

Introdução à Álgebra Linear

Um Curso de Nivelamento

Jorge Delgado

Depto. de Matemática Aplicada

Katia Frensel

Depto. de Geometria

Instituto de Matemática
UFF

Março de 2005





Conteúdo

Noções Preliminares	1
1. Corpos	1
2. Sistemas de Equações Lineares	3
3. Operações Elementares	4
4. Multiplicação de Matrizes	14
5. Matrizes Invertíveis	19
Espaços Vetoriais	29
1. Espaço Vetorial - definições e propriedades básicas	29
2. Subespaço Vetorial	33
3. Bases e Dimensão	38
4. Coordenadas	50
5. Equivalência por Linhas – resumo	55
Transformações Lineares	61
1. Transformação Linear - noções básicas	61
2. Funcionais Lineares	83
Polinômios	97
1. Álgebras - generalidades	97
2. Interpolação de Lagrange	104
3. Ideais de Polinômios	108
4. Decomposição de um polinômio em fatores primos	117

5. Determinantes	122
6. Propriedades dos Determinantes	134
Formas Canônicas - Preliminares	139
1. Formas Canônicas Elementares	139
2. Polinômios Anuladores	149
3. Subespaços Invariantes	156
4. Triangulação Simultânea e Diagonalização Simultânea	167
5. Decomposição em Soma Direta	170
6. Somas Diretas Invariantes	178
7. O Teorema da Decomposição Primária	183
Forma Canônica Racional	193
1. Subespaços cíclicos e anuladores	193
2. Decomposição Cíclica e a Forma Racional	197
Forma Canônica de Jordan	217
1. Forma Canônica Racional dos Operadores Nilpotentes	217
2. Cálculo da Forma Canônica dos Operadores Nilpotentes	219
3. Forma Canônica de Jordan	233
4. Forma Canônica de Jordan Real	242
5. Operadores Semi-Simples	264
Espaços Vetoriais com Produto Interno	269
1. Produto Interno - Definições básicas	269
2. Funcionais Lineares e Adjuntos	288
3. Operadores Unitários	298
4. Operadores Normais	307
5. Diagonalização simultânea de uma família comutativa de operadores normais	336
6. Formas	341

7. Formas Positivas	348
8. Resumo sobre matrizes positivas	354
9. Funções de operadores	356



Noções Preliminares

1. Corpos

Um *corpo comutativo* é um conjunto $\mathbb{K}$, cujos elementos são chamados *escalares*, com um par de operações:

$$\begin{aligned} + \quad & x \in \mathbb{K} \text{ e } y \in \mathbb{K} \implies x + y \in \mathbb{K} && \text{(adição)} \\ \cdot \quad & x \in \mathbb{K} \text{ e } y \in \mathbb{K} \implies xy = x \cdot y \in \mathbb{K} && \text{(multiplicação)} \end{aligned}$$

que satisfazem as seguintes propriedades:

1. A adição é *comutativa*: $x + y = y + x$, $\forall x, y \in \mathbb{K}$.
2. A adição é *associativa*: $x + (y + z) = (x + y) + z$, $\forall x, y, z \in \mathbb{K}$.
3. Existe um único elemento $0 \in \mathbb{K}$ (*zero*), tal que $x + 0 = x$, $\forall x \in \mathbb{K}$.
4. A cada $x \in \mathbb{K}$ corresponde um único elemento $-x \in \mathbb{K}$, tal que $x + (-x) = 0$.
5. A multiplicação é *comutativa*: $x \cdot y = y \cdot x$, $\forall x, y \in \mathbb{K}$.
6. A multiplicação é *associativa*: $x \cdot (y \cdot z) = (x \cdot y) \cdot z$, $\forall x, y, z \in \mathbb{K}$.
7. Existe um único elemento $1 \in \mathbb{K} - \{0\}$ (*um*), tal que $x \cdot 1 = x$, $\forall x \in \mathbb{K}$.
8. A cada $x \in \mathbb{K} - \{0\}$ (x não-nulo) corresponde um único elemento x^{-1} ou $\frac{1}{x}$ em $\mathbb{K}$, tal que $x \cdot x^{-1} = 1$.
9. A multiplicação é *distributiva* em relação à adição:
$$x \cdot (y + z) = x \cdot y + x \cdot z, \quad \forall x, y, z \in \mathbb{K}.$$

Para saber mais ...
Da história da Álgebra Linear, consulte as páginas:
http://www-history.mcs.st-andrews.ac.uk/history/HistTopics/Matrices_and_determinants.html
e
http://www-history.mcs.st-andrews.ac.uk/history/HistTopics/Abstract_linear_spaces.html

Definição 1.1

Um **subcorpo de** $\mathbb{K}$ é um subconjunto $\mathcal{F}$ de $\mathbb{K}$ que é um corpo com as operações de adição e multiplicação de $\mathbb{K}$.

Assim, dizer que $\mathcal{F}$ é um subcorpo do corpo $\mathbb{K}$ significa que 0 e 1 estão em $\mathcal{F}$ e que se x e y são elementos de $\mathcal{F}$, então, $x + y$, xy , $-x$ e x^{-1} (caso $x \neq 0$) são também elementos de $\mathcal{F}$.

Exemplo 1.1

O conjunto dos números complexos $\mathbb{C} = \{a + ib \mid a, b \in \mathbb{R}\}$ é um corpo com as operações de adição e multiplicação usuais. $\square$

Exemplo 1.2

O conjunto dos números reais $\mathbb{R}$ é um subcorpo de $\mathbb{C}$. $\square$

Exemplo 1.3

O conjunto dos números racionais $\mathbb{Q} = \{\frac{p}{q} \mid p \in \mathbb{Z}, q \in \mathbb{Z} - \{0\}\}$ é um subcorpo de $\mathbb{R}$ e de $\mathbb{C}$. Mais ainda, é fácil verificar que todo subcorpo de $\mathbb{C}$ deve conter $\mathbb{Q}$. $\square$

Exemplo 1.4

O conjunto $\{x + y\sqrt{2} \mid x, y \in \mathbb{Q}\}$ é um subcorpo de $\mathbb{C}$. $\square$

Observação 1.1

Num corpo $\mathbb{K}$ pode ser possível adicionar uma quantidade finita de parcelas iguais a 1 e obter 0, isto é, $1 + 1 + 1 + \dots + 1 = 0$.

Quando esse fato acontece, o menor natural k , tal que a soma de k parcelas 1 é igual a 0, é chamado a **característica** de $\mathbb{K}$. Quando tal número natural não existe, dizemos que o corpo tem **característica zero**.

Em todo o seguinte, vamos considerar apenas corpos de característica zero.

Exemplos ...
Os corpos $\mathbb{Z}_p = \mathbb{Z}/(p\mathbb{Z})$ com $p > 0$ inteiro primo têm característica p .

2. Sistemas de Equações Lineares

Seja $\mathbb{K}$ um corpo (como sempre, de característica zero) e seja o *sistema de m equações lineares a n incógnitas*

$$\begin{array}{ccccccc} A_{11}x_1 & + & A_{12}x_2 & + & \cdots & + & A_{1n}x_n & = & y_1 \\ A_{21}x_1 & + & A_{22}x_2 & + & \cdots & + & A_{2n}x_n & = & y_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ A_{m1}x_1 & + & A_{m2}x_2 & + & \cdots & + & A_{mn}x_n & = & y_m, \end{array} \quad (*)$$

onde $y_1, \dots, y_m$ e A_{ij} , $1 \leq i \leq m$ e $1 \leq j \leq n$, são elementos de $\mathbb{K}$.

Definição 2.1

- Uma n -upla $(x_1, x_2, \dots, x_n) \in \mathbb{K}^n$ que satisfaz o sistema é chamada uma *solução do sistema*.
- Se $(y_1, y_2, \dots, y_m) = (0, 0, \dots, 0)$, dizemos que o sistema é *homogêneo*.
- se $c_1, c_2, \dots, c_m \in \mathbb{K}$, dizemos que

$$\begin{aligned} & (c_1A_{11} + c_2A_{21} + \cdots + c_mA_{m1})x_1 \\ & + (c_1A_{12} + c_2A_{22} + \cdots + c_mA_{m2})x_2 \\ & + \cdots + (c_1A_{1n} + c_2A_{2n} + \cdots + c_mA_{mn})x_n = c_1y_1 + c_2y_2 + \cdots + c_my_m \end{aligned}$$

é uma *combinação linear* das equações do sistema.

Note que toda solução do sistema é, também, uma solução desta nova equação.

- Dizemos que dois sistemas de equações lineares são *equivalentes* se cada equação de um dos sistemas é combinação linear das equações do outro sistema.

Em particular, sistemas de equações lineares que são equivalentes possuem exatamente as mesmas soluções.

3. Operações Elementares

Podemos escrever o sistema $(\star)$ na forma matricial

$$AX = Y,$$

onde

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{pmatrix}_{m \times n}$$

é a *matriz dos coeficientes* do sistema,

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}_{n \times 1} \quad \text{e} \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}_{m \times 1}.$$

Vamos considerar três *operações elementares* sobre as linhas da matriz A que correspondem a formar combinações lineares das equações do sistema $AX = Y$:

1. multiplicação de uma linha de A por um escalar não-nulo;
2. substituição da p -ésima linha de A pelo resultado da soma da linha p -ésima mais c vezes a linha q -ésima, sendo $c \in \mathbb{K} - \{0\}$ e $p \neq q$ números entre 1 e m ;
3. transposição de duas linhas de A .

Uma operação elementar sobre linhas é uma função $\mathbf{e}$ que a cada matriz A de tamanho $m \times n$ associa outra matriz $\mathbf{e}(A)$ de mesmo tamanho. Para as três operações elementares acima, temos:

1. $\mathbf{e}(A)_{ij} = A_{ij}$ se $i \neq p$, $\mathbf{e}(A)_{pj} = c A_{pj}$.
2. $\mathbf{e}(A)_{ij} = A_{ij}$ se $i \neq p$, $\mathbf{e}(A)_{pj} = A_{pj} + c A_{qj}$.
3. $\mathbf{e}(A)_{ij} = A_{ij}$ se $i \neq p$, $i \neq q$, $\mathbf{e}(A)_{pj} = A_{qj}$ e $\mathbf{e}(A)_{qj} = A_{pj}$.

Teorema 3.1

Para cada operação elementar $\mathbf{e}$ existe uma operação elementar $\mathbf{e}_1$ do mesmo tipo, tal que, $\mathbf{e}_1(\mathbf{e}(A)) = \mathbf{e}(\mathbf{e}_1(A)) = A$.

Prova.

De fato,

- Se $\mathbf{e}$ é a multiplicação da p -ésima linha por c ($c \neq 0$), então $\mathbf{e}_1$ é a multiplicação da p -ésima linha por c^{-1} .
- Se $\mathbf{e}$ é a operação que substitui a p -ésima linha de A pela p -ésima linha mais c vezes a q -ésima linha, então $\mathbf{e}_1$ é a operação que substitui a p -ésima linha de A pela p -ésima linha de A mais $-c$ vezes a q -ésima linha de A .
- Se $\mathbf{e}$ é a operação que transpõe as linhas p e q de A , então $\mathbf{e}_1 = \mathbf{e}$. ■

Definição 3.1

Sejam A e B matrizes $m \times n$ sobre o corpo $\mathbb{K}$. Dizemos que B é *equivalente por linhas* a A se B pode ser obtida de A por uma seqüência finita de operações elementares sobre as linhas de A .

Observação 3.1

A equivalência por linhas no conjunto das matrizes $m \times n$ é uma relação de equivalência.

Teorema 3.2

Se A e B são matrizes $m \times n$ equivalentes por linhas, então os sistemas homogêneos $AX = 0$ e $BX = 0$ têm exatamente as mesmas soluções.

Prova.

Podemos passar de A para B por meio de uma seqüência finita de operações elementares sobre as linhas.

$$A = A_0 \rightarrow A_1 \rightarrow A_2 \rightarrow \cdots \rightarrow A_k = B.$$

Basta então observar que uma operação elementar sobre linhas não altera o conjunto de soluções.

De fato. Sejam C e D matrizes $m \times n$ tais que $\mathbf{e}(C) = D$.

Então, toda linha de D é uma combinação linear das linhas de C . Como $C = \mathbf{e}_1(\mathbf{e}(C)) = \mathbf{e}_1(D)$, toda linha de C é, também, uma combinação linear

das linhas de D. Logo, C e D são matrizes equivalentes e, portanto, os sistemas homogêneos associados a tais matrizes possuem as mesmas soluções. ■

Exemplo 3.1

Seja A a matriz 3×4

$$A = \begin{pmatrix} 2 & -1 & 3 & 2 \\ 1 & 4 & 0 & -1 \\ 2 & 6 & -1 & 5 \end{pmatrix}$$

sobre o corpo $\mathbb{Q}$.

Vamos efetuar uma seqüência finita de operações elementares sobre as linhas de A:

$$\begin{aligned} A &= \begin{pmatrix} 2 & -1 & 3 & 2 \\ 1 & 4 & 0 & -1 \\ 2 & 6 & -1 & 5 \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} 0 & -9 & 3 & 4 \\ 1 & 4 & 0 & -1 \\ 2 & 6 & -1 & 5 \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} 0 & -9 & 3 & 4 \\ 1 & 4 & 0 & -1 \\ 0 & -2 & -1 & 7 \end{pmatrix} \xrightarrow{(1)} \\ &\begin{pmatrix} 0 & -9 & 3 & 4 \\ 1 & 4 & 0 & -1 \\ 0 & 1 & 1/2 & -7/2 \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} 0 & -9 & 3 & 4 \\ 1 & 0 & -2 & 13 \\ 0 & 1 & 1/2 & -7/2 \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} 0 & 0 & 15/2 & -55/2 \\ 1 & 0 & -2 & 13 \\ 0 & 1 & 1/2 & -7/2 \end{pmatrix} \xrightarrow{(1)} \\ &\begin{pmatrix} 0 & 0 & 1 & -11/3 \\ 1 & 0 & -2 & 13 \\ 0 & 1 & 1/2 & -7/2 \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} 0 & 0 & 1 & -11/3 \\ 1 & 0 & 0 & 17/3 \\ 0 & 1 & 1/2 & -7/2 \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} 0 & 0 & 1 & -11/3 \\ 1 & 0 & 0 & 17/3 \\ 0 & 1 & 0 & -5/3 \end{pmatrix} = B \end{aligned}$$

Como as matrizes 3×4 A e B são equivalentes por linhas, os sistemas lineares homogêneos

$$\begin{cases} 2x_1 - x_2 + 3x_3 + 2x_4 = 0 \\ x_1 + 4x_2 - x_4 = 0 \\ 2x_1 + 6x_2 - x_3 + 5x_4 = 0 \end{cases}$$

e

$$\begin{cases} x_3 - \frac{11}{3}x_4 = 0 \\ x_1 + \frac{17}{3}x_4 = 0 \\ x_2 - \frac{5}{3}x_4 = 0 \end{cases}$$

têm as mesmas soluções. Logo, toda solução do sistema é da forma

$$\left(-\frac{17}{3}x_4, \frac{5}{3}x_4, \frac{11}{3}x_4, x_4\right), \quad x_4 \in \mathbb{R}.$$

Ou seja, toda solução é um múltiplo da solução $(-17, 5, 11, 3)$. $\square$

Definição 3.2

Uma matriz R de tamanho $m \times n$ é dita *linha-reduzida* ou *reduzida por linhas* se:

- a. a primeira entrada não-nula de cada linha não-nula de R é igual a 1;
- b. toda coluna de R que contém a primeira entrada não-nula de alguma linha tem todos as suas outras entradas nulas;

Exemplo 3.2

A matriz B obtida no exemplo anterior é reduzida por linhas. $\square$

Exemplo 3.3

A matriz identidade $n \times n$ I cujas entradas são:

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

Em alguns textos ...
A matriz identidade é também chamada matriz unidade, pois, como veremos mais adiante, atua como unidade multiplicativa para as matrizes.

$\square$

Exemplo 3.4

As matrizes

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} 0 & 2 & 1 \\ 1 & 0 & -3 \\ 0 & 0 & 0 \end{pmatrix}$$

não são reduzidas por linhas. $\square$

Teorema 3.3

Toda matriz $m \times n$ com entradas sobre um corpo $\mathbb{K}$ é equivalente por linhas a uma matriz reduzida por linhas.

Prova.

Seja A uma matriz $m \times n$. Se toda entrada na primeira linha de A é 0, a condição **a.** está satisfeita no que diz respeito à primeira linha.

Se a primeira linha tem uma entrada não-nula, seja k o menor inteiro positivo j para o qual $A_{1j} \neq 0$. Multiplicando a linha 1 por A_{1j}^{-1} , a condição

a. fica satisfeita em relação à linha 1.

Para cada $i \geq 2$, somamos $-A_{ik}$ vezes a linha 1 à linha i , para tornar as outras entradas na k -ésima coluna iguais a zero.

Consideremos agora a matriz que resultou das operações acima. Se toda entrada da linha 2 é nula, nada fazemos a essa linha. Mas se alguma entrada da linha 2 é diferente de zero, multiplicamos essa linha por um escalar de modo que a primeira entrada não-nula seja 1.

No caso em que a linha 1 tenha uma primeira entrada não-nula na coluna k , a primeira entrada não-nula na linha 2 aparece numa coluna k' , onde $k' \neq k$. Somando múltiplos adequados da linha 2 às diversas linhas, podemos fazer com que todas as entradas da coluna k' , com exceção do 1 na linha 2, sejam nulas.

Observe que ao efetuarmos essas últimas operações, não alteramos as entradas da linha 1 nas colunas $1, \dots, k$ e nenhuma entrada da coluna k . Além disso, se a linha 1 fosse nula, as operações com a linha 2 não afetariam a linha 1.

Trabalhando com uma linha de cada vez da maneira acima, chegaremos a uma matriz reduzida por linhas após um número finito de etapas. ■

Definição 3.3

Uma matriz R de tamanho $m \times n$ é dita *linha-reduzida à forma em escada* ou *reduzida por linhas à forma em escada*, se:

- a. R é reduzida por linhas;
- b. toda linha de R cujas entradas são todas nulas ocorre abaixo de todas as linhas que possuem uma entrada não-nula;
- c. se as linhas $1, \dots, r$ são as linhas não-nulas de R e se a primeira entrada não-nula da linha i ocorre na coluna k_i , $i = 1, \dots, r$, então $k_1 < k_2 < \dots < k_r$.

Isto é, existe um inteiro positivo r , $1 \leq r \leq m$ e r inteiros positivos $k_1, k_2, \dots, k_r$ com $1 \leq k_i \leq n$, tais que:

- a. $R_{ij} = 0$, para $i > r$ e $1 \leq j \leq n$;
- b. $R_{ij} = 0$, para $j < k_i$ e $1 \leq i \leq r$;

c. $R_{ik_j} = \delta_{ij}$, para $1 \leq i \leq r$ e $1 \leq j \leq r$;

d. $k_1 < \dots < k_r$.

Exemplo 3.5

A matriz identidade $n \times n$ e a matriz nula $m \times n$ são matrizes reduzidas por linhas à forma em escada. $\square$

Exemplo 3.6

A matriz

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

é uma matriz reduzida por linhas à forma em escada. $\square$

Teorema 3.4

Toda matriz é equivalente por linhas a uma matriz reduzida por linhas à forma em escada.

Prova.

Sabemos que toda matriz A é equivalente por linhas a uma matriz reduzida por linhas R . Efetuando um número finito de transposições das linhas de R , obtemos uma matriz equivalente por linhas a R , e portanto a A , que é reduzida por linhas à forma em escada. $\blacksquare$

Exemplo 3.7

A matriz A do exemplo 3.1 é equivalente por linhas à matriz B que está na forma reduzida por linhas. Essa matriz é, por sua vez, equivalente por linhas à matriz

$$C = \begin{pmatrix} 1 & 0 & 0 & 17/3 \\ 0 & 1 & 0 & -5/3 \\ 0 & 0 & 1 & -11/3 \end{pmatrix}$$

que é a forma reduzida por linhas à forma em escada. $\square$

Seja R uma matriz $m \times n$ reduzida por linhas à forma em escada e sejam $1, \dots, r$ as linhas não-nulas de R . Suponhamos que o primeiro elemento não-nulo da linha i , $1 \leq i \leq r$, ocorra na coluna k_i .

Então, o sistema $RX = 0$ consiste de r equações não-triviais. Além disso, a incógnita x_{k_i} aparecerá, com coeficiente não-nulo, apenas na i -ésima equação. Se indicarmos por $u_1, \dots, u_{n-r}$ as $n - r$ incógnitas que são diferentes de $x_{k_1}, \dots, x_{k_r}$, então as r equações não triviais do sistema são da forma:

$$\begin{array}{rcl} x_{k_1} & + & \sum_{j=1}^{n-r} C_{1j} u_j = 0 \\ \vdots & & \vdots \\ x_{k_r} & + & \sum_{j=1}^{n-r} C_{rj} u_j = 0 \end{array} \quad (1)$$

Todas as soluções do sistema $RX = 0$ são obtidas atribuindo valores arbitrários às incógnitas $u_1, \dots, u_{n-r}$ e calculando os valores correspondentes de $x_{k_1}, \dots, x_{k_r}$ por meio de (1).

Exemplo 3.8

Seja R a seguinte matriz 3×5 reduzida por linhas à forma em escada

$$R = \begin{pmatrix} 0 & 1 & -3 & 0 & 1/2 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Neste caso, temos $r = 2$, $k_1 = 2$ e $k_2 = 4$.

As equações não-triviais do sistema $RX = 0$ são

$$\begin{array}{lcl} x_2 - 3x_3 + \frac{1}{2}x_5 = 0 & \iff & x_2 = 3x_3 - \frac{1}{2}x_5 \\ x_4 + 2x_5 = 0 & \iff & x_4 = -2x_5 \end{array}$$

Atribuindo valores arbitrários $x_1 = a$, $x_3 = b$ e $x_5 = c$, vemos que qualquer solução do sistema $RX = 0$ é da forma $(a, 3b - \frac{1}{2}c, b, -2c, c)$. $\square$

Observação 3.2

Seja R uma matriz $m \times n$ reduzida por linhas à forma em escada. Se o número r de linhas não-nulas é menor que n , o sistema $RX = 0$ possui soluções não-triviais, isto é, soluções $(x_1, \dots, x_n)$ em que nem todo x_j é nulo.

Teorema 3.5

Se A é uma matriz $m \times n$ com $m < n$, então o sistema homogêneo $AX = 0$ admite soluções não-triviais.

Prova.

Seja R uma matriz $m \times n$ reduzida por linhas à forma em escada que seja equivalente por linhas à matriz A . Então, os sistemas $AX = 0$ e $RX = 0$ possuem as mesmas soluções.

Se r é o número de linhas não-nulas em R , então $r \leq m$. Como $m < n$, temos $r < n$. Logo, pela observação acima, $RX = 0$ e, portanto, $AX = 0$ admite soluções não-triviais. ■

Teorema 3.6

Se A é uma matriz quadrada $n \times n$, então, A é equivalente por linhas à matriz identidade $n \times n$ se, e somente se, o sistema $AX = 0$ possui apenas a solução trivial.

Prova.

($\Rightarrow$) Se A é equivalente por linhas à matriz identidade I , então, os sistemas $AX = 0$ e $IX = 0$ possuem as mesmas soluções. Logo, $AX = 0$ possui apenas a solução trivial.

($\Leftarrow$) Suponhamos agora que $AX = 0$ possua apenas a solução trivial $X = 0 = (0, 0, \dots, 0)$.

Seja R uma matriz reduzida por linhas à forma em escada equivalente por linhas à matriz A e seja r o número de linhas não-nulas de R .

Como $RX = 0$ possui apenas a solução trivial, temos que $r \geq n$. Mas como R possui n linhas, temos que $r = n$.

Além disso, como $k_1 < k_2 < \dots < k_n$ e R possui n colunas, temos que $k_i = i$ para cada $1 \leq i \leq n$.

Logo, R é a matriz identidade. ■

Consideremos agora o sistema não-homogêneo de equações lineares $AX = Y$.

A matriz completa A' do sistema $AX = Y$ é a matriz $m \times (n+1)$ cujas n primeiras colunas são as colunas de A e cuja última coluna é Y . Isto é,

$$A'_{ij} = A_{ij}, \quad \text{se } j \leq n$$

$$A'_{i(n+1)} = Y_i.$$

Efetuada uma seqüência finita de operações elementares sobre as linhas de A , obtemos uma matriz R reduzida por linhas à forma em escada. Efetuando a mesma seqüência de operações elementares sobre a matriz completa A' , obteremos uma matriz R' cujas primeiras n colunas são as colunas de R e cuja última coluna

$$Z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{pmatrix}$$

resulta de aplicar a mesma seqüência de operações sobre as linhas da matriz Y .

Da mesma forma que no caso homogêneo, podemos mostrar que os sistemas $AX = Y$ e $RX = Z$ são equivalentes e, portanto, admitem as mesmas soluções.

Se R possui r linhas não-nulas, com a primeira entrada não-nula da linha i ocorrendo na coluna k_i , $1 \leq i \leq r$, então as r primeiras equações de $RX = Y$ exprimem $x_{k_1}, \dots, x_{k_r}$ em termos das $(n - r)$ incógnitas restantes e dos escalares $z_1, \dots, z_r$. Além disso, as $(m - r)$ últimas equações são:

$$\begin{aligned} 0 &= z_{r+1} \\ 0 &= z_{r+2} \\ &\vdots \\ 0 &= z_m. \end{aligned}$$

Portanto, a condição para o sistema ter uma solução é que $z_i = 0$ para $r < i \leq m$.

Se tal condição é satisfeita, todas as soluções do sistema são obtidas atribuindo-se valores arbitrários às $(n - r)$ incógnitas x_i , $1 \leq i \leq n$, $i \notin \{k_1, \dots, k_r\}$, e calculando x_{k_i} a partir da i -ésima equação.

Exemplo 3.9

Seja $\mathbb{K}$ o corpo dos números racionais, e seja A a matriz

$$A = \begin{pmatrix} 1 & -2 & 1 \\ 2 & 1 & 1 \\ 0 & 5 & -1 \end{pmatrix}.$$

Para resolver o sistema $AX = Y$, para $Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$ dado, devemos aplicar

uma sequência de operações elementares sobre as linhas da matriz completa A' que torne A reduzida por linhas:

$$\begin{pmatrix} 1 & -2 & 1 & y_1 \\ 2 & 1 & 1 & y_2 \\ 0 & 5 & -1 & y_3 \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} 1 & -2 & 1 & y_1 \\ 0 & 5 & -1 & y_2 - 2y_1 \\ 0 & 5 & -1 & y_3 \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} 1 & -2 & 1 & y_1 \\ 0 & 5 & -1 & y_2 - 2y_1 \\ 0 & 0 & 0 & y_3 - y_2 + 2y_1 \end{pmatrix}$$

$$\xrightarrow{(1)} \begin{pmatrix} 1 & -2 & 1 & y_1 \\ 0 & 1 & -1/5 & (y_2 - 2y_1)/5 \\ 0 & 0 & 0 & y_3 - y_2 + 2y_1 \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} 1 & 0 & 3/5 & (y_1 + 2y_2)/5 \\ 0 & 1 & -1/5 & (y_2 - 2y_1)/5 \\ 0 & 0 & 0 & y_3 - y_2 + 2y_1 \end{pmatrix}$$

A condição para que o sistema tenha uma solução é $y_3 - y_2 + 2y_1 = 0$.

Se os escalares y_1 , y_2 , e y_3 satisfazem essa condição, todas as soluções são obtidas atribuindo-se um valor c a x_3 e depois calculando:

$$\begin{aligned} x_1 &= \frac{3}{5}c + \frac{1}{5}(y_1 + 2y_2) \\ x_2 &= -\frac{1}{5}c + \frac{1}{5}(y_2 - 2y_1). \end{aligned}$$

□

Observação 3.3

Suponhamos que as entradas da matriz A e os escalares $y_1, \dots, y_m$ estejam num subcorpo $\mathcal{F}$ do corpo $\mathbb{K}$.

Se o sistema $AX = Y$ admite uma solução com $x_1, \dots, x_n$ em $\mathbb{K}$, ele admite uma solução com $x_1, \dots, x_n$ em $\mathcal{F}$.

De fato, a condição para o sistema admitir uma solução, sobre cada um dos dois corpos, é que sejam válidas certas relações entre $y_1, \dots, y_m$ em $\mathcal{F}$ (a saber as relações $z_i = 0$, para $i > r$).

Por exemplo, se $AX = Y$ é um sistema de equações lineares no qual os escalares y_k e A_{ij} são números reais e, se existe uma solução na qual $x_1, \dots, x_n$ são números complexos, então existe uma solução com $x_1, \dots, x_n$ números reais.

4. Multipliação de Matrizes

Definição 4.1

Seja A uma matriz $m \times n$ sobre o corpo $\mathbb{K}$ e seja B uma matriz $n \times p$ sobre o corpo $\mathbb{K}$. O **produto** AB é a matriz C de tamanho $m \times p$ sobre o corpo $\mathbb{K}$, cujas entradas são

$$C_{ij} = \sum_{k=1}^n A_{ik} B_{kj}.$$

Observação 4.1

Se $\beta_1, \dots, \beta_n$ são as linhas da matriz B e $\gamma_1, \dots, \gamma_m$ são as linhas da matriz C , então

$$\gamma_i = A_{i1}\beta_1 + A_{i2}\beta_2 + \dots + A_{in}\beta_n.$$

Isto é, as linhas de C são combinações lineares das linhas de B .

De fato,

$$\begin{aligned} \gamma_i &= (C_{i1}, C_{i2}, \dots, C_{ip}) \\ &= \left(\sum_{k=1}^n A_{ik} B_{k1}, \sum_{k=1}^n A_{ik} B_{k2}, \dots, \sum_{k=1}^n A_{ik} B_{kp} \right) \\ &= \sum_{k=1}^n A_{ik} (B_{k1}, B_{k2}, \dots, B_{kp}) \\ &= \sum_{k=1}^n A_{ik} \beta_k. \end{aligned}$$

Exemplo 4.1

$$\text{a. } \begin{pmatrix} 5 & -1 & 2 \\ 0 & 7 & 2 \end{pmatrix}_{2 \times 3} = \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix}_{2 \times 2} \begin{pmatrix} 5 & -1 & 2 \\ 15 & 4 & 8 \end{pmatrix}_{2 \times 3}.$$

Neste caso:

$$\begin{aligned} \gamma_1 &= (5, -1, 2) = 1 \cdot (5, -1, 2) + 0 \cdot (15, 4, 8) \\ \gamma_2 &= (0, 7, 2) = -3 \cdot (5, -1, 2) + 1 \cdot (15, 4, 8). \end{aligned}$$

$$\text{b. } \begin{pmatrix} -2 & -4 \\ 6 & 12 \end{pmatrix}_{2 \times 2} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}_{2 \times 1} \begin{pmatrix} 2 & 4 \end{pmatrix}_{1 \times 2}.$$

Neste caso:

$$\begin{aligned} \gamma_1 &= (-2, -4) = -1 \cdot (2, 4) \\ \gamma_2 &= (6, 12) = 3 \cdot (2, 4). \end{aligned}$$

$$\text{c. } \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}_{3 \times 3} \begin{pmatrix} 1 & -5 & 2 \\ 2 & 3 & 4 \\ 9 & -1 & 3 \end{pmatrix}_{3 \times 3} = \begin{pmatrix} 2 & 3 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}_{3 \times 3}.$$

Neste caso:

$$\gamma_1 = (2, 3, 4) = 0 \cdot (1, -5, 2) + 1 \cdot (2, 3, 4) + 0 \cdot (9, -1, 3)$$

$$\gamma_2 = (0, 0, 0) = 0 \cdot (1, -5, 2) + 0 \cdot (2, 3, 4) + 0 \cdot (9, -1, 3)$$

$$\gamma_3 = (0, 0, 0) = 0 \cdot (1, -5, 2) + 0 \cdot (2, 3, 4) + 0 \cdot (9, -1, 3).$$

□

Observação 4.2

- O produto de matrizes não é comutativo.

De fato, considere o seguinte exemplo:

$$\text{Sejam } A = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \text{ e } B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Então

$$AB = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} \neq \begin{pmatrix} 4 & 1 \\ 3 & 1 \end{pmatrix} = BA.$$

- Se I é a matriz identidade $m \times m$ e A é uma matriz $m \times n$, então $IA = A$.
- Se I é a matriz identidade $n \times n$ e A é uma matriz $m \times n$, então $AI = A$.
- Se $\mathcal{O}^{k,m}$ é a matriz nula $k \times m$, então, para toda matriz A $m \times n$ e qualquer matriz B $p \times k$, temos $\mathcal{O}^{k,m}A = \mathcal{O}^{k,n}$ e $B\mathcal{O}^{k,m} = \mathcal{O}^{p,m}$.

Se B é uma matriz $n \times p$, as colunas de B são as $n \times 1$ matrizes $\beta_1, \dots, \beta_p$ dadas por:

$$\beta_j = \begin{pmatrix} B_{1j} \\ B_{2j} \\ \vdots \\ B_{nj} \end{pmatrix}_{n \times 1}, \quad 1 \leq j \leq p.$$

A matriz B é a sucessão dessas p colunas:

$$B = (\beta_1, \beta_2, \dots, \beta_p).$$

Seja A uma matriz $m \times n$. Então $A\beta_j$ é a j -ésima coluna da matriz $C = AB$, pois como

$$C_{ij} = \sum_{k=1}^n A_{ik} B_{kj},$$

então

$$C_j = \begin{pmatrix} C_{1j} \\ C_{2j} \\ \vdots \\ C_{mj} \end{pmatrix}_{m \times 1} = \begin{pmatrix} \sum_{k=1}^n A_{1k} B_{kj} \\ \sum_{k=1}^n A_{2k} B_{kj} \\ \vdots \\ \sum_{k=1}^n A_{mk} B_{kj} \end{pmatrix} = A \beta_j.$$

Assim, $AB = (A\beta_1, A\beta_2, \dots, A\beta_p)_{m \times p}$.

Teorema 4.1

Se A , B e C são matrizes sobre o corpo $\mathbb{K}$ tais que os produtos BC e $A(BC)$ são definidos, então estão definidos os produtos AB e $(AB)C$. Além disso,

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

Prova.

Suponhamos que B é uma matriz $n \times p$, que A é uma matriz com m linhas e que C é uma matriz com l colunas.

Como os produtos BC e $A(BC)$ estão definidos, temos que C é uma matriz $p \times l$ e que A é uma matriz $m \times n$, já que BC é uma matriz $n \times l$.

Assim, o produto AB existe e é uma matriz $m \times p$, e o produto $(AB)C$ existe e é uma matriz $m \times l$.

Além disso, usando a definição, temos:

$$\begin{aligned} [A(BC)]_{ij} &= \sum_{k=1}^n A_{ik} (BC)_{kj} = \sum_{k=1}^n A_{ik} \sum_{r=1}^p B_{kr} C_{rj} \\ &= \sum_{k=1}^n \sum_{r=1}^p A_{ik} B_{kr} C_{rj} = \sum_{r=1}^p \left(\sum_{k=1}^n A_{ik} B_{kr} \right) C_{rj} \\ &= \sum_{r=1}^p (AB)_{ir} C_{rj} \\ &= [(AB)C]_{ij}. \end{aligned}$$

Como queríamos demonstrar. ■

Observação 4.3

- Devido à propriedade associativa da multiplicação de matrizes, se A é uma matriz quadrada, o produto $A \cdot A \cdot \dots \cdot A$, de k fatores iguais a A , está definido sem ambigüidade e indicaremos esse produto por A^k , convencionando $A^0 = I$.
- A relação $A(BC) = (AB)C$ implica que combinações lineares de combinações lineares das linhas de C são, ainda, combinações lineares das linhas de C .

Definição 4.2

Uma matriz $m \times m$ é uma **matriz elementar** se ela pode ser obtida da matriz identidade $m \times m$ por meio de uma única operação elementar sobre as linhas.

Exemplo 4.2

As matrizes elementares 2×2 são:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}; \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}; \begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix} \text{ com } c \neq 0; \begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix} \text{ com } c \neq 0.$$

□

Teorema 4.2

Seja e uma operação elementar sobre linhas e seja E a matriz $m \times m$ elementar $E = e(I)$. Então, para toda matriz A $m \times n$, temos:

$$e(A) = EA.$$

Prova.

Analisemos, separadamente as três operações elementares.

- a.** Seja e a operação “substituição da linha r por c vezes a linha r , com $c \neq 0$ ”. Então,

$$E_{ik} = \begin{cases} \delta_{ik}, & \text{se } i \neq r \\ c\delta_{rk}, & \text{se } i = r \end{cases}$$

$$\text{Logo, } (EA)_{ij} = \sum_{k=1}^m E_{ik}A_{kj} = \begin{cases} A_{ij}, & \text{se } i \neq r \\ cA_{rj}, & \text{se } i = r \end{cases}, \text{ isto é, } EA = e(A).$$

- b.** Seja e a operação “transposição das linhas r e s , $r \neq s$ ”. Então,

$$E_{ik} = \begin{cases} \delta_{ik}, & \text{se } i \neq r \text{ e } i \neq s \\ \delta_{rk}, & \text{se } i = s \\ \delta_{sk}, & \text{se } i = r. \end{cases}$$

$$\text{Logo, } (EA)_{ij} = \sum_{k=1}^m E_{ik}A_{kj} = \begin{cases} A_{ij}, & \text{se } i \neq r \text{ e } j \neq s \\ A_{rj}, & \text{se } i = s \\ A_{sj}, & \text{se } i = r \end{cases}, \text{ isto é, } EA = \mathbf{e}(A).$$

c. Seja $\mathbf{e}$ a operação “substituição da linha r pela soma de c vezes a linha s mais a linha r ”. Então,

$$E_{ik} = \begin{cases} \delta_{ik}, & \text{se } i \neq r \\ \delta_{rk} + c\delta_{sk}, & \text{se } i = r. \end{cases}$$

$$\text{Logo, } (EA)_{ij} = \sum_{k=1}^m E_{ik}A_{kj} = \begin{cases} A_{ij}, & \text{se } i \neq r \\ A_{rj} + cA_{sj}, & \text{se } i = r \end{cases}, \text{ isto é, } EA = \mathbf{e}(A).$$

■

Corolário 4.1

Sejam A e B matrizes $m \times n$ sobre o corpo $\mathbb{K}$. Então, B é equivalente por linhas a A se, e somente se, $B = PA$, onde P é um produto de matrizes elementares $m \times m$.

Prova.

Suponhamos que $B = PA$, onde

$$P = E_s \cdot \dots \cdot E_2 \cdot E_1,$$

e as E_j são matrizes elementares $m \times m$. Então,

- E_1A é equivalente por linhas à matriz A ;
- $E_2E_1A = E_2(E_1A)$ é equivalente por linhas à matriz E_1A , e portanto a A ;

Continuando o raciocínio indutivamente, obtemos que $(E_s \cdot \dots \cdot E_1)A$ é equivalente por linhas à matriz A .

Suponhamos agora que B é equivalente por linhas à matriz A . Sejam $\mathbf{e}_1, \dots, \mathbf{e}_s$ operações elementares sobre linhas que aplicadas consecutivamente sobre a matriz A dão como resultado B . Se $E_1, \dots, E_s$ são as matrizes elementares correspondentes, então

$$B = \mathbf{e}_s(\mathbf{e}_{s-1}(\dots(\mathbf{e}_1(A))\dots)) = (E_s \cdot E_{s-1} \cdot \dots \cdot E_1) \cdot A.$$

Como queríamos demonstrar. ■

5. Matrizes Invertíveis

Definição 5.1

Seja A uma matriz quadrada, $n \times n$, sobre o corpo $\mathbb{K}$.

- uma matriz B $n \times n$ tal que $BA = I$ é dita uma *inversa de A pela esquerda*.
- uma matriz B $n \times n$ tal que $AB = I$ é dita uma *inversa de A pela direita*.
- uma matriz B $n \times n$ tal que $AB = BA = I$ é dita uma *inversa de A* . Neste caso, dizemos que A é *invertível*.

Lema 5.1

Se A possui uma inversa B pela esquerda e uma inversa C pela direita, então $B = C$ e A é invertível.

Prova.

Suponhamos que $BA = I$ e que $AC = I$. Então,

$$B = BI = B(AC) = (BA)C = IC = C.$$

Como $BA = AB = I$, A é invertível. ■

Observe que Se A possui inversa pela direita e pela esquerda, então A é invertível e a inversa (bilateral) é única.

Se A é invertível, designamos por A^{-1} a única matriz que é *inversa* (bilateral) de A .

Teorema 5.1

Sejam A e B matrizes $n \times n$ sobre o corpo $\mathbb{K}$.

1. Se A é invertível, então A^{-1} é invertível e $(A^{-1})^{-1} = A$.
2. Se A e B são invertíveis, então AB é invertível e $(AB)^{-1} = B^{-1}A^{-1}$.

Prova.

- Como $A^{-1}A = AA^{-1} = I$, temos que A^{-1} é invertível e $(A^{-1})^{-1} = A$.
- Como $(AB)(B^{-1}A^{-1}) = (B^{-1}A^{-1})(AB) = I$, temos que AB é invertível e $(AB)^{-1} = B^{-1}A^{-1}$. ■

Corolário 5.1

Um produto $A_1 A_2 \cdots A_k$, $k \in \mathbb{N}$, de matrizes invertíveis $n \times n$ é, também, invertível.

Teorema 5.2

As matrizes elementares são invertíveis e suas inversas são matrizes elementares do mesmo tipo.

Prova.

Seja E uma matriz elementar correspondente à operação elementar sobre linhas $\mathbf{e}$. Se $\mathbf{e}_1$ é a operação inversa de $\mathbf{e}$, e E_1 é a sua matriz elementar associada, então

$$\begin{aligned} EE_1 &= \mathbf{e}(E_1) = \mathbf{e}(\mathbf{e}_1(I)) = I, \\ E_1 E &= \mathbf{e}_1(E) = \mathbf{e}_1(\mathbf{e}(I)) = I, \end{aligned}$$

de modo que E é invertível e $E_1 = E^{-1}$. ■

Exemplo 5.1

As inversas das matrizes elementares 2×2 são:

$$\begin{aligned} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{-1} &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; & \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}^{-1} &= \begin{pmatrix} 1 & -c \\ 0 & 1 \end{pmatrix}; \\ \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}^{-1} &= \begin{pmatrix} 1 & 0 \\ -c & 1 \end{pmatrix}; & \begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix}^{-1} &= \begin{pmatrix} 1/c & 0 \\ 0 & 1 \end{pmatrix}, c \neq 0; \\ \begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix}^{-1} &= \begin{pmatrix} 1 & 0 \\ 0 & 1/c \end{pmatrix}, c \neq 0. \end{aligned}$$

□

Teorema 5.3

Se A é uma matriz $n \times n$, as seguintes afirmações são equivalentes:

1. A é invertível;
2. A é equivalente por linhas à matriz identidade;
3. A é um produto de matrizes elementares.

Prova.

Seja R uma matriz reduzida por linhas à forma em escada que seja equivalente por linhas à matriz A .

Então, existem matrizes elementares $E_1, \dots, E_k$, tais que,

$$R = E_k \cdot \dots \cdot E_1 \cdot A.$$

Como cada E_i , $1 \leq i \leq k$, é invertível, temos que,

$$A = E_1^{-1} \cdot \dots \cdot E_k^{-1} \cdot R.$$

Sendo o produto de matrizes invertíveis uma matriz invertível, temos que A é invertível se, e somente se, R é invertível.

Já que R é uma matriz quadrada que é reduzida por linhas à forma em escada, R é invertível se, e somente se, cada linha de R contém uma entrada não-nula, isto é, se, e somente se, $R = I$.

Assim, A é invertível se, e somente se, $R = I$, isto é, se, e somente se, A é equivalente por linhas à matriz identidade. Logo, as afirmações **1.** e **2.** são equivalentes.

Se $R = I$, temos que $A = E_1^{-1} \cdot \dots \cdot E_k^{-1}$, isto é, A é um produto de matrizes elementares. E se A é um produto de matrizes elementares, então A é invertível, pois toda matriz elementar é invertível. Logo, as afirmações **1.** e **3.** são equivalentes. ■

Corolário 5.2

Se A é uma matriz $n \times n$ invertível e se uma seqüência de operações elementares sobre linhas reduz A à matriz identidade I , então a mesma seqüência de operações sobre linhas, quando aplicada à matriz identidade I , produz A^{-1} .

Corolário 5.3

Sejam A e B matrizes $m \times n$. Então, B é equivalente por linhas a A se, e somente se, $B = PA$, onde P é uma matriz invertível $m \times m$.

Teorema 5.4

Seja A uma matriz $n \times n$. As seguintes afirmações são equivalentes:

- 1.** A é invertível.
- 2.** O sistema homogêneo $AX = 0$ possui apenas a solução trivial $X = 0$.

3. O sistema de equações $AX = Y$ possui uma única solução X para cada matriz $Y_{n \times 1}$.

Prova.

Pelo teorema 3.6, a afirmação 2. é equivalente a que a matriz A seja equivalente por linhas à matriz identidade I . Então, pelo teorema 5.3, as afirmações 1. e 2. são equivalentes.

Se A é inversível, a única solução do sistema $AX = Y$ é $X = A^{-1}Y$.

Suponhamos, agora, que $AX = Y$ possui uma solução para cada Y . Seja R uma matriz reduzida por linhas à forma em escada que seja equivalente por linhas à matriz A .

Para mostrar que $R = I$, basta mostrar que a última linha de R não é nula.

Seja $E = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ e seja P uma matriz $n \times n$ invertível, tal que $R = PA$. Como,

por hipótese, o sistema $AX = P^{-1}E$ tem solução, o sistema $RX = PAX = E$ tem solução. Logo, a última linha de R não é nula. Isto é, $R = I$.

Assim, pelo teorema 5.3, A é invertível. ■

Corolário 5.4

Uma matriz quadrada com inversa à esquerda ou à direita é invertível.

Prova.

Seja A uma matriz $n \times n$. Suponhamos que A possui uma inversa à esquerda, isto é, uma matriz B tal que $BA = I$. Então, o sistema $AX = 0$ possui somente a solução trivial, pois $X = IX = (BA)X = B(AX) = B0 = 0$.

Portanto, A é invertível.

Suponhamos agora que A possui uma inversa à direita, isto é, uma matriz C tal que $AC = I$. Então C possui uma inversa à esquerda e é, portanto, invertível, com inversa $A = C^{-1}$. Logo, A é invertível com inversa C . ■

Corolário 5.5

Seja $A = A_1 A_2 \cdots A_k$, onde $A_1, \dots, A_k$ são matrizes $n \times n$. Então, A é invertível se, e somente se, cada A_j é invertível.

Prova.

Já provamos que o produto de duas matrizes invertíveis é invertível. Então, por indução, podemos provar que o produto de um número finito de matrizes invertíveis é invertível.

Suponhamos agora que A é invertível e que X é uma solução do sistema $A_k X = 0$. Então, $AX = (A_1 A_2 \cdots A_{k-1}) A_k X = 0$.

Como A é invertível, temos que $X = 0$. Logo, o sistema $A_k X = 0$ possui apenas a solução trivial e A_k é, portanto, invertível.

Como $A_1 \cdots A_{k-1} = AA_k^{-1}$ é invertível, temos, usando o argumento anterior, que A_{k-1} é invertível.

Assim, podemos provar por indução, que cada matriz A_j é invertível. ■

Seja R uma matriz reduzida por linhas à forma em escada equivalente por linhas à matriz A e seja P uma matriz invertível tal que $R = PA$. Então, X é uma solução do sistema $AX = Y$ se, e somente se, X é uma solução do sistema $RX = PY$.

Seja A' a matriz completa do sistema $AX = Y$ com escalares arbitrários $y_1, \dots, y_m$ na última coluna.

Se efetuarmos sobre A' uma seqüência de operações elementares sobre linhas que reduza A a R , torna-se evidente o que é a matriz P .

Em particular, se A é uma matriz quadrada, esse processo mostrará se A é invertível ou não e, se for invertível, qual é a inversa P .

Exemplo 5.2

Seja $\mathbb{K}$ o corpo dos números racionais e seja $A = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}$.

Então,

$$\begin{pmatrix} 2 & -1 & y_1 \\ 1 & 3 & y_2 \end{pmatrix} \xrightarrow{(3)} \begin{pmatrix} 1 & 3 & y_2 \\ 2 & -1 & y_1 \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} 1 & 3 & y_2 \\ 0 & -7 & y_1 - 2y_2 \end{pmatrix} \\ \xrightarrow{(1)} \begin{pmatrix} 1 & 3 & y_2 \\ 0 & 1 & (-y_1 + 2y_2)/7 \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} 1 & 0 & (3y_1 + y_2)/7 \\ 0 & 1 & (-y_1 + 2y_2)/7 \end{pmatrix}$$

Então, A é inversível e $A^{-1} = \begin{pmatrix} 3/7 & 1/7 \\ -1/7 & 2/7 \end{pmatrix}$ ■

Como $P = E_k \cdot \dots \cdot E_1$ é o produto das matrizes elementares que correspondem às operações elementares efetuadas sobre as linhas de A para se obter R , podemos trabalhar com duas seqüências de matrizes, uma descrevendo a redução de A à matriz R e a outra registrando o efeito da mesma seqüência de operações sobre a matriz identidade para obter a matriz P .

Exemplo 5.3

Determinemos a matriz P tal que $R = PA$ é reduzida por linhas à forma em escada, onde

$$A = \begin{pmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{pmatrix}.$$

Aplicando operações elementares à matriz A e à matriz identidade simultaneamente, obtemos:

$$\begin{array}{ll}
 \begin{pmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{pmatrix} & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 \downarrow & \downarrow \\
 \begin{pmatrix} 1 & 1/2 & 1/3 \\ 0 & 1/12 & 1/12 \\ 0 & 1/12 & 1/45 \end{pmatrix} & \begin{pmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ -1/3 & 0 & 1 \end{pmatrix} \\
 \downarrow & \downarrow \\
 \begin{pmatrix} 1 & 1/2 & 1/3 \\ 0 & 1/12 & 1/12 \\ 0 & 0 & 1/180 \end{pmatrix} & \begin{pmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 1/6 & -1 & 1 \end{pmatrix} \\
 \downarrow & \downarrow \\
 \begin{pmatrix} 1 & 1/2 & 1/3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} & \begin{pmatrix} 1 & 0 & 0 \\ -6 & 12 & 0 \\ 30 & -180 & 180 \end{pmatrix} \\
 \downarrow & \downarrow \\
 \begin{pmatrix} 1 & 1/2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \begin{pmatrix} -9 & 60 & -60 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{pmatrix} \\
 \downarrow & \downarrow \\
 R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I & P = \begin{pmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{pmatrix}
 \end{array}$$

Logo, A é invertível e $P = A^{-1}$. $\square$



Espaços Vetoriais

1. Espaço Vetorial - definições e propriedades básicas

Definição 1.1

Um *espaço vetorial* sobre o corpo $\mathbb{K}$ é um conjunto E , cujos elementos são chamados *vetores*, e um par de operações:

$$\begin{aligned} \text{adição: } & + \quad x \in E, y \in E \implies x + y \in E \\ \text{multiplicação por escalares: } & \cdot \quad \lambda \in \mathbb{K}, x \in E \implies \lambda \cdot x \in E, \end{aligned}$$

Lembre que ...
Os elementos do corpo $\mathbb{K}$ são chamados *escalares*.

com as seguintes propriedades:

1. $(x + y) + z = x + (y + z)$, $\forall x, y, z \in E$ (propriedade associativa da adição).
2. $x + y = y + x$, $\forall x, y \in E$ (propriedade comutativa da adição).
3. Existe $\mathbf{0} \in E$, tal que $x + \mathbf{0} = x$, $\forall x \in E$ ($\mathbf{0}$ é o vetor nulo).
4. Para cada $x \in E$, existe $(-x) \in E$, tal que, $x + (-x) = \mathbf{0}$ (o vetor $(-x)$ é o simétrico ou inverso aditivo de x).
5. $(\lambda + \mu)x = \lambda x + \mu x$, $\forall x \in E, \forall \lambda, \mu \in \mathbb{K}$.
6. $\lambda(x + y) = \lambda x + \lambda y$, $\forall x, y \in E, \forall \lambda \in \mathbb{K}$.
7. $1 \cdot x = x$, $\forall x \in E$.
8. $(\lambda\mu)x = \lambda(\mu x)$, $\forall x \in E, \forall \lambda, \mu \in \mathbb{K}$.

Observação 1.1

1. O vetor nulo $\mathbf{0}$ é único, isto é, se $\mathbf{0}' \in E$ é tal que $\mathbf{0}' + x = x$, $\forall x \in E$, então $\mathbf{0}' = \mathbf{0}$.

De fato, $\mathbf{0}' = \mathbf{0}' + \mathbf{0} = \mathbf{0}$.

2. $\lambda \cdot \mathbf{0} = \mathbf{0}$, para todo $\lambda \in \mathbb{K}$.

Com efeito,

$$\begin{aligned}\lambda \mathbf{0} &= \lambda(\mathbf{0} + \mathbf{0}) = \lambda \mathbf{0} + \lambda \mathbf{0} \\ \Rightarrow \mathbf{0} &= \lambda \mathbf{0} + (-\lambda \mathbf{0}) = (\lambda \mathbf{0} + \lambda \mathbf{0}) + (-\lambda \mathbf{0}) = \lambda \mathbf{0} + (\lambda \mathbf{0} + (-\lambda \mathbf{0})) \\ \Rightarrow \mathbf{0} &= \lambda \mathbf{0} + \mathbf{0} = \lambda \mathbf{0}.\end{aligned}$$

3. $0 \cdot x = \mathbf{0}$, $\forall x \in E$.

De fato,

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

4. Sejam $\lambda \in \mathbb{K}$ e $x \in E$. Então, $\lambda x = \mathbf{0}$ se, e somente se, $\lambda = 0$ ou $x = \mathbf{0}$.

Para verificar essa afirmação, suponha que $\lambda x = \mathbf{0}$, com $\lambda \neq 0$. Então,

$$\mathbf{0} = \lambda^{-1}(\lambda x) = (\lambda^{-1}\lambda)x = 1 \cdot x = x.$$

5. Para cada $x \in E$ existe um único vetor $y \in E$, tal que $x + y = \mathbf{0}$.

De fato, $y = y + \mathbf{0} = y + (x + (-x)) = (y + x) + (-x) = \mathbf{0} + (-x) = -x$.

6. $(-1)x = -x$, $\forall x \in E$.

Para verificar essa afirmação, considere a seguinte cadeia de igualdades:

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

Definição 1.2

Designamos por $x - y$ o vetor $x + (-y)$.

Exemplo 1.1

1. Seja $\mathbb{K}$ um corpo. O conjunto

$$\mathbb{K}^n = \{(x_1, x_2, \dots, x_n) \mid x_1, x_2, \dots, x_n \in \mathbb{K}\},$$

é o espaço vetorial das n -úplas sobre $\mathbb{K}$, com as operações de adição e multiplicação por escalares de $\mathbb{K}$ definidas por:

Note que ...
0 representa o zero do corpo $\mathbb{K}$,
enquanto $\mathbf{0}$ é o vetor nulo do
espaço vetorial E .

$$\begin{aligned} + & : (x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n) \\ \cdot & : \lambda \cdot (x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n). \end{aligned}$$

Neste espaço, o vetor nulo é $\mathbf{0} = (0, \dots, 0)$.

Se $\mathbb{K} = \mathbb{R}$, então $\mathbb{K}^n = \mathbb{R}^n$, se $\mathbb{K} = \mathbb{C}$, então $\mathbb{K}^n = \mathbb{C}^n$ etc.

2. O conjunto das matrizes $m \times n$ com entradas no corpo $\mathbb{K}$:

$$\mathbb{K}^{m \times n} = \{A \mid A = (A_{ij})_{ij} \text{ é uma matriz com } A_{ij} \in \mathbb{K}\}$$

é um espaço vetorial com as operações:

$$\begin{aligned} + & : (A + B)_{ij} = A_{ij} + B_{ij} \\ \cdot & : (\lambda \cdot A)_{ij} = \lambda A_{ij}. \end{aligned}$$

Note que ...
 $\mathbb{K}^{1 \times n}$ se identifica com $\mathbb{K}^n$.

3. Seja S um conjunto não-vazio e $\mathbb{K}$ um corpo. O *conjunto das funções de S em $\mathbb{K}$* :

$$\mathcal{F}(S, \mathbb{K}) = \{f : S \rightarrow \mathbb{K}\}$$

é um espaço vetorial com as operações dadas por:

$$\begin{aligned} + & : (f + g)(p) = f(p) + g(p), \forall p \in S \\ \cdot & : (\lambda \cdot f)(p) = \lambda \cdot f(p), \forall \lambda \in \mathbb{K}, \forall p \in S. \end{aligned}$$

Neste caso, o vetor nulo é a função zero: $\mathcal{O}(p) = 0, \forall p \in S$ e, para cada $f \in \mathcal{F}(S, \mathbb{K})$, a função $-f \in \mathcal{F}(S, \mathbb{K})$ que a cada $p \in S$ faz corresponder o escalar $(-f)(p) = -(f(p))$, é o inverso aditivo de f .

Observe que os espaços vetoriais dos itens **1.** e **2.** são casos particulares deste exemplo (item **3.**). De fato,

- no item **1.** o conjunto S é $\{1, 2, \dots, n\}$;
- no item **2.** o conjunto S é $\{(i, j) \mid 1 \leq i \leq m, \text{ e } 1 \leq j \leq n\}$.

4. O conjunto das funções polinomiais na indeterminada x com coeficientes sobre o corpo $\mathbb{K}$

$$\mathbb{K}[x] = \{p(x) = c_0 + c_1x + \dots + c_nx^n \mid c_0, c_1, \dots, c_n \in \mathbb{K}, n \in \mathbb{N}\},$$

é um espaço vetorial com as operações definidas como no item **3.**

5. Se $\mathcal{F}$ é um subcorpo de $\mathbb{K}$, então $\mathbb{K}$ é, de maneira natural, um espaço vetorial sobre $\mathcal{F}$. Em particular, todo corpo é um espaço vetorial sobre si mesmo.

Por exemplo

Em geral ...
Você pode verificar que se S é um conjunto não-vazio e E é um espaço vetorial, então o conjunto $\mathcal{F}(S, E)$ que consiste das aplicações

$f : S \rightarrow E$
é um espaço vetorial com as operações de adição e multiplicação por escalares herdadas das operações correspondentes que fazem de E um espaço vetorial.

Atividade.

1. O conjunto $\mathbb{K}_n[x]$ que consiste das funções polinomiais na indeterminada x , grau menor ou igual a n e com coeficientes sobre o corpo $\mathbb{K}$ é um espaço vetorial?

2. O conjunto $\mathbb{K}[x, y]$ que consiste dos polinômios nas indeterminadas x e y com coeficientes no corpo $\mathbb{K}$ é um espaço vetorial? O que pode dizer sobre o caso em que o grau é menor ou igual a n ?

3. Generalize as conclusões de **2.** para o conjunto dos polinômios com uma quantidade finita de indeterminadas.

Justifique suas respostas.

- $\mathbb{C}$ é um espaço vetorial sobre $\mathbb{R}$,
- $\mathbb{R}$ é um espaço vetorial sobre $\mathbb{Q}$,
- $\mathbb{C}$ é um espaço vetorial sobre $\mathbb{Q}$,
- $\mathbb{C}$ é um espaço vetorial sobre $\mathbb{C}$ etc.

□

Observação 1.2

As propriedades associativa e comutativa da adição de vetores implicam que uma soma envolvendo um certo número de vetores independe da maneira pela qual estes vetores estão combinados ou associados. Logo, se $v_1, \dots, v_n$ são vetores em E , a soma deles pode ser escrita, sem ambigüidade, como

$$v_1 + v_2 + \dots + v_n.$$

Definição 1.3

Um vetor β do espaço vetorial E é uma **combinação linear** dos vetores $v_1, v_2, \dots, v_n \in E$, se existem escalares $c_1, c_2, \dots, c_n \in \mathbb{K}$, tais que

$$\beta = c_1 v_1 + c_2 v_2 + \dots + c_n v_n.$$

Observação 1.3

Usando as propriedades associativa e comutativa da adição e as propriedades distributivas da multiplicação por escalares, obtemos:

$$\sum_{i=1}^n c_i v_i + \sum_{i=1}^n d_i v_i = \sum_{i=1}^n (c_i + d_i) v_i,$$

$$c \sum_{i=1}^n c_i v_i = \sum_{i=1}^n (cc_i) v_i.$$

2. Subespaço Vetorial

Definição 2.1

Seja E um espaço vetorial sobre o corpo K . Um subconjunto $F \subset E$ é um **subespaço vetorial** de E se F for um espaço vetorial sobre $\mathbb{K}$ com as operações de adição de vetores e multiplicação de vetores por escalares que fazem de E um espaço vetorial. Ou seja, $F \subset E$ é um subespaço vetorial de E se:

1. $\mathbf{0} \in F$;
2. $v \in F \implies -v \in F$;
3. $u \in F$ e $v \in F \implies u + v \in F$;
4. $\lambda \in \mathbb{K}$ e $v \in F \implies \lambda v \in F$.

Proposição 2.1

Um subconjunto não-vazio $F \subset E$ é um subespaço vetorial de E se, e somente se, $\lambda v + w \in F, \forall v, w \in F, \forall \lambda \in \mathbb{K}$.

Prova.

$(\implies)$ é evidente.

$(\impliedby)$ Como $F \neq \emptyset$, seja $v \in F$. Logo, $(-1)v + v = \mathbf{0} \in F$.

Sejam $w \in F$ e $\lambda \in \mathbb{K}$. Então, $\lambda w + \mathbf{0} = \lambda w \in F$.

Em particular, $-w = (-1)w \in F$.

Finalmente, se $v, w \in F$, então $1v + w = v + w \in F$. Logo, F é um subespaço vetorial de E . ■

Exemplo 2.1

1. E é um subespaço de E .
2. $F = \{0\}$ é um subespaço de E , denominado o **subespaço nulo** de E .
3. O espaço $\mathbb{K}[x]$ das funções polinomiais com coeficientes no corpo $\mathbb{K}$ é um subespaço do espaço de todas as funções de $\mathbb{K}$ em $\mathbb{K}$.

4. Uma matriz $n \times n$ A sobre o corpo $\mathbb{K}$ é dita **simétrica** se $A_{ij} = A_{ji}$, $\forall i, j \in \{1, \dots, n\}$.

As matrizes simétricas formam um subespaço do espaço vetorial $\mathbb{K}^{n \times n}$ das matrizes $n \times n$ sobre o corpo $\mathbb{K}$.

5. Uma matriz $n \times n$ A sobre o corpo $\mathbb{C}$ dos números complexos é **hermitiana** se $A_{jk} = \overline{A_{kj}}$, $\forall j, k \in \{1, \dots, n\}$. Se A é uma matriz hermitiana, todas as entradas da diagonal $A_{11}, A_{22}, \dots, A_{nn}$, são números reais.

- O conjunto das matrizes hermitianas $n \times n$ não é um subespaço do espaço das matrizes $n \times n$ sobre o corpo $\mathbb{C}$, pois, por exemplo, a matriz identidade I é hermitiana, mas iI não é.
- O conjunto das matrizes hermitianas $n \times n$ é um espaço vetorial sobre o corpo $\mathbb{R}$ dos números reais.

□

Lema 2.1

Se A é uma matriz $m \times n$ sobre $\mathbb{K}$ e B, C são matrizes $n \times p$ sobre $\mathbb{K}$, então

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

onde $\lambda \in \mathbb{K}$ é um escalar qualquer.

Prova.

$$\begin{aligned} [A(\lambda B + C)]_{ij} &= \sum_{k=1}^n A_{ik}(\lambda B + C)_{kj} \\ &= \sum_{k=1}^n A_{ik}(\lambda B_{kj} + C_{kj}) \\ &= \sum_{k=1}^n (\lambda A_{ik}B_{kj} + A_{ik}C_{kj}) \\ &= \lambda \sum_{k=1}^n A_{ik}B_{kj} + \sum_{k=1}^n A_{ik}C_{kj} \\ &= \lambda[AB]_{ij} + [AC]_{ij}. \end{aligned}$$

■

Exemplo 2.2

Seja A uma matriz $m \times n$ sobre $\mathbb{K}$.

O subconjunto $\{X \in \mathbb{K}^{n \times 1} \mid AX = 0\}$ é um subespaço de $\mathbb{K}^{n \times 1}$.

De fato:

- Se $AX = AY = 0$, então $A(\lambda X + Y) = \lambda AX + AY = 0, \forall \lambda \in \mathbb{K}$.
- Se $X = 0$, então $AX = 0$. $\square$

Proposição 2.2

A interseção de uma coleção arbitrária $\{F_\alpha\}_{\alpha \in I}$ de subespaços de E é um subespaço de E .

Prova.

Seja $F = \bigcap_{\alpha \in I} F_\alpha$. Como $0 \in F_\alpha, \forall \alpha \in I$, temos que $0 \in F$. Logo, $F \neq \emptyset$.

Se $v, w \in F$, então $v, w \in F_\alpha, \forall \alpha \in I$.

Logo, $\lambda v + w \in F_\alpha, \forall \alpha \in I$, ou seja, $\lambda v + w \in F, \forall \lambda \in \mathbb{K} \text{ e } \forall v, w \in F$. $\blacksquare$

Definição 2.2

Seja $S \subset E$ um subconjunto. O subespaço de E gerado por S é a interseção de todos os subespaços de E que contém S .

Isto é, o subespaço de E gerado por S é o menor subespaço de E que contém S .

Observação 2.1

O subespaço gerado por $S = \emptyset$ é o subespaço nulo. O único vetor desse subespaço é o vetor nulo, 0 .

Proposição 2.3

Seja $S \subset E$ um subconjunto não-vazio. Então o subespaço F gerado por S é o conjunto de todas as combinações lineares de vetores em S .

Prova.

Seja W o conjunto de todas as combinações lineares de vetores de S .

Seja $\lambda \in \mathbb{K}$ e consideremos as combinações lineares:

$$v = \alpha_1 v_1 + \dots + \alpha_k v_k, w = \beta_1 w_1 + \dots + \beta_l w_l \in W,$$

onde $v_i \in S, \forall i \in \{1, \dots, k\}$ e $w_j \in S, \forall j \in \{1, \dots, l\}$.

Então, para $\lambda \in \mathbb{K}$, temos

$$\lambda v + w = \lambda \alpha_1 v_1 + \dots + \lambda \alpha_k v_k + \beta_1 w_1 + \dots + \beta_l w_l \in W,$$

pois $\lambda v + w$ é uma combinação linear de vetores em S .

Como $W \neq \emptyset$, já que $S \subset W$ e $S \neq \emptyset$, temos que W é um subespaço de E que contém S . Logo, $F \subset W$.

Seja $v \in W$. Então,

$$v = \alpha_1 v_1 + \dots + \alpha_k v_k,$$

onde $v_i \in S, \forall i \in \{1, \dots, k\}$. Como $S \subset F$ e F é um subespaço de E , temos que $v \in F$. Logo, $W \subset F$. Assim, $F = W$. ■

Exemplo 2.3

Seja $E = \mathbb{K}[x]$ o espaço vetorial das funções polinomiais de $\mathbb{K}$ em $\mathbb{K}$ e seja $S = \{f_n(x) = x^n \mid n \in \mathbb{N}\}$. Então, E é o subespaço gerado por S . □

Definição 2.3

Sejam $S_1, \dots, S_k$ subconjuntos do espaço vetorial E . O conjunto formado por todas as somas da forma

$$v_1 + v_2 + \dots + v_k, \quad v_i \in S_i, i = 1, \dots, k.$$

é chamado a **soma dos subconjuntos** $S_1, \dots, S_k$, e se designa por

$$S_1 + S_2 + \dots + S_k.$$

Proposição 2.4

Se $W_1, \dots, W_k$ são subespaços do espaço vetorial E , então $W_1 + \dots + W_k$ é um subespaço de E .

