j$, assumindo que $V$ é um espaço vetorial sobre um subcorpo dos complexos. (Sugestão para este último: use a função traço; qual é o traço de uma projeção?)

## 4.10 Fatoração de Polinômios

No Teorema 03, provamos que se $M$ é um ideal de $\mathbb{F}[x]$, então existe um único polinômio mônico $d$ em $\mathbb{F}[x]$ tal que $M$ é o ideal principal gerado por $d$, isto é, $M = d\mathbb{F}[x]$. Como conseqüência, vale o seguinte resultado:

**Corolário 08.** *Se $p_1, \ldots, p_k$ são polinômios sobre $\mathbb{F}$, não todos nulos, então existe um único polinômio mônico $d \in \mathbb{F}[x]$ tal que*

(a) *$d$ está no ideal gerado por $p_1, \ldots, p_k$, isto é, $d \in p_1\mathbb{F}[x] + \ldots + p_k\mathbb{F}[x]$;*

(b) *$d$ divide cada um dos polinômios $p_1, \ldots, p_k$.*

*Além disso, qualquer polinômio $d$ que satisfaz* (a) *e* (b) *satisfaz*

(c) *$d$ é um múltiplo polinomial de qualquer polinômio que divide os polinômios $p_1, \ldots, p_k$.*

**Prova:** Seja $d$ o gerador mônico do ideal $p_1\mathbb{F}[x] + \ldots + p_k\mathbb{F}[x]$. Como cada membro do ideal é divisível por $d$, em particular cada $p_i$ é divisível por $d$, o que prova (a) e (b).

Suponha que $f$ é um polinômio que divide $p_1, \ldots, p_k$. Então existem polinômios $g_1, \ldots, g_k$ tais que $p_i = fg_i$ para cada $i$. Também, como $d \in p_1\mathbb{F}[x] + \ldots + p_k\mathbb{F}[x]$, existem polinômios $q_1, \ldots, q_k$ tais que $d = p_1q_1 + \ldots + p_kq_k$. Logo

$$d = f(g_1q_1 + \ldots + g_kq_k).$$

Se $d'$ é outro polinômio que satisfaz (a) e (b), segue da definição de $d$ que $d' = fd$ para algum polinômio $f$, logo satisfaz (c). Se $d'$ é mônico, segue que $d' = d$. ■

**Definição 06.** Sejam $p_1, \ldots, p_k$ polinômios sobre um corpo $\mathbb{F}$, não todos nulos. O gerador mônico $d$ do ideal $p_1\mathbb{F}[x] + \ldots + p_k\mathbb{F}[x]$ é chamado o **máximo divisor comum** (mdc) de $p_1, \ldots, p_k$.

Dizemos que os polinômios $p_1, \ldots, p_k$ são **relativamente primos** se mdc $(p_1, \ldots, p_k) = 1$.

Observe que dizer que $p_1, \ldots, p_k$ são relativamente primos é equivalente a dizer que o ideal gerado por eles é todo o anel $\mathbb{F}[x]$. Em vista do Corolário 08 (c), quando os polinômios $p_1, \ldots, p_k$ são relativamente primos, eles não são simultaneamente divisíveis por nenhum outro polinômio diferente de 1.

**Definição 07.** Dizemos que um polinômio $f \in \mathbb{F}[x]$ é **redutível** sobre $\mathbb{F}$ se existem polinômios $g, h \in \mathbb{F}[x]$ de grau maior ou igual a 1 tais que $f = gh$. Caso contrário, dizemos que $f$ é **irredutível** sobre $\mathbb{F}$. Um polinômio irredutível não-escalar também é chamado um polinômio **primo** sobre $\mathbb{F}$.

Em outras palavras, dizer que $p$ é primo equivale a dizer que os únicos divisores de $p$ são $p$ e 1.

**Proposição 03.** *Suponha que $p$ é um polinômio primo que divide o produto $fg$. Então $p$ divide $f$ ou $p$ divide $g$.*

**Prova:** Sem perda de generalidade podemos assumir que $p$ é mônico. Seja $d = \text{mdc}\,(p, f)$. Como $p$ é primo, segue que $d = 1$ ou $d = p$. Se $d = p$, então $p$ divide $f$. Caso contrário, mostraremos que $d$ divide $g$. Como $\text{mdc}\,(p, f) = 1$, existem polinômios $h_1, h_2$ tais que

$$1 = ph_1 + fh_2.$$

Logo, multiplicando esta equação por $g$, obtemos

$$g = p\,(gh_1) + (fg)\,h_2,$$

Denote $k_1 = gh_1$. Como $p$ divide $fg$, temos $fg = ph_3$ para algum polinômio $h_3$. Portanto, se $k_2 = h_3h_2$, temos

$$g = pk_1 + pk_2 = p\,(k_1 + k_2)\,.$$

■

**Corolário 09.** *Se $p$ é um polinômio primo que divide o produto $f_1 \ldots f_k$, então $p$ divide algum dos polinômios $f_1, \ldots, f_k$.*

**Teorema 04.** *Se $\mathbb{F}$ é um corpo, então todo polinômio mônico não-escalar sobre $\mathbb{F}$ pode ser fatorado como um produto de polinômios primos mônicos sobre $\mathbb{F}$ de uma única maneira (a menos da ordem dos fatores).*

**Proposição 04.** *Seja $f$ um polinômio mônico não-escalar sobre $\mathbb{F}$ tal que*

$$f = p_1^{r_1} \ldots p_k^{r_k}$$

*é a fatoração prima de $f$. Para cada $i$, defina*

$$f_i = \frac{p}{p_i^{r_i}} = \prod_{\substack{j=1 \\ j \neq i}}^{k} p_j^{r_j}.$$

*Então $f_1, \ldots, f_k$ são relativamente primos.*

## 4.11 Teorema da Decomposição Primária e Teorema Espectral

Para operadores em geral, mesmo aqueles que não possuem nenhum autovalor, vale o seguinte resultado fundamental:

**Teorema 10.** (Teorema da Decomposição Primária) *Seja $T : V \longrightarrow V$ um operador linear sobre um espaço de dimensão finita $V$ sobre o corpo $\mathbb{F}$. Seja*

$$p = p_1^{r_1} \ldots p_k^{r_k}$$

*o polinômio mínimo de $T$, expresso como um produto de fatores primos distintos sobre $\mathbb{F}$. Seja $W_i = \ker p_i^{r_i}\,(T)$. Então*

(i) $V = W_1 \oplus \ldots \oplus W_k$.

(ii) *Cada $W_i$ é invariante sob $T$.*

(iii) *Se* $T_i = T|_{W_i}$*, então o polinômio mínimo de* $T_i$ *é* $p_i^{r_i}$*.*

**Prova:** A idéia da demonstração é usar os Teorema 8, encontrando as projeções associadas $E_i$, isto é $E_i$ é a identidade em $W_i$ e nula nos outros $W_j$, e que além disso comutam com $T$. Para isso, lembrando que operadores que podem ser escritos como polinômios em $T$ sempre comutam com $T$, encontraremos polinômios $h_i$ tais que $h_i(T)$ é a identidade em $W_i$ e nulo nos outros $W_j$.

Para cada $i$, defina

$$f_i = \frac{p}{p_i^{r_i}} = \prod_{\substack{j=1 \\ j \neq i}}^{k} p_j^{r_j}.$$

Como $p_1^{r_1}, \ldots, p_k^{r_k}$ são fatores primos distintos, os polinômios $f_1, \ldots, f_k$ são relativamente primos. Logo, existem polinômios $g_1, \ldots, g_k$ sobre $\mathbb{F}$ tais que

$$\sum_{i=1}^{k} f_i g_i = 1.$$

Defina

$$h_i = f_i g_i.$$

e

$$E_i = h_i(T).$$

Como $h_1 + \ldots + h_k = 1$, segue que $E_1 + \ldots + E_k = I$. Notando que se $i \neq j$ então o polinômio $f_i f_j$ é um múltiplo polinomial do polinômio mínimo $p$ (porque este produto contém todos os fatores de $p$), segue que

$$E_i E_j = [f_i g_i(T)][f_j g_j(T)] = f_i(T) f_j(T) g_i(T) g_j(T) = (f_i f_j)(T) g_i(T) g_j(T) = 0 g_i(T) g_j(T) = 0$$

se $i \neq j$. Segue dos Teoremas 7 e 8 que os operadores $E_i$ são projeções que correspondem a alguma soma direta do espaço $V$. Para provar (i) e (ii), basta mostrar que $\operatorname{im} E_i = W_i$.

Para provar isso, observe primeiro que $\operatorname{im} E_i \subset W_i$, pois se $v = E_i v$ então

$$p_i^{r_i}(T) v = p_i^{r_i}(T) E_i v = p_i^{r_i}(T) f_i(T) g_i(T) v = p(T) g_i(T) v = 0.$$

Reciprocamente, $W_i \subset \operatorname{im} E_i$. De fato, seja $v \in \ker p_i^{r_i}(T)$. Se $j \neq i$, então $f_j g_j$ é um múltiplo polinomial de $p_i^{r_i}$, logo $E_j v = 0$. De $E_1 + \ldots + E_k = I$ segue imediatamente que $E_i v = v$, logo $v \in \operatorname{im} E_i$.

Para terminar a demonstração do teorema, observe que $p_i^{r_i}(T_i) = 0$, já que por definição $W_i = \ker p_i^{r_i}(T)$, logo o polinômio mínimo de $T_i$ divide $p_i^{r_i}$. Reciprocamente, se $g = p_i^{s_i}$, $s_i \leqslant r_i$, é tal que $g(T_i) = 0$, então $g(T) f_i(T) = 0$. Em particular, $g f_i$ é divisível pelo polinômio mínimo de $T$, isto é, $p = p_i^{r_i} f_i$ divide $g f_i$, donde $p_i^{r_i}$ divide $g$ e portanto $s_i = r_i$. ■

A decomposição dada pelo Teorema 10 é chamada a **decomposição primária** de $T$.

**Corolário 4.** *Se* $E_1, \ldots, E_k$ *são as projeções associadas com a decomposição primária de* $T$*, então cada* $E_i$ *é um polinômio em* $T$*.*

*Conseqüentemente, se um operador linear* $S$ *comuta com* $T$*, então* $S$ *comuta com cada* $E_i$*, isto é, cada subespaço* $W_i$ *é invariante sob* $S$*. Em particular,* $T$ *e* $S$ *possuem a mesma decomposição em soma direta por subespaços invariantes.*

No caso em que o polinômio mínimo é um produto de fatores lineares, o Teorema da Decomposição Primária é às vezes chamado de *Teorema Espectral.*

**Corolário 5.** (Teorema Espectral) *Seja* $T : V \longrightarrow V$ *um operador linear sobre um espaço de dimensão finita* $V$ *sobre o corpo* $\mathbb{F}$*, tal que o polinômio mínimo para* $T$ *é um produto de fatores lineares*

$$p = (x - \lambda_1)^{r_1} \ldots (x - \lambda_k)^{r_k}$$

*com* $\lambda_1, \ldots, \lambda_k$ *distintos. Seja* $W_i = \ker (T - \lambda_i I)^{r_i}$*. Então*

(i) *$V = W_1 \oplus \ldots \oplus W_k$.*

(ii) *Cada $W_i$ é invariante sob $T$.*

(iii) *Se $T_i = T|_{W_i}$, então o polinômio mínimo de $T_i$ é $(x - \lambda_i)^{r_i}$.*

*Além disso, se*

$$p = (x - \lambda_1)^{d_1} \ldots (x - \lambda_k)^{d_k}$$

*é o polinômio característico de $T$, então $\dim W_i = d_i$.*

**Prova:** À exceção da última afirmação, o Teorema Espectral é o Teorema da Decomposição Primária para operadores cujos polinômios mínimos são completamente fatoráveis. Basta então provar esta última. Como o polinômio mínimo de $T_i$ é $(x - \lambda_i)^{r_i}$, o polinômio característico de $T_i$ é $(x - \lambda_i)^{e_i}$ onde $e_i = \dim W_i$. Usando a matriz em blocos $[T_i]_{\mathcal{B}_i}$ de $T$, onde $\mathcal{B}_i$ é uma base para $W_i$, é fácil ver que o polinômio característico de $T$ é o produto dos polinômios característicos dos operadores $T_i$. Portanto, necessariamente $e_i = d_i$. ■

Os elementos de $W_i = \ker (T - \lambda_i I)^{r_i}$ são chamados de **autovetores generalizados**. Em particular, operadores complexos possuem bases de autovetores generalizados. A decomposição primária de $T$ neste caso é também chamada **decomposição espectral**.

**Corolário 6.** *Sejam $T, S : V \longrightarrow V$ operadores lineares cujos polinômios mínimos são produtos de fatores lineares. Se $TS = ST$, então $T$ e $S$ possuem a mesma decomposição em soma direta por subespaços invariantes.*

No caso de operadores diagonalizáveis podemos dizer mais. Se dois operadores são simultaneamente diagonalizáveis, isto é, se existe uma mesma base em relação à qual as matrizes de ambos os operadores são diagonais, então os operadores comutam, pois matrizes diagonais comutam. Reciprocamente, a condição de comutatividade dos operadores é suficiente para garantir que eles são simultaneamente diagonalizáveis:

**Teorema 11.** *Sejam $T, S : V \longrightarrow V$ operadores lineares diagonalizáveis. Então, $T$ e $S$ são simultaneamente diagonalizáveis se e somente se $TS = ST$.*

O Teorema 11 se generaliza para uma família qualquer de operadores que comutam entre si, dois a dois.

Continue a considerar um operador $T : V \longrightarrow V$ cujo polinômio mínimo é um produto de fatores lineares, usando a notação do Corolário 5. Se $E_1, \ldots, E_k$ são as projeções associadas à decomposição espectral de $T$, defina um operador $D : V \longrightarrow V$ por

$$D = \lambda_1 E_1 + \ldots + \lambda_k E_k.$$

Pelo Teorema 9, $D$ é um operador diagonalizável. Considere o operador

$$N = T - D.$$

Como $T = TE_1 + \ldots + TE_k$, segue que

$$N = (T - \lambda_1 I) E_1 + \ldots + (T - \lambda_k I) E_k.$$

Usando os fatos que $E_i^2 = E_i$, $E_i E_j = 0$ se $i \neq j$, e que as projeções comutam com $T$, segue que

$$\begin{aligned} N^2 &= (T - \lambda_1 I)^2 E_1 + \ldots + (T - \lambda_k I)^2 E_k, \\ &\vdots \\ N^r &= (T - \lambda_1 I)^r E_1 + \ldots + (T - \lambda_k I)^r E_k. \end{aligned}$$

Em particular, se $r \geqslant r_i$, concluímos que

$$N^r = 0.$$

**Definição 12.** Seja $N : V \longrightarrow V$ um operador linear. Dizemos que $N$ é **nilpotente** se existe algum inteiro $r$ tal que $N^r = 0$.

**Teorema 12.** *Seja $T : V \longrightarrow V$ um operador linear sobre um espaço de dimensão finita $V$ sobre o corpo $\mathbb{F}$. Suponha que o polinômio mínimo de $T$ é um produto de fatores lineares. Então existe um único operador diagonalizável $D : V \longrightarrow V$ e um único operador nilpotente $N : V \longrightarrow V$ tais que*

$$T = D + N$$

*e*

$$DN = ND.$$

*Além disso, cada um deles é um polinômio em $T$.*

**Prova:** Em vista da discussão anterior, só falta provar a unicidade da decomposição (que $D$ e $N$ comutam segue do fato de ambos serem polinômios em $T$). Suponha que $T = D' + N'$, com $D'$ diagonalizável e $N'$ nilpotente, satisfazendo $D'N' = N'D'$. Mostraremos que $D' = D$ e $N' = N$.

Como $D'$ e $N'$ comutam entre si e $T = D' + N'$, segue que $D'$ e $N'$ comutam com $T$ e portanto com qualquer polinômio em $T$, em particular com $D$ e $N$. De $D + N = D' + N'$, segue que

$$D - D' = N' - N.$$

$D - D'$ é um operador diagonalizável. Como $D$ e $D'$ comutam, eles são simultaneamente diagonalizáveis. Como $N$ e $N'$ comutam e são nilpotentes, segue que $N' - N$ também é nilpotente, pois

$$(N' - N)^r = \sum_{i=0}^{r} (-1)^i \binom{r}{i} (N')^{r-i} N^i$$

de modo que se $r$ é suficientemente grande, todo termo no lado esquerdo da expressão será nulo, porque ou $(N')^{r-i} = 0$ ou $N^i = 0$ (ou ambos).

Em particular, $D - D'$ é um operador diagonalizável que também é nilpotente. Como o polinômio mínimo de um operador nilpotente é $x^r$ para algum $r$ e o polinômio mínimo de um operador diagonalizável é um produto de fatores lineares, segue que o polinômio mínimo de $D - D'$ é $x$, ou seja, $D - D'$ é o operador nulo. Portanto, $0 = D - D' = N' - N$. ■

**Corolário 7.** *Se $T$ é um operador linear complexo, então $T$ se decompõe de maneira única como a soma de um operador diagonalizável e um operador nilpotente que comutam. Além disso, eles são polinômios em $T$.*

## 4.12 Exercícios

19. Faça os seguintes exercícios do Capítulo 7 do livro-texto: 2, 6, 7, 8, 11, 12, 16, 26.

20. Seja $V$ um espaço vetorial de dimensão $n$ e $N : V \longrightarrow V$ um operador nilpotente. Então $N^n = 0$.

21. Dê um exemplo de duas matrizes $4 \times 4$ nilpotentes que possuem o mesmo polinômio mínimo mas que não são semelhantes.

22. Seja $V$ um espaço vetorial de dimensão finita e suponha que $T : V \longrightarrow V$ é um operador que comuta com todo operador diagonalizável. Mostre que $T$ é um múltiplo escalar do operador identidade.

23. Sejam $V$ um espaço vetorial de dimensão finita sobre $\mathbb{C}$, $T : V \longrightarrow V$ um linear e $D$ sua parte diagonal. Mostre que se $f \in \mathbb{C}[x]$, então a parte diagonal de $f(T)$ é $f(D)$.

24. Dada $A \in M_n(\mathbb{F})$, considere o operador linear $T : M_n(\mathbb{F}) \longrightarrow M_n(\mathbb{F})$ definido por $T(B) = AB - BA$. Mostre que se $A$ é uma matriz nilpotente, então $T$ é um operador nilpotente.

25. No Corolário 5, prove que se $p = (x - \lambda_1)^{d_1} \dots (x - \lambda_k)^{d_k}$ é o polinômio característico de $T$, então

## 4.13 Exercícios Computacionais

1. Ache os autovalores e correspondentes autoespaços das matrizes seguintes sobre $\mathbb{R}$ e sobre $\mathbb{C}$. Encontre os seus polinômios mínimos. Determine se a matriz é diagonalizável sobre $\mathbb{R}$ e sobre $\mathbb{C}$. Se não for, encontre a sua decomposição primária.

**(a)** $\begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ **(c)** $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$

**(d)** $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix}$ **(e)** $\begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ **(f)** $\begin{bmatrix} 3 & -3 & -4 \\ 0 & 3 & 5 \\ 0 & 0 & -1 \end{bmatrix}$

**(g)** $\begin{bmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{bmatrix}$ **(h)** $\begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 12 & 0 & 3 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}$ **(i)** $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$

**(j)** $\begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$ **(k)** $\begin{bmatrix} 1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ -2 & -2 & 2 & 1 \\ 1 & 1 & -1 & 0 \end{bmatrix}$

2. Sejam

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

Calcule os autovalores e correspondentes autoespaços de $AB$ e $BA$.

3. Calcule o polinômio mínimo das matrizes seguintes sobre $\mathbb{R}$ e sobre $\mathbb{C}$.

**(a)** $\begin{bmatrix} 3 & 0 & -4 & 0 \\ 0 & 3 & 5 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$ **(b)** $\begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$ **(c)** $\begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix}$

# Capítulo 5

# Forma Canônica de Jordan

## 5.1 Complexificação de um Espaço Vetorial

**Definição 1.** Seja $V$ um espaço vetorial real. A **complexificação** de $V$ é o espaço vetorial complexo

$$V_{\mathbb{C}} = \{u + iv : u, v \in V\},$$

com a soma de vetores e multiplicação de um vetor por um escalar complexo definidos de maneira natural, isto é,

$$(u_1 + iv_1) + (u_2 + iv_2) = (u_1 + u_2) + i(v_1 + v_2)$$

e

$$(a + ib)(u + iv) = (au - bv) + i(bu + av).$$

Vetores $w = u + iv$ em $V_{\mathbb{C}}$ tais que $u, v \in V$ e $v = 0$ são chamados **vetores reais**; se $u = 0$, eles são chamados **vetores imaginários puros**. Definimos o **vetor conjugado** de $w$ por

$$\overline{w} = \overline{(u + iv)} = u - iv.$$

**Definição 2.** Seja $V$ um espaço vetorial real e $T : V \longrightarrow V$ um operador linear. A **complexificação** de $T$ é o operador linear $T : V_{\mathbb{C}} \longrightarrow V_{\mathbb{C}}$ definido por

$$T_{\mathbb{C}}(u + iv) = Tu + iTv.$$

Observe que nem todo operador linear sobre $V_{\mathbb{C}}$ é a complexificação de um operador linear real sobre $V$.

**Proposição 1.** *Seja $V$ um espaço vetorial real e $T : V \longrightarrow V$ um operador linear. Então valem as seguintes afirmações:*

(i) *Toda base de $V$ é uma base de $V_{\mathbb{C}}$; em particular,* $\dim V = \dim V_{\mathbb{C}}$.