Além disso, $W_1 + \dots + W_k$ é o subespaço gerado pelo conjunto

$$S = \bigcup_{i=1}^k W_i.$$

Prova.

Sejam $v, w \in W_1 + \dots + W_k$.

Então, existem vetores $v_i, w_i \in W_i, 1 \leq i \leq k$, tais que

$$v = v_1 + \dots + v_k \quad \text{e} \quad w = w_1 + \dots + w_k.$$

Logo, o vetor

$$\lambda v + w = (\lambda v_1 + w_1) + \dots + (\lambda v_k + w_k)$$

pertence a $W_1 + \dots + W_k$, já que $\lambda v_i + w_i \in W_i$, $\forall i = 1, \dots, k$. Além disso, como $W_1 + \dots + W_k \neq \emptyset$, já que $\mathbf{0} \in W_1 + \dots + W_k$, temos que $W_1 + \dots + W_k$ é um subespaço de E .

Seja W o subespaço gerado pelo conjunto $S = \bigcup_{i=1}^k W_i$. Então, W é o conjunto de todas as combinações lineares de vetores em S .

Como

$$S = \bigcup_{i=1}^k W_i \subset W_1 + \dots + W_k$$

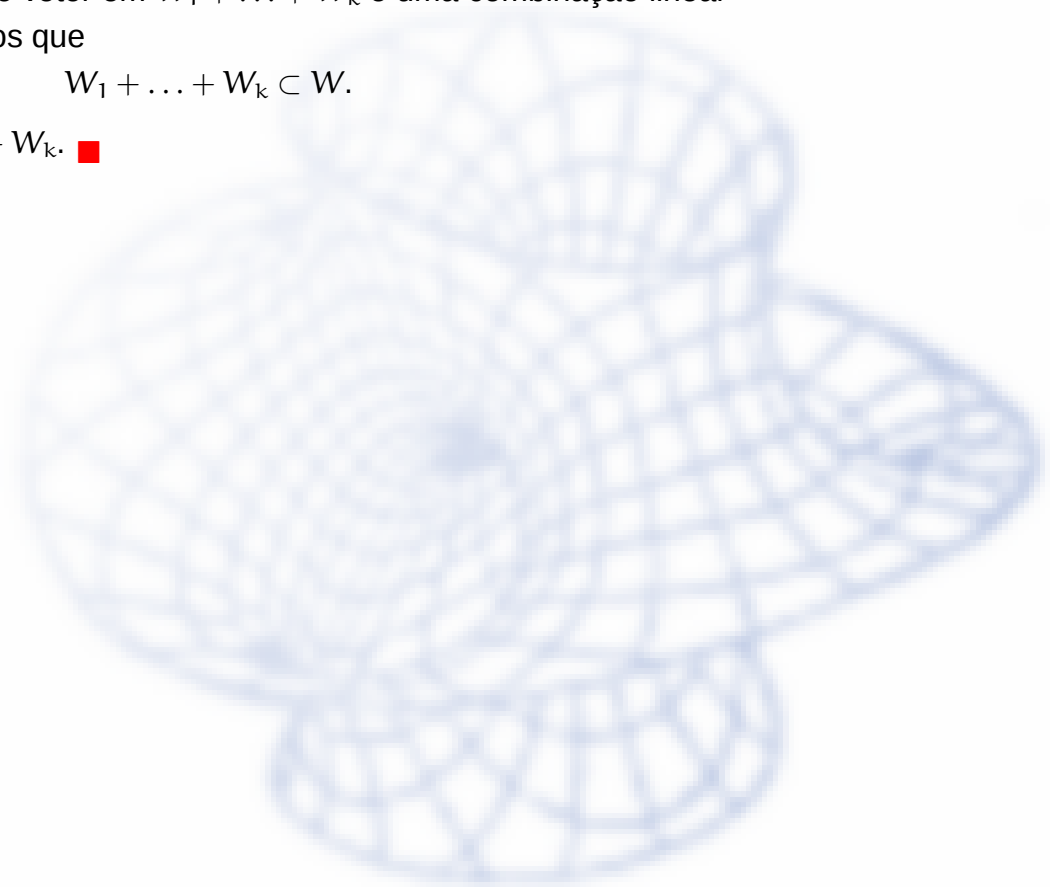
e $W_1 + \dots + W_k$ é um subespaço, temos que

$$W \subset W_1 + \dots + W_k.$$

Além disso, como todo vetor em $W_1 + \dots + W_k$ é uma combinação linear de vetores em S , temos que

$$W_1 + \dots + W_k \subset W.$$

Logo, $W = W_1 + \dots + W_k$. ■



3. Bases e Dimensão

Definição 3.1

Seja E um espaço vetorial sobre o corpo $\mathbb{K}$. Dizemos que um subconjunto $S \subset E$ é **linearmente dependente** (LD) se existem vetores distintos $v_1, \dots, v_n$ em S e escalares $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, não todos nulos, tais que

$$\lambda_1 v_1 + \dots + \lambda_n v_n = \mathbf{0}.$$

Um conjunto S que não é linearmente dependente é dito **linearmente independente** (LI). Isto é, S é LI se para toda coleção de vetores distintos $v_1, \dots, v_n$ em S e escalares $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, vale a implicação:

$$\lambda_1 v_1 + \dots + \lambda_n v_n = \mathbf{0} \implies \lambda_1 = \lambda_2 = \dots = \lambda_n = 0.$$

Observação 3.1

- Todo subconjunto $S \subset E$ que contém o vetor nulo é LD, pois $1 \cdot \mathbf{0} = \mathbf{0}$.
- Todo conjunto que contém um conjunto LD é LD.
- Todo subconjunto de um conjunto LI é LI.

Definição 3.2

Seja E um espaço vetorial. Uma **base** de E é um conjunto linearmente independente de vetores em E que geram o espaço E .

Definição 3.3

Dizemos que um espaço vetorial tem **dimensão finita** se ele possui uma base finita.

Observação 3.2

Um subconjunto $\{v_1, \dots, v_n\} \subset E$ é uma base de E se, e somente se, todo vetor $v \in E$ se expressa, de modo único, como combinação linear dos vetores $v_1, \dots, v_n$, ou seja, existem escalares únicos $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, tais que $v = \lambda_1 v_1 + \dots + \lambda_n v_n$.

Exemplo 3.1

Seja $\mathbb{K}$ um corpo e seja S o subconjunto de $\mathbb{K}^n$ formado pelos vetores

Verifique a afirmação! $\longrightarrow$

$$\begin{aligned} e_1 &= (1, 0, 0, \dots, 0, 0) \\ e_2 &= (0, 1, 0, \dots, 0, 0) \\ &\vdots \\ e_n &= (0, 0, 0, \dots, 0, 1). \end{aligned}$$

Se $x_1, \dots, x_n$ são escalares em $\mathbb{K}$, o vetor

$$x_1 e_1 + \dots + x_n e_n \text{ é o vetor } (x_1, \dots, x_n).$$

Assim, S gera $\mathbb{K}^n$. Como

$$x_1 e_1 + \dots + x_n e_n = (0, \dots, 0)$$

se, e somente se, $x_1 = x_2 = \dots = x_n = 0$, S é LI. Logo, S é uma base de $\mathbb{K}^n$, conhecida como a **base canônica** de $\mathbb{K}^n$. $\square$

Exemplo 3.2

Seja $\mathbb{K}$ um corpo e $\mathbb{K}^{m \times n}$ o espaço das matrizes $m \times n$ com entradas no corpo $\mathbb{K}$.

Considere as matrizes A^{ij} que tem 1 na entrada i, j e zero nas demais.

Se $B \in \mathbb{K}^{m \times n}$, então

$$B = \sum_{i,j} B_{ij} A^{ij}.$$

Logo, o conjunto $\{A^{ij} \mid 1 \leq i \leq m \text{ e } 1 \leq j \leq n\}$ é LI e gera o espaço $\mathbb{K}^{m \times n}$.

$\square$

Exemplo 3.3

Seja P uma matriz invertível $n \times n$ sobre o corpo $\mathbb{K}$. Então, as colunas $P_1, \dots, P_n$ formam uma base do espaço das matrizes colunas $\mathbb{K}^{n \times 1}$.

De fato, se X é uma matriz coluna $n \times 1$, então

$$PX = x_1 P_1 + \dots + x_n P_n.$$

Como $PX = 0$ admite apenas a solução trivial $X = 0$, $\{P_1, \dots, P_n\}$ é LI.

Seja $Y \in \mathbb{K}^{n \times 1}$ e seja $X = P^{-1}Y$. Então,

$$Y = PX = x_1 P_1 + \dots + x_n P_n, \text{ onde } X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Assim, $\{P_1, \dots, P_n\}$ gera o espaço $\mathbb{K}^{n \times 1}$.

Portanto, $\{P_1, \dots, P_n\}$ é uma base finita de $\mathbb{K}^{n \times 1}$. $\square$

Verifique que ...

O conjunto S é um subespaço de $\mathbb{K}^{n \times 1}$ conhecido como o espaço solução ou espaço das soluções do sistema linear homogêneo $AX = \mathbf{0}$.

Exemplo 3.4

Seja A uma matriz $n \times n$ com entradas no corpo $\mathbb{K}$ e seja

$$S = \{X \in \mathbb{K}^{n \times 1} \mid AX = \mathbf{0}\},$$

o conjunto de todas as soluções do sistema linear homogêneo $AX = \mathbf{0}$.

Seja R uma matriz reduzida por linhas à forma em escada que seja equivalente por linhas à matriz A . Então,

$$S = \{X \in \mathbb{K}^{n \times 1} \mid RX = \mathbf{0}\}.$$

Seja r o número de linhas não-nulas da matriz R , e suponhamos que as primeiras entradas não-nulas das linhas não-nulas de R ocorram nas colunas $k_1 < \dots < k_r$.

Seja $J = \{1, \dots, n\} - \{k_1, \dots, k_r\}$ o conjunto dos $n - r$ índices distintos de $k_1, \dots, k_r$.

Então, o sistema $RX = \mathbf{0}$ tem a forma

$$\begin{aligned} x_{k_1} + \sum_{j \in J} c_{1j} x_j &= 0 \\ &\vdots \\ x_{k_r} + \sum_{j \in J} c_{rj} x_j &= 0, \end{aligned}$$

onde $c_{ij} \in \mathbb{K}$.

Todas as soluções são obtidas atribuindo valores arbitrários aos x_j com $j \in J$ e calculando os valores correspondentes de $x_{k_1}, \dots, x_{k_r}$.

Para cada $j \in J$, seja E_j a solução obtida colocando $x_j = 1$ e $x_i = 0$ para todo $i \in J - \{j\}$.

Afirmção: os vetores E_j , $j \in J$, formam uma base de S .

Como a matriz coluna E_j possui um 1 na linha j e zero nas outras linhas indexadas por elementos de J , temos que $\{E_j \mid j \in J\}$ é um conjunto LI, já que

$$\sum_{j \in J} x_j E_j = \mathbf{0} \text{ se, e somente se, } x_j = 0 \forall j \in J.$$

Seja $T = \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix}$ uma matriz coluna que está no espaço solução S . Então,

$$N = \sum_{j \in J} t_j E_j,$$

também está no espaço solução e é uma solução tal que t_j ocupa a j -ésima linha, para todo $j \in J$.

Como só existe uma solução tal que $x_j = t_j$ para todo $j \in J$, temos que $T = N$.

Logo, $\{E_j \mid j \in J\}$ gera S .

Com isso provamos que $\{E_j \mid j \in J\}$ é uma base de S e que esse espaço tem dimensão finita. $\square$

Definição 3.4

Seja E um espaço vetorial de dimensão finita e seja $\{v_1, \dots, v_n\}$ uma base de E . Então, por definição, a **dimensão** de E , designada por $\dim(E)$, é

$$\dim E \stackrel{\text{def}}{=} n.$$

Demonstremos que o número $\dim E$ está bem definido, isto é, todas as bases de E possuem a mesma quantidade de vetores.

Lema 3.1

Seja $\{u_1, \dots, u_\ell\}$ um conjunto LI de vetores do espaço vetorial E tal que todo u_i é combinação linear dos vetores $w_1, \dots, w_m$ de E . Então, $\ell \leq m$.

Prova.

Vamos provar o lema usando indução em m .

1. Se $m = 1$, existem $\lambda_1, \dots, \lambda_\ell \in \mathbb{K}$, tais que $u_i = \lambda_i w_1$, para todo $1 \leq i \leq \ell$.

Como $u_i \neq \mathbf{0}$ para todo i , pois $\{u_1, \dots, u_\ell\}$ é LI, temos que $\lambda_i \neq 0$ para todo $i \in \{1, \dots, \ell\}$.

Suponhamos, por absurdo, que $\ell > 1$.

Então,

$$\frac{1}{\lambda_1} u_1 + \left(-\frac{1}{\lambda_2}\right) u_2 + 0u_3 + \dots + 0u_\ell = w_1 + (-w_1) = \mathbf{0},$$

o que é um absurdo, já que $\{u_1, \dots, u_\ell\}$ é LI.

2. Suponhamos que o lema vale para $m - 1$ e vamos provar que vale, também, para m .

Cada vetor u_i do conjunto linearmente independente $\{u_1, \dots, u_\ell\}$ é combinação

linear dos vetores $w_1, \dots, w_m$, isto é, existem $\lambda_i^{(j)} \in \mathbb{K}$, $1 \leq i \leq m$, $1 \leq j \leq \ell$, tais que

$$\begin{aligned} u_1 &= \lambda_1^{(1)} w_1 + \dots + \lambda_m^{(1)} w_m \\ u_2 &= \lambda_1^{(2)} w_1 + \dots + \lambda_m^{(2)} w_m \\ &\vdots \\ u_\ell &= \lambda_1^{(\ell)} w_1 + \dots + \lambda_m^{(\ell)} w_m \end{aligned}$$

Se $\lambda_1^j = 0$ para todo $j = 1, \dots, \ell$, temos que todo u_i , $i = 1, \dots, \ell$, é combinação linear dos vetores $w_2, \dots, w_m$. Logo, pela hipótese de indução, $\ell \leq m - 1 \leq m$.

Podemos supor então, sem perda de generalidade, que $\lambda_1^1 \neq 0$.

Sejam $\alpha = \frac{1}{\lambda_1^{(1)}}$ e

$$v_j = u_j - \alpha \lambda_1^j u_1, \quad j = 2, \dots, \ell.$$

Então, $v_j = \beta_2^{(j)} w_2 + \dots + \beta_m^{(j)} w_m$, onde $\beta_i^{(j)} \in \mathbb{K}$, $i = 2, \dots, m$ e $j = 2, \dots, \ell$.

Isto é, cada v_j é uma combinação linear de $w_2, \dots, w_m$.

Se provarmos que $\{v_2, \dots, v_\ell\}$ é LI teremos, pela hipótese de indução,

$$\ell - 1 \leq m - 1 \implies \ell \leq m.$$

Sejam $\mu_2, \dots, \mu_\ell \in \mathbb{K}$ tais que

$$\mu_2 v_2 + \dots + \mu_\ell v_\ell = \mathbf{0}.$$

Então,

$$\mathbf{0} = \mu_2 u_2 + \dots + \mu_\ell u_\ell - \alpha (\mu_2 \lambda_1^{(2)} + \dots + \mu_\ell \lambda_1^{(\ell)}) u_1.$$

Como $\{u_1, \dots, u_\ell\}$ é LI, temos que $\mu_2 = \dots = \mu_\ell = 0$.

Logo, $\{v_2, \dots, v_\ell\}$ é LI. ■

Como consequência do lema anterior, temos o seguinte lema.

Lema 3.2

Sejam $\{w_1, \dots, w_m\}$ um conjunto gerador e $\{u_1, \dots, u_\ell\}$ um conjunto LI do espaço vetorial E . Então $\ell \leq m$.

Teorema 3.1

Sejam $\{v_1, \dots, v_n\}$ e $\{w_1, \dots, w_m\}$ bases de um espaço vetorial E . Então, $m = n$.

Prova.

Como $\{v_1, \dots, v_n\}$ é um conjunto LI de vetores de E e $\{w_1, \dots, w_m\}$ é um conjunto gerador de E , temos $n \leq m$.

Analogamente, como $\{v_1, \dots, v_n\}$ é um conjunto gerador de E e $\{w_1, \dots, w_m\}$ é um conjunto LI de vetores de E , temos $n \geq m$.

Portanto, $n = m$. ■

Exemplo 3.5

1. O espaço vetorial $\mathbb{K}^n$ sobre o corpo $\mathbb{K}$ tem dimensão n : $\dim \mathbb{K}^n = n$.

Em particular, o espaço $\mathbb{C}^n$ sobre o corpo $\mathbb{C}$ tem dimensão n e o espaço $\mathbb{R}^n$ sobre o corpo $\mathbb{R}$ tem dimensão n .

2. O espaço vetorial $\mathbb{K}^{m \times n}$ das matrizes $m \times n$ com entradas no corpo $\mathbb{K}$ tem dimensão mn .

3. Se A é uma matriz $m \times n$, então o espaço solução do sistema homogêneo $AX = \mathbf{0}$ tem dimensão $n - r$, onde r é o número de linhas não-nulas de uma matriz linha reduzida à forma em escada e equivalente por linhas a A .

□

Exemplo 3.6

Seja S um conjunto qualquer e $\mathcal{F}(S, \mathbb{K})$ o conjunto de todas as funções de S em $\mathbb{K}$.

Sabemos, pelo item **3.** do exemplo **1.1**, que $\mathcal{F}(S, \mathbb{K})$ é um espaço vetorial sobre o corpo $\mathbb{K}$.

Afirmção: $\mathcal{F}(S, \mathbb{K})$ tem dimensão finita se, e somente se, S é um conjunto finito. Nesse caso, $\dim \mathcal{F}(S, \mathbb{K}) = \#S$, onde $\#S$ é o número de elementos de S .

De fato,

($\Leftarrow$) Suponhamos que $S = \{p_1, \dots, p_n\}$ é um conjunto finito.

Sejam $f_i : S \rightarrow \mathbb{K}$, $i = 1, \dots, n$, as funções definidas por:

$$f_i(p_j) = \delta_{ij}, \quad j = 1, \dots, n.$$

Se $f \in \mathcal{F}(S, \mathbb{K})$, então,

$$f = f(p_1)f_1 + \dots + f(p_n)f_n.$$

E se $f = \lambda_1 f_1 + \dots + \lambda_n f_n$, então

$$f(p_i) = \lambda_1 f_1(p_i) + \dots + \lambda_n f_n(p_i) = \lambda_i f_i(p_i) = \lambda_i,$$

para todo $i = 1, \dots, n$.

Logo, todo $f \in \mathcal{F}(S, \mathbb{K})$ se escreve de modo único como combinação linear das funções $f_1, \dots, f_n$.

Assim, $\{f_1, \dots, f_n\}$ é uma base de $\mathcal{F}(S, \mathbb{K})$ e $\dim \mathcal{F}(S, \mathbb{K}) = n = \#S$.

($\implies$) Se $\mathcal{F}(S, \mathbb{K})$ tem dimensão finita, todo subconjunto LI de $\mathcal{F}(S, \mathbb{K})$ tem um número finito de elementos.

Suponhamos, por absurdo, que S tem uma infinidade de elementos.

Consideremos o conjunto

$$T = \{f_p : S \rightarrow \mathbb{K} \mid f_p(p) = 1 \text{ e } f_p(q) = 0, \forall q \in S - \{p\}\}.$$

O conjunto T é LI, pois, se $p_1, \dots, p_n \in S$ são elementos distintos em S e $\lambda_1, \dots, \lambda_n$ são escalares em $\mathbb{K}$ tais que $\lambda_1 f_{p_1} + \dots + \lambda_n f_{p_n} = \mathbf{0}$, então

$$\lambda_1 f_{p_1}(p_i) + \dots + \lambda_n f_{p_n}(p_i) = \mathbf{0},$$

para $i = 1, \dots, n$. Como

$$\lambda_1 f_{p_1}(p_i) + \dots + \lambda_n f_{p_n}(p_i) = \lambda_i f_{p_i}(p_i) = \lambda_i,$$

temos que $\lambda_i = 0$, para todo $i = 1, \dots, n$.

Assim, T é infinito e LI, que contradiz o fato de $\mathcal{F}(S, \mathbb{K})$ ter dimensão finita.

□

Exemplo 3.7

Seja $\mathbb{K}$ um subcorpo do corpo dos números complexos $\mathbb{C}$, e seja E o espaço vetorial das funções polinomiais sobre $\mathbb{K}$. Isto é, E é o conjunto das funções $f : \mathbb{K} \rightarrow \mathbb{K}$ da forma

$$f(x) = c_0 + c_1 x + \dots + c_n x^n,$$

onde $n \in \mathbb{N}$ e $c_0, \dots, c_n \in \mathbb{K}$.

Observe que E é um subespaço vetorial do espaço $\mathcal{F}(\mathbb{K}, \mathbb{K})$ de todas as funções de $\mathbb{K}$ em $\mathbb{K}$.

Considere as funções polinomiais $f_k(x) = x^k$, $k \in \mathbb{N}$. O conjunto infinito $\{f_0, f_1, \dots, f_k, \dots\}$ é uma base de E .

De fato, dado $f(x) = c_0 + c_1 x + \dots + c_n x^n$ em E , temos que

$$f = c_0 f_0 + c_1 f_1 + \dots + c_n f_n.$$

Então, o conjunto $\{f_k \mid k \in \mathbb{N}\}$ gera o espaço E .

Além disso, conjunto $\{f_k \mid k \in \mathbb{N}\}$ é LI, pois se $c_0 f_0 + c_1 f_1 + \dots + c_n f_n = 0$, ou seja,

$$c_0 + c_1 x + \dots + c_n x^n = 0, \forall x \in \mathbb{K},$$

temos que $c_0 = c_1 = \dots = c_n = 0$, já que um polinômio de grau n com coeficientes complexos tem no máximo n raízes.

Observamos que o espaço E não tem uma base finita, pois se tivesse, todo conjunto LI de vetores de E teria uma quantidade finita de elementos.

Assim, E tem dimensão infinita. $\square$

Observação 3.3

Convencionamos que o **espaço nulo** $\{\mathbf{0}\}$ (gerado pelo conjunto vazio) tem dimensão zero.

Lema 3.3

Seja $\{w_1, \dots, w_m\} \subset E$ um conjunto LI e seja $w \in E$.

Então, $\{w, w_1, \dots, w_m\}$ é LI se, e somente se, w não é combinação linear dos vetores $w_1, \dots, w_m$.

Prova.

$(\implies)$ Suponhamos que w é uma combinação linear de $w_1, \dots, w_m$. Então, existem $\lambda_1, \dots, \lambda_m \in \mathbb{K}$ tais que $w = \lambda_1 w_1 + \dots + \lambda_m w_m$, ou seja,

$$1 \cdot w + (-\lambda_1)w_1 + \dots + (-\lambda_m)w_m = \mathbf{0},$$

Como o coeficiente 1 de w é diferente de zero, o conjunto $\{w, w_1, \dots, w_m\}$ é LD.

$(\impliedby)$ Suponhamos que o conjunto $\{w, w_1, \dots, w_m\}$ é LD. Então, existem escalares $\lambda, \lambda_1, \dots, \lambda_m$, não todos nulos, tais que

$$\lambda w + \lambda_1 w_1 + \dots + \lambda_m w_m = \mathbf{0}.$$

Se $\lambda = 0$, então $\lambda_1 w_1 + \dots + \lambda_m w_m = \mathbf{0}$, onde os coeficientes $\lambda_1, \dots, \lambda_m$ não são todos nulos, o que é um absurdo, já que $\{w_1, \dots, w_m\}$ é LI.

Logo, $\lambda \neq 0$ e

$$w = \left(-\frac{\lambda_1}{\lambda}\right) w_1 + \dots + \left(-\frac{\lambda_m}{\lambda}\right) w_m,$$

isto é, w é uma combinação linear dos vetores $w_1, \dots, w_m$. ■

Teorema 3.2

Seja E um espaço vetorial não-nulo e seja $\{u_1, \dots, u_n\}$ um conjunto de geradores de E . Então $\{u_1, \dots, u_n\}$ contém uma base de E .

Prova.

Vamos provar que existe $\{u_{i_1}, \dots, u_{i_\ell}\} \subset \{u_1, \dots, u_n\}$ tal que:

I. $\{u_{i_1}, \dots, u_{i_\ell}\}$ é LI.

II. $\{u_j, u_{i_1}, \dots, u_{i_\ell}\}$ não é LI, $\forall j \in \{1, \dots, n\}$.

Seja $i_1 \in \{1, \dots, n\}$ tal que $u_{i_1} \neq 0$. Então $\{u_{i_1}\}$ é LI.

Se $\{u_{i_1}\}$ satisfaz **II.** o processo termina.

Se $\{u_{i_1}\}$ não satisfaz **II.**, então existe $i_2 \in \{1, \dots, n\}$ tal que $\{u_{i_1}, u_{i_2}\}$ é LI.

Como esse processo é finito, com no máximo n etapas, podemos obter $\{u_{i_1}, \dots, u_{i_\ell}\} \subset \{u_1, \dots, u_n\}$ que satisfaz **I.** e **II.**

Vamos provar agora que $\{u_{i_1}, \dots, u_{i_\ell}\}$ é uma base de E .

Como $\{u_j, u_{i_1}, \dots, u_{i_\ell}\}$ não é LI, existem escalares $\lambda_1^{(j)}, \dots, \lambda_n^{(j)}$ tais que

$$u_j = \lambda_1^{(j)} u_{i_1} + \dots + \lambda_n^{(j)} u_{i_\ell}, \text{ para todo } j = 1, \dots, n.$$

Seja $v \in E$. Como $\{u_1, \dots, u_n\}$ gera o espaço E , existem escalares $\lambda_1, \dots, \lambda_n$, tais que

$$v = \lambda_1 u_1 + \dots + \lambda_n u_n.$$

Logo,

$$v = \sum_{j=1}^n \lambda_j u_j = \sum_{j=1}^n \lambda_j \sum_{k=1}^{\ell} \lambda_k^{(j)} u_{i_k} = \sum_{k=1}^{\ell} \left(\sum_{j=1}^n \lambda_j \lambda_k^{(j)} \right) u_{i_k},$$

isto é, v é uma combinação linear dos vetores $u_{i_1}, \dots, u_{i_\ell}$.

Logo, $\{u_{i_1}, \dots, u_{i_\ell}\}$ é uma base de E . ■

Teorema 3.3

Seja E um espaço vetorial de dimensão finita n e seja $\{w_1, \dots, w_m\} \subset E$ um subconjunto LI. Então $\{w_1, \dots, w_m\}$ está contido numa base de E :

$$\{w_1, \dots, w_m, w_{m+1}, \dots, w_{m+k}\}.$$

Prova.

Vamos provar que existe um conjunto de vetores $\{w_1, \dots, w_m, w_{m+1}, \dots, w_{m+k}\}$ tal que:

I. é linearmente independente;

II. $\{w_1, \dots, w_m, w_{m+1}, \dots, w_{m+k}, w\}$ não é LI, $\forall w \in E$.

• Se o conjunto LI $\{w_1, \dots, w_m\}$ não satisfaz **II.**, existe $w_{m+1} \in E$ tal que $\{w_1, \dots, w_m, w_{m+1}\}$ é LI.

Se $\{w_1, \dots, w_m, w_{m+1}\}$ não satisfaz **II.**, então existe $w_{m+2} \in E$ tal que $\{w_1, \dots, w_m, w_{m+1}, w_{m+2}\}$ é LI.

Este processo é finito, com no máximo $n - m$ etapas, pois todo subconjunto LI de E tem no máximo $n = \dim E$ elementos.

• Como $\{w_1, \dots, w_m, w_{m+1}, \dots, w_{m+k}\}$ satisfaz **I.** e **II.**, temos que este conjunto é LI e $\{v, w_1, \dots, w_{m+k}\}$ não é LI, $\forall v \in E$.

Assim, todo $v \in E$ é uma combinação linear dos vetores $w_1, \dots, w_{m+k}$.

Logo, $\{w_1, \dots, w_m, w_{m+1}, \dots, w_{m+k}\}$ é uma base de E que contém o conjunto LI $\{w_1, \dots, w_m\}$. ■

Corolário 3.1

Seja E um espaço vetorial sobre o corpo $\mathbb{K}$.

- Se $\{v_1, \dots, v_n\}$ é um conjunto de vetores LI e E tem dimensão finita, então $n \leq \dim E$.
- Se $\{v_1, \dots, v_n\}$ gera o espaço E , então E tem dimensão finita e $n \geq \dim E$.

Corolário 3.2

Seja E um espaço vetorial de dimensão finita n .

Se $\{v_1, \dots, v_n\}$ é um conjunto de vetores LI ou um conjunto gerador de E , então é uma base de E .

Prova.

• Suponhamos que $\{v_1, \dots, v_n\}$ é um conjunto LI. Se $\{v_1, \dots, v_n\}$ não gera E , existe $w \in E$ tal que $\{w, v_1, \dots, v_n\}$ é LI. Então, $n + 1 \leq \dim E = n$, o que é um absurdo.

• Se $\{v_1, \dots, v_n\}$ gera o espaço E , então existe um subconjunto $\{v_{i_1}, \dots, v_{i_k}\}$

de $\{v_1, \dots, v_n\}$ que é uma base de E . Como $\dim E = n = k$, temos que $\{v_1, \dots, v_n\}$ é uma base de E . ■

Teorema 3.4

Seja F um subespaço vetorial de um espaço E de dimensão finita. Então, F tem dimensão finita e $\dim F \leq \dim E$.

Além disso, $\dim F = \dim E$ se, e somente se, $F = E$.

Prova.

Podemos supor que $F \neq \{0\}$.

Vamos provar que existem vetores $v_i \in F$, $i = 1, \dots, k$ tais que:

I. $\{v_1, \dots, v_k\}$ é LI.

II. $\{v, v_1, \dots, v_k\}$ não é LI, $\forall v \in F$.

Como $F \neq \emptyset$, existe $v_1 \neq 0$ em F .

Então o conjunto $\{v_1\}$ é LI. Se $\{v_1\}$ não satisfaz II., existe $v_2 \in F$ tal que $\{v_1, v_2\}$ é LI.

Como todo conjunto LI de E tem no máximo $\dim E = n$ elementos, o processo pára em algum $k \leq n$.

Logo $\{v_1, \dots, v_k\}$ é uma base de F e $\dim F = k \leq \dim E$.

Se $\dim F = \dim E = n$ e $\{v_1, \dots, v_n\}$ é uma base de F , temos que $\{v_1, \dots, v_n\}$ é uma base de E . Logo $E = F$. ■

Teorema 3.5

Sejam F_1 e F_2 subespaços vetoriais de dimensão finita de um espaço vetorial E . Então $F_1 + F_2$ é um subespaço vetorial de dimensão finita de E , e

$$\dim(F_1 + F_2) = \dim F_1 + \dim F_2 - \dim(F_1 \cap F_2).$$

Prova.

Como $F_1 \cap F_2$ é um subespaço dos espaços F_1 e F_2 , $F_1 \cap F_2$ tem dimensão finita.

Se $\{v_1, \dots, v_k\}$ é uma base de $F_1 \cap F_2$, existem $w_1, \dots, w_m$ em F_1 , tais que $\{v_1, \dots, v_k, w_1, \dots, w_m\}$ é uma base de F_1 e existem $u_1, \dots, u_\ell$, tais que $\{v_1, \dots, v_k, u_1, \dots, u_\ell\}$ é uma base de F_2 .

Afirmção: $\{v_1, \dots, v_k, w_1, \dots, w_m, u_1, \dots, u_\ell\}$ é uma base de $F_1 + F_2$.

É claro que o conjunto acima gera o espaço $F_1 + F_2$. Basta, então, verificar que o conjunto é LI.

Sejam $\lambda_1, \dots, \lambda_k, \mu_1, \dots, \mu_m, \delta_1, \dots, \delta_\ell$ escalares tais que

$$\sum_{i=1}^k \lambda_i v_i + \sum_{j=1}^m \mu_j w_j + \sum_{s=1}^{\ell} \delta_s u_s = \mathbf{0} \quad (1)$$

Como

$$\sum_{s=1}^{\ell} \delta_s u_s = - \left(\sum_{i=1}^k \lambda_i v_i + \sum_{j=1}^m \mu_j w_j \right) \in F_1 \cap F_2,$$

existem escalares $\rho_1, \dots, \rho_k$ tais que,

$$\sum_{s=1}^{\ell} \delta_s u_s = \sum_{i=1}^k \rho_i v_i, \quad \text{ou seja} \quad \sum_{i=1}^k \rho_i v_i - \sum_{s=1}^{\ell} \delta_s u_s = \mathbf{0}.$$

Então, $\rho_1 = \dots = \rho_k = \delta_1 = \dots = \delta_\ell = 0$, já que $\{v_1, \dots, v_k, u_1, \dots, u_\ell\}$ é uma base de F_2 .

Como $\delta_1 = \dots = \delta_\ell = 0$, temos, por (1), que

$$\sum_{i=1}^k \lambda_i v_i + \sum_{j=1}^m \mu_j w_j = \mathbf{0}.$$

Logo, $\lambda_1 = \dots = \lambda_k = \mu_1 = \dots = \mu_m = 0$, já que $\{v_1, \dots, v_k, w_1, \dots, w_m\}$ é uma base de F_1 .

Provamos, assim, que $\{v_1, \dots, v_k, w_1, \dots, w_m, u_1, \dots, u_\ell\}$ é uma base de $F_1 + F_2$.

Então,

$$\dim(F_1 + F_2) = k + m + \ell = (k + m) + (k + \ell) - k,$$

ou seja,

$$\dim(F_1 + F_2) = \dim F_1 + \dim F_2 - \dim(F_1 \cap F_2).$$



4. Coordenadas

Definição 4.1

Se E é um espaço vetorial de dimensão finita, uma **base ordenada** de E é uma seqüência finita de vetores LI que geram E .

Se a seqüência $v_1, \dots, v_n$ é uma base ordenada de E , então o conjunto $\{v_1, \dots, v_n\}$ é uma base de E .

Notação. $\mathcal{B} = \{v_1, \dots, v_n\}$.

Dado um vetor $v \in E$, existem escalares únicos $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, denominados as **coordenadas de v em relação à base ordenada $\mathcal{B}$** , tais que

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

• Designamos por $[v]_{\mathcal{B}}$ a matriz $n \times 1$ cujas entradas são as coordenadas de v em relação à base ordenada $\mathcal{B}$. Isto é,

$$[v]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

Seja agora $\mathcal{B}' = \{v'_1, \dots, v'_n\}$ uma outra base ordenada de E . Então, para cada $j = 1, \dots, n$, existem escalares $A_{ij} \in \mathbb{K}$, tais que

$$v_j = \sum_{i=1}^n A_{ij} v'_i.$$

Sejam $\lambda'_1, \dots, \lambda'_n$ as coordenadas de um vetor v em relação à base ordenada $\mathcal{B}'$, isto é,

$$v = \lambda'_1 v'_1 + \dots + \lambda'_n v'_n.$$

Então,

$$\begin{aligned} v &= \sum_{j=1}^n \lambda_j v_j \\ &= \sum_{j=1}^n \lambda_j \sum_{i=1}^n A_{ij} v'_i \\ &= \sum_{j=1}^n \sum_{i=1}^n A_{ij} \lambda_j v'_i \\ \Rightarrow v &= \sum_{i=1}^n \left(\sum_{j=1}^n A_{ij} \lambda_j \right) v'_i. \end{aligned}$$

Note que...

Na definição ao lado, a palavra seqüência é crucial, pois o seu significado implica numa relação de precedência entre os vetores da base. Isto é, há o primeiro vetor, o segundo, e assim sucessivamente e sem ambigüidade.

Pela unicidade das coordenadas de um vetor em relação a uma base, temos que

$$\lambda'_i = \sum_{j=1}^n A_{ij} \lambda_j, \quad \text{para todo } 1 \leq i \leq n.$$

Logo,

$$[v]_{B'} = A [v]_B,$$

onde $A = (A_{ij})_{1 \leq i, j \leq n}$ é a matriz cuja j -ésima coluna é $A_j = [v_j]_{B'}$. A matriz A é chamada **matriz de mudança da base B para a base B'** .

Afirmção. A matriz A é invertível.

Com efeito,

- basta mostrar que o sistema homogêneo $AX = \mathbf{0}$ possui somente a solução trivial.

Se $AX = \mathbf{0}$, seja $v = x_1 v_1 + \dots + x_n v_n$, onde $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Então,

$$[v]_{B'} = A[v]_B = AX = \mathbf{0}.$$

Logo, $v = x_1 v_1 + \dots + x_n v_n = \mathbf{0}$. Como $\{v_1, \dots, v_n\}$ é LI, temos que $x_1 = \dots = x_n = \mathbf{0}$, ou seja, $X = \mathbf{0}$.

Observação 4.1

Como A é invertível, temos que:

$$[v]_B = A^{-1} [v]_{B'}.$$

Logo, $[v'_j]_B = A^{-1} [v'_j]_{B'} = A^{-1} e_j = j$ -ésima coluna da matriz A^{-1} .

Isto é, A^{-1} é a matriz de mudança da base B' para a base B .

Teorema 4.1

Sejam B uma matriz $n \times n$ invertível sobre o corpo $\mathbb{K}$, E um espaço vetorial sobre o corpo $\mathbb{K}$ e $B = \{v_1, \dots, v_n\}$ uma base ordenada de E .

Então, existe uma única base ordenada $B' = \{v'_1, \dots, v'_n\}$ de E tal que

$$[v]_B = B[v]_{B'}, \quad \forall v \in E.$$

Prova.

Caso exista B' , teremos que

$$[v'_j]_B = B[v'_j]_{B'} = B_j = j\text{-ésima coluna de } B.$$

Ou seja,

$$v'_j = \sum_{i=1}^n B_{ij} v_i.$$

Afirmção. $\{v'_1, \dots, v'_n\}$ é uma base de E .

De fato, como $\dim E = n$, basta mostrar que $v'_1, \dots, v'_n$ são LI.

Se $\mathbf{0} = x_1 v'_1 + \dots + x_n v'_n$ então,

$$\begin{aligned} \mathbf{0} &= \sum_{j=1}^n x_j v'_j = \sum_{j=1}^n x_j \sum_{i=1}^n B_{ij} v_i \\ \Rightarrow \mathbf{0} &= \sum_{i=1}^n \left(\sum_{j=1}^n B_{ij} x_j \right) v_i \\ \Rightarrow \mathbf{0} &= \sum_{j=1}^n B_{ij} x_j, \quad \forall i = 1, \dots, n \\ \Rightarrow BX &= \mathbf{0}, \quad \text{onde } X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \end{aligned}$$

Como B é invertível, temos que $X = \mathbf{0}$, ou seja, $x_1 = \dots = x_n = 0$. ■

Exemplo 4.1

Seja $\mathbb{K}$ um corpo e seja $x = (x_1, \dots, x_n)$ um vetor de $\mathbb{K}^n$. Se $\mathcal{B}$ é a base canônica de $\mathbb{K}^n$, $\mathcal{B} = \{e_1, \dots, e_n\}$, a matriz das coordenadas do vetor x em relação à base $\mathcal{B}$ é dada por

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

□

Exemplo 4.2

Para cada número real $\theta \in \mathbb{R}$, a matriz

$$P = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

é invertível e

$$P^{-1} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$$

Consideremos os vetores

$$v_1 = (\cos \theta, \sin \theta) \text{ e } v_2 = (-\sin \theta, \cos \theta),$$

e a base canônica $\mathcal{B} = \{e_1 = (1, 0), e_2 = (0, 1)\}$.

Como $[v_i]_{\mathcal{B}} = P_i$ (i -ésima coluna de P), $i = 1, 2$, temos que $\mathcal{B}' = \{v_1, v_2\}$ é uma base ordenada de $\mathbb{R}^2$ tal que

$$[v]_{\mathcal{B}} = P[v]_{\mathcal{B}'}, \quad \forall v \in \mathbb{R}^2.$$

Logo,

$$[v]_{\mathcal{B}'} = P^{-1}[v]_{\mathcal{B}}, \quad \forall v \in \mathbb{R}^2.$$

Então, se $v = (x, y) = xe_1 + ye_2$, temos que

$$[v]_{\mathcal{B}'} = \begin{pmatrix} x' \\ y' \end{pmatrix} = P^{-1}[v]_{\mathcal{B}} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

ou seja,

$$\begin{aligned} x' &= \cos \theta x + \sin \theta y \\ y' &= -\sin \theta x + \cos \theta y. \end{aligned}$$

□

Exemplo 4.3

A. Verifiquemos que os vetores $v_1 = (-1, 0, 0)$, $v_2 = (4, 2, 0)$ e $v_3 = (5, -3, 8)$ formam uma base de $\mathbb{R}^3$.

Para isso, basta mostrar que a matriz

$$A = \begin{pmatrix} -1 & 4 & 5 \\ 0 & 2 & -3 \\ 0 & 0 & 8 \end{pmatrix}$$

cujas j -ésimas colunas são $A_j = [v_j]_{\mathcal{B}}$, onde $\mathcal{B} = \{e_1, e_2, e_3\}$ é a base canônica de $\mathbb{R}^3$, é invertível.

De fato, vamos efetuar operações elementares na matriz A , e na matriz identidade, para tentar transformá-la numa matriz reduzida por linhas à forma em escada.

$$\begin{pmatrix} -1 & 4 & 5 \\ 0 & 2 & -3 \\ 0 & 0 & 8 \end{pmatrix} \longleftrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\downarrow \qquad \downarrow$$

$$\begin{pmatrix} 1 & -4 & -5 \\ 0 & 1 & -3/2 \\ 0 & 0 & 1 \end{pmatrix} \longleftrightarrow \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/8 \end{pmatrix}$$

$$\begin{array}{ccc} \downarrow & & \downarrow \\ \begin{pmatrix} 1 & 0 & -11 \\ 0 & 1 & -3/2 \\ 0 & 0 & 1 \end{pmatrix} & \longleftrightarrow & \begin{pmatrix} -1 & 2 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/8 \end{pmatrix} \\ \downarrow & & \downarrow \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \longleftrightarrow & \begin{pmatrix} -1 & 2 & 11/8 \\ 0 & 1/2 & 3/16 \\ 0 & 0 & 1/8 \end{pmatrix} \end{array}$$

Como a matriz A é equivalente por linhas à matriz identidade, temos que A é invertível e

$$A^{-1} = \begin{pmatrix} -1 & 2 & 11/8 \\ 0 & 1/2 & 3/16 \\ 0 & 0 & 1/8 \end{pmatrix}.$$

B. Determinemos, agora, as coordenadas x'_1, x'_2, x'_3 do vetor $v = (x_1, x_2, x_3)$ em relação à base $\mathcal{B}' = \{v_1, v_2, v_3\}$.

Como $[v]_{\mathcal{B}} = A[v]_{\mathcal{B}'}$, temos

$$[v]_{\mathcal{B}'} = A^{-1}[v]_{\mathcal{B}} = \begin{pmatrix} -1 & 2 & 11/8 \\ 0 & 1/2 & 3/16 \\ 0 & 0 & 1/8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

ou seja,

$$\begin{aligned} x'_1 &= -x_1 + 2x_2 + \frac{11}{8}x_3 \\ x'_2 &= \frac{1}{2}x_2 + \frac{3}{16}x_3 \\ x'_3 &= \frac{1}{8}x_3. \end{aligned}$$

Em particular, $v = (1, 2, 16) = 25v_1 + 5v_2 + 2v_3$. $\square$

5. Equivalência por Linhas – resumo

Definição 5.1

Seja A uma matriz $m \times n$ sobre o corpo $\mathbb{K}$. O *espaço-linha* de A é o subespaço de $\mathbb{K}^n$ gerado pelos vetores-linhas de A , e o *posto-linha* de A é a dimensão do espaço-linha de A .

Seja A uma matriz como na definição e seja P uma matriz $k \times m$ sobre $\mathbb{K}$. Então, o produto $B = PA$ é uma matriz $k \times n$ cujas linhas $B_1, \dots, B_k$ são combinações lineares dos vetores-linhas de A :

$$B_i = P_{i1}\alpha_1 + \dots + P_{im}\alpha_m.$$

Logo, o espaço-linha de B é um subespaço do espaço-linha de A .

Se a matriz P é uma matriz invertível (nesse caso $m = k$), ou seja, B e A são linha-equivalentes, então o espaço-linha de A coincide com o espaço-linha de B .

Teorema 5.1

Seja R uma matriz não-nula $m \times n$ reduzida por linhas à forma em escada. Então os vetores-linhas não-nulos de R formam uma base do espaço-linha de R .

Prova.

Como os vetores-linhas não-nulos de R , $\rho_i = (R_{i1}, \dots, R_{in})$, $i = 1, \dots, r$, geram o espaço-linha de R , basta mostrar que eles são LI.

Sejam $k_1 < \dots < k_r$ tais que:

1. $R_{ij} = 0$, se $j < k_i$;

2. $R_{ik_j} = \delta_{ij}$.

Seja $\beta = (b_1, \dots, b_n)$ um vetor do espaço-linha de R .

Então, $\beta = c_1\rho_1 + \dots + c_r\rho_r$.

Como $b_{k_j} = \sum_{i=1}^r c_i R_{ik_j} = \sum_{i=1}^r c_i \delta_{ij} = c_j$, temos que $\beta = b_{k_1}\rho_1 + \dots + b_{k_r}\rho_r$.

Em particular, se $c_1\rho_1 + \dots + c_r\rho_r = \mathbf{0}$, então c_j é a b_{k_j} coordenada do vetor nulo. Logo, $c_j = 0$, para todo $j = 1, \dots, r$. Assim, $\rho_1, \dots, \rho_k$ são LI ■

Teorema 5.2

Sejam m e n inteiros positivos e $\mathbb{K}$ um corpo.

Seja W um subespaço de $\mathbb{K}^n$ com $\dim W \leq m$. Então existe exatamente uma matriz $m \times n$ sobre $\mathbb{K}$, reduzida por linhas à forma em escada, cujo espaço-linha é W .

Prova.

Existência. Como $\dim W \leq m$, podemos tomar m vetores $\alpha_1, \dots, \alpha_m$ em W , que geram W . Seja A a matriz $m \times n$ que tem esses vetores como vetores-linhas.

Seja R uma matriz reduzida por linhas à forma em escada equivalente por linhas à matriz A .

Então, o espaço-linha de R é W .

Unicidade. Seja R uma matriz $m \times n$ arbitrária reduzida por linhas à forma em escada cujo espaço-linha é W . Sejam $\rho_1, \dots, \rho_r$ os vetores-linhas não-nulos de R e suponhamos que o primeiro elemento não-nulo de ρ_i ocorra na coluna k_i , $i = 1, \dots, r$.

Na demonstração do teorema anterior, vimos que se $\beta = (b_1, \dots, b_n) \in W$, então

$$\beta = \sum_{i=1}^r b_{k_i} \rho_i.$$

Assim, todo vetor $\beta \in W$ está determinado quando as suas coordenadas b_{k_j} , $j = 1, \dots, r$, são conhecidas.

Então, ρ_i é o único vetor em W cuja k_i -ésima coordenada é 1 e cuja k_j -ésima coordenada é nula para $j \neq i$.

Suponhamos que $\beta \neq \mathbf{0}$. Como $\beta = \sum_{i=1}^r b_{k_i} \rho_i$, temos que existe $s \in \{1, \dots, r\}$ tal que

$$\beta = \sum_{i=s}^r b_{k_i} \rho_i, \quad \text{com } b_{k_s} \neq 0.$$

Como $R_{ij} = 0$ se $i > s$ e $j \leq k_s$, temos que

$$(0, \dots, 0, b_{k_s}, \dots, b_n), \quad b_{k_s} \neq 0$$

Provamos, assim, que r e $k_1 < \dots < k_r$ são determinados apenas pelos vetores de W , pois $\{k_1, \dots, k_r\}$ é o conjunto dos inteiros positivos t ,

$1 \leq t \leq n$, tais que existe algum $\beta \neq 0$ em W , cuja primeira coordenada não-nula ocorre na coluna t .

Além disso, para cada $j = 1, \dots, r$, existe apenas um vetor em W com primeira coordenada não-nula igual a 1 na coluna k_j e k_i -ésima coluna nula se $i \in \{1, \dots, r\}$, $i \neq j$. Logo, os vetores não-nulos $\rho_1, \dots, \rho_r$ de R são determinados de modo único. ■

Corolário 5.1

Cada matriz $m \times n$ A é equivalente por linhas a exatamente uma matriz reduzida por linhas à forma em escada.

Prova.

Sejam R e R' matrizes reduzidas por linhas à forma em escada que sejam equivalentes por linhas à matriz A . Então, R e R' são equivalentes por linhas e têm, portanto, o mesmo espaço-linha. Logo, $R = R'$. ■

Corolário 5.2

Sejam A e B matrizes $m \times n$ sobre o corpo K . Então, A e B são equivalentes por linha se, e somente se, possuem o mesmo espaço-linha.

Prova.

Já sabemos que se A e B são equivalentes por linhas, então, possuem o mesmo espaço-linha.

Suponhamos que A e B possuem o mesmo espaço-linha. Seja R a matriz reduzida por linhas à forma em escada que é equivalente por linhas a A , e seja R' a matriz reduzida por linhas à forma em escada que é equivalente por linhas a B . Como o espaço-linha de A é igual ao espaço-linha de R e o espaço-linha de B é igual ao espaço-linha de R' e o espaço-linha de A é igual ao espaço-linha de B , temos que o espaço-linha de R é igual ao espaço-linha de R' . Logo, $R = R'$ e A é equivalente por linhas a B . ■

Exemplo 5.1

Seja W o subespaço de $\mathbb{R}^4$ gerado pelos vetores:

$$\begin{aligned}\alpha_1 &= (1, 2, 2, 1) \\ \alpha_2 &= (0, 2, 0, 1) \\ \alpha_3 &= (-2, 0, -4, 3).\end{aligned}$$

a. Demonstrar que $\{\alpha_1, \alpha_2, \alpha_3\}$ é uma base de W .

Seja A a matriz com vetores-linha $\alpha_1, \alpha_2, \alpha_3$. Então,

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

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \longrightarrow Q = \frac{1}{6} \begin{pmatrix} 6 & -6 & 0 \\ -2 & 5 & -1 \\ 4 & -4 & 2 \end{pmatrix}$$

onde R é a matriz reduzida por linhas à forma em escada equivalente por linhas a A e Q é a matriz invertível tal que $R = QA$.

Como o espaço-linha de A é igual ao espaço-linha de R e o posto de R é 3, temos que o posto de A é 3 e, portanto, $\{\alpha_1, \alpha_2, \alpha_3\}$ é uma base de W .

b. Seja $\beta = (b_1, b_2, b_3, b_4)$ um vetor de W . Determine as coordenadas de β em relação à base ordenada $\{\alpha_1, \alpha_2, \alpha_3\}$.

Seja, $\{\rho_1, \rho_2, \rho_3\}$ a base de W formada pelos vetores-linha de R . É fácil ver que o espaço gerado por ρ_1, ρ_2 e ρ_3 é formado pelos vetores β para os quais $b_3 = 2b_1$. Para um tal β , temos:

$$\begin{aligned} \beta &= b_1\rho_1 + b_2\rho_2 + b_4\rho_3 \\ &= (b_1, b_2, b_4) R \\ &= (b_1, b_2, b_4) QA \\ &= x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3, \end{aligned}$$

onde $(x_1, x_2, x_3) = (b_1, b_2, b_4)Q$, ou seja,

$$\begin{aligned} x_1 &= b_1 - \frac{1}{3}b_2 + \frac{2}{3}b_4 \\ x_2 &= -b_1 + \frac{5}{6}b_2 - \frac{2}{3}b_4 \\ x_3 &= -\frac{1}{6}b_2 + \frac{1}{3}b_4. \end{aligned} \tag{1}$$

c. Sejam

$$\begin{aligned} \alpha'_1 &= (1, 0, 2, 0) \\ \alpha'_2 &= (0, 2, 0, 1) \\ \alpha'_3 &= (0, 0, 0, 3). \end{aligned}$$

Mostrar que $\alpha'_1, \alpha'_2, \alpha'_3$ formam uma base de W .

Como os vetores $\alpha'_1, \alpha'_2, \alpha'_3$ são todos da forma (y_1, y_2, y_3, y_4) , com $y_3 = 2y_1$, temos que eles pertencem a W . É fácil verificar que esses vetores são LI e, portanto, formam uma base de W .

d. Determinar a matriz P de mudança da base $\mathcal{B}' = \{\alpha'_1, \alpha'_2, \alpha'_3\}$ para a base $\mathcal{B} = \{\alpha_1, \alpha_2, \alpha_3\}$.

Usando (1), temos que

$$\begin{aligned} x_1 &= 1 - \frac{1}{3}(0) + \frac{2}{3}(0) = 1 \\ x_2 &= -1 + \frac{2}{6}(0) - \frac{2}{3}(0) = -1 \\ x_3 &= -\frac{1}{6}(0) + \frac{1}{3}(0) = 0, \end{aligned}$$

são as coordenadas de $\alpha'_1 = (b_1, b_2, b_3, b_4) = (1, 0, 2, 0)$ na base $\mathcal{B}$, ou seja, $\alpha'_1 = 1\alpha_1 - 1\alpha_2$.

Analogamente, obtemos $\alpha'_2 = \alpha_2$ e $\alpha'_3 = 2\alpha_1 - 2\alpha_2 + \alpha_3$. Logo, como $P_j = [\alpha'_j]_{\mathcal{B}}$ é a j -ésima coluna da matriz de mudança de base P , temos que

$$P = \begin{pmatrix} 1 & 0 & 2 \\ -1 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

• Vamos agora resolver o mesmo exemplo usando a matriz B com vetores coluna α_1, α_2 e α_3 :

$$B = \begin{pmatrix} 1 & 0 & -2 \\ 2 & 2 & 0 \\ 2 & 0 & -4 \\ 1 & 1 & 3 \end{pmatrix}.$$

Para isto, determinemos $Y = (y_1, y_2, y_3, y_4)$ de modo que o sistema $BX = Y$

tenha solução.

Como

$$\begin{pmatrix} 1 & 0 & -2 & y_1 \\ 2 & 2 & 0 & y_2 \\ 2 & 0 & -4 & y_3 \\ 1 & 1 & 3 & y_4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & -2 & y_1 \\ 0 & 2 & 4 & y_2 - 2y_1 \\ 0 & 0 & 0 & y_3 - 2y_1 \\ 0 & 1 & 5 & y_4 - y_1 \end{pmatrix} \longrightarrow$$

$$\rightarrow \begin{pmatrix} 1 & 0 & -2 & y_1 \\ 0 & 0 & -6 & y_2 - 2y_4 \\ 0 & 1 & 5 & y_4 - y_1 \\ 0 & 0 & 0 & y_3 - 2y_1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & -2 & y_1 \\ 0 & 0 & 1 & -1/6(y_2 - 2y_4) \\ 0 & 1 & 5 & y_4 - y_1 \\ 0 & 0 & 0 & y_3 - 2y_1 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 1 & 0 & 0 & y_1 - 1/3y_2 + 2/3y_4 \\ 0 & 1 & 0 & -y_1 + 5/6y_2 - 2/3y_4 \\ 0 & 0 & 1 & 1/6(2y_4 - y_2) \\ 0 & 0 & 0 & y_3 - 2y_1 \end{pmatrix}.$$

Logo, a condição para que o sistema $BX = Y$ admita solução é $y_3 = 2y_1$.

Então, $\beta = (b_1, b_2, b_3, b_4) \in W$ se, e somente se, $b_3 = 2b_1$.

Seja

$$R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}_{4 \times 3}.$$

a matriz reduzida por linhas à forma em escada equivalente por linhas a B e seja

$$P = \begin{pmatrix} 1 & -1/3 & 0 & 2/3 \\ -1 & 5/6 & 0 & -2/3 \\ 0 & -1/6 & 0 & 2/6 \\ -2 & 0 & 1 & 0 \end{pmatrix}_{4 \times 4},$$

a matriz invertível tal que $R = PB$.

Sejam $\beta = (b_1, b_2, b_3, b_4) \in W$ e x_1, x_2, x_3 as coordenadas de β na base $\{\alpha_1, \alpha_2, \alpha_3\}$, isto é,

$$\beta = x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3.$$

Logo, $BX = Y$, onde $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ e $Y = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix}$.

Como $RX = PBX = PY$, temos que

$$\begin{aligned} x_1 &= b_1 - \frac{1}{3}b_2 + \frac{2}{3}b_4 \\ x_2 &= -b_1 + \frac{5}{6}b_2 - \frac{2}{3}b_4 \\ x_3 &= -\frac{1}{6}b_2 + \frac{2}{6}b_4. \end{aligned}$$

□



Transformações Lineares

1. Transformação Linear - noções básicas

Definição 1.1

Sejam V e W espaços vetoriais sobre o corpo $\mathbb{K}$. Uma **transformação linear** de V em W é uma função $T : V \longrightarrow W$ tal que:

$$\begin{aligned}T(v + w) &= T(v) + T(w), & \forall v, w \in V, \\T(\lambda v) &= \lambda T(v), & \forall v \in V, \lambda \in \mathbb{K}.\end{aligned}$$

Observação 1.1

a. $T(\lambda_1 v_1 + \dots + \lambda_n v_n) = \lambda_1 T(v_1) + \dots + \lambda_n T(v_n)$, $\forall v_1, \dots, v_n \in V$ e $\lambda_1, \dots, \lambda_n \in \mathbb{K}$.

b. $T(\mathbf{0}) = \mathbf{0}$ (com efeito, $T(\mathbf{0}) = T(\mathbf{0} + \mathbf{0}) = T(\mathbf{0}) + T(\mathbf{0}) \implies T(\mathbf{0}) = \mathbf{0}$).

Exemplo 1.1

Seja A uma matriz $m \times n$ sobre o corpo $\mathbb{K}$. Então,

$$\begin{array}{ccc}T : \mathbb{K}^{n \times 1} & \longrightarrow & \mathbb{K}^{m \times 1} \\X & \longmapsto & AX\end{array} \quad \text{e} \quad \begin{array}{ccc}T : \mathbb{K}^m & \longrightarrow & \mathbb{K}^n \\X & \longmapsto & XA = (A^t X^t)^t,\end{array}$$

são transformações lineares $\square$

Notação ...

A matriz *transposta* da matriz A , designada A^t , é a matriz cujas linhas são as respectivas colunas de A .

Observação 1.2

Toda transformação linear $T : \mathbb{K}^{n \times 1} \longrightarrow \mathbb{K}^{m \times 1}$ é da forma acima.

De fato, sejam $\{e_1, \dots, e_n\}$ a base canônica de $\mathbb{K}^{n \times 1}$, $\{\bar{e}_1, \dots, \bar{e}_m\}$ a base canônica de $\mathbb{K}^{m \times 1}$ e $A_{ij} \in \mathbb{K}$, $i = 1, \dots, m$, $j = 1, \dots, n$, tais que

$$T(e_j) = \sum_{i=1}^m A_{ij} \bar{e}_i.$$

Então,

$$\begin{aligned} T(X) &= T\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = T(x_1 e_1 + \dots + x_n e_n) \\ &= \sum_{j=1}^n x_j T(e_j) = \sum_{j=1}^n \sum_{i=1}^m x_j A_{ij} \bar{e}_i = \sum_{i=1}^m \left(\sum_{j=1}^n A_{ij} x_j \right) \bar{e}_i \\ &= AX. \end{aligned}$$

Exemplo 1.2

Seja $V = \mathcal{C}([a, b]; \mathbb{R})$ o espaço vetorial real das funções contínuas de $[a, b]$ em $\mathbb{R}$. Então, a transformação $T : V \rightarrow V$ definida por:

$$\begin{aligned} T : V &\rightarrow V & \text{onde} & & T_f : [a, b] &\rightarrow \mathbb{R} \\ f &\mapsto T(f) = T_f & & & x &\mapsto T_f(x) = \int_a^x f(s) ds, \end{aligned}$$

é linear. $\square$

Exemplo 1.3

Seja V o espaço vetorial dos polinômios com coeficientes no corpo $\mathbb{K}$. Então, a transformação **derivação** $T : V \rightarrow V$ definida por $f \mapsto Df$:

$$(Df)(x) = c_1 + 2c_2x + \dots + nc_nx^{n-1},$$

onde $f(x) = c_0 + c_1x + \dots + c_nx^n$, é linear. $\square$

Teorema 1.1

Seja V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{K}$ e seja $\{v_1, \dots, v_n\}$ uma base de V . Seja W um espaço vetorial sobre $\mathbb{K}$ e sejam $w_1, \dots, w_n$ vetores em W . Então, existe uma única transformação linear $L : V \rightarrow W$, tal que $L(v_i) = w_i$, para todo $i = 1, \dots, n$.

Prova.

Unicidade.

Sejam $L, T : V \rightarrow W$ transformações lineares tais que $L(v_i) = T(v_i) = w_i$, para todo $i = 1, \dots, n$.

Seja $v \in V$ um vetor arbitrário. Como $\{v_1, \dots, v_n\}$ é uma base de V , existem escalares únicos $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, tais que $v = \lambda_1 v_1 + \dots + \lambda_n v_n$.

Verifique ...
Que o conjunto $\mathcal{C}([a, b]; \mathbb{R})$ das funções contínuas do intervalo $[a, b]$ em $\mathbb{R}$ é um subespaço do espaço vetorial real $\mathcal{F}([a, b], \mathbb{R})$ que consiste de todas as funções do intervalo $[a, b]$ sobre $\mathbb{R}$.

Logo, usando a linearidade de T e L , obtemos

$$T(v) = T\left(\sum_{i=1}^n \lambda_i v_i\right) = \sum_{i=1}^n \lambda_i T(v_i) = \sum_{i=1}^n \lambda_i L(v_i) = L\left(\sum_{i=1}^n \lambda_i v_i\right) = L(v).$$

Existência.

Seja $v = x_1 v_1 + \dots + x_n v_n \in V$ um vetor arbitrário. Definimos a transformação $L : V \longrightarrow W$ por

$$L(v) = x_1 w_1 + \dots + x_n w_n.$$

Afirmção: L é linear.

Sejam $v = x_1 v_1 + \dots + x_n v_n$ e $w = y_1 v_1 + \dots + y_n v_n$ vetores de V e $\lambda \in \mathbb{K}$.

Então, $\lambda v + w = (\lambda x_1 + y_1)v_1 + \dots + (\lambda x_n + y_n)v_n \in V$ e

$$\begin{aligned} L(\lambda v + w) &= (\lambda x_1 + y_1)w_1 + \dots + (\lambda x_n + y_n)w_n \\ &= \lambda x_1 w_1 + y_1 w_1 + \dots + \lambda x_n w_n + y_n w_n \\ &= \lambda(x_1 w_1 + \dots + x_n w_n) + (y_1 w_1 + \dots + y_n w_n) \\ &= \lambda L(v) + L(w). \end{aligned}$$

ou seja, L é linear.

Além disso, $L(v_i) = w_i$ pela própria definição de L . ■

Definição 1.2

Seja $L : V \longrightarrow W$ uma transformação linear.

- $L(V) = \{w \in W \mid \exists v \in V; L(v) = w\}$ é chamado **imagem de L** .
- $L^{-1}(\mathbf{0}) = \{v \in V \mid L(v) = \mathbf{0}\}$ é chamado **núcleo de L** .

Observação 1.3

- $L(V)$ é um subespaço de W .

Sejam $\lambda \in \mathbb{K}$ e $w_1, w_2 \in L(V) \subset W$.

Então, existem $v_1, v_2 \in V$ tais que $L(v_1) = w_1$ e $L(v_2) = w_2$.

Logo,

$$L(\lambda v_1 + v_2) = \lambda L(v_1) + L(v_2) = \lambda w_1 + w_2,$$

ou seja, $\lambda w_1 + w_2 \in L(V)$.

- $L^{-1}(\mathbf{0})$ é um subespaço de V .

Sejam $\lambda \in \mathbb{K}$ e $v_1, v_2 \in L^{-1}(\mathbf{0})$.

Então, $L(v_1) = L(v_2) = \mathbf{0}$ e $L(\lambda v_1 + v_2) = \lambda L(v_1) + L(v_2) = \mathbf{0}$, ou seja, $\lambda v_1 + v_2 \in L^{-1}(\mathbf{0})$.

Lembre que ...

- Em geral, uma função é dita injetiva se transforma elementos distintos em elementos distintos.
- Os termos *injetora*, *injetiva* e *monomorfismo* são sinônimos.

Proposição 1.1

Uma transformação linear $L : V \rightarrow W$ é *injetora* se, e somente se, $L^{-1}(\mathbf{0}) = \{\mathbf{0}\}$.

Prova.

($\Rightarrow$) Suponhamos L injetora.

Seja $v \in V$. Então, $v \in L^{-1}(\mathbf{0})$ se, e somente se, $L(v) = \mathbf{0} = L(\mathbf{0})$ se, e somente se, $v = \mathbf{0}$. Logo, $L^{-1}(\mathbf{0}) = \{\mathbf{0}\}$.

($\Leftarrow$) Suponhamos que $L^{-1}(\mathbf{0}) = \{\mathbf{0}\}$.

Se $L(v) = L(w)$, temos $L(v - w) = L(v) - L(w) = \mathbf{0}$, ou seja, $v - w \in L^{-1}(\mathbf{0}) = \{\mathbf{0}\}$. Logo, $v - w = \mathbf{0}$, ou seja, $v = w$. ■

Definição 1.3

Seja $L : V \rightarrow W$ uma transformação linear. Se V tem dimensão finita, $\dim L(V)$ é o *posto* de L , e $\dim L^{-1}(\mathbf{0})$ é a *nulidade* de L .

Teorema 1.2

Seja $L : V \rightarrow W$ uma transformação linear. Se V é um espaço vetorial de dimensão finita, então $L(V)$ é de dimensão finita e

$$\dim V = \dim(L^{-1}(\mathbf{0})) + \dim(L(V)).$$

Prova.

Caso I. $L^{-1}(\mathbf{0}) = \{\mathbf{0}\}$.

Seja $\{v_1, \dots, v_n\}$ base de V . Vamos mostrar que $\{L(v_1), \dots, L(v_n)\}$ é uma base de $L(V)$.

1. $\{L(v_1), \dots, L(v_n)\}$ é LI.

Com efeito, temos

$$\begin{aligned}
 \lambda_1 L(v_1) + \dots + \lambda_n L(v_n) &= \mathbf{0} \\
 \Rightarrow L(\lambda_1 v_1 + \dots + \lambda_n v_n) &= \mathbf{0} \\
 \Rightarrow \lambda_1 v_1 + \dots + \lambda_n v_n &= \mathbf{0} \\
 \Rightarrow \lambda_1 = \dots = \lambda_n &= 0.
 \end{aligned}$$

2. $\{L(v_1), \dots, L(v_n)\}$ gera $L(V)$.