(ii) *A representação matricial de $T$ em relação a uma base $\mathcal{B}$ para $V$ é a representação matricial de $T_{\mathbb{C}}$ em relação à base $\mathcal{B}$ para $V_{\mathbb{C}}$.*

(iii) *Os polinômios característicos de $T$ e $T_{\mathbb{C}}$ são iguais e os polinômios mínimos de $T$ e $T_{\mathbb{C}}$ são iguais.*

(iv) *$\lambda$ é um autovalor de $T_{\mathbb{C}}$ se e somente se $\overline{\lambda}$ também é um autovalor de $T_{\mathbb{C}}$ e as multiplicidades algébricas de $\lambda$ e $\overline{\lambda}$ são iguais.*

(v) *Se $\widetilde{W}$ é um subespaço de $V_{\mathbb{C}}$ tal que se $w = u + iv \in \widetilde{W}$ então $\overline{w} = u - iv \in \widetilde{W}$ (isto é, $\widetilde{W}$ é invariante sob a operação de conjugação de vetores), então $\widetilde{W}$ possui uma base formada por vetores reais.*

(vi) *Se $\widetilde{W}$ é um subespaço de $V_{\mathbb{C}}$ tal que $\widetilde{W}$ possui uma base formada por vetores reais, então $\widetilde{W}$ é a complexificação de algum subespaço $W$ de $V$, isto é, $\widetilde{W} = W_{\mathbb{C}}$.*

**Prova:** (i) Se $\{x_1, \ldots, x_n\}$ é uma base de $V$, todos os vetores $u, v$ de $V$ se escrevem como combinação linear de $x_1, \ldots, x_n$, logo o mesmo vale para todos os vetores $u + iv$ de $V_{\mathbb{C}}$, pois se $u = u_1 x_1 + \ldots + u_n x_n$ e $v = v_1 x_1 + \ldots + v_n x_n$, então

$$u + iv = (u_1 + iv_1)\, x_1 + \ldots + (u_n + iv_n)\, x_n.$$

Além disso, $x_1, \ldots, x_n$ são também linearmente independentes em $V_{\mathbb{C}}$, pois se existem escalares complexos $\alpha_1 + i\beta_1, \ldots, \alpha_n + i\beta_n$ tais que

$$(\alpha_1 + i\beta_1)\, x_1 + \ldots + (\alpha_n + i\beta_n)\, x_n = 0,$$

então

$$(\alpha_1 x_1 + \ldots + \alpha_n x_n) + i\,(\beta_1 x_1 + \ldots + \beta_n x_n) = 0,$$

donde

$$\begin{aligned} \alpha_1 x_1 + \ldots + \alpha_n x_n &= 0, \\ \beta_1 x_1 + \ldots + \beta_n x_n &= 0, \end{aligned}$$

e portanto $\alpha_1 = \ldots = \alpha_n = 0$ e $\beta_1 = \ldots = \beta_n = 0$.
(ii) Como por (i) qualquer base $\mathcal{B}$ para $V$ é uma base para $V_{\mathbb{C}}$ também e $T_{\mathbb{C}} u = Tu$ quando $u$ é um vetor real, segue que a representação matricial de $T$ em relação a $\mathcal{B}$ é também a representação matricial de $T_{\mathbb{C}}$ em relação a $\mathcal{B}$. (iii) segue imediatamente de (ii).
(iv) Se $\lambda$ é um autovalor de $T_{\mathbb{C}}$ então $\lambda$ é uma raiz do polinômio característico $p$ de $T_{\mathbb{C}}$ de modo que $p(\lambda) = 0$. Mas o polinômio característico de $T_{\mathbb{C}}$ é também o polinômio característico de $T$, logo $p$ é um polinômio com coeficientes reais e portanto $\overline{p(\lambda)} = p\left(\overline{\lambda}\right)$. Assim, tomando o conjugado em ambos os lados da equação $p(\lambda) = 0$, concluímos que $p\left(\overline{\lambda}\right) = 0$. Além disso, dividindo sucessivamente por $(x - \lambda)\left(x - \overline{\lambda}\right)$, concluímos que a multiplicidade algébrica de $\lambda$ e $\overline{\lambda}$ é a mesma.
(v) Seja $\{z_1, \ldots, z_k\}$ uma base para $\widetilde{W}$, com $z_j = x_j + iy_j$ e $x_j, y_j \in V$ para todo $j$. Como $\overline{z_j} \in \widetilde{W}$ segue que

$$\begin{aligned} \frac{z_j + \overline{z_j}}{2} &= x_j \in \widetilde{W}, \\ \frac{z_j - \overline{z_j}}{2} &= y_j \in \widetilde{W}. \end{aligned}$$

Como os vetores $z_j$ são combinações lineares dos vetores $x_j, y_j$, concluímos que $x_1, \ldots, x_k, y_1, \ldots, yk$ são vetores reais que geram $\widetilde{W}$. Dentre estes vetores podemos escolher um subconjunto minimal L.I. que ainda gera $\widetilde{W}$, obtendo uma base de vetores reais para $\widetilde{W}$.
(vi) Seja $\{x_1, \ldots, x_k\}$ uma base para $\widetilde{W}$ formada por vetores reais $x_1, \ldots, x_k \in V$. Considere o subespaço $W$ de $V$ gerado pelos vetores $x_1, \ldots, x_k$. Como todo vetor de $\widetilde{W}$ se escreve como uma combinação linear

$$(\alpha_1 + i\beta_1)\, x_1 + \ldots + (\alpha_n + i\beta_n)\, x_n = (\alpha_1 x_1 + \ldots + \alpha_n x_n) + i\,(\beta_1 x_1 + \ldots + \beta_n x_n)$$

segue que $\widetilde{W} = W_{\mathbb{C}}$. ■
O espaço vetorial complexo $V_{\mathbb{C}}$ tem dimensão complexa $n = \dim V$. Porém, o espaço vetorial $V_{\mathbb{C}}$ também pode ser visto como um espaço vetorial real. Neste caso, $\dim V_{\mathbb{C}} = 2n$. De fato, se $\{x_1, \ldots, x_n\}$ é uma base de $V$, então $\{x_1, \ldots, x_n, ix_1, \ldots, ix_n\}$ é uma base para $V_{\mathbb{C}}$ sobre $\mathbb{R}$.

**Exemplo 0.** Nem todo subespaço $\widetilde{W}$ de $V_{\mathbb{C}}$ é a complexificação de algum subespaço $W$ de $V$. O subespaço $\widetilde{W} = \langle u + iv \rangle$ de $V_{\mathbb{C}}$ gerado pelo vetor $u + iv$ com $u, v \in V$, $u, v \neq 0$ e $u$ não é um múltiplo escalar de $v$, não é a complexificação de nenhum subespaço de $V$. De fato, segue da Proposição 1 (v) e (vi) que $\widetilde{W}$ é a complexificação de algum subespaço de $V$ se e somente se $\widetilde{W}$ é invariante por conjugação (porque um subespaço gerado por vetores reais é invariante por conjugação). Afirmamos que o vetor $u - iv \notin \widetilde{W}$. Com efeito, os vetores de $\widetilde{W}$ são da forma $(a + ib)(u + iv)$ para alguns $a, b \in \mathbb{R}$, donde um vetor típico de $\widetilde{W}$ é da forma $(au - bv) + i(bu + av)$. Para que tenhamos $u - iv \in \widetilde{W}$, é necessário que existam $a, b \in \mathbb{R}$ tais que

$$\begin{cases} au - bv = u \\ bu + av = -v \end{cases}.$$

Se $b = 0$, teríamos $a = 1$ na primeira equação e $a = -1$ na segunda, um absurdo. Se $b \neq 0$, segue da primeira e da segunda equação, respectivamente, que

$$\begin{aligned} v &= \frac{a-1}{b}u, \\ u &= -\frac{a+1}{b}v. \end{aligned}$$

Isso contraria o fato que $u$ e $v$ não são múltiplos escalares um do outro. $\square$

## 5.2 Forma de Jordan

Dado um operador linear, o objetivo de obter uma representação matricial para este operador *a mais diagonal possível* é obtido através da forma de Jordan. Nesta seção mostraremos que todos os operadores lineares cujos polinômios característicos se fatoram completamente, o que inclui operadores complexos, são representado por uma matriz na forma de Jordan..

O fato de um operador linear cujo polinômio característico é completamente fatorável deixar de ser diagonalizável não pode ser atribuído à falta de autovalores, já que todas as raízes do polinômio característico estão presentes. O problema está na falta de autovetores suficientes para produzir uma base para o espaço. Se existe um número suficiente de autovetores, então o operador é diagonalizável por definição e a sua forma de Jordan coincide com a sua forma diagonal. Caso contrário, para cada autovetor que faltar a forma de Jordan terá um 1 acima da diagonal, acima do autovalor correspondente.

**Definição 2.** Seja $J$ uma matriz sobre um corpo $\mathbb{F}$ e $\lambda_1, \ldots, \lambda_k$ seus autovalores distintos. Dizemos que $J$ está na **forma de Jordan** se

$$J = \begin{bmatrix} J_1 & 0 & \ldots & 0 \\ 0 & J_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & J_k \end{bmatrix}$$

com cada bloco $J_i$ na forma em blocos

$$J_i = \begin{bmatrix} J_1^i & 0 & \ldots & 0 \\ 0 & J_2^i & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & J_{n_i}^i \end{bmatrix}$$

em que os blocos $J_l^i$ tem a forma

$$J_l^i = \begin{bmatrix} \lambda_i & 1 & 0 & \ldots & \ldots & 0 \\ 0 & \lambda_i & 1 & \ldots & \ldots & 0 \\ 0 & 0 & \lambda_i & \ddots & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & \lambda_i & 1 \\ 0 & 0 & \ldots & \ldots & 0 & \lambda_i \end{bmatrix}$$

e seu tamanho decresce à medida que $l$ aumenta. O bloco $J^i$ é chamado um **bloco de Jordan** associado ao autovalor $\lambda_i$.

Se uma matriz $A$ é semelhante a uma matriz $J$ na forma de Jordan, dizemos que $J$ é a **forma de Jordan** de $A$.

Cada bloco de Jordan é uma matriz triangular com apenas um autovalor da matriz $A$ e apenas um autovetor. Quando o bloco tem tamanho $m$, o autovalor $\lambda_i$ é repetido $m$ vezes na diagonal e existem $m - 1$ 1's acima da diagonal. O mesmo autovalor $\lambda_i$ pode aparecer em vários blocos, o número de blocos distintos em que ele aparece correspondendo ao número de autovetores linearmente independentes correspondentes a $\lambda_i$. Em outras palavras, o número $n_i$ de blocos de Jordan que aparece na matriz $J_i$ corresponde à dimensão do autoespaço associado ao autovalor $\lambda_i$. Esta descrição ainda não permite determinar o tamanho dos diversos blocos de Jordan $J_l^i$ que aparecem em $J_i$. Isto será visto daqui a pouco.

Antes de demonstrar que toda matriz cujo polinômio característico é completamente fatorável é semelhante a uma matriz na forma de Jordan e que a forma de Jordan determina a classe de semelhança da matriz, construindo no processo um algoritmo para determinar a forma de Jordan exata de uma matriz, vamos considerar alguns exemplos:

**Exemplo 1.** As matrizes

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{e} \quad C = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

têm a mesma forma de Jordan

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

De fato, todas estas matrizes tem apenas o autovalor 1 e o autoespaço correspondente com dimensão 1. Logo, como as matrizes são $2 \times 2$, existe apenas um bloco de Jordan. De fato, vemos que as únicas formas de Jordan para matrizes $2 \times 2$ são

$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \quad \text{ou} \quad \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}.$$

O primeiro caso corresponde a uma matriz diagonalizável (e podemos ter ), enquanto que o segundo corresponde a uma matriz que possui um único autovalor com autoespaço correspondente de dimensão 1. Uma matriz real que não possui autovalores evidentemente não possuirá uma forma de Jordan. □

**Exemplo 2.** A matriz

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

tem a forma de Jordan

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

De fato, o único autovalor de $A$ é 0 e a dimensão de seu autoespaço é 1. □

**Exemplo 3.** A matriz

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

tem a forma de Jordan

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

De fato, o único autovalor de $B$ é 0 e a dimensão de seu autoespaço é 2. Os Exemplos 2 e 3 ilustram que também é fácil determinar as formas de Jordan de matrizes $3 \times 3$, bastando para isso encontrar os autovalores da matriz e as dimensões dos autoespaços associados. De fato, as únicas formas de Jordan possíveis para matrizes $3 \times 3$ são

$$\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}, \quad \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix} \quad \text{ou} \quad \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}.$$

O primeiro caso corresponde a uma matriz diagonalizável (os três autovalores podem ser distintos ou iguais, ou apenas dois dos autovalores são distintos); o segundo caso corresponde a um único autoespaço (se $\lambda_1 = \lambda_2$) ou dois autoespaços (se $\lambda_1 \neq \lambda_2$) com dimensão total 2, o terceiro espaço corresponde a um único autoespaço com dimensão 1. $\square$

**Exemplo 4.** Para matrizes $4 \times 4$ em diante, a estratégia de contar as dimensões dos autoespaços associados aos autovalores é em geral insuficiente para determinar a forma de Jordan de uma matriz.. Por exemplo, as matrizes a seguir, já dadas na forma de Jordan,

$$J_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{e} \quad J_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

têm ambas apenas um único autovalor, 0, cujo autoespaço tem dimensão 2, mas são formas de Jordan distintas por definição. De fato, $J_1$ e $J_2$ não são semelhantes porque $J_1$ tem polinômio mínimo $x^3$, enquanto que $J_2$ tem polinômio mínimo $x^2$. $\square$

**Exemplo 5.** Por outro lado, as matrizes $4 \times 4$ abaixo estão em forma de Jordan distintas

$$A = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix},$$

têm o mesmo polinômio característico $(x-2)^4$ e o mesmo polinômio mínimo $(x-2)^2$. Como elas representam formas de Jordan distintas, elas não são semelhantes. De fato, o autoespaço de $A$ associado ao autovalor 2 tem dimensão 2, enquanto que o autoespaço de $B$ associado ao autovalor 2 tem dimensão 3. $\square$

Observando a forma de Jordan, vemos que ela permite provar um resultado interessante:

**Proposição 2.** *Sejam*

$$p_c = (x - \lambda_1)^{d_1} \ldots (x - \lambda_k)^{d_k},$$
$$p_m = (x - \lambda_1)^{r_1} \ldots (x - \lambda_k)^{r_k},$$

*com $r_i \leqslant d_i$ para todo $i$. Então existe uma matriz que possui $p_c$ como polinômio característico e $p_m$ como polinômio mínimo.*

**Prova:** Basta tomar uma matriz em forma de Jordan em que o bloco associado ao autovalor $\lambda_i$ tem tamanho $d_i$ e possui o seu primeiro bloco de Jordan $J_1^i$ de tamanho $r_i$, pois o polinômio mínimo de $J_1^i$ é exatamente $(x - \lambda_i)^{r_i}$. ■

Seja $T : V \longrightarrow V$ um operador linear e suponha que $T$ possa ser representado na forma de Jordan $J$ em relação a alguma base $\mathcal{B}$ de $V$. Considere um dos blocos de Jordan de $J$ associados ao autovalor $\lambda$:

$$J_\lambda = \begin{bmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & \ldots & 0 \\ 0 & 0 & \lambda & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda & 1 \\ 0 & 0 & \ldots & 0 & \lambda \end{bmatrix}$$

Seja $\mathcal{B}' \subset \mathcal{B}$ a base do subespaço invariante $W$ associado a este bloco. Se $\mathcal{B}' = \{v_1, \ldots, v_r\}$, então

$$Tv_1 = \lambda v_1 \quad \text{e} \quad Tv_j = v_{j-1} + \lambda v_j \text{ para } j = 2, \ldots, r.$$

Dizemos que os vetores $v_1, \ldots, v_r$ formam uma **cadeia de Jordan** de comprimento $r$. Observe que $(T - \lambda I)\, v_j = v_{j-1}$ para $j = 2, \ldots, r$ e $(T - \lambda I)\, v_1 = 0$. Portanto,

$$\begin{aligned} (T - \lambda I)^r\, v_r &= 0, \\ (T - \lambda I)^{r-1}\, v_{r-1} &= 0, \\ &\vdots \\ (T - \lambda I)^2\, v_2 &= 0, \\ (T - \lambda I)\, v_1 &= 0. \end{aligned}$$

Temos $(T - \lambda I)^r\, v = 0$ para todo $v \in W$, o que implica que $(x - \lambda)^r$ é o polinômio mínimo para $T|_W$. Isso motiva a seguinte definição:

**Definição 3.** Seja $T : V \longrightarrow V$ um operador linear e $\lambda$ um autovalor de $T$. Dizemos que um vetor $v \in V$, $v \neq 0$, é $(T - \lambda I)$**-cíclico** se existir um inteiro positivo $r$ tal que

$$(T - \lambda I)^r\, v = 0.$$

O menor inteiro positivo com esta propriedade é chamado o **período** de $v$ relativo a $T - \lambda I$.

**Lema 1.** *Se $v$ é $(T - \lambda I)$-cíclico com período $r$, então os vetores*

$$v, (T - \lambda I)\, v, \ldots, (T - \lambda I)^{r-1}\, v$$

*são linearmente independentes.*

**Prova:** Denote $S = T - \lambda I$. Então temos que provar que $v, Sv, \ldots, S^{r-1}v$ são linearmente independentes. Sabemos que $S^r v = 0$ e $S^{r-1}v \neq 0$. Isso significa que $x^r$ é o $S$-anulador de $v$ (o $S$-anulador de um vetor é o seu $S$-condutor para o subespaço nulo). Os vetores $v, Sv, \ldots, S^{r-1}v$ serem linearmente dependentes é equivalente à existência de um polinômio não-nulo

$$f = a_0 + a_1 x + \ldots + a_{r-1} x^{r-1}$$

de grau $r - 1$ tal que

$$f(S)\, v = 0.$$

Mas isso contradiz o fato que $x^r$ é o $S$-anulador de $v$. ■

**Definição 4.** Dizemos que um subespaço vetorial $W \subset V$ é $(T - \lambda I)$**-cíclico** se existir algum vetor $v \in W$ e um autovalor $\lambda$ de $T$ tal que $v$ é $(T - \lambda I)$-cíclico de período $r$ e $W$ é gerado por $v, Tv, \ldots, T^{r-1}v$.

Segue do Lema 1 que se $W$ é um subespaço cíclico gerado pelo vetor cíclico $v$ de período $r$, então

$$\mathcal{B} = \left\{ (T - \lambda I)^{r-1} v, (T - \lambda I)^{r-2} v, \ldots, (T - \lambda I) v, v \right\}$$

é uma base para $W$. Em relação à base $\mathcal{B}$, a matriz de $T|_W$ é um bloco de Jordan. De fato, denotando

$$\begin{aligned}
v_1 &= (T - \lambda I)^{r-1} v, \\
v_2 &= (T - \lambda I)^{r-2} v, \\
&\vdots \\
v_j &= (T - \lambda I)^{r-j} v, \\
&\vdots \\
v_{r-1} &= (T - \lambda I) v, \\
v_r &= v,
\end{aligned}$$

temos

$$Tv_1 = T (T - \lambda I)^{r-1} v = (T - \lambda I + \lambda I) (T - \lambda I)^{r-1} v = (T - \lambda I)^r v + \lambda (T - \lambda I)^{r-1} v = \lambda v_1$$

e, para cada $j = 2, \ldots, r$,

$$Tv_j = T (T - \lambda I)^{r-j} v = (T - \lambda I + \lambda I) (T - \lambda I)^{r-j} v = (T - \lambda I)^{r-(j-1)} v + \lambda (T - \lambda I)^{r-j} v = v_{j-1} + \lambda v_j.$$

Provar a existência da forma de Jordan para um operador linear $T : V \longrightarrow V$ é portanto equivalente a encontrar uma decomposição em soma direta de $V$ por subespaços cíclicos. Procedemos a esta tarefa agora:

**Teorema 1.** (Forma de Jordan) *Seja $T : V \longrightarrow V$ um operador linear cujo polinômio característico se escreve como um produto de fatores lineares. Então existe uma base para $V$ em relação à qual $T$ é representado por uma matriz na forma de Jordan. A forma de Jordan de $T$ é única a menos da ordem de seus autovalores.*

**Prova:** (Existência) Sejam $\lambda_1, \ldots, \lambda_k$ os autovalores distintos de $T$ e

$$p = (x - \lambda_1)^{r_1} \ldots (x - \lambda_k)^{r_k}$$

o polinômio mínimo de $T$. Pelo Teorema da Decomposição Primária (Teorema Espectral), se $W_i = \ker (T - \lambda_i I)^{r_i}$ temos que $W_i$ é invariante sob $T$, $V = W_1 \oplus \ldots \oplus W_k$ e $(x - \lambda_i)^{r_i}$ é o polinômio mínimo de $T_i = T|_{W_i}$. Vamos provar que cada subespaço $W_i$ é a soma direta de subespaços $(T - \lambda_i I)$-cíclicos. No que se segue, omitiremos o índice $i$ por motivos de clareza de apresentação.

A demonstração será por indução na dimensão de $W$. Seja $n = \dim W$ e assuma que o teorema que queremos provar é válido para todo espaço vetorial de dimensão menor que $n$ (para $n = 1$ o resultado é trivial, pois toda matriz $1 \times 1$ já está na forma de Jordan). Temos $(T - \lambda I)^r = 0$ e $(T - \lambda I)^{r-1} \neq 0$ sobre $W$. O subespaço $(T - \lambda I) W$ tem dimensão estritamente menor que a de $W$. De fato, existe pelo menos um vetor $v \in W$ tal que $(T - \lambda I)^{r-1} v = w \neq 0$ e $0 = (T - \lambda I)^r v = (T - \lambda I) (T - \lambda I)^{r-1} v = (T - \lambda I) w$, isto é, $w$ é um autovetor de $T$ em $W$ associado a $\lambda$; em particular, $\dim (\ker (T - \lambda I)) \geqslant 1$ e segue do Teorema do Núcleo e da Imagem que

$$\dim W = \dim (\operatorname{im} (T - \lambda I)) + \dim (\ker (T - \lambda I)) > \dim (\operatorname{im} (T - \lambda I)).$$

Pela hipótese de indução, podemos escrever

$$(T - \lambda I)\, W = U_1 \oplus \ldots \oplus U_m,$$

onde cada subespaço $U_i$ é $(T - \lambda I)$-cíclico, gerado por algum vetor cíclico $u_i \in U_i$ de período $s_i$.

Seja $v_i \in W$ tal que $(T - \lambda I)\, v_i = u_i$. Então cada vetor $v_i$ é um vetor cíclico, pois se $(T - \lambda I)^{s_i} u_i = 0$ então $(T - \lambda I)^{s_i+1} v_i = 0$. Seja $V_i$ o subespaço cíclico de $W$ gerado pelos vetores

$$v_i, (T - \lambda I)\, v_i, \ldots, (T - \lambda I)^{s_i - 1} v_i, (T - \lambda I)^{s_i} v_i.$$

Considere o subespaço $V' = V_1 + \ldots + V_m$. Afirmamos que

$$V' = V_1 \oplus \ldots \oplus V_m, \tag{5.1}$$

isto é, os subespaços cíclicos $V_1, \ldots, V_m$ são L.I. De fato, suponhamos que existam vetores $w_1 \in V_1, \ldots, w_m \in V_m$ tais que

$$w_1 + \ldots + w_m = 0.$$

Queremos mostrar que cada $w_i = 0$. Para isso, lembre-se que todo vetor em $V_i$ é da forma $f_i\,(T - \lambda I)\, v_i$ para algum polinômio $f_i$ de grau $\leqslant s_i$. Logo, podemos escrever

$$f_1\,(T - \lambda I)\, v_1 + \ldots + f_m\,(T - \lambda I)\, v_m = 0. \tag{5.2}$$

Aplicando $T - \lambda I$ e observando que $(T - \lambda I)\, f_i\,(T - \lambda I) = f_i\,(T - \lambda I)\,(T - \lambda I)$, obtemos

$$f_1\,(T - \lambda I)\, u_1 + \ldots + f_m\,(T - \lambda I)\, u_m = 0.$$

Mas os espaços $U_1, \ldots, U_m$ são linearmente independentes, logo

$$f_1\,(T - \lambda I)\, u_1 = \ldots = f_m\,(T - \lambda I)\, u_m = 0.$$

Segue que o polinômio mínimo $x^{s_i}$ divide $f_i$; em particular, $x$ é um fator de $f_i$, ou seja, $f_i = x g_i$ para algum polinômio $g_i$, o que implica

$$f_i\,(T - \lambda I) = (T - \lambda I)\, g_i\,(T - \lambda I) = g_i\,(T - \lambda I)\,(T - \lambda I)\,.$$

Substituindo esta expressão em (6.28), obtemos

$$g_1\,(T - \lambda I)\, u_1 + \ldots + g_m\,(T - \lambda I)\, u_m = 0$$

Novamente,

$$g_1\,(T - \lambda I)\, u_1 = \ldots = g_m\,(T - \lambda I)\, u_m = 0$$

donde $x^{s_i}$ divide $g_i$, o que por sua vez implica que $x^{s_i+1}$ divide $f_i$ e, como $(T - \lambda I)^{s_i+1} v_i = 0$, concluímos que

$$f_1\,(T - \lambda I)\, v_1 = \ldots = f_m\,(T - \lambda I)\, v_m = 0.$$

Agora mostraremos que

$$W = V' + \ker\,(T - \lambda I)\,. \tag{5.3}$$

De fato, note que $(T - \lambda I)\, W = (T - \lambda I)\, V'$ pois todo vetor de $(T - \lambda I)\, W$ é da forma

$$f_1\,(T - \lambda I)\, u_1 + \ldots + f_m\,(T - \lambda I)\, u_m = (T - \lambda I)\,[f_1\,(T - \lambda I)\, v_1 + \ldots + f_m\,(T - \lambda I)\, v_m] \in (T - \lambda I)\, V'$$

para alguns polinômios $f_i$. Daí, dado $v \in W$, temos $(T - \lambda I)\, v = (T - \lambda I)\, v'$ para algum vetor $v' \in V'$ e portanto

$$(T - \lambda I)\,(v - v') = 0,$$

donde $v - v' \in \ker(T - \lambda I)$. Escrevendo $v = v' + (v - v')$, provamos (5.3).

Esta soma, no entanto, não é necessariamente direta. Porém, tomando uma base de Jordan $\mathcal{B}'$ para $V'$, podemos estender $\mathcal{B}'$ a uma base para $V$ adicionando vetores $v_1, \ldots, v_s$ de $\ker(T - \lambda I)$. Cada $v_j$ satisfaz $(T - \lambda I)\, v_j = 0$, logo é um autovetor de $T$ e o espaço de dimensão 1 gerado por $v_j$ é obviamente um espaço cíclico. Obtemos então a decomposição em soma direta de subespaços cíclicos de $W$ desejada:

$$W = V' \oplus \langle v_1 \rangle \oplus \ldots \oplus \langle v_s \rangle$$

A unicidade da forma de Jordan será provada na seção a seguir (veja o Teorema 2 e a discussão que lhe precede). ■

Observe que como a demonstração do Teorema 1 foi por indução, ele ainda não nos diz como obter a forma de Jordan no caso geral (e muito menos a base de vetores de $V$ em relação à qual o operador $T$ é representado por uma matriz na forma de Jordan; em outras palavras, ele não nos diz como obter as cadeias de Jordan), apenas garante que todo operador linear cujo polinômio característico é completamente fatorável possui uma forma de Jordan.

**Corolário 1.** *Seja $A$ uma matriz complexa. Então $A$ é semelhante a uma matriz na forma de Jordan, única a menos da ordem de seus autovalores.*

**Corolário 2.** *Matrizes que possuem formas de Jordan são semelhantes se e somente se elas possuem a mesma forma de Jordan.*

## 5.3 Cálculo da Forma de Jordan

A demonstração por indução da existência da forma de Jordan (para operadores cujos polinômios característicos são completamente fatoráveis) não produz um algoritmo para o seu cálculo. Nesta seção vamos considerar alguns algoritmos. Seja $T$ um operador com polinômio característico

$$p_c = (x - \lambda_1)^{d_1} \ldots (x - \lambda_k)^{d_k}$$

e polinômio mínimo

$$p_m = (x - \lambda_1)^{r_1} \ldots (x - \lambda_k)^{r_k},$$

onde $\lambda_1, \ldots, \lambda_k$ são os autovalores distintos de $T$. Usando o Teorema da Decomposição Primária, sabemos que existe uma base $\mathcal{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_k\}$ para $T$ em relação à qual a matriz de $T$ assume a forma diagonal em blocos

$$J = \begin{bmatrix} J_1 & 0 & \ldots & 0 \\ 0 & J_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & J_k \end{bmatrix}.$$

Cada bloco $J_i$ é a representação matricial de $T|_{W_i}$ com relação à base $\mathcal{B}_i$ de $W_i$, onde $W_i = \ker(T - \lambda_i I)^{r_i}$ é o autoespaço generalizado associado ao autovalor $\lambda_i$ e $\dim W_i = d_i$; portanto, cada bloco $J_i$ tem tamanho $d_i \times d_i$. Queremos escrever cada bloco $J_i$ na forma diagonal em blocos de Jordan

$$J_i = \begin{bmatrix} J_1^i & 0 & \ldots & 0 \\ 0 & J_2^i & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & J_{n_i}^i \end{bmatrix}$$

em que os blocos $J_l^i$ tem a forma

$$J_l^i = \begin{bmatrix} \lambda_i & 1 & 0 & \dots & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & \dots & 0 \\ 0 & 0 & \lambda_i & \ddots & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \lambda_i & 1 \\ 0 & 0 & \dots & \dots & 0 & \lambda_i \end{bmatrix}.$$

Na discussão que se segue freqüentemente omitiremos o índice $i$ objetivando maior clareza na exposição. Para cada autovalor $\lambda = \lambda_i$ defina

$$\delta_j = \dim \left[ \ker (T - \lambda I)^j \right], \quad \text{para } j = 1, \dots, r$$

onde $r = r_i$. Os números $\delta_j$ são chamados **índices de deficiência** do autovalor $\lambda$. Observe que $\delta_1 = \dim [\ker (T - \lambda I)]$ é exatamente a dimensão do autoespaço associado ao autovalor $\lambda$, isto é, o número máximo de autovetores independentes associados ao autovalor $\lambda$, enquanto que $\delta_r = \dim [\ker (T - \lambda I)^r] = d_i = d$ é a dimensão do autoespaço generalizado $W = W_i$ associado a $\lambda$. O índice de deficiência $\delta_j$ é o número de colunas nulas na matriz escalonada reduzida de $(T - \lambda I)^j |_W$, daí o nome; esta matriz tem tamanho $d \times d$, e o número de deficiência varia desde o número mínimo de deficiência $\delta_1$ até o número máximo $\delta_r = d$, caso em que a matriz de $(T - \lambda I)^r |_W$ é a matriz nula.

Defina agora para cada autovalor $\lambda = \lambda_i$

$$\nu_j = \text{ número de blocos de Jordan de tamanho } j \times j \text{ em } J_i, \quad \text{para } j = 1, \dots, r.$$

Sabemos que não existem blocos de Jordan de tamanho maior que $r$ porque o polinômio mínimo de $T_W$ é $(x - \lambda)^r$. Temos

$$\delta_1 = \nu_1 + \nu_2 + \dots + \nu_r, \tag{5.4}$$

pois $\delta_1$ é o número total de blocos de Jordan presentes em $J_i$ (lembre-se que cada um dos autovetores linearmente independentes dá origem a um bloco de Jordan). Em seguida, considere $\delta_2 = \dim \left[ \ker (T - \lambda I)^2 \right]$. Cada bloco de Jordan $1 \times 1$ contribui um vetor para a base $\ker (T - \lambda I)^2$, cada bloco de Jordan $2 \times 2$ contribui dois vetores para $\ker (T - \lambda I)^2$, enquanto que blocos $j \times j$ com $j \geqslant 3$ contribuem também dois vetores; de fato,

$$\begin{array}{lcl} (T - \lambda I) v_1 = 0 & & (T - \lambda I)^2 v_1 = 0 \\ (T - \lambda I) v_2 = v_1, & & (T - \lambda I)^2 v_2 = 0, \\ (T - \lambda I) v_3 = v_2, & & (T - \lambda I)^2 v_3 = v_1, \\ (T - \lambda I) v_4 = v_3, & \Rightarrow & (T - \lambda I)^2 v_4 = v_2, \\ \vdots & & \vdots \\ (T - \lambda I) v_j = v_{j-1} & & (T - \lambda I)^2 v_j = v_{j-2}. \end{array}.$$

Portanto,

$$\delta_2 = \nu_1 + 2\nu_2 + \dots + 2\nu_r. \tag{5.5}$$

Considere agora $\delta_3 = \dim \left[ \ker (T - \lambda I)^3 \right]$. Cada bloco de Jordan $1 \times 1$ contribui um vetor para a base $\ker (T - \lambda I)^3$, cada bloco de Jordan $2 \times 2$ contribui dois vetores para $\ker (T - \lambda I)^3$, cada bloco de Jordan $3 \times 3$ contribui três vetores para $\ker (T - \lambda I)^3$, enquanto que blocos $j \times j$ com $j \geqslant 4$ contribuem também três vetores; de fato,

$$\begin{array}{lll}
(T-\lambda I)\,v_1 = 0 & (T-\lambda I)^2\,v_1 = 0 & (T-\lambda I)^3\,v_1 = 0 \\
(T-\lambda I)\,v_2 = v_1, & (T-\lambda I)^2\,v_2 = 0, & (T-\lambda I)^3\,v_2 = 0, \\
(T-\lambda I)\,v_3 = v_2, & (T-\lambda I)^2\,v_3 = v_1, & (T-\lambda I)^3\,v_3 = 0, \\
(T-\lambda I)\,v_4 = v_3, \quad \Rightarrow & (T-\lambda I)^2\,v_4 = v_2, \quad \Rightarrow & (T-\lambda I)^3\,v_4 = v_1, \\
\vdots & \vdots & \vdots \\
(T-\lambda I)\,v_j = v_{j-1} & (T-\lambda I)^2\,v_j = v_{j-2}. & (T-\lambda I)^2\,v_j = v_{j-3}.
\end{array}.$$

Assim,

$$\delta_3 = \nu_1 + 2\nu_2 + 3v_3 + \ldots + 3\nu_r. \tag{5.6}$$

Em geral,

$$\delta_j = \nu_1 + 2\nu_2 + 3v_3 + \ldots + jv_j \ldots + j\nu_r = \sum_{l=1}^{j-1} l\nu_l + j\sum_{l=j}^{r} \nu_l. \tag{5.7}$$

Desta forma, obtemos um sistema com $r$ equações a $r$ incógnitas:

$$\left\{\begin{array}{lcl}
\nu_1 + \nu_2 + \ldots + \nu_r & = & \delta_1 \\
\nu_1 + 2\nu_2 + \ldots + 2\nu_r & = & \delta_2 \\
\nu_1 + 2\nu_2 + 3v_3 + \ldots + 3\nu_r & = & \delta_3 \\
\vdots & & \vdots \\
\nu_1 + 2\nu_2 + 3v_3 + \ldots + (r-1)\,v_{r-1} + (r-1)\,\nu_r & = & \delta_{r-1} \\
\nu_1 + 2\nu_2 + 3v_3 + \ldots + (r-1)\,v_{r-1} + r\nu_r & = & \delta_r
\end{array}\right..$$

Os valores dos índices de deficiência $\delta_1, \delta_2, \ldots, \delta_r$ devem ser calculados diretamente a partir do operador $T$. Observe que a matriz do sistema acima

$$\begin{bmatrix}
1 & 1 & 1 & \ldots & 1 \\
1 & 2 & 2 & \ldots & 2 \\
1 & 2 & 3 & \ldots & 3 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & 2 & 3 & \ldots & r
\end{bmatrix}$$

possui uma inversa com forma bastante simples

$$\begin{bmatrix}
2 & -1 & 0 & \ldots & \ldots & \ldots & 0 \\
-1 & 2 & -1 & 0 & \ldots & \ldots & 0 \\
0 & -1 & 2 & -1 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\
0 & \ldots & 0 & -1 & 2 & -1 & 0 \\
0 & \ldots & \ldots & 0 & -1 & 2 & -1 \\
0 & \ldots & \ldots & \ldots & 0 & -1 & 1
\end{bmatrix},$$

ou seja, uma matriz tridiagonal com $-1$'s nas diagonais secundárias e 2's na diagonal principal, exceto pelo último elemento da diagonal principal que é igual a 1. Por exemplo, para $r = 5$, a matriz do sistema e sua inversa são

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

A verificação deste fato pode ser feita simplesmente multiplicando as duas matrizes (e também pode-se provar que ambas as matrizes possuem determinante igual a 1; verifique, calculando o determinante por escalonamento). Em particular, existe uma única solução $\nu_1, \nu_2, \dots, \nu_r$ para o sistema, *o que prova a unicidade da forma de Jordan*. Resumimos a discussão acima no seguinte teorema:

**Teorema 2.** *Seja $T : V \longrightarrow V$ um operador linear cujo polinômio característico é*

$$p_c = (x - \lambda_1)^{d_1} \dots (x - \lambda_k)^{d_k}$$

*e cujo polinômio mínimo é*

$$p_m = (x - \lambda_1)^{r_1} \dots (x - \lambda_k)^{r_k}.$$

*Então a forma de Jordan de $T$ é completamente determinada pelos índices de deficiência dos autovalores de $T$*

$$\delta_j^i = \dim \left[ \ker (T - \lambda_i I)^j \right], \quad j = 1, \dots, r_i, \quad i = 1, \dots, k.$$

*Mais precisamente, se $\nu_j^i$ é o número de blocos de Jordan de tamanho $j \times j$ associados ao autovalor $\lambda_i$, para $j = 1, \dots, r_i$, temos*

$$\begin{bmatrix} \nu_1^i \\ \nu_2^i \\ \nu_3^i \\ \\ \nu_{r_i-2}^i \\ \nu_{r_i-1}^i \\ \nu_{r_i}^i \end{bmatrix} = \begin{bmatrix} 2 & -1 & & & & & \\ -1 & 2 & -1 & & & & \\ & -1 & 2 & -1 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & -1 & 2 & -1 & \\ & & & & -1 & 2 & -1 \\ & & & & & -1 & 1 \end{bmatrix} \begin{bmatrix} \delta_1^i \\ \delta_2^i \\ \delta_3^i \\ \\ \delta_{r_i-2}^i \\ \delta_{r_i-1}^i \\ \delta_{r_i}^i \end{bmatrix}.$$

**Exemplo 6.** Encontre a forma de Jordan para a matriz

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

O polinômio característico de $A$ é $(x-1)^4$. Temos

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

de modo que o autoespaço associado ao autovalor 1 tem base $\{(1, -1, 0, 0), (1, 0, 0, -1)\}$ e portanto $\delta_1 = 2$. Em seguida,

$$(A - I)^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

de modo que $\delta_2 = 4$ e o polinômio mínimo de $A$ é $(x-1)^2$. Segue que não há blocos de Jordan de tamanho maior que 2 e

$$\begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}.$$

Concluímos que a forma de Jordan da matriz $A$ é

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

□

**Exemplo 7.** Encontre a forma de Jordan para a matriz

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

O polinômio característico de $B$ é $(x-1)^4$. Temos

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

de modo que o autoespaço associado ao autovalor 1 tem base $\{(1,-1,0,0),(1,0,0,-1)\}$ e portanto $\delta_1 = 2$. Em seguida,

$$(B-I)^2 = \begin{bmatrix} -1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

de modo que $\delta_2 = 3$. Finalmente,

$$(B-I)^3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

donde $\delta_3 = 4$ e o polinômio mínimo de $B$ é $(x-1)^3$. Segue que não há blocos de Jordan de tamanho maior que 3 e

$$\begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{bmatrix} = \begin{matrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{matrix} \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Concluímos que a forma de Jordan da matriz $B$ é

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

□

**Exemplo 8.** Encontre a forma de Jordan para a matriz

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

O polinômio característico de $C$ é $x\,(x-2)^5$. É claro que temos um único bloco de Jordan de tamanho 1 para o autovalor 0. Para o autovalor 2 temos

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

de modo que o autoespaço associado ao autovalor 2 tem base $\{(1,1,0,0,0,0)\,,(0,0,1,-1,0,0)\}$ e portanto $\delta_1 = 2$. Em seguida,

$$(C-2I)^2 = \begin{bmatrix} 0 & 0 & 2 & 2 & 0 & 0 \\ 0 & 0 & 2 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 \\ 0 & 0 & 0 & 0 & -2 & 2 \end{bmatrix}, \quad \text{de modo que } \delta_2 = 4;$$

$$(C-2I)^3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -4 & 4 \\ 0 & 0 & 0 & 0 & 4 & -4 \end{bmatrix}, \quad \text{de modo que } \delta_3 = 5.$$

Como a dimensão do autoespaço generalizado associado ao autovalor 2 é 5, concluímos que o polinômio mínimo do operador representado pela matriz $C$ restrito a este subespaço é $(x-2)^3$. Segue que não há blocos de Jordan de tamanho maior que 3 associados ao autovalor 2 e

$$\begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{bmatrix} = \begin{matrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{matrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{bmatrix} = \begin{matrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{matrix} \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Concluímos que a forma de Jordan da matriz $C$ é

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

$\square$

## 5.4 Base de Jordan

Nas aplicações, muitas vezes é necessário também obter uma **base de Jordan**, isto é, uma base em relação à qual a matriz do operador está na forma de Jordan. Isso deve ser feito separadamente para cada autovalor. O trabalho é facilitado se a forma de Jordan do operador é obtida antecipadamente, através do algoritmo obtido na seção anterior, de modo que sabemos exatamente quais e quantos são os blocos de Jordan de determinado tamanho da forma de Jordan do operador. A partir daí, um procedimento para obter uma base de Jordan para cada bloco de Jordan é encontrar o vetor cíclico que gera o bloco, começando pelo(s) bloco(s) de maior tamanho. Vejamos alguns exemplos:

**Exemplo 9.** Encontre a forma de Jordan e sua respectiva base de Jordan para a matriz

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

O polinômio característico de $A$ é $(x-2)^3$. Temos

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

$$(A-2I)^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Logo $\delta_1 = 2$, $\delta_2 = 3$ e

$$\begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

de modo que a forma de Jordan da matriz $A$ é

$$J = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Uma base para $\ker(A-2I)$ é dada pelos vetores

$$u_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{e} \quad u_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Buscamos um vetor $v_2 \in \ker(A-2I)^2 = \mathbb{R}^3$ tal que $v_2 \notin \ker(A-2I)$, mas $(A-2I)\,v_2 \in \ker(A-2I)$, isto é, tal que $v_2$ não seja um autovetor mas $(A-2I)\,v_2$ é um autovetor. Então $\{v_1 = (A-2I)\,v_2, v_2\}$ será uma base para o bloco de Jordan

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

Resolvendo a equação

$$(A-2I)\,v = a_1 u_1 + a_2 u_2$$

para $v = (x_1, x_2, x_3)$, obtemos

$$\begin{bmatrix} x_2 \\ 0 \\ -x_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ 0 \\ a_2 \end{bmatrix}$$

donde $x_2 = a_1$ e $x_2 = -a_2$, isto é, $a_1 = -a_2$. Ou seja,

$$v = \begin{bmatrix} a \\ b \\ -a \end{bmatrix}$$

é o formato de um vetor $v$ cuja imagem por $A - 2I$ é um autovetor. Escolhendo $a = 0$ e $b = 1$, segue que

$$v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{e} \quad v_1 = (A - 2I)\, v_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

Para completar, escolhemos algum autovetor que seja linearmente independente de $v_1$, por exemplo $u_1$. Portanto, uma base de Jordan para $A$ é dada por

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

Em particular,

$$P = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

é a matriz de mudança de coordenadas, de modo que $J = P^{-1}AP$, isto é,

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

$\square$

**Exemplo 10.** Vimos no Exemplo 6 que a forma de Jordan para a matriz

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

é

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

Uma base para o núcleo de

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

é dada pelos vetores

$$u_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} \quad \text{e} \quad u_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}.$$

Além disso, vimos também que
$$(A-I)^2 = 0.$$
Buscamos vetores linearmente independentes $v_2, v_4 \in \ker(A-I)^2 = \mathbb{R}^4$ tal que $v_2, v_4 \notin \ker(A-I)$, mas $(A-I)\,v_2$, $(A-I)\,v_4 \in \ker(A-I)$, isto é, tal que $v_2, v_4$ não sejam autovetores mas $(A-I)\,v_2$, $(A-I)\,v_4$ são autovetores. Então $\{v_1 = (A-I)\,v_2, v_2\}$ e $\{v_3 = (A-I)\,v_4, v_4\}$ serão bases para os dois blocos de Jordan
$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$
que aparecem na forma de Jordan de $A$. Resolvendo a equação
$$(A-I)\,v = a_1 u_1 + a_2 u_2$$
para $v = (x_1, x_2, x_3, x_4)$, temos
$$\begin{aligned}
\begin{bmatrix} -x_1 - x_2 - 2x_3 - x_4 \\ x_1 + x_2 + x_3 + x_4 \\ 0 \\ x_3 \end{bmatrix} &= \begin{bmatrix} -1 & -1 & -2 & -1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = a_1 \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix} \\
&= \begin{bmatrix} a_1 + a_2 \\ -a_1 \\ 0 \\ -a_2 \end{bmatrix}.
\end{aligned}$$
Escolhendo $a_1 = 1$ e $a_2 = 0$, obtemos
$$\begin{bmatrix} -x_1 - x_2 - 2x_3 - x_4 \\ x_1 + x_2 + x_3 + x_4 \\ 0 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$$
e daí, escolhendo $x_1 = -1$ e $x_4 = 0$ obtemos $x_2 = x_3 = 0$. Segue que
$$v_2 = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{e} \quad v_1 = (A-I)\,v_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}.$$
Escolhendo agora $a_1 = 0$ e $a_2 = 1$, obtemos
$$\begin{bmatrix} -x_1 - x_2 - 2x_3 - x_4 \\ x_1 + x_2 + x_3 + x_4 \\ 0 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$
e daí, escolhendo $x_1 = 1$ e $x_4 = 0$ obtemos $x_2 = 0$ e $x_3 = -1$. Segue que
$$v_4 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} \quad \text{e} \quad v_3 = (A-I)\,v_4 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}.$$