Sejam $v \in L(V)$ e $u = \lambda_1 v_1 + \dots + \lambda_n v_n \in V$ tal que $L(u) = v$. Então,

$$v = L(u) = L(\lambda_1 v_1 + \dots + \lambda_n v_n) = \lambda_1 L(v_1) + \dots + \lambda_n L(v_n).$$

Caso II. $L^{-1}(\mathbf{0}) \neq \{\mathbf{0}\}$.

• Seja $\{v_1, \dots, v_n\}$ uma base de V e seja $w \in L(V)$.

Então existe $v = \lambda_1 v_1 + \dots + \lambda_n v_n \in V$ tal que $L(v) = w$, ou seja $w = \lambda_1 L(v_1) + \dots + \lambda_n L(v_n)$.

Logo, $\{L(v_1), \dots, L(v_n)\}$ gera o espaço $L(V)$, que tem, portanto, dimensão finita.

• Sejam $\{u_1, \dots, u_k\}$ base de $L^{-1}(\mathbf{0})$ e $\{w_1 = L(v_1), \dots, w_m = L(v_m)\}$ base de $L(V)$.

Afirmção. $\{u_1, \dots, u_k, v_1, \dots, v_m\}$ é uma base de V .

1. Vejamos primeiro que o conjunto é LI.

Sejam $\lambda_1, \dots, \lambda_k, \mu_1, \dots, \mu_m \in \mathbb{K}$ tais que

$$\lambda_1 u_1 + \dots + \lambda_k u_k + \mu_1 v_1 + \dots + \mu_m v_m = \mathbf{0}.$$

Então,

$$\begin{aligned}
 L(\lambda_1 u_1 + \dots + \lambda_k u_k + \mu_1 v_1 + \dots + \mu_m v_m) &= \mathbf{0} \\
 \Rightarrow \lambda_1 L(u_1) + \dots + \lambda_k L(u_k) + \mu_1 L(v_1) + \dots + \mu_m L(v_m) &= \mathbf{0} \\
 \Rightarrow \lambda_1 \mathbf{0} + \dots + \lambda_k \mathbf{0} + \mu_1 w_1 + \dots + \mu_m w_m &= \mathbf{0} \\
 \Rightarrow \mu_1 w_1 + \dots + \mu_m w_m &= \mathbf{0} \\
 \Rightarrow \mu_1 = \dots = \mu_m &= 0 \\
 \Rightarrow \lambda_1 u_1 + \dots + \lambda_k u_k &= \mathbf{0} \\
 \Rightarrow \lambda_1 = \dots = \lambda_k &= 0.
 \end{aligned}$$

2. Vejamos agora que o conjunto gera V .

Seja $v \in V$. Como $L(v) \in L(V)$, existem $\mu_1, \dots, \mu_m \in \mathbb{K}$ tais que

$$L(v) = \mu_1 w_1 + \dots + \mu_m w_m.$$

Logo,

$$L\left(v - \sum_{j=1}^m \mu_j v_j\right) = L(v) - \sum_{j=1}^m \mu_j L(v_j) = L(v) - \sum_{j=1}^m \mu_j w_j = \mathbf{0},$$

ou seja, $v - \sum_{j=1}^m \mu_j v_j \in L^{-1}(\mathbf{0})$. Então, existem $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ tais que

$$v - \sum_{j=1}^m \mu_j v_j = \sum_{i=1}^k \lambda_i u_i,$$

isto é,

$$v = \lambda_1 u_1 + \dots + \lambda_k u_k + \mu_1 v_1 + \dots + \mu_m v_m.$$

Portanto, $\{u_1, \dots, u_k, v_1, \dots, v_m\}$ é uma base de V e

$$\dim V = k + m = \dim L^{-1}(\mathbf{0}) + \dim L(V).$$

Como queríamos demonstrar. ■

Teorema 1.3

Seja A uma matriz $m \times n$ com entradas no corpo $\mathbb{K}$. Então,

$$\text{posto-linha}(A) = \text{posto-coluna}(A).$$

Prova.

Seja a transformação linear

$$\begin{aligned} T: \mathbb{K}^{n \times 1} &\longrightarrow \mathbb{K}^{m \times 1} \\ X &\longmapsto AX. \end{aligned}$$

Como para todo $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^{n \times 1}$, vale $AX = x_1 A_1 + \dots + x_n A_n$, onde

$A_1, \dots, A_n$ são as colunas de A , temos que o espaço-coluna de A é a imagem da transformação T . Portanto, $\text{posto-coluna}(A) = \text{posto}(T)$.

Como $X \in T^{-1}(\mathbf{0})$ se, e somente se, $T(X) = AX = \mathbf{0}$, temos que $T^{-1}(\mathbf{0})$ é o espaço solução S do sistema homogêneo $AX = \mathbf{0}$.

Pelo teorema anterior, temos que

$$\dim S + \text{posto-coluna}(A) = n. \quad (1)$$

Já provamos também que se r é a dimensão do espaço-linha de A , então $n - r$ é a dimensão do espaço solução S .

Logo,

$$\dim S + \text{posto-linha}(A) = n. \quad (2)$$

De (1) e (2), obtemos que

$$\text{posto-linha}(A) = \text{posto-coluna}(A).$$



Definição 1.4

Sejam V e W espaços vetoriais sobre o corpo $\mathbb{K}$, T e L transformações lineares de V em W e $c \in \mathbb{K}$. Definimos as transformações $L + T$ e cL de V em W , da seguinte maneira:

$$\begin{aligned}(L + T)(v) &= L(v) + T(v), \\ (cL)(v) &= cL(v),\end{aligned} \quad \forall v \in V.$$

Proposição 1.2

As transformações $L + T$ e cL são lineares.

Prova.

Sejam $u, v \in V$ e $\lambda \in \mathbb{K}$. Então:

$$\begin{aligned}\bullet \quad (L + T)(\lambda u + v) &= L(\lambda u + v) + T(\lambda u + v) \\ &= \lambda L(u) + L(v) + \lambda T(u) + T(v) \\ &= \lambda(L(u) + T(u)) + L(v) + T(v) \\ &= \lambda(L + T)(u) + (L + T)(v)\end{aligned}$$

e

$$\begin{aligned}\bullet \quad (cL)(\lambda u + v) &= c(L(\lambda u + v)) \\ &= c(\lambda L(u) + L(v)) \\ &= c\lambda L(u) + cL(v) \\ &= \lambda(cL(u)) + cL(v) \\ &= \lambda(cL)(u) + (cL)(v).\end{aligned}$$

Como queríamos demonstrar. 

Definição 1.5

Se V e W são espaços vetoriais, designamos por $\mathcal{L}(V, W)$ o conjunto cujos elementos são todas as transformações lineares de V em W :

$$\mathcal{L}(V, W) = \{L : V \longrightarrow W \mid L \text{ é transformação linear}\}.$$

Teorema 1.4

Sejam V e W espaços vetoriais sobre o corpo $\mathbb{K}$. O conjunto $\mathcal{L}(V, W)$ com as operações de adição e multiplicação por escalares definidas acima é um espaço vetorial sobre $\mathbb{K}$.

Prova.

Atividade. ■

Teorema 1.5

Sejam V um espaço vetorial de dimensão n sobre o corpo $\mathbb{K}$ e W um espaço vetorial de dimensão m sobre o corpo K . Então, o espaço $\mathcal{L}(V, W)$ tem dimensão finita igual a mn .

Prova.

Sejam $\mathcal{B} = \{v_1, \dots, v_n\}$ e $\mathcal{B}' = \{w_1, \dots, w_m\}$ bases ordenadas de V e W , respectivamente.

Para cada $i = 1, \dots, m$ e $j = 1, \dots, n$, seja $L_{ij} \in \mathcal{L}(V, W)$ tal que

$$\begin{cases} L_{ij}(v_j) = w_i \\ L_{ij}(v_k) = \mathbf{0} \quad \text{se } k \neq j, \end{cases} \implies L_{ij}(v_k) = \delta_{jk} w_i.$$

Vamos mostrar que $\{L_{ij} \mid i = 1, \dots, m \text{ e } j = 1, \dots, n\}$ é uma base de $\mathcal{L}(V, W)$.

Seja $L : V \longrightarrow W$ uma transformação linear e sejam $A_{ij} \in \mathbb{K}$ tais que,

$$L(v_j) = \sum_{i=1}^m A_{ij} w_i, \quad 1 \leq j \leq n$$

• Verifiquemos que $L = \sum_{j=1}^n \sum_{i=1}^m A_{ij} L_{ij}$.

Seja $U = \sum_{j=1}^n \sum_{i=1}^m A_{ij} L_{ij}$. Então,

$$\begin{aligned} U(v_k) &= \sum_{j=1}^n \sum_{i=1}^m A_{ij} L_{ij}(v_k) \\ &= \sum_{i=1}^m \sum_{j=1}^n A_{ij} L_{ij}(v_k) \\ &= \sum_{i=1}^m A_{ik} w_i \\ &= L(v_k). \end{aligned}$$

Logo, $U = L$ e, portanto, $\{L_{ij} \mid i = 1, \dots, m \text{ e } j = 1, \dots, n\}$ gera $L(V, W)$.

• Vejamos agora que $\{L_{ij} \mid i = 1, \dots, m \text{ e } j = 1, \dots, n\}$ é LI.

Se $\sum_{j=1}^n \sum_{i=1}^m A_{ij} L_{ij} = \mathcal{O}$, então

$$\begin{aligned} \sum_{j=1}^n \sum_{i=1}^m A_{ij} L_{ij}(v_k) &= \mathbf{0} \quad \forall k \in \{1, \dots, n\} \\ \Rightarrow \sum_{i=1}^m \sum_{j=1}^n A_{ij} L_{ij}(v_k) &= \sum_{i=1}^m A_{ik} w_i = \mathbf{0} \quad \forall k \in \{1, \dots, n\} \\ \Rightarrow A_{ik} &= 0 \quad \forall k \in \{1, \dots, n\} \text{ e } \forall i \in \{1, \dots, m\}. \end{aligned}$$

← Ao lado ...
Estamos designando $\mathcal{O}$ a transformação linear nula, ou seja, $\mathcal{O} : V \rightarrow W$ é dada por $\mathcal{O}(v) = \mathbf{0}$, $\forall v \in V$.

Mostrando o desejado. ■

Proposição 1.3

Sejam V, W, Z espaços vetoriais sobre o corpo $\mathbb{K}$, e $L : V \rightarrow W, T : W \rightarrow Z$ transformações lineares.

Então, **a função composta** $T \circ L : V \rightarrow Z$ é linear, onde

$$T \circ L(v) = T(L(v)), \quad \forall v \in V.$$

Notação ...
A composta $T \circ L$ de duas transformações lineares costuma ser escrita também usando a notação de justaposição TL .

Prova.

Sejam $v, w \in V$ e $\lambda \in \mathbb{K}$. Então,

$$\begin{aligned} (T \circ L)(\lambda v + w) &= T(L(\lambda v + w)) \\ &= T(\lambda L(v) + L(w)) \\ &= \lambda T(L(v)) + T(L(w)) \\ &= \lambda(T \circ L)(v) + (T \circ L)(w). \end{aligned}$$

Logo, $T \circ L$ é uma transformação linear. ■

Exemplo 1.4

Sejam A uma matriz $m \times n$ com entradas no corpo $\mathbb{K}$ e B uma matriz $p \times m$ com entradas em $\mathbb{K}$.

Consideremos as transformações lineares $T : \mathbb{K}^{n \times 1} \rightarrow \mathbb{K}^{m \times 1}$ e $U : \mathbb{K}^{m \times 1} \rightarrow \mathbb{K}^{p \times 1}$, dadas por $T(X) = AX$ e $U(Y) = BY$.

Então, $U \circ T : \mathbb{K}^{n \times 1} \rightarrow \mathbb{K}^{p \times 1}$ é dada por

$$(U \circ T)(X) = U(T(X)) = U(AX) = BAX.$$

Logo, $U \circ T$ é **a multiplicação à esquerda pela matriz produto** BA . □

Definição 1.6

Se V é um espaço vetorial sobre o corpo $\mathbb{K}$, um **operador linear** sobre V é uma transformação linear de V em V .

Observação 1.4

- Se $U, T \in \mathcal{L}(V, V)$ então $U \circ T \in \mathcal{L}(V, V)$. Assim, o espaço possui uma **multiplicação** definida sobre si por meio da composição.
- Se $T \in \mathcal{L}(V, V)$, usaremos a notação $T^0 = I$ e $T^n = T \circ \dots \circ T$ (n vezes), para $n = 1, 2, 3, \dots$
- A operação de composição não é comutativa, isto é, nem sempre $L \circ T = T \circ L$.

Exemplo 1.5

Seja V o espaço dos polinômios com coeficientes em $\mathbb{K}$ e sejam D e M os operadores:

$$D : V \longrightarrow V$$

$$p(x) = c_0 + c_1x + \dots + c_nx^n \longmapsto D(p)(x) = c_1 + 2c_2x + \dots + nc_nx^{n-1},$$

e

$$M : V \longrightarrow V$$

$$p(x) = c_0 + c_1x + \dots + c_nx^n \longmapsto M(p)(x) = xp(x) = c_0x + c_1x^2 + \dots + c_nx^{n+1}.$$

Então,

$$(DM)(p)(x) = D(M(p))(x) = D(q(x)),$$

onde $q(x) = c_0x + c_1x^2 + \dots + c_nx^{n+1}$.

Logo,

$$(DM)(p)(x) = c_0 + 2c_1x + \dots + (n+1)c_nx^n,$$

De modo análogo, obtemos

$$(MD)(p)(x) = M(D(p))(x) = c_1x + 2c_2x^2 + \dots + nc_nx^n.$$

Logo, $DM - MD = I \neq \mathcal{O}$. $\square$

Proposição 1.4

Sejam V um espaço vetorial sobre o corpo $\mathbb{K}$, U , T_1 e T_2 operadores lineares sobre V e $\lambda \in \mathbb{K}$. Então,

- $IU = UI = U$;
- $U(T_1 + T_2) = UT_1 + UT_2$; $(T_1 + T_2)U = T_1U + T_2U$;

$$c. c(UT_1) = (cU)T_1 = U(cT_1).$$

Prova.

Atividade. ■

Seja $L : V \longrightarrow W$ uma transformação linear *bijetora*, isto é, L é injetora e sobrejetora. Então existe $L^{-1} : W \longrightarrow V$ tal que $LL^{-1} = I_W$ e $L^{-1}L = I_V$.

Lembre que ...
Uma transformação linear $L : V \longrightarrow W$ é sobrejetora, ou sobrejetiva, se a sua imagem é todo W , isto é, $L(V) = W$.

Proposição 1.5

Se $L : V \longrightarrow W$ é uma transformação linear bijetora, então L^{-1} é linear.

Prova.

Sejam $w_1, w_2 \in W$ e $\lambda \in \mathbb{K}$. Sejam $v_1 = L^{-1}(w_1)$ e $v_2 = L^{-1}(w_2)$. Então, $L(v_1) = w_1$, $L(v_2) = w_2$ e

$$L(\lambda v_1 + v_2) = \lambda L(v_1) + L(v_2) = \lambda w_1 + w_2,$$

isto é, $L^{-1}(\lambda w_1 + w_2) = \lambda v_1 + v_2 = \lambda L^{-1}(w_1) + L^{-1}(w_2)$. ■

Observação 1.5

Sejam $L : V \longrightarrow W$ e $T : W \longrightarrow Z$ transformações lineares invertíveis.

Então, $T \circ L : V \longrightarrow Z$ é invertível e $(T \circ L)^{-1} = L^{-1} \circ T^{-1}$.

Definição 1.7

Dizemos que a transformação linear $T : V \longrightarrow W$ é *não-singular* se T é injetora.

Proposição 1.6

A transformação linear $T : V \longrightarrow W$ é não-singular se, e somente se, transforma subconjuntos linearmente independentes de V em subconjuntos linearmente independentes de W .

Prova.

($\implies$) Seja S um subconjunto LI de V e sejam $v_1, \dots, v_n \in S$.

Então, os vetores $T(v_1), \dots, T(v_n)$ são LI, pois se

$$\lambda_1 T(v_1) + \dots + \lambda_n T(v_n) = \mathbf{0},$$

temos

$$\begin{aligned} T(\lambda_1 v_1 + \dots + \lambda_n v_n) &= \mathbf{0} \\ \implies \lambda_1 v_1 + \dots + \lambda_n v_n &= \mathbf{0} \quad (T \text{ é não-singular}) \\ \implies \lambda_1 = \dots = \lambda_n &= 0. \end{aligned}$$

($\Leftarrow$) Seja $v \neq 0$ em V . Como o subconjunto $\mathcal{S} = \{v\}$ é LI, temos que $T(\mathcal{S}) = \{T(v)\}$ é LI. Em particular $T(v) \neq 0$, isto é, T é não-singular. ■

Observação 1.6

A existência de uma inversa à esquerda (ou à direita) não implica que a transformação seja invertível (veja o seguinte exemplo).

Exemplo 1.6

Considere as transformações lineares:

$$\begin{aligned} T: \mathbb{R}^2 &\longrightarrow \mathbb{R} & L: \mathbb{R} &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto x & x &\longmapsto (x, 0). \end{aligned} \quad \text{e}$$

Então, $T \circ L(x) = x = I(x)$ e $L \circ T(x, y) = (x, 0)$.

Logo,

- L é injetora, não é sobrejetora, tem inversa à esquerda e não é invertível.
- T é sobrejetora, não é injetora, tem inversa à direita e não é invertível.

□

Teorema 1.6

Sejam V e W espaços vetoriais de dimensão finita sobre o corpo $\mathbb{K}$ tais que $\dim V = \dim W$.

Se $T: V \longrightarrow W$ é uma transformação linear, as seguintes afirmações são equivalentes:

1. T é invertível;
2. T é injetora;
3. T é sobrejetora;
4. Se $\{v_1, \dots, v_n\}$ é uma base qualquer de V , então $\{T(v_1), \dots, T(v_n)\}$ é uma base de W .
5. Existe uma base $\{v_1, \dots, v_n\}$ de V tal que $\{T(v_1), \dots, T(v_n)\}$ é uma base de W .

Prova.

1 $\Rightarrow$ 2: Segue da definição.

2 $\Rightarrow$ 3: Seja $\{v_1, \dots, v_n\}$ uma base de V . Como $\{v_1, \dots, v_n\}$ é LI e T é injetora, temos que $\{T(v_1), \dots, T(v_n)\}$ é LI. Como $\dim V = \dim W = n$, tal conjunto é uma base de W .

Dado $w \in W$ arbitrário e $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tais que

$$w = \lambda_1 T(v_1) + \dots + \lambda_n T(v_n).$$

Então, $w = T(\lambda_1 v_1 + \dots + \lambda_n v_n)$, ou seja, T é sobrejetora.

3 $\Rightarrow$ 4: Seja $\{v_1, \dots, v_n\}$ uma base de V .

Como $T(V) = W$, para todo $w \in W$ existe $v \in V$ tal que $T(v) = w$. Sejam $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tais que $v = \lambda_1 v_1 + \dots + \lambda_n v_n$.

Então, $w = T(v) = \lambda_1 T(v_1) + \dots + \lambda_n T(v_n)$.

Logo, $\{T(v_1), \dots, T(v_n)\}$ gera o espaço W . Como W tem dimensão n , $\{T(v_1), \dots, T(v_n)\}$ é uma base.

4 $\Rightarrow$ 5: É óbvio.

5 $\Rightarrow$ 1: Seja $\{v_1, \dots, v_n\}$ uma base de V tal que $\{T(v_1), \dots, T(v_n)\}$ é uma base de W .

Seja $w \in W$. Então, existem $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tais que

$$w = \lambda_1 T(v_1) + \dots + \lambda_n T(v_n) = T(\lambda_1 v_1 + \dots + \lambda_n v_n),$$

ou seja, existe $v = \lambda_1 v_1 + \dots + \lambda_n v_n$, tal que $T(v) = w$.

Logo, T é sobrejetora.

Seja agora $v = \lambda_1 v_1 + \dots + \lambda_n v_n$, tal que $T(v) = \mathbf{0}$. Então,

$$\lambda_1 T(v_1) + \dots + \lambda_n T(v_n) = \mathbf{0} \Rightarrow \lambda_1 = \dots = \lambda_n = 0 \Rightarrow v = \mathbf{0}.$$

Ou seja, T é injetora.

Como T é injetora e sobrejetora, T é invertível. ■

Definição 1.8

Se V e W são espaços vetoriais sobre o corpo $\mathbb{K}$, uma transformação linear bijetora (i.e. injetora e sobrejetora) $T : V \rightarrow W$ é denominada um **isomorfismo** de V em W .

Se existe um isomorfismo de V em W dizemos que V é **isomorfo** a W .

Observação 1.7

O isomorfismo é uma relação de equivalência no conjunto dos espaços vetoriais.

De fato,

- V é isomorfo a V , pois a transformação identidade $I : V \rightarrow V, v \mapsto v$ é bijetora e linear.
- Se $T : V \rightarrow W$ é um isomorfismo, então $T^{-1} : W \rightarrow V$ é um isomorfismo. Ou seja, se V é isomorfo a W , então W é isomorfo a V .
- Se $T : V \rightarrow W$ e $L : W \rightarrow Z$ são isomorfismos, então, $L \circ T : V \rightarrow Z$ é um isomorfismo. Ou seja, se V é isomorfo a W e W é isomorfo a Z , então V é isomorfo a Z .

Nota ...
Em virtude da observação ao lado,
se V é isomorfo a W costuma-se
dizer que V e W são isomorfos.

Teorema 1.7

Todo espaço vetorial V de dimensão n sobre o corpo $\mathbb{K}$ é isomorfo a $\mathbb{K}^n$.

Prova.

Seja $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base de V e seja $\mathcal{B}' = \{e_1, \dots, e_n\}$ base canônica de $\mathbb{K}^n$.

Seja $T : V \rightarrow \mathbb{K}^n$ tal que $L(v_i) = e_i, i = 1, \dots, n$. Então,

$$T(x_1v_1 + \dots + x_nv_n) = (x_1, \dots, x_n).$$

Como T transforma a base $\mathcal{B}$ na base $\mathcal{B}'$ de $\mathbb{K}^n$ e $\dim V = n = \dim \mathbb{K}^n$, temos que T é um isomorfismo. ■

Observação 1.8

Espaços vetoriais de dimensão finita são isomorfos **se, e somente se**, tem a mesma dimensão.

Definição 1.9

Seja V um espaço vetorial de dimensão n sobre o corpo $\mathbb{K}$ e seja W um espaço vetorial de dimensão m sobre $\mathbb{K}$. Sejam $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base de V e $\mathcal{B}' = \{w_1, \dots, w_m\}$ uma base de W .

Seja $L : V \rightarrow W$ uma transformação linear e sejam $A_{ij} \in \mathbb{K}$ tais que

$$L(v_j) = \sum_{i=1}^m A_{ij}w_i, \quad \forall j \in \{1, \dots, n\}.$$

A matriz de L em relação às bases ordenadas $\mathcal{B}$ e $\mathcal{B}'$ é a matriz $m \times n$:

$$[L]_{\mathcal{B}'\mathcal{B}} = (A_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{pmatrix}.$$

Note que ...

As entradas da j -ésima coluna da matriz ao lado são os coeficientes dos vetores $w_1, \dots, w_m$ da base $\mathcal{B}'$ de W , que figuram na expressão de $L(v_j)$.

Proposição 1.7

Seja $L : V \longrightarrow W$ uma transformação linear como acima. Então, para todo $v \in V$, temos

$$[L(v)]_{\mathcal{B}'} = [L]_{\mathcal{B}'\mathcal{B}}[v]_{\mathcal{B}}.$$

Prova.

Seja $v = \lambda_1 v_1 + \dots + \lambda_n v_n$, isto é, $[v]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$.

Então,

$$L(v) = L\left(\sum_{j=1}^n \lambda_j v_j\right) = \sum_{j=1}^n \lambda_j L(v_j) = \sum_{j=1}^n \lambda_j \sum_{i=1}^m A_{ij} w_i = \sum_{i=1}^m \left(\sum_{j=1}^n A_{ij} \lambda_j\right) w_i,$$

ou seja, $\sum_{j=1}^n A_{ij} \lambda_j$ é a i -ésima coordenada de $L(v)$ na base ordenada $\mathcal{B}'$.

Logo, $[L(v)]_{\mathcal{B}'} = [L]_{\mathcal{B}'\mathcal{B}}[v]_{\mathcal{B}}$. ■

Observação 1.9

$[L(v_j)]_{\mathcal{B}'}$ é a j -ésima coluna da matriz $[L]_{\mathcal{B}'\mathcal{B}}$.

Casos particulares:

- $[v]_{\mathcal{B}'} = [I(v)]_{\mathcal{B}'} = [I]_{\mathcal{B}'\mathcal{B}}[v]_{\mathcal{B}}$.

ou

- $[v]_{\mathcal{B}} = [I(v)]_{\mathcal{B}} = [I]_{\mathcal{B}\mathcal{B}'}[v]_{\mathcal{B}'}$.

Reciprocamente, dada uma matriz A $m \times n$, existe uma transformação linear $A^{\mathcal{B}'\mathcal{B}} : V \longrightarrow W$ tal que $[A^{\mathcal{B}'\mathcal{B}}]_{\mathcal{B}'\mathcal{B}} = A$.

De fato, seja $v = \lambda_1 v_1 + \dots + \lambda_n v_n$. Definimos $A^{\mathcal{B}'\mathcal{B}} : V \longrightarrow W$ da seguinte maneira:

$$A^{B'B}(v) = \sum_{i=1}^m \left(\sum_{j=1}^n A_{ij} \lambda_j \right) w_i.$$

Proposição 1.8

A aplicação $\mathcal{M} : \mathcal{L}(V, W) \rightarrow \mathbb{K}^{m \times n}$ que a cada transformação linear $L : V \rightarrow W$ associa a matriz $[L]_{B'B}$ é um isomorfismo.

Prova.

Sejam $L : V \rightarrow W$, $T : V \rightarrow W$ transformações lineares e $\lambda \in \mathbb{K}$.

Sejam $[L]_{B'B} = A = [A_1, \dots, A_n]$ e $[T]_{B'B} = B = [B_1, \dots, B_n]$, onde $A_j = [L(v_j)]_{B'}$ e $B_j = [T(v_j)]_{B'}$ são as j -ésimas colunas de A e B .

Então,

$$\begin{aligned} [(\lambda L + T)(v_j)]_{B'} &= [\lambda L(v_j) + T(v_j)]_{B'} \\ &= \lambda [L(v_j)]_{B'} + [T(v_j)]_{B'} \\ &= \lambda A_j + B_j, \end{aligned}$$

isto é,

$$[\lambda L + T]_{B'B} = \lambda A + B.$$

Logo, a aplicação que leva L em $[L]_{B'B}$ é linear.

Como $([L]_{B'B})^{B'B} = L$ e $[A^{B'B}]_{B'B} = A$, temos que a aplicação linear $\mathbb{K}^{m \times n} \rightarrow \mathcal{L}(V, W)$ que leva a matriz A na transformação linear $A^{B'B}$ é a inversa da aplicação linear $\mathcal{M} : \mathcal{L}(V, W) \rightarrow \mathbb{K}^{m \times n}$ que leva L em $[L]_{B'B}$. Logo, $\mathcal{M}$ é um isomorfismo. ■

Exemplo 1.7

Sejam A uma matriz $m \times n$ sobre o corpo $\mathbb{K}$ e $L : \mathbb{K}^{n \times 1} \rightarrow \mathbb{K}^{m \times 1}$ a transformação linear que leva X em AX .

Sejam $B = \{e_1, \dots, e_n\}$ e $B' = \{\bar{e}_1, \dots, \bar{e}_m\}$ as bases canônicas de $\mathbb{K}^{n \times 1}$ e $\mathbb{K}^{m \times 1}$, respectivamente.

Então, $L(e_j) = Ae_j = A_j$ é a j -ésima coluna da matriz A .

Como $A_j = \sum_{i=1}^m A_{ij} \bar{e}_i$, temos que $A = [L]_{B'B}$. □

Proposição 1.9

Sejam V , W e Z espaços vetoriais sobre o corpo $\mathbb{K}$ de dimensões n , m e p , respectivamente.

Sejam $T : V \longrightarrow W$ e $U : W \longrightarrow Z$ transformações lineares. Então,

$$[U \circ T]_{\mathcal{B}''\mathcal{B}} = [U]_{\mathcal{B}''\mathcal{B}'} [T]_{\mathcal{B}'\mathcal{B}},$$

onde $\mathcal{B} = \{v_1, \dots, v_n\}$, $\mathcal{B}' = \{w_1, \dots, w_m\}$ e $\mathcal{B}'' = \{z_1, \dots, z_p\}$ são bases de V , W e Z respectivamente.

Prova.

Sejam $A = [T]_{\mathcal{B}'\mathcal{B}}$, $B = [U]_{\mathcal{B}''\mathcal{B}'}$ e $C = [UT]_{\mathcal{B}''\mathcal{B}}$.

Então,

$$\begin{aligned} (U \circ T)(v_j) &= U(T(v_j)) = U\left(\sum_{i=1}^m A_{ij} w_i\right) = \sum_{i=1}^m A_{ij} U(w_i) \\ &= \sum_{i=1}^m A_{ij} \sum_{k=1}^p B_{ki} z_k = \sum_{k=1}^p \left(\sum_{i=1}^m B_{ki} A_{ij}\right) z_k \\ &= \sum_{k=1}^p (BA)_{kj} z_k. \end{aligned}$$

Logo, $C_{kj} = (BA)_{kj}$, $\forall k = 1, \dots, p$ e $\forall j = 1, \dots, n$.

Ou seja, $C = BA$. ■

Proposição 1.10

Sejam V e W espaços vetoriais de dimensão finita. A transformação linear $L : V \longrightarrow W$ é um isomorfismo (ou seja, é invertível) se, e somente se, $[L]_{\mathcal{B}'\mathcal{B}}$ é uma matriz invertível, quaisquer que sejam as bases $\mathcal{B}$ de V e $\mathcal{B}'$ de W . Além disso, se $[L]_{\mathcal{B}'\mathcal{B}}$ é invertível, então $[L]_{\mathcal{B}'\mathcal{B}}^{-1} = [L^{-1}]_{\mathcal{B}\mathcal{B}'}$.

Prova.

($\implies$) Seja $T : W \longrightarrow V$ tal que $T \circ L = I_V$ e $L \circ T = I_W$, isto é, $T = L^{-1}$.

Então,

$$[T]_{\mathcal{B}\mathcal{B}'} [L]_{\mathcal{B}'\mathcal{B}} = [I_V]_{\mathcal{B}\mathcal{B}} = I \quad \text{e} \quad [L]_{\mathcal{B}'\mathcal{B}} [T]_{\mathcal{B}\mathcal{B}'} = [I_W]_{\mathcal{B}'\mathcal{B}'} = I,$$

ou seja, $[L]_{\mathcal{B}'\mathcal{B}}$ é invertível e $[T]_{\mathcal{B}\mathcal{B}'} = [L^{-1}]_{\mathcal{B}\mathcal{B}'}$ é a sua inversa.

($\impliedby$) Sejam $A = [L]_{\mathcal{B}'\mathcal{B}}$ e $B = A^{-1}$. Seja $T : W \longrightarrow V$ a transformação linear $B^{\mathcal{B}\mathcal{B}'}$, ou seja, $[T]_{\mathcal{B}\mathcal{B}'} = B$.

Então,

$$[T \circ L]_{\mathcal{B}\mathcal{B}} = [T]_{\mathcal{B}\mathcal{B}'} [L]_{\mathcal{B}'\mathcal{B}} = BA = I.$$

Logo, $(T \circ L)(v_j) = v_j, \forall j = 1, \dots, n$, onde $\mathcal{B} = \{v_1, \dots, v_n\}$.

Ou seja, $T \circ L = I_V$.

De modo análogo, temos que

$$[L \circ T]_{\mathcal{B}'\mathcal{B}'} = [L]_{\mathcal{B}'\mathcal{B}} [T]_{\mathcal{B}\mathcal{B}'} = AB = I.$$

Logo, $(L \circ T)(w_i) = w_i, \forall i = 1, \dots, m$, onde $\mathcal{B}' = \{w_1, \dots, w_m\}$.

Ou seja, $L \circ T = I_W$.

Obtemos, assim, que L é invertível. ■

Proposição 1.11

Sejam $\mathcal{B}_1$ e $\mathcal{B}'_1$ bases de V e $\mathcal{B}_2$ e $\mathcal{B}'_2$ bases de W , e $L : V \rightarrow W$ uma transformação linear. Então,

$$[L]_{\mathcal{B}'_2\mathcal{B}'_1} = [I_W]_{\mathcal{B}'_2\mathcal{B}_2} [L]_{\mathcal{B}_2\mathcal{B}_1} [I_V]_{\mathcal{B}_1\mathcal{B}'_1}.$$

Ou seja, o seguinte diagrama é comutativo:

$$\begin{array}{ccc} V_{\mathcal{B}_1} & \xrightarrow{L} & W_{\mathcal{B}_2} \\ & \uparrow [I_V]_{\mathcal{B}_1\mathcal{B}'_1} & \downarrow [I_W]_{\mathcal{B}'_2\mathcal{B}_2} \\ V_{\mathcal{B}'_1} & \xrightarrow{L} & W_{\mathcal{B}'_2} \end{array}$$

Prova.

Como $[I_W]_{\mathcal{B}'_2\mathcal{B}_2} [L]_{\mathcal{B}_2\mathcal{B}_1} = [I_W \circ L]_{\mathcal{B}'_2\mathcal{B}_1} = [L]_{\mathcal{B}'_2\mathcal{B}_1}$, temos

$$[I_W]_{\mathcal{B}'_2\mathcal{B}_2} [L]_{\mathcal{B}_2\mathcal{B}_1} [I_V]_{\mathcal{B}_1\mathcal{B}'_1} = [L]_{\mathcal{B}'_2\mathcal{B}_1} [I_V]_{\mathcal{B}_1\mathcal{B}'_1} = [L \circ I_V]_{\mathcal{B}'_2\mathcal{B}'_1} = [L]_{\mathcal{B}'_2\mathcal{B}'_1}.$$

■

Observação 1.10

As matrizes $[I_W]_{\mathcal{B}'_2\mathcal{B}_2}$ e $[I_V]_{\mathcal{B}_1\mathcal{B}'_1}$ são invertíveis.

Notação

Se $L : V \rightarrow V$ é um operador sobre V e $\mathcal{B}$ é uma base de V , escrevemos

$$[L]_{\mathcal{B}} \stackrel{\text{def}}{=} [L]_{\mathcal{B}\mathcal{B}}.$$

Note que ...
A matriz $[I_W]_{\mathcal{B}'_2\mathcal{B}_2}$ é a matriz de mudança de base da base $\mathcal{B}_2$ para a base $\mathcal{B}'_2$ de W . Analogamente, a matriz $[I_V]_{\mathcal{B}_1\mathcal{B}'_1}$ muda a base $\mathcal{B}_1$ para a base $\mathcal{B}'_1$ de V .

Em particular: se $\mathcal{B}$ e $\mathcal{B}'$ são bases de V e $L : V \rightarrow V$ é um operador sobre V , então,

$$[L]_{\mathcal{B}'} = [I]_{\mathcal{B}'\mathcal{B}}[L]_{\mathcal{B}}[I]_{\mathcal{B}\mathcal{B}'}.$$

Como, além disso, $[I]_{\mathcal{B}'\mathcal{B}} = ([I]_{\mathcal{B}\mathcal{B}'})^{-1}$, temos que

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

onde $P = [I]_{\mathcal{B}\mathcal{B}'}$, ou seja, se $\mathcal{B}' = \{v'_1, \dots, v'_n\}$, então $P_j = [v'_j]_{\mathcal{B}}$ é a j -ésima coluna de P .

Definição 1.10

Sejam A e B matrizes $n \times n$ sobre o corpo $\mathbb{K}$.

Dizemos que **B é semelhante a A sobre $\mathbb{K}$** se existe uma matriz $n \times n$ invertível P com entradas no corpo $\mathbb{K}$, tal que

$$B = P^{-1}AP.$$

Observação 1.11

Sejam $\mathcal{B}$ e $\mathcal{B}'$ bases de um espaço vetorial V sobre K e $L : V \rightarrow V$ um operador linear. Então $[L]_{\mathcal{B}'}$ é semelhante a $[L]_{\mathcal{B}}$.

Sejam, agora, A e B matrizes semelhantes $n \times n$ sobre o corpo $\mathbb{K}$.

Sejam V um espaço vetorial de dimensão n e $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base de V .

Consideremos o operador linear $L : V \rightarrow V$ tal que $[L]_{\mathcal{B}} = A$. Se $B = P^{-1}AP$, definimos os vetores

$$v'_j = \sum_{i=1}^n P_{ij}v_i, \quad j = 1, \dots, n.$$

Afirmção: $\mathcal{B}' = \{v'_1, \dots, v'_n\}$ é uma base de V .

De fato:

Seja $U : V \rightarrow V$ o operador linear tal que $U(v_j) = v'_j, j = 1, \dots, n$.

Como $[U]_{\mathcal{B}\mathcal{B}} = P$ e P é invertível, temos que U é invertível. Logo, $\mathcal{B}'$ é uma base de V .

Além disso, $[I]_{\mathcal{B}\mathcal{B}'} = P$. Logo,

$$[L]_{\mathcal{B}'} = [I]_{\mathcal{B}'\mathcal{B}}[L]_{\mathcal{B}}[I]_{\mathcal{B}\mathcal{B}'} = P^{-1}[L]_{\mathcal{B}}P = P^{-1}AP = B.$$

Isto é, se A e B são semelhantes, existem bases $\mathcal{B}$ e $\mathcal{B}'$ do espaço

Note que ...
A matriz P é a matriz de mudança da base $\mathcal{B}'$ para a base $\mathcal{B}$.

V , tais que $[L]_{\mathcal{B}} = A$ e $[L]_{\mathcal{B}'} = B$.

Exemplo 1.8

Seja T o operador linear sobre $\mathbb{R}^3$ definido por

$$T(x_1, x_2, x_3) = (3x_1 + x_3, -2x_1 + x_2, -x_1 + 2x_2 + 4x_3).$$

a. Determinar a matriz de T em relação à base canônica $\mathcal{B} = \{e_1, e_2, e_3\}$ de $\mathbb{R}^3$.

Solução: Como

$$\begin{aligned} T(e_1) &= T(1, 0, 0) = (3, -2, -1) \\ T(e_2) &= T(0, 1, 0) = (0, 1, 2) \\ T(e_3) &= T(0, 0, 1) = (1, 0, 4), \end{aligned}$$

temos que,

$$[T]_{\mathcal{B}} = \begin{pmatrix} 3 & 0 & 1 \\ -2 & 1 & 0 \\ -1 & 2 & 4 \end{pmatrix}.$$

b. Determinar a matriz de T em relação à base ordenada $\mathcal{B}' = \{w_1, w_2, w_3\}$, onde $w_1 = (1, 0, 1)$, $w_2 = (-1, 2, 1)$ e $w_3 = (2, 1, 1)$.

Solução: Seja $P = [I]_{\mathcal{B}\mathcal{B}'} = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

Então,

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

Determinemos P^{-1} .

$$\begin{array}{ccc} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} & \longleftrightarrow & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ \downarrow & & \downarrow \\ \begin{pmatrix} 1 & -1 & 2 \\ 0 & 2 & 1 \\ 0 & 2 & -1 \end{pmatrix} & \longleftrightarrow & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \\ \downarrow & & \downarrow \\ \begin{pmatrix} 1 & -1 & 2 \\ 0 & 2 & 1 \\ 0 & 0 & -2 \end{pmatrix} & \longleftrightarrow & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix} \\ \downarrow & & \downarrow \end{array}$$

$$\begin{array}{ccc}
 \downarrow & & \downarrow \\
 \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 1/2 \\ 0 & 0 & 1 \end{pmatrix} & \longleftrightarrow & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 1/2 & 1/2 & -1/2 \end{pmatrix} \\
 \downarrow & & \downarrow \\
 \begin{pmatrix} 1 & 0 & 5/2 \\ 0 & 1 & 1/2 \\ 0 & 0 & 1 \end{pmatrix} & \longleftrightarrow & \begin{pmatrix} 1 & 1/2 & 0 \\ 0 & 1/2 & 0 \\ 1/2 & 1/2 & -1/2 \end{pmatrix} \\
 \downarrow & & \downarrow \\
 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \longleftrightarrow & \begin{pmatrix} -1/4 & -3/4 & 5/4 \\ -1/4 & 1/4 & 1/4 \\ 1/2 & 1/2 & -1/2 \end{pmatrix}
 \end{array}$$

Isto é,

$$P^{-1} = \begin{pmatrix} -1/4 & -3/4 & 5/4 \\ -1/4 & 1/4 & 1/4 \\ 1/2 & 1/2 & -1/2 \end{pmatrix} = -\frac{1}{4} \begin{pmatrix} 1 & 3 & -5 \\ 1 & -1 & -1 \\ -2 & -2 & 2 \end{pmatrix}.$$

Logo,

$$\begin{aligned}
 [T]_{B'} &= -\frac{1}{4} \begin{pmatrix} 1 & 3 & -5 \\ 1 & -1 & -1 \\ -2 & -2 & 2 \end{pmatrix} \begin{pmatrix} 3 & 0 & 1 \\ -2 & 1 & 0 \\ -1 & 2 & 4 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} \\
 &= -\frac{1}{4} \begin{pmatrix} 1 & 3 & -5 \\ 1 & -1 & -1 \\ -2 & -2 & 2 \end{pmatrix} \begin{pmatrix} 4 & -2 & 7 \\ -2 & 4 & -3 \\ 3 & 9 & 4 \end{pmatrix} \\
 &= -\frac{1}{4} \begin{pmatrix} -17 & -35 & -22 \\ 3 & -15 & 6 \\ 2 & 14 & 0 \end{pmatrix}.
 \end{aligned}$$

Ou seja,

$$\begin{aligned}
 T(w_1) &= \frac{17}{4}w_1 - \frac{3}{4}w_2 - \frac{2}{4}w_3 \\
 T(w_2) &= \frac{35}{4}w_1 + \frac{15}{4}w_2 - \frac{14}{4}w_3 \\
 T(w_3) &= \frac{22}{4}w_1 - \frac{6}{4}w_2.
 \end{aligned}$$

c. Verificar que o operador T é invertível e determinar T^{-1} .

Solução: Como sabemos que T é invertível se, e somente se, $[T]_B$ é uma matriz invertível, vamos verificar que $[T]_B$ é invertível e determinar sua inversa:

$$\begin{aligned}
 [T]_B &= \begin{pmatrix} 3 & 0 & 1 \\ -2 & 1 & 0 \\ -1 & 2 & 4 \end{pmatrix} \longleftrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 &\quad \downarrow \quad \quad \quad \downarrow \\
 &\begin{pmatrix} 1 & 0 & 1/3 \\ -2 & 1 & 0 \\ -1 & 2 & 4 \end{pmatrix} \longleftrightarrow \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 &\quad \downarrow \quad \quad \quad \downarrow \\
 &\begin{pmatrix} 1 & 0 & 1/3 \\ 0 & 1 & 2/3 \\ 0 & 2 & 13/3 \end{pmatrix} \longleftrightarrow \begin{pmatrix} 1/3 & 0 & 0 \\ 2/3 & 1 & 0 \\ 1/3 & 0 & 1 \end{pmatrix} \\
 &\quad \downarrow \quad \quad \quad \downarrow \\
 &\begin{pmatrix} 1 & 0 & 1/3 \\ 0 & 1 & 2/3 \\ 0 & 0 & 3 \end{pmatrix} \longleftrightarrow \begin{pmatrix} 1/3 & 0 & 0 \\ 2/3 & 1 & 0 \\ -1 & -2 & 1 \end{pmatrix} \\
 &\quad \downarrow \quad \quad \quad \downarrow \\
 &\begin{pmatrix} 1 & 0 & 1/3 \\ 0 & 1 & 2/3 \\ 0 & 0 & 1 \end{pmatrix} \longleftrightarrow \begin{pmatrix} 3/9 & 0 & 0 \\ 6/9 & 9/9 & 0 \\ -1/3 & -2/3 & 1/3 \end{pmatrix} \\
 &\quad \downarrow \quad \quad \quad \downarrow \\
 &\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \longleftrightarrow \begin{pmatrix} 4/9 & 2/9 & -1/9 \\ 8/9 & 13/9 & -2/9 \\ -1/3 & -2/3 & 1/3 \end{pmatrix} \\
 &\quad \quad \quad \downarrow \\
 &\quad \quad \quad \frac{1}{9} \begin{pmatrix} 4 & 2 & -1 \\ 8 & 13 & -2 \\ -3 & -6 & 3 \end{pmatrix} = ([T]_B)^{-1}.
 \end{aligned}$$

Logo, $[T^{-1}]_B = ([T]_B)^{-1}$ e $T^{-1} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ é dada por:

$$T^{-1}(x_1, x_2, x_3) = \frac{1}{9} (4x_1 + 2x_2 - x_3, 8x_1 + 13x_2 - 2x_3, -3x_1 - 6x_2 + 3x_3).$$

□

2. Funcionais Lineares

Definição 2.1

Seja V um espaço vetorial sobre o corpo $\mathbb{K}$. Um **funcional linear** sobre V é uma transformação linear $f : V \longrightarrow \mathbb{K}$.

Exemplo 2.1

Seja $f : \mathbb{K}^n \longrightarrow \mathbb{K}$ um funcional linear. Então,

$$f(x_1, \dots, x_n) = f\left(\sum_{i=1}^n x_i e_i\right) = \sum_{i=1}^n x_i f(e_i),$$

onde $\{e_1, \dots, e_n\}$ é a base canônica de $\mathbb{K}^n$. Todo funcional linear sobre $\mathbb{K}^n$ é da forma:

$$f(x_1, \dots, x_n) = a_1 x_1 + \dots + a_n x_n,$$

onde $a_i \in \mathbb{K}$, $i = 1, \dots, n$, e $[a_1, \dots, a_n]$ é a matriz de f em relação à base canônica de $\mathbb{K}^n$ e à base $\{1\}$ de $\mathbb{K}$. $\square$

Exemplo 2.2

Seja A uma matriz $n \times n$ sobre o corpo $\mathbb{K}$. O **traço** de A é o escalar

$$\text{Tr } A = A_{11} + A_{22} + \dots + A_{nn}.$$

A **função traço**

$$\begin{aligned} \text{Tr} : \mathbb{K}^{n \times n} &\longrightarrow \mathbb{K} \\ A &\longmapsto \text{Tr } A, \end{aligned}$$

é um funcional linear, pois

$$\text{Tr}(\lambda A + B) = \sum_{i=1}^n (\lambda A + B)_{ii} = \lambda \sum_{i=1}^n A_{ii} + \sum_{i=1}^n B_{ii} = \lambda \text{Tr } A + \text{Tr } B.$$

$\square$

Exemplo 2.3

Seja $\mathcal{C}([a, b], \mathbb{R})$ o espaço das funções contínuas definidas no intervalo $[a, b]$ com valores reais. Então,

$$\begin{aligned} \mathcal{I} : \mathcal{C}([a, b], \mathbb{R}) &\longrightarrow \mathbb{R} \\ f &\longmapsto \int_a^b f(x) dx. \end{aligned}$$

é um funcional linear. $\square$

Definição 2.2

O espaço $\mathcal{L}(V, K)$ de todos os funcionais lineares sobre V é chamado de **espaço dual** de V e se designa por V^* .

Sabemos que se V tem dimensão finita, então V^* também é, e temos $\dim V^* = \dim V$.

Teorema 2.1

Seja V um espaço vetorial sobre $\mathbb{K}$ e seja $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base de V . Então, existe uma única base $\mathcal{B}^* = \{f_1, \dots, f_n\}$ de V^* tal que $f_i(v_j) = \delta_{ij}$, $\forall i, j = 1, \dots, n$.

Para cada funcional linear $f \in V^*$, temos

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

e para todo $v \in V$, temos

$$v = \sum_{i=1}^n f_i(v) v_i.$$

Prova.

Dada a base $\mathcal{B} = \{v_1, \dots, v_n\}$, sejam $f_i : V \rightarrow \mathbb{K}$, $i = 1, \dots, n$, os funcionais lineares que satisfazem

$$f_i(v_j) = \delta_{ij}, \forall j = 1, \dots, n.$$

Para mostrar que $\mathcal{B}^* = \{f_1, \dots, f_n\}$ é uma base de V^* , basta mostrar que os funcionais $f_1, \dots, f_n$ são LI.

De fato, suponhamos que

$$\sum_{i=1}^n c_i f_i = \mathcal{O}.$$

Então,

$$0 = \sum_{i=1}^n c_i f_i(v_j) = \sum_{i=1}^n c_i \delta_{ij} = c_j, \quad \forall j = 1, \dots, n.$$

Logo, $\mathcal{B}^* = \{f_1, \dots, f_n\}$ é uma base de V^* .

Seja, agora, $f \in V^*$. Então, existem $c_1, \dots, c_n \in \mathbb{K}$, tais que

$$f = \sum_{i=1}^n c_i f_i.$$

Ao lado ...
Estamos designando $\mathcal{O}$ o funcional linear zero definido por:
 $\mathcal{O}(v) = 0, \forall v \in V$.

Como $f(v_j) = \sum_{i=1}^n c_i f_i(v_j) = c_j$, $\forall j = 1, \dots, n$, temos que

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

Seja $v \in V$. Então, existem $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, tais que $v = \sum_{i=1}^n \lambda_i v_i$.

Como $f_i(v) = \lambda_i$, temos que $v = \sum_{i=1}^n f_i(v) v_i$. ■

Definição 2.3

A base $\mathcal{B}^* = \{f_1, \dots, f_n\}$ de V^* é dita **dual** da base $\mathcal{B} = \{v_1, \dots, v_n\}$ de V .

Exemplo 2.4

Seja V o espaço das funções polinomiais de $\mathbb{R}$ em $\mathbb{R}$ de grau ≤ 2 . Sejam t_1, t_2, t_3 três números reais distintos e L_1, L_2, L_3 os funcionais lineares sobre V definidos por $L_i(p) = p(t_i)$, $\forall p \in V$.

Esses funcionais são LI, pois se

$$c_1 L_1 + c_2 L_2 + c_3 L_3 = \mathcal{O},$$

temos que $c_1 L_1(p) + c_2 L_2(p) + c_3 L_3(p) = 0$, $\forall p \in V$.

Fazendo $p = 1$, $p = x$ e $p = x^2$, temos que:

$$\begin{aligned} c_1 + c_2 + c_3 &= 0 \\ c_1 t_1 + c_2 t_2 + c_3 t_3 &= 0 \\ c_1 t_1^2 + c_2 t_2^2 + c_3 t_3^2 &= 0. \end{aligned} \quad (1)$$

Como

$$\begin{aligned} \begin{pmatrix} 1 & 1 & 1 \\ t_1 & t_2 & t_3 \\ t_1^2 & t_2^2 & t_3^2 \end{pmatrix} &\longrightarrow \begin{pmatrix} 1 & 1 & 1 \\ 0 & t_2 - t_1 & t_3 - t_1 \\ 0 & t_2^2 - t_1^2 & t_3^2 - t_1^2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & \frac{t_3 - t_1}{t_2 - t_1} \\ 0 & t_2 + t_1 & \frac{t_3^2 - t_1^2}{t_2 - t_1} \end{pmatrix} \\ &\longrightarrow \begin{pmatrix} 1 & 0 & \frac{t_2 - t_1 - (t_3 - t_1)}{t_2 - t_1} \\ 0 & 1 & \frac{t_3 - t_1}{t_2 - t_1} \\ 0 & t_2 + t_1 & \frac{t_3^2 - t_1^2}{t_2 - t_1} \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & \frac{t_2 - t_3}{t_2 - t_1} \\ 0 & 1 & \frac{t_3 - t_1}{t_2 - t_1} \\ 0 & 0 & \frac{(t_3 - t_1)(t_3 - t_2)}{t_2 - t_1} \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \end{aligned}$$

temos que a matriz $\begin{pmatrix} 1 & 1 & 1 \\ t_1 & t_2 & t_3 \\ t_1^2 & t_2^2 & t_3^2 \end{pmatrix}$ é invertível.

Logo, o sistema (1) possui somente a solução trivial $c_1 = c_2 = c_3 = 0$.

Como $\dim V^* = \dim V = 3$ o conjunto LI $\{L_1, L_2, L_3\}$ é uma base de V^* .

Qual é a base $\{p_1, p_2, p_3\}$ de V cuja base dual é $\{L_1, L_2, L_3\}$?

Tal base deve satisfazer

$$L_i(p_j) = p_j(t_i) = \delta_{ij}, \quad \forall i, j = 1, \dots, n.$$

Essas funções polinomiais são dadas por:

$$\begin{aligned} p_1(x) &= \frac{(x - t_2)(x - t_3)}{(t_1 - t_2)(t_1 - t_3)} \\ p_2(x) &= \frac{(x - t_1)(x - t_3)}{(t_2 - t_1)(t_2 - t_3)} \\ p_3(x) &= \frac{(x - t_1)(x - t_2)}{(t_3 - t_1)(t_3 - t_2)}. \end{aligned}$$

Assim, para todo $p \in V$, temos que

$$p = L_1(p) p_1 + L_2(p) p_2 + L_3(p) p_3,$$

ou seja,

$$p = p(t_1) p_1 + p(t_2) p_2 + p(t_3) p_3.$$

Portanto, se c_1, c_2, c_3 são números reais arbitrários, existe uma única função polinomial p sobre $\mathbb{R}$, de grau ≤ 2 , tal que $p(t_i) = c_i$, $i = 1, 2, 3$.

Essa função polinomial é

$$p = c_1 p_1 + c_2 p_2 + c_3 p_3.$$

□

Seja V um espaço vetorial de dimensão n e seja $f : V \rightarrow \mathbb{K}$ um funcional linear não-nulo. Então, **o núcleo de f tem dimensão $n - 1$** , pois o posto de f é 1.

Definição 2.4

Todo subespaço de dimensão $n - 1$ de um espaço vetorial de dimensão n é chamado um **hiperplano** ou um **subespaço de codimensão 1**.

Observação 2.1

Todo subespaço W de codimensão 1 de um espaço vetorial V de dimensão n é o núcleo de um funcional linear não-nulo.

De fato, seja $\{v_1, \dots, v_{n-1}\}$ base de W e $v_n \in V$ tal que $\{v_1, \dots, v_{n-1}, v_n\}$ seja uma base de V .

Seja $f : V \rightarrow \mathbb{K}$ o funcional linear dado por

$$f(v_i) = \begin{cases} 0, & \text{se } i \neq n \\ 1, & \text{se } i = n. \end{cases}$$

Então, f é um funcional linear não-nulo e $v = \lambda_1 v_1 + \dots + \lambda_{n-1} v_{n-1} + \lambda_n v_n$ pertence ao núcleo de f se, e somente se,

$$f(v) = \lambda_n = 0.$$

Ou seja, v pertence ao núcleo de f se, e só se, $v = \lambda_1 v_1 + \dots + \lambda_n v_n \in W$.

Definição 2.5

Seja $S \subset V$ um subconjunto do espaço vetorial V .

O **anulador de S** é o conjunto S^0 de todos os funcionais lineares $f \in V^*$ tais que $f(v) = 0, \forall v \in S$.

Ou seja,

$$S^0 = \{f \in V^* \mid f(v) = 0, \forall v \in S\}.$$

Observação 2.2

- S^0 é um subespaço vetorial de V (verifique!).
- Se $S = \{0\}$, então $S^0 = V$.
- Se $S = V$, então $S^0 = \{0\}$, onde $0 : V \rightarrow \mathbb{K}$ é o funcional linear nulo.

Teorema 2.2

Seja V um espaço vetorial de dimensão finita e seja W um subespaço de V . Então,

$$\dim W + \dim W^0 = \dim V.$$

Prova.

Seja $\{v_1, \dots, v_k\}$ uma base de W e sejam $v_{k+1}, \dots, v_n$ vetores em V tais que $\mathcal{B} = \{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$ seja uma base de V .

Seja $\{f_1, \dots, f_n\}$ a base de V^* dual da base $\mathcal{B}$.

Vamos mostrar que $\{f_{k+1}, \dots, f_n\}$ é uma base de W^0 .

Como $f_j(v) = 0, j \geq k+1$, para todo $v \in W$, temos que $f_j \in W^0, \forall j \geq k+1$.

Sendo que os funcionais lineares $f_{k+1}, \dots, f_n$ são LI, basta mostrar que $\{f_{k+1}, \dots, f_n\}$ gera W^0 .

Seja $f \in V^*$. Então,

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

Logo, se $f \in W^0$, temos que $f(v_i) = 0$, para $i = 1, \dots, k$. Portanto,

$$f = \sum_{i=k+1}^n f(v_i) f_i.$$

Como queríamos demonstrar. ■

Corolário 2.1

Todo subespaço vetorial W de dimensão k de um espaço vetorial V de dimensão n é a interseção de $n - k$ hiperplanos de V .

Prova.

Sejam $f_{k+1}, \dots, f_n$ os funcionais lineares acima. Então $W \subset \bigcap_{j=k+1}^n N_{f_j}$,

onde $N_{f_j} = f_j^{-1}(0)$ é o núcleo de f_j .

Se $v \in \bigcap_{j=k+1}^n N_{f_j}$, temos que $f_j(v) = 0$, para todo $j = k+1, \dots, n$, isto é, se

$v = \lambda_1 v_1 + \dots + \lambda_n v_n$, então $\lambda_j = 0$, para $j = k+1, \dots, n$.

Logo, $v = \lambda_1 v_1 + \dots + \lambda_k v_k \in W$.

Assim, $W = \bigcap_{j=k+1}^n N_{f_j} = \{v \in V; f_j(v) = 0, \forall j = k+1, \dots, n\}$. ■

Corolário 2.2

Sejam W_1, W_2 subespaços vetoriais de um espaço vetorial V de dimensão finita.

Então $W_1 = W_2$ se, e somente se, $W_1^0 = W_2^0$.

Prova.

$(\implies)$ óbvio.

($\Leftarrow$) Suponhamos que existe $v \in W_1$ tal que $v \notin W_2$. Seja $\{v_1, \dots, v_k\}$ uma base de W_2 .

Como $v \notin W_2$, temos que $\{v_1, \dots, v_k, v\}$ é LI.

Seja $v_{k+1} = v$, e sejam $v_{k+1}, \dots, v_n \in V$ tais que $\{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$ é uma base de V .

Seja $f: V \rightarrow \mathbb{K}$ um funcional linear, tal que $f(v) = f(v_{k+1}) = 1$ e $f(v_i) = 0$, para $i \neq k+1$.

Então, $f \in W_2^0$ e $f \notin W_1^0$, ou seja, $W_1^0 \neq W_2^0$. ■

Pergunta:

Toda base de V^* é a base dual de alguma base de V ?

Seja $v \in V$. Consideremos a função $L_v: V^* \rightarrow \mathbb{K}$ definida por $L_v(f) = f(v)$, $\forall f \in V^*$.

A função L_v definida ao lado é chamada a *avaliação em v* e, segundo a afirmação, resulta ser um funcional linear sobre o espaço dual V^* .

Afirmação. L_v é linear.

De fato, para quaisquer $f, g \in V^*$ e $\lambda \in \mathbb{K}$, temos

$$L_v(\lambda f + g) = (\lambda f + g)(v) = \lambda f(v) + g(v) = \lambda L_v(f) + L_v(g).$$

Então, $L_v \in (V^*)^*$.

Designaremos $(V^*)^*$ apenas por V^{**} . Esse espaço é chamado o *duplo dual* ou *bidual* de V .

Teorema 2.3

Seja V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{K}$. Então,

$$\begin{aligned} \mathcal{L}: V &\rightarrow V^{**} \\ v &\mapsto \mathcal{L}(v) = L_v, \end{aligned}$$

é um isomorfismo linear de V em V^{**} .

Prova.

- $\mathcal{L}$ é linear.

Com efeito, dados $v, w \in V$, $\lambda \in \mathbb{K}$ e $f \in V^*$ arbitrários. Temos que

$$\begin{aligned} \mathcal{L}(\lambda v + w)(f) &= L_{\lambda v + w}(f) = f(\lambda v + w) = \lambda f(v) + f(w) \\ &= \lambda L_v(f) + L_w(f) = \lambda \mathcal{L}(v)(f) + \mathcal{L}(w)(f) \\ &= (\lambda \mathcal{L}(v) + \mathcal{L}(w))(f). \end{aligned}$$

Portanto, $\mathcal{L}(\lambda v + w) = \lambda \mathcal{L}(v) + \mathcal{L}(w)$, isto é, $\mathcal{L}$ é linear.

- $\mathcal{L}$ é injetora.

Seja $v \in V - \{0\}$. Vamos mostrar que $\mathcal{L}(v) = L_v \neq 0$, ou seja, que existe $f \in V^*$, tal que $L_v(f) = f(v) \neq 0$.

Como $v \neq 0$, existem $v_2, \dots, v_n \in V$, onde $n = \dim V$, tais que $\{v = v_1, v_2, \dots, v_n\}$ é uma base de V .

Seja $f : V \rightarrow \mathbb{K}$ o funcional linear definido por :

$$f(v_i) = \begin{cases} 1 & \text{se } i = 1 \\ 0 & \text{se } i \in \{2, \dots, n\}. \end{cases}$$

Então, $f(v) = L_v(f) = 1 \neq 0$.

Provamos assim que a aplicação $\mathcal{L} : V \rightarrow V^{**}$ é linear e injetora.

Como V é de dimensão finita, temos

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

Logo, $\mathcal{L}$ é uma transformação linear invertível, ou seja, é um isomorfismo de V em V^{**} . ■

Corolário 2.3

Seja V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{K}$. Toda base de V^* é a base dual de alguma base de V .

Prova.

Seja $\mathcal{B}^* = \{f_1, \dots, f_n\}$ uma base de V^* . e seja $\mathcal{B}^{**} = \{L_1, \dots, L_n\}$ a base de V^{**} dual da base $\mathcal{B}^*$, isto é,

$$L_i(f_j) = \delta_{ij}, \forall i, j = 1, \dots, n.$$

Pelo teorema anterior, para cada $i = 1, \dots, n$, existe $v_i \in V$, tal que $L_i = L_{v_i}$. Como a aplicação $v \mapsto L_v$ é um isomorfismo linear de V em V^{**} , temos que $\mathcal{B} = \{v_1, \dots, v_n\}$ é uma base de V . Além disso, temos

$$\delta_{ij} = L_i(f_j) = L_{v_i}(f_j) = f_j(v_i), \forall i, j = 1, \dots, n,$$

ou seja, $\mathcal{B}^*$ é a base dual da base $\mathcal{B}$. ■

Observação 2.3

- Devido ao teorema anterior, geralmente identificamos v com L_v e dizemos que V é o espaço dual de V^* ou que os espaços V e V^* são duais um do outro de uma maneira natural.

- Seja $B \subset V^*$. Então B^0 é, por definição, um subespaço de V^{**} formado por todos os funcionais lineares $L \in V^{**}$ tal que $L(f) = 0, \forall f \in B$.

Mas como todo $L \in V^{**}$ é da forma L_v para algum $v \in V$, podemos definir B^0 como sendo o conjunto:

$$B^0 = \{v \in V \mid f(v) = 0, \quad \forall f \in B\}.$$

- Assim, B^0 é a interseção de todos os núcleos dos funcionais $f \in B$.
- Já vimos, também, que se W é um subespaço de V , então

$$W = \{v \in V \mid f(v) = 0, \quad \forall f \in W^0\}.$$

Logo, $(W^0)^0 = W$.

Teorema 2.4

Se S é um subconjunto qualquer de um espaço vetorial V de dimensão finita, então $(S^0)^0$ é o subespaço gerado por S .

O duplo anulador de um subconjunto S , designado S^{00} , é por definição $(S^0)^0$. Análogamente, $W^{00} = (W^0)^0$.

Prova.

Seja W o subespaço gerado por S . Então, $S^0 = W^0$.

Logo, $(S^0)^0 = (W^0)^0 = W$ é o subespaço gerado por S .

Outra demonstração:

Como $\dim W + \dim W^0 = \dim V$ e $\dim W^0 + \dim (W^0)^0 = \dim V^* = \dim V$, temos que $\dim (W^0)^0 = \dim W$.

Além disso, como $W \subset (W^0)^0$, concluímos que $W = (W^0)^0$. ■

Definição 2.6

Seja V um espaço vetorial. Um subconjunto N de V é um *hiperplano* de V se:

1. N é um subespaço próprio de V , isto é, $N \neq V$.
2. Se W é um subespaço de V que contém N , então $W = N$ ou $W = V$.

Em suma: $N \subset V$ é um hiperplano de V se N é um subespaço próprio maximal de V .

Teorema 2.5

Se f é um funcional linear não-nulo sobre o espaço vetorial V , então o núcleo de f é um hiperplano de V .

Reciprocamente, todo hiperplano de V é o núcleo de um funcional linear não nulo sobre V .

Prova.

Seja $f : V \rightarrow \mathbb{K}$ um funcional linear não-nulo sobre V e N_f o seu núcleo.

Seja $v \in V$ tal que $f(v) \neq 0$.

Sejam $w \in V$ e $c = \frac{f(w)}{f(v)}$. Então, o vetor $u = w - cv$ pertence a N_f , pois

$$f(u) = f(w - cv) = f(w) - cf(v) = 0.$$

Isto é, todo vetor $w \in V$ pertence ao subespaço gerado por N_f e v .

Provamos, assim, que se $v \notin N_f$, o espaço gerado por N_f e v é o espaço V . Logo, N_f é um subespaço próprio maximal.

Seja agora N um hiperplano em V e seja v um vetor que não pertence a N . Como N é um hiperplano, o subespaço gerado por N e v é todo o espaço V .

Logo, todo vetor w em V se escreve na forma

$$w = u + cv, \text{ com } c \in \mathbb{K} \text{ e } u \in N.$$

O vetor $u \in N$ e o escalar $c \in \mathbb{K}$ são determinados de modo único por w .

De fato, se também tivéssemos

$$w = u' + c'v, \text{ com } c' \in \mathbb{K} \text{ e } u' \in N,$$

então, $u - u' = (c' - c)v$. Se $c \neq c'$ obteríamos que $v \in N$, o que não é possível pela própria escolha de v . Portanto, $c = c'$ e $u = u'$.

A aplicação $f : V \rightarrow \mathbb{K}$ que a cada $w = u + cv$ faz corresponder o escalar c é linear e $N_f = N$ (verifique!). ■

Lema 2.1

Se f e g são funcionais lineares sobre um espaço vetorial V , então g é múltiplo de f se, e somente se, $N_f \subset N_g$.

Prova.

($\Rightarrow$) óbvio.

($\Leftarrow$) Se $f = \mathcal{O}$, então $g = \mathcal{O}$. Logo, g é múltiplo de f .

Se $f \neq \mathcal{O}$, o núcleo de f , N_f , é um hiperplano de V .

Seja $v \in V$ tal que $f(v) \neq 0$, e seja $c = \frac{g(v)}{f(v)}$.

O funcional $h = g - cf$ é nulo em N_f , pois, como $N_f \subset N_g$, tanto f como g são nulos em N_f e $h(v) = g(v) - cf(v) = 0$. Logo, h é nulo no espaço gerado por N_f e v , que é todo o espaço V . Logo, $h = 0$.

Isto é, $g = cf$, ou seja, g é múltiplo de f . ■

Teorema 2.6

Sejam $g, f_1, \dots, f_r$ funcionais lineares sobre um espaço vetorial V e sejam $N, N_1, \dots, N_r$ seus respectivos núcleos.