Portanto, uma base de Jordan para $A$ é dada por

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

Em particular,

$$P = \begin{bmatrix} 1 & -1 & 1 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

é a matriz de mudança de coordenadas, de modo que $J = P^{-1}AP$, isto é,

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

□

**Exemplo 11.** Vimos no Exemplo 7 que a forma de Jordan para a matriz

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

é

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

Uma base para o núcleo de

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

é dada pelos vetores

$$u_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} \quad \text{e} \quad u_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}.$$

Uma base para o núcleo de

$$(B - I)^2 = \begin{bmatrix} -1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

é dada pelos vetores

$$u_1, u_2 \quad \text{e} \quad u_3 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}.$$

Além disso, vimos também que
$$(B-I)^3 = 0.$$
Buscamos um vetor linearmente independente $v_3 \in \ker(B-I)^3 = \mathbb{R}^4$ tal que $v_3 \notin \ker(B-I)$ e $v_3 \notin \ker(B-I)^2$ mas $(B-I)\,v_3 \in \ker(B-I)^2$ e $(B-I)^2\,v_3 \in \ker(B-I)$. Então
$$\left\{v_1 = (B-I)^2\,v_3, v_2 = (B-I)\,v_3, v_3\right\}$$
será uma base para o bloco de Jordan
$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$
que aparece na forma de Jordan de $B$. Resolvendo as equações
$$\begin{aligned} (B-I)\,v &= a_1u_1 + a_2u_2 + a_3u_3, \\ (B-I)^2\,v &= a_4u_1 + a_5u_2, \end{aligned}$$
para $v = (x_1, x_2, x_3, x_4)$, temos
$$\begin{aligned} \begin{bmatrix} -x_1 - 2x_2 - x_3 - x_4 \\ x_1 + x_2 + x_3 + x_4 \\ x_2 \\ 0 \end{bmatrix} &= \begin{bmatrix} -1 & -2 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = a_1 \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix} + a_3 \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} a_1 + a_2 + a_3 \\ -a_3 \\ -a_1 \\ -a_2 \end{bmatrix}, \end{aligned}$$
$$\begin{aligned} \begin{bmatrix} -x_1 - x_2 - x_3 - x_4 \\ 0 \\ x_1 + x_2 + x_3 + x_4 \\ 0 \end{bmatrix} &= \begin{bmatrix} -1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = a_4 \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} + a_5 \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} a_4 + a_5 \\ 0 \\ -a_4 \\ -a_5 \end{bmatrix}. \end{aligned}$$
Obtemos as relações $a_2 = a_5 = 0$, $a_4 = -a_3$ e
$$\begin{aligned} x_2 &= -a_1, \\ -x_1 - 2x_2 - x_3 - x_4 &= a_1 + a_3, \\ x_1 + x_2 + x_3 + x_4 &= -a_3, \end{aligned}$$
o que se simplifica para
$$\begin{aligned} x_2 &= -a_1, \\ x_1 + x_3 + x_4 &= a_1 - a_3. \end{aligned}$$
Escolhendo $a_1 = 1$ e $a_3 = 1$, podemos tomar $x_1 = 0, x_3 = 0$ e segue que $x_2 = -1, x_4 = 0$.
$$v_3 = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \quad v_2 = (A-I)\,v_3 = \begin{bmatrix} 2 \\ -1 \\ -1 \\ 0 \end{bmatrix} \quad \text{e} \quad v_1 = (A-I)^2\,v_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}.$$

Finalmente, escolhemos um autovetor $v_4$ linearmente independente de $v_1$:

$$v_4 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}.$$

Portanto, uma base de Jordan para $B$ é dada por

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix} \right\}.$$

Em particular,

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

é a matriz de mudança de coordenadas, de modo que $J = P^{-1}BP$, isto é,

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

$\square$

Um algoritmo pra encontrar uma base de Jordan pode ser descrito em linhas gerais da seguinte forma:

1. Primeiro encontramos uma base $\{v_1^1, \dots, v_{\delta_1}^1\}$ para $\ker(T - \lambda I)$, isto é, vetores linearmente independentes que geram o autoespaço associado ao autovalor $\lambda$.

2. Em seguida, se $\delta_2 > \delta_1$, encontramos uma base $\{V_1^1, \dots, V_{\delta_1}^1\}$ para $\ker(T - \lambda I)$ tal que
$$(T - \lambda I) v_j^2 = V_j^1$$
tem $\delta_2 - \delta_1$ soluções linearmente independentes $v_1^2, \dots, v_{\delta_2-\delta_1}^2$. Então
$$\{V_1^1, \dots, V_{\delta_1}^1\} \cup \{v_1^2, \dots, v_{\delta_2-\delta_1}^2\}$$
é uma base para $\ker(T - \lambda I)^2$.

3. Se $\delta_3 > \delta_2$, encontramos uma base $\{V_1^2, \dots, V_{\delta_2}^2\}$ para $\ker(T - \lambda I)^2$ tal que
$$(T - \lambda I) v_j^3 = V_j^2$$
tem $\delta_3 - \delta_2$ soluções linearmente independentes $v_1^3, \dots, v_{\delta_3-\delta_2}^3$.
Se, para $j = 1, \dots, \delta_2 - \delta_1$, temos $V_j^2 = \sum\limits_{j=1}^{\delta_2-\delta_1} a_i^j v_i^2$, tome $\widetilde{V}_j^1 = \sum\limits_{j=1}^{\delta_2-\delta_1} a_i^j V_i^1$. Para $j = \delta_2 - \delta_1 + 1, \dots, \delta_1$ tome $\widetilde{V}_j^1 = V_j^1$. Então
$$\left\{\widetilde{V}_1^1, \dots, \widetilde{V}_{\delta_1}^1\right\} \cup \{V_1^2, \dots, V_{\delta_2-\delta_1}^2\} \cup \{v_1^3, \dots, v_{\delta_3-\delta_2}^3\}$$
é uma base para $\ker(T - \lambda I)^3$.

4. Continue este processo até obter uma base para $W$.

## 5.5 Forma de Jordan Real

**Teorema 3.** (Forma de Jordan Real) *Seja $T : V \longrightarrow V$ um operador linear real. Então existe uma base para $V$ em relação à qual $T$ é representado por uma matriz diagonal em blocos, com autovalores reais dando origem aos blocos de Jordan usuais e os autovalores complexos dando origem a blocos da forma (chamados blocos de Jordan reais)*

$$\begin{bmatrix} D_{a,b} & I_2 & 0 & \dots & \dots & 0 \\ 0 & D_{a,b} & I_2 & \dots & \dots & 0 \\ 0 & 0 & D_{a,b} & \ddots & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & D_{a,b} & I_2 \\ 0 & 0 & \dots & \dots & 0 & D_{a,b} \end{bmatrix}$$

*onde*

$$D_{a,b} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \quad e \quad I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

*sendo que $a + ib$ é um autovalor complexo de $T_{\mathbb{C}}$. Esta matriz é única a menos da ordem dos blocos.*

**Prova:** Considere o operador complexificado $T_{\mathbb{C}} : V_{\mathbb{C}} \longrightarrow V_{\mathbb{C}}$. Sejam $\lambda_1, \dots, \lambda_k$ os autovalores distintos de $T_{\mathbb{C}}$ e $p = (x - \lambda_1)^{r_1} \dots (x - \lambda_k)^{r_k}$ o polinômio mínimo de $T_{\mathbb{C}}$. Aplicando o Teorema da Decomposição Primária como no Teorema 1, escrevemos $V_{\mathbb{C}} = W_1 \oplus \dots \oplus W_k$ onde $W_i = \ker (T - \lambda_i I)^{r_i}$ e $W_i$ é invariante sob $T_{\mathbb{C}}$. Os autovalores reais de $T_{\mathbb{C}}$ dão origem aos blocos de Jordan usuais para $T$ considerados anteriormente, já que os autoespaços generalizados associados a um autovalor real possuem uma base de vetores reais. Vamos considerar os autoespaços generalizados associados aos autovalores complexos. Seja $\lambda = a + ib$ um autovalor de $T_{\mathbb{C}}$, $b \neq 0$. Pela Proposição 1 (iv), o conjugado $\overline{\lambda} = a - bi$ também é um autovalor de $T_{\mathbb{C}}$. Denotemos os autoespaços generalizados associados a estes dois autovalores por $\widetilde{W}_\lambda$ e $\widetilde{W}_{\overline{\lambda}}$.

Observe que se $\mathcal{B}_\lambda = \{u_1 + iv_1, \dots, u_m + iv_m\}$ é uma base para $\widetilde{W}_\lambda$, então $\mathcal{B}_{\overline{\lambda}} = \{u_1 - iv_1, \dots, u_m - iv_m\}$ é uma base para $\widetilde{W}_{\overline{\lambda}}$. De fato, pelo Teorema Espectral estes autoespaços generalizados possuem a mesma dimensão (já que raízes conjugadas têm a mesma dimensão algébrica). Além disso, $\mathcal{B}_\lambda$ é linearmente independente em $V_{\mathbb{C}}$ se e somente se $\mathcal{B}_{\overline{\lambda}}$ é linearmente independente em $V_{\mathbb{C}}$. De fato, como

$$(a_j - ib_j)(u_j - iv_j) = \overline{(a_j + ib_j)(u_j + iv_j)},$$

segue que

$$\sum_{j=1}^{m} (a_j - ib_j)(u_j - iv_j) = \overline{\sum_{j=1}^{m} (a_j + ib_j)(u_j + iv_j)},$$

de modo que existe uma combinação linear não trivial dos vetores de $\mathcal{B}_\lambda$ produzindo o vetor nulo se e somente se existe uma combinação linear não trivial dos vetores de $\mathcal{B}_{\overline{\lambda}}$ produzindo o vetor nulo. Por fim, observe que $u + iv \in \widetilde{W}_\lambda$ se e somente se $u - iv \in \widetilde{W}_{\overline{\lambda}}$, porque $(T_{\mathbb{C}} - \lambda I)^r (u + iv) = 0$ se e somente se $\overline{(T_{\mathbb{C}} - \lambda I)^r (u + iv)} = \left(T_{\mathbb{C}} - \overline{\lambda} I\right)^r (u - iv) = 0$.

Afirmamos que

$$\mathcal{B} = \{u_1, v_1, \dots, u_m, v_m\}$$

é uma base de vetores reais para a soma direta $\widetilde{W}_\lambda \oplus \widetilde{W}_{\overline{\lambda}}$. Para ver isso, é só observar que este conjunto gera $\widetilde{W}_\lambda \oplus \widetilde{W}_{\overline{\lambda}}$ e tem o número de elementos igual à $\dim \left(\widetilde{W}_\lambda \oplus \widetilde{W}_{\overline{\lambda}}\right) = \dim \widetilde{W}_\lambda + \dim \widetilde{W}_{\overline{\lambda}}$. Segue da Proposição 1 (vi) e (i) que $\widetilde{W}_\lambda \oplus \widetilde{W}_{\overline{\lambda}}$ é a complexificação de algum subespaço $W_{\lambda\overline{\lambda}}$ de $V$ que tem $\mathcal{B}$ como base.

Agora, escolha uma base $\mathcal{B}_\lambda$ tal que os vetores $w_1 = u_1 + iv_1, \dots, w_m = u_m + iv_m$ formam uma cadeia de Jordan para o operador complexo $T_{\mathbb{C}}$, isto é,

$$\begin{aligned} T_{\mathbb{C}} w_1 &= \lambda w_1, \\ T_{\mathbb{C}} w_j &= \lambda w_j + w_{j-1}, \quad \text{para } j = 2, \dots, m. \end{aligned}$$

Então

$$\begin{aligned} Tu_1 + iTv_1 &= T_{\mathbb{C}} (u_1 + iv_1) = T_{\mathbb{C}} w_1 = \lambda w_1 = (a + ib)(u_1 + iv_1) \\ &= (au_1 - bv_1) + i (bu_1 + av_1), \end{aligned}$$

de modo que

$$\begin{aligned} Tu_1 &= au_1 - bv_1, \\ Tv_1 &= bu_1 + av_1, \end{aligned}$$

originando o primeiro bloco

$$D_{a,b} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

no bloco de Jordan real associado ao subespaço $W_{\lambda\overline{\lambda}}$. Além disso,

$$\begin{aligned} Tu_j + iTv_j &= T_{\mathbb{C}} (u_j + iv_j) = T_{\mathbb{C}} w_j = \lambda w_j + w_{j-1} = (a + ib)(u_j + iv_j) + u_{j-1} + iv_{j-1} \\ &= (au_j - bv_j + u_{j-1}) + i (bu_j + av_j + v_{j-1}), \end{aligned}$$

originando os blocos

$$\begin{bmatrix} I_2 \\ D_{a,b} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ a & b \\ -b & a \end{bmatrix}.$$

■

## 5.6 Exercícios

1. Ache a forma de Jordan e uma base de Jordan para as matrizes de (a) até (i). Ache a forma de Jordan real e uma base de Jordan real para as matrizes de (j) até (l).

**(a)** $\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$ **(c)** $\begin{bmatrix} 3 & -3 & -4 \\ 0 & 3 & 5 \\ 0 & 0 & -1 \end{bmatrix}$ **(d)** $\begin{bmatrix} 2 & 1 & 2 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$

**(e)** $\begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 12 & 0 & 3 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}$ **(f)** $\begin{bmatrix} 1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ -2 & -2 & 2 & 1 \\ 1 & 1 & -1 & 0 \end{bmatrix}$ **(g)** $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 1 & 0 & 2 \end{bmatrix}$ **(h)** $\begin{bmatrix} 2 & 1 & 4 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

**(i)** $\begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ **(j)** $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ **(k)** $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix}$ **(l)** $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$

# Capítulo 6

# Espaços com Produto Interno

Neste capítulo, consideraremos apenas espaços vetoriais sobre um corpo $\mathbb{K}$, onde $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

## 6.1 Produto Interno

**Definição 1.** Seja $V$ um espaço vetorial sobre $\mathbb{K}$. Um **produto interno** em $V$ é uma função $\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbb{K}$ que satisfaz as seguintes propriedades:

(i) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ para todos $x, y, z \in V$;

(ii) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$ para todos $x, y \in V$ e para todo $\alpha \in \mathbb{K}$;

(iii) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ para todos $x, y \in V$;

(iv) $\langle x, x \rangle > 0$ para todo $x \neq 0$.

Um espaço vetorial real de dimensão finita dotado de um produto interno é freqüentemente chamado um **espaço euclidiano**. Um espaço vetorial complexo dotado de um produto interno é freqüentemente chamado um **espaço hermitiano** e o seu produto interno é às vezes chamado de **produto hermitiano**.

As condições (ii) e (iv) implicam que

$$\langle x, x \rangle \geqslant 0 \text{ para todo } x \in V \tag{6.1}$$

e

$$\langle x, x \rangle = 0 \text{ se e somente se } x = 0. \tag{6.2}$$

Além disso, as condições (ii) e (iii) implicam

$$\langle x, \alpha y \rangle = \overline{\alpha} \langle x, y \rangle \text{ para todos } x, y \in V \text{ e para todo } \alpha \in \mathbb{K}, \tag{6.3}$$

pois

$$\langle x, \alpha y \rangle = \overline{\langle \alpha y, x \rangle} = \overline{\alpha \langle y, x \rangle} = \overline{\alpha} \overline{\langle y, x \rangle} = \overline{\alpha} \langle x, y \rangle .$$

Em resumo,

$$\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle \text{ para todos } x, y, z \in V \text{ e para todos } \alpha, \beta \in \mathbb{K}, \tag{6.4}$$

enquanto que

$$\langle x, \alpha y + \beta z \rangle = \overline{\alpha} \langle x, y \rangle + \overline{\beta} \langle x, z \rangle \text{ para todos } x, y, z \in V \text{ e para todos } \alpha, \beta \in \mathbb{K}. \tag{6.5}$$

A necessidade de tomar conjugados no caso complexo em (iii) justifica-se para assegurar consistência com (iv). De fato, se valesse simplesmente $\langle x, y\rangle = \langle y, x\rangle$ e (iv), teríamos

$$0 < \langle ix, ix\rangle = i^2 \langle x, x\rangle = -\langle x, x\rangle < 0.$$

A condição (iv) tem prioridade sobre a comutatividade do produto interno, isto é, abdicamos da comutatividade $\langle x, y\rangle = \langle y, x\rangle$ a fim de que (iv) valha, porque é esta última propriedade que nos permite definir uma noção de norma de vetores a partir do produto interno, como veremos na próxima seção.

Observamos que o produto interno complexo é completamente determinado por sua parte real. De fato, se $V$ é um espaço vetorial complexo com produto interno $\langle \cdot, \cdot\rangle$, então

$$\langle x, y\rangle = \operatorname{Re}\langle x, y\rangle + i \operatorname{Im}\langle x, y\rangle .$$

Por outro lado, se $z \in \mathbb{C}$ então $\operatorname{Im} z = \operatorname{Re}(-iz)$, logo

$$\operatorname{Im}\langle x, y\rangle = \operatorname{Re}(-i\langle x, y\rangle) = \operatorname{Re}(\langle x, iy\rangle).$$

Portanto,

$$\langle x, y\rangle = \operatorname{Re}\langle x, y\rangle + i \operatorname{Re}(\langle x, iy\rangle). \tag{6.6}$$

**Exemplo 1.** Definimos um produto interno em $\mathbb{K}^n$ da seguinte forma. Se $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$, então

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y_i}. \tag{6.7}$$

Este é o chamado **produto interno canônico** em $\mathbb{K}^n$. $\square$

**Exemplo 2.** Identificando $\mathbb{R}^n$ com o espaço das matrizes reais $n \times 1$ (matrizes colunas reais), dada uma matriz real $n \times n$ invertível $A$, definimos um produto interno em $\mathbb{R}^n$ por

$$\langle x, y\rangle = y^t \left(A^t A\right) x. \tag{6.8}$$

Note que $A^t A$ é uma matriz simétrica. Quando $A = I$, obtemos o produto interno canônico em $\mathbb{R}^n$. Vamos verificar apenas a propriedade (iii) e (iv) da Definição 1, já que a verificação das demais propriedades é imediata. Como a matriz $A^t A$ é simétrica e a transposta de uma matriz $1 \times 1$ (um número) é ela próprio, temos

$$\langle x, y\rangle = y^t \left(A^t A\right) x = \left[x^t \left(A^t A\right) y\right]^t = x^t \left(A^t A\right) y = \langle y, x\rangle .$$

Em seguida, observando que

$$\left(A^t A\right)_{ij} = \sum_{r=1}^{n} a^t_{ir} a_{rj} = \sum_{r=1}^{n} a_{ri} a_{rj}.$$

Daí,

$$\begin{aligned}
\langle x, x\rangle = x^t \left(A^t A\right) x &= \sum_{i=1}^{n} x_{1i} \left[\left(A^t A\right) x\right]_{i1} = \sum_{i=1}^{n} x_{1i} \left[\sum_{j=1}^{n} \left(A^t A\right)_{ij} x_{j1}\right] = \sum_{i=1}^{n} x_{1i} \left[\sum_{j=1}^{n}\sum_{r=1}^{n} a_{ri} a_{rj} x_{j1}\right] \\
&= \sum_{i,j,r=1}^{n} x_i x_j a_{ri} a_{rj} = \sum_{r=1}^{n} \left(\sum_{i=1}^{n} a_{ri} x_i\right)\left(\sum_{j=1}^{n} a_{rj} x_j\right) = \sum_{r=1}^{n} \left|\sum_{i=1}^{n} a_{ri} x_i\right|^2 \geqslant 0.
\end{aligned}$$

Agora, como $A$ é invertível, se $x \neq 0$ existe pelo menos algum $r$ tal que

$$\sum_{i=1}^{n} a_{ri} x_i \neq 0$$

(este somatório é exatamente o $r$-ésimo elemento da matriz $Ax$) logo $\langle x, x\rangle > 0$. $\square$

**Exemplo 3.** Dada uma matriz complexa $A$ $n \times n$, definimos a sua **transposta conjugada** $A^*$ por

$$(A^*)_{ij} = \overline{A}_{ji}. \tag{6.9}$$

Observe que se $A$ possui apenas entradas reais, sua transposta conjugada coincide com sua transposta. Identificando $\mathbb{C}^n$ com o espaço das matrizes reais $n \times 1$ (matrizes colunas complexas), dada uma matriz complexa $n \times n$ invertível $A$, definimos um produto interno em $\mathbb{C}^n$ por

$$\langle x, y \rangle = y^* A^* A x. \tag{6.10}$$

Note que $A^*A$ é uma matriz **hermitiana**, isto é, uma matriz $B$ que satisfaz $B^* = B$ (matrizes hermitianas reais são matrizes simétricas). Quando $A = I$, obtemos o produto interno canônico em $\mathbb{C}^n$. Observe que

$$(A^*A)_{ij} = \sum_{r=1}^{n} a_{ir}^* a_{rj} = \sum_{r=1}^{n} \overline{a}_{ri} a_{rj},$$

de modo que

$$\begin{aligned}
\langle x, x \rangle &= x^* (A^*A) x = \sum_{i=1}^{n} \overline{x}_{1i} \left[ (A^*A) x \right]_{i1} = \sum_{i=1}^{n} \overline{x}_{1i} \left[ \sum_{j=1}^{n} (A^*A)_{ij} x_{j1} \right] = \sum_{i=1}^{n} \overline{x}_{1i} \left[ \sum_{j=1}^{n} \sum_{r=1}^{n} \overline{a}_{ri} a_{rj} x_{j1} \right] \\
&= \sum_{i,j,r=1}^{n} \overline{x}_i x_j \overline{a}_{ri} a_{rj} = \sum_{r=1}^{n} \left( \sum_{i=1}^{n} \overline{a_{ri} x_i} \right) \left( \sum_{j=1}^{n} a_{rj} x_j \right) = \sum_{r=1}^{n} \overline{\left( \sum_{i=1}^{n} a_{ri} x_i \right)} \left( \sum_{j=1}^{n} a_{rj} x_j \right) \\
&= \sum_{r=1}^{n} \left| \sum_{i=1}^{n} a_{ri} x_i \right|^2 > 0
\end{aligned}$$

porque $A$ é invertível, como argumentado no exemplo anterior. □

**Exemplo 4.** Mais geralmente, se $V$ é um espaço vetorial e $W$ é um espaço vetorial com produto interno, ambos sobre o mesmo corpo $\mathbb{K}$, se $T : V \longrightarrow V$ é um isomorfismo, definimos um produto interno em $V$ a partir do produto interno em $W$ por

$$\langle x, y \rangle_V := \langle Tx, Ty \rangle_W. \tag{6.11}$$

Dizemos que $\langle \cdot, \cdot \rangle_V$ é o produto interno em $V$ **induzido** pelo produto interno em $W$ através do isomorfismo $T$. Na verdade, é suficiente que $T$ seja uma aplicação linear injetiva para esta definição fazer sentido (pois $T$ é um isomorfismo de $V$ sobre a sua imagem). □