Então g é uma combinação linear de $f_1, \dots, f_r$ se, e somente se,

$$N_1 \cap \dots \cap N_r \subset N.$$

Prova.

Se $g = c_1 f_1 + \dots + c_r f_r$ e $f_i(v) = 0$ para cada $i = 1, \dots, r$, então, $g(v) = 0$. Portanto, $N_1 \cap \dots \cap N_r \subset N$.

Vamos provar a recíproca usando indução sobre o número r .

(1) A propriedade vale para $r = 1$ em virtude do lema anterior.

(2) Suponhamos o resultado válido para $r = k - 1$.

Sejam $g, f_1, \dots, f_k$ funcionais lineares tais que $N_1 \cap \dots \cap N_k \subset N_g$.

Sejam $g', f'_1, \dots, f'_{k-1}$ as restrições de $g, f_1, \dots, f_{k-1}$ ao subespaço N_k .

Se $v \in N_k$ e $f_i(v) = 0, \forall i = 1, \dots, k-1$, então $v \in N_1 \cap \dots \cap N_{k-1} \cap N_k \subset N_g$.

Logo, $g(v) = 0$.

Pela hipótese de indução, existem escalares $c_1, \dots, c_{k-1}$ tais que,

$$g' = c_1 f'_1 + \dots + c_{k-1} f'_{k-1}.$$

Seja $h = g - \sum_{i=1}^{k-1} c_i f_i$. Então, h é um funcional linear sobre V tal que

$h(v) = 0$, para todo $v \in N_k$.

Pelo lema anterior, existe $c_k \in \mathbb{K}$, tal que $h = c_k f_k$, ou seja,

$$g = \sum_{i=1}^k c_i f_i.$$

■

Definição 2.7

A **transposta** de uma transformação linear $T : V \longrightarrow W$ é a aplicação

$$\begin{aligned} T^t : W^* &\longrightarrow V^* \\ g &\longmapsto T^t(g) = g \circ T. \end{aligned}$$

Observação 2.4

A transposta $T^t : W^* \longrightarrow V^*$ da transformação linear $T : V \longrightarrow W$ é uma transformação linear.

De fato, se $g_1, g_2 \in W^*$ e $\lambda \in \mathbb{K}$, temos

$$\begin{aligned} T^t(\lambda g_1 + g_2) &= (\lambda g_1 + g_2) \circ T \\ &= \lambda(g_1 \circ T) + g_2 \circ T \\ &= \lambda T^t(g_1) + T^t(g_2). \end{aligned}$$

Proposição 2.1

Sejam V e W espaços vetoriais sobre o corpo $\mathbb{K}$ e seja $T : V \longrightarrow W$ uma transformação linear. O núcleo de $T^t : W^* \longrightarrow V^*$ é o anulador da imagem $T(V)$ de T .

Além disso, se V e W têm dimensão finita, então:

1. $\text{posto}(T^t) = \text{posto}(T)$;
2. A imagem de T^t é o anulador do núcleo de T : $T^t(W^*) = (N_T)^0$.

Prova.

Se $g \in W^*$ pertence ao núcleo de T^t , então $T^t(g) = g \circ T = \mathcal{O}$, ou seja, $g(T(v)) = 0, \forall v \in V$.

Logo, $N_{T^t} = (T(V))^0$.

Suponhamos agora, que V e W têm dimensão finita, sendo $\dim V = n$ e $\dim W = m$.

Seja $r = \text{posto}(T) = \dim T(V)$. Então $\dim T(V)^0 = m - r$. Pela primeira afirmação do teorema, temos que a nulidade de T^t é $m - r$.

Então, $\text{posto}(T^t) = m - (m - r) = r$, ou seja, $\text{posto}(T^t) = \text{posto}(T)$.

Seja agora $N = N_T$ o núcleo de T e seja $f = T^t(g) = g \circ T$ um funcional na imagem de T^t .

Se $v \in N$, temos que $f(v) = (g \circ T)(v) = g(T(v)) = g(\mathbf{0}) = 0$, ou seja,

$$f = T^t(g) \in N^0.$$

Como $T^t(W^*) \subset N^0$ e

$$\dim N^0 = n - \dim N = \text{posto}(T) = \text{posto}(T^t),$$

obtemos que $T^t(W^*) = N^0$. ■

Definição 2.8

Se A é uma matriz $m \times n$ com entradas no corpo $\mathbb{K}$, a **transposta** de A é a matriz $n \times m$ A^t , definida por $A_{ij}^t = A_{ji}$.

Proposição 2.2

Sejam V e W espaços vetoriais de dimensão finita sobre o corpo $\mathbb{K}$.

Seja $\mathcal{B}$ uma base de V com base dual $\mathcal{B}^*$, e seja Ω uma base de W com base dual Ω^* . Então,

$$[T^t]_{\mathcal{B}^* \Omega^*} = ([T]_{\Omega \mathcal{B}})^t.$$

Lembre que, se
 $T: V_{\mathcal{B}} \rightarrow W_{\Omega}$
 $\Rightarrow T^t: W_{\Omega}^* \rightarrow V_{\mathcal{B}}^*$

Prova.

Sejam a matriz $m \times n$ $A = [T]_{\Omega \mathcal{B}}$ e a matriz $n \times m$ $B = [T^t]_{\mathcal{B}^* \Omega^*}$.

Sejam

$$\begin{aligned} \mathcal{B} &= \{v_1, \dots, v_n\}, & \Omega &= \{w_1, \dots, w_m\}, \\ \mathcal{B}^* &= \{f_1, \dots, f_n\}, & \Omega^* &= \{g_1, \dots, g_m\}. \end{aligned}$$

Por definição, temos

$$\begin{aligned} T(v_j) &= \sum_{i=1}^m A_{ij} w_i, & j &= 1, \dots, n \\ T^t(g_j) &= \sum_{i=1}^n B_{ij} f_i, & j &= 1, \dots, m. \end{aligned}$$

Por outro lado,

$$\begin{aligned} T^t(g_j)(v_i) &= g_j(T(v_i)) = g_j\left(\sum_{k=1}^m A_{ki} w_k\right) \\ &= \sum_{k=1}^m A_{ki} g_j(w_k) = \sum_{k=1}^m A_{ki} \delta_{jk} \\ &= A_{ji}. \end{aligned}$$

Sabemos que para todo funcional $f \in V^*$,

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

Logo,

$$T^t(g_j) = \sum_{i=1}^n T^t g_j(v_i) f_i = \sum_{i=1}^n A_{ji} f_i$$

ou seja, $B_{ij} = A_{ji}$. ■

Teorema 2.7

Seja A uma matriz $m \times n$ sobre o corpo $\mathbb{K}$.

Então, o posto-linha de A é igual ao posto-coluna de A .

Prova.

Seja $T : \mathbb{K}^{n \times 1} \longrightarrow \mathbb{K}^{m \times 1}$ a transformação linear dada por $T(X) = AX$.

Sejam $\mathcal{B}$ e Ω as bases canônicas de $\mathbb{K}^{n \times 1}$ e $\mathbb{K}^{m \times 1}$, respectivamente.

Então, $A = [T]_{\mathcal{B}\Omega}$.

O posto-coluna de A é o posto da transformação T , pois a imagem de T consiste de todas as m -úplas que são combinações lineares dos vetores-colunas de A .

Em relação às bases duais Ω^* e $\mathcal{B}^*$, a aplicação transposta T^t é representada pela matriz A^t . Como os vetores-colunas de A^t são os vetores-linhas de A , temos que o posto-coluna de A^t é igual ao posto-linha de A .

Além disso, como a transformação linear $L : (\mathbb{K}^{n \times 1})^* \longrightarrow \mathbb{K}^{n \times 1}$ que leva f em $[f]_{\mathcal{B}^*}$ é um isomorfismo e

$$\begin{aligned} L(T^t(W^*)) &= L(\{x_1 T^t(g_1) + \dots + x_m T^t(g_m) \mid x_1, \dots, x_m \in \mathbb{K}\}) \\ &= \{x_1 \beta_1 + \dots + x_m \beta_m \mid x_1, \dots, x_m \in \mathbb{K}\} \end{aligned}$$

é o subespaço gerado pelos vetores-colunas β_j , $j = 1, \dots, m$, da matriz A^t , temos que

$$\begin{aligned} \text{posto}(T^t) &= \dim T^t(W^*) = \text{posto-coluna}(A^t) \\ &= \text{posto-linha}(A). \end{aligned}$$

Mas, pelo teorema anterior, $\text{posto}(T) = \text{posto}(T^t)$.

Logo, $\text{posto-coluna}(A) = \text{posto}(T) = \text{posto}(T^t) = \text{posto-linha}(A)$. ■



Polinômios

1. Álgebras - generalidades

Definição 1.1

Uma **álgebra linear** sobre o corpo $\mathbb{K}$ é um espaço vetorial $\mathcal{A}$ sobre $\mathbb{K}$ com uma operação adicional, chamada **multiplicação de vetores**, que a cada par de vetores $\alpha, \beta \in \mathcal{A}$ associa um vetor de $\mathcal{A}$, designado $\alpha \cdot \beta$ ou $\alpha\beta$, satisfazendo as seguintes propriedades:

1. associatividade:

$$\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma.$$

2. distributividade em relação à adição:

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= \alpha \cdot \beta + \alpha \cdot \gamma, \\ (\alpha + \beta) \cdot \gamma &= \alpha \cdot \gamma + \beta \cdot \gamma.\end{aligned}$$

3. Para cada $c \in \mathbb{K}$,

$$c(\alpha \cdot \beta) = (c\alpha) \cdot \beta = \alpha \cdot (c\beta).$$

Além disso, dizemos que

- $\mathcal{A}$ é uma **álgebra com elemento unidade** se existir um vetor $\mathbf{1} \in \mathcal{A}$ tal que $\mathbf{1} \cdot \alpha = \alpha \cdot \mathbf{1} = \alpha, \forall \alpha \in \mathcal{A}$.
- $\mathcal{A}$ é uma **álgebra comutativa** se $\alpha \cdot \beta = \beta \cdot \alpha, \forall \alpha, \beta \in \mathcal{A}$.

Exemplo 1.1

a. Um corpo $\mathbb{K}$ é uma álgebra linear comutativa com elemento unidade.

b. O conjunto $\mathbb{K}^{n \times n}$ das matrizes $n \times n$ sobre o corpo $\mathbb{K}$ é uma álgebra linear com elemento unidade (a matriz identidade) e que é não-comutativa se $n \geq 2$.

c. O conjunto $\mathcal{L}(V, V)$ dos operadores lineares sobre o espaço vetorial V com a composição de operadores como sendo o produto, é uma álgebra linear com elemento unidade que é não-comutativa se $\dim V \geq 2$.

d. Seja $\mathbb{K}$ um corpo e $\mathcal{F}(\mathbb{N}, \mathbb{K})$ o espaço vetorial de todas as funções de $\mathbb{N}$ em $\mathbb{K}$. Denotamos esse espaço vetorial por $\mathbb{K}^\infty$ e seus vetores $f \in \mathcal{F}(\mathbb{N}, \mathbb{K})$ são escritos como seqüências infinitas $f = (f_0, f_1, \dots, f_n, \dots)$, onde $f_i = f(i)$, para todo $i \in \mathbb{N}$.

Se $f = (f_0, f_1, \dots, f_n, \dots)$ e $g = (g_0, g_1, \dots, g_n, \dots)$ são vetores de $\mathbb{K}^\infty$ e $\lambda \in \mathbb{K}$, já sabemos que

$$\lambda f + g = (\lambda f_0 + g_0, \lambda f_1 + g_1, \dots, \lambda f_n + g_n, \dots).$$

Vamos agora definir um produto de vetores em $\mathbb{K}^\infty$.

Dados $f = (f_0, f_1, \dots, f_n, \dots)$, $g = (g_0, g_1, \dots, g_n, \dots) \in \mathbb{K}^\infty$, definimos o vetor produto $f \cdot g$ por:

$$(f \cdot g)_n = \sum_{i=0}^{\infty} f_i g_{n-i}, \quad n \in \mathbb{N}.$$

Assim,

$$f \cdot g = (f_0 g_0, f_0 g_1 + f_1 g_0, f_0 g_2 + f_1 g_1 + f_2 g_0, \dots).$$

• Como

$$(f \cdot g)_n = \sum_{i=0}^{\infty} f_i g_{n-i} = \sum_{i=0}^{\infty} g_i f_{n-i} = (g \cdot f)_n, \quad \forall n \in \mathbb{N},$$

temos que a multiplicação em $\mathbb{K}^\infty$ é comutativa.

• A multiplicação é, também, distributiva.

De fato, se $h = (h_0, h_1, \dots, h_n, \dots) \in \mathbb{K}^\infty$, então

$$\begin{aligned} (f \cdot (g + h))_n &= \sum_{i=0}^n f_i (g + h)_{n-i} = \sum_{i=0}^n f_i (g_{n-i} + h_{n-i}) \\ &= \sum_{i=0}^n f_i g_{n-i} + \sum_{i=0}^n f_i h_{n-i} \\ &= (f \cdot g)_n + (f \cdot h)_n, \quad \forall n \in \mathbb{N}, \end{aligned}$$

ou seja, $f \cdot (g + h) = f \cdot g + f \cdot h$.

De modo análogo, podemos provar que $(f + g) \cdot h = f \cdot h + g \cdot h$.

- A multiplicação satisfaz a propriedade:

$$c(f \cdot g) = (cf) \cdot g = f \cdot (cg), \quad \forall c \in \mathbb{K}, \forall f, g \in \mathbb{K}^\infty.$$

De fato:

$$(c(f \cdot g))_n = c(f \cdot g)_n = c \sum_{i=0}^n f_i g_{n-i} = \sum_{i=0}^n (cf)_i g_{n-i} = ((cf) \cdot g)_n$$

e

$$(c(f \cdot g))_n = c(f \cdot g)_n = c \sum_{i=0}^n f_i g_{n-i} = \sum_{i=0}^n f_i (cg_{n-i}) = (f \cdot (cg))_n,$$

para todo $n \in \mathbb{N}$, ou seja $c(f \cdot g) = (cf) \cdot g = f \cdot (cg)$.

- A multiplicação é associativa.

De fato:

$$((f \cdot g) \cdot h)_n = \sum_{i=0}^n (f \cdot g)_i h_{n-i} = \sum_{i=0}^n \left(\sum_{j=0}^i f_j g_{i-j} \right) h_{n-i}.$$

Como (ver diagrama ao lado)

$$\{(i, j) \mid i \in \{0, \dots, n\} \text{ e } j \in \{0, \dots, i\}\} = \{(i, j) \mid j \in \{0, \dots, n\} \text{ e } i \in \{j, \dots, n\}\},$$

temos que

$$((f \cdot g) \cdot h)_n = \sum_{j=0}^n \left(\sum_{i=j}^n f_j g_{i-j} \right) h_{n-i} = \sum_{j=0}^n f_j \left(\sum_{i=j}^n g_{i-j} h_{n-i} \right).$$

Fazendo $k = i - j$, temos que k varia de 0 a $n - j$ quando i varia de j a n ,
e $n - i = n - (k + j) = (n - j) - k$.

Logo, para todo $n \in \mathbb{N}$, temos

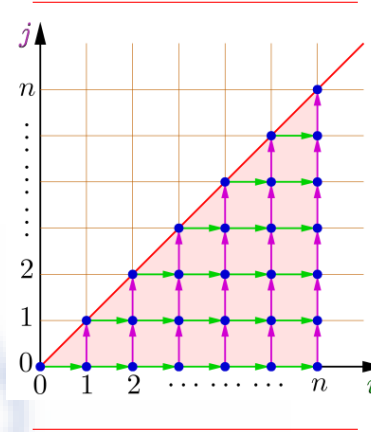
$$\begin{aligned} ((f \cdot g) \cdot h)_n &= \sum_{j=0}^n f_j \left(\sum_{i=j}^n g_{i-j} h_{n-i} \right) = \sum_{j=0}^n f_j \sum_{k=0}^{n-j} g_k h_{(n-j)-k} \\ &= \sum_{j=0}^n f_j (g \cdot h)_{n-j} = (f \cdot (g \cdot h))_n. \end{aligned}$$

Ou seja, $(f \cdot g) \cdot h = f \cdot (g \cdot h)$.

- O vetor $\mathbf{1} = (1, 0, 0, \dots)$ é o elemento unidade da multiplicação.

De fato,

$$(1 \cdot f)_n = \sum_{i=0}^n (1)_i (f)_{n-i} = 1 f_{n-0} + 0 f_{n-1} + 0 f_{n-2} + \dots = f_n, \quad \forall n \in \mathbb{N},$$



ou seja, $\mathbf{1} \cdot f = f \cdot \mathbf{1} = f$.

Portanto, $\mathbb{K}^\infty$ com as operações definidas acima é uma álgebra linear comutativa com elemento unidade sobre o corpo $\mathbb{K}$.

Notação: Seja $x = (0, 1, 0, \dots) \in \mathbb{K}^\infty$. O produto de x por si mesmo n vezes será indicado por x^n e convencionamos que $x^0 = \mathbf{1} = (1, 0, 0, \dots)$.

Então, observe que

$$x^2 = (0, 0, 1, 0, 0, \dots), \quad x^3 = (0, 0, 0, 1, 0, \dots),$$

e, em geral, para todo $n \in \mathbb{N}$, temos

$$(x^n)_k = \begin{cases} 1 & \text{se } k = n \\ 0 & \text{se } k \neq n \end{cases}, \forall k \in \mathbb{N}. \quad \text{Ou seja, } (x^n)_k = \delta_{nk}, \forall k \in \mathbb{N}.$$

De fato,

$$\bullet (x \cdot x)_n = \sum_{i=0}^n (x)_i (x)_{n-i} = (x)_1 (x)_{n-1} = \begin{cases} 1, & \text{se } n = 2 \\ 0, & \text{se } n \neq 2, \end{cases}$$

ou seja, $x^2 = (0, 0, 1, 0, 0, \dots)$.

• Suponhamos, por indução, que $(x^k)_n = 0$, se $n \neq k$ e $(x^k)_k = 1$.

Então,

$$(x^{k+1})_n = (x^k \cdot x)_n = \sum_{i=0}^n (x^k)_i (x)_{n-i} = \begin{cases} 0, & \text{se } n < k; \\ (x^k)_k \cdot (x)_{n-k} = x_{n-k}, & \text{se } n \geq k. \end{cases}$$

Mas, como $(x)_{n-k} = 1$ se, e só se, $n = k + 1$ e $(x)_{n-k} = 0$ se $n = k$ e $n \geq k + 2$, temos que

$$(x^{k+1})_n = \begin{cases} 0, & \text{se } n \neq k+1 \\ 1, & \text{se } n = k+1. \end{cases}$$

□

Observação 1.1

Como o conjunto $\{\mathbf{1}, x, x^2, x^3, \dots\}$ é LI, $\mathbb{K}^\infty$ não tem dimensão finita.

Definição 1.2

Seja $\mathbb{K}[x]$ o subespaço de $\mathbb{K}^\infty$ gerado pelos vetores $\mathbf{1}, x, x^2, \dots, x^n, \dots$. Os vetores de $\mathbb{K}[x]$ são denominados *polinômios* em $\mathbb{K}$ (ou, com coeficientes em $\mathbb{K}$).

Se $f = (f_0, f_1, \dots, f_n, \dots) \in \mathbb{K}[x]$ é um polinômio não-nulo, existe $n \geq 0$ tal que $f_n \neq 0$ e $f_k = 0, \forall k > n$.

Esse inteiro n , que é único para cada $f \in \mathbb{K}[x]$, é denominado o **grau** de f .

Então, se f é um polinômio não-nulo de grau n ,

$$\begin{aligned} f &= (f_0, f_1, \dots, f_n, 0, 0, \dots) \\ &= f_0x^0 + f_1x + f_2x^2 + \dots + f_nx^n \\ &= f_0\mathbf{1} + f_1x + f_2x^2 + \dots + f_nx^n. \end{aligned}$$

- Os escalares $f_0, f_1, \dots, f_n$ são denominados os **coeficientes** de f .
- Os polinômios da forma cx^0 , ou seja $c\mathbf{1}$, são chamados **polinômios constantes**. Frequentemente, escreveremos apenas c em vez de cx^0 e/ou $c\mathbf{1}$.
- Um polinômio não-nulo f de grau n , tal que $f_n = 1$ é dito um **polinômio unitário**.

Observação 1.2

Ao polinômio nulo não será atribuído grau algum.

Teorema 1.1

O conjunto dos polinômios sobre um corpo $\mathbb{K}$ é uma álgebra linear comutativa com elemento unidade sobre $\mathbb{K}$.

Lema 1.1

Sejam f, g polinômios não-nulos sobre $\mathbb{K}$. Então,

- $f \cdot g$ é um polinômio não-nulo e $\text{grau}(f \cdot g) = \text{grau}(f) + \text{grau}(g)$.
- $f \cdot g$ é um polinômio unitário se f e g são polinômios unitários.
- $f \cdot g$ é um polinômio constante se, e somente se, f e g são constantes.
- Se $f + g \neq \mathbf{0}$, então $\text{grau}(f + g) \leq \max\{\text{grau}(f), \text{grau}(g)\}$.

Prova.

Suponhamos que $\text{grau}(f) = m$ e $\text{grau}(g) = n$ e seja $k \in \mathbb{N}$. Então,

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

Para que $f_i g_{m+n+k-i} \neq 0$ é necessário que $i \leq m$ e $m+n+k-i \leq n$. Logo, é necessário que $m+k \leq i \leq m$, o que implica que $k=0$ e $i=m$.

Assim,

$$(f \cdot g)_{m+n} = f_m g_n$$

e

$$(f \cdot g)_{m+n+k} = 0, \text{ se } k > 0.$$

Logo, $f \cdot g$ é um polinômio de grau $m + n$.

Além disto, se $f_m = g_n = 1$, $(f \cdot g)_{m+n} = 1$, ou seja, se f e g são unitários, então, também, $f \cdot g$ é unitário.

Do anterior, temos que

$$\text{grau}(f) = \text{grau}(g) = 0 \iff \text{grau}(f) + \text{grau}(g) = 0 \iff \text{grau}(f \cdot g) = 0,$$

ou seja, f e g são polinômios constantes se, e somente se, $f \cdot g$ é um polinômio constante.

Suponhamos agora que $f + g \neq \mathbf{0}$ e seja $k = \max\{\text{grau}(f), \text{grau}(g)\}$.

Então,

$$(f + g)_\ell = f_\ell + g_\ell = 0, \quad \forall \ell > k.$$

Logo, $\text{grau}(f + g) \leq k$. ■

Corolário 1.1

Sejam $f, g, h \in \mathbb{K}[x]$. Se $f \neq \mathbf{0}$ e $f \cdot g = f \cdot h$, então $g = h$.

Prova.

Como $f \cdot g = f \cdot h$, temos $f \cdot (g - h) = \mathbf{0}$, e como $f \neq \mathbf{0}$, segue de **a.** do lema anterior, que $g - h = \mathbf{0}$. Ou seja $g = h$. ■

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

Então, pela definição da multiplicação de polinômios, temos que:

$$f \cdot g = \sum_{s=0}^{m+n} \left(\sum_{r=0}^s f_r g_{s-r} \right) x^s.$$

No caso particular em que $f = c x^m$ e $g = d x^n$, com $c, d \in \mathbb{K}$, temos

$$f \cdot g = cd x^{m+n}. \quad (1)$$

Dessa relação e das propriedades distributivas de $\mathbb{K}[x]$, segue que

$$f \cdot g = \sum_{i=0}^m \sum_{j=0}^n f_i g_j x^i x^j = \sum_{i,j} f_i g_j x^{i+j}.$$

Definição 1.3

Seja $\mathcal{A}$ uma álgebra linear com elemento unidade sobre o corpo $\mathbb{K}$. Indicaremos o elemento unidade de $\mathcal{A}$ por $\mathbf{1}$ e faremos a convenção $\alpha^0 = \mathbf{1}$, para todo $\alpha \in \mathcal{A}$. Então, a cada polinômio $f = \sum_{i=0}^n f_i x^i$ sobre $\mathbb{K}$ e a cada $\alpha \in \mathcal{A}$, associamos o elemento $f(\alpha) \in \mathcal{A}$, denominado a **avaliação de f em α** , segundo a regra:

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

Exemplo 1.2

Seja $\mathbb{K} = \mathbb{C}$ e seja $f = x^2 + x + 2 \in \mathbb{C}[x]$. Então,

• Se $\mathcal{A} = \mathbb{C}$ e $\alpha = 2 \in \mathcal{A}$, $f(\alpha) = 2^2 + 2 + 2 = 8$.

• Se $\mathcal{A} = \mathbb{C}^{2 \times 2}$ e $\alpha = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$, temos

$$f(\alpha) = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} + 2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix} + \begin{pmatrix} 3 & 1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 6 & 3 \\ 6 & 6 \end{pmatrix}$$

• se $\mathcal{A} = \mathcal{L}(\mathbb{C}^2, \mathbb{C}^2)$ e $\alpha = T : \mathbb{C}^2 \longrightarrow \mathbb{C}^2$ é o operador linear

$$T(z_1, z_2) = (2z_1 + 3z_2, 4z_2),$$

então, $f(\alpha) = T^2 + T + 2I \in \mathcal{A}$ é o operador linear

$$\begin{aligned} f(T)(z_1, z_2) &= T(T(z_1, z_2)) + T(z_1, z_2) + 2I(z_1, z_2) \\ &= T(2z_1 + 3z_2, 4z_2) + (2z_1 + 3z_2, 4z_2) + (2z_1, 2z_2) \\ &= (2(2z_1 + 3z_2) + 3(4z_2), 4(4z_2)) + (4z_1 + 3z_2, 6z_2) \\ &= (4z_1 + 18z_2, 16z_2) + (4z_1 + 3z_2, 6z_2) \\ &= (8z_1 + 21z_2, 22z_2). \end{aligned}$$

□

Teorema 1.2

Seja $\mathcal{A}$ uma álgebra linear com elemento unidade sobre $\mathbb{K}$, e sejam $\alpha \in \mathcal{A}$, $c \in \mathbb{K}$, f e g polinômios sobre $\mathbb{K}$. Então,

a. $(cf + g)(\alpha) = cf(\alpha) + g(\alpha),$

b. $(f \cdot g)(\alpha) = f(\alpha)g(\alpha).$

Prova.

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

Suponhamos, também, que $n \geq m$. Então,

$$cf + g = \sum_{i=0}^n (cf_i + g_i) x^i, \quad \text{onde } g_j = 0 \text{ se } j > m.$$

Logo,

$$\begin{aligned} (cf + g)(\alpha) &= \sum_{i=0}^n (cf_i + g_i) \alpha^i = c \sum_{i=0}^n f_i \alpha^i + \sum_{i=0}^n g_i \alpha^i \\ &= c \sum_{i=0}^n f_i \alpha^i + \sum_{i=0}^m g_i \alpha^i \\ &= cf(\alpha) + g(\alpha). \end{aligned}$$

Como $f \cdot g = \sum_{i,j} f_i g_j x^{i+j}$ (ver pag. 100), por (1) temos que $f \cdot g(\alpha) =$

$\sum_{i,j} f_i g_j \alpha^{i+j}$. Logo

$$f \cdot g(\alpha) = \left(\sum_{i=0}^n f_i \alpha^i \right) \left(\sum_{j=0}^m g_j \alpha^j \right) = f(\alpha) g(\alpha).$$



2. Interpolação de Lagrange

Seja V o subespaço de $\mathbb{K}[x]$ gerado pelos polinômios $1, x, x^2, \dots, x^n$, ou seja, V é o subespaço de $\mathbb{K}[x]$ formado por todos os polinômios de grau $\leq n$, junto com o polinômio nulo. Então, **$\dim V = n + 1$** .

Sejam $t_0, t_1, \dots, t_n \in \mathbb{K}$, $n + 1$ escalares distintos, e $L_i : V \rightarrow \mathbb{K}$ as aplicações definidas por $L_i(f) = f(t_i)$, $i = 0, 1, \dots, n$.

Pelo item **a.** do teorema anterior, temos que cada L_i é um funcional linear, ou seja, $L_i \in V^*$.

Vamos mostrar que $\{L_0, L_1, \dots, L_n\}$ é uma base do espaço vetorial V^* dual de V .

Para isso, basta mostrar que existe uma base $\{p_0, p_1, \dots, p_n\}$ de V tal que $\{L_0, L_1, \dots, L_n\}$ seja a sua base dual, ou seja, $L_i(p_j) = \delta_{ij} = P_j(t_i)$, para $i, j = 0, 1, \dots, n$.

Consideremos os polinômios

$$p_j = \prod_{\substack{i=1 \\ i \neq j}}^n \frac{x - t_i}{t_j - t_i}, \quad j = 0, 1, \dots, n.$$

Então, já que cada $p_j \in V$ é um polinômio de grau n , e $p_j(t_k) = \delta_{jk}$, pelo item **b.** do teorema anterior.

Afirmção: Os polinômios $p_0, p_1, \dots, p_n$ são LI.

De fato, se $\sum_{j=0}^n c_j p_j = \mathbf{0}$, temos que $\sum_{j=0}^n c_j p_j(t_i) = 0$, para todo $i = 0, 1, \dots, n$.

Como $p_j(t_i) = \delta_{ji}$, temos que $c_i = 0$, para todo $i = 0, 1, \dots, n$.

Sendo $\dim V = n + 1$, temos que $\{p_0, p_1, \dots, p_n\}$ é uma base de V dual da base $\{L_0, L_1, \dots, L_n\}$.

Então, se $f \in V$, temos a chamada **fórmula de interpolação de Lagrange**:

$$f = \sum_{i=0}^n f(t_i) p_i = \sum_{i=0}^n f(t_i) \prod_{\substack{k=1 \\ k \neq i}}^n \frac{x - t_k}{t_i - t_k}$$

Tomando $f = x^j$, obtemos,

$$x^j = \sum_{i=0}^n t_i^j p_i.$$

Então, a matriz de mudança de base da base $\{\mathbf{1}, x, x^2, \dots, x^n\}$ para a base $\{p_0, p_1, \dots, p_n\}$, é

$$\begin{pmatrix} 1 & t_0 & t_0^2 & \dots & t_0^n \\ 1 & t_1 & t_1^2 & \dots & t_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_n & t_n^2 & \dots & t_n^n \end{pmatrix}$$

conhecida como a **matriz de Vandermonde**. Essa matriz é, portanto, invertível.

Seja, agora f um polinômio sobre $\mathbb{K}$, e seja $\tilde{f} : \mathbb{K} \rightarrow \mathbb{K}$ a função polinomial que leva t em $f(t)$.

Observação 2.1

Pode acontecer que $\tilde{f} = \tilde{g}$ para dois polinômios f e g tais que $f \neq g$.

De fato, seja $\mathbb{K} = \mathbb{Z}_2$ e sejam $f = x^2 + 1$ e $g = x^3 + 1$ polinômios em $\mathbb{Z}_2[x]$.

Como $\tilde{f}(0) = \tilde{g}(0) = 1$ e $\tilde{f}(1) = \tilde{g}(1) = 0$, temos que $\tilde{f} = \tilde{g}$, mas $f \neq g$.

Veremos que isso não ocorre quando $\mathbb{K}$ é um corpo com um número infinito de elementos.

Para descrever de maneira precisa a relação entre os polinômios e as funções polinomiais, devemos definir o produto de duas funções polinomiais.

Definição 2.1

Sejam f e g polinômios sobre $\mathbb{K}$. O *produto das funções polinomiais* $\tilde{f}$ e $\tilde{g}$, associadas a f e g , respectivamente, é a função polinomial $\tilde{f} \cdot \tilde{g} : \mathbb{K} \rightarrow \mathbb{K}$ dada por:

$$(\tilde{f} \cdot \tilde{g})(t) = \tilde{f}(t) \cdot \tilde{g}(t), \quad \forall t \in \mathbb{K}.$$

Pela parte **b.** do teorema **1.2**, temos que $(f \cdot g)(t) = f(t)g(t)$. Logo,

$$(\widetilde{f \cdot g})(t) = (f \cdot g)(t) = f(t)g(t) = \tilde{f}(t) \cdot \tilde{g}(t), \quad \forall t \in \mathbb{K}.$$

Assim,

$$\widetilde{f \cdot g} = \tilde{f} \cdot \tilde{g}.$$

Com a multiplicação definida acima, o espaço vetorial das funções polinomiais sobre $\mathbb{K}$ é uma álgebra linear comutativa com elemento unidade.

De fato,

$$\tilde{f} \cdot (\tilde{g} \cdot \tilde{h}) = \tilde{f}(\widetilde{g \cdot h}) = \widetilde{f \cdot (g \cdot h)} = \widetilde{(f \cdot g) \cdot h} = (\widetilde{f \cdot g}) \cdot \tilde{h} = (\tilde{f} \cdot \tilde{g}) \cdot \tilde{h}.$$

Definição 2.2

Duas álgebras lineares $\mathcal{A}$ e $\tilde{\mathcal{A}}$ sobre o corpo $\mathbb{K}$ são ditas *isomorfas* se existe uma aplicação $\alpha \mapsto \tilde{\alpha}$ de $\mathcal{A}$ em $\tilde{\mathcal{A}}$, tal que:

$$\begin{aligned}\widetilde{(c\alpha + \beta)} &= c\widetilde{\alpha} + \widetilde{\beta}, \quad \forall \alpha, \beta \in \mathcal{A}, \quad \forall c \in \mathbb{K}. \\ \widetilde{(\alpha\beta)} &= \widetilde{\alpha}\widetilde{\beta},\end{aligned}$$

Exemplo 2.1

Seja V um espaço vetorial de dimensão n sobre o corpo $\mathbb{K}$ e seja $\mathcal{B}$ uma base ordenada de V .

Sejam $T, L \in \mathcal{L}(V, V)$ operadores lineares sobre V e $c \in \mathbb{K}$.

Como $[cT + L]_{\mathcal{B}} = c[T]_{\mathcal{B}} + [L]_{\mathcal{B}}$ e $[T \circ L]_{\mathcal{B}} = [T]_{\mathcal{B}}[L]_{\mathcal{B}}$, temos que a aplicação

$$\begin{aligned}\mathcal{L}(V, V) &\longrightarrow \mathbb{K}^{n \times n} \\ T &\longmapsto [T]_{\mathcal{B}}\end{aligned}$$

é um isomorfismo da álgebra dos operadores lineares sobre V na álgebra das matrizes $n \times n$ com entradas em $\mathbb{K}$.

Seja $U \in \mathcal{L}(V, V)$ e seja $f = \sum_{i=0}^n c_i x^i$ um polinômio de grau n com coeficientes em $\mathbb{K}$. Então,

$$f(U) = \sum_{i=0}^n c_i U^i,$$

e como $T \longmapsto [T]_{\mathcal{B}}$ é uma aplicação linear, temos

$$[f(U)]_{\mathcal{B}} = \sum_{i=0}^n c_i [U^i]_{\mathcal{B}}.$$

Além disso, como $[T \circ L]_{\mathcal{B}} = [T]_{\mathcal{B}}[L]_{\mathcal{B}}$ para quaisquer $L, T \in \mathcal{L}(V, V)$, temos que

$$[f(U)]_{\mathcal{B}} = \sum_{i=0}^n c_i [U]_{\mathcal{B}}^i,$$

ou seja,

$$[f(U)]_{\mathcal{B}} = f([U]_{\mathcal{B}}).$$

□

Teorema 2.1

Seja $\mathbb{K}$ um corpo com um número infinito de elementos.

Então, a aplicação $f \longmapsto \tilde{f}$ é um isomorfismo da álgebra dos polinômios sobre $\mathbb{K}$ na álgebra das funções polinomiais sobre $\mathbb{K}$.

Prova.

Como já sabemos que a aplicação é sobrejetora, e que

$$\begin{aligned}\widetilde{(cf + g)} &= c\tilde{f} + \tilde{g} \\ \widetilde{(f \cdot g)} &= \tilde{f} \cdot \tilde{g},\end{aligned}$$

basta mostrar que a aplicação é injetora. Para tanto é suficiente, pela linearidade, demonstrar que $\tilde{f} = \mathcal{O}$ implica $f = \mathbf{0}$.

Seja f um polinômio de grau $\leq n$ tal que $\tilde{f} = \mathcal{O}$, e sejam $t_0, t_1, \dots, t_n$ elementos distintos de $\mathbb{K}$.

Pela fórmula de interpolação de Lagrange, temos que

$$f = \sum_{i=0}^n f(t_i) p_i.$$

Como $\tilde{f} = \mathcal{O}$, temos $f(t_i) = 0$, para todo $i = 0, 1, \dots, n$. Logo, $f = \mathbf{0}$. ■

3. Ideais de Polinômios

Lema 3.1

Sejam f e d polinômios não-nulos sobre o corpo $\mathbb{K}$ tais que

$$\text{grau}(d) \leq \text{grau}(f).$$

Então existe um polinômio g sobre $\mathbb{K}$, tal que

$$f - gd = 0 \quad \text{ou} \quad \text{grau}(f - gd) < \text{grau}(f).$$

Prova.

Suponhamos que

$$f = \sum_{i=0}^m a_i x^i = a_m x^m + \sum_{i=0}^{m-1} a_i x^i, \quad a_m \neq 0$$

e

$$d = \sum_{j=0}^n b_j x^j = b_n x^n + \sum_{j=0}^{n-1} b_j x^j, \quad b_n \neq 0.$$

Tome $g = \frac{a_m}{b_n} x^{m-n}$, que faz sentido, pois $m \geq n$. Então,

$$f - gd = f - \frac{a_m}{b_n} x^{m-n} d = 0 \quad \text{ou} \quad \text{grau}(f - gd) < \text{grau}(f).$$

■

Teorema 3.1

Sejam f e d polinômios sobre um corpo $\mathbb{K}$, sendo d um polinômio não-nulo. Então existem polinômios q e r tais que:

$$(1) \quad f = qd + r,$$

$$(2) \quad r = 0 \text{ ou } \text{grau}(r) < \text{grau}(d).$$

Os polinômios q e r que satisfazem (1) e (2) são únicos.

Prova.

Existência: Se $f = 0$ ou $\text{grau}(f) < \text{grau}(d)$, basta tomar $q = 0$ e $r = f$.

Suponhamos, então, que $f \neq 0$ e $\text{grau}(f) \geq \text{grau}(d)$.

Então, pelo lema anterior, existe um polinômio $g \in \mathbb{K}[x]$ tal que $f - dg = 0$ ou $\text{grau}(f - dg) < \text{grau}(f)$.

Se $f - dg \neq 0$ e $\text{grau}(f - dg) \geq \text{grau}(d)$, existe $h \in \mathbb{K}[x]$ tal que $f - dg - dh = 0$ ou $\text{grau}(f - d(g + h)) < \text{grau}(f - dg) < \text{grau}(f)$.

Após um número finito de passos, no máximo $\text{grau}(f) - \text{grau}(d) + 1$ no caso em que d não é um polinômio constante, chegamos a um polinômio $q \in \mathbb{K}[x]$ e a um polinômio $r \in \mathbb{K}[x]$, tais que

$$f - dq = r, \text{ com } r = 0 \text{ ou } \text{grau}(r) < \text{grau}(d).$$

No caso em que d é um polinômio constante, basta tomar

$$q = \frac{1}{d} \cdot f \text{ e } r = 0.$$

Unicidade: Suponhamos que existem outros polinômios $q_1, r_1 \in \mathbb{K}[x]$ tais que $f = q_1d + r_1$, onde $r_1 = 0$ ou $\text{grau}(r_1) < \text{grau}(d)$.

Como $f = qd + r = q_1d + r_1$, temos que $d(q - q_1) = r_1 - r$.

Se $q - q_1 \neq 0$ então, $d(q - q_1) \neq 0$ e

$$\text{grau}(d(q - q_1)) = \text{grau}(d) + \text{grau}(q - q_1) \geq \text{grau}(d).$$

Mas, como $\text{grau}(d(q - q_1)) = \text{grau}(r_1 - r) < \text{grau}(d)$, chegamos a uma contradição ($\text{grau}(d) < \text{grau}(d)$).

Logo, $q - q_1 = 0$ e $r_1 - r = 0$, ou seja, $q_1 = q$ e $r_1 = r$. ■

Definição 3.1

Seja d um polinômio não-nulo sobre $\mathbb{K}$ e $f \in \mathbb{K}[x]$. Se existe $q \in \mathbb{K}[x]$ tal que $f = qd$, dizemos que d *divide* f , que f é *divisível* por d , ou que f

é **múltiplo** de d , e denominamos q o **quociente** de f por d . Escrevemos, também, $q = \frac{f}{d}$.

Proposição 3.1

Seja $f \in \mathbb{K}[x]$ e seja $c \in \mathbb{K}$. Então, f é divisível por $x - c$ se, e somente se, $f(c) = 0$.

Prova.

Pelo teorema anterior, existe $q \in \mathbb{K}[x]$ e um polinômio constante r , tais que

$$f = (x - c)q + r.$$

Logo, $f(c) = 0$ se, e somente se, $r = 0$. ■

Definição 3.2

Um escalar $c \in \mathbb{K}$ é uma **raiz** ou um **zero** de um polinômio $f \in \mathbb{K}[x]$, se $f(c) = 0$.

Corolário 3.1

Um polinômio f não-nulo de grau n sobre um corpo $\mathbb{K}$ tem no máximo n raízes em $\mathbb{K}$.

Prova.

O resultado é óbvio quando f tem grau 0 ou 1.

Suponhamos que $\text{grau}(f) > 1$.

Vamos proceder por indução sobre o grau n de f .

Suponhamos que o resultado vale para polinômios de grau $n - 1$.

Seja $c \in \mathbb{K}$ uma raiz de f . Então, existe $q \in \mathbb{K}[x]$ de grau $n - 1$ tal que $f = (x - c)q$. Como, por hipótese de indução, q possui no máximo $n - 1$ raízes, temos que f possui no máximo n raízes. ■

A **derivada** do polinômio $f = c_0 + c_1x + c_2x^2 + \dots + c_nx^n$ é o polinômio

$$Df = f' = c_1 + 2c_2x + \dots + nc_nx^{n-1}.$$

A derivação é linear, isto é, D é um operador linear sobre $\mathbb{K}[x]$. Existem também as derivadas formais de ordem dois $D^2f = f''$, de ordem três $D^3f = f'''$ etc.

Teorema 3.2

Se f é um polinômio sobre $\mathbb{K}$ de grau $\leq n$, então

$$f = \sum_{k=0}^n \frac{D^k f(c)}{k!} (x - c)^k,$$

onde c é um escalar em $\mathbb{K}$.

Prova.

Pela fórmula do binômio de Newton, temos que

$$x^m = (c + (x - c))^m = \sum_{k=0}^m \binom{m}{k} c^{m-k} (x - c)^k,$$

onde $\binom{m}{k} = \frac{m!}{k!(m-k)!} = \frac{m(m-1)(m-2)\dots(m-k+1)}{1 \cdot 2 \cdot 3 \cdot \dots \cdot k}$. Então,

$$x^m = \sum_{k=0}^m \frac{D^k x^m(c)}{k!} (x - c)^k,$$

que é a fórmula de Taylor para $f = x^m$.

Se $f = \sum_{m=0}^n a_m x^m$, temos que $D^k f(c) = \sum_{m=0}^n a_m D^k (x^m)(c)$, já que a deri-

vada de ordem k é linear.

Logo,

$$\begin{aligned} \sum_{k=0}^n \frac{D^k f(c)}{k!} (x - c)^k &= \sum_{k=0}^n \sum_{m=0}^n a_m \frac{D^k (x^m)(c)}{k!} (x - c)^k \\ &= \sum_{m=0}^n a_m \sum_{k=0}^n \frac{D^k (x^m)(c)}{k!} (x - c)^k \\ &= \sum_{m=0}^n a_m x^m = f \end{aligned}$$

**Observação 3.1**

- Os polinômios $\mathbf{1}, (x - c), \dots, (x - c)^n$ são LI.

De fato: Procedemos com a demonstração usando indução sobre n .

1. $\{1\}$ é LI, já que $\mathbf{1} \neq \mathbf{0}$.

2. Suponhamos que $\{\mathbf{1}, (x - c), \dots, (x - c)^{n-1}\}$ é LI e que

$$a_n(x-c)^n + a_{n-1}(x-c)^{n-1} + \dots + a_1(x-c) + a_0 = \mathbf{0}.$$

Logo,

$$a_n(x-c)^n = -(a_{n-1}(x-c)^{n-1} + \dots + a_1(x-c) + a_0).$$

Então, $a_n = 0$, pois se $a_n \neq 0$, teríamos que $a_n(x-c)^n$ seria um polinômio de grau n , enquanto $a_{n-1}(x-c)^{n-1} + \dots + a_0$ é um polinômio de grau $\leq n-1$.

Como $a_n = 0$, temos

$$a_{n-1}(x-c)^{n-1} + \dots + a_1(x-c) + a_0 = \mathbf{0},$$

Logo, $a_{n-1} = \dots = a_1 = a_0 = 0$.

• A fórmula de Taylor nos fornece, então, o único modo de escrever um polinômio f como combinação linear dos polinômios $(x-c)^k$, $0 \leq k \leq n$.

Definição 3.3

Se c for uma raiz do polinômio f , a **multiplicidade de c como raiz de f** é o maior inteiro positivo r tal que $(x-c)^r$ divide f .

Teorema 3.3

Seja f um polinômio sobre $\mathbb{K}$ de grau n . Então o escalar $c \in \mathbb{K}$ é uma raiz de f de multiplicidade r se, e só se,

$$\begin{aligned} (D^k f)(c) &= 0, & 0 \leq k \leq r-1; \\ (D^r f)(c) &\neq 0. \end{aligned}$$

Prova.

Suponhamos que f tem multiplicidade r . Então existe $g \in \mathbb{K}[x]$ tal que $f = (x-c)^r g$, com $g(c) \neq 0$, pois se $g(c) = 0$, f seria divisível por $(x-c)^{r+1}$.

Pela fórmula de Taylor aplicada a g , temos que

$$g = \sum_{k=0}^{n-r} \frac{D^k g(c)}{k!} (x-c)^k.$$

Logo,

$$f = \sum_{k=0}^{n-r} \frac{D^k g(c)}{k!} (x-c)^{r+k}.$$

Como existe apenas uma maneira de escrever f como combinação linear dos polinômios $(x-c)^k$, $0 \leq k \leq n$, temos que

$$\frac{D^m f(c)}{m!} = \begin{cases} 0, & \text{se } 0 \leq m \leq r-1 \\ \frac{D^{m-r} g(c)}{(m-r)!}, & \text{se } r \leq m \leq n \end{cases}$$

Logo, $D^m f(c) = 0$ se $0 \leq m \leq r-1$ e $D^r f(c) = r! g(c) \neq 0$.

Suponhamos agora que $D^k f(c) = 0$, $0 \leq k \leq r-1$ e $D^r f(c) \neq 0$. Então, pela fórmula de Taylor,

$$f = (x-c)^r \sum_{k=r}^n \frac{D^k f(c)}{k!} (x-c)^{k-r},$$

ou seja, existe $g = \sum_{k=r}^n \frac{D^k f(c)}{k!} (x-c)^{k-r} \in \mathbb{K}[x]$, tal que $f = (x-c)^r g$.

Observe que $g(c) = \frac{D^r f(c)}{r!} \neq 0$.

Suponhamos que r não é o maior inteiro positivo tal que $(x-c)^r$ divide f .

Então, existe $h \in \mathbb{K}[x]$, tal que $f = (x-c)^{r+1} h$. Sendo $f = (x-c)^r g = (x-c)^{r+1} h$, temos que $(x-c)^r (g - (x-c)h) = 0$.

Mas, como $x-c \neq 0$, temos que $g - (x-c)h = 0$, ou seja, $g = (x-c)h$.

Assim, $g(c) = 0$, o que é uma contradição. ■

Definição 3.4

Seja $\mathbb{K}$ um corpo. Um *ideal* em $\mathbb{K}[x]$ é um subespaço $\mathbb{I}$ de $\mathbb{K}[x]$ tal que $fg \in \mathbb{I}$ para todo $f \in \mathbb{K}[x]$ e todo $g \in \mathbb{I}$.

Exemplo 3.1

Seja $d \in \mathbb{K}[x]$. Então $\mathbb{I} = d \cdot \mathbb{K}[x]$, o conjunto de todos os múltiplos de d é um ideal.

De fato, $\mathbb{I}$ é não vazio, pois contém $d = d \cdot 1$.

- Se $f, g \in \mathbb{K}[x]$ e $c \in \mathbb{K}$, então $c(d \cdot f) + d \cdot g = d \cdot (cf + g) \in \mathbb{I}$.

Logo, $\mathbb{I}$ é um subespaço de $\mathbb{K}[x]$.

- Se $h \in \mathbb{K}[x]$ e $d \cdot f \in \mathbb{I}$, então $(d \cdot f) \cdot h = d \cdot (f \cdot h) \in \mathbb{I}$. Logo $\mathbb{I}$ é um ideal.

O ideal $\mathbb{I} = d \cdot \mathbb{K}[x]$ é chamado o *ideal principal gerado por d* . □

Exemplo 3.2

Sejam $d_1, \dots, d_n \in \mathbb{K}[x]$. Então a soma $\mathbb{I}$ dos subespaços $d_i \cdot \mathbb{K}[x]$,

$i = 1, \dots, n$, é um subespaço e, também, um ideal.

De fato, se $p \in \mathbb{I}$, existem $f_1, \dots, f_n \in \mathbb{K}[x]$ tais que $p = d_1 \cdot f_1 + \dots + d_n \cdot f_n$.

Se $g \in \mathbb{K}[x]$, temos que

$$p \cdot g = (d_1 \cdot f_1 + \dots + d_n \cdot f_n) \cdot g = d_1 \cdot (f_1 \cdot g) + \dots + d_n \cdot (f_n \cdot g),$$

ou seja, $p \cdot g \in \mathbb{I}$. Dizemos que $\mathbb{I} = d_1 \cdot \mathbb{K}[x] + \dots + d_n \cdot \mathbb{K}[x]$ é o *ideal gerado pelos polinômios* $d_1, \dots, d_n$. $\square$

Exemplo 3.3

Seja $\mathbb{K}$ um subcorpo dos números complexos e consideremos o ideal

$$\mathbb{I} = (x + 2)\mathbb{K}[x] + (x^2 + 8x + 16)\mathbb{K}[x],$$

Então, $\mathbb{I} = \mathbb{K}[x]$. De fato, como

$$(x^2 + 8x + 16) - x(x + 2) = 6x + 16 \in \mathbb{I},$$

temos que $(6x + 16) - 6 \cdot (x + 2) = 4 \in \mathbb{I}$. Assim, o polinômio constante **1** pertence a $\mathbb{I}$, bem como todos os seus múltiplos. $\square$

Teorema 3.4

Seja $\mathbb{I}$ um ideal não-nulo de $\mathbb{K}[x]$. Então existe um único polinômio unitário $d \in \mathbb{K}[x]$ tal que $\mathbb{I} = d \cdot \mathbb{K}[x]$.

Prova.

Existência. Como $\mathbb{I}$ é não-vazio, existe em $\mathbb{I}$ um polinômio d de grau mínimo e unitário. Vamos mostrar que $\mathbb{I} = d \cdot \mathbb{K}[x]$.

Seja $f \in \mathbb{I}$. Então existem $q, r \in \mathbb{K}[x]$ tais que $f = q \cdot d + r$, onde $r = \mathbf{0}$ ou $\text{grau}(r) < \text{grau}(d)$.

Como $d \in \mathbb{I}$, temos que $d \cdot q \in \mathbb{I}$ e, portanto, $r = f - q \cdot d \in \mathbb{I}$. Sendo d um polinômio em $\mathbb{I}$ de grau mínimo, não podemos ter $\text{grau}(r) < \text{grau}(d)$. Logo, $r = \mathbf{0}$ e $f = d \cdot q$.

Assim, $\mathbb{I} = d \cdot \mathbb{K}[x]$.

Unicidade. Sejam $d, d_1 \in \mathbb{K}[x]$ polinômios unitários tais que

$$\mathbb{I} = d \cdot \mathbb{K}[x] = d_1 \mathbb{K}[x].$$

Então, existem $q, q_1 \in \mathbb{K}[x]$ tais que $d = d_1 \cdot q_1$ e $d_1 = d \cdot q$. Logo, $d = d \cdot q \cdot q_1$ e, portanto,

$$\text{grau}(d) = \text{grau}(d) + \text{grau}(q) + \text{grau}(q_1).$$

Assim, $\text{grau}(q) = \text{grau}(q_1) = 0$. Ou seja, q e q_1 são polinômios constantes. Como d e d_1 são unitários, $q = q_1 = 1$, isto é, $d = d_1$. ■

Corolário 3.2

Sejam $p_1, \dots, p_n \in \mathbb{K}[x]$ polinômios não todos nulos. Então, existe um único polinômio unitário d tal que:

(a) d pertence ao ideal gerado por $p_1, \dots, p_n$;

(b) d divide p_i , $i = 1, \dots, n$.

Além disso, todo polinômio d que satisfaz **(a)** e **(b)**, satisfaz também:

(c) d é divisível por todo polinômio que divide cada um dos polinômios $p_1, \dots, p_n$.

Prova.

Seja $d \in \mathbb{K}[x]$ o polinômio unitário tal que $d \cdot \mathbb{K}[x] = p_1 \cdot \mathbb{K}[x] + \dots + p_n \cdot \mathbb{K}[x]$.

Então, d pertence ao ideal gerado por $p_1, \dots, p_n$ e cada p_i , $i = 1, \dots, n$, pertence ao ideal $d \cdot \mathbb{K}[x]$.

Logo, d divide cada um dos polinômios p_i .

Suponhamos agora que f seja um polinômio que divide cada um dos polinômios $p_1, \dots, p_n$. Então, existem polinômios $h_1, \dots, h_n$, tais que $p_i = h_i \cdot f$, $i = 1, \dots, n$.

Além disso, como $d \in p_1 \cdot \mathbb{K}[x] + \dots + p_n \cdot \mathbb{K}[x]$, existem polinômios $q_1, \dots, q_n$, tais que

$$d = p_1 \cdot q_1 + \dots + p_n \cdot q_n.$$

Logo, $d = f \cdot (h_1 \cdot q_1 + \dots + h_n \cdot q_n)$, ou seja, d é divisível por f .

Seja d' um polinômio unitário que satisfaz **(a)** e **(b)**.

Por **(a)**,

$$d' \in d \cdot \mathbb{K}[x] = p_1 \cdot \mathbb{K}[x] + \dots + p_n \cdot \mathbb{K}[x],$$

ou seja, existe $q \in \mathbb{K}[x]$, tal que $d' = d \cdot q$.

Como d' satisfaz **(b)** e d satisfaz **(c)**, d é divisível por d' , ou seja, existe $h \in \mathbb{K}[x]$ tal que $d = d' \cdot h$. Logo, $d = d \cdot q \cdot h$ e, portanto, $q = h = 1$.

Assim, $d' = d$. ■

Definição 3.5

Sejam $p_1, \dots, p_n$ polinômios em $\mathbb{K}[x]$ não todos nulos. Então, o único polinômio unitário $d \in \mathbb{K}[x]$ tal que

$$d \cdot \mathbb{K}[x] = p_1 \cdot \mathbb{K}[x] + \dots + p_n \cdot \mathbb{K}[x],$$

é chamado o **máximo divisor comum (m.d.c.)** de $p_1, \dots, p_n$, e é indicado por $(p_1, \dots, p_n)$.

Se $(p_1, \dots, p_n) = \mathbf{1}$, ou seja, $p_1 \cdot \mathbb{K}[x] + \dots + p_n \cdot \mathbb{K}[x] = \mathbf{1} \cdot \mathbb{K}[x] = \mathbb{K}[x]$, dizemos que $p_1, \dots, p_n$ são **primos entre si** ou **relativamente primos**.

Exemplo 3.4

No exemplo anterior, vimos que

$$\mathbb{K}[x] = (x+2) \cdot \mathbb{K}[x] + (x^2+8x+16) \cdot \mathbb{K}[x].$$

Logo, $(x+2, x^2+8x+16) = \mathbf{1}$. $\square$

Exemplo 3.5

Verifiquemos que $((x-2)^2 \cdot (x+i), (x-2) \cdot (x^2+1)) = (x-2) \cdot (x+i)$, onde $\mathbb{K}$ é o corpo dos números complexos.

De fato, o ideal

$$((x-2)^2 \cdot (x+i)) \cdot \mathbb{K}[x] + ((x-2) \cdot (x^2+1)) \cdot \mathbb{K}[x]$$

contém

$$(x-2)^2 \cdot (x+i) - (x-2) \cdot (x^2+1) = (x+2) \cdot (x+i) \cdot (i-2).$$

Logo, contém o polinômio $(x-2) \cdot (x+i)$, que é unitário e divide

$$(x-2)^2 \cdot (x+i) \quad \text{e} \quad (x-2) \cdot (x^2+1).$$

Como o polinômio $(x-2) \cdot (x+i)$ satisfaz **(a)** e **(b)** do corolário anterior, temos que $(x-2)(x+i) = ((x-2)^2(x+i), (x-2)(x^2+1))$. $\square$

4. Decomposição de um polinômio em fatores primos

Definição 4.1

Dizemos que um polinômio $f \in \mathbb{K}[x]$ é **reduzível sobre $\mathbb{K}[x]$** se existem polinômios $g, h \in \mathbb{K}[x]$, de grau ≥ 1 , tais que $f = g \cdot h$. Caso contrário, dizemos que o polinômio f é **irreduzível**. Se $f \in \mathbb{K}[x]$ é irreduzível e não-constante, f é dito um **polinômio primo sobre $\mathbb{K}$** .

Exemplo 4.1

O polinômio $x^2 + 1$ é reduzível sobre o corpo $\mathbb{C}$ dos números complexos, pois

$$x^2 + 1 = (x + i)(x - i).$$

Por outro lado, $x^2 + 1$ é irreduzível sobre o corpo $\mathbb{R}$ dos números reais, pois se

$$x^2 + 1 = (ax + b) \cdot (a'x + b'),$$

com $a, b, a', b' \in \mathbb{R}$, então

$$aa' = 1, ab' + a'b = 0 \text{ e } bb' = 1.$$

Logo, $0 = a \frac{1}{b} + \frac{1}{a} b = \frac{a^2 + b^2}{ab}$, ou seja, $a^2 + b^2 = 0$. Como $a, b \in \mathbb{R}$, temos que $a = b = 0$, o que é uma contradição, já que $aa' = 1$ e $bb' = 1$.

□

Teorema 4.1

Sejam f, g e p polinômios sobre o corpo $\mathbb{K}$. Se p é um polinômio primo que divide o produto fg , então p divide f ou p divide g .

Prova.

Podemos supor, sem perda de generalidade, que p é unitário. Como p é primo, os únicos divisores unitários de p são **1** e p .

Seja d o m.d.c de p e f . Então, $d = \mathbf{1}$ ou $d = p$.

Se $d = p$, então p divide f .

Se $d = \mathbf{1}$, isto é, p e f são primos entre si, existem polinômios $h, \ell \in \mathbb{K}[x]$, tais que $1 = h \cdot f + \ell \cdot p$.

Multiplicando essa igualdade por g , obtemos que $g = h \cdot f \cdot g + \ell \cdot p \cdot g$. Como p divide $f \cdot g$, temos que p divide $h \cdot f \cdot g$. Além disso, p divide $\ell \cdot p \cdot g$. Logo, p divide g . ■

Corolário 4.1

Se p é um polinômio primo que divide o produto $f_1 \cdot \dots \cdot f_n$, então p divide um dos polinômios $f_1, \dots, f_n$.

Prova.

Faremos a demonstração por indução sobre n .

Pelo teorema anterior, o resultado é válido para $n = 2$.

Suponhamos que o resultado seja válido para $n = k$ e que p divide $f_1 \cdot \dots \cdot f_{k+1}$.

Como p divide $f_1 \cdot \dots \cdot f_k \cdot f_{k+1}$ temos, pelo teorema anterior, que p divide f_{k+1} ou p divide $f_1 \cdot \dots \cdot f_k$.

Isto é, p divide f_{k+1} ou, pela hipótese de indução, p divide f_j , para algum $j = 1, \dots, k$.

Assim, em qualquer caso, p divide algum f_j , $j = 1, \dots, k + 1$. ■

Teorema 4.2

Todo polinômio $f \in \mathbb{K}[x]$ unitário e não-constante pode ser decomposto como um produto de polinômios primos em $\mathbb{K}[x]$ de uma única maneira, a menos da ordem dos fatores.

Prova.

Provaremos o resultado por indução sobre o grau de f .

Se $\text{grau}(f) = 1$, então f é irredutível e, portanto, primo.

Suponhamos que o teorema seja válido para polinômios de grau $< n$ e que $\text{grau } f = n + 1$.

Se f é primo, então nada temos a provar.

Se f não é primo, isto é, f é redutível, existem polinômios $g, h \in \mathbb{K}[x]$ não-constantes e unitários, tais que $f = g \cdot h$.

Como $\text{grau}(g) < n$ e $\text{grau}(h) < n$, temos, por hipótese de indução, que g e h podem ser decompostos como produtos de polinômios primos unitários

em $\mathbb{K}[x]$. Logo, f também pode ser decomposto como um produto de polinômios primos.

Suponhamos agora, que

$$f = p_1 \cdot \dots \cdot p_m = q_1 \cdot \dots \cdot q_n,$$

onde $p_1, \dots, p_m, q_1, \dots, q_n \in \mathbb{K}[x]$ são polinômios primos unitários.

Se $m = 1$ ou $n = 1$, então $m = n = 1$.

Suponhamos então que $m > 1$ e $n > 1$.

Como p_1 divide $q_1 \cdot \dots \cdot q_n$, temos que p_1 divide algum q_j , $j = 1, \dots, n$.

Sendo p_1 e q_j primos unitários, temos que $p_1 = q_j$.

Reordenando os polinômios q_i , caso seja necessário, podemos supor que $p_1 = q_1$.

Logo,

$$p_1 \cdot p_2 \cdot \dots \cdot p_m = p_1 \cdot q_2 \cdot \dots \cdot q_n,$$

e, portanto, $p_2 \cdot \dots \cdot p_m = q_2 \cdot \dots \cdot q_n$.

O teorema se segue agora pela hipótese de indução, pois esse polinômio tem grau menor que $\text{grau}(f)$. ■

Observação 4.1

Na decomposição de um polinômio unitário não-constante em produto de fatores primos unitários, alguns fatores podem repetir-se. Assim, se $p_1, \dots, p_r$ são os fatores primos unitários distintos que ocorrem na decomposição de um polinômio unitário não-constante f , então

$$f = p_1^{n_1} \cdot \dots \cdot p_r^{n_r},$$

sendo o expoente n_i o número de vezes que o fator primo p_i ocorre nessa decomposição. Essa decomposição de f em produto de potências de fatores primos é única, a menos da ordem dos fatores, e se denomina a *decomposição primária de f* .

Assim, todo divisor unitário de f é da forma

$$p_1^{\gamma_1} \cdot \dots \cdot p_r^{\gamma_r},$$

onde $0 \leq \gamma_i \leq n_i$, $\forall i = 1, \dots, r$.

Em particular, o m.d.c de uma coleção finita de polinômios não-constantes e unitários $f_1, \dots, f_n$ é

$$p_1^{s_1} \cdot \dots \cdot p_k^{s_k},$$

onde $p_1, \dots, p_k$ são os polinômios primos que aparecem nas decomposições de todos os polinômios $f_1, \dots, f_n$ e, para cada $i = 1, \dots, k$, o expoente s_i é o menor dos expoentes com que o polinômio p_i aparece nas decomposições dos polinômios $f_1, \dots, f_n$.

Se os polinômios $f_1, \dots, f_n$ não possuem fatores primos em comum, eles são primos entre si.

Exemplo 4.2

Sejam $a, b, c \in \mathbb{K}$ escalares distintos. Então, $x - a$, $x - b$ e $x - c$ são polinômios primos unitários distintos em $\mathbb{K}[x]$.

Logo,

$$((x - b)^n \cdot (x - c)^s, (x - a)^m \cdot (x - c)^{s+1}) = (x - c)^s,$$

e

$$((x - b)^n \cdot (x - c)^s, (x - a)^m \cdot (x - c)^s, (x - b)^n(x - a)^m) = 1.$$

□

Teorema 4.3

Seja f um polinômio unitário não-constante sobre o corpo $\mathbb{K}$ e seja

$$f = p_1^{n_1} \cdot \dots \cdot p_k^{n_k},$$

a decomposição de f em (produto de potências de) fatores primos.

Para cada j , $1 \leq j \leq k$, seja

$$f_j = \frac{f}{p_j^{n_j}} = \prod_{i \neq j} p_i^{n_i},$$

Então, $f_1, \dots, f_k$ são primos entre si.

Teorema 4.4

Seja f um polinômio sobre o corpo $\mathbb{K}$ com derivada f' . Então, f é um produto de polinômios primos distintos se, e somente se, f e f' são primos entre si.

Prova.

($\Leftarrow$) Suponhamos que na decomposição de f em fatores primos algum polinômio primo p esteja repetido, ou seja, $f = p^2 \cdot h$, para algum $h \in \mathbb{K}[x]$.

Então, $f' = 2p \cdot p' \cdot h + p^2 \cdot h'$. Logo, p também divide f' , não sendo, portanto, f e f' primos entre si.

($\implies$) Suponhamos que $f = p_1 \cdot \dots \cdot p_k$, onde $p_1, \dots, p_k$ são polinômios primos unitários e distintos.

Seja

$$f_j = \frac{f}{p_j} = \prod_{i \neq j} p_i.$$

Então,

$$f' = p'_1 \cdot f_1 + \dots + p'_k \cdot f_k.$$

Seja p um polinômio primo que divide f e f' . Então, $p = p_i$ para algum $i = 1, \dots, k$. Como p_i divide f_j , para $j \neq i$, e p_i divide f' , temos que $p = p_i$ divide $p'_i \cdot f_i$.

Então, p divide p'_i ou p divide f_i .

Como $p = p_i$ não aparece na decomposição $\prod_{j \neq i} p_j$, temos que p_i não divide f_i .

Também p_i não divide p'_i , pois $\text{grau}(p_i) > \text{grau}(p'_i)$.

Com isso verificamos que nenhum polinômio primo divide f e f' simultaneamente. Ou seja f e f' são primos entre si. ■

Definição 4.2

O corpo $\mathbb{K}$ é dito **algebricamente fechado** se todo polinômio primo sobre $\mathbb{K}$ tem grau 1, isto é, se todo polinômio primo unitário sobre $\mathbb{K}$ é da forma $x - c$.

Maneiras equivalentes para definir um corpo algebricamente fechado:

- Um corpo $\mathbb{K}$ é algebricamente fechado se todo polinômio unitário não-constante $f \in \mathbb{K}[x]$ se expressa na forma:

$$f = (x - c_1)^{\nu_1} \cdot \dots \cdot (x - c_k)^{\nu_k},$$

onde $c_1, \dots, c_k \in \mathbb{K}$ são escalares distintos e $\nu_1, \dots, \nu_k$ são inteiros positivos.

- Um corpo $\mathbb{K}$ é algebricamente fechado se todo polinômio não-constante $f \in \mathbb{K}[x]$ possui uma raiz em $\mathbb{K}$. Ou seja, existe $c \in \mathbb{K}$ tal que $f(c) = 0$.

Os escalares $c_1, \dots, c_k$ são as raízes de f e os inteiros $\nu_1, \dots, \nu_k$ são as respectivas multiplicidades dessas raízes.