**Exemplo 5.** Definimos um produto interno em $M_n(\mathbb{K})$ por

$$\langle A, B \rangle = \sum_{i,j=1}^{n} a_{ij} \overline{b}_{ij}. \tag{6.12}$$

Usando a transposta conjugada, este produto interno pode ser definido em termos da função traço:

$$\langle A, B \rangle = \operatorname{tr}(AB^*). \tag{6.13}$$

No caso em que $\mathbb{K} = \mathbb{R}$, temos

$$\langle A, B \rangle = \operatorname{tr}\left(AB^t\right).$$

□

**Exemplo 6.** Se $C([0,1];\mathbb{K})$ denota o espaço das funções contínuas no intervalo $[0,1]$ com valores em $\mathbb{K}$, definimos um produto interno neste espaço de dimensão infinita por

$$\langle f,g\rangle = \int_0^1 f(t)\,\overline{g(t)}\,dt. \tag{6.14}$$

□

**Proposição 1.** *Seja $V$ um espaço vetorial sobre $\mathbb{K}$ de dimensão $n$. Então todo produto interno em $V$ é determinado por uma matriz hermitiana. Mais especificamente, fixe uma base de $V$ e escreva os vetores em coordenadas com relação a esta base. Se $\langle\cdot,\cdot\rangle$ é um produto interno em $V$, então existe uma matriz hermitiana $H$ sobre $\mathbb{K}$ tal que*

$$\langle x,y\rangle = y^*Hx, \tag{6.15}$$

*$H$ é invertível e satisfaz $x^*Hx > 0$ para todo $x \in V$. Reciprocamente, se $H$ é uma matriz hermitiana invertível que satisfaz $x^*Hx > 0$ para todo $x \in V$, então a equação acima define um produto interno em $V$.*

**Prova:** Seja $\mathcal{B} = \{x_1,\dots,x_n\}$ uma base para $V$. Então afirmamos que

$$H_{ij} = \langle x_j, x_i\rangle.$$

Com efeito, se $x = \sum_{i=1}^n a_i x_i$ e $y = \sum_{j=1}^n b_j x_j$, temos

$$\begin{aligned}\langle x,y\rangle &= \left\langle \sum_{i=1}^n a_i x_i, y\right\rangle = \sum_{i=1}^n a_i \langle x_i, y\rangle = \sum_{i=1}^n a_i \left\langle x_i, \sum_{j=1}^n b_j x_j\right\rangle = \sum_{i=1}^n\sum_{j=1}^n a_i \overline{b}_j \langle x_i, x_j\rangle \\ &= \sum_{j=1}^n\sum_{i=1}^n \overline{b}_j \langle x_i, x_j\rangle a_i = \sum_{j=1}^n \overline{b}_j \sum_{i=1}^n H_{ji} a_{i1} = \sum_{j=1}^n \overline{b}_{1j} (Hx)_{j1} = y^*Hx.\end{aligned}$$

Como $x^*Hx = \langle x,x\rangle > 0$ para todo $x \neq 0$, em particular $\ker H = \{0\}$ e portanto $H$ é invertível.

Reciprocamente, se $H$ é uma matriz hermitiana invertível que satisfaz $x^*Hx > 0$ para todo $x \in V$ e definimos

$$\langle x,y\rangle = y^*Hx,$$

este é um produto interno em $V$. É fácil ver que as propriedades (i), (ii) e (iv) da Definição 1 são válidas. Para verificar (iii), observe que

$$\langle x,y\rangle = y^*(A^*A)x = [x^*(A^*A)y]^* = \overline{x^*(A^*A)y} = \overline{\langle y,x\rangle}.$$

■

Em particular, se $V$ é um espaço vetorial real,

$$\langle x,y\rangle = y^tAx$$

para alguma matriz real simétrica $A$.

## 6.2 Norma

**Definição 2.** Seja $V$ um espaço vetorial sobre $\mathbb{K}$. Uma **norma** em $V$ é uma função $\|\cdot\| : V \longrightarrow [0,+\infty)$ que satisfaz as seguintes propriedades:

(i) $\|x\| > 0$ para todo $x \neq 0$;

(ii) $\|\alpha x\| = |\alpha| \|x\|$ para todo $x \in V$ e para todo $\alpha \in \mathbb{K}$;

(iii) (Desigualdade Triangular) $\|x + y\| \leqslant \|x\| + \|y\|$ para todos $x, y \in V$.

Um espaço vetorial dotado de uma norma é chamado um **espaço normado**.

Uma norma pode ser definida a partir de um produto interno, enquanto que um produto interno pode ser definido a partir de uma norma somente se esta satisfaz a identidade do paralelogramo, como veremos a seguir.

**Definição 3.** Seja $V$ um espaço vetorial com produto interno. Se $\langle x, y\rangle = 0$, dizemos que $x$ e $y$ são **vetores ortogonais**.

**Proposição 2.** (Desigualdade de Cauchy-Schwarz) *Seja $V$ um espaço vetorial com produto interno $\langle \cdot, \cdot \rangle$ e $\|\cdot\|$ a norma derivada deste produto interno. Então*

$$|\langle x, y\rangle| \leqslant \|x\| \|y\| \tag{6.16}$$

*para todos $x, y \in V$.*

**Prova:** Se $x = \alpha y$, então

$$|\langle x, y\rangle| = |\langle \alpha y, y\rangle| = |\alpha \langle y, y\rangle| = |\alpha| \|y\|^2 = |\alpha| \|y\| \|y\| = \|x\| \|y\|,$$

ou seja, a igualdade na desigualdade de Cauchy-Schwarz é atingida quando um dos vetores é múltiplo escalar do outro.

Se $x$ não é múltiplo escalar de $y$, em particular $x \neq 0$ e podemos tomar a *componente ortogonal* de $y$ à direção $x$, isto é, o vetor

$$z = y - \frac{\langle y, x\rangle}{\|x\|^2} x$$

(o vetor $\dfrac{\langle y, x\rangle}{\|x\|^2} x$ é a *componente tangencial* de $y$ na direção $x$). De fato, temos

$$\left\langle y - \frac{\langle y, x\rangle}{\|x\|^2} x, x \right\rangle = \langle y, x\rangle - \left\langle \frac{\langle y, x\rangle}{\|x\|^2} x, x \right\rangle = \langle y, x\rangle - \frac{\langle y, x\rangle}{\|x\|^2} \langle x, x\rangle = 0.$$

Daí, segue que

$$\begin{aligned}
0 \leqslant \|z\|^2 &= \left\langle y - \frac{\langle y, x\rangle}{\|x\|^2} x, y - \frac{\langle y, x\rangle}{\|x\|^2} x \right\rangle = \left\langle y - \frac{\langle y, x\rangle}{\|x\|^2} x, y \right\rangle \\
&= \|y\|^2 - \frac{\langle y, x\rangle}{\|x\|^2} \langle x, y\rangle = \|y\|^2 - \frac{|\langle x, y\rangle|^2}{\|x\|^2},
\end{aligned}$$

donde o resultado desejado. ■

**Proposição 3.** *Seja $V$ um espaço vetorial com produto interno. Então*

$$\|x\| = \sqrt{\langle x, x\rangle} \tag{6.17}$$

*define uma norma em $V$.*

**Prova:** A condição (i) da Definição 2 decorre imediatamente da condição (iv) da Definição 1. A condição (ii) da Definição 2 decorre de

$$\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x\rangle} = \sqrt{\alpha\overline{\alpha} \langle x, x\rangle} = \sqrt{\alpha^2 \langle x, x\rangle} = |\alpha| \sqrt{\langle x, x\rangle} = |\alpha| \|x\| .$$

Finalmente, a desigualdade triangular é provada usando a desigualdade de Cauchy-Schwarz (deduzida na Proposição 2 a seguir):

$$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y\rangle = \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle = \langle x,x\rangle + \langle x,y\rangle + \overline{\langle x,y\rangle} + \langle y,y\rangle \\ &= \|x\|^2 + 2\operatorname{Re}\langle x,y\rangle + \|y\|^2 \\ &\leqslant \|x\|^2 + 2\operatorname{Re}|\langle x,y\rangle| + \|y\|^2 \\ &= \|x\|^2 + 2|\langle x,y\rangle| + \|y\|^2 \\ &\leqslant \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2 .\end{aligned}$$

■

A norma definida na Proposição 1 é chamada a **norma derivada do produto interno** ou *norma induzida pelo produto interno.*

Em particular, se $V$ é um espaço vetorial real, segue da desigualdade de Cauchy-Schwarz que

$$-1 \leqslant \frac{\langle x, y\rangle}{\|x\| \|y\|} \leqslant 1.$$

**Definição 4.** Seja $V$ um espaço vetorial real. Dados dois vetores $x, y \in V$ definimos o seu ângulo $\measuredangle(x, y)$ por

$$\measuredangle(x, y) = \arccos \frac{\langle x, y\rangle}{\|x\| \|y\|}.$$

Em particular, se $x, y$ são vetores ortogonais, então$\measuredangle(x, y) = \pi/2$.

**Proposição 4.** (Teorema de Pitágoras) *Seja $V$ um espaço vetorial com produto interno $\langle\cdot,\cdot\rangle$ e $\|\cdot\|$ a norma derivada deste produto interno. Se $x, y \in V$ são vetores ortogonais, temos*

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 . \tag{6.18}$$

*Reciprocamente, se $V$ é um espaço vetorial real e $x, y \in V$ são vetores que satisfazem a identidade de Pitágoras $\|x+y\|^2 = \|x\|^2 + \|y\|^2$, então $x, y$ são vetores ortogonais.*

**Prova:** Temos

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle = \langle x,x\rangle + \langle y,y\rangle = \|x\|^2 + \|y\|^2 .$$

Se $V$ é um espaço vetorial real, então

$$\|x+y\|^2 = \langle x,x\rangle + 2\langle x,y\rangle + \langle y,y\rangle = \|x\|^2 + 2\langle x,y\rangle + \|y\|^2 .$$

Logo, se $x, y$ satisfazem a identidade de Pitágoras, necessariamente $\langle x, y\rangle = 0$. ■

**Proposição 5.** (Identidades Polares) *Seja $V$ um espaço vetorial com produto interno $\langle\cdot,\cdot\rangle$ e $\|\cdot\|$ a norma derivada deste produto interno. Se $V$ é um espaço vetorial real, então*

$$\langle x, y\rangle = \frac{1}{4}\|x+y\|^2 - \frac{1}{4}\|x-y\|^2 . \tag{6.19}$$

*Se $V$ é um espaço vetorial complexo, então*

$$\langle x, y\rangle = \frac{1}{4}\|x+y\|^2 - \frac{1}{4}\|x-y\|^2 + \frac{i}{4}\|x+iy\|^2 - \frac{i}{4}\|x-iy\|^2 . \tag{6.20}$$

**Prova:** Se $V$ é um espaço vetorial real, então

$$\frac{1}{4}\left\|x+y\right\|^2-\frac{1}{4}\left\|x-y\right\|^2=\frac{1}{4}\left(\langle x,x\rangle+2\langle x,y\rangle+\langle y,y\rangle\right)-\frac{1}{4}\left(\langle x,x\rangle+2\langle x,y\rangle+\langle y,y\rangle\right)=\langle x,y\rangle .$$

Se $V$ é um espaço vetorial complexo, então

$$\begin{aligned}
&\frac{1}{4}\left\|x+y\right\|^2-\frac{1}{4}\left\|x-y\right\|^2+\frac{i}{4}\left\|x+iy\right\|^2-\frac{i}{4}\left\|x-iy\right\|^2\\
&=\frac{1}{4}\left(\langle x,x\rangle+\langle x,y\rangle+\langle y,x\rangle+\langle y,y\rangle\right)-\frac{1}{4}\left(\langle x,x\rangle-\langle x,y\rangle-\langle y,x\rangle+\langle y,y\rangle\right)\\
&+\frac{i}{4}\left(\langle x,x\rangle-i\langle x,y\rangle+i\langle y,x\rangle+\langle y,y\rangle\right)-\frac{i}{4}\left(\langle x,x\rangle+i\langle x,y\rangle-i\langle y,x\rangle+\langle y,y\rangle\right)\\
&=\frac{1}{2}\langle x,y\rangle+\frac{1}{2}\langle y,x\rangle+\frac{1}{4}\langle x,y\rangle-\frac{1}{4}\langle y,x\rangle+\frac{1}{4}\langle x,y\rangle-\frac{1}{4}\langle y,x\rangle\\
&=\langle x,y\rangle .
\end{aligned}$$

■

**Proposição 6.** (Identidade do Paralelogramo) *Seja $V$ um espaço vetorial com produto interno $\langle\cdot,\cdot\rangle$ e $\|\cdot\|$ a norma derivada deste produto interno. Então*

$$\left\|x+y\right\|^2+\left\|x-y\right\|^2=2\left(\left\|x\right\|^2+\left\|y\right\|^2\right). \tag{6.21}$$

**Prova:** Temos

$$\begin{aligned}
\left\|x+y\right\|^2+\left\|x-y\right\|^2&=\left(\langle x,x\rangle+\langle x,y\rangle+\langle y,x\rangle+\langle y,y\rangle\right)+\left(\langle x,x\rangle-\langle x,y\rangle-\langle y,x\rangle+\langle y,y\rangle\right)\\
&=2\left(\left\|x\right\|^2+\left\|y\right\|^2\right).
\end{aligned}$$

■

**Proposição 7.** *Seja $V$ um espaço vetorial normado, cuja norma $\|\cdot\|$ satisfaz a identidade do paralelogramo*

$$\left\|x+y\right\|^2+\left\|x-y\right\|^2=2\left(\left\|x\right\|^2+\left\|y\right\|^2\right).$$

*Se $V$ é um espaço vetorial real, então a identidade polar*

$$\langle x,y\rangle:=\frac{1}{4}\left\|x+y\right\|^2-\frac{1}{4}\left\|x-y\right\|^2$$

*define um produto interno $\langle\cdot,\cdot\rangle$ em $V$ tal que a sua norma é derivada dele.*

*Se $V$ é um espaço vetorial complexo, então a identidade polar*

$$\langle x,y\rangle:=\frac{1}{4}\left\|x+y\right\|^2-\frac{1}{4}\left\|x-y\right\|^2+\frac{i}{4}\left\|x+iy\right\|^2-\frac{i}{4}\left\|x-iy\right\|^2=\frac{1}{4}\sum_{n=1}^{4}i^n\left\|x+i^ny\right\|^2$$

*define um produto interno $\langle\cdot,\cdot\rangle$ em $V$ tal que a sua norma é derivada dele.*

**Prova:** Seja $V$ um espaço vetorial real normado onde vale a identidade do paralelogramo. Vamos verificar que

$$\langle x,y\rangle:=\frac{1}{4}\left\|x+y\right\|^2-\frac{1}{4}\left\|x-y\right\|^2$$

satisfaz todas as condições da Definição 1 para ser um produto interno real em $V$.

(i) Temos

$$\begin{aligned}
\langle x,z\rangle+\langle y,z\rangle &= \frac{1}{4}\left\|x+z\right\|^2-\frac{1}{4}\left\|x-z\right\|^2+\frac{1}{4}\left\|y+z\right\|^2-\frac{1}{4}\left\|y-z\right\|^2 \\
&= \frac{1}{4}\left(\left\|x+z\right\|^2+\left\|y+z\right\|^2\right)-\frac{1}{4}\left(\left\|x-z\right\|^2+\left\|y-z\right\|^2\right) \\
&= \frac{1}{8}\left(\left\|x+z+y+z\right\|^2+\left\|x+z-(y+z)\right\|^2\right) \\
&\quad -\frac{1}{8}\left(\left\|x-z+y-z\right\|^2+\left\|x-z-(y-z)\right\|^2\right) \\
&= \frac{1}{8}\left(\left\|x+z+y+z\right\|^2+\left\|x-y\right\|^2\right)-\frac{1}{8}\left(\left\|x-z+y-z\right\|^2+\left\|x-y\right\|^2\right) \\
&= \frac{1}{8}\left(\left\|x+z+y+z\right\|^2\right)-\frac{1}{8}\left(\left\|x-z+y-z\right\|^2\right) \\
&= \frac{1}{8}\left(2\left\|x+y+z\right\|^2+2\left\|z\right\|^2-\left\|(x+y+z)-z\right\|^2\right) \\
&\quad -\frac{1}{8}\left(2\left\|x+y-z\right\|^2+2\left\|z\right\|^2-\left\|(x+y-z)+z\right\|^2\right) \\
&= \frac{1}{8}\left(2\left\|x+y+z\right\|^2-\left\|x+y\right\|^2\right)-\frac{1}{8}\left(2\left\|x+y-z\right\|^2-\left\|x+y\right\|^2\right) \\
&= \frac{1}{4}\left\|x+y+z\right\|^2-\frac{1}{4}\left\|x+y-z\right\|^2 \\
&= \langle x+y,z\rangle\,.
\end{aligned}$$

(ii) Se $\alpha=n\in\mathbb{N}$, por iteração de (i) obtemos

$$\langle nx,y\rangle=n\,\langle x,y\rangle\,;$$

por exemplo, para $n=2$ temos

$$\langle 2x,y\rangle=\langle x+x,y\rangle=\langle x,y\rangle+\langle x,y\rangle=2\,\langle x,y\rangle\,.$$

Se $n=-1$, notando que

$$\langle 0,y\rangle=\frac{1}{4}\left\|0+y\right\|^2-\frac{1}{4}\left\|0-y\right\|^2=\frac{1}{4}\left\|y\right\|^2-\frac{1}{4}\left\|-y\right\|^2=\frac{1}{4}\left\|y\right\|^2-\frac{1}{4}\left\|y\right\|^2=0,$$

escrevemos

$$0=\langle 0,y\rangle=\langle x-x,y\rangle=\langle x,y\rangle+\langle -x,y\rangle\,,$$

de modo que

$$\langle -x,y\rangle=-\,\langle x,y\rangle\,.$$

Daí, se $n\in\mathbb{N}$,

$$\langle -nx,y\rangle=\langle n\,(-x)\,,y\rangle=n\,\langle -x,y\rangle=(-1)\,n\,\langle x,y\rangle=-n\,\langle x,y\rangle\,.$$

Portanto, $\langle \alpha x,y\rangle=\alpha\,\langle x,y\rangle$ para todo $\alpha\in\mathbb{Z}$. Em seguida, para provar que

$$\left\langle \frac{1}{n}x,y\right\rangle=\frac{1}{n}\,\langle x,y\rangle$$

para todo $n\in\mathbb{N}$, notamos que

$$\langle x,y\rangle=\left\langle \sum\frac{1}{n}x,y\right\rangle=\sum\left\langle \frac{1}{n}x,y\right\rangle=n\left\langle \frac{1}{n}x,y\right\rangle.$$

Reunindo os dois resultados, concluímos que $\langle \alpha x, y\rangle = \alpha \langle x, y\rangle$ para todo $\alpha \in \mathbb{Q}$. Para obter o resultado geral para qualquer $\alpha \in \mathbb{R}$, observe que a função norma é contínua em um espaço normado e como o produto interno foi definido a partir da norma, ele também é uma função contínua. Assim, dado qualquer $\alpha \in \mathbb{R}$, tomamos uma seqüência $(\alpha_n) \subset \mathbb{Q}$ tal que $\alpha_n \to \alpha$ e obtemos

$$\begin{array}{ccc} \langle \alpha_n x, y\rangle & = & \alpha_n \langle x, y\rangle \\ \downarrow & & \downarrow \\ \langle \alpha x, y\rangle & & \alpha \langle x, y\rangle \end{array}$$

(iii) Temos

$$\langle x, y\rangle = \frac{1}{4}\|x+y\|^2 - \frac{1}{4}\|x-y\|^2 = \frac{1}{4}\|y+x\|^2 - \frac{1}{4}\|y-x\|^2 = \langle y, x\rangle .$$

(iv) Se $x \neq 0$, temos

$$\langle x, x\rangle = \frac{1}{4}\|x+x\|^2 - \frac{1}{4}\|x-x\|^2 = \frac{1}{4}\|2x\|^2 = \|x\|^2 > 0.$$

O caso complexo fica como exercício para o aluno. ■

## 6.3 Bases Ortonormais e Projeções Ortogonais

**Definição 5.** Seja $V$ um espaço vetorial com produto interno. Dizemos que um conjunto $S \subset V$ é **ortogonal** se os elementos de $S$ são dois a dois ortogonais. Um conjunto ortogonal cujos vetores têm todos norma 1 é chamado um conjunto **ortonormal**.

O vetor nulo é o único vetor ortogonal a todos os vetores de $V$.

**Proposição 8.** *Se um vetor $y$ é a combinação linear dos vetores ortogonais não-nulos $x_1, \ldots, x_m$, então*

$$y = \sum_{k=1}^{m} \frac{\langle y, x_k\rangle}{\|x_k\|^2} x_k. \tag{6.22}$$

*Em particular, se os vetores $x_1, \ldots, x_m$ são ortonormais, segue que*

$$y = \sum_{k=1}^{m} \langle y, x_k\rangle x_k. \tag{6.23}$$

**Prova:** Se $y = \sum_{j=1}^{m} \alpha_j x_j$, então

$$\langle y, x_k\rangle = \left\langle \sum_{j=1}^{m} \alpha_j x_j, x_k \right\rangle = \sum_{j=1}^{m} \alpha_j \langle x_j, x_k\rangle = \alpha_k \|x_k\|^2 .$$

■

**Corolário 1.** *Qualquer conjunto ortogonal de vetores não-nulos é linearmente independente.*

**Prova:** Se $x_1, \ldots, x_m$ são vetores ortogonais não-nulos e $0 = \sum_{j=1}^{m} \alpha_j x_j$, segue da proposição anterior que $\alpha_j = \dfrac{\langle 0, x_k\rangle}{\|x_k\|^2} = 0$. ■

**Teorema 1.** (Processo de Ortogonalização de Gram-Schmidt) *Todo espaço vetorial com produto interno possui uma base ortonormal.*

**Prova:** Seja $\mathcal{B}' = \{y_1, \ldots, y_n\}$ uma base para um espaço vetorial com produto interno $V$. Construiremos uma base ortogonal $\mathcal{B} = \{x_1, \ldots, x_n\}$ a partir de $\mathcal{B}'$ usando o algoritmo chamado *processo de ortogonalização de Gram-Schmidt*. Para obter uma base ortonormal a partir da base ortogonal $\mathcal{B}$ basta dividir cada vetor de $\mathcal{B}$ pela sua norma.

Primeiro tomamos $x_1 = y_1$. Indutivamente, suponha obtidos vetores ortogonais $x_1, \ldots, x_m$ tais que

$$\{x_1, \ldots, x_k\}$$

é uma base para o subespaço gerado pelos vetores $y_1, \ldots, y_k$, $1 \leqslant k \leqslant m$. Para construir o vetor $x_{m+1}$, consideremos a *projeção ortogonal* do vetor $y_{m+1}$ sobre o subespaço gerado pelos vetores $x_1, \ldots, x_m$:

$$\sum_{j=1}^{m} \frac{\langle y_{m+1}, x_j \rangle}{\|x_j\|^2} x_j.$$

Tomamos

$$x_{m+1} = y_{m+1} - \sum_{j=1}^{m} \frac{\langle y_{m+1}, x_j \rangle}{\|x_j\|^2} x_j.$$

Então $x_{m+1} \neq 0$, caso contrário $x_{m+1}$ está no subespaço gerado por $y_1, \ldots, y_m$ e portanto é uma combinação linear destes vetores. Além disso, para todo $1 \leqslant k \leqslant m$ temos

$$\langle x_{m+1}, x_k \rangle = \langle y_{m+1}, x_k \rangle - \sum_{j=1}^{m} \frac{\langle y_{m+1}, x_j \rangle}{\|x_j\|^2} \langle x_j, x_k \rangle = \langle y_{m+1}, x_k \rangle - \frac{\langle y_{m+1}, x_k \rangle}{\|x_k\|^2} \langle x_k, x_k \rangle = 0.$$

Assim, $\{x_1, \ldots, x_{m+1}\}$ é um conjunto de $m+1$ vetores ortogonais que geram o subespaço $\langle y_1, \ldots, y_{m+1} \rangle$ de dimensão $m+1$, logo é uma base para este. ■

A maior vantagem em se usar bases ortonormais é a facilidade em trabalhar com coordenadas em relação a estas bases: como vimos na Proposição 8, para obter as coordenadas de um vetor em relação a uma base ortonormal basta calcular os produtos internos dele em relação aos vetores da base.

**Proposição 9.** *Seja $V$ um espaço vetorial com produto interno de dimensão finita e $T : V \longrightarrow V$ um operador linear. Se $\mathcal{B} = \{x_1, \ldots, x_n\}$ é uma base ortonormal para $V$ e $A = [T]_{\mathcal{B}}$, então*

$$a_{ij} = \langle Tx_j, x_i \rangle .$$

**Prova:** Por definição,

$$Tx_j = \sum_{k=1}^{n} a_{kj} x_k.$$

Logo,

$$\langle Tx_j, x_i \rangle = \left\langle \sum_{k=1}^{n} a_{kj} x_k, x_i \right\rangle = \sum_{k=1}^{n} a_{kj} \langle x_k, x_i \rangle = a_{ij}.$$

■

**Definição 6.** Seja $V$ um espaço vetorial com produto interno e $W \subset V$ um subespaço de $V$. O **complemento ortogonal** de $W$ é o subespaço

$$W^{\perp} = \{y \in V : \langle y, x \rangle = 0 \text{ para todo } x \in W\} .$$

A linearidade do produto interno com relação à primeira variável garante que $W^{\perp}$ é um subespaço vetorial. Observe que o complemento ortogonal de qualquer subconjunto $S$ de $V$ é um subespaço vetorial de $V$, mesmo se o próprio $S$ não for.

**Proposição 10.** *Para qualquer subespaço $W \subset V$ temos*

$$V = W \oplus W^{\perp}.$$

**Prova:** Seja $\{x_1, \ldots, x_m\}$ uma base ortogonal para $W$. Dado $x \in V$ defina

$$y = \sum_{k=1}^{m} \frac{\langle x, x_k \rangle}{\|x_k\|^2} x_k \in W.$$

Então $z = x - y \in W^{\perp}$ porque $\langle x - y, x_j \rangle = 0$ para todo $j$:

$$\langle x - y, x_j \rangle = \left\langle x - \sum_{k=1}^{m} \frac{\langle x, x_k \rangle}{\|x_k\|^2} x_k, x_j \right\rangle = \langle x, x_j \rangle - \sum_{k=1}^{m} \frac{\langle x, x_k \rangle}{\|x_k\|^2} \langle x_k, x_j \rangle = \langle x, x_j \rangle - \frac{\langle x, x_j \rangle}{\|x_j\|^2} \langle x_j, x_j \rangle = 0.$$

Logo, $x = y + z$ com $y \in W$ e $z \in W^{\perp}$. Além disso, se $x \in W \cap W^{\perp}$, então $\langle x, x \rangle = 0$, logo $x = 0$. ■

**Definição 7.** Seja $V$ um espaço vetorial com produto interno. Dizemos que uma projeção $E : V \longrightarrow V$ é uma projeção ortogonal se

$$\ker E \perp \operatorname{im} E.$$

Na decomposição em soma direta

$$\begin{aligned} V &= \operatorname{im} E \oplus^{\perp} \ker E \\ x &= y + z, \end{aligned}$$

o vetor $y$ é chamado a **projeção ortogonal** de $x$ no subespaço $W := \operatorname{im} E$.

**Corolário 2.** *Seja $V$ um espaço vetorial com produto interno, $W \subset V$ um subespaço de $V$ e $\{x_1, \ldots, x_m\}$ uma base ortogonal para $W$. Dado $x \in V$, a projeção ortogonal de $x$ em $W$ é o vetor*

$$y = \sum_{k=1}^{m} \frac{\langle x, x_k \rangle}{\|x_k\|^2} x_k. \tag{6.24}$$

**Definição 8.** Seja $V$ um espaço vetorial com produto interno e $W \subset V$ um subespaço de $V$. Dado $x \in V$, a **melhor aproximação** de $x$ por vetores de $W$ é o vetor $y$ que satisfaz

$$\|x - y\| \leqslant \|x - z\|$$

para todos os vetores $z \in W$.

Em outras palavras,

$$\|x - y\| = \inf_{z \in W} \|x - z\|.$$

**Teorema 2.** *Seja $V$ um espaço vetorial com produto interno e $W \subset V$ um subespaço de $V$. Dado $x \in V$, o vetor em $W$ que melhor aproxima $x$ é a projeção ortogonal de $x$ em $W$.*

**Prova:** Escreva

$$x = y + y^{\perp}$$

onde $y$ é a projeção ortogonal de $x$ em $W$ e $y^{\perp} \in W^{\perp}$. Dado $z \in W$, temos

$$x - z = y - z + y^{\perp},$$

e como $x - z$ é ortogonal ao vetor $y^{\perp}$, segue da identidade de Pitágoras que

$$\|x - z\|^2 = \|y - z\|^2 + \left\|y^{\perp}\right\|^2.$$

Em particular, o valor mínimo para $\|x - z\|$ é atingido quando $\|y - z\| = 0$, isto é, quando $z = y$. ■

## 6.4 Operador Adjunto

Em um espaço vetorial com produto interno de dimensão finita, todo funcional linear é derivado do produto interno:

**Teorema 3.** (Teorema da Representação de Riesz) *Seja $V$ um espaço vetorial com produto interno de dimensão finita e $f \in V^*$ um funcional linear. Então existe um único vetor $x \in V$ tal que*

$$f(y) = \langle y, x \rangle \quad \text{para todo } y \in V.$$

*Esta correspondência determina um isomorfismo canônico entre $V$ e $V^*$.*

**Prova:** Seja $\{x_1, \ldots, x_m\}$ uma base ortonormal para $V$. Se $y \in V$, então

$$y = \sum_{k=1}^{m} \langle y, x_k \rangle x_k.$$

Logo,

$$f(y) = \sum_{k=1}^{m} \langle y, x_k \rangle f(x_k) = \sum_{k=1}^{m} \left\langle y, \overline{f(x_k)} x_k \right\rangle = \left\langle y, \sum_{k=1}^{m} \overline{f(x_k)} x_k \right\rangle.$$

Tome

$$x = \sum_{k=1}^{m} \overline{f(x_k)} x_k.$$

Se $x' \in V$ é outro vetor tal que $f(y) = \langle y, x' \rangle$ para todo $y \in V$, então $\langle y, x \rangle = \langle y, x' \rangle$ para todo $y \in V$, donde $\langle y, x - x' \rangle = 0$. Tomando $y = x - x'$, concluímos que $x - x' = 0$. ■

Observe que o vetor $x$ do Teorema da Representação de Riesz está no complemento ortogonal de $\ker f$ ou, mais precisamente, $\langle x \rangle = \ker f^{\perp}$.

**Definição 9.** Sejam $V, W$ espaços vetoriais com produto interno e $T : V \longrightarrow W$ uma aplicação. Dizemos que uma aplicação $T^* : W \longrightarrow V$ é a **adjunta** de $T$ se

$$\langle Tx, y \rangle = \langle x, T^* y \rangle \quad \text{para todos } x \in V, y \in W.$$

Observe que também temos

$$\langle x, Ty \rangle = \langle T^* x, y \rangle,$$

pois

$$\langle x, Ty \rangle = \overline{\langle Ty, x \rangle} = \overline{\langle y, T^* x \rangle} = \langle T^* x, y \rangle.$$

**Lema 1.** *Sejam $V, W$ espaços vetoriais com produto interno e $T : V \longrightarrow W$ uma aplicação linear. Se a adjunta existir, ela é única e linear.*

**Prova:** Sejam $y_1, y_2 \in W$ e $\alpha, \beta \in \mathbb{K}$. Então, para todo $x \in V$ temos

$$\begin{aligned} \langle x, T^* (\alpha y_1 + \beta y_2) \rangle &= \langle Tx, \alpha y_1 + \beta y_2 \rangle = \overline{\alpha} \langle Tx, y_1 \rangle + \overline{\beta} \langle Tx, y_2 \rangle = \overline{\alpha} \langle x, T^* y_1 \rangle + \overline{\beta} \langle x, T^* y_2 \rangle \\ &= \langle x, \alpha T^* y_1 + \beta T^* y_2 \rangle, \end{aligned}$$

o que implica, como no argumento para provar a unicidade usado na demonstração do teorema anterior,

$$T^* (\alpha y_1 + \beta y_2) = \alpha T^* y_1 + \beta T^* y_2.$$

O mesmo argumento também estabelece a unicidade de $T^*$. ■

Observe que na demonstração do Lema 1 não usamos o fato de $T$ ser linear para provar que a adjunta $T^*$ é linear. Segue que, quando existe, a adjunta de qualquer aplicação é necessariamente linear. Por outro lado, como a adjunta da adjunta de $T$ é a própria aplicação $T$, porque

$$\langle T^* x, y \rangle = \langle x, Ty \rangle ,$$

como vimos logo acima, concluímos que $T$ (adjunta de uma aplicação) deve ser linear. Em outras palavras, para que a adjunta de uma aplicação $T$ exista, $T$ já deve ser uma aplicação linear. Assim, não há realmente nenhum ganho em generalidade em definir a adjunta de uma aplicação arbitrária ao invés de definir apenas a adjunta de aplicações lineares, pois as únicas aplicações que possuem adjuntas são as aplicações lineares.

**Teorema 4.** *Sejam $V, W$ espaços vetoriais com produto interno de dimensão finita. Então toda aplicação linear $T : V \longrightarrow W$ possui uma única adjunta linear.*

**Prova:** Para cada $y \in W$, a aplicação $x \mapsto \langle Tx, y \rangle$ é um funcional linear em $V^*$. Pelo Teorema de Representação de Riesz existe um único vetor $z \in X$ tal que

$$\langle Tx, y \rangle = \langle x, z \rangle \quad \text{para todo } x \in V.$$

Definimos uma aplicação $T^* : W \longrightarrow V$ adjunta de $T$ por

$$T^* y = z.$$

Pelo Lema 1, $T^*$ é única e linear. ■

**Proposição 11.** *Seja $V$ um espaço vetorial com produto interno de dimensão finita e $T : V \longrightarrow V$ um operador linear. Se $\mathcal{B} = \{x_1, \ldots, x_n\}$ é uma base ortonormal para $V$ e*

$$A = [T]_{\mathcal{B}} ,$$

*então*

$$[T^*]_{\mathcal{B}} = A^*.$$

*Em outras palavras, em relação a uma base ortonormal, a matriz do operador adjunto $T^*$ é a transposta conjugada da matriz do operador $T$.*

**Prova:** Seja $B = [T^*]_{\mathcal{B}}$. Pela Proposição 9 temos

$$\begin{aligned} a_{ij} &= \langle Tx_j, x_i \rangle , \\ b_{ij} &= \langle T^* x_j, x_i \rangle . \end{aligned}$$

Logo,

$$b_{ij} = \langle T^* x_j, x_i \rangle = \overline{\langle x_i, T^* x_j \rangle} = \overline{\langle Tx_i, x_j \rangle} = \overline{a}_{ji}.$$

■

**Definição 10.** Seja $V$ um espaço vetorial com produto interno de dimensão finita e $T : V \longrightarrow V$ um operador linear. Dizemos que $T$ é um operador **auto-adjunto** se

$$T = T^*.$$

**Corolário 0.** *Seja $V$ um espaço vetorial com produto interno de dimensão finita e $T : V \longrightarrow V$ um operador linear auto-adjunto. Se $\mathcal{B} = \{x_1, \ldots, x_n\}$ é uma base ortonormal para $V$ então $A = [T]_{\mathcal{B}}$ é uma matriz hermitiana.*

**Proposição 12.** *Seja $V$ um espaço vetorial com produto interno de dimensão finita e $T, U : V \longrightarrow V$ operadores lineares. Então*

(i) $(T+U)^* = T^* + U^*$;

(ii) $(\alpha T)^* = \overline{\alpha} T^*$;

(iii) $(TU)^* = U^* T^*$;

(iv) $(T^*)^* = T$;

(v) $\left(T^{-1}\right)^* = (T^*)^{-1}$, *se $T$ é invertível.*

**Teorema 5.** *Sejam $V, W$ espaços vetoriais com produto interno e $T : V \longrightarrow W$ uma aplicação linear. Então*

(i) $\ker T^* = (\operatorname{im} T)^{\perp}$;

(ii) $\ker T = (\operatorname{im} T^*)^{\perp}$;

(iii) $\dim(\operatorname{im} T) = \dim(\operatorname{im} T^*)$.

*Em particular,*

$$V = \ker T^* \oplus^{\perp} \operatorname{im} T.$$

**Prova:** (i)

$$\begin{aligned} y \in \ker T^* &\Leftrightarrow T^* y = 0 \\ &\Leftrightarrow \langle x, T^* y \rangle = 0 \quad \text{para todo } x \in V \\ &\Leftrightarrow \langle Tx, y \rangle = 0 \quad \text{para todo } x \in V \\ &\Leftrightarrow y \text{ é ortogonal a } \operatorname{im} T \\ &\Leftrightarrow y \in (\operatorname{im} T)^{\perp}. \end{aligned}$$

(ii) prova-se de maneira análoga. (iii) $\ker T^* = (\operatorname{im} T)^{\perp}$ implica $\operatorname{im} T = (\ker T^*)^{\perp}$, logo

$$\dim(\operatorname{im} T) = \dim(\ker T^*)^{\perp} = \dim W - \dim(\ker T^*) = \dim(\operatorname{im} T^*).$$

■

Segue do Teorema 5 que a equação

$$Tx = y$$

tem solução se e somente se

$$y \in (\ker T^*)^{\perp},$$

pois $(\ker T^*)^{\perp} = \operatorname{im} T$. Em outras palavras,

**Teorema 6.** (Alternativa de Fredholm) *Seja $V$ um espaço vetorial com produto interno de dimensão finita e $T : V \longrightarrow V$ um operador linear. Então vale apenas uma e somente uma das alternativas a seguir:*

$$\textit{ou } Tx = y \textit{ tem solução,}$$

$$\textit{ou } T^* z = 0 \textit{ tem solução } z \textit{ tal que } \langle y, z \rangle \neq 0.$$

## 6.5 Operadores Unitários e Isometrias

O objetivo principal desta seção é estudar operadores lineares que preservam o produto interno e, conseqüentemente, a norma de um espaço vetorial.

**Definição 11.** Sejam $V, W$ espaços vetoriais com produto interno e $T : V \longrightarrow W$ uma aplicação. Dizemos que $T$ **preserva o produto interno** se

$$\langle Tx, Ty \rangle = \langle x, y \rangle \quad \text{para todos } x, y \in V. \tag{6.25}$$

**Definição 12.** Sejam $V, W$ espaços vetoriais normados e $T : V \longrightarrow W$ uma aplicação. Dizemos que $T$ é uma **isometria** (ou que $T$ preserva a norma) se

$$\|Tx - Ty\| = \|x - y\| \quad \text{para todos } x, y \in V. \tag{6.26}$$

**Teorema 7.** *Sejam $V, W$ espaços vetoriais com produto interno e norma derivada deste. Seja $T : V \longrightarrow W$ uma isometria tal que $T(0) = 0$. Então*

$$T(x + y) = T(x) + T(y)$$

*para todos $x, y \in V$. Se, além disso, $V, W$ forem espaços vetoriais reais, então $T$ é linear.*

**Prova:** Temos

$$\|Tx\| = \|x\| \quad \text{para todo } x \in V. \tag{6.27}$$

Logo,

$$\begin{aligned}
&\langle z, z \rangle - \langle z, w \rangle - \langle w, z \rangle + \langle w, w \rangle \\
&= \langle z - w, z - w \rangle \\
&= \|z - w\|^2 \\
&= \|Tz - Tw\|^2 \\
&= \langle Tz - Tw, Tz - Tw \rangle \\
&= \langle Tz, Tz \rangle - \langle Tz, Tw \rangle - \langle Tw, Tz \rangle + \langle Tw, Tw \rangle \\
&= \langle z, z \rangle - \langle Tz, Tw \rangle - \langle Tw, Tz \rangle + \langle w, w \rangle ,
\end{aligned}$$

donde

$$\langle Tz, Tw \rangle + \langle Tw, Tz \rangle = \langle z, w \rangle + \langle w, z \rangle \tag{6.28}$$

para todos $z, w \in V$. Portanto,

$$\begin{aligned}
\|T(x + y) - T(x) - T(y)\|^2 &= \|T(x + y)\|^2 + \|Tx\|^2 + \|Ty\|^2 \\
&\quad + \langle T(x + y), Tx \rangle + \langle Tx, T(x + y) \rangle \\
&\quad + \langle T(x + y), Ty \rangle + \langle Ty, T(x + y) \rangle \\
&\quad + \langle Tx, Ty \rangle + \langle Ty, Tx \rangle \\
&= \|x + y\|^2 + \|x\|^2 + \|y\|^2 \\
&\quad + \langle x + y, x \rangle + \langle x, x + y \rangle \\
&\quad + \langle x + y, y \rangle + \langle y, x + y \rangle \\
&\quad + \langle x, y \rangle + \langle y, x \rangle \\
&= \|(x + y) - x - y\|^2 \\
&= 0.
\end{aligned}$$

Se $V, W$ são espaços vetoriais reais, (6.28) implica

$$\langle Tx, Ty \rangle = \langle x, y \rangle \tag{6.29}$$

para todos $x, y \in V$. Logo,

$$\langle T(\alpha x), Ty \rangle = \langle \alpha x, y \rangle = \alpha \langle x, y \rangle = \alpha \langle Tx, Ty \rangle = \langle \alpha Tx, Ty \rangle$$

ou

$$\langle T(\alpha x) - \alpha Tx, Ty \rangle = 0$$

para todo $Ty \in \operatorname{im} T$. Como $T(\alpha x) - \alpha Tx \in \operatorname{im} T$, temos $T(\alpha x) - \alpha Tx = T(y)$ para algum $y$ e concluímos que $T(\alpha x) = \alpha Tx$. ■
Uma translação $T(x) = x + a$ é um exemplo de uma isometria que não satisfaz $T(0) = 0$.

**Corolário 1.** *Sejam $V, W$ espaços vetoriais reais com produto interno e norma derivada deste. Então toda isometria entre $V$ e $W$ é a composta de uma aplicação linear e uma translação.*

**Prova:** Seja $T : V \longrightarrow W$ uma isometria. Seja $a = T(0)$. Então $U : V \longrightarrow W$ definida por $U(x) = T(x) - a$ é uma isometria que satisfaz $U(0) = 0$ e o resultado segue do Teorema 7. ■

**Teorema 8.** *Sejam $V, W$ espaços vetoriais com produto interno e norma derivada deste. Seja $T : V \longrightarrow W$ uma aplicação linear. As seguintes afirmações são equivalentes:*

(i) *$T$ é uma isometria.*

(ii) *$T$ preserva o produto interno.*

(iii) *$T^*T = I$.*

**Prova:** (i) $\Rightarrow$ (ii) Se $V, W$ são reais, pela identidade de polarização real temos

$$\langle x, y \rangle = \frac{1}{4} \|x + y\|^2 - \frac{1}{4} \|x - y\|^2 = \frac{1}{4} \|Tx + Ty\|^2 - \frac{1}{4} \|Tx - Ty\|^2 = \langle Tx, Ty \rangle .$$

(ii) $\Rightarrow$ (i) Pois

$$\|Tx - Ty\|^2 = \|T(x - y)\|^2 = \langle T(x - y), T(x - y) \rangle = \langle x - y, x - y \rangle = \|x - y\|^2 .$$

(ii) $\Rightarrow$ (iii) Dado $y \in V$, como

$$\langle x, y \rangle = \langle Tx, Ty \rangle = \langle x, T^*Ty \rangle$$

para todo $x \in V$, segue que $T^*Ty = y$ para todo $y \in V$, isto é, $T^*T = I$.
(iii) $\Rightarrow$ (ii) Pois

$$\langle x, y \rangle = \langle x, T^*Ty \rangle = \langle Tx, Ty \rangle .$$

■

**Definição 13.** Seja $V$ um espaço vetorial com produto interno e $T : V \longrightarrow V$ um operador linear. Dizemos que $T$ é um **operador unitário** se

$$TT^* = T^*T = I. \tag{6.30}$$

**Corolário 2.** *Seja $V$ um espaço vetorial com produto interno e $T : V \longrightarrow V$ um operador unitário. Então $T$ é uma isometria e preserva o produto interno. Em particular, $T$ leva conjuntos ortonormais em conjuntos ortonormais.*

Um operador unitário real é chamado um operador ortogonal. Neste caso,

$$TT^t = T^tT = I. \tag{6.31}$$

Analogamente, dizemos que uma matriz $A$ é **unitária** se $AA^* = A^*A = I$ e que uma matriz real é ortogonal se $AA^t = A^tA = I$. Uma matriz unitária é portanto uma matriz cuja inversa é a sua transposta conjugada e uma matriz ortogonal é uma matriz cuja inversa é a sua transposta.