Exemplo 4.3

O corpo $\mathbb{C}$ dos números complexos é algebricamente fechado (*Teorema fundamental da Álgebra*). $\square$

Exemplo 4.4

Se f é um polinômio unitário com coeficientes reais e c é uma raiz complexa de f , então $\bar{c}$ é também uma raiz de f .

Portanto, o conjunto das raízes de f é da forma:

$$\{t_1, \dots, t_s, c_1, \dots, c_r, \bar{c}_1, \dots, \bar{c}_r\},$$

onde $t_1, \dots, t_s \in \mathbb{R}$ e $c_1, \dots, c_r \in \mathbb{C}$.

Logo, f se decompõe em $\mathbb{C}[x]$ sob a forma:

$$f = (x - t_1) \cdot \dots \cdot (x - t_s) \cdot (x - c_1) \cdot (x - \bar{c}_1) \cdot \dots \cdot (x - c_r) \cdot (x - \bar{c}_r),$$

ou seja, f se decompõe em $\mathbb{R}[x]$ sob a forma:

$$f = (x - t_1) \cdot \dots \cdot (x - t_s) \cdot p_1 \cdot \dots \cdot p_r,$$

onde $p_i = (x - c_i) \cdot (x - \bar{c}_i) = x^2 - (c_i + \bar{c}_i)x + |c_i|^2$ é um polinômio primo de grau 2 em $\mathbb{R}[x]$, $i = 1, \dots, r$.

Assim, todo polinômio primo em $\mathbb{R}[x]$ tem grau 1 ou grau 2. $\square$

5. Determinantes

Definição 5.1

Um *anel* é um conjunto $\mathbb{F}$, munido de duas operações:

$$\begin{aligned} e \quad (x, y) &\longrightarrow x + y && \text{(adição)} \\ (x, y) &\longrightarrow x \cdot y = xy && \text{(multiplicação),} \end{aligned}$$

que satisfazem as seguintes propriedades:

1. $\mathbb{F}$ é um grupo comutativo em relação à adição;

2. A multiplicação é associativa:

$$(xy)z = x(yz), \quad \forall x, y, z \in \mathbb{F};$$

3. A multiplicação é distributiva em relação à adição:

$$\begin{aligned} x(y + z) &= xy + xz \\ (x + y)z &= xz + yz, \end{aligned} \quad \forall x, y, z \in \mathbb{F}.$$

Lembre que um conjunto G é um grupo em relação a uma operação $\star : G \times G \longrightarrow G$ se a operação é associativa, possui elemento neutro e todo elemento de G possui inverso em relação à operação. Além disso, quando a operação é comutativa, o grupo é dito *comutativo* ou *Abeliano*.

Se $xy = yx$ para todos $x, y \in \mathbb{F}$, dizemos que o anel $\mathbb{F}$ é **comutativo**.

Se existe um elemento $\mathbf{1} \in \mathbb{F}$ tal que $\mathbf{1}x = x\mathbf{1} = x \ \forall x \in \mathbb{F}$, dizemos que $\mathbb{F}$ é um **anel com elemento unidade** (o elemento $\mathbf{1}$).

Observação 5.1

Um corpo é um anel comutativo com elemento unidade diferente de zero que satisfaz a propriedade adicional de que a cada x não-nulo corresponde um elemento x^{-1} tal que $x^{-1}x = 1$.

Exemplo 5.1

O conjunto dos números inteiros, com as operações usuais de adição e multiplicação, é um anel comutativo com elemento unidade, mas não é um corpo. $\square$

Exemplo 5.2

O conjunto dos polinômios sobre um corpo, com a adição e multiplicação que definimos para polinômios, é um anel comutativo com elemento unidade. $\square$

Exemplo 5.3

Se $\mathbb{F}$ é um anel comutativo com elemento unidade, então o conjunto das matrizes $m \times n$ com entradas em $\mathbb{F}$, que denotamos $\mathbb{F}^{m \times n}$, com as operações

$$\begin{aligned}(A + B)_{ij} &= A_{ij} + B_{ij} \\ (CD)_{ij} &= \sum_{k=1}^n C_{ik}D_{kj}.\end{aligned}$$

é um anel comutativo com elemento unidade. $\square$

Definição 5.2

Seja $\mathbb{F}$ um anel comutativo com elemento unidade. Uma função $D : \mathbb{F}^{n \times n} \rightarrow \mathbb{F}$ é **n -linear** se para cada $i = 1, \dots, n$, D é uma função linear da i -ésima linha, quando as outras $n - 1$ linhas são mantidas fixas.

Se $D : \mathbb{F}^{n \times n} \rightarrow \mathbb{F}$ é uma função e se $\alpha_1, \dots, \alpha_n$ são as linhas da matriz A , escrevemos $D(A) = D(\alpha_1, \dots, \alpha_n)$. Dizer que D é n -linear significa que, para todo $i = 1, \dots, n$,

$$D(\alpha_1, \dots, \alpha_{i-1}, c\alpha_i + \alpha'_i, \alpha_{i+1}, \dots, \alpha_n) = c D(\alpha_1, \dots, \alpha_{i-1}, \alpha_i, \alpha_{i+1}, \dots, \alpha_n) + D(\alpha_1, \dots, \alpha_{i-1}, \alpha'_i, \alpha_{i+1}, \dots, \alpha_n),$$

Exemplo 5.4

Sejam $k_1, \dots, k_n$ inteiros positivos, $1 \leq k_i \leq n$, e seja $a \in \mathbb{F}$. Para cada matriz A $n \times n$ sobre $\mathbb{F}$, definimos

$$D(A) = aA(1, k_1) \cdot \dots \cdot A(n, k_n)$$

Então, a função D é n -linear. De fato, considerando D como uma função da i -ésima linha de A , com as outras linhas fixas, podemos escrever

$$D(\alpha_i) = A(i, k_i)b,$$

onde b é um elemento fixo de $\mathbb{F}$. Seja $\alpha'_i = (A'_{i1}, \dots, A'_{in})$.

Então,

$$D(c\alpha_i + \alpha'_i) = (cA(i, k_i) + A'(i, k_i))b = cD(\alpha_i) + D(\alpha'_i).$$

Logo, D é uma função n -linear.

Caso particular: o produto dos elementos da diagonal $D(A) = A_{11} \cdot \dots \cdot A_{nn}$ é uma função n -linear sobre $\mathbb{F}^{n \times n}$. $\square$

Exemplo 5.5

Seja D uma função **bilinear** (2-linear) sobre as matrizes 2×2 com entradas em $\mathbb{F}$.

Fazendo $e_1 = (1, 0)$ e $e_2 = (0, 1)$, temos

$$\begin{aligned} D(A) &= D(A_{11}e_1 + A_{12}e_2, A_{21}e_1 + A_{22}e_2) \\ &= A_{11}D(e_1, A_{21}e_1 + A_{22}e_2) + A_{12}D(e_2, A_{21}e_1 + A_{22}e_2) \\ &= A_{11}A_{21}D(e_1, e_1) + A_{11}A_{22}D(e_1, e_2) \\ &\quad + A_{12}A_{21}D(e_2, e_1) + A_{12}A_{22}D(e_2, e_2). \end{aligned}$$

$\square$

Lema 5.1

Uma combinação linear de funções n -lineares sobre $\mathbb{F}^{n \times n}$ é n -linear.

Prova.

Basta mostrar que uma combinação linear de duas funções n -lineares é n -linear.

Ao lado, $A(i, j)$ designa a entrada A_{ij} da matriz A .

Sejam D e E funções n -lineares, e sejam $a, b \in \mathbb{F}$.

A combinação linear $aD + bE$ é definida por

$$(aD + bE)(A) = aD(A) + bE(A).$$

Então, fixando todas as linhas exceto a linha i , temos

$$\begin{aligned} (aD + bE)(c\alpha_i + \alpha'_i) &= aD(c\alpha_i + \alpha'_i) + bE(c\alpha_i + \alpha'_i) \\ &= acD(\alpha_i) + aD(\alpha'_i) + bcE(\alpha_i) + bE(\alpha'_i) \\ &= c(aD + bE)(\alpha_i) + (aD + bE)(\alpha'_i). \end{aligned}$$

Como queríamos demonstrar. ■

Exemplo 5.6

Seja $\mathbb{F}$ um anel comutativo com elemento unidade e seja D a função definida sobre as matrizes 2×2 com entradas em $\mathbb{F}$ por

$$D(A) = A_{11}A_{22} - A_{12}A_{21}.$$

Como $D = D_1 + D_2$, onde $D_1(A) = A_{11}A_{22}$ e $D_2(A) = -A_{12}A_{21}$, temos, pelo lema anterior, que D é uma função bilinear.

Vejamos algumas propriedades dessa função:

- Se I é a matriz identidade 2×2 , então $D(I) = \mathbf{1}$, isto é, $D(e_1, e_2) = \mathbf{1}$.
- Se as duas linhas de A são iguais, então

$$D(A) = A_{11}A_{22} - A_{12}A_{21} = A_{11}A_{12} - A_{12}A_{11} = 0.$$

- Se A' é a matriz obtida de uma matriz $A \in \mathbb{F}^{2 \times 2}$ permutando suas linhas, então

$$\begin{aligned} D(A') &= A'_{11}A'_{22} - A'_{12}A'_{21} \\ &= A_{21}A_{12} - A_{22}A_{11} = -D(A). \end{aligned}$$

□

Definição 5.3

Uma função n -linear D é dita **alternada**, se as seguintes condições são satisfeitas:

- (a) $D(A) = 0$ sempre que duas linhas de A são iguais.
- (b) Se A' é a matriz obtida permutando duas linhas de A , então

$$D(A') = -D(A).$$

Observação 5.2

Demonstraremos abaixo que toda função n –linear que satisfaz (a) também satisfaz (b).

Por outro lado, se D satisfaz a condição (b) e A é uma matriz com duas linhas iguais, então $D(A) = -D(A)$, ou seja, $D(A) + D(A) = 0$. Podemos, então, concluir que D satisfaz a condição (a) se, por exemplo, $\mathbb{F}$ é um corpo onde $1 + 1 \neq 0$, mas em geral (a) não é uma consequência de (b).

Definição 5.4

Seja $\mathbb{F}$ um anel comutativo com elemento unidade e seja n um inteiro positivo. Suponhamos que D seja uma função de $\mathbb{F}^{n \times n}$ em $\mathbb{F}$. Dizemos que D é uma *função determinante* se D é n –linear, alternada e $D(I) = 1$.

A existência e a unicidade da função determinante é evidente para matrizes 1×1 , $A = [a]$, sobre $\mathbb{F}$. Basta tomar $D(A) = a$.

Pelo exemplo 5.5, sabemos que toda função bilinear sobre as matrizes 2×2 sobre $\mathbb{F}$ é da forma

$$\begin{aligned} D(A) &= A_{11}A_{21}D(e_1, e_1) + A_{11}A_{22}D(e_1, e_2) \\ &\quad + A_{12}A_{21}D(e_2, e_1) + A_{12}A_{22}D(e_2, e_2). \end{aligned}$$

Se D é alternada, temos que

$$D(e_1, e_1) = D(e_2, e_2) = 0 \quad \text{e} \quad D(e_2, e_1) = -D(e_1, e_2).$$

Logo, $D(A) = (A_{11}A_{22} - A_{12}A_{21})D(e_1, e_2)$.

Se, além disso, $D(I) = D(e_1, e_2) = 1$, temos que

$$D(A) = A_{11}A_{22} - A_{12}A_{21},$$

é a única função determinante das matrizes 2×2 com entradas no anel $\mathbb{F}$.

Antes de provarmos a existência e a unicidade da função determinante das matrizes $n \times n$ com entradas no anel $\mathbb{F}$, precisamos provar alguns resultados preliminares.

Lema 5.2

Seja D uma função n –linear sobre as matrizes $n \times n$ com entradas no anel $\mathbb{F}$. Suponhamos que $D(A) = 0$ sempre que duas linhas adjacentes na matriz A sejam iguais. Então D é alternada.

Prova.

- Primeiro vamos provar que $D(A') = -D(A)$ sendo A' obtida transpondo-se duas linhas adjacentes da matriz A .

Sejam $\alpha_1, \dots, \alpha_n$ as linhas de A e suponhamos que

$$A' = (\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \alpha_i, \alpha_{i+2}, \dots, \alpha_n),$$

isto é, A' é obtida transpondo-se as linhas α_i e α_{i+1} da matriz A .

Então,

$$\begin{aligned} & D(\alpha_1, \dots, \alpha_{i-1}, \alpha_i + \alpha_{i+1}, \alpha_i + \alpha_{i+1}, \alpha_{i+2}, \dots, \alpha_n) \\ = & D(\alpha_1, \dots, \alpha_{i-1}, \alpha_i, \alpha_i, \alpha_{i+2}, \dots, \alpha_n) \\ & + D(\alpha_1, \dots, \alpha_{i-1}, \alpha_i, \alpha_{i+1}, \alpha_{i+2}, \dots, \alpha_n) \\ & + D(\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \alpha_i, \alpha_{i+2}, \dots, \alpha_n) \\ = & D(\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \alpha_{i+1}, \alpha_{i+2}, \dots, \alpha_n). \end{aligned}$$

Por hipótese,

$$\begin{aligned} D(\alpha_1, \dots, \alpha_{i-1}, \alpha_i, \alpha_i, \alpha_{i+2}, \dots, \alpha_n) &= 0 \\ D(\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \alpha_{i+1}, \alpha_{i+2}, \dots, \alpha_n) &= 0 \\ D(\alpha_1, \dots, \alpha_{i-1}, \alpha_i + \alpha_{i+1}, \alpha_i + \alpha_{i+1}, \alpha_{i+2}, \dots, \alpha_n) &= 0. \end{aligned}$$

Logo,

$$D(\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \alpha_i, \alpha_{i+2}, \dots, \alpha_n) = -D(\alpha_1, \dots, \alpha_{i-1}, \alpha_i, \alpha_{i+1}, \alpha_{i+2}, \dots, \alpha_n),$$

ou seja, $D(A') = -D(A)$.

- Seja agora B a matriz obtida transpondo-se as linhas i e j da matriz A , sendo $i < j$.

Podemos obter B a partir de A por uma sucessão de transposições de pares de linhas adjacentes.

Primeiro transpomos as linhas i e $i + 1$ e continuamos até que as linhas estejam na ordem

$$\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_j, \alpha_i, \alpha_{j+1}, \dots, \alpha_n.$$

Para isso, foram necessárias $k = j - i$ transposições de linhas adjacentes.

Agora, temos que deslocar α_j para a i -ésima posição, realizando $j - 1 - i = k - 1$ transposições de linhas adjacentes.

Dessa forma, obtemos B a partir de A por meio de $k + k - 1 = 2k - 1$ transposições de linhas adjacentes.

Assim,

$$D(B) = (-1)^{2k-1} D(A) = -D(A).$$

- Seja agora A uma matriz com duas linhas iguais, digamos $\alpha_i = \alpha_j$, com $i < j$.

Se $j = i + 1$, $D(A) = 0$, pois A tem duas linhas adjacentes iguais.

Se $j > i + 1$, a matriz B , obtida transpondo as linhas j e $i + 1$ da matriz A , tem duas linhas adjacentes iguais e, portanto, $D(B) = 0$. Mas, como,

$$D(A) = -D(B),$$

temos que $D(A) = 0$. ■

Definição 5.5

Se $n > 1$ e A é uma matriz $n \times n$ com entradas no anel $\mathbb{F}$, indicamos por $A(i|j)$ a matriz $(n-1) \times (n-1)$ obtida de A , retirando-se a sua i -ésima linha e a sua j -ésima coluna.

Se D é uma função $(n-1)$ -linear e A é uma matriz $n \times n$, escrevemos $D_{ij} = D(A(i|j))$.

Teorema 5.1

Seja $n > 1$ e seja D uma função $(n-1)$ -linear alternada sobre as matrizes $(n-1) \times (n-1)$ com entradas no anel $\mathbb{F}$.

Para cada j , $1 \leq j \leq n$, a função E_j definida por

$$E_j(A) = \sum_{i=1}^n (-1)^{i+j} A_{ij} D_{ij}(A)$$

é uma função n -linear alternada sobre as $n \times n$ matrizes A .

Se D é uma função determinante, então cada E_j também o é.

Prova.

Como D é $(n-1)$ -linear e $D_{ij}(A)$ independe da i -ésima linha, temos que D_{ij} é uma função linear de todas as linhas, exceto a i -ésima. Mas, como a função $A \mapsto A_{ij}$ é linear com respeito apenas à i -ésima linha de A , temos que $A_{ij} D_{ij}$ é uma função n -linear de A .

Logo, E_j é n -linear, pois uma combinação linear de funções n -lineares é n -linear.

Para mostrar que E_j é alternada, basta mostrar, pelo lema anterior, que $E_j(A) = 0$ sempre que A tiver duas linhas adjacentes iguais.

Suponhamos que $\alpha_k = \alpha_{k+1}$. Se $i \neq k$ e $i \neq k+1$, a matriz $A(i|j)$ tem duas linhas iguais e, portanto, $D_{ij}(A) = 0$.

Logo,

$$E_j(A) = (-1)^{k+j} A_{kj} D_{kj}(A) + (-1)^{k+1+j} A_{(k+1)j} D_{(k+1)j}(A).$$

Como $\alpha_k = \alpha_{k+1}$, temos que

$$A_{kj} = A_{(k+1)j} \quad \text{e} \quad A(k|j) = A(k+1|j).$$

Então, $E_j(A) = 0$.

Suponhamos, agora, que D seja uma função determinante, ou seja $D(I^{n-1}) = 1$, onde estamos designando I^{n-1} a matriz identidade $(n-1) \times (n-1)$.

Se I^n é a matriz identidade $n \times n$, temos que $I^n(j|j) = I^{n-1}$, $1 \leq j \leq n$, e $I^n_{ij} = 0$, se $i \neq j$.

Logo,

$$E_j(I^n) = D(I^n(j|j)) = D(I^{n-1}) = 1,$$

ou seja, E_j é uma função determinante. ■

Corolário 5.1

Seja $\mathbb{F}$ um anel comutativo com elemento unidade e seja n um inteiro positivo. Então existe uma função determinante sobre $\mathbb{F}^{n \times n}$.

Prova.

Já provamos a existência da função determinante para $n = 1$ e $n = 2$. O corolário segue por indução, pois o teorema anterior nos diz como construir uma função determinante sobre matrizes $n \times n$, a partir de uma função determinante sobre matrizes $(n-1) \times (n-1)$.

O nosso objetivo agora é mostrar a *unicidade* da função determinante.

Suponhamos que D seja uma função n -linear alternada sobre as matrizes $n \times n$ sobre $\mathbb{F}$. Seja A uma matriz $n \times n$ com entradas em $\mathbb{F}$ que tem por linhas $\alpha_1, \dots, \alpha_n$, e sejam $e_1, \dots, e_n$ as linhas da matriz identidade $n \times n$.

Como

$$\alpha_i = \sum_{j=1}^n A(i,j) e_j, \quad 1 \leq i \leq n,$$

temos que

$$\begin{aligned} D(A) &= D\left(\sum_{j=1}^n A(1, j)e_j, \alpha_2, \dots, \alpha_n\right) \\ &= \sum_{j=1}^n A(1, j)D(e_j, \alpha_2, \dots, \alpha_n). \end{aligned}$$

Se, agora, substituirmos α_2 por $\sum_{k=1}^n A(2, k)e_k$, temos que

$$D(e_j, \alpha_2, \dots, \alpha_n) = \sum_{k=1}^n A(2, k)D(e_j, e_k, \dots, \alpha_n)$$

Assim,

$$D(A) = \sum_{k,j} A(1, j)A(2, k)D(e_j, e_k, \alpha_3, \dots, \alpha_n).$$

Ou seja, se substituirmos cada linha α_i por $\sum_{k=1}^n A(i, k)e_k$, $i = 1, \dots, n$, obteremos que

$$D(A) = \sum_{k_1, \dots, k_n=1}^n A(1, k_1)A(2, k_2) \dots A(n, k_n)D(e_{k_1}, e_{k_2}, \dots, e_{k_n}).$$

Como D é alternada, $D(e_{k_1}, e_{k_2}, \dots, e_{k_n}) = 0$ sempre que dois dos índices k_i são iguais. ■

Definição 5.6

Uma seqüência $(k_1, \dots, k_n)$ de inteiros positivos menores ou iguais a n , com a propriedade de não existirem dois k_i iguais, é denominada uma *permutação de grau n* .

Uma permutação de grau n pode ser definida como uma função bijetora σ do conjunto $\{1, 2, \dots, n\}$ em si mesmo. Tal função σ corresponde à n -upla $(\sigma_1, \sigma_2, \dots, \sigma_n)$ e é simplesmente uma regra para ordenar $1, 2, \dots, n$ de outra maneira.

Assim,

$$D(A) = \sum_{\sigma} A(1, \sigma_1) \dots A(n, \sigma_n)D(e_{\sigma_1}, \dots, e_{\sigma_n}),$$

onde a soma é estendida a todas as permutações σ distintas de grau n .

Como a sequência $(\sigma_1, \dots, \sigma_n)$ pode ser obtida da sequência $(1, \dots, n)$ após um número finito m , $0 \leq m \leq n$, de transposições de pares de elementos, e D é alternada, temos que

$$D(e_{\sigma_1}, \dots, e_{\sigma_n}) = (-1)^m D(e_1, \dots, e_n).$$

Em particular, se D é uma função determinante,

$$D(e_{\sigma_1}, \dots, e_{\sigma_n}) = (-1)^m,$$

onde m depende somente de σ e não de D .

Com isto, podemos provar um fato básico sobre permutações.

Proposição 5.1

O número de transposições de pares de elementos usadas para passar da sequência $(1, 2, \dots, n)$ para a sequência $(\sigma_1, \dots, \sigma_n)$ é sempre par ou sempre ímpar.

Prova.

Seja D uma função determinante sobre as matrizes $n \times n$ sobre $\mathbb{F}$, cuja existência já foi provada.

Seja σ uma permutação de grau n e suponhamos que passamos de $(1, 2, \dots, n)$ a $(\sigma_1, \dots, \sigma_n)$ por meio de m transposições de pares (i, j) , $i \neq j$. Então, $D(e_{\sigma_1}, \dots, e_{\sigma_n}) = (-1)^m$.

Se $D(e_{\sigma_1}, \dots, e_{\sigma_n}) = 1$, m tem que ser par, e se $D(e_{\sigma_1}, \dots, e_{\sigma_n}) = -1$, m tem que ser ímpar. ■

Definição 5.7

Se o número de transposições de pares usadas para passar da sequência $(1, \dots, n)$ à sequência $(\sigma_1, \dots, \sigma_n)$ é sempre par (ímpar) dizemos que **a permutação é par (ímpar)**.

Define-se, também, o **sinal** de uma permutação por

$$\text{sinal } \sigma = \begin{cases} 1, & \text{se } \sigma \text{ é par} \\ -1, & \text{se } \sigma \text{ é ímpar.} \end{cases}$$

Teorema 5.2

Seja $\mathbb{F}$ um anel comutativo com elemento unidade e seja n um inteiro positivo. Então existe exatamente uma função determinante sobre o conjunto

das matrizes $n \times n$ com entradas em $\mathbb{F}$, que designamos $\det$ e é dada por

$$\det(A) = \sum_{\sigma} (\text{sinal } \sigma) A(1, \sigma_1) \dots A(n, \sigma_n),$$

sendo a soma estendida a todas as permutações distintas σ de grau n .

Se D é uma função n -linear alternada arbitrária sobre $\mathbb{F}^{n \times n}$, então

$$D(A) = \det(A) \cdot D(I)$$

para toda matriz $A \in \mathbb{F}^{n \times n}$.

Prova.

Já verificamos que se D é uma função n -linear alternada sobre $\mathbb{F}^{n \times n}$, então

$$D(A) = \sum_{\sigma} A(1, \sigma_1) \dots A(n, \sigma_n) D(e_{\sigma_1}, \dots, e_{\sigma_n}),$$

sendo a soma estendida a todas as permutações distintas σ de grau n .

Logo, como $D(e_{\sigma_1}, \dots, e_{\sigma_n}) = \text{sinal } \sigma$, temos

$$D(A) = \sum_{\sigma} (\text{sinal } \sigma) A(1, \sigma_1) \dots A(n, \sigma_n) D(I), \quad (I)$$

Provamos, assim, que

$$D(A) = \sum_{\sigma} (\text{sinal } \sigma) A(1, \sigma_1) \dots A(n, \sigma_n)$$

é a única função determinante sobre $\mathbb{F}^{n \times n}$, que denotaremos por $\det(A)$.

Se D é uma função n -linear alternada sobre $\mathbb{F}^{n \times n}$, então, por (I),

$$D(A) = \det(A) \cdot D(I).$$



Observação 5.3

Existem exatamente $n! = 1 \cdot 2 \cdot \dots \cdot n$ permutações de grau n , pois se σ é uma tal permutação, existem n escolhas possíveis para σ_1 ; $n - 1$ possibilidades para σ_2 ; $n - 2$ possibilidades para σ_3 , e assim por diante.

A fórmula

$$\det(A) = \sum_{\sigma} (\text{sinal } \sigma) A(1, \sigma_1) \dots A(n, \sigma_n)$$

fornece $\det(A)$ como uma soma de $n!$ termos, um para cada permutação

σ de grau n . Um termo genérico é um produto $A(1, \sigma_1) \dots A(n, \sigma_n)$ de n entradas da matriz A , uma entrada de cada linha e uma de cada coluna, acompanhado de um sinal $+$ ou $-$, conforme a permutação σ seja par ou ímpar.

Teorema 5.3

Seja $\mathbb{F}$ um anel comutativo com elemento unidade e sejam A e B matrizes $n \times n$ com entradas em $\mathbb{F}$. Então,

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

Prova.

Definimos a função $D(A) = \det(AB)$. Indicando por $\alpha_1, \dots, \alpha_n$ as linhas de A , temos que

$$D(\alpha_1, \dots, \alpha_n) = \det(\alpha_1 B, \dots, \alpha_n B).$$

Como $(c\alpha_i + \alpha'_i)B = c(\alpha_i B) + (\alpha'_i B)$ e a função $\det$ é n -linear, concluímos que D é n -linear.

Se $\alpha_i = \alpha_j$, então $\alpha_i B = \alpha_j B$, e já que $\det$ é alternada, temos

$$D(\alpha_1, \dots, \alpha_n) = 0,$$

ou seja, D é alternada.

Sendo D uma função n -linear alternada, do teorema anterior segue que

$$D(A) = \det(A) \cdot D(I).$$

Mas $D(I) = \det(IB) = \det(B)$. Portanto,

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

Como queríamos demonstrar. ■

Observação 5.4

1. Como as permutações são funções bijetoras do conjunto $\{1, \dots, n\}$ sobre si mesmo, podemos definir o *produto das permutações σ e τ* como sendo a função composta

$$\sigma \circ \tau(i) = \sigma(\tau(i)), \quad i = 1, \dots, n.$$

Se $\mathcal{E}$ indica a *permutação idêntica* (ou identidade), $E(i) = i$, então cada σ possui uma inversa σ^{-1} tal que $\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = \mathcal{E}$.

Então, o conjunto das permutações de grau n , com o produto dado pela composição, é um grupo, denominado *grupo simétrico de grau n* .

2. Sejam σ e τ permutações de grau n , e sejam $e_1, \dots, e_n$ as linhas da matriz identidade $n \times n$.

Seja A a matriz de linhas $e_{\tau_1}, \dots, e_{\tau_n}$ e seja B a matriz de linhas $e_{\sigma_1}, \dots, e_{\sigma_n}$. A i -ésima linha da matriz A tem exatamente um elemento não-nulo, a saber o 1 na coluna τ_i .

Assim, a i -ésima linha da matriz AB é $e_{\tau_i} B = e_{\sigma(\tau_i)}$, pois $e_{\tau_i} B$ é a τ_i -ésima linha da matriz B . Logo, $AB = (e_{\sigma\tau_1}, \dots, e_{\sigma\tau_n})$.

Como $\det(A) = \text{sinal } \tau$, $\det(B) = \text{sinal } \sigma$ e $\det(AB) = \text{sinal } (\sigma\tau)$, temos, pelo teorema anterior, que

$$\text{sinal } (\sigma \circ \tau) = (\text{sinal } \tau) \cdot (\text{sinal } \sigma).$$

Então, $\sigma \circ \tau$ é uma permutação par se σ e τ são ambas pares ou ambas ímpares, enquanto $\sigma \circ \tau$ é ímpar se uma das permutações é par e a outra é ímpar.

6. Propriedades dos Determinantes

(1) $\det(A^t) = \det(A)$

De fato, sendo

$$\begin{aligned} \det(A^t) &= \sum_{\sigma} (\text{sinal } \sigma) A^t(1, \sigma_1) \dots A^t(n, \sigma_n) \\ &= \sum_{\sigma} (\text{sinal } \sigma) A(\sigma_1, 1) \dots A(\sigma_n, n), \end{aligned}$$

e $A(\sigma_i, i) = A(j, \sigma^{-1}j)$, para $\sigma(i) = \sigma_i = j$, temos

$$\det(A^t) = \sum_{\sigma} (\text{sinal } \sigma) A(1, \sigma^{-1}(1)) \dots A(n, \sigma^{-1}(n)).$$

Além disso, como $\sigma \cdot \sigma^{-1}$ é a permutação idêntica, temos que

$$(\text{sinal } \sigma)(\text{sinal } \sigma^{-1}) = 1,$$

ou seja, $\text{sinal } \sigma^{-1} = \text{sinal } \sigma$.

Logo,

$$\det(A^t) = \sum_{\sigma} (\text{sinal } \sigma^{-1}) A(1, \sigma^{-1}(1)) \dots A(n, \sigma^{-1}(n)) = \det(A),$$

pois quando σ percorre todas as permutações de grau n , σ^{-1} também o faz.

Como consequência da igualdade $\det(A^t) = \det(A)$, temos que *a função $\det(A)$ é, também, uma função n -linear alternada das colunas da matriz A .*

(2) Se a matriz B é obtida da matriz A somando-se um múltiplo de uma linha (coluna) a outra linha (coluna), então $\det(B) = \det(A)$.

De fato, sejam $\alpha_1, \dots, \alpha_n$ as linhas de A e seja B a matriz obtida de A somando-se $c\alpha_j$ a α_i , onde $i < j$.

Como $\det$ é uma função linear da i -ésima linha, temos

$$\det(B) = \det(A) + c \det(\alpha_1, \dots, \alpha_{i-1}, \alpha_j, \alpha_{i+1}, \dots, \alpha_j, \dots, \alpha_n) = \det(A).$$

(3) Seja $\begin{pmatrix} A & B \\ O & C \end{pmatrix}$ uma matriz $n \times n$ na forma de blocos, onde A é uma matriz $r \times r$, C é uma matriz $s \times s$ (com $r + s = n$), B é uma matriz $r \times s$ e O indica a matriz nula $s \times r$. Então

$$\det \begin{pmatrix} A & B \\ O & C \end{pmatrix} = \det(A) \cdot \det(C)$$

De fato, definamos a função

$$D(A, B, C) = \det \begin{pmatrix} A & B \\ O & C \end{pmatrix}.$$

Se fixarmos A e B , então D é alternada e s -linear como uma função das linhas de C . Assim, pelo teorema 5.2,

$$D(A, B, C) = \det(C) \cdot D(A, B, I),$$

onde I é a matriz identidade $s \times s$. Subtraindo das linhas de B múltiplos das linhas de I , obtemos, pela propriedade anterior, que

$$D(A, B, I) = D(A, O, I).$$

Como $D(A, O, I)$ é alternada e r -linear como uma função das linhas de A , temos, pelo teorema 5.2, que

$$D(A, O, I) = \det(A) \cdot D(I, O, I).$$

$$\text{Mas } D(I, O, I) = \det \begin{pmatrix} I^r & O \\ O & I^s \end{pmatrix} = \det(I^n) = 1.$$

Logo,

$$\begin{aligned} D(A, B, C) &= \det(C) \cdot D(A, B, I) = \det(C) \cdot D(A, O, I) \\ &= \det(C) \cdot \det(A) \cdot D(I, O, I) = \det(A) \cdot \det(C). \end{aligned}$$

Por um raciocínio análogo, ou tomando transpostas, verificamos que

$$\det \begin{pmatrix} A & O \\ B & C \end{pmatrix} = \det(A) \det(C).$$

Exemplo 6.1

Seja $A \in \mathbb{Q}^{4 \times 4}$ a matriz

$$A = \begin{pmatrix} 1 & -1 & 2 & 3 \\ 2 & 2 & 0 & 2 \\ 4 & 1 & -1 & -1 \\ 1 & 2 & 3 & 0 \end{pmatrix}.$$

Subtraindo das linhas 2, 3 e 4 múltiplos convenientes da primeira linha, obtemos a matriz

$$\begin{pmatrix} 1 & -1 & 2 & 3 \\ 0 & 4 & -4 & -4 \\ 0 & 5 & -9 & -13 \\ 0 & 3 & 1 & -3 \end{pmatrix},$$

que tem, pela propriedade (2), o mesmo determinante que a matriz A .

Subtraindo da linha 3, $\frac{5}{4}$ da linha 2 e subtraindo da linha 4, $\frac{3}{4}$ da linha 2, obtemos a matriz

$$B = \begin{pmatrix} 1 & -1 & 2 & 3 \\ 0 & 4 & -4 & -4 \\ 0 & 0 & -4 & -8 \\ 0 & 0 & 4 & 0 \end{pmatrix},$$

cujo determinante é igual ao da matriz A . Como B é uma matriz em forma de blocos, temos que

$$\det(A) = \det(B) = \det \begin{pmatrix} 1 & -1 \\ 0 & 4 \end{pmatrix} \cdot \det \begin{pmatrix} -4 & -8 \\ 4 & 0 \end{pmatrix} = 4 \times 32 = 128.$$

□

No teorema 5.1, provamos que se $n > 1$ e se D é uma função determinante sobre as matrizes $(n-1) \times (n-1)$, então

$$E_j(A) = \sum_{i=1}^n (-1)^{i+j} A_{ij} D_{ij}(A)$$

é uma função determinante sobre as matrizes $n \times n$, para todo $j = 1, \dots, n$.

Pela unicidade da função determinante, temos, para cada $j = 1, \dots, n$,

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

O escalar

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

é chamado o *cofator* i, j da matriz A .

A fórmula acima para $\det(A)$ é denominada o *desenvolvimento de $\det(A)$ pelos cofatores da j -ésima coluna*, ou o *desenvolvimento pelos menores da j -ésima coluna*.

Assim, a fórmula acima nos diz que, para cada $j = 1, \dots, n$,

$$\det(A) = \sum_{i=1}^n A_{ij} C_{ij}, \quad (I)$$

onde o cofator C_{ij} é $(-1)^{i+j}$ vezes o determinante da matriz $(n-1) \times (n-1)$ obtida de A retirando-se a i -ésima linha e a j -ésima coluna de A .

Se $j \neq k$, então,

$$\sum_{i=1}^n A_{ik} C_{ij} = 0. \quad (II)$$

De fato, seja B a matriz obtida de A substituindo a sua j -ésima coluna pela k -ésima coluna. Como B tem duas colunas iguais e $B(i|j) = A(i|j)$, temos que

$$\begin{aligned} 0 = \det(B) &= \sum_{i=1}^n (-1)^{i+j} B_{ij} \det(B(i|j)) \\ &= \sum_{i=1}^n (-1)^{i+j} A_{ik} \det(A(i|j)) \\ &= \sum_{i=1}^n A_{ik} C_{ij}. \end{aligned}$$

Então, por (I) e (II), temos que

$$\sum_{i=1}^n A_{ik} C_{ij} = \delta_{jk} \det(A). \quad (III)$$

A matriz transposta da matriz de cofatores de A é chamada a *matriz adjunta de A* e se designa por **Adj A** , ou seja,

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

Por (III), temos que

$$\sum_{i=1}^n (\text{Adj } A)_{ji} A_{ik} = \delta_{jk} \det(A),$$

ou seja,

$$(\text{Adj } A) \cdot A = \det(A) \cdot I. \quad (\text{IV})$$

Vamos provar, agora, que $A \cdot (\text{Adj } A) = \det(A) \cdot I$.

Como $A^t(i|j) = A(j|i)^t$, temos

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

isto é, o cofator i, j da matriz A^t é o cofator j, i da matriz A . Logo, $\text{Adj}(A^t) = (\text{Adj } A)^t$.

Assim, por (IV), temos que

$$\text{Adj}(A^t) \cdot A^t = \det(A^t) \cdot I = \det(A) \cdot I,$$

e transpondo,

$$A \cdot (\text{Adj}(A^t))^t = \det(A) \cdot I,$$

ou seja,

$$A \cdot (\text{Adj } A) = \det(A) \cdot I,$$

pois $(\text{Adj}(A^t))^t = \text{Adj } A$.

Resumindo, temos

$$A \cdot (\text{Adj } A) = (\text{Adj } A) \cdot A = \det(A) \cdot I$$

Da mesma forma que para matrizes sobre um corpo, definimos as matrizes invertíveis com entradas num anel comutativo com elemento unidade.

Definição 6.1

Seja $\mathbb{F}$ um anel comutativo com elemento unidade, uma matriz $A \in \mathbb{F}^{n \times n}$ é dita *invertível sobre $\mathbb{F}$* se existe uma matriz $B \in \mathbb{F}^{n \times n}$, dita *inversa de A* , tal que $AB = BA = I$.

Observação 6.1

Se a matriz $A \in \mathbb{F}^{n \times n}$ possui uma matriz inversa, então ela é única e se designa A^{-1} .

De fato, se $BA = AC = I$, então

$$C = IC = (BA)C = B(AC) = BI = B.$$

Como

$$A \cdot (\text{Adj } A) = (\text{Adj } A) \cdot A = \det(A) \cdot I,$$

temos que *se $\det(A)$ é invertível em $\mathbb{F}$, então A é uma matriz invertível sobre $\mathbb{F}$* e

$$A^{-1} = (\det(A))^{-1} \cdot \text{Adj } A.$$

Reciprocamente, se A é invertível sobre $\mathbb{F}$, $\det(A)$ é um elemento invertível do anel $\mathbb{F}$, pois se $BA = I$, temos:

$$1 = \det(I) = \det(B) \cdot \det(A).$$

Teorema 6.1

Seja $\mathbb{F}$ um anel comutativo com elemento unidade e seja $A \in \mathbb{F}^{n \times n}$. Então A é invertível sobre $\mathbb{F}$ se, e somente se, $\det(A) \in \mathbb{F}$ é um elemento invertível. Nesse caso, temos

$$A^{-1} = (\det(A))^{-1} \cdot \text{Adj } A$$

Em particular, *uma matriz $n \times n$ com entradas num corpo é invertível se, e somente se, seu determinante é diferente de zero.*

Exemplo 6.2

Seja $\mathbb{F}$ o anel dos inteiros e seja $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

Então, $\det(A) = -2$ e $\text{Adj } A = \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix}$.

Assim, como $\det(A) \neq \pm 1$, A não é invertível sobre o anel dos inteiros, mas é invertível como uma matriz sobre o corpo dos números racionais, sendo, nesse caso,

$$A^{-1} = \frac{1}{-2} \text{Adj } A = -\frac{1}{2} \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & -1 \\ 3/2 & -1/2 \end{pmatrix}.$$

□

Observação 6.2

Se $\mathbb{F}$ é um anel comutativo com elemento unidade e $A \in \mathbb{F}^{n \times n}$ possui uma inversa à esquerda ou à direita, então é invertível.

De fato, suponhamos que A possui uma inversa B à direita. Como $AB = I$, temos que $\det(A) \cdot \det(B) = 1$, ou seja, $\det(A)$ possui um elemento inverso em $\mathbb{F}$. Logo, pelo teorema acima, A é invertível e

$$B = A^{-1} = (\det(A))^{-1} \cdot \text{Adj } A.$$

Observação 6.3

Se $\mathbb{K}$ é um corpo, os únicos elementos do anel de polinômios $\mathbb{K}[x]$ que são invertíveis são os polinômios constantes não-nulos.

De fato, se $f, g \in \mathbb{K}[x]$ e $fg = 1$, temos que $\text{grau}(f) + \text{grau}(g) = 0$.

Assim, $\text{grau}(f) = \text{grau}(g) = 0$, ou seja f e g são polinômios constantes não-nulos.

Portanto, *uma matriz $n \times n$ com entradas em $\mathbb{K}[x]$ é invertível sobre $\mathbb{K}[x]$ se, e somente se, seu determinante é um polinômio constante não-nulo.*

Exemplo 6.3

Consideremos as matrizes $A, B \in (\mathbb{R}[x])^{2 \times 2}$ dadas por:

$$A = \begin{pmatrix} x^2 + x & x + 1 \\ x - 1 & 1 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} x^2 - x & x + 2 \\ x^2 - 2x - 3 & x \end{pmatrix}.$$

Então,

$$\begin{aligned} \det(A) &= (x^2 + x) \cdot 1 - (x + 1)(x - 1) = x + 1, \\ \det(B) &= (x^2 - x)x - (x + 2)(x^2 - 2x - 3) = -6. \end{aligned}$$

Assim, A não é invertível sobre $\mathbb{R}[x]$ e B é invertível sobre $\mathbb{R}[x]$.

Mais ainda, como

$$\text{Adj } A = \begin{pmatrix} 1 & -x - 1 \\ -x + 1 & x^2 + x \end{pmatrix} \quad \text{e} \quad \text{Adj } B = \begin{pmatrix} x & -x - 2 \\ -x^2 + 2x - 3 & x^2 - 1 \end{pmatrix},$$

temos que $(\text{Adj } A)A = (x + 1)I$ e $(\text{Adj } B)B = -6I$, ou seja,

$$B^{-1} = -\frac{1}{6} \begin{pmatrix} x & -x - 2 \\ -x^2 + 2x - 3 & x^2 - 1 \end{pmatrix}.$$

□

Observação 6.4

Matrizes semelhantes tem o mesmo determinante. Isto é, se P é invertível e $B = P^{-1}AP$, então

$$\det B = \det(P^{-1}AP) = \det(P^{-1}) \cdot \det(A) \cdot \det(P) = \det(A),$$

pois, como $P^{-1}P = I$, $\det(P) \cdot \det(P^{-1}) = \det(I) = 1$.

Definição 6.2

Seja T um operador linear sobre um espaço vetorial V de dimensão finita. Definimos o **determinante de T** como sendo o determinante de $[T]_{\mathcal{B}}$, onde $\mathcal{B}$ é uma base ordenada de V .

O determinante de T está bem definido, pois se $\mathcal{B}'$ é outra base ordenada de V , então $[T]_{\mathcal{B}}$ e $[T]_{\mathcal{B}'}$ são semelhantes, e possuem, portanto, o mesmo determinante.

Observação 6.5

Seja $\mathbb{K}$ um corpo e seja $A \in \mathbb{K}^{n \times n}$. Consideremos o sistema de equações lineares $AX = Y$ para uma n -upla $Y = (y_1, \dots, y_n)$ dada.

Então,

$$(\text{Adj } A)AX = (\text{Adj } A)Y,$$

ou seja,

$$\det(A)X = (\text{Adj } A)Y.$$

Assim, se $X = (x_1, \dots, x_n)$, temos

$$\det(A)x_j = \sum_{i=1}^n (\text{Adj } A)_{ji}y_i = \sum_{i=1}^n (-1)^{i+j}y_i \det(A(i|j)) = \det(B_j),$$

onde B_j é a matriz $n \times n$ obtida substituindo a j -ésima coluna de A por Y .

Se $\det(A) \neq 0$, temos que o sistema $AX = Y$ possui uma única solução para X , dada por:

$$x_j = \frac{\det(B_j)}{\det(A)}, \quad j = 1, \dots, n$$

Ou ainda

$$X = \frac{1}{\det(A)} (\det(B_1), \dots, \det(B_n))$$

Formas Canônicas - Preliminares

1. Formas Canônicas Elementares

Definição 1.1

Seja V um espaço vetorial sobre o corpo $\mathbb{K}$ e seja T um operador linear sobre V . Um **autovalor de T** é um escalar $\lambda \in \mathbb{K}$ tal que existe um vetor $v \in V$ não-nulo tal que $T(v) = \lambda v$.

As terminologias **valor característico**, **valor próprio**, **raiz característica** e **autovalor** são equivalentes.

Se λ é um autovalor de T , então

- todo vetor $w \in V$ tal que $T(w) = \lambda w$ é chamado um **autovetor de T associado ao autovalor λ** .
- a coleção de todos os autovetores associados a um autovalor λ de T é denominado o **autoespaço de T associado a λ** .

As terminologias **vetor característico**, **vetor próprio** e **autovetor** são equivalentes. Também, as terminologias **espaço característico** e **autoespaço** são equivalentes.

O autoespaço de T associado a um autovalor λ é de fato um subespaço de V , pois

$$\{v \in V \mid Tv = \lambda v\} = \{v \in V \mid (T - \lambda I)v = \mathbf{0}\} = \text{núcleo}(T - \lambda I).$$

Assim, $\lambda \in \mathbb{K}$ é um autovalor de T se, e somente se, $T - \lambda I$ não é injetora, ou seja, o subespaço $\{v \in V \mid (T - \lambda I)v = \mathbf{0}\}$ não é o subespaço nulo.

Lema 1.1

Seja L um operador linear sobre um espaço vetorial V de dimensão finita.

Então, L é invertível se, e somente se, $\det(L) \neq 0$.

Prova.

Seja $\mathcal{B}$ uma base ordenada do espaço vetorial V . Como L é invertível se, e somente se, $[L]_{\mathcal{B}}$ é invertível, temos que L é invertível se, e somente se, $\det([L]_{\mathcal{B}}) \neq 0$, ou seja, se, e só se, $\det(L) \neq 0$. ■

Teorema 1.1

Seja T um operador linear sobre um espaço vetorial V de dimensão finita e seja λ um escalar. Então, as seguintes afirmações são equivalentes.

(a) λ é um autovalor de T .

(b) $T - \lambda I$ não é invertível.

(c) $\det(T - \lambda I) = 0$.

Prova.

Basta observar que sendo $T - \lambda I$ um operador sobre um espaço vetorial de dimensão finita, temos que $T - \lambda I$ não é injetiva se, e só se, $T - \lambda I$ não é invertível. ■

Definição 1.2

Seja $\mathbb{K}$ um corpo e seja $A \in \mathbb{K}^{n \times n}$. Um **autovalor de A em $\mathbb{K}$** é um escalar $\lambda \in \mathbb{K}$ tal que a matriz $A - \lambda I$ não é invertível.

Então, λ é um autovalor de A se, e só se,

$$\det(A - \lambda I) = \det(\lambda I - A) = 0.$$

Definição 1.3

O polinômio $p_c(x) = \det(xI - A)$, onde $xI - A$ é uma matriz com entradas em $\mathbb{K}[x]$, é denominado o **polinômio característico** da matriz A .

Observação 1.1

Os autovalores de A em $\mathbb{K}$ são as raízes do polinômio característico da matriz A .

Observação 1.2

O polinômio característico p_c de uma matriz $n \times n$ é um polinômio unitário de grau n .

De fato, pela fórmula do determinante de uma matriz em função de suas entradas

$$\det(B) = \sum_{\sigma} (\text{sinal } \sigma) B(1, \sigma_1) \dots B(n, \sigma_n),$$

temos que

$$\det(xI - A) = (x - A_{11}) \dots (x - A_{nn}) + \sum_{\sigma \neq \text{id}} (\text{sinal } \sigma) q_{\sigma}(x),$$

onde $q_{\sigma}(x)$ são polinômios de grau $\leq n - 1$.

Lema 1.2

Matrizes semelhantes têm o mesmo polinômio característico.

Prova.

Se $B = P^{-1}AP$, então

$$\begin{aligned} \det(xI - B) &= \det(xI - P^{-1}AP) \\ &= \det(P^{-1}(xI - A)P) \\ &= \det(P^{-1}) \det(xI - A) \det(P) \\ &= \det(xI - A), \end{aligned}$$

pois $\det(P^{-1}) \cdot \det(P) = \det(P^{-1}P) = \det(I) = 1$. ■

Definição 1.4

Seja T um operador linear sobre um espaço vetorial V de dimensão finita e seja $\mathcal{B}$ uma base ordenada de V . O **polinômio característico do operador** T é, por definição, o polinômio característico da matriz $[T]_{\mathcal{B}}$.

Observe que, em virtude do lema, o polinômio característico de um operador depende apenas do operador e não da matriz representante de T utilizada para determiná-lo. Isto é, qualquer que seja a base $\mathcal{B}$, o polinômio característico de T é $\det(xI - [T]_{\mathcal{B}})$.

Observação 1.3

Os **autovalores** do operador T são as raízes do polinômio característico de T .

De fato, λ é autovalor de T se, e só se, $T - \lambda I$ não é invertível. Ou seja, se, e só se, $[T - \lambda I]_{\mathcal{B}} = [T]_{\mathcal{B}} - \lambda I$ não é invertível. Logo, λ é autovalor de T se, e somente se, $p_C(\lambda) = \det(\lambda I - [T]_{\mathcal{B}}) = 0$.

Observação 1.4

Como o polinômio característico tem grau n , T possui no máximo n autovalores diferentes. Mas, T pode não ter nenhum autovalor.

Exemplo 1.1

Seja T o operador linear sobre $\mathbb{R}^2$ que é representado, em relação à base canônica, pela matriz

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

O polinômio característico de T (ou de A) é

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

Como esse polinômio não possui raízes reais, T não possui autovalores.

Mas se U é o operador linear sobre $\mathbb{C}^2$ que é representado pela matriz A em relação à base canônica, então U possui dois autovalores: i e $-i$. $\square$

Assim, ao discutirmos os autovalores de uma matriz A , precisamos tomar o cuidado de estipular o corpo envolvido.

Exemplo 1.2

Seja A a matriz 3×3 real

$$A = \begin{pmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{pmatrix}.$$

O polinômio característico de A é

$$\det(xI - A) = \det \begin{pmatrix} x-3 & -1 & 1 \\ -2 & x-2 & 1 \\ -2 & -2 & x \end{pmatrix} = x^3 - 5x^2 + 8x - 4 = (x-1)(x-2)^2.$$

Logo, 1 e 2 são os autovalores de A .

Seja T o operador linear sobre $\mathbb{R}^3$ representado por A em relação à base canônica.

Determinemos os autovetores de T associados aos autovalores 1 e 2.

Como

$$A - 1I = A - I = \begin{pmatrix} 2 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 2 & -1 \end{pmatrix},$$

tem posto dois, temos que o núcleo de $T - I$ tem dimensão 1, ou seja, o

espaço dos autovetores associados ao autovalor 1 tem dimensão 1.

Sendo $(T - I)(1, 0, 2) = (0, 0, 0)$, temos que $v_1 = (1, 0, 2)$ gera o núcleo de $T - I$. Logo,

$$\{v \in \mathbb{R}^3 \mid Tv = v\} = \{k v_1 \mid k \in \mathbb{R}\}.$$

Consideremos agora a matriz

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

Como o posto de $A - 2I$ tem dimensão dois, o espaço dos autovetores de T associados ao autovalor 2 tem dimensão 1.

Sendo $(T - 2I)(1, 1, 2) = (0, 0, 0)$, temos que

$$\{v \mid Tv = 2v\} = \{k v_2 \mid k \in \mathbb{R}\},$$

onde $v_2 = (1, 1, 2)$. $\square$

Definição 1.5

Seja T um operador linear sobre o espaço V de dimensão finita. Dizemos que T é **diagonalizável** se existe uma base de V formada por autovetores de T .

Ou seja, T é diagonalizável se seus autovetores geram o espaço V .

Se $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ é uma base de autovetores de V e $T(\alpha_i) = c_i \alpha_i$, $i = 1, \dots, n$, então

$$[T]_{\mathcal{B}} = \begin{pmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_n \end{pmatrix},$$

é uma matriz diagonal.

Exemplo 1.3

O operador linear T sobre $\mathbb{R}^2$ do exemplo 1.1 não é diagonalizável, pois T não possui autovalores em $\mathbb{R}$. $\square$

Exemplo 1.4

O operador linear T sobre $\mathbb{R}^3$ do exemplo 1.2 apesar de possuir dois au-

tovalores 1 e 2 no corpo $\mathbb{R}$, não é diagonalizável, pois os espaços dos autovetores associados a esses autovalores têm dimensão 1. Portanto, não há possibilidade de formar uma base de $\mathbb{R}^3$ constituída de autovetores de T . $\square$

Suponhamos que o operador $T : V \rightarrow V$ seja diagonalizável e que $c_1, \dots, c_k$ são os autovalores distintos de T . Seja $\mathcal{B}$ uma base ordenada de V formada por autovetores de T . Então $[T]_{\mathcal{B}}$ é uma matriz diagonal, cujos elementos da diagonal são os escalares c_i , cada um repetido um determinado número de vezes. Se c_i está repetido d_i vezes, podemos, reordenando a base $\mathcal{B}$, caso necessário, fazer com que a matriz do operador tenha a forma em blocos:

$$[T]_{\mathcal{B}} = \begin{pmatrix} c_1 I_{d_1} & O & \cdots & O \\ O & c_2 I_{d_2} & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & c_k I_{d_k} \end{pmatrix},$$

onde I_j é a matriz unidade $d_j \times d_j$.

Nesse caso, o polinômio característico de T é o produto de fatores lineares:

$$p_c(x) = (x - c_1)^{d_1} \cdots (x - c_k)^{d_k}.$$

Portanto, se o corpo $\mathbb{K}$ não for algebricamente fechado, estaremos observando uma propriedade especial de T , ao dizermos que seu polinômio característico tem uma tal fatoração.

Também podemos observar que a multiplicidade d_i do autovalor c_i como raiz do polinômio característico é igual à dimensão do espaço de autovetores associado ao autovalor c_i .

De fato, como $[T - c_i I]_{\mathcal{B}}$ é uma matriz diagonal com d_i zeros na diagonal, temos que a dimensão do núcleo de $T - c_i I$ é igual a d_i .

A relação entre a dimensão do autoespaço e a multiplicidade do autovalor como uma raiz do polinômio característico nos fornecerá uma maneira simples de verificar se um operador dado é ou não diagonalizável.

Lema 1.3

Suponhamos que $T\alpha = c\alpha$. Se f é um polinômio arbitrário, então

$$f(T)(\alpha) = f(c)\alpha.$$

Prova.

Seja $f = a_n x^n + \dots + a_1 x + a_0$. Então,

$$\begin{aligned} f(T)(\alpha) &= (a_n T^n + \dots + a_1 T + a_0 I)(\alpha) \\ &= a_n T^n(\alpha) + \dots + a_1 T(\alpha) + a_0 I(\alpha) \\ &= a_n c^n \alpha + \dots + a_1 c \alpha + a_0 \alpha \\ &= (a_n c^n + \dots + a_1 c + a_0) \alpha \\ &= f(c) \alpha. \end{aligned}$$



Lema 1.4

Seja T um operador linear sobre um espaço vetorial V de dimensão finita.

Sejam $c_1, \dots, c_k$ os autovalores distintos de T e seja W_i , $i = 1, \dots, k$, o espaço dos autovetores associado ao autovalor c_i . Se $W = W_1 + \dots + W_k$ e $\mathcal{B}_i$ é uma base ordenada de W_i , então $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ é uma base ordenada de W . Em particular, $\dim W = \dim W_1 + \dots + \dim W_k$.

Prova.

Seja $v_1 + \dots + v_k = \mathbf{0}$, onde $v_i \in W_i$, $i = 1, \dots, k$, e seja f um polinômio arbitrário.

Como $T(v_i) = c_i v_i$, temos, pelo lema anterior, que

$$\mathbf{0} = f(T)(\mathbf{0}) = f(T)v_1 + \dots + f(T)v_k = f(c_1)v_1 + \dots + f(c_k)v_k.$$

Sejam $f_1, \dots, f_k$ polinômios tais que

$$f_i(c_j) = \delta_{ij} = \begin{cases} 1, & i = j; \\ 0, & i \neq j. \end{cases}$$

Então,

$$\mathbf{0} = f_i(T)(\mathbf{0}) = \sum_{j=1}^k f_i(c_j)v_j = \sum_{j=1}^k \delta_{ij}v_j = v_i.$$

Como $W = W_1 + \dots + W_k$ é o espaço gerado por todos os autovetores de T , temos que $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ gera W , onde $\mathcal{B}_i$ é uma base de W_i , $i = 1, \dots, k$.

Seja $\mathcal{B}_i = \{v_1^i, \dots, v_{n_i}^i\}$, $i = 1, \dots, k$, e seja

$$a_1^1 v_1^1 + \dots + a_{n_1}^1 v_{n_1}^1 + a_1^2 v_1^2 + \dots + a_{n_2}^2 v_{n_2}^2 + \dots + a_1^k v_1^k + \dots + a_{n_k}^k v_{n_k}^k = \mathbf{0},$$

uma combinação linear nula dos vetores de $\mathcal{B}$.

Fazendo $v_i = a_1^i v_1^i + \dots + a_{n_i}^i v_{n_i}^i \in W_i$, temos que $v_1 + \dots + v_k = \mathbf{0}$. Pelo provado acima, $v_i = \mathbf{0}$ para cada $i = 1, \dots, k$.

Logo, $a_1^i v_1^i + \dots + a_{n_i}^i v_{n_i}^i = \mathbf{0}$ para todo $i = 1, \dots, k$. Como $\{v_1^i, \dots, v_{n_i}^i\} = \mathcal{B}_i$ é um conjunto LI, temos que $a_1^i = \dots = a_{n_i}^i = 0$ para todo $i = 1, \dots, k$. Provando, assim, que $\mathcal{B}$ é um conjunto LI de vetores de V .

Como $\mathcal{B}$ é LI e gera W , temos que $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ é uma base de W e, portanto,

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



Teorema 1.2

Seja T um operador linear sobre um espaço vetorial V de dimensão finita. Sejam $c_1, \dots, c_k$ os autovalores distintos de T e seja W_i o núcleo de $T - c_i I$. As seguintes afirmações são equivalentes:

(a) T é diagonalizável.

(b) O polinômio característico de T é

$$p_c = (x - c_1)^{d_1} \dots (x - c_k)^{d_k},$$

onde $d_k = \dim W_k$.

(c) $\dim V = \dim W_1 + \dots + \dim W_k$.

Prova.

Já provamos **(a) $\implies$ (b)**.

(b) $\implies$ (c) Suponhamos que $p_c = (x - c_1)^{d_1} \dots (x - c_k)^{d_k}$.

Então, $\text{grau}(p_c) = \dim V = d_1 + \dots + d_k = \dim W_1 + \dots + \dim W_k$.

Pelo lema anterior, $\dim(W_1 + \dots + W_k) = \dim V$.

Logo, $V = W_1 + \dots + W_k$.

(c) $\implies$ (a) Suponhamos, agora, que $\dim V = \dim W_1 + \dots + \dim W_k$.

Como $\dim W = \dim W_1 + \dots + \dim W_k$, onde $W = W_1 + \dots + W_k$, temos que $\dim W = \dim V$.

Então, $V = W$, ou seja, os autovetores de T geram V . ■

O análogo do teorema acima para *matrizes* pode ser formulado da seguinte maneira:

Teorema 1.3

Seja A uma matriz $n \times n$ com entradas no corpo $\mathbb{K}$ e sejam $c_1, \dots, c_k$ os autovalores distintos de A em $\mathbb{K}$. Para cada i , seja W_i o espaço dos vetores colunas em $\mathbb{K}^{n \times 1}$ tais que

$$(A - c_i I)X = \mathbf{0},$$

e seja B_i uma base ordenada de W_i .

Os vetores das bases $B_1, \dots, B_k$ podem ser reunidos para formar as colunas $P_1, \dots, P_\ell$ de uma matriz P .

Então, a matriz A é semelhante sobre $\mathbb{K}$ a uma matriz diagonal se, e só se, $\ell = n$, ou seja, se, e só se, P é uma matriz quadrada.

Nesse caso, P é invertível e $P^{-1}AP$ é diagonal.

Exemplo 1.5

Seja T o operador linear sobre $\mathbb{R}^3$ representado em relação à base canônica pela matriz

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

Vamos calcular o polinômio característico de A por meio de operações elementares sobre linhas e colunas da matriz $xI - A \in (\mathbb{R}[x])^{3 \times 3}$:

$$\begin{aligned} \det(xI - A) &= \det \begin{pmatrix} x-5 & 6 & 6 \\ 1 & x-4 & -2 \\ -3 & 6 & x+4 \end{pmatrix} = \det \begin{pmatrix} x-5 & 0 & 6 \\ 1 & x-2 & -2 \\ -3 & -x+2 & x+4 \end{pmatrix} \\ &= (x-2) \det \begin{pmatrix} x-5 & 0 & 6 \\ 1 & 1 & -2 \\ -3 & -1 & x+4 \end{pmatrix} = (x-2) \det \begin{pmatrix} x-5 & 0 & 6 \\ 1 & 1 & -2 \\ -2 & 0 & x+2 \end{pmatrix} \\ &= (x-2) \det \begin{pmatrix} x-5 & 6 \\ -2 & x+2 \end{pmatrix} = (x-2) ((x-5)(x+2) + 12) . \\ &= (x-2)(x^2 - 3x + 2) = (x-2)^2(x-1) . \end{aligned}$$

Então, 1 e 2 são os autovalores de T , e

$$A - I = \begin{pmatrix} 4 & -6 & -6 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{pmatrix} \quad \text{e} \quad A - 2I = \begin{pmatrix} 3 & -6 & -6 \\ -1 & 2 & 2 \\ 3 & -6 & -6 \end{pmatrix} .$$

Como $A - I$ não é invertível e posto $(A - I) \geq 2$, pois $(4, -6, -6)$ e $(-1, 3, 2)$ são LI, temos que posto $(A - I) = 2$.

Além disso, é claro que posto $(A - 2I) = 1$.

Sejam W_1 e W_2 os espaços dos autovetores associados aos autovalores 1 e 2, respectivamente.

Como $\dim W_1 = 3 - 2 = 1$ e $\dim W_2 = 3 - 1 = 2$, temos, pelo teorema anterior, que T é diagonalizável, pois $\dim W_1 + \dim W_2 = 1 + 2 = \dim \mathbb{R}^3$.

É fácil ver que o núcleo de $T - I$ é gerado pelo vetor $\alpha_1 = (3, -1, 3)$, e assim, $\{\alpha_1\}$ é uma base de W_1 .

O núcleo de $T - 2I$, isto é, o espaço W_2 , é formado pelos vetores (x_1, x_2, x_3) tais que $x_1 = 2x_2 + 2x_3$. Logo, por exemplo, os vetores

$$\alpha_2 = (2, 1, 0), \quad \text{e} \quad \alpha_3 = (2, 0, 1),$$

formam uma base de W_2 .

Então, $\mathcal{B} = \{\alpha_1, \alpha_2, \alpha_3\}$ é uma base de $\mathbb{R}^3$ e $[T]_{\mathcal{B}}$ é a matriz diagonal

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Assim, $D = P^{-1}AP$, onde

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

é a matriz de mudança de base da base $\mathcal{B}$ para a base canônica. $\square$

2. Polinômios Anuladores

Seja V um espaço vetorial sobre o corpo $\mathbb{K}$ e seja $T : V \longrightarrow V$ um operador sobre V .

Então, o conjunto

$$\mathbb{I} = \{p \in \mathbb{K}[x] \mid p(T) = O\},$$

é um ideal da álgebra $\mathbb{K}[x]$ dos polinômios com coeficientes em $\mathbb{K}$.

De fato, sejam $p, q \in \mathbb{K}[x]$ e seja $\lambda \in \mathbb{K}$. Como

$$(\lambda p + q)(T) = \lambda p(T) + q(T),$$

e

$$(p \cdot q)(T) = p(T) \cdot q(T),$$

temos que $\lambda p + q \in \mathbb{I}$ se $p, q \in \mathbb{I}$ e $p \cdot q \in \mathbb{I}$ se $p \in \mathbb{I}$.

O ideal $\mathbb{I}$ pode ser o ideal nulo, isto é, o ideal formado apenas pelo polinômio nulo. Mas veremos agora que isto não ocorre quando o espaço vetorial V é de dimensão finita.

De fato, se $\dim V = n$, temos que $\dim \mathcal{L}(V, V) = n^2$. Logo, os $n^2 + 1$ operadores $I, T, \dots, T^{n^2}$, são LD. Isto é, existem escalares $c_0, c_1, \dots, c_{n^2}$ em $\mathbb{K}$ não todos nulos tais que

$$c_0 I + c_1 T + \dots + c_{n^2} T^{n^2} = O.$$

Então, o polinômio

$$p = c_0 + c_1 x + \dots + c_{n^2} x^{n^2}$$

é um polinômio não-nulo de grau $\leq n^2$ que pertence ao ideal $\mathbb{I}$.

Como $\mathbb{I}$ é um ideal não-nulo, existe um único polinômio unitário $p \in \mathbb{K}[x]$ tal que $\mathbb{I} = p \mathbb{K}[x]$.

Ou seja, se $f \in \mathbb{K}[x]$, temos: $f(T) = O$ se, e somente se, existe $q \in \mathbb{K}[x]$ tal que $f = p \cdot q$.

Definição 2.1

Seja V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{K}$ e seja $T : V \longrightarrow V$ um operador linear. O **polinômio minimal de T** , designado p_m , é o único gerador unitário do ideal dos polinômios com coeficientes em $\mathbb{K}$ que anulam T .

Observação 2.1

Nenhum polinômio sobre $\mathbb{K}$ que anule T tem grau menor que o de p_m .

Definição 2.2

Se A é uma matriz $n \times n$ com entradas no corpo $\mathbb{K}$, definimos o **polinômio minimal de A** como sendo o único gerador unitário do ideal formado pelos polinômios de $\mathbb{K}[x]$ que anulam A .

Observação 2.2

Se o operador T é representado, em relação a alguma base ordenada $\mathcal{B}$ de V , pela matriz $A = [T]_{\mathcal{B}}$, então T e A têm o mesmo polinômio minimal.

De fato, se $f \in \mathbb{K}[x]$, temos que

$$f(T) = O \iff [f(T)]_{\mathcal{B}} = \mathbf{0} \iff f([T]_{\mathcal{B}}) = O \iff f(A) = O.$$

Observação 2.3

Seja $\mathbb{F}$ um subcorpo do corpo $\mathbb{K}$ e suponhamos que A seja uma matriz $n \times n$ com entradas no corpo $\mathbb{F}$. Então, o polinômio minimal de A , quando considerada como uma matriz de $\mathbb{F}^{n \times n}$, é igual ao polinômio minimal de A , quando considerada como uma matriz de $\mathbb{K}^{n \times n}$.

Seja $p_m \in \mathbb{F}[x]$ o polinômio minimal de A , considerada como uma matriz de $\mathbb{F}^{n \times n}$.

Como $p_m \in \mathbb{K}[x]$ e $p_m(A) = O$, temos que o polinômio minimal $\overline{p_m} \in \mathbb{K}[x]$ de A considerada como uma matriz de $\mathbb{K}^{n \times n}$, tem grau $\leq \text{grau}(p_m)$.

Seja $\overline{p_m} = x^k + a_{k-1}x^{k-1} + \dots + a_1x + a_0$. Então,

$$A^k + a_{k-1}A^{k-1} + \dots + a_1A + a_0I = O.$$

Ou seja, o vetor $(a_{k-1}, \dots, a_0) \in \mathbb{K}^k$ é solução de um sistema de n^2 equações lineares de k incógnitas da forma $BX = Y$, onde $B \in \mathbb{F}^{n^2 \times k}$ e $Y \in \mathbb{F}^{n^2 \times 1}$.

Então, pelo que foi provado anteriormente, o sistema $BX = Y$ possui uma solução $(b_{k-1}, \dots, b_0) \in \mathbb{F}^k$. Ou seja, o polinômio

$$q = x^k + b_{k-1}x^{k-1} + \dots + b_1x + b_0 \in \mathbb{F}[x]$$

anula a matriz A . Logo, $\text{grau}(q) = \text{grau}(\overline{p_m}) \geq \text{grau}(p_m)$.

Então, $\text{grau}(p_m) = \text{grau}(\overline{p_m})$.

Além disso, como $\overline{p_m} \in \mathbb{K}[x]$ é o gerador unitário do ideal de $\mathbb{K}[x]$ dos polinômios que anulam A , $p_m \in \mathbb{K}[x]$, $p_m(A) = O$ e p_m é unitário, temos que $p_m = \overline{p_m}$.

Teorema 2.1

Seja T um operador linear sobre um espaço vetorial V de dimensão n (ou, seja A uma matriz $n \times n$). Os polinômios característico e minimal de T (respectivamente, de A) possuem as mesmas raízes, a menos de multiplicidade.

Prova.

• Seja p_m o polinômio minimal de T e seja $c \in \mathbb{K}$ tal que $p_m(c) = 0$. Então, existe $q \in \mathbb{K}[x]$ tal que $p_m = (x - c)q$.

Como $\text{grau}(q) < \text{grau}(p_m)$, temos, pela definição de polinômio minimal, que $q(T) \neq O$. Existe, portanto, $v \in V$ tal que $q(T)(v) \neq 0$.

Seja $w = q(T)(v)$. Então,

$$0 = p_m(T)(v) = (T - cI)q(T)(v) = (T - cI)(w).$$

Ou seja, c é autovalor de T .

• Suponhamos, agora, que c é um autovalor de T e que $T(v) = cv$, $v \in V - \{0\}$.

Então, pelo lema 1.3,

$$p_m(T)(v) = p_m(c)v.$$

Como $p_m(T) = O$ e $v \neq 0$, temos que $p_m(c) = 0$. ■

Observação 2.4

Seja T um operador linear diagonalizável e sejam $c_1, \dots, c_k$ os autovalores distintos de T . Então

$$p = (x - c_1) \dots (x - c_k)$$

é o polinômio minimal de T .

De fato, se v é um autovetor de T , então, algum dos operadores $T - c_1I, \dots, T - c_kI$ leva v em 0 . Portanto,

$$p(T) = (T - c_1I) \dots (T - c_kI)(v) = 0,$$

para todo autovetor v . Como V possui uma base formada por autovetores, temos que $p(T) = O$.

Além disso, pelo teorema anterior, $c_1, \dots, c_k$ são as raízes do polinômio minimal. Logo, $p = (x - c_1) \dots (x - c_k)$ é o polinômio minimal do operador diagonalizável T .

Provamos, assim, que *se T é um operador linear diagonalizável, o polinômio minimal de T é um produto de fatores lineares distintos*. Como veremos mais tarde, essa propriedade caracteriza os operadores diagonalizáveis.

Exemplo 2.1

Seja T o operador linear sobre $\mathbb{R}^3$ representado em relação à base canônica, pela matriz

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

No exemplo 1.5 vimos que o operador T é diagonalizável e que o seu polinômio característico é $(x - 2)^2(x - 1)$. Logo, pela observação anterior, $p_m = (x - 2)(x - 1)$ é o polinômio minimal de T . $\square$

Exemplo 2.2

Seja T o operador linear sobre $\mathbb{R}^3$ representado, em relação à base canônica, pela matriz

$$A = \begin{pmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{pmatrix}$$

Vimos, no exemplo 1.2, que o operador T não é diagonalizável e que $(x - 1)(x - 2)^2$ é o seu polinômio característico.

Sabemos, pelo teorema anterior, que 1 e 2 são as raízes do polinômio minimal p_m de T . Então, $p = (x - 1)^k(x - 2)^\ell$, para alguns $k \geq 1$ e $\ell \geq 1$.

Como

$$(A - I)(A - 2I) = \begin{pmatrix} 2 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & -1 \\ 2 & 2 & -2 \end{pmatrix} = \begin{pmatrix} 2 & 0 & -1 \\ 2 & 0 & -1 \\ 4 & 0 & -2 \end{pmatrix} \quad (I)$$

não é a matriz nula, temos que $(x - 1)(x - 2)$ não é o polinômio minimal de T .

Assim, obtemos que o grau do polinômio minimal de T é, pelo menos, três

e, então, os possíveis candidatos a serem o polinômio minimal de T são $(x-1)^2(x-2)$ e $(x-1)(x-2)^2$.

Por (I), temos

$$(A - I)(A - 2I)^2 = \begin{pmatrix} 2 & 0 & -1 \\ 2 & 0 & -1 \\ 4 & 0 & -2 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & -1 \\ 2 & 2 & -2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Logo, o polinômio minimal de T é o seu polinômio característico. $\square$

Exemplo 2.3

Seja T o operador linear sobre $\mathbb{R}^2$ representado, em relação à base canônica, pela matriz

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

Como vimos no exemplo 1.1, o polinômio característico de T é $x^2 + 1$, que não possui raízes reais.

Considerando A como uma matriz 2×2 complexa, temos que $(x+i)(x-i)$ é o seu polinômio característico, com raízes i e $-i$.

Logo, i e $-i$ são, também, raízes do polinômio minimal de A , considerada tanto como uma matriz real quanto como uma matriz complexa.

Sendo

$$A^2 + I = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

temos que $x^2 + 1$ é o polinômio minimal de A , considerada como uma matriz real.

Portanto, $x^2 + 1$ é o polinômio minimal do operador T . $\square$

Teorema 2.2

Seja V um espaço vetorial de dimensão finita n sobre o corpo $\mathbb{K}$ e seja T um operador linear sobre V .

Se $p_c \in \mathbb{K}[x]$ é o polinômio característico de T , então $p_c(T) = O$.

Em outras palavras, *o polinômio minimal divide o polinômio característico* de T .

Prova.

Seja $L = \{q(T) \mid q \in \mathbb{K}[x]\}$.

Pelo visto nas seções anteriores, L é uma álgebra linear comutativa com elemento unidade sobre o corpo $\mathbb{K}$. Em particular, L é um anel comutativo com elemento unidade.

Seja $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base de V e seja A a matriz que representa T em relação à base $\mathcal{B}$. Então,

$$T(v_i) = \sum_{j=1}^n A_{ji} v_j, \quad 1 \leq i \leq n.$$

Essas equações podem ser escritas como

$$\sum_{j=1}^n (\delta_{ij} T - A_{ji} I) v_j = \mathbf{0}, \quad 1 \leq i \leq n.$$

Seja B a matriz $n \times n$ sobre L com entradas

$$B_{ij} = \delta_{ij} T - A_{ji} I.$$

Quando $n = 2$, temos

$$B = \begin{pmatrix} T - A_{11}I & -A_{21}I \\ -A_{12}I & T - A_{22}I \end{pmatrix},$$

e

$$\begin{aligned} \det(B) &= (T - A_{11}I)(T - A_{22}I) - A_{12}A_{21}I \\ &= T^2 - (A_{11} + A_{22})T + (A_{11}A_{22} - A_{12}A_{21})I \\ &= p_c(T), \end{aligned}$$

onde

$$p_c = x^2 - (\text{Tr } A)x + \det(A)$$

é o polinômio característico de T .

Para $n > 2$, temos também que

$$\det(B) = p_c(T),$$

pois o polinômio característico de T , p_c , é o determinante da matriz $xI - A$, cujas entradas são os polinômios

$$(xI - A)_{ij} = \delta_{ij}x - A_{ij}.$$

Como $\det(xI - A) = \det((xI - A)^t)$, temos, também, que p_c é o determinante da matriz $xI - A^t$, cujas entradas são os polinômios

$$(xI - A^t)_{ij} = \delta_{ij}x - A_{ji}.$$

Pela definição de B , temos que

$$\sum_{j=1}^n B_{ij}v_j = \mathbf{0}, \quad 1 \leq i \leq n. \quad (I)$$

Seja $\tilde{B} = \text{Adj } B$. Então, por (I), temos que

$$\sum_{j=1}^n \tilde{B}_{ki}B_{ij}v_j = \mathbf{0}, \quad 1 \leq i, k \leq n.$$

Somando em relação a i , temos

$$\sum_{i=1}^n \sum_{j=1}^n \tilde{B}_{ki}B_{ij}v_j = \mathbf{0} \implies \sum_{j=1}^n \left(\sum_{i=1}^n \tilde{B}_{ki}B_{ij} \right) v_j = \mathbf{0}. \quad (II)$$

Como $\tilde{B}B = \det(B)I$, temos que

$$\sum_{i=1}^n \tilde{B}_{ki}B_{ij} = \delta_{kj} \det(B).$$

Logo, por (II), temos

$$\sum_{j=1}^n \delta_{kj} \det(B)v_j = \mathbf{0} \implies \det(B)v_k = \mathbf{0}, \quad 1 \leq k \leq n.$$

Assim, provamos que $p_c(T) = \det(B) = 0$, ou seja, que o polinômio característico de T anula T . ■

Observação 2.5

Como o polinômio minimal p_m divide o polinômio característico p_c e os dois polinômios possuem as mesmas raízes, temos que, se p_c se fatora como

$$p_c = (x - c_1)^{d_1} \dots (x - c_k)^{d_k},$$

onde $c_1, \dots, c_k$ são as raízes distintas e $d_j \geq 1$ para todo j , então

$$p_m = (x - c_1)^{r_1} \dots (x - c_k)^{r_k},$$

com $1 \leq r_j \leq d_j$, para todo $j = 1, \dots, k$.

3. Subespaços Invariantes

Definição 3.1

Seja V um espaço vetorial sobre o corpo $\mathbb{K}$ e T um operador linear sobre V . Dizemos que um subespaço W de V é *invariante por T* , se para todo vetor $v \in W$, o vetor $T(v)$ pertence, também, a W .

Ou seja, W é um subespaço invariante por T se $T(W) \subset W$.

Exemplo 3.1

Seja $T : V \rightarrow V$ um operador linear. Então, os seguintes subespaços de V são invariantes por T :

- $W = V$;
- $W = \{\mathbf{0}\}$;
- $W = T(V) = \text{Im}(T)$ (a imagem de T);
- $W = T^{-1}(\mathbf{0}) = \text{Ker}(T)$ (o núcleo de T). $\square$

Exemplo 3.2

Seja $\mathbb{K}$ um corpo e seja D o operador derivação sobre o espaço $\mathbb{K}[x]$ dos polinômios com coeficientes em $\mathbb{K}$.

Seja n um inteiro positivo e seja W o subespaço de $\mathbb{K}[x]$ formado pelos polinômios de grau $\leq n$.

Então W é invariante por D . $\square$

Exemplo 3.3

Sejam $T, U : V \rightarrow V$ operadores lineares sobre V que comutam, isto é $TU = UT$.

Então, $W = U(V)$, a imagem de U , e $N = U^{-1}(\mathbf{0})$, o núcleo de U , são subespaços invariantes por T .

De fato, se $v \in W$, existe $w \in V$ tal que $U(w) = v$. Logo,

$$T(v) = T(U(w)) = U(T(w)) \in W.$$

Analogamente, se $v \in N$, $U(v) = \mathbf{0}$. Logo,

$$U(T(v)) = T(U(v)) = T(\mathbf{0}) = \mathbf{0},$$

ou seja, $T(v) \in N$. $\square$

Um tipo particular de operador que comuta com T é um operador $U = g(T)$, onde g é um polinômio.

Por exemplo, $U = T - cI$, onde c é um autovalor de T .

Nesse caso, o núcleo de U , que é *o espaço dos autovetores associados ao autovalor c , é invariante por T*

Exemplo 3.4

Seja T o operador linear sobre $\mathbb{R}^2$ que é representado em relação à base canônica pela matriz

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

Então, os únicos subespaços de $\mathbb{R}^2$ que são invariantes por T são $\mathbb{R}^2$ e o subespaço nulo.

De fato, qualquer outro subespaço invariante W teria dimensão 1, ou seja, W seria gerado por um vetor v não-nulo.

Sendo W invariante por T , teríamos $T(v) = \lambda v$, ou seja T teria um autovalor em $\mathbb{R}$, o que é um absurdo, pois já verificamos, no exemplo 1.1, que T não possui autovalores reais. $\square$

Observação 3.1

Quando o subespaço W é invariante por T , T induz um operador linear T_W sobre o espaço W , definido por $T_W(v) = T(v)$, para todo $v \in W$.

Suponhamos que V tem dimensão finita e que $\mathcal{B} = \{v_1, \dots, v_n\}$ é uma base ordenada de V tal que $\mathcal{B}' = \{v_1, \dots, v_r\}$ seja uma base ordenada de W , sendo $r = \dim W$.

Seja $A = [T]_{\mathcal{B}}$. Então,

$$T(v_j) = \sum_{i=1}^n A_{ij} v_i, \quad 1 \leq j \leq n.$$

Como W é invariante por T , temos que $A_{ij} = 0$, para todos $1 \leq j \leq r$ e $i \geq r+1$. Assim,

$$A = \begin{pmatrix} B & C \\ O & D \end{pmatrix},$$

onde $B = [T_W]_{\mathcal{B}'}$, é uma matriz $r \times r$, C é uma matriz $r \times (n-r)$ e D é uma

matriz $(n - r) \times (n - r)$.

Lema 3.1

Seja W um subespaço invariante por T . Então, o polinômio característico do operador T_W divide o polinômio característico de T , e o polinômio minimal de T_W divide o polinômio minimal de T .

Prova.

Seja $A = \begin{pmatrix} B & C \\ O & D \end{pmatrix}$, onde $A = [T]_B$ e $B = [T_W]_{B'}$.

Por causa da forma em blocos da matriz A , temos

$$\det(xI - A) = \det(xI - B) \det(xI - D).$$

Ou seja, o polinômio característico $\det(xI - B)$ de T_W divide o polinômio característico $\det(xI - A)$ de T .

Como a k -ésima potência da matriz A tem a forma em blocos

$$A^k = \begin{pmatrix} B^k & C_k \\ O & D^k \end{pmatrix},$$

onde C_k é uma matriz $r \times (n - r)$, temos que qualquer polinômio que anula A , também anula B .

Assim, o polinômio minimal de B divide o polinômio minimal de A . ■

Observação 3.2

Seja T um operador linear sobre um espaço vetorial V de dimensão finita n . Sejam $c_1, \dots, c_k$ os autovalores distintos de T e seja W_i , $i = 1, \dots, k$ o espaço dos autovetores associados ao autovalor c_i .

Seja $W = W_1 + \dots + W_k$ o subespaço gerado por todos os autovetores de T .

Se $\mathcal{B}_i$, $i = 1, \dots, k$ é uma base ordenada de W_i , já sabemos, pelo lema 1.4, que $\mathcal{B}' = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ é uma base ordenada de W e que $\dim W = \dim W_1 + \dots + \dim W_k$.

Então, se $\mathcal{B}' = \{v_1, \dots, v_r\}$, onde $r = \dim W$, temos que

$$T(v_i) = t_i v_i, \quad i = 1, \dots, r,$$

onde $(t_1, \dots, t_r) = (c_1, \dots, c_1, \dots, c_k, \dots, c_k)$ com cada c_i repetido $\dim W_i$ vezes.

O subespaço W é invariante por T , pois todo $v \in W$ pode ser escrito na forma

$$v = \lambda_1 v_1 + \dots + \lambda_r v_r,$$

logo,

$$T(v) = \lambda_1 t_1 v_1 + \dots + \lambda_r t_r v_r.$$

Sejam $v_{r+1}, \dots, v_n$ vetores de V tais que $\mathcal{B} = \{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ seja uma base de V .

A matriz de T em relação à base $\mathcal{B}$ é da forma em blocos

$$[T]_{\mathcal{B}} = \begin{pmatrix} [T_W]_{\mathcal{B}'} & C \\ O & D \end{pmatrix},$$

onde

$$[T_W]_{\mathcal{B}'} = \begin{pmatrix} t_1 & 0 & \dots & 0 \\ 0 & t_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & t_r \end{pmatrix}.$$

é uma matriz $r \times r$ diagonal.

Assim, $g = (x - c_1)^{\ell_1} \dots (x - c_k)^{\ell_k}$, com $\ell_i = \dim W_i$, é o polinômio característico de T_W .

Como g divide o polinômio característico p_c de T , temos que **a multiplicidade de c_i como raiz de p_c é no mínimo $\dim W_i$.**

Definição 3.2

Seja V um espaço vetorial sobre o corpo $\mathbb{K}$.

Seja W um subespaço de V invariante pelo operador $T : V \rightarrow V$ e seja v um vetor do espaço V .

O **T -condutor de v em W** é o conjunto $S_T(v; W)$ formado de todos os polinômios $g \in \mathbb{K}[x]$ tais que $g(T)(v) \in W$.

Em particular, se $W = \{0\}$, o T -condutor de v é denominado o **T -anulador de v .**

Observação 3.3

Como o operador T permanecerá fixo durante a maioria das discussões, abandonaremos o índice T e escreveremos $S(v; W)$ em vez de $S_T(v; W)$, e

tal conjunto será denominado apenas *condutor de v em W* .

Lema 3.2

Se W é um subespaço invariante por T , então W é invariante por qualquer polinômio em T .

Assim, para todo $v \in V$, o condutor $S(v; W)$ é um ideal na álgebra dos polinômios $\mathbb{K}[x]$.

Prova.

Se $w \in W$, então $T(w) \in W$. Como $T^n(w) = T(T^{n-1}(w))$, podemos provar, por indução, que $T^n(w) \in W$, para todo $n \in \mathbb{N}$.

Assim, $a_0w + a_1T(w) + \dots + a_nT^n(w) \in W$ quaisquer que sejam os escalares $a_0, a_1, \dots, a_n \in \mathbb{K}$, ou seja $p(T)(w) \in W$ para todo $p \in \mathbb{K}[x]$.

O condutor $S(v; W)$ é um subespaço de $\mathbb{K}[x]$, pois se $p(T)(v) \in W$ e $g(T)(w) \in W$, então

$$(cp + g)(T)(v) = cp(T)(v) + g(T)(v) \in W.$$

Como W é invariante por T , temos que se g é um polinômio em $S(v; W)$, isto é, se $g(T)(v) \in W$ e se $f \in \mathbb{K}[x]$ é um polinômio arbitrário, então

$$(fg)(T)(v) = (f(T) \circ g(T))(v) = f(T)(g(T)(v)) \in W,$$

pois W é invariante por qualquer polinômio em T . ■

Observação 3.4

O único gerador unitário do ideal $S(v; W)$ é também denominado *T -condutor de v em W* (ou *T -anulador de v* , no caso em que $W = \{0\}$).

O T -condutor de v em W é o polinômio unitário g de menor grau tal que $g(T)(v) \in W$.

Observe que o polinômio minimal pertence ao ideal $S(v; W)$. Logo, todo T -condutor (ou T -anulador) divide o polinômio minimal.

Definição 3.3

Um operador linear T sobre um espaço vetorial de dimensão finita é *triangulável* se existe uma base ordenada $\mathcal{B}$ de V tal que a matriz $[T]_{\mathcal{B}}$ é triangular.

Lema 3.3

Seja V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{K}$. Seja T um operador linear sobre V tal que o polinômio minimal de T seja um produto de fatores lineares

$$p_m = (x - c_1)^{r_1} \dots (x - c_k)^{r_k}, \quad c_i \in \mathbb{K}.$$

Seja W um subespaço próprio de V (i.e. $W \neq V$) invariante por T . Então existe um vetor v em V tal que:

(a) $v \notin W$;

(b) $(T - cI)(v) \in W$, para algum autovalor c do operador T .

Prova.

Seja w um vetor arbitrário de V que não pertence a W e seja g o T -condutor de w em W .

Como w não está em W , o polinômio g não é constante e divide o polinômio minimal p_m . Então,

$$g = (x - c_1)^{s_1} \dots (x - c_k)^{s_k},$$

onde pelo menos um dos inteiros $s_1, \dots, s_k$ é positivo.

Seja j tal que $s_j > 0$. Então $(x - c_j)$ divide g :

$$g = (x - c_j) h.$$

Pela definição de g , o vetor $v = h(T)(w)$ não pertence a W , pois $\text{grau}(h) < \text{grau}(g)$.

Mas,

$$(T - c_j I)(v) = (T - c_j I)(h(T)(w)) = ((T - c_j I) h(T))(w) = g(T)(w)$$

pertence a W . ■

Teorema 3.1

Seja V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{K}$ e seja T um operador linear sobre V . Então T é triangulável se, e somente se, o polinômio minimal de T é um produto de polinômios lineares sobre K .

Prova.

Suponhamos que o polinômio minimal de T é um produto de fatores lineares:

$$p_m = (x - c_1)^{r_1} \dots (x - c_k)^{r_k}.$$

Aplicaremos o lema anterior, para obter uma base $\mathcal{B} = \{v_1, \dots, v_n\}$ de V tal que

$$T(v_j) = a_{1j}v_1 + \dots + a_{jj}v_j, \quad 1 \leq j \leq n,$$

ou seja $T(v_j)$ pertence ao subespaço gerado por $v_1, \dots, v_j$.

Assim,

$$[T]_{\mathcal{B}} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & a_{nn} \end{pmatrix}.$$

é uma matriz triangular superior.

Começamos aplicando o lema anterior ao subespaço $W_1 = \{0\}$. Então, existe $v_1 \neq 0$ e $a_{11} \in \mathbb{K}$ tal que $(T - a_{11}I)(v_1) \in W_1$, ou seja, $T(v_1) = a_{11}v_1$.

Seja W_2 o subespaço gerado por v_1 . Esse subespaço é invariante por T , pois $T(v_1) = a_{11}v_1 \in W_2$.

Pelo lema anterior, existe $v_2 \notin W_1$ e $a_{22} \in \mathbb{K}$, tais que $(T - a_{22}I)(v_2) \in W_2$, ou seja, existe $a_{12} \in \mathbb{K}$, tal que $T(v_2) = a_{12}v_1 + a_{22}v_2$.

Seja, agora, W_3 o subespaço gerado por v_1 e v_2 que é invariante por T , pois $T(v_1) = a_{11}v_1 \in W_3$ e $T(v_2) = a_{12}v_1 + a_{22}v_2 \in W_3$. Pelo lema anterior, existe $v_3 \notin W_2$ e $a_{33} \in \mathbb{K}$, tais que $(T - a_{33}I)(v_3) \in W_3$, ou seja, existem $a_{13}, a_{23} \in \mathbb{K}$ tais que

$$T(v_3) = a_{13}v_1 + a_{23}v_2 + a_{33}v_3.$$

Prosseguindo desta maneira, obtemos vetores $v_1, \dots, v_n$ em V tais que

$$T(v_j) = a_{1j}v_1 + \dots + a_{jj}v_j, \quad j = 1, \dots, n.$$

Observe, também, que $\{v_1, \dots, v_n\}$ é um conjunto LI, e portanto, uma base de V , pois, em cada passo j , o vetor v_j não pertence ao espaço gerado por $v_1, \dots, v_{j-1}$.

Suponhamos, agora, que T é triangulável, ou seja, que existe uma base $\mathcal{B}$ de V tal que

$$[T]_{\mathcal{B}} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Então, $xI - [T]_{\mathcal{B}}$ é também uma matriz triangular, cujo determinante, o polinômio característico p_c , é o produto das entradas da diagonal, ou seja,

$$p_c = (x - a_{11})(x - a_{22}) \cdots (x - a_{nn}).$$

Os elementos da diagonal $a_{11}, \dots, a_{nn}$ são os autovalores c_i e aparecem repetidos d_i vezes, ou seja

$$p_c = (x - c_1)^{d_1} \cdots (x - c_k)^{d_k}.$$

Como o polinômio minimal p_m divide p_c , temos que p_m é da forma

$$p_m = (x - c_1)^{r_1} \cdots (x - c_k)^{r_k},$$

onde $1 \leq r_j \leq d_j$ para todo $j = 1, \dots, k$. ■

Corolário 3.1

Seja $\mathbb{K}$ um corpo algebricamente fechado. Então, toda matriz $n \times n$ sobre $\mathbb{K}$ é semelhante sobre $\mathbb{K}$ a uma matriz triangular.

Prova.

Seja T o operador sobre $\mathbb{K}^n$ representado em relação à base canônica $\mathcal{B}$ por uma matriz A .

Como o corpo $\mathbb{K}$ é algebricamente fechado, o polinômio minimal de T é um produto de polinômios lineares sobre $\mathbb{K}$. Então, pelo teorema anterior, existe uma base $\mathcal{B}'$ de $\mathbb{K}^n$ tal que $[T]_{\mathcal{B}'}$ é triangular.

Como $A = [T]_{\mathcal{B}}$ e $[T]_{\mathcal{B}'}$ são semelhantes, temos que A é semelhante a uma matriz triangular. ■

Teorema 3.2

Seja V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{K}$ e seja T um operador linear sobre V . Então, T é diagonalizável se, e somente se, o polinômio minimal de T tem a forma

$$p_m = (x - c_1) \cdots (x - c_k),$$

onde $c_1, \dots, c_k$ são elementos distintos de $\mathbb{K}$.

Prova.

Já provamos que se T é diagonalizável então o polinômio minimal de T é um produto de fatores lineares distintos.

Para demonstrar a recíproca, suponhamos que o subespaço W de V , formado por todos os autovetores de T , seja próprio, isto é, $W \neq V$.

Pelo lema anterior, existe um vetor $v \notin W$ e um autovalor c_j de T tais que o vetor

$$w = (T - c_j I)(v) \in W.$$

Como $w \in W$, existem autovetores $w_1, \dots, w_k$ tais que

$$w = w_1 + \dots + w_k,$$

onde $T(w_i) = c_i w_i$, $1 \leq i \leq k$.

Portanto,

$$h(T)(w) = h(c_1)w_1 + \dots + h(c_k)w_k \in W,$$

para todo polinômio $h \in \mathbb{K}[x]$.

Sendo $p_m = (x - c_j)q$ e $(q - q(c_j))(c_j) = 0$, existe um polinômio $h \in \mathbb{K}[x]$ tal que

$$q - q(c_j) = (x - c_j)h.$$

Temos

$$(q(T) - q(c_j)I)(v) = q(T)(v) - q(c_j)(v) = h(T)(T - c_j I)(v) = h(T)(w) \in W.$$

Como

$$0 = p_m(T)(v) = (T - c_j I)q(T)(v),$$

temos que $q(T)(v)$ é um autovetor de T e, portanto, $q(T)(v) \in W$.

Assim,

$$q(c_j)(v) = q(T)(v) - h(T)(w) \in W.$$

Como $v \notin W$, temos que $q(c_j) = 0$, que é uma contradição, pois, por hipótese, todas as raízes de p_m têm multiplicidade igual a 1. ■

Observação 3.5

Seja T um operador linear sobre um espaço vetorial V de dimensão finita. Se seu polinômio característico

$$p_c = (x - c_1)^{d_1} \dots (x - c_k)^{d_k},$$

pode ser decomposto como um produto de polinômios lineares, temos

dois métodos para determinar se T é ou não diagonalizável. Um método consiste em verificar se existem, para cada i , d_i autovetores independentes associados ao autovalor c_i . O outro método consiste em verificar se $(T - c_1 I) \cdot (T - c_k I)$ é ou não o operador nulo.

Observação 3.6

Seja A uma matriz triangular

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Então, o polinômio característico de A é dado por:

$$p_c = \det(xI - A) = \det \begin{pmatrix} x - a_{11} & -a_{12} & -a_{13} & \cdots & -a_{1n} \\ 0 & x - a_{22} & -a_{23} & \cdots & -a_{2n} \\ 0 & 0 & x - a_{33} & \cdots & -a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & x - a_{nn} \end{pmatrix}.$$

ou seja, $p_c = (x - a_{11}) \cdots (x - a_{nn})$.

Como,

$$Ae_j = \sum_{k=1}^j a_{kj} e_k,$$

temos que

$$p_c(A)(e_j) = (A - a_{11}I)(A - a_{22}I) \cdots (A - a_{nn}I)(e_j) = \mathbf{0},$$

para todo $j = 1, \dots, n$.

De fato, como $(A - a_{11}I)(e_1) = \mathbf{0}$, temos que $p_c(A)(e_1) = \mathbf{0}$, pois as matrizes $(A - a_{ii}I)$, $i = 1, \dots, n$, comutam entre si.

Supondo que $(A - a_{11}I)(A - a_{22}I) \cdots (A - a_{jj}I)(e_i) = \mathbf{0}$ para todo $i = 1, \dots, j$, temos que

$$(A - a_{11}I)(A - a_{22}I) \cdots (A - a_{jj}I)(A - a_{(j+1)(j+1)}I)(e_i) = \mathbf{0},$$

para todo $i = 1, \dots, j + 1$.

De fato:

- $(A - a_{11}I) \dots (A - a_{jj}I)(A - a_{(j+1)(j+1)}I)(e_i)$
 $= (A - a_{(j+1)(j+1)}I) ((A - a_{11}I) \dots (A - a_{jj}I)(e_i))$
 $= (A - a_{(j+1)(j+1)}I)(\mathbf{0}) = \mathbf{0}, \quad \text{para todo } i = 1, \dots, j.$
- $(A - a_{11}I) \dots (A - a_{jj}I)(A - a_{(j+1)(j+1)}I)(e_{j+1}) = \mathbf{0}$
 pois $(A - a_{(j+1)(j+1)}I)(e_{j+1}) = \sum_{k=1}^j a_{k(j+1)} e_k$ e
 $(A - a_{11}I) \dots (A - a_{jj}I)(e_k) = \mathbf{0}, \quad \text{para todo } k = 1, \dots, j.$

Provamos, assim, que

$$p_c(A)(e_j) = (A - a_{11}I) \dots (A - a_{nn}I)(e_j) = \mathbf{0},$$

para todo $j = 1, \dots, n$, ou seja, $p_c(A) = O$.

Observação 3.7

Como todo operador linear T sobre um espaço vetorial V de dimensão finita sobre um corpo $\mathbb{K}$ algebricamente fechado possui uma base $\mathcal{B}$ tal que $[T]_{\mathcal{B}}$ é triangular, temos que $p_c(T) = O$, onde p_c é o polinômio característico de T .

Assim, obtemos uma outra demonstração do Teorema de Cayley-Hamilton para operadores sobre um espaço vetorial sobre um corpo algebricamente fechado.

Mas, como todo corpo $\mathbb{K}$ é um subcorpo de um corpo algebricamente fechado $\mathbb{K}'$, temos que se A é uma matriz $n \times n$ sobre $\mathbb{K}$, então A é semelhante em $\mathbb{K}'$ a uma matriz triangular A' . Além disso, já provamos que matrizes semelhantes possuem o mesmo polinômio característico e que uma matriz triangular anula o seu polinômio característico.

Logo, como $A = B^{-1}A'B$, para alguma matriz B invertível, e $f(A) = f(A')$ para todo polinômio f , temos que

$$p_c(A) = p_c(A') = O,$$

onde p_c é o polinômio característico de A .

Provamos, assim, que se A é uma matriz $n \times n$ sobre um corpo arbitrário, então $p_c(A) = O$, onde p_c é o polinômio característico de A . Ou seja, se T é um operador linear sobre um espaço vetorial V de dimensão finita sobre um corpo $\mathbb{K}$ arbitrário, então $p_c(T) = O$, onde p_c é o polinômio característico de T .

Acabamos de apresentar, portanto, uma outra demonstração do Teorema de Cayley-Hamilton, mas, para isso, usamos o fato que todo corpo é um subcorpo de um corpo algebricamente fechado.

4. Triangulação Simultânea e Diagonalização Simultânea

Seja V um espaço vetorial de dimensão finita e seja $\mathcal{F}$ uma família de operadores lineares sobre V .

Um subespaço S é *invariante por $\mathcal{F}$* se W é invariante por cada operador em $\mathcal{F}$.

Lema 4.1

Seja $\mathcal{F}$ uma família comutativa de operadores lineares trianguláveis. Seja W um subespaço próprio de V , invariante por $\mathcal{F}$. Então existe um vetor $v \in V$ tal que:

- (a) $v \notin W$;
- (b) para cada T em $\mathcal{F}$, o vetor $T(v)$ está no subespaço gerado por v e W .

Prova.

Seja $\{T_1, \dots, T_r\}$ um subconjunto linearmente independente maximal de $\mathcal{F}$, isto é, uma base do subespaço gerado por $\mathcal{F}$. Se v é um vetor que satisfaz (b) para cada T_i , $i = 1, \dots, r$, então (b) é verdadeiro para todo operador que seja uma combinação linear de $T_1, \dots, T_r$.

Pelo lema 3.3, existe $v_1 \notin W$ e existe $c_1 \in \mathbb{K}$, tal que $(T - c_1 I)(v_1) \in W$, pois, como T_1 é triangulável, seu polinômio minimal é um produto de polinômios lineares.

Seja $V_1 = \{v \in V \mid (T_1 - c_1 I)(v) \in W\}$. Como $W \subset V_1$ e $v_1 \in V_1$, $v_1 \notin W$, temos que W é um subespaço próprio de V_1 .

Além disso, V_1 é invariante por $\mathcal{F}$. De fato, se $T \in \mathcal{F}$ e $v \in V_1$, então

$$(T_1 - c_1 I)(T(v)) = T(T_1 - c_1 I)(v) \in W,$$

pois $(T_1 - c_1 I)(v) \in W$ e W é invariante por T .

Logo, $T(v) \in V_1$, para todo $v \in V_1$ e para todo $T \in \mathcal{F}$.

Seja U_2 o operador linear sobre V_1 obtido pela restrição de T_2 ao subespaço V_1 .

Como o polinômio minimal de T_2 é um produto de fatores lineares e o polinômio minimal de U_2 divide o polinômio minimal de T_2 , temos que o polinômio minimal de U_2 é também um produto de fatores lineares.

Aplicando o lema 3.3 ao operador U_2 e ao subespaço próprio W de V_1 invariante por U_2 , obtemos um vetor $v_2 \in V_1$, tal que $v_2 \notin W$, e um escalar $c_2 \in \mathbb{K}$, tal que $(U_2 - c_2I)(v_2) = (T_2 - c_2I)(v_2) \in W$.

Logo, $v_2 \notin W$, $(T_1 - c_1I)(v_2) \in W$ e $(T_2 - c_2I)(v_2) \in W$, ou seja, $v_2 \notin W$, mas $T_1(v_2)$ e $T_2(v_2)$ pertencem ao subespaço gerado por v_2 e W .

Seja $V_2 = \{v \in V_1 \mid (T_2 - c_2I)(v) \in W\}$. Como $W \subset V_2$ e $v_2 \in V_2$, $v_2 \notin W$, temos que W é um subespaço próprio de V_2 .

Se $T \in \mathcal{F}$ e $v \in V_2$, temos que

$$(T_2 - c_2)(T(v)) = T(T_2 - c_2I)(v) \in W,$$

pois $(T_2 - c_2I)(v) \in W$ e W é invariante por T .

Logo, $T(v) \in V_2$ para todo $v \in V_2$ e para todo $T \in \mathcal{F}$.

Seja U_3 a restrição de T_3 ao subespaço V_2 . Como T_3 é triangulável, seu polinômio minimal é um produto de fatores lineares. Assim, o polinômio minimal de U_3 , que divide o polinômio minimal de T_3 é, também, um produto de fatores lineares.

Aplicando o lema 3.3 ao operador U_3 e ao subespaço W de V_2 invariante por U_3 , obtemos um vetor $v_3 \in V_2$, tal que $v_3 \notin W$, e um escalar $c_3 \in \mathbb{K}$, tal que $(U_3 - c_3I)(v_3) = (T_3 - c_3I)(v_3) \in W$. Logo, $v_3 \notin W$, $(T_1 - c_1I)(v_3) \in W$, pois $v_3 \in V_2 \subset V_1$, $(T_2 - c_2I)(v_3) \in W$, pois $v_3 \in V_2$, e $(T_3 - c_3I)(v_3) \in W$.

Prosseguindo desta maneira, chegaremos a um vetor $v = v_r \notin W$ tal que $(T_j - c_jI)(v) \in W$, para todo $j = 1, \dots, r$. Ou seja, $v \notin W$ e $T_j(v)$ pertence ao subespaço gerado por v e W , para todo $j = 1, \dots, r$. ■

Teorema 4.1

Seja V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{K}$. Seja $\mathcal{F}$ uma família comutativa de operadores lineares trianguláveis sobre V . Então, existe uma base ordenada $\mathcal{B}$ de V tal que $[T]_{\mathcal{B}}$ é uma matriz triangular superior para todo operador $T \in \mathcal{F}$.

Prova.

Devido ao lema anterior, a demonstração deste teorema é igual à demonstração do teorema 3.1, colocando $\mathcal{F}$ no lugar de T . ■

Corolário 4.1

Seja $\mathcal{F}$ uma família comutativa de matrizes $n \times n$ sobre um corpo $\mathbb{K}$ algebricamente fechado. Então, existe uma matriz P invertível, com entradas em $\mathbb{K}$, tal que $P^{-1}AP$ é triangular superior para toda matriz $A \in \mathcal{F}$.

Observação 4.1

A condição da família $\mathcal{F}$ ser comutativa é apenas suficiente, mas não necessária para a existência de uma base $\mathcal{B}$ tal que $[T]_{\mathcal{B}}$ é triangular superior para todo $T \in \mathcal{F}$, pois dois operadores trianguláveis, ou duas matrizes triangulares superiores, nem sempre comutam.

Teorema 4.2

Seja $\mathcal{F}$ uma família comutativa de operadores lineares diagonalizáveis sobre o espaço vetorial V de dimensão finita. Então existe uma base ordenada $\mathcal{B}$ de V tal que $[T]_{\mathcal{B}}$ é diagonal para todo operador $T \in \mathcal{F}$.

Prova.

Provaremos o teorema por indução sobre a dimensão de V .

Se $\dim V = 1$, nada há por demonstrar. Suponhamos que o teorema seja válido para espaços vetoriais de dimensão menor que n e suponhamos que $\dim V = n$.

Seja $T \in \mathcal{F}$ um operador que não seja múltiplo do operador identidade. Sejam $c_1, \dots, c_k$ os autovalores distintos de T , e seja W_i , $i = 1, \dots, k$, o núcleo de $(T - c_i I)$.

Para todo operador U que comuta com T , W_i é invariante por U , pois

$$TU(v) = UT(v) = U(c_i v) = c_i U(v),$$

para todo $v \in W_i$.

Seja $\mathcal{F}_i$ a família de operadores lineares sobre W_i , obtida pela restrição dos operadores em $\mathcal{F}$ ao subespaço invariante W_i .

Cada operador em $\mathcal{F}_i$ é diagonalizável, pois seu polinômio minimal divide o polinômio minimal do operador correspondente em $\mathcal{F}$, sendo, portanto,

um produto de fatores lineares distintos.

Como $\dim W_i < \dim V$, os operadores em $\mathcal{F}_i$ podem ser diagonalizados simultaneamente, ou seja, W_i possui uma base $\mathcal{B}_i$ formada por vetores que são autovetores de todos os operadores de $\mathcal{F}_i$.

Como T é diagonalizável, temos que $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ é uma base de V formada por autovetores de todos os operadores de $\mathcal{F}$, ou seja, $[L]_{\mathcal{B}}$ é diagonal para todo operador $L \in \mathcal{F}$ ■

Observação 4.2

No teorema anterior, a comutatividade da família $\mathcal{F}$ é uma condição não só suficiente mas também necessária para a diagonalização simultânea dos operadores em $\mathcal{F}$, pois duas matrizes diagonais sempre comutam.

5. Decomposição em Soma Direta

Definição 5.1

Seja V um espaço vetorial. Os subespaços $W_1, \dots, W_k$ de V são chamados *independentes*, se

$$v_1 + \dots + v_k = \mathbf{0},$$

com $v_i \in W_i$, implica que $v_i = \mathbf{0}$, para todo $i = 1, \dots, k$.

Se $W = W_1 + \dots + W_k$, então todo vetor $v \in W$ pode ser expresso como

$$v = v_1 + \dots + v_k, \quad v_i \in W_i.$$

Se, além disso, $W_1, \dots, W_k$ são independentes, todo vetor $v \in W$ pode ser expresso *de modo único* como uma tal soma.

De fato, se v se escreve, também, como $v = w_1 + \dots + w_k$, $w_i \in W_i$, temos que

$$v - v = (v_1 - w_1) + \dots + (v_k - w_k) = \mathbf{0},$$

com $v_i - w_i \in W_i$, $i = 1, \dots, k$. Logo, $v_i = w_i$ para todo $i = 1, \dots, k$, pois $W_1, \dots, W_k$ são independentes.

Definição 5.2

Sejam $W_1, \dots, W_k$ subespaços de um espaço vetorial V .

Dizemos que o subespaço W é a **soma direta** de $W_1, \dots, W_k$ e escrevemos

$$W = W_1 \oplus \dots \oplus W_k$$

se as seguintes condições são satisfeitas

- $W = W_1 + \dots + W_k$;
- $W_1, \dots, W_k$ são independentes.

Teorema 5.1

Seja V um espaço vetorial de dimensão finita e sejam $W_1, \dots, W_k$ subespaços de V .

Se $W = W_1 + \dots + W_k$, as seguintes afirmações são equivalentes:

- (a) $W = W_1 \oplus W_2 \oplus \dots \oplus W_k$;
- (b) $W_j \cap (W_1 + \dots + W_{j-1}) = \{\mathbf{0}\}$, $2 \leq j \leq k$;
- (c) Se $\mathcal{B}_i$ é uma base ordenada de W_i , $1 \leq i \leq k$, então $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ é uma base de W .
- (d) $\dim W = \sum_{i=1}^k \dim W_i$.

Prova.

(a) $\implies$ (b) Se $v \in W_j \cap (W_1 + \dots + W_{j-1})$, existem vetores $v_1, \dots, v_{j-1}$, $v_i \in W_i$, $i = 1, \dots, j-1$, tais que

$$v = v_1 + \dots + v_{j-1}.$$

Como $W_1, \dots, W_k$ são independentes e

$$v_1 + \dots + v_{j-1} + (-v) + \mathbf{0} + \dots + \mathbf{0} = \mathbf{0},$$

temos que $v_1 = v_2 = \dots = v_{j-1} = v = \mathbf{0}$.

(b) $\implies$ (c) Como $W = W_1 + \dots + W_k$ e $\mathcal{B}_i$ é uma base de W_i , $i = 1, \dots, k$, temos que $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ gera o subespaço W . Também, toda combinação linear nula de vetores de $\mathcal{B}$ tem a forma

$$v_1 + \dots + v_k = \mathbf{0},$$

onde cada v_i é uma combinação linear de vetores da base $\mathcal{B}_i$, $i = 1, \dots, k$.

Seja j o maior inteiro dentre $1, 2, \dots, k$, tal que $v_j \neq \mathbf{0}$. Então

$$\mathbf{0} = v_1 + \dots + v_j, \text{ com } v_j \neq \mathbf{0}.$$

Logo, $v_j = -v_1 - \dots - v_{j-1}$ é um vetor não nulo em $W_j \cap (W_1 + \dots + W_{j-1})$,

o que contradiz (b).

Assim, $v_i = \mathbf{0}$ para todo $i = 1, \dots, k$.

Como cada v_i é uma combinação linear de vetores da base $\mathcal{B}_i$, temos que essa combinação linear só pode ser a trivial. Logo, $\mathcal{B}$ é uma base de W .

(c) $\Rightarrow$ (d) é óbvio.

(d) $\Rightarrow$ (a) Seja $\mathcal{B}_i$ uma base de W_i , $i = 1, \dots, k$. Como $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$

gera o subespaço $W = W_1 + \dots + W_k$ e $\dim W = \sum_{i=1}^k \dim W_i = \sum_{i=1}^k \#\mathcal{B}_i$,

temos que $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ é uma base de W .

Então, se $v_1 + \dots + v_k = \mathbf{0}$, com $v_i \in W_i$, temos que existem escalares $\lambda_j^i \in \mathbb{K}$, $j = 1, \dots, k$, $i = 1, \dots, \dim W_i = n_i$, tais que

$$\lambda_1^1 v_1^1 + \lambda_2^1 v_2^1 + \dots + \lambda_{n_1}^1 v_{n_1}^1 + \dots + \lambda_1^k v_1^k + \dots + \lambda_{n_k}^k v_{n_k}^k = \mathbf{0},$$

onde $\mathcal{B}_i = \{v_1^i, \dots, v_{n_i}^i\}$ e $v_i = \sum_{j=1}^{n_i} \lambda_j^i v_j^i$.

Como $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ é uma base de W , temos que $\lambda_j^i = 0$, para todo $j = 1, \dots, k$ e todo $i = 1, \dots, n_j$, ou seja $v_j = \mathbf{0}$, para todo $j = 1, \dots, k$. Logo, $W_1, \dots, W_i$ são independentes, ou seja, $W = W_1 \oplus \dots \oplus W_k$. ■

Exemplo 5.1

Seja n um inteiro positivo e seja $\mathbb{K}^{n \times n}$ o espaço vetorial das matrizes $n \times n$ sobre um subcorpo $\mathbb{K}$ do corpo dos números complexos.

Seja W_1 o subespaço das matrizes *simétricas*, isto é, as matrizes A tais que $A^t = A$.

Seja W_2 o subespaço das matrizes *antisimétricas*, isto é, matrizes A tais que $A^t = -A$. Então,

$$\mathbb{K}^{n \times n} = W_1 \oplus W_2$$

De fato, se $A \in \mathbb{K}^{n \times n}$, temos

$$A = A_1 + A_2,$$

onde $A_1 = \frac{1}{2}(A + A^t) \in W_1$ e $A_2 = \frac{1}{2}(A - A^t) \in W_2$.

Se $A = B_1 + B_2$, onde $B_1 \in W_1$ e $B_2 \in W_2$, temos que

Ao lado estamos designando $\#\mathcal{B}_i$ o número de elementos do conjunto $\mathcal{B}_i$, isto é, a cardinalidade do conjunto $\mathcal{B}_i$.

A identidade ao lado diz que toda matriz se expressa de forma única como a soma de uma matriz simétrica com uma matriz antisimétrica.

$$A^t = B_1^t + B_2^t = B_1 - B_2.$$

Logo, $A + A^t = 2B_1$, ou seja $B_1 = \frac{1}{2}(A + A^t) = A_1$, e

$$B_2 = A - B_1 = A - \frac{1}{2}(A + A^t) = \frac{1}{2}(A - A^t) = A_2,$$

provando, assim, a unicidade da decomposição. $\square$

Observação 5.1

Seja T um operador linear sobre um espaço vetorial V de dimensão finita.

Sejam $c_1, \dots, c_k$ os autovalores distintos de T e seja W_i , $i = 1, \dots, k$, o espaço dos autovetores associado ao autovalor c_i .

Já provamos que $W_1, \dots, W_k$ são LI. Logo, $V = W_1 \oplus \dots \oplus W_k$.

Como consequência desse fato e do teorema 1.2, obtemos o seguinte resultado:

O operador $T : V \longrightarrow V$ é diagonalizável $\iff V = W_1 \oplus \dots \oplus W_k$

Corolário 5.1

Seja V um espaço vetorial de dimensão finita e sejam $W_1, \dots, W_k$ subespaços de V . Então:

(a) $V = W_1 \oplus \dots \oplus W_k \iff V = W_1 + \dots + W_k$ e $W_j \cap (W_1 + \dots + W_{j-1}) = \{0\}$, $\forall j = 2, \dots, k$.

(b) $V = W_1 \oplus \dots \oplus W_k \iff V = W_1 + \dots + W_k$ e $\dim(V) = \sum_{i=1}^k \dim(W_i)$.

(c) $V = W_1 \oplus \dots \oplus W_k \iff \dim(V) = \sum_{i=1}^k \dim(W_i)$ e $W_j \cap (W_1 + \dots + W_{j-1}) = \{0\}$, $\forall j = 2, \dots, k$.

Prova.

Pelo lema anterior, só falta provar que se $\dim(V) = \sum_{i=1}^k \dim(W_i)$ e

$W_j \cap (W_1 + \dots + W_{j-1}) = \{0\}$, $2 \leq j \leq k$, então $V = W_1 \oplus \dots \oplus W_k$.

De fato, ao demonstrarmos que (b) $\implies$ (c) no teorema 5.1, provamos que se $W_j \cap (W_1 + \dots + W_{j-1}) = \{0\}$, $2 \leq j \leq k$, então $B_1 \cup \dots \cup B_k$ é um conjunto

de vetores LI, onde $\mathcal{B}_i$ é uma base de W_i . Tal conjunto é uma base de V , pois possui $\sum_{i=1}^k \dim(W_i) = \dim(V)$ vetores.

Logo, $V = W_1 + \dots + W_k$ e, portanto, $V = W_1 \oplus \dots \oplus W_k$. ■

Proposição 5.1

Sejam V e W espaços vetoriais sobre o corpo $\mathbb{K}$. Então, o conjunto $V \times W = \{(v, w) \mid v \in V \text{ e } w \in W\}$ com as operações

$$(v, w) + (v', w') = (v + v', w + w')$$

$$\lambda(v, w) = (\lambda v, \lambda w)$$

é um espaço vetorial.

Prova.

Exercício. ■

Proposição 5.2

Se $\dim V = n$ e $\dim W = m$, então $\dim(V \times W) = m + n$.

Prova.

Sejam $\{v_1, \dots, v_n\}$ uma base de V e $\{w_1, \dots, w_m\}$ uma base de W . Então, o conjunto

$$\mathcal{B} = \{(v_1, \mathbf{0}), \dots, (v_n, \mathbf{0}), (\mathbf{0}, w_1), \dots, (\mathbf{0}, w_m)\},$$

é uma base de $V \times W$.

De fato, se $(v, w) \in V \times W$, existem escalares $\lambda_1, \dots, \lambda_n, \delta_1, \dots, \delta_m \in \mathbb{K}$ tais que

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n \quad \text{e} \quad w = \delta_1 w_1 + \dots + \delta_m w_m.$$

Logo,

$$(v, w) = \lambda_1(v_1, \mathbf{0}) + \dots + \lambda_n(v_n, \mathbf{0}) + \delta_1(\mathbf{0}, w_1) + \dots + \delta_m(\mathbf{0}, w_m),$$

ou seja, o conjunto $\mathcal{B}$ gera o espaço $V \times W$.

Além disso, $\mathcal{B}$ é um conjunto LI, pois se

$$\lambda_1(v_1, \mathbf{0}) + \dots + \lambda_n(v_n, \mathbf{0}) + \delta_1(\mathbf{0}, w_1) + \dots + \delta_m(\mathbf{0}, w_m) = (\mathbf{0}, \mathbf{0}),$$

temos que

$$(\lambda_1 v_1 + \dots + \lambda_n v_n, \delta_1 w_1 + \dots + \delta_m w_m) = (\mathbf{0}, \mathbf{0}).$$

Então, $\lambda_1 v_1 + \dots + \lambda_n v_n = \delta_1 w_1 + \dots + \delta_m w_m = \mathbf{0}$, o que implica $\lambda_1 = \dots = \lambda_n = \delta_1 = \dots = \delta_m = 0$. ■

Proposição 5.3

Sejam W_1 e W_2 subespaços de um espaço vetorial de dimensão finita.

Seja $B : W_1 \times W_2 \longrightarrow V$ definida por $B(v, w) = v + w$. Então B é linear, e B é um isomorfismo se, e só se, $V = W_1 \oplus W_2$.

Prova.

Sejam $(v_1, w_1), (v_2, w_2) \in W_1 \times W_2$ e $\lambda \in \mathbb{K}$. Então,

$$\begin{aligned} B(\lambda(v_1, w_1) + (v_2, w_2)) &= B(\lambda v_1 + v_2, \lambda w_1 + w_2) \\ &= \lambda v_1 + v_2 + \lambda w_1 + w_2 \\ &= \lambda(v_1 + w_1) + (v_2 + w_2) \\ &= \lambda B(v_1, w_1) + B(v_2, w_2), \end{aligned}$$

ou seja B é linear.

Provaremos, agora, que B é um isomorfismo se, e só se, $V = W_1 \oplus W_2$.

($\implies$) Temos que $\dim(V) = \dim(W_1 \times W_2) = \dim(W_1) + \dim(W_2)$. Além disso, se $v \in W_1 \cap W_2$, então, $(v, -v) \in W_1 \times W_2$ e $B(v, -v) = \mathbf{0}$. Como B é injetora, temos que $v = \mathbf{0}$. Logo, $W_1 \cap W_2 = \{\mathbf{0}\}$. Então, pelo item (c) do corolário 5.1, concluímos $V = W_1 \oplus W_2$.

($\impliedby$) Se $V = W_1 \oplus W_2$, temos que B é sobrejetora, pois $B(W_1 \times W_2) = W_1 + W_2 = V$. Além disso, como

$$\dim(V) = \dim(W_1) + \dim(W_2) = \dim(W_1 \times W_2),$$

temos que B é um isomorfismo. ■

Definição 5.3

Seja V um espaço vetorial. Dizemos que um operador linear E sobre V é uma *projeção* se $E^2 = E$.

Seja $E : V \longrightarrow V$ uma projeção, e sejam $E(V)$ a imagem de E e $\text{Ker}(E)$ o núcleo de E . Valem as seguintes afirmativas:

(1) $v \in E(V) \iff E(v) = v$.

De fato, se $E(v) = v$ então, claramente, $v \in E(V)$. Verifiquemos a suficiência da condição.

Seja $w \in E(V)$. Então existe $v \in V$ tal que $E(v) = w$.

Logo, $E(w) = E(E(v)) = E(v) = w$.

$$(2) V = E(V) \oplus \text{Ker}(E).$$

Com efeito, seja $v \in V$, então $v = E(v) + v - E(v)$. Como $E(v) \in E(V)$ e $E(v - E(v)) = E(v) - E^2(v) = \mathbf{0}$, ou seja $v - E(v) \in \text{Ker}(E)$, temos que $V = E(V) + \text{Ker}(E)$.

Além disso, se $v \in E(V) \cap \text{Ker}(E)$, então $E(v) = v$ e $E(v) = \mathbf{0}$, isto é, $v = \mathbf{0}$. Logo, $E(V) \cap \text{Ker}(E) = \{\mathbf{0}\}$ e, portanto, $V = E(V) \oplus \text{Ker}(E)$.

Observação 5.2

Seja $V = W_1 \oplus W_2$ e seja $E : V \rightarrow V$ definida por $E(v) = v_2$, onde $v = v_1 + v_2$ com $v_1 \in W_1$ e $v_2 \in W_2$.

Então E é um operador linear e é uma projeção, chamada a **projeção sobre W_2 segundo W_1** .

De fato, se $v = v_1 + v_2$ e $w = w_1 + w_2$, onde $v_1, w_1 \in W_1$ e $v_2, w_2 \in W_2$, então

$$E(\lambda v + w) = E((\lambda v_1 + w_1) + (\lambda v_2 + w_2)) = \lambda v_2 + w_2 = \lambda E(v) + E(w).$$

Além disso, $E(E(v)) = E(E(v_1 + v_2)) = E(v_2) = v_2 = E(v)$, ou seja $E^2 = E$ e $E(V) = W_2$.

Observação 5.3

Toda projeção $E : V \rightarrow V$ é diagonalizável.

De fato, como $V = E(V) + \text{Ker}(E)$, temos que $\mathcal{B} = \{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ é uma base de V , onde $\{v_1, \dots, v_r\}$ é uma base de $E(V)$ e $\{v_{r+1}, \dots, v_n\}$ é uma base de $\text{Ker}(E)$. Logo,

$$[E]_{\mathcal{B}} = \begin{pmatrix} I_{r \times r} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix},$$

pois $E(v_i) = v_i$, para $i = 1, \dots, r$, e $E(v_j) = \mathbf{0}$, para $j = r+1, \dots, n$.

Observação 5.4

Suponhamos que $V = W_1 \oplus \dots \oplus W_k$. Para cada $j = 1, \dots, k$, podemos definir a aplicação $E_j : V \rightarrow V$ por $E_j(v) = v_j$, onde $v = v_1 + \dots + v_k$, $v_i \in W_i$, $i = 1, \dots, k$.

É fácil verificar que cada E_j é linear, que $E_j(V) = W_j$ e que $E_j^2 = E_j$.

O núcleo de E_j é $W_1 + \dots + W_{j-1} + W_{j+1} + \dots + W_k$, pois a afirmação de que $E_j(v) = \mathbf{0}$ significa que v é uma soma de vetores dos espaços $W_1, \dots, W_{j-1}, W_{j+1}, \dots, W_k$.

Como $E_j(v) = v_j$, $j = 1, \dots, k$, onde $v = v_1 + \dots + v_k$, temos que

$$v = E_1(v) + \dots + E_k(v),$$

para todo vetor $v \in V$. Logo, $I = E_1 + \dots + E_k$.

Note também que se $i \neq j$, então $E_i E_j = \mathcal{O}$ (operador nulo), pois a imagem W_j de E_j está contida no núcleo de E_i .

Teorema 5.2

Se $V = W_1 \oplus \dots \oplus W_k$, então existem k operadores $E_1, \dots, E_k$ sobre V tais que:

(a) $E_j^2 = E_j$ (E_j é uma projeção), para $j = 1, \dots, k$.

(b) $E_i E_j = \mathcal{O}$, para $i \neq j$.

(c) $I = E_1 + \dots + E_k$.

(d) $E_j(V) = W_j$, $j = 1, \dots, k$.

Reciprocamente, se $E_1, \dots, E_k$ são k operadores lineares sobre V que satisfazem as condições (a), (b) e (c), então $V = W_1 \oplus \dots \oplus W_k$, onde $W_j = E_j(V)$, para $j = 1, \dots, k$.

Prova.

Precisamos provar apenas a recíproca. Sejam $E_1, \dots, E_k$ operadores lineares sobre V que satisfazem as condições (a), (b) e (c), e seja $W_j = E_j(V)$, $j = 1, \dots, k$. Então,

$$V = W_1 + \dots + W_k,$$

pois, por (c), temos que $v = E_1(v) + \dots + E_k(v)$, para todo $v \in V$.

Essa expressão é única, pois se $v = v_1 + \dots + v_k$, com $v_i \in W_i = E_i(V)$, $v_i = E_i(w_i)$, temos, por (a) e (b), que

$$E_j(v) = \sum_{i=1}^k E_j(v_i) = \sum_{i=1}^k E_j E_i(w_i) = E_j^2(w_j) = E_j(w_j) = v_j.$$

Logo, $V = W_1 \oplus \dots \oplus W_k$. ■

6. Somas Diretas Invariantes

Seja $V = W_1 \oplus \dots \oplus W_k$, onde cada subespaço W_i é invariante por um operador linear T dado. Seja $T_i = T|_{W_i}$, $i = 1, \dots, k$. Então, se $v = v_1 + \dots + v_k$, com $v_i \in W_i$, $i = 1, \dots, k$,

$$T(v) = T_1(v_1) + \dots + T_k(v_k).$$

Dizemos, nesse caso, que T é a **soma direta** dos operadores $T_1, \dots, T_k$.

Seja $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ uma base ordenada de V , onde $\mathcal{B}_i$ é uma base ordenada de W_i , $i = 1, \dots, k$. Então, se $A_i = [T_i]_{\mathcal{B}_i}$, temos que

$$[T]_{\mathcal{B}} = \begin{pmatrix} A_1 & \mathcal{O} & \dots & \mathcal{O} \\ \mathcal{O} & A_2 & \dots & \mathcal{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{O} & \mathcal{O} & \dots & A_k \end{pmatrix},$$

em que A_i é uma matriz $d_i \times d_i$, $d_i = \dim(W_i)$, e os símbolos $\mathcal{O}$ são blocos retangulares de vários tamanhos com entradas nulas.

Teorema 6.1

Seja T um operador linear sobre o espaço V e sejam W_i e E_i , $i = 1, \dots, k$, como no Teorema 5.2. Então, cada W_i é invariante por T se, e somente se, T comuta com cada uma das projeções E_i , $i = 1, \dots, k$.

Prova.

($\Leftarrow$) Suponhamos que T comuta com cada E_i . Seja $v \in W_i$. Então, $E_i(v) = v$ e

$$T(v) = T(E_i(v)) = E_i(T(v)).$$

Ou seja, $T(v) \in W_i$. Logo, cada W_i é invariante por T .

($\Rightarrow$) Suponhamos, agora, que cada W_i é invariante por T . Seja $v \in V$. Então,

$$v = E_1(v) + \dots + E_k(v) \Rightarrow T(v) = T(E_1(v)) + \dots + T(E_k(v)).$$

Como $E_i(v) \in W_i$ e esse subespaço é invariante por T , devemos ter $T(E_i(v)) = E_i(w_i)$, para algum $w_i \in W_i$.

Assim,

$$E_j(T(E_i(v))) = E_j(E_i(w_i)) = \begin{cases} \mathbf{0}, & \text{se } j \neq i; \\ E_j(w_j), & \text{se } j = i. \end{cases}$$

Logo,

$$E_j T(v) = E_j(T(E_1(v))) + \dots + E_j(T(E_k(v))) = E_j(w_j) = T(E_j(v)).$$

Como isso vale para todo $v \in V$, temos que $E_j T = T E_j$, $j = 1, \dots, k$. ■

Teorema 6.2

Seja T um operador linear sobre um espaço V de dimensão finita. Se T é diagonalizável e $c_1, \dots, c_k$ são os autovalores distintos de T , então existem operadores lineares $E_1, \dots, E_k$ sobre V tais que:

- (a) $T = c_1 E_1 + \dots + c_k E_k$.
- (b) $I = E_1 + \dots + E_k$.
- (c) $E_i E_j = \mathcal{O}$, se $i \neq j$.
- (d) $E_j^2 = E_j$ (i.e. E_j é uma projeção).
- (e) A imagem de E_j é o autoespaço de T associado a c_j .

Reciprocamente, se existem k escalares distintos $c_1, \dots, c_k$ e k operadores lineares não-nulos $E_1, \dots, E_k$ satisfazendo as condições (a), (b) e (c), então T é diagonalizável, $c_1, \dots, c_k$ são os autovalores distintos de T e as condições (d) e (e) são, também, satisfeitas.

Prova.

Suponha que T seja diagonalizável, com autovalores distintos $c_1, \dots, c_k$. Seja W_i o autoespaço associado ao autovalor c_i . Então, pelo Teorema 5.2, (b), (c), (d) e (e) são satisfeitas. Basta, então, verificar que (a) é satisfeita.

Seja $v \in V$.

Então,

$$v = E_1(v) + \dots + E_k(v).$$

e

$$T(v) = T(E_1(v)) + \dots + T(E_k(v)) = c_1 E_1(v) + \dots + c_k E_k(v),$$

ou seja, $T = c_1 E_1 + \dots + c_k E_k$.

Suponhamos, agora, que sejam dados um operador linear T , escalares distintos $c_1, \dots, c_k$ e operadores não-nulos $E_1, \dots, E_k$ satisfazendo (a), (b) e (c).

Como $E_i E_j = \mathcal{O}$, para $i \neq j$, temos, multiplicando ambos os membros

de $I = E_1 + \dots + E_k$ por E_i , que $E_i = E_i^2$. E, multiplicando a identidade $T = c_1 E_1 + \dots + c_k E_k$ por E_i , resulta que $TE_i = c_i E_i^2 = c_i E_i$, o que mostra que todo vetor na imagem de E_i está no núcleo de $T - c_i I$.

Como $E_i \neq \mathcal{O}$, existe um vetor não-nulo no núcleo de $T - c_i I$, ou seja c_i é um autovalor de T . Além disso, os escalares c_i , $i = 1, \dots, k$ são os únicos autovalores de T . De fato, se c é um escalar arbitrário, temos

$$T - cI = (c_1 - c)E_1 + \dots + (c_k - c)E_k.$$

Logo, se $(T - cI)(v) = \mathbf{0}$, devemos ter $(c_i - c)E_i(v) = \mathbf{0}$, $i = 1, \dots, k$, pois $V = E_1(V) \oplus \dots \oplus E_k(V)$, já que $E_i E_j = \mathcal{O}$, $i \neq j$, $E_j^2 = E_j$ e $I = E_1 + \dots + E_k$. Se v não é o vetor nulo, existe i tal que $E_i(v) \neq \mathbf{0}$, de modo que $c_i - c = 0$ para tal i .

O operador T é diagonalizável, pois todo vetor não-nulo na imagem de E_i é um autovetor e todo vetor v se escreve na forma

$$v = E_1(v) + \dots + E_k(v),$$

ou seja, os autovetores de T geram V .

Resta provar que $E_i(V) = \text{Ker}(T - c_i I)$.

Se $v \in \text{Ker}(T - c_i I)$, isto é, se $T(v) = c_i v$, então

$$\sum_{j=1}^k (c_j - c_i) E_j(v) = \mathbf{0}.$$

Logo, $(c_j - c_i) E_j(v) = \mathbf{0}$ para todo $j = 1, \dots, k$.

Assim, $E_j(v) = \mathbf{0}$ para $j \neq i$, ou seja, $v = E_i(v)$, mostrando que $\text{Ker}(T - c_i I) \subset E_i(V)$.

Como já provamos que $E_i(V) \subset \text{Ker}(T - c_i I)$, temos que $\text{Ker}(T - c_i I) = E_i(V)$, para todo $i = 1, \dots, k$. ■

Observação 6.1

Se $g = a_0 + a_1 x + \dots + a_n x^n$ é um polinômio em $\mathbb{K}[x]$, então

$$g(T) = g(c_1)E_1 + \dots + g(c_k)E_k,$$

ou seja,

$$g(T) = a_n \sum_{i=1}^k c_i^n E_i + \dots + a_0 \sum_{i=1}^k E_i.$$

De fato, basta provar que

$$T^m = \sum_{i=1}^k c_i^m E_i,$$

para todo inteiro positivo m .

Faremos a demonstração desse fato usando indução sobre m .

Se $m = 1$, temos que $T = \sum_{i=1}^k c_i E_i$.

Suponhamos o resultado válido para o inteiro m : $T^m = \sum_{i=1}^k c_i^m E_i$.

Então,

$$\begin{aligned} T^{m+1} &= T^m T = \left(\sum_{i=1}^k c_i^m E_i \right) \left(\sum_{j=1}^k c_j E_j \right) \\ &= \sum_{i=1}^k \sum_{j=1}^k c_i^m c_j E_i E_j = \sum_{i=1}^k c_i^{m+1} E_i, \end{aligned}$$

pois $E_i E_j = \mathcal{O}$, se $i \neq j$, e $E_i^2 = E_i$, para todo $i = 1, \dots, k$.

Observação 6.2

Se p_j são os polinômios de Lagrange correspondentes aos escalares $c_1, \dots, c_k$, i.e.

$$p_j = \prod_{i \neq j} \frac{x - c_i}{c_j - c_i},$$

temos que

$$p_j(T) = \sum_{i=1}^k p_j(c_i) E_i = E_i,$$

pois $p_j(c_i) = \delta_{ij}$. Logo, *as projeções E_i são polinômios em T .*

Observação 6.3

Daremos agora uma outra demonstração do seguinte resultado:

T é diagonalizável se, e somente se, o seu polinômio minimal tem a forma

$$p_m = (x - c_1) \dots (x - c_k),$$

onde $c_1, \dots, c_k \in \mathbb{K}$ são escalares distintos.