**Proposição 13.** *Seja $V$ um espaço vetorial com produto interno de dimensão finita e $T : V \longrightarrow V$ um operador unitário. Então*

$$|\det T| = 1.$$

*Em particular, se $V$ é real então $\det T = \pm 1$.*

**Prova:** Pois

$$\det(TT^*) = \det I = 1$$

e

$$\det T^* = \det \overline{T^t} = \det \overline{T}^t = \det \overline{T} = \overline{\det T},$$

logo

$$\det(TT^*) = \det T \det T^* = \det T \overline{\det T} = |\det T|^2 .$$

■

Em particular, uma aplicação linear que preserva normas preserva volumes.

**Proposição 14.** *Seja $V$ um espaço vetorial com produto interno de dimensão finita e $T : V \longrightarrow V$ um operador linear. Então $T$ é unitário se e somente se a sua matriz em relação a uma base ortonormal é uma matriz unitária.*

## 6.6 Operadores Normais

O objetivo principal desta seção é resolver o seguinte problema: em que condições um operador linear possui uma base ortonormal em relação à qual a sua matriz é diagonal.

Vamos começar obtendo condições necessárias para isso acontecer. Se $\mathcal{B} = \{x_1, \ldots, x_n\}$ é uma base ortonormal que diagonaliza o operador $T$, então $T$ é representado nesta base por uma matriz diagonal com entradas diagonais $\lambda_1, \ldots, \lambda_n$. O operador adjunto $T^*$ é representado nesta mesma base pela matriz transposta conjugada, ou seja uma matriz diagonal com entradas diagonais $\overline{\lambda}_1, \ldots, \overline{\lambda}_n$. Em particular, como ambos operadores são representados por matrizes diagonais em relação a uma mesma base, segue que eles comutam: $TT^* = T^*T$. Veremos no final da seção que esta condição também é suficiente, pelo menos no caso complexo. No caso real, esta condição não é suficiente, pois pode ocorrer que $T$ nem possua autovalores (por exemplo, quase todas as rotações em $\mathbb{R}^2$).

**Definição 13.** Seja $V$ um espaço vetorial com produto interno e $T : V \longrightarrow V$ um operador linear. Dizemos que $T$ é um **operador normal** se

$$TT^* = T^*T. \tag{6.32}$$

Exemplos de operadores normais são operadores auto-adjuntos e operadores unitários. Analogamente, dizemos que uma matriz $A$ é **normal** se $AA^* = A^*A$.

Primeiro trataremos do caso especial dos operadores auto-adjuntos.

**Teorema 9.** *Seja $V$ um espaço vetorial com produto interno e $T : V \longrightarrow V$ um operador linear auto-adjunto. Então todo autovalor de $T$ é real.*

*Além disso, autovetores de $T$ associados a autovalores distintos são ortogonais.*

**Prova:** Suponha que $\lambda$ é um autovalor de $T$. Seja $v$ um autovetor não-nulo associado a $\lambda$. Então

$$\lambda \langle v, v \rangle = \langle \lambda v, v \rangle = \langle Tv, v \rangle = \langle v, Tv \rangle = \langle v, \lambda v \rangle = \overline{\lambda} \langle v, v \rangle$$

e portanto

$$\lambda = \overline{\lambda},$$

isto é, $\lambda \in \mathbb{R}$. Sejam $\lambda_1, \lambda_2$ autovalores reais distintos de $T$ e $v_1, v_2$ autovetores não-nulos associado a $\lambda_1, \lambda_2$, respectivamente. Então

$$\lambda_1 \langle v_1, v_2 \rangle = \langle \lambda_1 v_1, v_2 \rangle = \langle Tv_1, v_2 \rangle = \langle v_1, Tv_2 \rangle = \langle v_1, \lambda_2 v_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle$$

e como $\lambda_1 \neq \lambda_2$, concluímos que

$$\langle v_1, v_2 \rangle = 0.$$

■

**Teorema 10.** *Seja $V$ um espaço vetorial com produto interno e $T : V \longrightarrow V$ um operador linear auto-adjunto. Então $T$ possui um autovalor real.*

**Prova:** Se $V$ é um espaço vetorial complexo, não há necessidade de fazer nada, pois todo operador complexo possui autovalores (pois seu polinômio característico possui raízes) e pelo teorema anterior seus autovalores são reais. Portanto, assuma que $V$ é um espaço vetorial real. Escolha uma base ortonormal $\mathcal{B}$ para $V$ e seja $A = [T]_{\mathcal{B}}$. Como $T$ é um operador auto-adjunto, isto é, que satisfaz $T^* = T$, a matriz $A$ é uma matriz hermitiana (isto é, simétrica, pois é uma matriz real) pois também satisfaz $A^* = A$. Se considerada como uma matriz de um operador complexo em um espaço com produto interno complexo seu polinômio característico possui uma raiz complexa $\lambda$, logo $\det(\lambda I - A) = 0$ e a matriz $\lambda I - A$ é singular e possui núcleo não-trivial; por exemplo, $A$ é a matriz do operador $U$ em $M_{n \times 1}(\mathbb{C})$ definido por $UX = AX$, onde $n = \dim V$. Mas como $A$ é hermitiana, o operador $U$ é auto-adjunto e segue do teorema anterior que $\lambda$ é real. Como $\lambda I - A$ é uma matriz real cujo determinante é nulo, ela possui núcleo real não-trivial, portanto $\lambda$ é um autovalor de $A$. ■

**Proposição 15.** *Seja $V$ um espaço vetorial com produto interno e $T : V \longrightarrow V$ um operador linear. Se $W \subset V$ é um subespaço invariante sob $T$, então $W^\perp$ é invariante sob $T^*$.*

**Prova:** Se $y \in W^\perp$, então $\langle x, y \rangle = 0$ para todo $x \in W$. Logo

$$\langle x, T^* y \rangle = \langle Tx, y \rangle = 0$$

para todo $x \in W$, pois $Tx \in W$. Portanto, $T^* y \in W^\perp$. ■

**Teorema 11.** *Seja $V$ um espaço vetorial com produto interno de dimensão finita e $T : V \longrightarrow V$ um operador linear auto-adjunto. Então $T$ é diagonalizável através de uma base ortonormal.*

**Prova:** A demonstração será por indução em $\dim V$. Pelo Teorema 10, $T$ possui um autovalor real. Se $\dim V = 1$, isso dá uma base ortonormal para $V$ constituída de autovetores de $T$. Assuma o teorema verdadeiro para espaços com produto interno de dimensão menor que $n$ e seja $\dim V = n$. Novamente usando o Teorema 10, seja $x_1$ um autovetor de norma 1 associado a um autovalor real de $T$. Seja $W = \langle x_1 \rangle$. Então $W$ é invariante sob $T$ e pela proposição anterior $W^\perp$ é invariante sob $T^* = T$ . Mas $\dim W^\perp = n - 1$, logo pela hipótese de indução existe uma base ortonormal $\{x_2, \ldots, x_n\}$ para $W^\perp$ de autovetores de $T|_{W^\perp}$ e portanto de $T$. Como $V = W \oplus W^\perp$, segue que $\{x_1, x_2, \ldots, x_n\}$ é uma base ortonormal para $V$ de autovetores de $T$. ■

**Corolário 3.** *Seja $A$ uma matriz hermitiana. Então existe uma matriz unitária $P$ tal que*

$$D = P^* A P$$

*é uma matriz diagonal.*

**Corolário 4.** *Seja $V$ um espaço vetorial real com produto interno de dimensão finita. Então um operador linear sobre $V$ é diagonalizável através de uma base ortonormal se e somente se ele é auto-adjunto.*

**Prova:** Se $T : V \longrightarrow V$ é um operador linear diagonalizável através de uma base ortonormal, então existe uma matriz ortogonal $P$ tal que

$$D = P^t A P.$$

Em particular,

$$A = PDP^t$$

e

$$A^t = \left(PDP^t\right)^t = \left(P^t\right)^t D^t P^t = PDP^t = A.$$

■

Voltemos agora a considerar operadores normais. Em primeiro lugar, vamos caracterizar os operadores normais.

**Lema 2.** *Seja $V$ um espaço vetorial complexo com produto interno e $T : V \longrightarrow V$ um operador linear. Então*

$$\langle Tx, x\rangle = 0 \quad \textit{para todo } x \in V$$

*se e somente se*

$$T = 0.$$

**Prova:** Suponha $\langle Tz, z\rangle = 0$ para todo $z \in V$. Dados quaisquer $x, y \in V$, temos

$$\begin{aligned}
0 &= \langle T(x+y), x+y\rangle = \langle Tx, x\rangle + \langle Tx, y\rangle + \langle Ty, x\rangle + \langle Ty, y\rangle = \langle Tx, y\rangle + \langle Ty, x\rangle, \\
0 &= \langle T(x+iy), x+iy\rangle = \langle Tx, x\rangle + \langle Tx, iy\rangle + \langle iTy, x\rangle + \langle T(iy), iy\rangle = -i\langle Tx, y\rangle + i\langle Ty, x\rangle,
\end{aligned}$$

ou seja,

$$\begin{aligned}
\langle Tx, y\rangle + \langle Ty, x\rangle &= 0, \\
\langle Tx, y\rangle - \langle Ty, x\rangle &= 0,
\end{aligned}$$

Portanto, somando as duas equações concluímos que

$$\langle Tx, y\rangle = 0$$

para todo $y \in V$, logo $Tx = 0$. Como $x$ é arbitrário, isso implica $T = 0$. A recíproca é imediata. ■

**Lema 3.** *Seja $V$ um espaço vetorial com produto interno e $T : V \longrightarrow V$ um operador linear. Se $T$ é auto-adjunto, então*

$$\langle Tx, x\rangle \in \mathbb{R} \quad \textit{para todo } x \in V.$$

*Se $V$ é um espaço vetorial complexo e $\langle Tx, x\rangle \in \mathbb{R}$ para todo $x \in V$, então $T$ é auto-adjunto.*

**Prova:** Se $T$ é auto-adjunto, então

$$\langle Tx, x\rangle = \langle x, Tx\rangle = \overline{\langle Tx, x\rangle},$$

logo $\langle Tx, x\rangle \in \mathbb{R}$. Reciprocamente, se $V$ é um espaço vetorial complexo e $\langle Tx, x\rangle \in \mathbb{R}$ para todo $x \in V$, então

$$\langle Tx, x\rangle = \overline{\langle Tx, x\rangle} = \langle x, Tx\rangle = \langle T^* x, x\rangle,$$

de modo que

$$\langle (T - T^*)x, x\rangle = 0 \quad \text{para todo } x \in V.$$

Pelo lema anterior, segue que $T = T^*$. ■

**Lema 4.** *Seja $V$ um espaço vetorial real com produto interno e $T : V \longrightarrow V$ um operador linear. Então*

$$\langle Tx, x\rangle = 0 \quad \textit{para todo } x \in V$$

*se e somente se*

$$T^* = -T.$$

**Prova:** Se $\langle Tx, x\rangle = 0$ para todo $x \in V$, então

$$\begin{aligned}
0 &= \langle T(x+y), x+y\rangle = \langle Tx, x\rangle + \langle Tx, y\rangle + \langle Ty, x\rangle + \langle Ty, y\rangle \\
&= \langle Tx, y\rangle + \langle Ty, x\rangle = \langle Tx, y\rangle + \langle x, Ty\rangle = \langle Tx, y\rangle + \langle T^*x, y\rangle \\
&= \langle (T+T^*)\, x, y\rangle
\end{aligned}$$

para todos $x, y \in V$. Portanto, $T = -T^*$. Reciprocamente, se $T^* = -T$, então

$$\langle Tx, x\rangle = \langle x, T^*x\rangle = \langle x, -Tx\rangle = -\langle Tx, x\rangle,$$

logo $\langle Tx, x\rangle = 0$. ■
Um operador $T$ que satisfaz $T^* = -T$ é chamado um **operador antiauto-adjunto**.

**Teorema 12.** *Seja $V$ um espaço vetorial com produto interno e $T : V \longrightarrow V$ um operador linear. Então $T$ é normal se e somente se*

$$\|Tx\| = \|T^*x\| \quad \textit{para todo } x \in V.$$

*Em particular, se $T$ é normal, segue que*

$$V = \ker T \oplus^{\perp} \operatorname{im} T.$$

**Prova:** Seja $T$ normal. Então

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle x, T^*Tx\rangle = \langle x, TT^*x\rangle = \langle T^*x, T^*x\rangle = \|T^*x\|^2.$$

Reciprocamente, se $\|Tx\| = \|T^*x\|$ para todo $x \in V$, então

$$\langle T^*Tx, x\rangle = \langle Tx, Tx\rangle = \|Tx\|^2 = \|T^*x\|^2 = \langle T^*x, T^*x\rangle = \langle TT^*x, x\rangle$$

para todo $x \in V$, o que implica

$$\langle (T^*T - TT^*)\, x, x\rangle = 0$$

Se $V$ é um espaço vetorial complexo, segue imediatamente do Lema 2 que $T^*T = TT^*$. Se $V$ é um espaço vetorial real, segue do Lema 2 que $T^*T - TT^*$ é um operador antiauto-adjunto; como $T^*T - TT^*$ também é um operador auto-adjunto, pois

$$(T^*T - TT^*)^* = (T^*T)^* - (TT^*)^* = T^*(T^*)^* - (T^*)^*T^* = T^*T - TT^*,$$

concluímos que $T^*T - TT^* = 0$.

Se $T$ é um operador normal, segue que $\ker T = \ker T^*$. Como

$$V = \ker T^* \oplus^{\perp} \operatorname{im} T,$$

obtemos

$$V = \ker T \oplus^{\perp} \operatorname{im} T.$$

■
Se $V = \ker T \oplus^{\perp} \operatorname{im} T$, não é necessariamente verdade que $T$ é um operador normal, pois $T$ pode ser definido de qualquer forma arbitrária em $\operatorname{im} T$.

**Teorema 13.** *Seja $V$ um espaço vetorial com produto interno de dimensão finita e $T : V \longrightarrow V$ um operador normal. Então $v$ é um autovetor para $T$ associado ao autovalor $\lambda$ se e somente se $v$ é um autovetor para $T^*$ associado ao autovalor $\overline{\lambda}$.*

**Prova:** Se $\lambda$ é qualquer escalar, então $T - \lambda I$ também é um operador normal, pois $(T - \lambda I)^* = T^* - \overline{\lambda} I$ e portanto

$$\begin{aligned}(T - \lambda I)(T - \lambda I)^* &= (T - \lambda I)\left(T^* - \overline{\lambda} I\right) = TT^* - \overline{\lambda} T - \lambda T^* + \lambda\overline{\lambda} I \\ &= T^*T - \lambda T^* - \overline{\lambda} T + \overline{\lambda}\lambda I = \left(T^* - \overline{\lambda} I\right)(T - \lambda I) \\ &= (T - \lambda I)^* (T - \lambda I).\end{aligned}$$

Segue do teorema anterior que

$$\|(T - \lambda I) v\| = \left\|\left(T^* - \overline{\lambda} I\right) v\right\|,$$

logo $(T - \lambda I) v = 0$ se e somente se $\left(T^* - \overline{\lambda} I\right) v = 0$. ■

**Teorema 14.** *Seja $V$ um espaço vetorial complexo com produto interno de dimensão finita. Então todo operador linear sobre $V$ é triangularizável através de uma base ortonormal.*

**Prova:** A demonstração será por indução em $\dim V$. O resultado é obviamente verdadeiro se $\dim V = 1$. Assuma o teorema verdadeiro para espaços vetoriais complexos com produto interno de dimensão menor que $n$ e seja $\dim V = n$. Como $V$ é um espaço vetorial complexo, existe um autovetor $x_n$ de norma 1 associado a um autovalor complexo de $T^*$. Seja $W = \langle x_n \rangle$. Então $W$ é invariante sob $T^*$ e pela Proposição 15 $W^\perp$ é invariante sob $T$ . Mas $\dim W^\perp = n - 1$, logo pela hipótese de indução existe uma base ortonormal $\{x_1, \ldots, x_{n-1}\}$ de $W^\perp$ em relação a qual a matriz de $T|_{W^\perp}$ é triangular. Como $V = W \oplus W^\perp$, segue que $\{x_1, \ldots, x_{n-1}, x_n\}$ é uma base ortonormal para $V$ em relação à qual a matriz de $T$ é triangular superior, pois $Tx_n$ é simplesmente uma combinação linear de $x_1, \ldots, x_{n-1}, x_n$. ■

**Corolário 5.** *Seja $A$ uma matriz complexa. Então existe uma matriz unitária $U$ tal que*

$$T = U^* A U$$

*é uma matriz triangular.*

**Teorema 15.** *Seja $V$ um espaço vetorial complexo com produto interno de dimensão finita. Então um operador linear sobre $V$ é diagonalizável através de uma base ortonormal se e somente se ele é normal.*

**Prova:** Já vimos no início desta seção que todo operador diagonalizável através de uma base ortonormal é normal. Para provar a recíproca, usando o Teorema 14 já sabemos que existe uma base ortonormal $\mathcal{B} = \{x_1, \ldots, x_n\}$ em relação à qual a matriz do operador linear normal $T : V \longrightarrow V$ é triangular. O resultado seguirá se provarmos que toda matriz triangular normal é diagonal.

E, de fato, seja $A = [T]_{\mathcal{B}}$. Como $A$ é triangular, temos

$$Tx_1 = a_{11} x_1.$$

Pelo Teorema 13,

$$T^* x_1 = \overline{a}_{11} x_1.$$

Por outro lado,

$$T^* x_1 = \sum_{i=1}^n (A^*)_{i1} x_i = \sum_{i=1}^n \overline{a}_{1i} x_i.$$

Portanto, $\overline{a}_{1i} = 0$ para todo $i \geqslant 2$, o que implica $a_{1i} = 0$ para todo $i \geqslant 2$.

Agora, em particular, $a_{12} = 0$ e o fato de $A$ ser triangular implicam que

$$Tx_2 = a_{22} x_2.$$

Usando o mesmo argumento concluímos que $a_{2i} = 0$ para todo $i \geqslant 3$. Continuando com este argumento provamos que $Tx_j = a_{jj} x_j$ para todo $j$, logo $A$ é diagonal. ■

**Corolário 6.** *Seja $A$ uma matriz normal. Então existe uma matriz unitária $U$ tal que*

$$D = U^* A U$$

*é uma matriz diagonal.*

Leia a Seção 10.4 do livro-texto.

## 6.7 Teoria Espectral para Operadores Auto-adjuntos

**Teorema 16.** *Seja $T : V \longrightarrow V$ um operador auto-adjunto sobre um espaço real com produto interno de dimensão finita ou um operador normal sobre um espaço vetorial complexo com produto interno de dimensão finita. Sejam $\lambda_1, \ldots, \lambda_k$ os autovalores distintos de $T$. Seja $W_i$ o autoespaço associado a $\lambda_i$ e $E_i : V \longrightarrow V$ a projeção ortogonal de $V$ sobre $W_i$. Então $W_i$ é ortogonal a $W_j$ se $i \neq j$,*

$$V = W_1 \oplus^{\perp} \ldots \oplus^{\perp} W_k$$

*e*

$$T = \lambda_1 E_1 + \ldots + \lambda_k E_k.$$

**Prova:** Sejam $v_i \in W_i$ e $v_j \in W_j$ com $i \neq j$. Então, usando o Teorema 13, temos

$$\lambda_i \langle v_i, v_j \rangle = \langle \lambda_i v_i, v_j \rangle = \langle T v_i, v_j \rangle = \langle v_i, T^* v_j \rangle = \left\langle v_i, \overline{\lambda}_j v_j \right\rangle = \lambda_j \langle v_i, v_j \rangle$$

e como $\lambda_i \neq \lambda_2$, concluímos que

$$\langle v_i, v_j \rangle = 0.$$

Do Teorema 15 segue imediatamente que $V = W_1 +^{\perp} \ldots +^{\perp} W_k$. Se $x_i \in W_i$ e

$$x_1 + \ldots + x_k = 0,$$

então

$$0 = \langle x_j, 0 \rangle = \langle x_j, x_1 + \ldots + x_k \rangle = \|x_j\|^2,$$

logo os espaços $W_i$ são L.I. e a soma é direta. Finalmente, como a soma é direta, temos

$$E_1 + \ldots + E_k = I,$$

donde

$$T = TI = TE_1 + \ldots + TE_k = \lambda_1 E_1 + \ldots + \lambda_k E_k.$$

■

Esta decomposição é chamada a **resolução espectral** do operador $T$. O **espectro** de um operador linear é o conjunto de seus autovalores.

**Definição 14.** Seja $V$ um espaço vetorial com produto interno e $T : V \longrightarrow V$ um operador linear auto-adjunto. Dizemos que $T$ é **positivo definido** se

$$\langle Tx, x \rangle > 0 \tag{6.33}$$

para todo $x \in V$. Se

$$\langle Tx, x \rangle \geqslant 0 \tag{6.34}$$

para todo $x \in V$, dizemos que $T$ é **positivo semidefinido**.

Dado um operador linear invertível $T$, o operador $T^*T$ é sempre auto-adjunto, positivo definido, pois

$$(T^*T)^* = T^*(T^*)^* = T^*T,$$
$$\langle T^*Tx, x\rangle = \langle Tx, Tx\rangle = \|Tx\|^2 > 0.$$

Se $T$ não é invertível, então $T^*T$ é apenas auto-adjunto, positivo semidefinido.