Se T é diagonalizável, $T = c_1 E_1 + \dots + c_k E_k$, então,

$$g(T) = g(c_1) E_1 + \dots + g(c_k) E_k,$$

para todo polinômio $g \in \mathbb{K}[x]$.

Logo, $g(T) = \mathcal{O}$ se, e somente se, $g(c_i) = 0$, para todo $i = 1, \dots, k$.

Assim,

$$p_m = (x - c_1) \dots (x - c_k),$$

é o polinômio minimal de T .

Suponhamos agora que T seja um operador linear sobre V com polinômio minimal

$$p_m = (x - c_1) \dots (x - c_k),$$

onde $c_1, \dots, c_k \in \mathbb{K}$ são escalares distintos.

Sejam

$$p_j = \prod_{i \neq j} \frac{x - c_i}{c_j - c_i},$$

os polinômios de Lagrange correspondentes aos escalares $c_1, \dots, c_k$.

Então, $p_j(c_i) = \delta_{ij}$ e

$$g = g(c_1) p_1 + \dots + g(c_k) p_k,$$

para todo polinômio g de grau $\leq k - 1$.

Tomando os polinômios $g = 1$ e $g = x$, obtemos que

$$\begin{aligned} 1 &= p_1 + \dots + p_k \\ x &= c_1 p_1 + \dots + c_k p_k. \end{aligned} \tag{I}$$

Observe que a segunda igualdade só é válida se $k > 1$. Mas, se $k = 1$, T é um múltiplo do operador identidade, sendo, portanto, diagonalizável.

Seja $E_j = p_j(T)$. De (I), temos que

$$I = p_1(T) + \dots + p_k(T) = E_1 + \dots + E_k,$$

e

$$T = c_1 p_1(T) + \dots + c_k p_k(T) = c_1 E_1 + \dots + c_k E_k.$$

Observe que, se $i \neq j$, então $p_i p_j$ é divisível pelo polinômio minimal p_m , pois $p_i p_j$ contém $x - c_\ell$ como fator, para todo $\ell = 1, \dots, k$.

Assim,

$$E_i E_j = p_i(T) p_j(T) = p_i p_j(T) = \mathcal{O}, \quad \text{se } i \neq j.$$

Além disso, $E_i \neq \mathcal{O}$, $i = 1, \dots, k$, pois $E_i = p_i(T)$ e p_i é um polinômio de grau menor que o polinômio minimal p_m .

Como $T = c_1 E_1 + \dots + c_k E_k$, $I = E_1 + \dots + E_k$, $E_i E_j = \mathcal{O}$, $i \neq j$, $E_i \neq \mathcal{O}$, $i = 1, \dots, k$, e $c_1, \dots, c_k$ são distintos, temos, pelo Teorema 6.2, que T é diagonalizável.

7. O Teorema da Decomposição Primária

Ao estudar um operador T usando seus autovalores, iremos nos confrontar com dois problemas. Primeiro, T poderá não ter nenhum valor característico, que é uma deficiência do corpo dos escalares, a saber, o corpo não é algebricamente fechado. Segundo, mesmo que o polinômio característico se decomponha num produto de polinômios de grau 1, podem não existir autovetores suficientes para gerar o espaço V , que é, evidentemente, uma deficiência do operador T .

Por exemplo, seja T o operador sobre $\mathbb{K}^3$ representado, em relação à base canônica, pela matriz

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

O polinômio característico de T é $(x - 2)^2(x + 1)$.

Como

$$A - 2I = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -3 \end{pmatrix} \quad \text{e} \quad A + I = \begin{pmatrix} 3 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

temos que

$$(A - 2I)(A + I) = \begin{pmatrix} 0 & 0 & 0 \\ 3 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq \mathcal{O}.$$

Logo, $p_m = (x - 2)^2(x + 1)$ é o polinômio minimal de T .

Como posto $(A - 2I) = \text{posto}(A + I) = 2$, temos que

$$\dim(\text{Ker}(T - 2I)) = \dim(\text{Ker}(T + I)) = 1,$$

ou seja,

$$\text{Ker}(T - 2I) \oplus \text{Ker}(T - I) \neq \mathbb{K}^3.$$

Mas,

$$\text{Ker}((T - 2I)^2) \oplus \text{Ker}(T + I) = \mathbb{K}^3,$$

pois $e_3 \in \text{Ker}(T + I)$ e $e_1, e_2 \in \text{Ker}((T - 2I)^2)$, uma vez que

$$(A - 2I)^2 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -3 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 9 \end{pmatrix}.$$

Mostraremos, no caso geral, que se

$$p_m = (x - c_1)^{r_1} \dots (x - c_k)^{r_k}$$

é o polinômio minimal de T , onde $c_1, \dots, c_k$ são escalares distintos, então

$$V = \text{Ker}((T - c_1)^{r_1}) \oplus \dots \oplus \text{Ker}((T - c_k)^{r_k}).$$

Teorema 7.1

Seja V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{K}$ e seja T um operador linear sobre V .

Seja $p_m = p_1^{r_1} \dots p_k^{r_k}$ o polinômio minimal de T , onde os $p_i \in \mathbb{K}[x]$ são polinômios primos unitários distintos e os r_i são inteiros positivos.

Seja $W_i = \text{Ker}(p_i^{r_i}(T))$, $i = 1, \dots, k$.

Então,

(a) $V = W_1 \oplus \dots \oplus W_k.$

(b) $T(W_i) \subset W_i$, $i = 1, \dots, k.$

(c) Se $T_i = T|_{W_i}$, então o polinômio minimal de T_i é $p_i^{r_i}.$

Prova.

Para cada i , seja

$$f_i = \frac{p}{p_i^{r_i}} = \prod_{j \neq i} p_j^{r_j},$$

Como $p_1, \dots, p_k$ são polinômios primos distintos, os polinômios $f_1, \dots, f_k$ são primos entre si.

Assim, existem polinômios $g_1, \dots, g_k$ tais que

$$\sum_{i=1}^k f_i g_i = 1.$$

Note, também, que se $i \neq j$, então $f_i f_j$ é divisível pelo polinômio p_m , pois

$f_i f_j$ contém $p_i^{r_i}$ como fator, $n = 1, \dots, k$.

Seja $h_i = f_i g_i$ e seja $E_i = h_i(T) = f_i(T)g_i(T)$.

Como $h_1 + \dots + h_k = 1$ e p_m divide $f_i f_j$, para $i \neq j$, temos que

$$\begin{aligned} E_i + \dots + E_k &= I, \\ E_i E_j &= \mathcal{O}, \text{ se } i \neq j. \end{aligned}$$

Assim, $E_1, \dots, E_k$ são projeções que correspondem a alguma decomposição do espaço V em soma direta.

Vamos mostrar que $E_i(V) = W_i$, $i = 1, \dots, k$.

Seja $v \in E_i(V)$, então $E_i(v) = v$ e

$$p_i^{r_i}(T)(v) = p_i^{r_i}(T)f_i(T)g_i(T)(v) = p_m(T)g_i(T)(v) = \mathbf{0}.$$

Ou seja, $E_i(V) \subset W_i$, $i = 1, \dots, k$.

Seja, agora, $v \in W_i = \text{Ker}(p_i^{r_i}(T))$.

Se $j \neq i$, $f_j g_j$ é divisível por $p_i^{r_i}$ e, então, $f_j(T)g_j(T)(v) = \mathbf{0}$, ou seja, $E_j(v) = \mathbf{0}$, se $j \neq i$. Logo, $v = E_i(v)$, isto é, $v \in E_i(V)$.

Pelo Teorema 5.2, temos que

$$V = W_1 \oplus \dots \oplus W_k.$$

É óbvio que cada W_i é invariante por T , pois se $v \in W_i = \text{Ker}(p_i^{r_i}(T))$, então $p_i^{r_i}(T(v)) = T(p_i^{r_i}(T)(v)) = \mathbf{0}$.

Se $T_i = T|_{W_i}$, então $p_i^{r_i}(T_i) = \mathcal{O}$, pois, por definição, $p_i^{r_i}(T)$ se anula no subespaço W_i . Logo, o polinômio minimal de T_i divide $p_i^{r_i}$.

Seja g um polinômio arbitrário tal que $g(T_i) = \mathcal{O}$.

Então, $g(T)f_i(T) = \mathbf{0}$, pois $g(T)(v) = \mathbf{0}$, para todo $v \in W_i$ e $f_i(T)(v) = \mathbf{0}$, para todo $v \in W_j$, $j \neq i$.

Assim, o polinômio minimal p_m de T divide $g f_i$, ou seja, $p_m = p_i^{r_i} f_i$ divide $g f_i$. Logo, $p_i^{r_i}$ divide g , pois f_i não possui em sua decomposição o fator primo p_i .

Portanto, o polinômio minimal de T_i é $p_i^{r_i}$. ■

Corolário 7.1

Se $E_1, \dots, E_k$ são as projeções associadas à decomposição primária de T , então cada E_i é um polinômio em T e, conseqüentemente, se um opera-

dor U comuta com T , então comuta com cada um dos E_i , ou seja, cada subespaço W_i é invariante por U .

Observação 7.1

Seja T um operador linear sobre V com polinômio minimal

$$p_m = (x - c_1)^{r_1} \dots (x - c_k)^{r_k},$$

ou seja, os polinômios primos p_i são da forma $x - c_i$.

Seja $W_i = \text{Ker}((T - c_i I)^{r_i}) = E_i(V)$, $i = 1, \dots, k$, e seja o operador

$$D = c_1 E_1 + \dots + c_k E_k.$$

Então, pelo Teorema 6.2, D é um operador diagonalizável, que denominamos a *parte diagonalizável de T* , sendo $c_1, \dots, c_k$ os autovalores distintos de D e $W_i = \text{Ker}((T - c_i I)^{r_i})$ o autoespaço de D associado ao autovalor c_i .

Consideremos o operador $N = T - D$. Como

$$\begin{aligned} T &= T E_1 + \dots + T E_k \\ D &= c_1 E_1 + \dots + c_k E_k, \end{aligned}$$

temos que

$$N = (T - c_1 I) E_1 + \dots + (T - c_k I) E_k.$$

Logo,

$$N^2 = (T - c_1 I)^2 E_1 + \dots + (T - c_k I)^2 E_k,$$

e, de modo geral,

$$N^r = (T - c_1 I)^r E_1 + \dots + (T - c_k I)^r E_k.$$

Quando $r \geq r_i$, para todo i , temos que $N^r = \mathcal{O}$, pois o operador $(T - c_i I)^r E_i = \mathcal{O}$, já que $E_i(V) = \text{Ker}((T - c_i I)^{r_i})$.

Definição 7.1

Seja N um operador linear sobre o espaço vetorial V de dimensão finita sobre o corpo $\mathbb{K}$. Dizemos que N é *nilpotente* se existe um inteiro positivo k tal que $N^k = \mathcal{O}$. O menor inteiro positivo r tal que $N^r = \mathcal{O}$ é denominado a *ordem de N* ou o *índice de nilpotência de N* .

Teorema 7.2

Seja V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{K}$ e seja T um operador linear sobre V .

Suponhamos que o polinômio minimal de T se decompõe em $\mathbb{K}$ como um produto de fatores lineares. Então, existe um operador diagonalizável D sobre V e um operador nilpotente N sobre V tais que

(a) $T = D + N$.

(b) $DN = ND$.

O operador diagonalizável D e o operador nilpotente N são determinados de modo único por (a) e (b), e cada um deles é um polinômio em T .

Prova.

Já provamos que podemos escrever $T = D + N$, onde $D = c_1 E_1 + \dots + c_k E_k$ é diagonalizável e $N = T - D$ é nilpotente.

Como cada E_i é um polinômio em T , temos que D e N são polinômios em T e, portanto, comutam entre si.

Suponhamos, agora, que $T = D' + N'$, onde D' é diagonalizável, N' é nilpotente e $D'N' = N'D'$.

Como D' e N' comutam entre si e $T = D' + N'$, temos que D' e N' comutam com T . Sendo D e N polinômios em T , temos que D' e N' comutam com D e N . Logo, os quatro operadores D , N , D' e N' comutam entre si.

Temos

$$D + N = T = D' + N' \implies D - D' = N' - N.$$

Como D e D' são operadores diagonalizáveis que comutam, eles são simultaneamente diagonalizáveis e, portanto, $D - D'$ é diagonalizável.

Como N e N' são operadores nilpotentes que comutam, $N' - N$ é nilpotente. De fato, como $NN' = N'N$, temos que

$$(N' - N)^r = \sum_{j=0}^r \binom{r}{j} (N')^{r-j} (-N)^j,$$

o que pode ser provado usando um argumento indutivo sobre r junto com a fórmula do binômio de Newton.

Seja $r' = \max\{r_1, r_2\}$, onde $(N')^{r_1} = \mathcal{O}$ e $N^{r_2} = \mathcal{O}$.

Então,

$$(N' - N)^{2r'} = \sum_{j=0}^{2r'} \binom{2r'}{j} (N')^{2r'-j} (-N)^j = \mathcal{O},$$

pois se $j \geq r'$, $(-N)^j = \mathcal{O}$, e se $j < r'$, temos $2r' - j > r'$, logo $(N')^{2r'-j} = \mathcal{O}$.

Como $D - D' = N' - N$, temos que $D - D'$ é um operador diagonalizável e nilpotente. Então, $D - D' = \mathcal{O}$, ou seja $D = D'$, e, portanto, $N = N'$ (ver observação abaixo). ■

Observação 7.2

De fato, podemos provar que o único operador L que é simultaneamente diagonalizável e nilpotente é o operador nulo de, pelo menos, duas maneiras:

- Suponha que $L^r = \mathcal{O}$ e que λ é um autovalor de L . Então, existe $v \neq \mathbf{0}$ tal que $L(v) = \lambda v$. Logo, $L^r(v) = \lambda^r v = \mathbf{0}$. Como $v \neq \mathbf{0}$, devemos ter $\lambda = 0$. Assim, o espaço V é gerado pelos autovetores associados ao autovalor 0, ou seja, $L = \mathcal{O}$.
- Sendo L nilpotente, seu polinômio minimal é da forma x^s , para algum $s \geq 1$. Mas como L é também diagonalizável, seu polinômio minimal é um produto de fatores lineares distintos. Logo, x é o polinômio minimal de L , ou seja $L = \mathcal{O}$.

Corolário 7.2

Seja V um espaço vetorial de dimensão finita sobre um corpo algebricamente fechado $\mathbb{K}$.

Então, todo operador linear T sobre V se escreve como uma soma de um operador diagonalizável D com um operador nilpotente N , que comutam. Esses operadores D e B são únicos e são polinômios em T .

Lema 7.1

Se N é um operador nilpotente, então $I + N$ é um isomorfismo.

Prova.

Seja $L = \sum_{j=0}^r (-1)^j N^j$, onde $N^{r+1} = \mathcal{O}$ e $N^r \neq \mathcal{O}$. Então,

$$(I + N)L = L(I + N) = \sum_{j=0}^r (-1)^j N^j + \sum_{j=0}^r (-1)^j N^{j+1} = I.$$

Logo, $I + N$ é invertível e

Observe a analogia entre o operador inverso de $I + N$ ao lado e a série de potências $(1 - x)^{-1} =$

$$\sum_{j=0}^{\infty} x^j.$$

$$(I + N)^{-1} = \sum_{j=0}^r (-1)^j N^j.$$



Observação 7.3

Se N é nilpotente e $\mu \neq 0$, então $\mu I + N$ é um isomorfismo.

De fato, como $\mu I + N = \mu \left(I + \frac{1}{\mu} N \right)$ e $\frac{1}{\mu} N$ é nilpotente, temos, pelo lema anterior, que $I + \frac{1}{\mu} N$ é um isomorfismo. Logo, $\mu I + N = \mu \left(I + \frac{1}{\mu} N \right)$.

Lema 7.2

Seja V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{K}$ e seja N um operador nilpotente sobre V .

Suponha que o polinômio característico de N se escreve como um produto de fatores lineares em $\mathbb{K}[x]$. Então, x^n é o polinômio característico de N , onde $n = \dim(V)$.

Prova.

De fato, se $p_c = (x - c_1)^{d_1} \dots (x - c_k)^{d_k}$ é o polinômio característico de N , $c_1, \dots, c_k$ são os valores característicos distintos de N .

Mas como N é nilpotente, 0 (o zero) é o único valor característico de N . Logo, $p_c = x^n$. ■

Observação 7.4

Se V é um espaço vetorial de dimensão n sobre o corpo $\mathbb{K}$ e N é um operador nilpotente sobre V , então $p_c = x^n$ é o polinômio característico de N .

Com efeito, seja $\overline{\mathbb{K}}$ um corpo algebricamente fechado do qual $\mathbb{K}$ é um subcorpo. Seja $\mathcal{B}$ uma base ordenada de V e $A = [N]_{\mathcal{B}}$.

Considerando A como uma matriz sobre $\overline{\mathbb{K}}$, temos, pelo lema anterior, que x^n é o seu polinômio característico. Mas, como o polinômio característico de N é $\det(xI - N) = \det(xI - A)$, e esse determinante independe se estamos considerando A como uma matriz com entradas em $\mathbb{K}$ ou em $\overline{\mathbb{K}}$, temos que x^n é o polinômio característico de N .

Teorema 7.3

Seja T um operador linear sobre o espaço vetorial V de dimensão finita, cujo polinômio característico é

$$p_c = (x - c_1)^{d_1} \dots (x - c_k)^{d_k},$$

e cujo polinômio minimal é

$$p_m = (x - c_1)^{r_1} \dots (x - c_k)^{r_k}.$$

Seja $W_i = \text{Ker}(T - c_i I)^{r_i}$. Então,

- (a) $W_i = \{v \in V \mid \exists n \in \mathbb{N}^*; (T - c_i I)^n(v) = \mathbf{0}\}$;
- (b) $\dim(W_i) = d_i, i = 1, \dots, k$;
- (c) $(T - c_i I)|_{W_i}$ é nilpotente de *ordem* r_i , ou seja, $((T - c_i I)|_{W_i})^{r_i} \equiv \mathcal{O}$ e $((T - c_i I)|_{W_i})^{r_i-1} \neq \mathcal{O}$.

Prova.

Vamos mostrar primeiro que $W_i = \overline{W_i}$, onde

$$\overline{W_i} = \{v \in V \mid \exists n \in \mathbb{N}^*; (T - c_i I)^n(v) = \mathbf{0}\}.$$

Seja $v \in W_i = \text{Ker}(T - c_i I)^{r_i}$. Então, $(T - c_i I)^{r_i}(v) = \mathbf{0}$.

Logo, $v \in \overline{W_i}$, isto é, $W_i \subset \overline{W_i}$.

Seja $v \in \overline{W_i}$. Então, existe $n \in \mathbb{N}^*$, tal que $(T - c_i I)^n(v) = \mathbf{0}$.

Como $V = W_1 \oplus \dots \oplus W_k$, existem vetores $v_j \in W_j, j = 1, \dots, k$, tais que $v = v_1 + \dots + v_k$.

Logo,

$$\begin{aligned} \mathbf{0} = (T - c_i I)^n(v) &= \sum_{j=1}^k (T_j - c_i I)^n(v_j) \\ &= \sum_{j=1}^k ((c_j - c_i)I + (T_j - c_j I))^n(v_j). \end{aligned}$$

Como $(T_j - c_i I)^n(v_j) \in W_j$, e os subespaços $W_1, \dots, W_k$ são LI, temos que $((c_j - c_i)I + (T_j - c_j I))^n(v_j) = \mathbf{0}, j = 1, \dots, k$.

Pelo lema anterior, $(c_j - c_i)I + (T_j - c_j I)$ é um isomorfismo para $j \neq i$, pois $c_j - c_i \neq 0$ e $T_j - c_j I$ é nilpotente.

Logo, $((c_j - c_i)I + (T_j - c_j I))^n$ é um isomorfismo, se $j \neq i$.

Como $((c_j - c_i)I + (T_j - c_j I))^n(v_j) = \mathbf{0}$, temos que $v_j = \mathbf{0}$ para $j \neq i$. Assim,

$v = v_i \in W_i$, ou seja, $\overline{W_i} \subset W_i$.

Vamos mostrar, agora, que $\dim(W_i) = \dim(\overline{W_i}) = d_i$, onde d_i é a multiplicidade de c_i como raiz do polinômio característico p_c de T .

Como $p_c = (x - c_1)^{d_1} \dots (x - c_k)^{d_k}$ é o polinômio característico de T , temos

$$\begin{aligned} \widetilde{p}_c &= \det(xI - (T - c_i I)) = \det((x + c_i)I - T) \\ &= (x + c_i - c_i)^{d_i} \prod_{j \neq i} (x + c_i - c_j)^{d_j} \\ &= x^{d_i} \prod_{j \neq i} (x - (c_j - c_i))^{d_j}, \end{aligned}$$

é o polinômio característico de $T - c_i I$ que, como vemos na última expressão, é um produto de fatores lineares.

Como o polinômio característico de $N_i = T_i - c_i I$ divide o polinômio característico de $T - c_i I$, temos que o polinômio característico de N_i é um produto de fatores lineares.

Seja $n_i = \dim(W_i)$. Como $N_i = T_i - c_i I$ é nilpotente, temos, pelo lema 7.2, que x^{n_i} é o seu polinômio característico.

Logo, o polinômio característico de $T_i = N_i + c_i I$ é

$$\det(xI - T_i) = \det(xI - (N_i + c_i I)) = \det((x - c_i)I - N_i) = (x - c_i)^{n_i}.$$

Mas, como $(x - c_i)^{n_i}$ divide p_c , temos que $n_i \leq d_i$. Logo, $n_i = d_i$, pois

$$n_1 + \dots + n_k = d_1 + \dots + d_k,$$

ou seja, $\dim(W_i) = d_i$.

Já sabemos que $(T_i - c_i I)^{r_i} = \mathcal{O}$, pois $T_i = T|_{W_i}$ e $W_i = \text{Ker}(T - c_i I)^{r_i}$. Suponhamos, por absurdo, que $(T_i - c_i I)^{v_i} = \mathcal{O}$, para algum $v_i < r_i$.

Pelo Teorema da Decomposição Primária, temos que $(x - c_i)^{r_i}$ é o polinômio minimal de T_i . Logo, x^{r_i} é o polinômio minimal de $N_i = T_i - c_i I$.

Como $N_i^{v_i} = \mathcal{O}$, temos que x^{r_i} divide x^{v_i} , o que é uma contradição, pois $v_i < r_i$.

Então, $(T - c_i I)^{r_i} = \mathcal{O}$ e $(T_i - c_i I)^{r_i-1} \neq \mathcal{O}$, ou seja, $T_i - c_i I$ é um operador nilpotente de ordem r_i . ■

Forma Canônica Racional

1. Subespaços cíclicos e anuladores

Definição 1.1

Seja V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{K}$, e seja $v \in V$. O **subespaço T -cíclico gerado por v** é o subespaço

$$Z(v, T) = \{g(T)(v) \mid g \in \mathbb{K}[x]\}.$$

Se $Z(v, T) = V$, dizemos que v é um **vetor cíclico** de T .

Observação 1.1

Seja W o menor subespaço de V que contém v e é invariante por T , ou seja, W é a interseção de todos os subespaços de V que contém v e são invariantes por T .

Então, $W = Z(v, T)$. De fato, se $g(T)(v) \in Z(v, T)$, $g \in \mathbb{K}[x]$, temos $g(T)(v) \in W$, pois W é invariante por T e contém v . Logo, $Z(v, T) \subset W$.

Mas, como $Z(v, T)$ é um subespaço invariante por T que contém v , temos que $W \subset Z(v, T)$.

- $Z(v, T)$ é o subespaço gerado pelos vetores $T^k(v)$, $k \geq 0$.

Então, v é um vetor cíclico de T se, e somente se, $\{T^k(v) \mid k \geq 0\}$ gera o espaço V .

- $Z(v, T)$ é unidimensional se, e somente se, v é um autovetor de T .
- Se $T = I$ é o operador identidade, todo vetor não nulo gera um subespaço cíclico unidimensional. Logo, se $\dim(V) > 1$, I não possui um vetor cíclico.
- Seja T um operador nilpotente sobre um espaço vetorial de dimensão

n sobre o corpo $\mathbb{K}$, tal que $T^n = \mathcal{O}$ e $T^{n-1} \neq \mathcal{O}$. Seja $v \in V$ tal que $T^{n-1}(v) \neq \mathbf{0}$. Então,

$$\{v, T(v), \dots, T^{n-1}(v)\},$$

é uma base de V , ou seja, v é um vetor cíclico de N .

De fato, se

$$\begin{aligned} & \alpha_0 v + \alpha_1 T(v) + \dots + \alpha_{n-1} T^{n-1}(v) = \mathbf{0} \\ \implies & \alpha_0 T^{n-1}(v) + \alpha_1 T^n(v) + \dots + \alpha_{n-1} T^{2n-2}(v) = \mathbf{0} \\ \implies & \alpha_0 T^{n-1}(v) = \mathbf{0} \implies \alpha_0 = 0. \end{aligned}$$

Ou seja, $\alpha_1 T(v) + \dots + \alpha_{n-1} T^{n-1}(v) = \mathbf{0}$. Logo,

$$\begin{aligned} & \alpha_1 T^{n-1}(v) + \dots + \alpha_{n-1} T^{2n-3}(v) = \mathbf{0} \\ \implies & \alpha_1 T^{n-1}(v) = \mathbf{0} \implies \alpha_1 = 0. \end{aligned}$$

Prosseguindo dessa maneira, obtemos que $\alpha_0 = \alpha_1 = \dots = \alpha_{n-1} = 0$, isto é, $\{v, T(v), \dots, T^{n-1}(v)\}$ é LI, e, portanto, uma base de V .

Exemplo 1.1

Seja T o operador linear sobre $\mathbb{R}^2$ representado, com respeito à base canônica, pela matriz

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Como $A \neq \mathcal{O}$, $A^2 = \mathcal{O}$ e $A(e_1) = e_2 \neq \mathbf{0}$, temos que $\{e_1, A(e_1) = e_2\}$ é uma base de $\mathbb{R}^2$, ou seja, e_1 é um vetor cíclico de T . $\square$

Definição 1.2

Seja V um espaço vetorial sobre o corpo $\mathbb{K}$.

Definimos o **T -anulador** de um vetor $v \in V$ como sendo o ideal em $\mathbb{K}[x]$ dado por

$$M(v, T) = \{g \in \mathbb{K}[x] \mid g(T)(v) = \mathbf{0}\}.$$

O único polinômio unitário p_v que gera esse ideal é também denominado o T -anulador de v .

Observação 1.2

Seja p_m o polinômio minimal de T . Como $p_m(T) = \mathcal{O}$, temos $p_m(T)(v) = \mathbf{0}$. Logo, $p_m \in M(v, T)$ e p_v divide p_m .

Observação 1.3

O grau de p_v é zero se, e somente se, $v = \mathbf{0}$.

Teorema 1.1

Seja v um vetor não nulo e seja p_v o T -anulador do vetor $v \in V$. Então:

(a) $\text{grau}(p_v) = \dim(Z(v, T))$.

(b) Se $\text{grau}(p_v) = k$, então $\mathcal{B}_v = \{v, T(v), \dots, T^{k-1}(v)\}$ é uma base de $Z(v, T)$.

(c) Se $T_v = T|_{Z(v, T)}$, então p_v é o polinômio minimal de T_v .

(d) A matriz de T_v em relação à base $\mathcal{B}_v$ é

$$[T_v]_{\mathcal{B}_v} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{k-1} \end{pmatrix},$$

A matriz ao lado é a matriz associada ao polinômio unitário p_v .

onde $p_v = a_0 + a_1x + \dots + a_{k-1}x^{k-1} + x^k$.

Prova.

Seja $g \in \mathbb{K}[x]$. Então, existem $q, r \in \mathbb{K}[x]$ tais que $g = p_v q + r$, onde $r \equiv 0$ ou $\text{grau}(r) < \text{grau}(p_v) = k$.

Como $p_v(T)q(T)(v) = \mathbf{0}$, temos que

$$g(T)(v) = r(T)(v).$$

Sendo $r \equiv 0$ ou $\text{grau}(r) < k$, temos que $g(T)(v)$ é uma combinação linear de $v, T(v), \dots, T^{k-1}(v)$. Como $g(T)(v)$ é um vetor arbitrário de $Z(v, T)$, obtemos que $\{v, T(v), \dots, T^{k-1}(v)\}$ gera o subespaço $Z(v, T)$.

Além disso, $\{v, T(v), \dots, T^{k-1}(v)\}$ é LI, pois, caso contrário, existiria um polinômio não-nulo de grau menor que k que anulava v , o que é um absurdo, pois $\text{grau}(p_v) = k$.

Seja $T_v = T|_{Z(v, T)}$ e p o polinômio minimal de T_v .

Como $p_v(T)(v) = \mathbf{0}$, temos que

$$p_v(T)(T^j(v)) = T^j(p_v(T)(v)) = \mathbf{0}, \forall j \in \mathbb{N}.$$

Logo, $p_v(T_v) = \mathcal{O}$. Então, p divide p_v . (I)

Por outro lado, como $p(T_v) = \mathcal{O}$, temos que $\mathbf{0} = p(T_v)(v) = p(T)(v)$, ou

seja, $p \in M(v, T)$. Logo, p_v divide p .

(III)

De (I) e (II), temos que $p = p_v$, pois ambos são polinômios unitários.

Se $p_v = a_0 + a_1x + \dots + a_{k-1}x^{k-1} + x^k$, temos que

$$T^k(v) = -a_0v - a_1T(v) - \dots - a_{k-1}T^{k-1}(v).$$

Fazendo $v_i = T^{i-1}(v)$, $i = 1, \dots, k$ e $\mathcal{B}_v = \{v_1, \dots, v_k\}$ é uma base de $Z(v, T)$ tal que

$$T(v_i) = T(T^{i-1}(v)) = T^i(v) = v_{i+1}, \text{ se } i = 1, \dots, k-1,$$

e

$$T(v_k) = T(T^{k-1}(v)) = T^k(v) = -a_0v_1 - a_1v_2 - \dots - a_{k-1}v_k.$$

Logo,

$$[T_v]_{\mathcal{B}_2} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & 0 & \dots & 0 & -a_2 \\ 0 & 0 & 1 & \dots & 0 & -a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -a_{k-1} \end{pmatrix}.$$



Observação 1.4

Se T é um operador linear sobre o espaço V que possui um vetor cíclico v , então seu polinômio minimal é igual ao seu polinômio característico.

De fato, pelo Teorema acima, $V = Z(v, T)$ e p_v é o polinômio minimal p_m de T . Como $\text{grau}(p_v) = \dim(Z(v, T)) = \dim(V) = \text{grau}(p_c)$, sendo p_c o polinômio característico de T , e como p_m divide p_c , temos que

$$p_m = p_c = p_v.$$

Provaremos depois que para todo operador T existe um vetor $v \in V$ cujo T -anulador é o polinômio minimal de T .

Com esse resultado, podemos provar que se o polinômio minimal é igual ao polinômio característico de T , então T possui um vetor cíclico.

De fato, seja $v \in V$ tal que $p_v = p_m$, sendo p_m o polinômio minimal de T . Como $p_m = p_c$, sendo p_c o polinômio característico de T , temos que

$$\dim(Z(v, T)) = \text{grau}(p_v) = \text{grau}(p_c) = \dim(V),$$

ou seja, $Z(v, T) = V$.

Corolário 1.1

Se T é um operador linear sobre o espaço V de dimensão finita, então T possui um vetor cíclico se, e só se, existe uma base ordenada $\mathcal{B}$ de V tal que $[T]_{\mathcal{B}}$ é a matriz associada a um polinômio unitário p . Nesse caso, $p = p_m = p_c$, onde p_m é o polinômio minimal e p_c é o polinômio característico de T .

Prova.

A implicação $(\implies)$ já foi provada.

$(\impliedby)$ Como $[T]_{\mathcal{B}}$ é uma matriz $n \times n$, onde $n = \dim(V)$, temos que $\text{grau}(p) = n$.

Se $\mathcal{B} = \{v_1, \dots, v_n\}$, temos que

$$v_2 = T(v_1), v_3 = T(v_2), \dots, v_i = T(v_{i-1}), \dots, v_n = T(v_{n-1}).$$

Assim, $v_i = T^{i-1}(v_1)$, para todo $i = 1, \dots, n$, pois $v_1 = T^{1-1}(v_1)$, e se $v_{j-1} = T^{j-2}(v_1)$, então $v_j = T(v_{j-1}) = T(T^{j-2}(v_1)) = T^{j-1}(v_1)$.

Logo, $\{v_1, T(v_1), \dots, T^{n-1}(v_1)\}$ gera o espaço V e, portanto, v_1 é um vetor cíclico de T . ■

2. Decomposição Cíclica e a Forma Racional

Seja T um operador sobre um espaço V e seja W um subespaço de V invariante por T . Suponha que existe um subespaço W' de V invariante por T tal que $V = W \oplus W'$.

Seja $v \in V$. Então, existe $w \in W$ e $w' \in W'$, tais que $v = w + w'$. Se $g \in \mathbb{K}[x]$,

$$g(T)(v) = g(T)(w) + g(T)(w').$$

Como $g(T)(w) \in W$ e $g(T)(w') \in W'$, temos que $g(T)(v) \in W$ se, e somente se, $g(T)(w') = \mathbf{0}$, isto é, se, e somente se, $g(T)(v) = g(T)(w)$.

Ou seja, se $g(T)(v) \in W$, então existe $w \in W$ tal que

$$g(T)(v) = g(T)(w).$$

Definição 2.1

Sejam T um operador linear sobre o espaço V e W um subespaço de V .

Dizemos que W é **T –admissível** se:

- (a) W é invariante por T .
- (b) Se $f(T)(v) \in W$, então existe $w \in W$ tal que $f(T)(v) = f(T)(w)$.

Observação 2.1

Pelo provado acima, temos que se W é invariante e tem um subespaço suplementar invariante, então W é T –admissível. Uma das consequências do teorema que provaremos a seguir é a recíproca dessa afirmação, ou seja, se W é T –admissível, então, W possui um subespaço suplementar invariante.

Teorema 2.1

Sejam T um operador sobre um espaço vetorial V de dimensão finita e W_0 um subespaço próprio T –admissível de V . Existem vetores não-nulos $v_1, \dots, v_r \in V$ com respectivos T –anuladores $p_1, \dots, p_r$ tais que

- (i) $V = W_0 \oplus Z(v_1, T) \oplus \dots \oplus Z(v_r, T)$,
- (ii) p_k divide p_{k-1} , $k = 2, \dots, r$.

Além disso, o inteiro r e os anuladores $p_1, \dots, p_r$ são determinados de modo único por (i) e (ii) e pelo fato de que nenhum w_k é nulo.

Prova.

Vamos fazer a demonstração em quatro partes.

Parte I. Existem vetores não-nulos $w_1, \dots, w_r \in V$ tais que

- (a) $V = W_0 + Z(w_1, T) + \dots + Z(w_r, T)$,
- (b) se $1 \leq k \leq r$ e

$$W_k = W_0 + Z(w_1, T) + \dots + Z(w_k, T),$$

o condutor $p_k = s(w_k, W_{k-1})$ tem grau máximo entre todos os T –condutores no espaço W_{k-1} , isto é, para cada k

$$\text{grau}(p_k) = \max_{v \in V} \text{grau}(s(v, W_{k-1})).$$

- Se W é um subespaço próprio T –invariante, então

$$0 < \max_{v \in V} \text{grau}(s(v, W)) \leq \dim(V).$$

De fato, se $v \notin W$, $\text{grau}(s(v, W)) \geq 1$. E como o polinômio minimal p_m de T pertence a $S(v, W)$, para todo $v \in V$, pois $p_m(T)(v) = \mathbf{0}$, temos que

$s(v, W)$ divide p_m .

Logo,

$$\text{grau}(s(v, W)) \leq \text{grau}(p_m) \leq \dim(V) = n.$$

Seja $w \in V$ tal que $\text{grau}(s(w, W)) = \max_{v \in V} \text{grau}(s(v, W))$.

O subespaço $W + Z(w, T)$ é, então, T -invariante e possui dimensão maior que $\dim W$, pois $w \notin W$.

Apliquemos esse processo a $W = W_0$ para obter w_1 . Então,

$$\text{grau}(s(w_1, W_0)) = \max_{v \in V} \text{grau}(s(v, W_0)),$$

$$\dim(W_0 + Z(w_1, T)) > \dim(V),$$

e $W_0 + Z(w_1, T)$ é T -invariante.

Se $W_1 = W_0 + Z(w_1, T)$ é próprio, existe $w_2 \in V$ tal que

$$\text{grau}(s(w_2, W_1)) = \max_{v \in V} \text{grau}(s(v, W_1)),$$

$$\dim(W_0 + Z(w_1, T) + Z(w_2, T)) > \dim(W_1) > \dim(W),$$

e $W_0 + Z(w_1, T) + Z(w_2, T)$ é T -invariante.

Prosseguindo desta maneira, devemos obter $V = W_r$ em, no máximo, $\dim(V) = n$ passagens.

Parte II. Fixemos $k \in \{1, \dots, r\}$. Seja $w \in V$ e seja $f = s(w, W_{k-1})$. Se

$$f(T)(w) = w_0 + \sum_{i=1}^{k-1} g_i(T)(w_i), \quad \text{com } w_0 \in W_0,$$

então f divide g_i , $i = 1, \dots, k-1$, e $w_0 = f(T)(u_0)$, para algum $u_0 \in W_0$.

Se $k = 1$, $f(T)(w) = w_0 \in W_0$. Como W_0 é T -admissível, existe $u_0 \in W_0$, tal que $f(T)(w) = f(T)(u_0)$.

Suponhamos $k > 1$ e apliquemos o algoritmo da divisão:

$$g_i = f h_i + r_i, \quad \text{com } r_i \equiv 0 \text{ ou } \text{grau}(r_i) < \text{grau}(f_i).$$

Queremos mostrar que $r_i \equiv 0$, $i = 1, \dots, k-1$.

Seja $u = w - \sum_{i=1}^{k-1} h_i(T)(w_i)$. Então $u - w \in W_{k-1}$.

Afirmção: $f = s(u, W_{k-1})$.

De fato, $q \in s(u, W_{k-1}) \iff q(T)(u) \in W_{k-1} \iff q(T)(w) \in W_{k-1} \iff q \in s(w, W_{k-1})$.

Além disso,

$$\begin{aligned} f(T)(u) &= f(T)(w) - \sum_{i=1}^{k-1} (f h_i)(T)(w_i) \\ &= w_0 + \sum_{i=1}^{k-1} g_i(T)(w_i) - \sum_{i=1}^{k-1} (f h_i)(T)(w_i) \\ &= w_0 + \sum_{i=1}^{k-1} r_i(T)(w_i). \end{aligned}$$

Suponhamos que existe $r_i, i = 1, \dots, k-1$, diferente de zero.

Seja $j = \max\{i \mid r_i \neq 0\}$. Então,

$$f(T)(u) = w_0 + \sum_{i=1}^j r_i(T)(w_i),$$

com $r_j \neq 0$ e $\text{grau}(r_j) < \text{grau}(f)$.

Seja $p = s(u, W_{j-1})$. Como $W_{j-1} \subset W_{k-1}$, o condutor $f = s(u, W_{k-1})$ divide p , ou seja, $p = f g$, para algum $g \in \mathbb{K}[x]$.

Então,

$$p(T)(u) = g(T) f(T)(u) = g(T)(w_0) + g(T) r_j(T)(w_j) + \sum_{i=1}^{j-1} (g r_i)(T)(w_i).$$

Como $p(T)(u) \in W_{j-1}$ e $g(T)(w_0) + \sum_{i=1}^{j-1} (g r_i)(T)(w_i) \in W_{j-1}$, temos que

$$g(T) r_j(T)(w_j) \in W_{j-1}.$$

Logo,

$$\begin{aligned} \text{grau}(g r_j) &\geq \text{grau}(s(w_j, W_{j-1})) = \text{grau}(p_j) \\ &\geq \text{grau}(s(u, W_{j-1})) = \text{grau}(p) = \text{grau}(f g). \end{aligned}$$

Então, $\text{grau}(r_j) \geq \text{grau}(f)$, que é uma contradição.

Como $r_i = 0, i = 1, \dots, k-1$, temos que f divide $g_i, i = 1, \dots, k-1$, e $f(T)(u) = w_0 \in W_0$. Sendo W_0 T -admissível, existe $u_0 \in W_0$, tal que $f(T)(u) = f(T)(u_0) = w_0$.

Essa parte II nos diz, em particular, que $W_1, \dots, W_r$ são subespaços T -admissíveis.

De fato, seja $w \in V$ e seja $f = s(w, W_{j-1})$, $j = 1, \dots, r$. Então, pela parte II, temos

$$f(T)(w) = f(T)(u_0) + \sum_{i=1}^{j-1} f(T) h_i(T)(w_i),$$

onde $u_0 \in W_0$.

Suponhamos que $q(T)(w) \in W_{j-1}$, com $q \in \mathbb{K}[x]$.

Então, f divide q , ou seja, $q = f h$, $h \in \mathbb{K}[x]$.

Logo,

$$\begin{aligned} q(T)(w) &= h(T) f(T)(w) = h(T) f(T) \left(u_0 + \sum_{i=1}^{j-1} h_i(T)(w_i) \right) \\ &= q(T) \left(u_0 + \sum_{i=1}^{j-1} h_i(T)(w_i) \right), \end{aligned}$$

onde $u_0 + \sum_{i=1}^{j-1} h_i(T)(w_i) \in W_{j-1}$.

Parte III. Existem vetores não-nulos $v_1, \dots, v_r$ em V que satisfazem as condições (a) e (b) do teorema.

Aplicando o resultado da parte II ao vetor $w = w_k$, $k = 1, \dots, r$, e ao T -condutor $f = p_k = s(w_k, W_{k-1})$, temos que

$$p_k(T)(w_k) = p_k(T)(u_0) + \sum_{i=1}^{k-1} p_k(T) h_i(T)(w_i),$$

onde $u_0 \in W_0$. Seja

$$v_k = w_k - u_0 - \sum_{i=1}^{k-1} h_i(T)(w_i).$$

Como $w_k - v_k \in W_{k-1}$, temos que $s(v_k, W_{k-1}) = s(w_k, W_{k-1}) = p_k$. Além disso, como $p_k(T)(v_k) = 0$, temos que

$$W_{k-1} \cap Z(v_k, T) = \{0\}.$$

De fato, seja $w \in W_{k-1} \cap Z(v_k, T)$. Então, existe $g \in \mathbb{K}[x]$ tal que $w = g(T)(v_k)$. Como $g(T)(v_k) \in W_{k-1}$, temos que p_k divide g , ou seja, existe $h \in \mathbb{K}[x]$ tal que $g = h p_k$.

Logo, $w = g(T)(v_k) = h(T) p_k(T)(v_k) = 0$.

Afirmção: $p_k = s(v_k, W_{k-1})$ é o T -anulador de v_k .

Se $q(T)(v_k) = 0$, então $q(T)(v_k) \in W_{k-1}$. Logo, p_k divide o T -anulador de v_k . Como $p_k(T)(v_k) = 0$, o T -anulador de v_k divide p_k . Então, p_k é o T -anulador de v_k .

Afirmção: $p_k = s(v_k, W_{k-1})$ é o T -anulador de v_k .

Se $q(T)(v_k) = \mathbf{0}$, então $q(T)(v_k) \in W_{k-1}$. Logo, p_k divide o T -anulador de v_k . Como $p_k(T)(v_k) = \mathbf{0}$, o T -anulador de v_k divide p_k . Então, p_k é o T -Anulador de v_k .

Afirmção: $W_1 = W_0 \oplus Z(v_1, T)$.

Como já provamos que $W_0 \cap Z(v_1, T) = \{\mathbf{0}\}$, basta provar que

$$W_1 = W_0 + Z(w_1, T) = W_0 + Z(v_1, T).$$

Sendo $v_1 = w_1 - u_0 \in W_0 + Z(w_1, T)$, temos que $Z(v_1, T) \subset W_0 + Z(w_1, T)$. Logo, $W_0 + Z(v_1, T) \subset W_1$.

De modo análogo, como $w_1 = u_0 + v_1 \in W_0 + Z(v_1, T)$, temos que $Z(w_1, T) \subset W_0 + Z(v_1, T)$. Logo, $W_0 + Z(w_1, T) \subset W_0 + Z(v_1, T)$, ou seja, $W_1 \subset W_0 + Z(v_1, T)$.

Afirmção: Suponha que $W_j = W_0 \oplus Z(v_1, T) \oplus \dots \oplus Z(v_j, T)$. Vamos mostrar que $W_{j+1} = W_0 \oplus Z(v_1, T) \oplus \dots \oplus Z(v_j, T) \oplus Z(v_{j+1}, T)$.

Como $W_{j+1} = W_j + Z(w_{j+1}, T)$ e $W_j \cap Z(v_{j+1}, T) = \{\mathbf{0}\}$ basta provar que $W_{j+1} = W_j + Z(v_{j+1}, T)$.

Sendo $v_{j+1} = w_{j+1} - u_j$, onde $u_j = u_0 + \sum_{i=1}^j h_i(T)(w_i) \in W_j$, temos que

$Z(v_{j+1}, T) \subset W_j + Z(w_{j+1}, T)$. Logo,

$$W_j + Z(v_{j+1}, T) \subset W_j + Z(w_{j+1}, T).$$

De modo análogo, como $w_{j+1} = v_{j+1} + u_j$, temos que

$$Z(w_{j+1}, T) \subset W_j + Z(v_{j+1}, T).$$

Logo,

$$W_j + Z(w_{j+1}, T) \subset W_j + Z(v_{j+1}, T).$$

Vamos, agora, provar a **parte (b)** do Teorema.

Fazendo $w = v_2$ e $f = s(v_2, W_1) = p_2$ na parte II da demonstração, temos que p_2 divide p_1 .

De fato, como $p_2(T)(v_2) = \mathbf{0}$, $p_1(T)(v_1) = \mathbf{0}$ e $v_1 = w_1 - u_0$, $u_0 \in W_0$, temos

$$p_2(T)(v_2) = p_1(T)(v_1) = p_1(T)(-u_0) + p_1(T)(w_1),$$

onde $p_1(T)(-u_0) \in W_0$. Logo, pela parte II, p_2 divide p_1 .

Vamos provar que p_3 divide p_2 , fazendo $w = v_3$ e $f = s(v_3, W_2) = p_3$.

De fato, como $p_3(T)(v_3) = p_2(T)(v_2) = p_1(T)(v_1) = \mathbf{0}$, $v_1 = -u_0 + w_1$ e $v_2 = -u'_0 + w_2 - h(T)(w_1)$, onde $u_0, u'_0 \in W_0$, temos que

$$\begin{aligned} p_3(T)(v_3) &= p_1(T)(v_1) + p_2(T)(v_2) \\ &= p_1(T)(-u_0) + p_1(T)(w_1) + p_2(T)(-u'_0) \\ &\quad + p_2(T)(w_2) - p_2(T)h(T)(w_1) \\ &= (p_1(T)(-u_0) + p_2(T)(-u'_0)) \\ &\quad + (p_1 - p_2h)(T)(w_1) + p_2(T)(w_2). \end{aligned}$$

Como $p_1(T)(-u_0) + p_2(T)(-u'_0) \in W_0$, temos, pela parte II, que $p_3 \mid p_2$.

Prosseguindo desta maneira, podemos provar que $p_j \mid p_{j-1}$, isto é, p_j divide p_{j-1} , $j = 1, \dots, r$. Então

$$p_r \mid p_{r-1} \mid \dots \mid p_2 \mid p_1.$$

Parte IV. Unicidade. O número r e os polinômios $p_1, \dots, p_r$ são determinados de modo único pelas condições do Teorema.

Suponhamos que, além dos vetores não-nulos $v_1, \dots, v_r$, existem outros vetores não-nulos $u_1, \dots, u_s$ com respectivos T -anuladores $g_1, \dots, g_s$, tais que

$$\begin{cases} V = W_0 \oplus Z(u_1, T) \oplus \dots \oplus Z(u_s, T) \\ g_k \mid g_{k-1}, \quad k = 1, \dots, s. \end{cases}$$

Mostraremos que $r = s$ e $p_i = g_i$ para todo i .

Seja $S(V, W_0) = \{g \in \mathbb{K}[x] \mid g(T)(V) \subset W_0\}$. É fácil ver que $S(V, W_0)$ é um ideal não-nulo, pois W_0 é T -invariante e o polinômio minimal pertence a $S(V, W_0)$.

Seja g o gerador unitário de $S(V, W_0)$. Vamos mostrar que $g_1 = g$.

Seja $v \in V$. Então, existe $w_0 \in W$ e existem $f_1, \dots, f_s \in \mathbb{K}[x]$, tais que

$$v = w_0 + f_1(T)(u_1) + \dots + f_s(T)(u_s).$$

Portanto, $g_1(T)(v) = g_1(T)(w_0) + g_1(T)f_1(T)(u_1) + \dots + g_1(T)f_s(T)(u_s)$.

Como g_i divide g_1 , $i = 1, \dots, s$, temos que $g_1(T)f_k(T)(u_k) = \mathbf{0}$, $k = 1, \dots, s$. Logo, $g_1(T)(v) = g_1(T)(w_0) \in W_0$. Ou seja, $g_1 \in S(V, W_0)$. Então, g divide g_1 .

Afirmção: $g_1 = s(u_1, W_0)$.

Seja $f = s(u_1, W_0)$. Como $g_1 \in S(u_1, W_0)$, temos que f divide g_1 . Por outro lado, $f(T)(u_1) \in W_0 \cap Z(u_1, T) = \{\mathbf{0}\}$. Logo, $f(T)(u_1) = \mathbf{0}$.

Então, g_1 divide f . Assim, $g_1 = f = s(u_1, W_0)$.

Como $S(V, W_0) \subset S(u_1, W_0)$, temos que g_1 divide g .

Logo, $g = g_1$. Pelos mesmos argumentos, podemos provar que $g = p_1$. Assim, $p_1 = g_1$.

Observação 2.2

Quando $W_0 = \{\mathbf{0}\}$, temos que o gerador unitário g do ideal $S(V, \{\mathbf{0}\})$ é o polinômio minimal de T . Logo, acabamos de provar que: *se T é um operador linear sobre um espaço vetorial de dimensão finita V , existe $v \in V$ tal que o T -anulador de v é o polinômio minimal de T .*

Antes de continuar com a demonstração do Teorema 2.1, provaremos os seguintes resultados:

(1) $f(T)(Z(v, T)) = Z(f(T)(v), T)$.

- Se $w \in Z(v, T)$, existe $g \in \mathbb{K}[x]$ tal que $w = g(T)(v)$. Logo,

$$f(T)(w) = f(T)g(T)(v) = g(T)f(T)(v),$$

ou seja, $f(T)(w) \in Z(f(T)(v), T)$.

Se $w \in Z(f(T)(v), T)$, existe $g \in \mathbb{K}[x]$ tal que $w = g(T)f(T)(v) = f(T)g(T)(v)$. Logo, como $g(T)(v) \in Z(v, T)$, temos que $w \in f(T)(Z(v, T))$.

(2) Se $V = V_1 \oplus \dots \oplus V_k$, onde cada V_i é T -invariante, então

$$f(T)(V) = f(T)(V_1) \oplus \dots \oplus f(T)(V_k).$$

- Seja $v \in f(T)(V)$. Então, existe $w \in V$ tal que $v = f(T)(w)$. Sejam $w_i \in V_i$, $i = 1, \dots, k$, tais que $w = w_1 + \dots + w_k$.

Então, $v = f(T)(w) = f(T)(w_1) + \dots + f(T)(w_k)$. Como $f(T)(w_i) \in f(T)(V_i)$, $i = 1, \dots, k$, temos que $f(T)(V) \subset f(T)(V_1) + \dots + f(T)(V_k)$.

Seja, agora, $v \in f(T)(V_1) + \dots + f(T)(V_k)$. Então, existem $v_i \in V_i$, $i = 1, \dots, k$, tais que

$$v = f(T)(v_1) + \dots + f(T)(v_k) = f(T)(v_1 + \dots + v_k).$$

Ou seja, $v \in f(T)(V)$.

Logo, $f(T)(V) = f(T)(V_1) + \dots + f(T)(V_k)$.

• Vamos, agora, provar que:

$$(f(T)(V_1) + \dots + f(T)(V_{j-1})) \cap f(T)(V_j) = \{0\}, \quad j = 2, \dots, k.$$

Seja $v \in (f(T)(V_1) + \dots + f(T)(V_{j-1})) \cap f(T)(V_j)$.

Então, existem v_i , $i = 1, \dots, j$, tais que

$$v = f(T)(v_1) + \dots + f(T)(v_{j-1}) = f(T)(v_j),$$

ou seja, $v = f(T)(v_1 + \dots + v_{j-1}) = f(T)(v_j)$.

Mas, como cada V_i é T -invariante, temos que

$$f(T)(v_1 + \dots + v_{j-1}) \subset V_1 + \dots + V_{j-1},$$

e $f(T)(v_j) \in V_j$.

Logo, $v = 0$, pois $(V_1 + \dots + V_{j-1}) \cap V_j = \{0\}$.

(3) Se v e w têm o mesmo T -anulador, então $f(T)(v)$ e $f(T)(w)$ têm o mesmo T -anulador e, portanto,

$$\dim(Z(f(T)(v), T)) = \dim(Z(f(T)(w), T)).$$

• Seja $g \in \mathbb{K}[x]$ um polinômio pertencente ao T -anulador de $f(T)(v)$.

Então, $g(T)f(T)(v) = 0$, ou seja gf pertence ao T -anulador de v .

Logo, p_v divide gf , ou seja, existe $h \in \mathbb{K}[x]$, tal que $gf = p_v h$. Mas, como

$p_v = p_w$, temos que $gf = p_w h = hp_w$. Assim,

$$g(T)f(T)(w) = h(T)p_w(T)(w) = 0,$$

ou seja, g pertence ao T -anulador de $f(T)(w)$.

De modo análogo, podemos provar que se $\bar{g} \in \mathbb{K}[x]$ pertence ao T -anulador de w , então $\bar{g}$ pertence ao T -anulador de v .

Podemos, finalmente, concluir a demonstração do Teorema 2.1.

Mostraremos que $r = s$ e $p_i = g_i$, $i = 2, \dots, r$.

Seja $r \geq 2$, $\dim(W_0) + \dim(Z(v_1, T)) < \dim(V)$. Como $p_1 = g_1$,

$$\dim(Z(v_1, T)) = \text{grau}(p_1) = \text{grau}(g_1) = \dim(Z(v_1, T)).$$

Logo,

$$\dim(W_0) + \dim(Z(v_1, T)) < \dim(V).$$

Então, $s \geq 2$. Mostraremos, agora, que $p_2 = g_2$.

De (2), provado acima, temos que

$$\begin{aligned} p_2(T)(V) &= p_2(T)(W_0) \oplus p_2(T)(Z(v_1, T)) \oplus p_2(T)(Z(v_2, T)) \\ &\quad \oplus \dots \oplus p_r(T)(Z(v_r, T)). \end{aligned}$$

Logo, por (1),

$$p_2(T)(V) = p_2(T)(W_0) \oplus Z(p_2(T)(v_1), T),$$

pois

$$p_2(T)(Z(v_i, T)) = Z(p_2(T)(v_i), T) = \{0\},$$

já que p_i divide p_2 , $i \geq 2$ ($\implies p_2(T)(v_i) = 0$, $i \geq 2$).

Também, por (2) e (1),

$$p_2(T)(V) = p_2(T)(W_0) \oplus Z(p_2(T)(u_1), T) \oplus \dots \oplus Z(p_2(T)(u_j), T).$$

Como v_1 e u_1 têm o mesmo T -anulador, temos, por (3), que $p_2(T)(v_1)$ e $p_2(T)(u_1)$ tem o mesmo T -anulador e

$$\dim(Z(p_2(T)(v_1), T)) = \dim(Z(p_2(T)(u_1), T)).$$

Logo, $Z(p_2(T)(u_i), T) = \{0\}$, se $i \geq 2$, ou seja, $p_2(T)(u_i) = 0$, se $i \geq 2$.

Em particular, $p_2(T)(u_2) = 0$. Logo, g_2 divide p_2 .

Se invertermos o argumento, podemos mostrar que p_2 divide g_2 . Logo, $g_2 = p_2$.

Prosseguindo desta maneira, chegaremos que $r = s$ e $p_i = g_i$, $i = 1, \dots, r$. ■

Corolário 2.1

Se T é um operador linear sobre um espaço vetorial V de dimensão finita, então, todo subespaço T -admissível possui um subespaço suplementar que também é T -invariante.

Prova.

Seja W_0 um subespaço T -admissível de V . Se $W_0 = V$, o seu suplementar é $\{0\}$.

Se W_0 é próprio, seja $W'_0 = Z(v_1, T) \oplus \dots \oplus Z(v_r, T)$, dado pelo Teorema anterior. Então W'_0 é T -invariante e $V = W_0 \oplus W'_0$. ■

Corolário 2.2

Seja T um operador linear sobre um espaço vetorial V de dimensão finita.

Então:

(a) Existe um vetor $v \in V$ tal que o T -anulador de v é o polinômio minimal de T .

(b) T possui um vetor cíclico se, e somente se, os polinômios característicos e minimal de T são iguais.

Prova.

(a) Fazendo $W_0 = \{0\}$ no Teorema anterior, temos que

$$V = Z(v_1, T) \oplus \dots \oplus Z(v_r, T),$$

onde $p_r | p_{r-1} | \dots | p_2 | p_1$.

Seja p_m o polinômio minimal de T .

Como $p_m(T)(v) = 0, \forall v \in V$, temos que $p_i | p_m, i = 1, \dots, r$.

Em particular, $p_1 | p_m$.

Seja, agora, $v \in V$. Então, existem $w_i \in Z(v_i, T), i = 1, \dots, r$, tais que

$$v = w_1 + \dots + w_r.$$

Logo,

$$p_1(T)(v) = p_1(T)(w_1) + \dots + p_1(T)(w_r) = 0,$$

pois $p_i | p_1$ e $p_i(T)(v_i) = 0, i = 1, \dots, r$.

Assim, $p_m | p_1$. Portanto, $p_m = p_1$.

(b) Se T possui um vetor cíclico v , então

$$V = Z(v, T).$$

Logo, por (a), $p_m = p_v$, onde p_m é o polinômio minimal de T . Mas, como

$$\text{grau}(p_v) = \dim(Z(v, T)) = \dim(V),$$

temos que p_m é o polinômio característico p_c de T .

Se $p_m = p_c$, temos que

$$\text{grau}(p_m) = \text{grau}(p_c) = \dim(V).$$

Logo,

$$\text{grau}(p_1) = \text{grau}(p_m) = \dim(V).$$

Assim $r = 1$ na decomposição cíclica, ou seja, $V = Z(v_1, T)$. Então v_1 é um vetor cíclico de T . ■

Teorema 2.2

Seja T um operador linear sobre um espaço vetorial de dimensão finita. Sejam p_m e p_c os polinômios minimal e característico de T , respectivamente. Então:

(a) p_m divide p_c .

(b) p_m e p_c possuem os mesmos fatores primos a menos de multiplicidades.

(c) Se

$$p_m = f_1^{r_1} \dots f_k^{r_k},$$

é a decomposição de p_m em fatores primos, então

$$p_c = f_1^{d_1} \dots f_k^{d_k},$$

onde d_i é a nulidade de $f_i(T)^{r_i}$ dividida pelo grau de f_i , ou seja,

$$d_i = \frac{\dim(\text{Ker}(f_i(T)^{r_i}))}{\text{grau}(f_i)}.$$

Prova.

Seja $V = Z(v_1, T) \oplus \dots \oplus Z(v_r, T)$ a decomposição cíclica do operador T , e seja p_i o T -anulador de v_i , $i = 1, \dots, r$.

Seja $T_i = T|_{Z(v_i, T)}$. Como v_i é um vetor cíclico de T_i , temos, pelo Corolário anterior, que p_i é o polinômio minimal e o polinômio característico de T_i .

Logo, como cada $Z(v_i, T)$ é T -invariante, temos $p_c = p_1 \dots p_r$.

Temos, também, que $p_m = p_1$. Logo, p_m divide p_c , ou seja, todo fator primo de p_m é um fator primo de p_c .

Se f é um fator primo de p_c , então f divide $p_m = p_1 \dots p_r$. Logo, existe i tal que f divide p_i . Como p_i divide p_1 , $i = 1, \dots, r$, temos que p_i divide $p_1 = p_m$. Então, f divide p_m .

Pelo Teorema da Decomposição Primária,

$$V = V_1 \oplus \dots \oplus V_k,$$

onde $p_m = f_1^{r_1} \dots f_k^{r_k}$ é a decomposição de p_m em fatores primos e

$$V_i = \text{Ker}(p_i^{r_i}), \quad i = 1, \dots, k.$$

Como o polinômio minimal de $T_i = T|_{V_i}$ é $f_i^{r_i}$, aplicando (b) ao operador T_i , temos que $f_i^{d_i}$, $d_i \geq r_i$, é o polinômio característico de T_i . Logo, $d_i \text{ grau}(f_i) = \dim(V_i)$, ou seja,

$$d_i = \frac{\dim(\text{Ker}(f_i^{r_i}(T)))}{\text{grau}(f_i)}.$$

Como cada V_i é T -invariante, temos que $p_c = f_1^{d_1} \dots f_k^{d_k}$. ■

Observação 2.3

Se o polinômio minimal de T

$$p_m = (x - c_1)^{r_1} \dots (x - c_k)^{r_k},$$

é um produto de potências de fatores lineares distintos, então

$$p_c = (x - c_1)^{d_1} \dots (x - c_k)^{d_k},$$

onde $d_i = \dim(\text{Ker}(T - c_i I)^{r_i})$, $i = 1, \dots, k$.

Corolário 2.3

Se T é um operador linear nilpotente sobre um espaço vetorial V de dimensão n , então x^n é o seu polinômio característico.

Prova.

Seja k um inteiro positivo tal que $T^k = \mathcal{O}$. Então, p_m divide x^k . Logo, x é o único fator primo de p_m e de p_c . Como p_c tem grau n , temos que $p_c = x^n$. ■

Observação 2.4

Seja T um operador linear sobre um espaço vetorial de dimensão finita n .

Sejam $v_1, \dots, v_r$ vetores não-nulos de V tais que

$$(1) V = Z(v_1, T) \oplus \dots \oplus Z(v_r, T);$$

$$(2) p_{i+1} \text{ divide } p_i, \quad i = 1, \dots, r-1.$$

onde p_i é o T -anulador de v_i , $i = 1, \dots, r$.

Como $T_i = T|_{Z(v_i, T)}$ possui o vetor cíclico v_i , temos que p_i é o polinômio minimal de T_i e, também, o polinômio característico de T_i .

Seja $\mathcal{B}_i = \{v_i, T(v_i), \dots, T^{k_i-1}(v_i)\}$ uma base de $Z(v_i, T)$, onde

$$k_i = \dim(Z(v_i, T)) = \text{grau}(p_i), \quad i = 1, \dots, r.$$

Então, $[T_i]_{\mathcal{B}_i}$ é a matriz associada ao polinômio unitário não-constante p_i , $i = 1, \dots, r$.

Se $\mathcal{B}$ é a base ordenada de V obtida pela reunião das $\mathcal{B}_i$ ordenadas da forma $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_r$, temos que

$$[T]_{\mathcal{B}} = \begin{pmatrix} A_1 & \mathcal{O} & \cdots & \mathcal{O} \\ \mathcal{O} & A_2 & \cdots & \mathcal{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{O} & \mathcal{O} & \cdots & A_r \end{pmatrix},$$

onde A_i é a matriz $k_i \times k_i$ associada ao polinômio unitário p_i .

Note que: O polinômio característico de T é $p_1 \cdot p_2 \cdot \dots \cdot p_r$.

Definição 2.2

Dizemos que a matriz $n \times n$ A está na **forma racional** se

$$A = \begin{pmatrix} A_1 & \mathcal{O} & \cdots & \mathcal{O} \\ \mathcal{O} & A_2 & \cdots & \mathcal{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{O} & \mathcal{O} & \cdots & A_r \end{pmatrix},$$

onde cada A_i é a matriz associada a um polinômio unitário não-constante p_i , tais que p_{i+1} divide p_i , $i = 1, \dots, r-1$.

Teorema 2.3

Seja B uma matriz $n \times n$ com entradas no corpo $\mathbb{K}$. Então, B é semelhante sobre $\mathbb{K}$ a uma única matriz na forma racional.

Prova.

Seja T o operador linear sobre $\mathbb{K}^n$ que é representado por B em relação à base canônica. Pela observação anterior, existe uma base de $\mathbb{K}^n$ em relação à qual T é representado por uma matriz A na forma racional. Então, B é semelhante à matriz A , que está na forma racional.

Suponhamos que B seja semelhante sobre $\mathbb{K}$ a uma outra matriz C na forma racional. Então, existe uma matriz P invertível tal que $C = P^{-1}BP$.

Os vetores colunas de P formam uma base $\mathcal{B}$ de $\mathbb{K}^n$ tal que

$$[T]_{\mathcal{B}} = C = \begin{pmatrix} C_1 & \mathcal{O} & \cdots & \mathcal{O} \\ \mathcal{O} & C_2 & \cdots & \mathcal{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{O} & \mathcal{O} & \cdots & C_s \end{pmatrix},$$

onde cada C_i é uma matriz $r_i \times r_i$ associada a um polinômio unitário não-constante $g_i = x^{r_i} + c_{r_i-1}^i x^{r_i-1} + \dots + c_0^i$, de modo que g_{i+1} divide g_i , $i = 1, \dots, s-1$.

Assim, existem vetores $w_1, \dots, w_s$ pertencentes à base $\mathcal{B}$ tais que:

- $\mathcal{B} = \{w_1, T(w_1), \dots, T^{r_1-1}(w_1), w_2, T(w_2), \dots, T^{r_2-1}(w_2), \dots, w_s, T(w_s), \dots, T^{r_s-1}(w_s)\};$
- $T^{r_i}(w_i) = -c_0^i w_i - c_1^i T(w_i) - \dots - c_{r_i-1}^i T^{r_i-1}(w_i).$

Logo, $\mathcal{B}_i = \{w_i, T(w_i), \dots, T^{r_i-1}(w_i)\}$ é uma base de $Z(w_i, T)$ e g_i é o T -anulador de w_i , pois $g_i(T)(w_i) = 0$ e $\text{grau}(g_i) = \dim(Z(w_i, T)) = r_i$.

Assim, $\mathbb{K}^n = Z(w_1, T) \oplus \dots \oplus Z(w_s, T)$ e g_{i+1} divide g_i , $i = 1, \dots, s-1$.

Pela unicidade da decomposição cíclica, $s = r$ e $g_i = p_i$, $i = 1, \dots, r$.

Portanto, $C = A$. ■

Observação 2.5

A forma racional de um operador T sobre um espaço vetorial V de dimensão finita é a matriz $[T]_{\mathcal{B}}$, onde $\mathcal{B}$ é uma base de V tal que $[T]_{\mathcal{B}}$ está na forma racional.

Exemplo 2.1

Seja T um operador linear sobre um espaço vetorial V de dimensão dois.

Se o polinômio minimal de T tem grau 1, $p_m = x - c$, então $T = cI$.

Sejam $v_1, v_2 \in V$ vetores LI. Então

$$V = Z(v_1, T) \oplus Z(v_2, T), \quad p_1 = p_2 = x - c \quad \text{e} \quad [T]_{\mathcal{B}} = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix},$$

onde $\mathcal{B} = \{v_1, v_2\}$.