**Lema 5.** *Toda projeção ortogonal é um operador auto-adjunto.*

**Prova:** Seja $E : V \longrightarrow V$ uma projeção. Então existe uma base $\mathcal{B}' = \{x_1, \ldots, x_m\}$ para $\operatorname{im} E$ e uma base $\mathcal{B}'' = \{x_{m+1}, \ldots, x_n\}$ para $\ker E$ tais que $\mathcal{B} = \{x_1, \ldots, x_n\}$ é uma base para $V$ que diagonaliza $E$ e

$$[E]_{\mathcal{B}} = \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix}.$$

Usando o processo de ortogonalização de Gram-Schmidt, podemos escolher $\mathcal{B}'$ e $\mathcal{B}''$ ortonormais. Se $E$ é uma projeção ortogonal, então $\mathcal{B}' \perp \mathcal{B}''$ e segue imediatamente que $\mathcal{B}$ é uma base ortonormal para $V$ que diagonaliza $E$. Daí, pela Proposição 11,

$$[E^*]_{\mathcal{B}} = ([E]_{\mathcal{B}})^* = \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix}^* = \begin{bmatrix} I_m^* & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix} = [E]_{\mathcal{B}},$$

e portanto $E$ é auto-adjunto. ■

**Lema 6.** *Seja $T : V \longrightarrow V$ um operador normal diagonalizável sobre um espaço real com produto interno de dimensão finita. Então $T$ é auto-adjunto.*

**Prova:** Por definição, existe uma base para $V$ constituída por autovetores de $T$. Afirmamos que podemos tomar esta base ortonormal. De fato, como vimos na demonstração do Teorema 16, se $T$ é um operador normal, autoespaços correspondentes a autovalores distintos são ortogonais. Usando o processo de ortogonalização de Gram-Schmidt em cada autoespaço, concluímos como no lema anterior que podemos obter uma base ortonormal de autovetores de $T$. O resultado segue agora do Corolário 4. ■

**Teorema 17.** *Seja $T : V \longrightarrow V$ um operador normal diagonalizável sobre um espaço com produto interno de dimensão finita. Então, os autovalores de $T$ são*

(i) *reais, se e somente se $T$ é auto-adjunto;*

(ii) *não-negativos, se e somente se $T$ é positivo semidefinido;*

(iii) *positivos, se e somente se $T$ é positivo definido;*

(iv) *de módulo 1, se e somente se $T$ é unitário.*

**Prova:** Pelo Lema 6 e pelo Teorema 15, $T$ satisfaz as hipóteses do Teorema Espectral. Seja

$$T = \lambda_1 E_1 + \ldots + \lambda_k E_k$$

a resolução espectral de $T$. Segue do Lema 5 que a resolução espectral de $T^*$ é

$$T^* = \overline{\lambda}_1 E_1 + \ldots + \overline{\lambda}_k E_k.$$

$T$ é auto-adjunto se e somente se $T = T^*$, isto é, se e somente se

$$\left(\lambda_1 - \overline{\lambda}_1\right) E_1 + \ldots + \left(\lambda_k - \overline{\lambda}_k\right) E_k = 0.$$

Como $E_i E_j = 0$ se $i \neq j$ e $E_i$ não é o operador nulo, concluímos que $T$ é auto-adjunto se e somente se $\lambda_i = \overline{\lambda}_i$, isto é, se e somente se os autovalores de $T$ são reais.

Agora,

$$\langle Tx, x\rangle = \left\langle \sum_{i=1}^{k} \lambda_i E_i x, \sum_{j=1}^{k} E_j x \right\rangle = \sum_{i,j=1}^{k} \lambda_i \langle E_i x, E_j x\rangle = \sum_{i,j=1}^{k} \lambda_i \|E_i x\|^2,$$

logo, tomando $x \in W_i$ para $i = 1, \ldots, k$, concluímos que $\langle Tx, x\rangle \geqslant 0$ para todo $x \in V$ se e somente se $\lambda_i \geqslant 0$ para todo $i$ e que $\langle Tx, x\rangle > 0$ para todo $x \in V$ se e somente se $\lambda_i > 0$ para todo $i$.

Finalmente, temos

$$TT^* = (\lambda_1 E_1 + \ldots + \lambda_k E_k)\left(\overline{\lambda}_1 E_1 + \ldots + \overline{\lambda}_k E_k\right) = |\lambda_1|^2 E_1 + \ldots + |\lambda_k|^2 E_k.$$

Se $|\lambda_1| = \ldots = |\lambda_k| = 1$, então $TT^* = I$ e $T$ é unitário. Reciprocamente, se $TT^* = I$, então

$$|\lambda_1|^2 E_1 + \ldots + |\lambda_k|^2 E_k = I,$$

donde, multiplicando a equação por $E_j$, obtemos

$$E_j = |\lambda_j|^2 E_j,$$

o que implica $|\lambda_j| = 1$. ■

## 6.8 Métodos Variacionais

Métodos variacionais são extremamente importantes em várias áreas da Matemática, Pura e Aplicada. Em Álgebra Linear Numérica, eles são a base de vários métodos eficientes importantes para a resolução de sistemas lineares e de obtenção de autovalores de matrizes simétricas.

**Teorema 18.** (Método Variacional para a Solução de Sistemas Lineares) *Seja $A$ uma matriz simétrica positiva definida. Então a solução do sistema*

$$Ax = b$$

*é dada pelo ponto $x$ que minimiza o funcional*

$$f(y) = \frac{1}{2} y^t A y - y^t b. \tag{6.35}$$

**Prova:** Como $A$ é uma matriz simétrica positiva definida, em particular $A$ é invertível e existe uma solução $x$ para o sistema $Ax = b$. Observe que

$$y^t A x = \left(y^t A x\right)^t = x^t A^t y = x^t A y.$$

Daí,

$$\begin{aligned}
f(y) - f(x) &= \frac{1}{2} y^t A y - y^t b - \frac{1}{2} x^t A x + x^t b = \frac{1}{2} y^t A y - y^t A x - \frac{1}{2} x^t A x + x^t A x \\
&= \frac{1}{2} y^t A y - y^t A x + \frac{1}{2} x^t A x \\
&= \frac{1}{2} y^t A y - \frac{1}{2} y^t A x - \frac{1}{2} x^t A y + \frac{1}{2} x^t A x \\
&= \frac{1}{2} y^t A (y - x) - \frac{1}{2} x^t A (y - x) \\
&= \frac{1}{2} (y - x)^t A (y - x).
\end{aligned}$$

Como $A$ é positiva definida, segue que

$$(y-x)^t A(y-x) = \langle A(y-x), (y-x)\rangle \geqslant 0$$

e

$$(y-x)^t A(y-x) = 0$$

se e somente se $y = x$. Portanto,

$$f(y) > f(x)$$

para todo $y \neq x$ e o mínimo de $f$ ocorre em $x$. ■

Observe que definindo um produto interno a partir da matriz simétrica positiva definida $A$ da maneira usual por $\langle x, y\rangle = y^t Ax$, o funcional $f$ pode ser escrito na forma

$$f(y) = \frac{1}{2}\langle y, y\rangle - y^t b.$$

Outra maneira de enxergar o resultado do teorema anterior é observar que o gradiente do funcional $f$ é

$$\nabla f(y) = Ay - b;$$

se $x$ é um ponto de mínimo temos $\nabla f(x) = 0$, ou seja,

$$Ax = b.$$

Este método variacional é a base do Método do Gradiente Conjugado para a resolução de sistemas lineares envolvendo matrizes simétricas positivas definidas, que aparecem freqüentemente nas aplicações.

Veremos agora que o menor autovalor de um operador linear pode ser encontrado como o mínimo de um certo funcional, enquanto que o seu maior autovalor é o máximo deste mesmo funcional:

**Teorema 19.** (Princípio de Rayleigh) *Seja $V$ um espaço vetorial com produto interno de dimensão $n$ e $T : V \longrightarrow V$ um operador linear auto-adjunto. Sejam $\lambda_1 \leqslant \ldots \leqslant \lambda_n$ os autovalores de $T$, de modo que $\lambda_1$ é o menor autovalor de $T$ e $\lambda_n$ é o maior autovalor de $T$. Então*

$$\lambda_1 = \min_{\substack{x\in V\\ x\neq 0}} \frac{\langle Tx, x\rangle}{\|x\|^2} \tag{6.36}$$

*e*

$$\lambda_n = \max_{\substack{x\in V\\ x\neq 0}} \frac{\langle Tx, x\rangle}{\|x\|^2} \tag{6.37}$$

**Prova:** Seja $\mathcal{B} = \{v_1, \ldots, v_n\}$ uma base ortonormal de autovetores de $T$ correspondentes aos autovalores $\lambda_1 \leqslant \ldots \leqslant \lambda_n$ de $T$. Então, para todo $x = \sum_{i=1}^{n} x_i v_i \in V$ temos

$$\begin{aligned}
\lambda_1 \|x\|^2 &= \sum_{i=1}^{n} \lambda_1 x_i^2 \\
&\leqslant \sum_{i=1}^{n} \lambda_i x_i^2 = \sum_{i,j=1}^{n} \lambda_i x_i x_j \langle v_i, v_j\rangle = \sum_{i,j=1}^{n} \langle \lambda_i x_i v_i, x_j v_j\rangle \\
&= \left\langle \sum_{i=1}^{n} \lambda_i x_i v_i, \sum_{j=1}^{n} x_j v_j \right\rangle = \left\langle \sum_{i=1}^{n} x_i T v_i, \sum_{j=1}^{n} x_j v_j \right\rangle = \left\langle T\left(\sum_{i=1}^{n} x_i v_i\right), \sum_{j=1}^{n} x_j v_j \right\rangle \\
&= \langle Tx, x\rangle .
\end{aligned}$$

Portanto, para todo $x \in V$, $x \neq 0$, vale

$$\lambda_1 \leqslant \frac{\langle Tx, x\rangle}{\|x\|^2}.$$

O mínimo é atingido em $x = v_1$ ou em qualquer outro autovetor de $T$ associado a $\lambda_1$. Da mesma forma, obtemos

$$\lambda_n \|x\|^2 = \sum_{i=1}^{n} \lambda_n x_i^2 \geqslant \sum_{i=1}^{n} \lambda_i x_i^2 = \langle Tx, x\rangle .$$

■

O quociente

$$\frac{\langle Tx, x\rangle}{\|x\|^2}$$

é chamado o **quociente de Rayleigh**.

Os outros autovalores de $T$, $\lambda_2, \ldots, \lambda_{n-1}$, são pontos de sela e podem ser encontrado através de um princípio de minimax:

**Teorema 20.** (Princípio de Minimax para Autovalores) *Seja $V$ um espaço vetorial com produto interno de dimensão $n$ e $T : V \longrightarrow V$ um operador linear auto-adjunto. Sejam $\lambda_1 \leqslant \ldots \leqslant \lambda_n$ os autovalores de $T$. Então*

$$\lambda_j = \min_{\substack{W \subset V \text{ subespaço} \\ \dim W = j}} \left( \max_{\substack{x \in W \\ \|x\|=1}} \langle Tx, x\rangle \right). \tag{6.38}$$

**Prova:** Seja $W \subset V$ um subespaço de dimensão $j$. Primeiro mostraremos que

$$\max_{\substack{x \in W \\ \|x\|=1}} \langle Tx, x\rangle \geqslant \lambda_j.$$

Seja $\mathcal{B} = \{v_1, \ldots, v_n\}$ uma base ortonormal de autovetores de $T$ correspondentes aos autovalores $\lambda_1, \ldots, \lambda_n$. Seja $Z = \langle v_1, \ldots, v_{j-1}\rangle$. Como $Z^\perp = \langle v_j, \ldots, v_n\rangle$, temos

$$n \geqslant \dim\left(W + Z^\perp\right) = \dim W + \dim Z^\perp - \dim\left(W \cap Z^\perp\right) = j + n - (j-1) - \dim\left(W \cap Z^\perp\right),$$

de modo que

$$\dim\left(W \cap Z^\perp\right) \geqslant 1$$

e existe um vetor $x \in W \cap Z^\perp$ tal que $\|x\| = 1$. Escrevendo $x = \sum_{k=j}^{n} x_k v_k$, temos $|x| = \sum_{k=j}^{n} |x_k|^2 = 1$, donde

$$\begin{aligned}
\langle Tx, x\rangle &= \left\langle \sum_{k=j}^{n} x_k T v_k, \sum_{l=j}^{n} x_l v_l \right\rangle = \left\langle \sum_{k=j}^{n} x_k \lambda_k v_k, \sum_{l=j}^{n} x_l v_l \right\rangle = \sum_{k,l=j}^{n} \lambda_k x_k x_l \langle v_k, v_l\rangle \\
&= \sum_{k=j}^{n} \lambda_k |x_k|^2 \geqslant \lambda_j \sum_{k=j}^{n} |x_k|^2 = \lambda_j.
\end{aligned}$$

Para completar a demonstração, devemos encontrar um subespaço $W \subset V$ de dimensão $j$ tal que $\langle Tx, x\rangle \leqslant \lambda_j$ para todo $x \in W$ com $\|x\| = 1$. Tomemos $W = \langle v_1, \ldots, v_j\rangle$. Temos

$$\begin{aligned}
\langle Tx, x\rangle &= \left\langle \sum_{k=1}^{j} x_k T v_k, \sum_{l=1}^{j} x_l v_l \right\rangle = \left\langle \sum_{k=1}^{j} x_k \lambda_k v_k, \sum_{l=1}^{j} x_l v_l \right\rangle = \sum_{k,l=1}^{j} \lambda_k x_k x_l \langle v_k, v_l\rangle \\
&= \sum_{k=1}^{j} \lambda_k |x_k|^2 \leqslant \lambda_j \sum_{k=1}^{j} |x_k|^2 = \lambda_j.
\end{aligned}$$

O minimax é atingido em $v_j$. ■

## 6.9 Formas Bilineares e Sesquilineares

**Definição 15.** Seja $V$ um espaço vetorial sobre $\mathbb{K}$. Uma **forma sesquilinear** em $V$ é uma função $B : V \times V \longrightarrow \mathbb{K}$ que satisfaz

(i) $B(\alpha x + \beta y, z) = \alpha B(x, z) + \beta B(y, z)$ para todos $x, y, z \in V$ e para todos $\alpha, \beta \in \mathbb{K}$;

(ii) $B(x, \alpha y + \beta z) = \overline{\alpha} B(x, y) + \overline{\beta} B(x, z)$ para todos $x, y, z \in V$ e para todos $\alpha, \beta \in \mathbb{K}$.

Se $\mathbb{K} = \mathbb{R}$, dizemos simplesmente que $B$ é uma **forma bilinear**.

Uma forma bilinear é **simétrica** se $B(x, y) = B(y, x)$ para todos $x, y \in V$. Uma forma sesquilinear é **hermitiana** se $B(x, y) = \overline{B(y, x)}$ para todos $x, y \in V$.

**Teorema 21.** *Seja $V$ um espaço vetorial com produto interno de dimensão finita sobre $\mathbb{K}$ e $B$ uma forma sesquilinear. Então existe um único operador $T$ em $V$ tal que*

$$B(x, y) = \langle Tx, y \rangle .$$

*Esta correspondência determina um isomorfismo canônico entre o espaço das formas sesquilineares e o espaço dos operadores $\mathcal{L}(V)$.*

*Além disso, a forma $B$ é hermitiana se e somente se $T$ é auto-adjunto.*

**Prova:** Fixe um vetor $y \in V$. Então

$$f(x) = B(x, y)$$

é um funcional linear, logo pelo Teorema de Representação de Riesz existe um único $z \in V$ tal que

$$B(x, y) = \langle x, z \rangle$$

para todo $x \in V$. Defina uma aplicação $U : V \longrightarrow V$ por $Uy = z$. Afirmamos que $U$ é um operador linear. De fato,

$$\begin{aligned} \langle x, U(\beta_1 y_1 + \beta_2 y_2) \rangle &= B(x, \beta_1 y_1 + \beta_2 y_2) = \beta_1 B(x, y_1) + \overline{\beta}_2 B(x, y_2) \\ &= \beta_1 \langle x, U(y_1) \rangle + \overline{\beta}_2 \langle x, U(y_2) \rangle \\ &= \langle x, \beta_1 U(y_1) + \beta_2 U(y_2) \rangle \end{aligned}$$

para todo $x \in V$. Tomando $T = U^*$, segue que

$$B(x, y) = \langle x, Uy \rangle = \langle U^* x, y \rangle = \langle Tx, y \rangle .$$

O argumento usual mostra que $T$ é única.

Se $B$ é hermitiana, para todos $x, y \in V$ temos

$$\langle Tx, y \rangle = B(x, y) = \overline{B(y, x)} = \overline{\langle Ty, x \rangle} = \langle x, Ty \rangle ,$$

isto é,

$$T^* = T.$$

■

**Corolário 7.** *Se $\mathcal{B} = \{v_1, \ldots, v_n\}$ é uma base ortonormal para $V$, então toda forma sesquilinear pode ser escrita na forma*

$$B(x, y) = B\left( \sum_{i=1}^{n} x_i v_i, \sum_{j=1}^{n} y_j v_j \right) = \sum_{i,j=1}^{n} a_{ij} x_i \overline{y}_j$$

*para alguns $a_{ij} \in \mathbb{K}$.*

**Prova:** Escreva

$$B(x,y) = \langle Tx, y\rangle = \left\langle T\left(\sum_{i=1}^{n} x_i v_i\right), \sum_{j=1}^{n} y_j v_j \right\rangle = \sum_{i,j=1}^{n} x_i \overline{y}_j \langle Tv_i, v_j\rangle$$

e defina $a_{ij} = \langle Tv_i, v_j\rangle$. ■

**Definição 16.** Seja $V$ um espaço vetorial com produto interno de dimensão finita sobre $\mathbb{K}$ e $B$ uma forma sesquilinear. A função $q : V \longrightarrow \mathbb{K}$ definida por

$$q(x) = B(x,x)$$

é chamada a **forma quadrática** induzida por $B$.

Se $B$ é hermitiana [simétrica], dizemos que a forma quadrática é **hermitiana** [**simétrica**].

Segue do Corolário 7 que se $\mathcal{B} = \{v_1, \ldots, v_n\}$ é uma base ortonormal para $V$, toda forma quadrática pode ser escrita na forma

$$q(x) = q(x_1, \ldots, x_n) = q\left(\sum_{i=1}^{n} x_i v_i\right) = \sum_{i,j=1}^{n} a_{ij} x_i \overline{x}_j$$

para alguns $a_{ij} \in \mathbb{K}$. Observe que quando $B$ é hermitiana, $q(x) \in \mathbb{R}$ para todo $x \in V$, pois $B(x,x) = \overline{B(x,x)}$ implica $B(x,x) \in \mathbb{R}$. Vale a recíproca:

**Proposição 16.** *Seja $V$ um espaço vetorial complexo com produto interno. Então uma forma sesquilinear $B$ é hermitiana se e somente se $B(x,x) \in \mathbb{R}$ para todo $x \in V$.*

**Prova:** Assuma $B(x,x) \in \mathbb{R}$ para todo $x \in V$. Dados $x, y \in V$ precisamos mostrar que $B(x,y) = \overline{B(y,x)}$. Temos

$$B(x+y, x+y) = B(x,x) + B(x,y) + B(y,x) + B(y,y).$$

Como $B(x+y, x+y), B(x,x), B(y,y) \in \mathbb{R}$ segue que

$$B(x,y) + B(y,x) \in \mathbb{R}.$$

Da mesma forma, escrevendo

$$B(x+iy, x+iy) = B(x,x) - iB(x,y) + iB(y,x) - B(y,y),$$

concluímos que

$$-iB(x,y) + iB(y,x) \in \mathbb{R}.$$

Números reais são iguais a seus conjugados, logo

$$\begin{aligned} B(x,y) + B(y,x) &= \overline{B(x,y)} + \overline{B(y,x)}, \\ -iB(x,y) + iB(y,x) &= i\overline{B(x,y)} - i\overline{B(y,x)}. \end{aligned}$$

Somanda a primeira equação à segunda multiplicada por $i$, obtemos

$$2B(x,y) = 2\overline{B(y,x)},$$

donde segue o resultado. ■

Formas quadráticas hermitianas podem ser expressas em formas diagonais, isto é, formas que não possuem produtos mistos:

**Teorema 22.** (Diagonalização de Formas Quadráticas) *Seja $V$ um espaço vetorial com produto interno de dimensão $n$ sobre $\mathbb{K}$. Dada uma forma quadrática hermitiana $q$ em $V$, existe uma base ortonormal $\mathcal{B} = \{v_1, \ldots, v_n\}$ para $V$ tal que*

$$q(x) = q(x_1, \ldots, x_n) = q\left(\sum_{i=1}^{n} x_i v_i\right) = \sum_{i=1}^{n} \lambda_i x_i^2$$

*para alguns $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Além disso, se $B$ é a forma hermitiana que induz $q$, temos também*

$$B(x, y) = B\left(\sum_{i=1}^{n} x_i v_i, \sum_{j=1}^{n} y_j v_j\right) = \sum_{i=1}^{n} \lambda_i x_i \overline{y}_j.$$

**Prova:** Seja $B$ a forma sesquilinear hermitiana tal que $q(x) = B(x, x)$ e $T$ o operador linear tal que $B(x, y) = \langle Tx, y \rangle$ para todos $x, y \in V$. Pelo Teorema 19 $T$ é auto-adjunto, logo existe uma base ortonormal $\mathcal{B} = \{v_1, \ldots, v_n\}$ para $V$ que diagonaliza $T$, isto é,

$$Tv_j = \lambda_j v_j$$

com $\lambda_j \in \mathbb{R}$, para $j = 1, \ldots, n$. Segue que

$$\begin{aligned} B\left(\sum_{i=1}^{n} x_i v_i, \sum_{j=1}^{n} y_j v_j\right) &= \left\langle T\left(\sum_{i=1}^{n} x_i v_i\right), \sum_{j=1}^{n} y_j v_j \right\rangle = \sum_{i,j=1}^{n} x_i \overline{y}_j \langle Tv_i, v_j \rangle \\ &= \sum_{i,j=1}^{n} x_i \overline{y}_j \lambda_i \langle v_i, v_j \rangle = \sum_{i=1}^{n} \lambda_i x_i \overline{y}_j. \end{aligned}$$

A expressão para $q$ é obtida imediatamente a partir da expressão para $B$. ■

# Referências Bibliográficas


[1] BUENO, HAMILTON P., *Álgebra Linear: um segundo curso*, IMPA, 2006.

[2] HIRSCH, M. e SMALE, S., *Differential Equations, Dynamical Systems and Linear Algebra*, Academic Press, New York, 1974.

[3] HOFFMAN, K. e KUNZE, R., *Linear Algebra,* 2nd. Ed., Prentice Hall, 1971.

[4] LANG, S., *Álgebra Linear*, 3a. Ed., Editora Ciência Moderna, 2004.

[5] STRANG, G., *Linear Algebra and its Applications*, 3rd. Ed., Brooks Cole, 1988.