Se o polinômio minimal $p_m = x^2 + ax + b$ tem grau dois, então $p_c = p_m$. Logo, T possui um vetor cíclico v_1 e uma base $\{v_1, T(v_1)\}$ tal que $V = Z(v_1, T)$ e

$$[T]_{\mathcal{B}} = \begin{pmatrix} 0 & -b \\ 1 & -a \end{pmatrix}.$$

Para matrizes, essa análise diz que toda matriz 2×2 sobre o corpo $\mathbb{K}$ é semelhante sobre $\mathbb{K}$ a exatamente uma matriz dos tipos:

$$\begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} \quad \text{ou} \quad \begin{pmatrix} 0 & -b \\ 1 & -a \end{pmatrix}.$$

□

Exemplo 2.2

Seja T o operador linear sobre $\mathbb{R}^3$ que é representado, em relação à base canônica pela matriz

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

Vimos, no exemplo 2.1, que $p_m = (x - 1)(x - 2)^2$ é o polinômio característico de T e $p_m = (x - 1)(x - 2) = x^2 - 3x + 2$ é o polinômio minimal de T .

Como, na decomposição cíclica, o T -anulador do primeiro vetor v_1 é o polinômio minimal, temos que $\dim(Z(v_1, T)) = 2$. Portanto, na decomposição, aparece apenas mais um vetor v_2 , sendo $\dim(Z(v_2, T)) = 1$, ou seja, v_2 é um vetor característico de T . O T -anulador de v_2 é $p_2 = x - 2$, pois $p_c = p_1 p_2$.

Então, existe uma base ordenada $\mathcal{B}$ de V tal que

$$[T]_{\mathcal{B}} = \begin{pmatrix} 0 & -2 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Como $\dim(Z(v, T))$ é igual a 1 ou a 2, se $v \neq 0$, pois $\mathbb{R}^3$ tem dimensão 3 e T não possui vetor cíclico, já que $p_m \neq p_c$, temos que $e_1 = (1, 0, 0)$ gera um subespaço T -cíclico de dimensão 2, pois $T(e_1) = (5, -1, 3)$ não é múltiplo de e_1 . Podemos, então, fazer $v_1 = e_1$.

O subespaço $Z(e_1, T)$ consiste dos vetores

$$ae_1 + bT(e_1) = (a + 5b, -b, 3b),$$

$a, b \in \mathbb{R}$, ou seja, consiste de todos os vetores (x_1, x_2, x_3) tais que $x_3 = -3x_2$.

Devemos achar agora um vetor $v_2 \neq 0$, tal que $T(v_2) = 2v_2$ e $v_2 \notin Z(v_1, T)$.

Se $v = (x_1, x_2, x_3)$, pode-se verificar facilmente que $T(v) = 2v$ se, e somente se, $x_1 = 2x_2 + 2x_3$.

Assim, $v_2 = (2, 1, 0)$ satisfaz $T(v_2) = 2v_2$ e $v_2 \notin Z(v_1, T)$. Podemos, então, tomar

$$\mathcal{B} = \{e_1, T(e_1), v_2\} = \{(1, 0, 0), (5, -1, 3), (2, 1, 0)\}.$$

□

Observação 2.6

Seja T um operador linear diagonalizável sobre um espaço vetorial V de dimensão finita. Sejam $c_1, \dots, c_k$ os valores característicos distintos de T e seja W_i o espaço dos vetores característicos associados ao valor característico c_i . Então,

$$V = W_1 \oplus \dots \oplus W_k,$$

e

$$p_c = (x - c_1)^{d_1} \dots (x - c_k)^{d_k},$$

onde $d_i = \dim(W_i)$.

Vamos, agora, determinar a decomposição cíclica de T a partir de uma base que diagonaliza a matriz de T .

Seja $v \in V$. Então existem $v_i \in W_i$, $i = 1, \dots, k$, tais que $v = v_1 + \dots + v_k$.

Como $T(v_i) = c_i v_i$, temos que

$$f(T)(v) = f(c_1)(v_1) + \dots + f(c_k)(v_k) \quad (I)$$

para todo polinômio $f \in \mathbb{K}[x]$. Sabemos que dados escalares arbitrários $t_1, \dots, t_k$, existe um polinômio f tal que $f(c_i) = t_i$, $i = 1, \dots, k$. Logo, $Z(v, T)$ é o subespaço gerado pelos vetores $v_1, \dots, v_k$.

Temos, por (I), que $f(T)(v) = \mathbf{0}$ se, e somente se, $f(c_i)(v_i) = 0$, $i = 1, \dots, k$. Portanto, $f(T)(v) = 0$ se, e somente se $f(c_i) = 0$, para todo i tal que $v_i \neq \mathbf{0}$.

Assim, o T -anulador de v é o produto

$$\prod_{i; v_i \neq \mathbf{0}} (x - c_i).$$

Seja $\mathcal{B}_i = \{v_1^i, \dots, v_{d_i}^i\}$ base de W_i e seja $r = \max_{1 \leq i \leq k} d_i$.

Definimos $v_1, \dots, v_r$ por

$$v_j = \sum_{i; d_i \geq j} v_j^i, \quad 1 \leq j \leq r.$$

Logo, o subespaço $Z(v_j, T)$ é gerado pelos vetores v_j^i tais que $d_i \geq j$, e o T -anulador de v_j é

$$p_j = p_{v_j} = \prod_{i; d_i \geq j} (x - c_i),$$

sendo $p_{v_1} = \prod_{i=1}^n (x - c_i) = p_m$.

Logo, $V = Z(v_1, T) \oplus \dots \oplus Z(v_r, T)$, pois cada v_j^i , $i = 1, \dots, k$, $j = 1 \dots, d_i$, pertence a um e somente um dos subespaços $Z(v_1, T), \dots, Z(v_r, T)$, e $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_r$ é uma base de V .

Além disso, p_{j+1} divide p_j para todo $j = 1, \dots, r-1$, pois se $x - c_i$ é fator de p_{j+1} , então $x - c_i$ é um fator de p_j , já que $d_i \geq j+1$ implica que $d_i \geq j$.





Forma Canônica de Jordan

1. Forma Canônica Racional dos Operadores Nilpotentes

Seja N um operador linear nilpotente de ordem k sobre um espaço vetorial V de dimensão n , ou seja, $N^k = \mathcal{O}$ e $N^{k-1} \neq \mathcal{O}$, onde $k \leq n$.

Pelo Teorema da Decomposição Cíclica, existe um número inteiro positivo r e r vetores não-nulos $v_1, \dots, v_r$ em V com N -anuladores $p_1, \dots, p_r$ tais que

$$V = Z(v_1, N) \oplus \dots \oplus Z(v_r, N),$$

e p_{i+1} divide p_i , $i = 1, \dots, r-1$.

Como N é nilpotente de ordem k , temos que $p_m = x^k$ é o polinômio minimal de N . Assim, $p_1 = x^k$ e $p_i = x^{k_i}$, onde

$$k_1 = k \geq k_2 \geq \dots \geq k_r.$$

A matriz associada ao polinômio x^{k_i} é a matriz $k_i \times k_i$

$$A_i = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix},$$

que é denominada **matriz nilpotente elementar de ordem k_i** , cujo polinômio minimal, que é igual ao seu polinômio característico, é x^{k_i} .

Então, pelo Teorema da Decomposição Cíclica, existe uma base ordenada $\mathcal{B}$, tal que

$$[N]_{\mathcal{B}} = \begin{pmatrix} A_1 & \mathcal{O} & \cdots & \mathcal{O} \\ \mathcal{O} & A_2 & \cdots & \mathcal{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{O} & \mathcal{O} & \cdots & A_r \end{pmatrix},$$

onde cada A_i é uma matriz nilpotente elementar de ordem k_i , $k_1 = k$, $k_1 + \dots + k_r = n$ e $k_1 \geq k_2 \geq \dots \geq k_r$. Esses inteiros positivos r e $k_1, \dots, k_r$, determinam a forma racional do operador N .

Observação 1.1

O número r de blocos é igual à dimensão do núcleo de N , ou melhor, $\{N^{k_1-1}(v_1), \dots, N^{k_r-1}(v_r)\}$ é uma base de $\text{Ker}(N)$.

De fato, seja $v \in \text{Ker}(N)$. Como $V = Z(v_1, N) \oplus \dots \oplus Z(v_r, N)$, existem polinômios $f_i \in \mathbb{K}[x]$, tais que:

$$v = f_1(N)(v_1) + \dots + f_r(N)(v_r),$$

sendo $\text{grau}(f_i) \leq k_i - 1$.

Como $N(v) = 0$, temos que

$$\begin{aligned} 0 &= N(v) = N f_1(N)(v_1) + \dots + N f_r(N)(v_r) \\ &= (x f_1)(N)(v_1) + \dots + (x f_r)(N)(v_r). \end{aligned}$$

Logo, $(x f_i)(N)(v_i) = 0$, $i = 1, \dots, r$. Como x^{k_i} é o N -anulador de v_i , temos que x^{k_i} divide $x f_i$.

Sendo $\text{grau}(f_i) \leq k_i - 1$, temos que $\text{grau}(f_i) = k_i - 1$ e $f_i = c_i x^{k_i-1}$, para algum $c_i \in \mathbb{K}$. Assim,

$$v = c_1 N^{k_1-1}(v_1) + \dots + c_r N^{k_r-1}(v_r).$$

Observe que cada $N^{k_i-1}(v_i) \in \text{Ker}(N)$, pois $N(N^{k_i-1}(v_i)) = N^{k_i}(v_i) = 0$, já que x^{k_i} é o N -anulador de v_i .

Então, $\{N^{k_1-1}(v_1), \dots, N^{k_r-1}(v_r)\}$ é uma base do núcleo de N , pois os vetores $N^{k_i-1}(v_i)$, $i = 1, \dots, r$, são LI e geram $\text{Ker}(N)$.

A matriz

$$A = \begin{pmatrix} A_1 & \mathcal{O} & \cdots & \mathcal{O} \\ \mathcal{O} & A_2 & \cdots & \mathcal{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{O} & \mathcal{O} & \cdots & A_r \end{pmatrix},$$

dada acima, é chamada **forma canônica do operador nilpotente N** .

2. Cálculo da Forma Canônica dos Operadores Nilpotentes

Teorema 2.1

Seja N um operador linear nilpotente de ordem n sobre um espaço vetorial V de dimensão finita. Seja ν_k o número de blocos $k \times k$ na forma canônica de N e $\delta_k = \dim(\text{Ker}(N^k))$, $k = 1, \dots, n$.

Então:

$$(1) \delta_m = \sum_{1 \leq k < m} k\nu_k + m \sum_{m \leq j \leq n} \nu_j, \quad m = 1, \dots, n.$$

(2) Temos

$$\begin{aligned} \nu_1 &= 2\delta_1 - \delta_2; \\ \nu_k &= -\delta_{k-1} + 2\delta_k - \delta_{k+1}, \quad 1 < k < n \\ \nu_n &= -\delta_{n-1} + \delta_n. \end{aligned}$$

Fazendo $\delta_0 = 0$ e $\delta_{n+1} = \delta_n$, podemos escrever as três relações acima na forma geral

$$\nu_k = -\delta_{k-1} + 2\delta_k - \delta_{k+1}$$

Prova.

Seja $\delta_1 = \dim(\text{Ker}(N))$. Como δ_1 é igual ao número de blocos, temos que

$$\delta_1 = \nu_1 + \dots + \nu_n,$$

provando (1) no caso $m = 1$.

Seja $\delta_2 = \dim(\text{Ker}(N^2))$. Como $\text{Ker}(N) \subset \text{Ker}(N^2)$, temos que cada bloco 1×1 contribui com uma dimensão ao $\text{Ker}(N^2)$.

Se $k \geq 2$, cada bloco $k \times k$ contribui com duas dimensões ao $\text{Ker}(N^2)$. De fato, se $\{w, N(w), \dots, N^{k-1}(w)\}$ é a parte da base $\mathcal{B}$ de V que dá origem a um bloco de ordem $k \times k$, temos, fazendo $w_j = N^{j-1}(w)$, $j = 1, \dots, k$, que

$$N^2(w_j) = N^2(N^{j-1}(w)) = N^{j+1}(w) = \mathbf{0},$$

se, e somente se, $j + 1 \geq k$, ou seja $j = k - 1$ ou $j = k$.

Logo, $\delta_2 = \nu_1 + 2\nu_2 + 2\nu_3 + \dots + 2\nu_n$, provando (1) no caso $m = 2$.

Caso geral.

Seja $\delta_k = \dim(\text{Ker}(N^k))$. Cada bloco $j \times j$, com $j \leq k - 1$, contribui com j dimensões ao $\text{Ker}(N^k)$, pois $\text{Ker}(N^j) \subset \text{Ker}(N^k)$ e x^j é o polinômio unitário associado a um bloco $j \times j$.

Se $j \geq k$, cada bloco $j \times j$ contribui com k dimensões ao $\text{Ker}(N^k)$. De fato, se $\{w, N(w), \dots, N^{j-1}(w)\}$ é a parte da base $\mathcal{B}$ de V que dá origem a um bloco de ordem $k \times k$, temos, fazendo $w_i = N^{i-1}(w)$, $i = 1, \dots, j$, que

$$N^k(w_i) = N^k(N^{i-1}(w)) = N^{k+i-1}(w) = \mathbf{0}.$$

se, e só se, $k + i - 1 \geq j$, ou seja $i \geq j - k + 1$.

Então, $N^k(w_i) = \mathbf{0}$ para k valores

$$j - (k - 1), j - (k - 2), \dots, j - (k - k)$$

de i . Provamos, assim, que

$$\delta_k = 1v_1 + 2v_2 + \dots + (k - 1)v_{k-1} + k(v_k + \dots + v_n).$$

Como

$$\delta_{k-1} = v_1 + 2v_2 + \dots + (k - 1)v_{k-1} + (k - 1)(v_k + \dots + v_n),$$

e

$$\delta_k = v_1 + 2v_2 + \dots + (k - 1)v_{k-1} + kv_k + k(v_{k+1} + \dots + v_n),$$

temos que

$$\delta_k - \delta_{k-1} = v_k + \dots + v_n.$$

Logo,

$$\begin{aligned} 2\delta_1 - \delta_2 &= \delta_1 - (\delta_2 - \delta_1) = v_1 + \dots + v_n - (v_2 + \dots + v_n) = v_1 \\ -\delta_{k-1} + 2\delta_k - \delta_{k+1} &= (\delta_k - \delta_{k-1}) - (\delta_{k+1} - \delta_k) \\ &= v_k + \dots + v_n - (v_{k+1} + \dots + v_n) \\ &= v_k, \quad \text{se } 1 < k < n \\ -\delta_{n-1} + \delta_n &= v_n. \end{aligned}$$



Exemplo 2.1

Seja T o operador linear sobre $\mathbb{R}^3$ que é representado em relação à base canônica pela matriz

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

Como

$$A^2 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix},$$

e

$$A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

temos que T é um operador nilpotente de ordem 3, com polinômio minimal x^3 .

Assim, a forma canônica de T possui somente um bloco de ordem 3×3 :

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

Ou seja, a matriz A é semelhante à matriz B . $\square$

Exemplo 2.2

Seja T o operador linear sobre $\mathbb{R}^4$ representado em relação à base canônica pela matriz

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 6 & 7 & 0 & 0 \\ 8 & 9 & 0 & 0 \end{pmatrix}.$$

Como $\delta_1 = \dim(\text{Ker}(T)) = 2$, temos que a matriz de T na forma canônica tem 2 blocos. Surgem, então, duas possibilidades:

- um bloco 3×3 e um bloco 1×1 ;
- dois blocos 2×2 .

Mas, como $A^2 = \mathcal{O}$, temos que $p_m = x^2$ é o polinômio minimal de T , ou seja, T é nilpotente de ordem 2.

Logo, a matriz B na forma canônica de T tem dois blocos 2×2 , ou seja,

$$B = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

sendo A semelhante a B . $\square$

Exemplo 2.3

Seja T o operador linear sobre $\mathbb{R}^5$ que é representado em relação à base canônica pela matriz

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 3 & 0 \end{pmatrix}.$$

Como $A^2 = 0$, temos que $p_m = x^2$ é o polinômio minimal de T e, portanto, T é nilpotente de ordem 2.

Além disso, $\delta_1 = \dim(\text{Ker}(N)) = 3$, ou seja, a matriz de T na forma canônica tem 3 blocos, sendo o primeiro deles uma matriz 2×2 .

Logo, a matriz B , forma canônica de A , tem dois blocos 2×2 e um bloco 1×1 , ou seja,

$$B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

sendo A semelhante a B $\square$

Exemplo 2.4

Determine todas as formas canônicas possíveis de uma matriz 10×10 A , cujo polinômio minimal é $p_m = x^3$ e cujo posto é 6. Determine, também, para cada caso possível, a dimensão de $\text{Ker}(A^2)$.

Como $\text{posto}(A) = 6$, temos que $\delta_1 = \dim(\text{Ker}(A)) = 4$, ou seja, a forma canônica de A tem 4 blocos.

Além disso, $\delta_3 = \dim(\text{Ker}(A^3)) = 10$, pois x^3 é o polinômio minimal de A .

Sendo

$$\begin{aligned} \nu_1 &= 2\delta_1 - \delta_2 = 8 - \delta_2 \geq 0 \\ \nu_2 &= -\delta_1 + 2\delta_2 - \delta_3 = 2\delta_2 - 14 \geq 0 \\ \nu_3 &= -\delta_2 + \delta_3 = -\delta_2 + 10 \geq 0. \end{aligned}$$

temos que $7 \leq \delta_2 \leq 8$.

Se $\delta_2 = \text{Ker}(N^2) = 7$, $\nu_1 = 1$, $\nu_2 = 0$ e $\nu_3 = 3$, ou seja, a matriz na forma canônica tem 3 blocos 3×3 e um bloco 1×1 , sendo, portanto,

$$\left(\begin{array}{ccc|ccc|ccc|c} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right).$$

Se $\delta_2 = \text{Ker}(A^2) = 8$, então $\nu_1 = 0$, $\nu_2 = 2$ e $\nu_3 = 2$, ou seja, a matriz na forma canônica tem dois blocos 3×3 e dois blocos 2×2 , sendo, portanto,

$$\left(\begin{array}{ccc|ccc|cc|cc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{array} \right).$$

□

Vamos, agora, mostrar um método para determinar uma base $\mathcal{B}$ de V de modo que $[N]_{\mathcal{B}}$ está na forma canônica.

Suponhamos que N seja um operador nilpotente de ordem k , ou seja, $N^k = \mathcal{O}$ e $N^{k-1} \neq \mathcal{O}$.

Afirmção: $N^{k-1}(V) \subsetneq N^{k-2}(V) \subsetneq \dots \subsetneq N(V) \subsetneq N^0(V) = V$.

De fato, temos que, para $j \leq k-1$, $N^{k-j}(V) \subset N^{k-j-1}(V)$.

Resta mostrar que $N^{k-j}(V) \neq N^{k-j-1}(V)$.

Seja

$$N : N^{k-j-1}(V) \longrightarrow N^{k-j}(N) \subset N^{k-j-1}(V).$$

Se

$$N^{k-j}(V) = N(N^{k-j-1}(V)) = N^{k-j-1}(V),$$

teríamos que $N|_{N^{k-j-1}(V)}$ seria sobrejetora e, portanto, injetora, que é uma contradição, conforme provaremos abaixo.

Seja $w \in V$ tal que $N^{k-1}(w) \neq \mathbf{0}$. Como

$$N^{k-1}(w) = N^{k-j-1}(N^j(w)) \in N^{k-j-1}(V) \text{ e } N(N^{k-1}(w)) = \mathbf{0},$$

temos que $N|_{N^{k-j-1}(V)}$ não é injetora.

Seja $\{u_1, \dots, u_{v_k}\}$ uma base de $N^{k-1}(V)$.

Então, existem $w_1^1, \dots, w_{v_k}^1 \in V$ tais que $N^{k-1}(w_i^1) = u_i$, $i = 1, \dots, v_k$.

Além disso,

$$\begin{aligned} &\{w_1^1, N(w_1^1), \dots, N^{k-1}(w_1^1)\} \cup \{w_2^1, N(w_2^1), \dots, N^{k-1}(w_2^1)\} \cup \dots \\ &\cup \{w_{v_k}^1, N(w_{v_k}^1), \dots, N^{k-1}(w_{v_k}^1)\}, \end{aligned}$$

é um conjunto LI, pois se

$$w = \sum_{i=1}^{v_k} a_i^1 w_i^1 + \sum_{i=1}^{v_k} a_i^2 N(w_i^1) + \dots + \sum_{i=1}^{v_k} a_i^k N^{k-1}(w_i^1) = \mathbf{0},$$

temos que $N^{k-1}(w) = \mathbf{0}$, ou seja,

$$\sum_{i=1}^{v_k} a_i^1 N^{k-1}(w_i^1) = \mathbf{0},$$

pois

$$N^{k-1} \left(\sum_{i=1}^{v_k} a_i^{j+1} N^j(w_i^1) \right) = \sum_{i=1}^{v_k} a_i^{j+1} N^{k+j-1}(w_i^1) = \mathbf{0},$$

já que $k+j-1 \geq k$, para $j \geq 1$.

Logo, $a_i^1 = 0$, $i = 1, \dots, v_k$, pois

$$\{N^{k-1}(w_1^1), \dots, N^{k-1}(w_{v_k}^1)\}$$

é um conjunto LI.

Aplicando N^{k-2} ao vetor

$$\sum_{i=1}^{v_k} a_i^2 N(w_i^1) + \dots + \sum_{i=1}^{v_k} a_i^k N^{k-1}(w_i^1) = \mathbf{0},$$

obtemos que $a_1^2 = \dots = a_{v_k}^2 = \mathbf{0}$.

Prosseguindo dessa maneira, chegaremos que $a_i^j = 0$, $i = 1, \dots, v_k$, $j = 1, \dots, k$.

• Os vetores do conjunto

$$\{N^{k-1}(w_1^1), \dots, N^{k-1}(w_{v_k}^1)\} \cup \{N^{k-2}(w_1^1), \dots, N^{k-2}(w_{v_k}^1)\}$$

são LI e pertencem ao subespaço $N^{k-2}(V)$.

Se esses vetores não geram $N^{k-2}(V)$, existem vetores $u_1, \dots, u_{v_{k-1}}$ tais que $N^{k-2}(u_1), \dots, N^{k-2}(u_{v_{k-1}})$ completam a base acima.

Afirmção: Podemos substituir cada u_i por um vetor w_i^2 de modo que os vetores

$$N^{k-2}(w_1^2), \dots, N^{k-2}(w_{v_{k-1}}^2)$$

completem a base acima e, além disso, $N^{k-1}(w_i^2) = \mathbf{0}$, $i = 1, \dots, v_{k-1}$.

De fato, como

$$N(N^{k-2}(u_i)) = N^{k-1}(u_i) \in N^{k-1}(V),$$

existem escalares $\lambda_1^i, \dots, \lambda_{v_k}^i$, tais que

$$N(N^{k-2}(u_i)) = N^{k-1}(u_i) = \lambda_1^i N^{k-1}(w_1^1) + \dots + \lambda_{v_k}^i N^{k-1}(w_{v_k}^1).$$

Então

$$N(N^{k-2}(u_i - \lambda_1^i w_1^1 - \dots - \lambda_{v_k}^i w_{v_k}^1)) = \mathbf{0}.$$

Fazendo

$$w_i^2 = u_i - \lambda_1^i w_1^1 - \dots - \lambda_{v_k}^i w_{v_k}^1,$$

temos que $N^{k-1}(w_i^2) = \mathbf{0}$ e

$$N^{k-2}(w_i^2) = N^{k-2}(u_i) - \lambda_1^i N^{k-2}(w_1^1) - \dots - \lambda_{v_k}^i N^{k-2}(w_{v_k}^1).$$

Logo, o subespaço gerado por

$$\begin{aligned} &\{N^{k-1}(w_1^1), \dots, N^{k-1}(w_{v_k}^1)\} \cup \{N^{k-2}(w_1^1), \dots, N^{k-2}(w_{v_k}^1)\} \\ &\cup \{N^{k-2}(w_1^2), \dots, N^{k-2}(w_{v_{k-1}}^2)\} \end{aligned}$$

é igual ao subespaço $N^{k-2}(V)$, gerado por

$$\begin{aligned} &\{N^{k-1}(w_1^1), \dots, N^{k-1}(w_{v_k}^1)\} \cup \{N^{k-2}(w_1^1), \dots, N^{k-2}(w_{v_k}^1)\} \\ &\cup \{N^{k-2}(u_1), \dots, N^{k-2}(u_{v_{k-1}})\}. \end{aligned}$$

Além disso,

$$\{N^{k-1}(w_1^1), \dots, w_1^1\} \cup \dots \cup \{N^{k-1}(w_{v_k}^1), \dots, w_{v_k}^1\} \\ \cup \{N^{k-2}(w_1^2), \dots, w_1^2\} \cup \dots \cup \{N^{k-2}(w_{v_{k-1}}^2), \dots, w_{v_{k-1}}^2\}$$

é um conjunto LI, pois uma combinação linear desses vetores se escreve na forma

$$N^{k-1}(v_0) + N^{k-2}(v_1) + \dots + v_{k-1} + N^{k-2}(w_0) \\ + N^{k-3}(w_1) + \dots + w_{k-2} = \mathbf{0}, \quad (I)$$

onde

$$v_i = a_1^i w_1^1 + \dots + a_{v_k}^i w_{v_k}^1, \quad i = 0, 1, \dots, k-1,$$

e

$$w_j = b_1^j w_1^2 + \dots + b_{v_{k-1}}^j w_{v_{k-1}}^2, \quad j = 0, 1, \dots, k-2.$$

Aplicando N^{k-1} à expressão (I), obtemos que $N^{k-1}(v_{k-1}) = \mathbf{0}$, pois $N^{k-1}(w_j) = \mathbf{0}$, $j = 0, 1, \dots, k-2$, e

$$N^{k-1}(N^{k-(i+1)}(v_i)) = N^{2k-i-2}(v_i) = \mathbf{0}, \quad i = 0, 1, \dots, k-2,$$

já que $k-1 + k-(i+1) = k+k-2-i \geq k$.

Como

$$\mathbf{0} = N^{k-1}(v_{k-1}) = a_1^{k-1} N^{k-1}(w_1^1) + \dots + a_{v_k}^{k-1} N^{k-2}(w_{v_k}^1)$$

e

$$\{N^{k-1}(w_1^1), \dots, N^{k-1}(w_{v_k}^1)\},$$

é LI, temos que $a_i^{k-1} = 0$, $i = 0, 1, \dots, k-1$, ou seja, $v_{k-1} = \mathbf{0}$.

Assim,

$$N^{k-1}(v_0) + \dots + N(v_{k-2}) + N^{k-2}(w_0) + \dots + w_{k-2} = \mathbf{0}. \quad (II)$$

Aplicando N^{k-2} à expressão (II), obtemos que

$$N^{k-1}(v_{k-2}) + N^{k-2}(w_{k-2}) = \mathbf{0},$$

ou seja,

$$\mathbf{0} = a_1^{k-2} N^{k-1}(w_1^1) + \dots + a_{v_k}^{k-2} N^{k-1}(w_{v_k}^1) + b_1^{k-2} N^{k-2}(w_1^2) \\ + \dots + b_{v_{k-1}}^{k-2} N^{k-2}(w_{v_{k-1}}^2).$$

Como

$$\{N^{k-1}(w_1^1), \dots, N^{k-1}(w_{v_k}^1), N^{k-2}(w_1^2), \dots, N^{k-2}(w_{v_{k-1}}^2)\}$$

é um conjunto LI, temos que

$$a_1^{k-2} = \dots = a_{v_k}^{k-2} = b_1^{k-2} = \dots = b_{v_{k-1}}^{k-2} = 0.$$

Então, $v_{k-2} = w_{k-2} = \mathbf{0}$.

Prosseguindo dessa maneira, chegamos que

$$N^{k-1}(v_0) + N^{k-2}(w_0) = \mathbf{0}.$$

Logo,

$$a_1^0 = \dots = a_{v_k}^0 = b_1^0 = \dots = b_{v_{k-1}}^0 = 0,$$

pois

$$\{N^{k-1}(w_1^1), \dots, N^{k-1}(w_{v_k}^1)\} \cup \{N^{k-2}(w_1^2), \dots, N^{k-2}(w_{v_{k-1}}^2)\},$$

é um conjunto LI.

• Os vetores do conjunto

$$\begin{aligned} &\{N^{k-1}(w_1^1), \dots, N^{k-1}(w_{v_k}^1)\} \cup \{N^{k-2}(w_1^1), \dots, N^{k-2}(w_{v_k}^1)\} \\ &\cup \{N^{k-3}(w_1^1), \dots, N^{k-3}(w_{v_k}^1)\} \cup \{N^{k-2}(w_1^2), \dots, N^{k-2}(w_{v_{k-1}}^2)\} \\ &\cup \{N^{k-3}(w_1^2), \dots, N^{k-3}(w_{v_{k-1}}^2)\} \end{aligned}$$

são LI e pertencem ao subespaço $N^{k-3}(V)$.

Se esses vetores não geram $N^{k-3}(V)$, existem vetores $w_1^3, \dots, w_{v_{k-2}}^3$ tais que os vetores $N^{k-3}(w_1^3), \dots, N^{k-3}(w_{v_{k-2}}^3)$ completam a base acima.

Podemos supor, também, como fizemos anteriormente, que

$$N(N^{k-3}(w_i^3)) = N^{k-2}(w_i^3) = \mathbf{0}, \quad i = 1, \dots, v_{k-2}.$$

Continuando dessa maneira até chegarmos a $V = N^0(V) \supsetneq N(V)$, obteremos a base procurada.

$$\begin{array}{ccccccc}
 V = N^0(V) & \longrightarrow & w_1^1, \dots, w_{\gamma_k}^1, & w_1^2, \dots, w_{\gamma_{k-1}}^2, & w_1^3, \dots, w_{\gamma_{k-2}}^3, & \dots, & w_1^{k-1}, \dots, w_{\gamma_2}^{k-1}, & [w_1^k, \dots, w_{\gamma_1}^k] \\
 \mathcal{N} & & & & & & & \downarrow N & \downarrow N \\
 \mathcal{N}(V) & \longrightarrow & N(w_1^1), \dots, N(w_{\gamma_k}^1), & N(w_1^2), \dots, N(w_{\gamma_{k-1}}^2), & N(w_1^3), \dots, N(w_{\gamma_{k-2}}^3), & \dots, & [N(w_1^{k-1}), \dots, N(w_{\gamma_2}^{k-1})] & 0 \\
 \mathcal{N} & & & & & & & \downarrow N & \downarrow N \\
 \vdots & & \vdots & \vdots & \vdots & & 0 & 0 & 0 \\
 \vdots & & \vdots & \vdots & \vdots & & & & \\
 \mathcal{N}^{k-3}(V) & \longrightarrow & N^{k-3}(w_1^1), \dots, N^{k-3}(w_{\gamma_k}^1), & N^{k-3}(w_1^2), \dots, N^{k-3}(w_{\gamma_{k-1}}^2), & [N^{k-3}(w_1^3), \dots, N^{k-3}(w_{\gamma_{k-2}}^3)] & & & & \\
 \mathcal{N} & & & & & & & \downarrow N & \downarrow N \\
 \mathcal{N}^{k-2}(V) & \longrightarrow & N^{k-2}(w_1^1), \dots, N^{k-2}(w_{\gamma_k}^1), & [N^{k-2}(w_1^2), \dots, N^{k-2}(w_{\gamma_{k-1}}^2)] & 0 & & 0 & & \\
 \mathcal{N} & & & & & & & \downarrow N & \downarrow N \\
 \mathcal{N}^{k-1}(V) & \longrightarrow & N^{k-1}(w_1^1), \dots, N^{k-1}(w_{\gamma_k}^1), & 0 & 0 & & 0 & &
 \end{array}$$

Note que:

Os vetores $w_1^1, \dots, w_{\gamma_k}^1$ determinam os γ_k subespaços cíclicos $Z(w_1^1, N), \dots, Z(w_{\gamma_k}^1, N)$ de dimensão k e, portanto, os blocos $k \times k$ da forma canônica de Jordan de N .

De modo geral, os vetores $w_1^j, \dots, w_{\gamma_{k-(j-1)}}^j, j = 1, \dots, k$, determinam os $\gamma_{k-(j-1)}$ subespaços cíclicos $Z(w_1^j, N), \dots, Z(w_{\gamma_{k-(j-1)}}^j, N)$ de dimensão $k - (j - 1)$ e, portanto, os blocos $(k - (j - 1)) \times (k - (j - 1))$ da forma canônica de Jordan de N .

Exemplo 2.5

Considere o operador N linear sobre $\mathbb{R}^5$ que é representado em relação à base canônica pela matriz

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 3 & 0 \end{pmatrix}.$$

Como $A^2 = \mathcal{O}$, temos que N é um operador nilpotente de ordem 2.

Sendo $N(e_1) = e_2$, $N(e_2) = \mathbf{0}$, $N(e_3) = 2e_5$, $N(e_4) = 3e_5$ e $N(e_5) = \mathbf{0}$, o espaço $N(\mathbb{R}^5)$ é gerado por e_2 e e_5 e o conjunto $\{e_1, N(e_1), e_3, N(e_3)\}$ é LI.

Como $\dim(N^0(\mathbb{R}^5)) = 5$, devemos determinar um vetor $w_3 \in \mathbb{R}^5$ que complete esse conjunto LI a uma base de $\mathbb{R}^5$ e tal que $N(w_3) = \mathbf{0}$.

Sendo $\text{Ker}(N) = \{(x, y, z, w, t) \mid x = 0, \text{ e } 2z + 3w = 0\}$, temos que, para $w_3 = (0, 0, 3, -2, 0)$, $N(w_3) = \mathbf{0}$ e $\mathcal{B} = \{w_1, N(w_1), w_2, N(w_2), w_3\}$ é uma base de $\mathbb{R}^5$, onde $w_1 = e_1$ e $w_2 = e_3$.

Logo,

$$\mathbb{R}^5 = Z(w_1, N) \oplus Z(w_2, N) \oplus Z(w_3, N),$$

é uma decomposição cíclica de N , onde $p_1 = p_2 = p_m = x^2$ e $p_3 = x$ são os N -anuladores de w_1 , w_2 e w_3 , respectivamente.

Temos, também, que

$$[N]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & & & \\ 1 & 0 & & & \\ & & 0 & 0 & \\ & & 1 & 0 & \\ & & & & 0 \end{pmatrix},$$

é a matriz do operador N na forma canônica racional. $\square$

Exemplo 2.6

Seja N o operador linear sobre $\mathbb{R}^5$ que é representado em relação à base canônica pela matriz

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

Como

$$A^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{e} \quad A^4 = \mathcal{O},$$

temos que N é um operador nilpotente de ordem 4.

Como $\{2e_3 = N^3(e_1)\}$ é uma base de $N^3(\mathbb{R}^5)$, temos que $\{N^3(e_1), N^2(e_1) = 2e_4\}$ é um conjunto LI contido em $N^2(\mathbb{R}^5)$.

Sendo $\dim(N^2(\mathbb{R}^5)) = 2$, $\{N^3(e_1), N^2(e_1)\}$ é uma base de $N^2(\mathbb{R}^5)$.

Temos, também, que o conjunto LI $\{N^3(e_1), N^2(e_1), N(e_1) = e_2\}$ é uma base de $N(\mathbb{R}^5)$, pois $\dim(N(\mathbb{R}^5)) = 3$.

Como $\{N^3(e_1), N^2(e_1), N(e_1), e_1\}$ é LI e $\dim(\mathbb{R}^5) = 5$, devemos determinar $w_2 \in \mathbb{R}^5$ que completa a base e satisfaz $N(w_2) = \mathbf{0}$.

Sendo $\text{Ker}(N) = \{(x, y, z, w, t) \mid x = y = w = 0\}$, temos que para $w_2 = e_5$, $N(w_2) = \mathbf{0}$ e $\mathcal{B} = \{e_1, N(e_1), N^2(e_1), N^3(e_1), w_2\}$ é uma base de $\mathbb{R}^5$.

Assim, $\mathbb{R}^5 = Z(w_1, N) \oplus Z(w_2, N)$, onde $w_1 = e_1$, é uma decomposição cíclica de N . Além disso, $p_1 = p_m = x^4$ e $p_2 = x$ são os N -anuladores de w_1 e w_2 , respectivamente.

A matriz de N em relação à base $\mathcal{B}$

$$[N]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

é a matriz do operador N na forma canônica racional. $\square$

Exemplo 2.7

Seja $V = \mathcal{P}^6(\mathbb{R}) = \{f \in \mathbb{R}[x] \mid \text{grau}(f) \leq 6\}$ e seja $N = D^2$ o operador linear sobre $\mathcal{P}^6(\mathbb{R})$ dado pela **derivação de ordem 2**.

Temos que $N(V) = \mathcal{P}^4(\mathbb{R})$ (polinômios reais de grau ≤ 4);

$N^2(V) = D^4(\mathcal{P}^6(\mathbb{R})) = \mathcal{P}^2(\mathbb{R})$ (polinômios reais de grau ≤ 2);

$N^3(V) = D^6(\mathcal{P}^6(\mathbb{R})) = \{f \in \mathbb{R}[x] \mid \text{grau}(f) = 0 \quad \text{ou} \quad f \equiv \mathcal{O}\}$

e $N^4(V) = D^8(\mathcal{P}^6(\mathbb{R})) = \{\mathcal{O}\}$.

Logo, N é um operador nilpotente de ordem 4 e $\{N^3(x^6) = D^6(x^6) = 6!\}$ é uma base de $N^3(V)$.

Assim, $\{N^3(x^6), N^2(x^6) = D^4(x^6) = 360x^2\}$ é um conjunto LI de vetores de $N^2(V)$. Como $N^2(V) = \mathcal{P}^2(\mathbb{R})$ tem dimensão 3, devemos achar um polinômio w_2 de modo que $N(N^2(w_2)) = D^6(w_2) = 0$ e $\{N^3(x^6), N^2(x^6), N^2(w_2)\}$ seja uma base de $N^2(V)$.

Podemos tomar $w_2 = x^5$. Nesse caso, $N^2(x^5) = D^4(x^5) = 120x$.

Temos que $\{N^3(x^6), N^2(x^6), N(x^6), N^2(x^5), N(x^5) = D^2(x^5) = 20x^3\}$ é um conjunto LI de vetores de $N(V) = \mathcal{P}^4(\mathbb{R})$.

Como $\dim(N(V)) = 5$, temos que

$$\{N^3(x^6), N^2(x^6), N(x^6), N^2(x^5), N(x^5)\},$$

é uma base de $N(V)$.

Como $\mathcal{B} = \{x^6, N(x^6), N^2(x^6), N^3(x^6), x^5, N(x^5), N^2(x^5)\}$ é um conjunto LI de vetores de $N^0(V)$ e $\dim(V) = 7$, temos que esse conjunto é uma base de V .

Logo, $\mathcal{P}^6(\mathbb{R}) = Z(w_1, N) \oplus Z(w_2, N)$, $w_1 = x^6$ e $w_2 = x^5$, é uma decomposição cíclica de $N = D^2$, onde $p_1 = p_m = x^4$ e $p_2 = x^3$ são os N -anuladores de w_1 e w_2 , respectivamente. Além disso,

$$[N]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

é a matriz de N na forma canônica racional. $\square$

Exemplo 2.8

Determine os blocos da forma canônica racional de um operador nilpotente N de ordem 5 sobre o espaço vetorial V , sabendo que

$$\dim(N^4(V)) = 3, \dim(N^2(V)) = 11, \dim(N(V)) = 16 \text{ e } \dim(V) = 23.$$

Como $\dim(N^4(V)) = 3$, existem vetores w_1^1, w_2^1 e w_3^1 tais que

$$\mathcal{B}_1 = \{N^4(w_1^1), N^4(w_2^1), N^4(w_3^1)\}$$

seja uma base de $N^4(V)$.

Com isso, já sabemos que a matriz de N na forma canônica racional tem 3 blocos 5×5 .

Como

$$\mathcal{B}_2 = \{N^4(w_1^1), N^3(w_1^1), N^4(w_2^1), N^3(w_2^1), N^4(w_3^1), N^3(w_3^1)\}$$

é um conjunto de vetores LI de $N^3(V)$ e $\dim(N^3(V)) = 6$, temos que esse conjunto é uma base de $N^3(V)$.

Como

$$\begin{aligned} \mathcal{B}_3 = \{ & N^4(w_1^1), N^3(w_1^1), N^2(w_1^1), N^4(w_2^1), N^3(w_2^1), N^2(w_2^1), \\ & N^4(w_3^1), N^3(w_3^1), N^2(w_3^1) \} \end{aligned}$$

é um conjunto de vetores LI de $N^2(V)$ e $\dim(N^2(V)) = 11$, existem dois vetores w_1^3 e w_2^3 tais que

$$N(N^2(w_1^3)) = N^3(w_1^3) = \mathbf{0}, \quad N(N^2(w_2^3)) = N^3(w_2^3) = \mathbf{0}$$

e $\mathcal{B}_3 \cup \{N^2(w_1^3), N^2(w_2^3)\}$ é uma base de $N^2(V)$. Com isso, temos que a matriz de N na forma canônica racional tem dois blocos 3×3 .

Como

$$\begin{aligned} \mathcal{B}_4 = \{ & N^4(w_1^1), N^3(w_1^1), N^2(w_1^1), N(w_1^1), N^4(w_2^1), N^3(w_2^1), N^2(w_2^1), N(w_2^1), \\ & N^4(w_3^1), N^3(w_3^1), N^2(w_3^1), N(w_3^1), N^2(w_1^3), N(w_1^3), N^2(w_2^3), N(w_2^3) \} \end{aligned}$$

é um conjunto LI de vetores de $N(V)$ e $\dim(N(V)) = 16$, temos que $\mathcal{B}_4$ é uma base de $N(V)$.

Como

$$\mathcal{B}_5 = \mathcal{B}_4 \cup \{w_1^1, w_2^1, w_3^1, w_1^3, w_2^3\}$$

é um conjunto LI de vetores de $N^0(V) = V$ e $\dim(V) = 23$, temos que existem dois vetores w_1^5 e w_2^5 tais que $N(w_1^5) = N(w_2^5) = \mathbf{0}$ e $\mathcal{B}_5 \cup \{w_1^5, w_2^5\}$ é uma base de V .

Logo, a matriz de N na forma canônica racional tem 2 blocos 1×1 .

Outra maneira de resolver:

Seja $\delta_k = \dim(\text{Ker}(N^k))$ e seja ν_k o número de blocos $k \times k$ na forma canônica racional de N .

Como $\dim(V) = 23$, $\dim(N(V)) = 16$, $\dim(N^2(V)) = 11$, $\dim(N^3(V)) = 6$, $\dim(N^4(V)) = 3$ e $\dim(N^5(V)) = 0$, temos que

$$\begin{aligned} \delta_1 &= 23 - 16 = 7, & \delta_2 &= 23 - 11 = 12, & \delta_3 &= 23 - 6 = 17, \\ \delta_4 &= 23 - 3 = 20, & \text{e } \delta_5 &= 23 - 0 = 23 & (\delta_6 &= 23). \end{aligned}$$

Então,

$$\begin{aligned} \gamma_1 &= 2\delta_1 - \delta_2 = 2 \times 7 - 12 = 2, \\ \gamma_2 &= -\delta_1 + 2\delta_2 - \delta_3 = -7 + 2 \times 12 - 17 = 0, \\ \gamma_3 &= -\delta_2 + 2\delta_3 - \delta_4 = -12 + 2 \times 17 - 20 = 2, \\ \gamma_4 &= -\delta_3 + 2\delta_4 - \delta_5 = -17 + 2 \times 20 - 23 = 0, \\ \gamma_5 &= -\delta_4 + 2\delta_5 - \delta_6 = -20 + 2 \times 23 - 23 = 3. \end{aligned}$$

Assim, a forma canônica de N tem dois blocos 1×1 , dois blocos 3×3 e três blocos 5×5 . $\square$

3. Forma Canônica de Jordan

Seja T um operador linear sobre um espaço vetorial V de dimensão finita n . Suponhamos que o polinômio característico de T se decomponha como um produto de fatores primos lineares

$$p_c = (x - c_1)^{d_1} \dots (x - c_k)^{d_k},$$

onde $c_1, \dots, c_k$ são os autovalores distintos de T e $d_i \geq 1$, $i = 1, \dots, k$. Então, o polinômio minimal de T é

$$p_m = (x - c_1)^{r_1} \dots (x - c_k)^{r_k},$$

onde $1 \leq r_i \leq d_i$, $i = 1, \dots, k$.

Se $W_i = \text{Ker}(T - c_i I)^{r_i}$, o Teorema da Decomposição Primária nos diz que:

- $V = W_1 \oplus \dots \oplus W_k$,
- $\dim(W_i) = d_i$,
- $(x - c_i)^{r_i}$ é o polinômio minimal de $T_i = T|_{W_i}$.

Seja $N_i = T_i - c_i I$.

Então, como já provamos anteriormente, N_i é nilpotente e x^{r_i} é o seu polinômio minimal.

Seja $\mathcal{B}_i$ uma base de W_i tal que $[N_i]_{\mathcal{B}_i}$ está na forma canônica racional. Como $T_i = c_i I + N_i$, $[T_i]_{\mathcal{B}_i} = c_i I + [N_i]_{\mathcal{B}_i}$, onde I é a matriz identidade $d_i \times d_i$.

Seja $V_i = \text{Ker}(T - c_i I)$ o autoespaço de T associado ao autovalor c_i . Como $\dim(\text{Ker}(N_i)) = \dim(\text{Ker}(T_i - c_i I)) = \dim(\text{Ker}(T - c_i I))$, pois $\text{Ker}(T - c_i I) \subset W_i$, temos que se $\dim(\text{Ker}(T - c_i I)) = n_i$, onde $n_i \leq d_i$, a matriz na forma racional de N_i tem n_i , a matriz na forma racional de T_i tem n_i blocos da forma

$$\begin{pmatrix} c_i & 0 & 0 & \cdots & 0 & 0 \\ 1 & c_i & 0 & \cdots & 0 & 0 \\ 0 & 1 & c_i & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & c_i & 0 \\ 0 & 0 & 0 & \cdots & 1 & c_i \end{pmatrix},$$

sendo que o primeiro bloco é uma matriz $r_i \times r_i$.

Ou seja,

$$[T_i]_{\mathcal{B}_i} = \begin{pmatrix} J_1^i & & \\ & J_2^i & \\ & & \ddots \\ & & & J_{n_i}^i \end{pmatrix}_{d_i \times d_i},$$

onde cada J_j^i é da forma

$$J_j^i = \begin{pmatrix} c_i & 0 & 0 & \cdots & 0 & 0 \\ 1 & c_i & 0 & \cdots & 0 & 0 \\ 0 & 1 & c_i & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & c_i & 0 \\ 0 & 0 & 0 & \cdots & 1 & c_i \end{pmatrix}.$$

Definição 3.1

Uma *matriz elementar de Jordan com autovalor c* é uma matriz da forma

$$J = \begin{pmatrix} c & 0 & 0 & \cdots & 0 & 0 \\ 1 & c & 0 & \cdots & 0 & 0 \\ 0 & 1 & c & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & c & 0 \\ 0 & 0 & 0 & \cdots & 1 & c \end{pmatrix}.$$

Seja $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B}_k$. Então,

$$[T]_{\mathcal{B}} = \begin{pmatrix} A_1 & & \\ & A_2 & \\ & & \ddots \\ & & & A_k \end{pmatrix}_{n \times n},$$

onde cada A_i é uma matriz $d_i \times d_i$ da forma

$$A_i = [T_i]_{\mathcal{B}_i} = \begin{pmatrix} J_1^i & & \\ & J_2^i & \\ & & \ddots \\ & & & J_{n_i}^i \end{pmatrix}_{d_i \times d_i},$$

sendo cada J_j^i uma matriz elementar de Jordan com autovalor c_i . Além disto, dentro de cada A_i , as dimensões das matrizes J_j^i diminuem a medida que j aumenta.

Dizemos que uma matriz que satisfaz as condições acima está sob a **forma canônica de Jordan**.

Acabamos de mostrar que se T é um operador cujo **polinômio característico se decompõe em um produto de fatores lineares sobre o corpo dos escalares**, então existe uma base ordenada de V em relação à qual T é representado por uma matriz que está sob a forma canônica de Jordan.

Unicidade: Se duas matrizes semelhantes estão sob a forma canônica de Jordan, então elas podem diferir apenas quanto à ordem dos autovalores.

Seja $\mathcal{B}'$ uma base de V tal que $[T]_{\mathcal{B}'} = \tilde{A}$ também esteja na forma canônica de Jordan.

Seja

$$[T]_{\mathcal{B}'} = \tilde{A} = \begin{pmatrix} \tilde{A}_1 & & \\ & \tilde{A}_2 & \\ & & \ddots \\ & & & \tilde{A}_s \end{pmatrix}_{n \times n},$$

onde cada matriz $\tilde{A}_i$ é da forma

$$\tilde{A}_i = \begin{pmatrix} \tilde{J}_1^i & & \\ & \tilde{J}_2^i & \\ & & \ddots \\ & & & \tilde{J}_{s_i}^i \end{pmatrix}_{\ell_i \times \ell_i},$$

sendo cada $\tilde{J}_j^i$ uma matriz elementar de Jordan com autovalor λ_i , ($\lambda_i \neq \lambda_j$, para $i \neq j$).

Como $\tilde{A}$ é uma matriz triangular inferior, seu polinômio característico é da forma

$$\tilde{p}_c = (x - \lambda_1)^{\ell_1} \dots (x - \lambda_s)^{\ell_s}.$$

Sendo A e $\tilde{A}$ semelhantes, o polinômio característico de A

$$p_c = (x - c_1)^{d_1} \dots (x - c_k)^{\ell_s},$$

é, também, o polinômio característico de $\tilde{A}$.

Logo, $s = k$, $\{c_1, \dots, c_k\} = \{\lambda_1, \dots, \lambda_k\}$ e, se $c_i = \lambda_j$, então $d_i = \ell_j$.

Seja $\{c_{\sigma_1}, \dots, c_{\sigma_k}\}$ uma reordenação dos escalares $\{c_1, \dots, c_k\}$ tal que $\lambda_i = c_{\sigma_i}$, $i = 1, \dots, k$.

Vamos mostrar que $\tilde{A}_i = A_{\sigma_i}$, $i = 1, \dots, k$.

Seja $B' = \{v_1^1, \dots, v_{\ell_1}^1, v_1^2, \dots, v_{\ell_2}^2, \dots, v_1^k, \dots, v_{\ell_k}^k\}$ e seja $\tilde{W}_i$ o subespaço gerado por $B'_i = \{v_1^i, \dots, v_{\ell_i}^i\}$, $i = 1, \dots, k$. Então, cada subespaço $\tilde{W}_i$ é T -invariante devido à forma da matriz $\tilde{A} = [T]_{B'}$; $[T|_{\tilde{W}_i}]_{B'_i} = \tilde{A}_i$ e

$$V = \tilde{W}_1 \oplus \dots \oplus \tilde{W}_k.$$

Afirmção: $\tilde{W}_i = \text{Ker}(T - \lambda_i I)^n$, $i = 1, \dots, k$.

De fato,

$$[T - \lambda_i I]_{B'} = \tilde{A} - \lambda_i I = \begin{pmatrix} \tilde{A}_1 - \lambda_i I & & & \\ & \ddots & & \\ & & \tilde{A}_i - \lambda_i I & \\ & & & \ddots \\ & & & & \tilde{A}_k - \lambda_i I \end{pmatrix},$$

onde $\tilde{A}_i - \lambda_i I$ é uma matriz nilpotente de ordem menor ou igual a $\ell_i = d_{\sigma_i}$, e $\tilde{A}_j - \lambda_i I$ é invertível se $j \neq i$, pois $\tilde{A}_j - \lambda_i I = (\lambda_j - \lambda_i)I + N_j$, sendo N_j nilpotente e $\lambda_j \neq \lambda_i$. Logo,

$$[(T - \lambda_i I)^n]_{B'} = (\tilde{A} - \lambda_i I)^n = \begin{pmatrix} (\tilde{A}_1 - \lambda_i I)^n & & & \\ & \ddots & & \\ & & (\tilde{A}_i - \lambda_i I)^n & \\ & & & \ddots \\ & & & & (\tilde{A}_k - \lambda_i I)^n \end{pmatrix}.$$

Na matriz ao lado, note que $(\tilde{A}_i - \lambda_i I)^n \equiv \mathcal{O}$.

Provamos, assim, que todo vetor da base $\mathcal{B}'_i$ pertence ao núcleo de $(T - \lambda_i I)^n$, ou seja, $\tilde{W}_i \subset \text{Ker}((T - \lambda_i I)^n)$ e que $\dim(\text{Ker}((T - \lambda_i I)^n)) = \ell_i$. Mas como $\dim(\tilde{W}_i) = \ell_i$, temos que $\tilde{W}_i = \text{Ker}((T - \lambda_i I)^n)$.

Fazendo o mesmo com os subespaços T -invariantes W_i obtidos da base $\mathcal{B}$, vemos que $W_j = \text{Ker}((T - c_j I)^n)$, $j = 1, \dots, k$. Como $\lambda_i = c_{\sigma_i}$, temos que

$$\tilde{W}_i = \text{Ker}((T - \lambda_i I)^n) = \text{Ker}((T - c_{\sigma_i} I)^n) = W_{\sigma_i}.$$

Logo,

$$[(T - c_{\sigma_i} I)|_{\tilde{W}_i}]_{\mathcal{B}_{\sigma_i}} = A_{\sigma_i} - c_{\sigma_i} I,$$

e

$$[(T - c_{\sigma_i} I)|_{\tilde{W}_i}]_{\mathcal{B}'_i} = \tilde{A}_i - c_{\sigma_i} I = \tilde{A}_i - \lambda_i I,$$

estão na forma canônica racional. Pela unicidade, temos $\tilde{A}_i = A_{\sigma_i}$.

Observação 3.1

(1) Seja $\mathbb{K}$ um corpo algebricamente fechado e T um operador linear sobre um espaço vetorial V de dimensão finita sobre $\mathbb{K}$. Então existe uma base $\mathcal{B}$ tal que $[T]_{\mathcal{B}}$ está na forma canônica de Jordan.

(2) O operador T é diagonalizável se, e somente se, $n_i = d_i$, ou seja, $\dim(\text{Ker}(T - c_i I))$ é a multiplicidade de c_i como raiz de p_c , para todo $i = 1, \dots, k$.

(3) Se o operador T é nilpotente, a forma canônica de Jordan de T é igual a sua forma racional.

Exemplo 3.1

Seja T um operador linear sobre $\mathbb{C}^2$. Então o polinômio característico de T pode ser $(x - c_1)(x - c_2)$, com $c_1 \neq c_2$, ou $(x - c)^2$.

No primeiro caso, T é diagonalizável e é representado em relação a alguma base ordenada $\mathcal{B}$ por

$$[T]_{\mathcal{B}} = \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix}.$$

No segundo caso, o polinômio minimal de T pode ser $x - c$ ou $(x - c)^2$.

Se $p_m = x - c$, temos que $T = cI$ e

$$[T]_{\mathcal{B}} = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix},$$

para qualquer base $\mathcal{B}$ de $\mathbb{C}^2$.

Se $p_m = (x - c)^2$, existe uma base $\mathcal{B}$ de $\mathbb{C}^2$ tal que

$$[T]_{\mathcal{B}} = \begin{pmatrix} c & 0 \\ 1 & c \end{pmatrix}.$$

Assim, toda matriz 2×2 sobre o corpo $\mathbb{C}$ é semelhante a uma matriz de um dos dois tipos

$$\begin{pmatrix} c & 0 \\ 1 & c \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix},$$

podendo ser $c_1 \neq c_2$ ou $c_1 = c_2$. $\square$

Exemplo 3.2

Seja A a matriz 3×3 com entradas em $\mathbb{C}$ dada por

$$A = \begin{pmatrix} 2 & 0 & 0 \\ a & 2 & 0 \\ b & c & -1 \end{pmatrix}.$$

O polinômio característico de A é $(x - 2)^2(x + 1)$.

Se o polinômio minimal de A é $(x - 2)^2(x + 1)$, então A é semelhante à matriz

$$\tilde{A} = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

na forma canônica de Jordan.

Se $p_m = (x - 2)(x - 1)$, A é diagonalizável e, portanto, semelhante à matriz

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

Como

$$(A - 2I)(A + I) = \begin{pmatrix} 0 & 0 & 0 \\ 3a & 0 & 0 \\ ac & 0 & 0 \end{pmatrix},$$

temos que A é semelhante a uma matriz diagonal se, e somente se, $a = 0$.

$\square$

Exemplo 3.3

Seja A a matriz 4×4 dada por

$$\tilde{A} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & a & 2 \end{pmatrix}.$$

O polinômio característico de A é $(x-2)^4$ e seu polinômio minimal é $(x-2)^2$ pois A é formada por dois blocos 2×2 .

O primeiro bloco da matriz na forma canônica semelhante a A tem tamanho 2×2 , e o número de blocos é $\dim(\text{Ker}(A - 2I))$.

Como

$$A - 2I = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & a & 0 \end{pmatrix},$$

$$\text{temos que } \dim(\text{Ker}(A - 2I)) = \begin{cases} 2, & \text{se } a \neq 0, \\ 3, & \text{se } a = 0. \end{cases}$$

Logo, A é semelhante a

$$\tilde{A} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix}, \quad \text{se } a \neq 0,$$

e é semelhante a

$$\tilde{A} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}, \quad \text{se } a = 0.$$

Observe que essas matrizes possuem polinômios característicos e minimais iguais, mas não são semelhantes, pois possuem formas canônicas de Jordan diferentes. $\square$

Exemplo 3.4

Seja A a matriz 3×3 dada por

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

Então,

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

Logo, $p_m = p_c$ e A é diagonalizável, ou seja, A é semelhante à matriz

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

□

Exemplo 3.5

Seja A a matriz 4×4 dada por

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 2 & -1 & 1 & 0 \\ 2 & 1 & 1 & 3 \end{pmatrix}.$$

Como $p_c = (x-1)(x-2)(x-3)^2$ e $\dim(\text{Ker}(A-3I)) = 1$, pois

$$A - 3I = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 2 & -1 & -2 & 0 \\ 2 & 1 & 1 & 0 \end{pmatrix},$$

temos que $p_m = p_c = (x-1)(x-2)(x-3)^2$.

Logo, a matriz na forma canônica de Jordan semelhante a A é

$$\tilde{A} = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

□

Exemplo 3.6

Determine todas as matrizes na forma canônica de Jordan com polinômio característico

$$p_c = (x-2)^3(x-3)^4(x-1),$$

e polinômio minimal

$$p_m = (x - 2)^2(x - 3)^2(x - 1).$$

Determine, também, em cada caso possível, as dimensões dos autoespaços E_1 , E_2 e E_3 associados aos autovalores 1, 2 e 3, respectivamente.

Podemos ver facilmente que $\dim(E_1) = 1$ e $\dim(E_2) = 2$.

Como o primeiro bloco associado ao autovalor 3 é 2×2 e 3 tem multiplicidade 4 como raiz do polinômio característico, podemos ter $\dim(E_3) = 2$ ou $\dim(E_3) = 3$.

Caso 1. Se $\dim(E_3) = 2$, a matriz na forma canônica de Jordan é

$$\begin{pmatrix} 3 & 0 & & & & & & \\ 1 & 3 & & & & & & \\ & & 3 & 0 & & & & \\ & & 1 & 3 & & & & \\ & & & & 2 & 0 & & \\ & & & & 1 & 2 & & \\ & & & & & & 2 & \\ & & & & & & & 1 \end{pmatrix}_{8 \times 8}.$$

Nas matrizes ao lado, os espaços vazios indicam zeros.

Caso 2. Se $\dim(E_3) = 3$, a matriz na forma canônica de Jordan é

$$\begin{pmatrix} 3 & 0 & & & & & & \\ 1 & 3 & & & & & & \\ & & 3 & & & & & \\ & & & 3 & & & & \\ & & & & 2 & 0 & & \\ & & & & 1 & 2 & & \\ & & & & & & 2 & \\ & & & & & & & 1 \end{pmatrix}_{8 \times 8}.$$

□

Exemplo 3.7

Seja A a matriz 3×3

$$A = \begin{pmatrix} 2 & 1 & 3 \\ -1 & -2 & 1 \\ 3 & -1 & 4 \end{pmatrix}.$$

Então,

$$\begin{aligned}
 p_c &= \det(xI - A) = \det \begin{pmatrix} x-2 & -1 & -3 \\ 1 & x+2 & -1 \\ -3 & 1 & x-4 \end{pmatrix} \\
 &= (x-2)((x+2)(x-4) + 1) + (-(x-4) + 3) - 3(1 + 3(x+2)) \\
 &= (x-2)(x^2 - 2x - 7) + (-x + 7) - 3(3x + 7) \\
 &= (x-2)(x^2 - 2x - 7) + (-10x - 14) \\
 &= x^3 - 2x^2 - 7x - 2x^2 + 4x + 14 - 10x - 14 \\
 &= x^3 - 4x^2 - 13x \\
 &= x(x^2 - 4x - 13).
 \end{aligned}$$

Como p_c tem três raízes reais distintas, $p_m = p_c$ e A é diagonalizável.

Como $p_m = p_c$, a matriz na forma canônica racional que é semelhante a A é

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 13 \\ 0 & 1 & 4 \end{pmatrix}.$$

□

4. Forma Canônica de Jordan Real

Seja V um espaço vetorial sobre o corpo $\mathbb{R}$ dos números reais. O **complexificado de V** é o espaço vetorial sobre o corpo $\mathbb{C}$ dos números complexos

$$\hat{V} = \{(u, v) \mid u, v \in V\},$$

com as operações de adição e multiplicação por escalares definidas por:

$$\begin{aligned}
 (u, v) + (u', v') &= (u + u', v + v') \\
 (a + ib)(u, v) &= (au - bv, bu + av).
 \end{aligned}$$

O conjunto $\hat{V}$ com essas operações é de fato um espaço vetorial sobre $\mathbb{C}$. provaremos apenas que

$$((a + ib)(c + id))(u, v) = (a + ib)((c + id)(u, v)),$$

ficando a verificação das outras propriedades como exercício:

$$\begin{aligned}
 ((a + ib)(c + id))(u, v) &= ((ac - bd) + i(bc + ad))(u, v) \\
 &= ((ac - bd)u - (bc + ad)v, (bc + ad)u + (ac - bd)v) \\
 &= (a(cu - dv) - b(du + cv), b(cu - dv) + a(du + cv)) \\
 &= (a + ib)(cu - dv, du + cv) \\
 &= (a + ib)((c + id)(u, v)).
 \end{aligned}$$

Observação 4.1

(1) Em particular, $(u, v) = (u, \mathbf{0}) + i(v, \mathbf{0})$.

(2) Se $a \in \mathbb{R}$, $a(u, v) = (a u, a v)$.

Afirmção: $\dim_{\mathbb{C}}(\hat{V}) = \dim_{\mathbb{R}}(V)$.

Seja $\mathcal{B} = \{w_1, \dots, w_n\}$ uma base de V . Vamos provar que

$$\hat{\mathcal{B}} = \{(w_1, \mathbf{0}), \dots, (w_n, \mathbf{0})\}$$

é uma base de $\hat{V}$.

Seja $(u, v) \in \hat{V}$. Então, existem escalares $a_i \in \mathbb{R}$ e $b_i \in \mathbb{R}$, tais que

$$u = a_1 w_1 + \dots + a_n w_n \quad \text{e} \quad v = b_1 w_1 + \dots + b_n w_n.$$

Assim,

$$\begin{aligned}
 (a_1 + ib_1)(w_1, \mathbf{0}) + \dots + (a_n + ib_n)(w_n, \mathbf{0}) \\
 &= (a_1 w_1, b_1 w_1) + \dots + (a_n w_n, b_n w_n) \\
 &= (a_1 w_1 + \dots + a_n w_n, b_1 w_1 + \dots + b_n w_n) \\
 &= (u, v).
 \end{aligned}$$

Se $(u, v) = (\mathbf{0}, \mathbf{0})$, ou seja, se

$$(a_1 + ib_1)(w_1, \mathbf{0}) + \dots + (a_n + ib_n)(w_n, \mathbf{0}) = (\mathbf{0}, \mathbf{0}),$$

temos que $a_1 w_1 + \dots + a_n w_n = \mathbf{0} = b_1 w_1 + \dots + b_n w_n$.

Logo, $a_1 = \dots = a_n = 0 = b_1 = \dots = b_n$.

Ou seja, os vetores $(w_1, \mathbf{0}), \dots, (w_n, \mathbf{0})$ são LI e geram o espaço vetorial $\hat{V}$.

- Seja T um operador linear sobre o espaço vetorial V sobre $\mathbb{R}$.

Definimos a aplicação

$$\begin{aligned}
 \hat{T}: \hat{V} &\longrightarrow \hat{V} \\
 (u, v) &\longrightarrow \hat{T}(u, v) = (T(u), T(v)).
 \end{aligned}$$

Afirmção: $\hat{T}$ é linear sobre $\mathbb{C}$.

De fato,

$$\begin{aligned}\hat{T}((a+ib)(u,v) + (u',v')) &= \hat{T}((au-bv, bu+av) + (u',v')) \\ &= \hat{T}(au-bv+u', bu+av+v') \\ &= (T(au-bv+u'), T(bu+av+v')) \\ &= (aT(u)-bT(v)+T(u'), bT(u)+aT(v)+T(v')) \\ &= (aT(u)-bT(v), bT(u)+aT(v)) + (T(u'), T(v')) \\ &= (a+ib)(T(u), T(v)) + (T(u'), T(v')) \\ &= (a+ib)\hat{T}(u,v) + \hat{T}(u',v')\end{aligned}$$

Afirmção: $\hat{p}_c = p_c$, onde p_c é o polinômio característico de T e $\hat{p}_c$ é o polinômio característico de $\hat{T}$.

Seja $\mathcal{B} = \{w_1, \dots, w_n\}$ uma base de V e seja $\hat{\mathcal{B}} = \{(w_1, \mathbf{0}), \dots, (w_n, \mathbf{0})\}$ base de $\hat{V}$.

Sendo $T(w_j) = \sum_{i=1}^n a_{ij}w_i$, $j = 1, \dots, n$, temos que

$$\hat{T}(w_j, \mathbf{0}) = (T(w_j), \mathbf{0}) = \left(\sum_{i=1}^n a_{ij}w_i, \mathbf{0} \right) = \sum_{i=1}^n a_{ij}(w_i, \mathbf{0}).$$

Ou seja, $[\hat{T}]_{\hat{\mathcal{B}}} = [T]_{\mathcal{B}}$. Logo,

$$\hat{p}_c = \det(xI - [\hat{T}]_{\hat{\mathcal{B}}}) = \det(xI - [T]_{\mathcal{B}}) = p_c.$$

Afirmção: $\hat{p}_m = p_m$, onde p_m é o polinômio minimal de T e $\hat{p}_m$ é o polinômio minimal de $\hat{T}$.

Seja $f \in \mathbb{R}[x]$, então, $f(\hat{T}) = \widehat{f(T)}$. De fato, se $f = a_kx^k + \dots + a_1x + a_0$; $a_0, \dots, a_k \in \mathbb{R}$, temos que

$$\begin{aligned}f(\hat{T})(u,v) &= (a_k(\hat{T})^k + \dots + a_1\hat{T} + a_0I)(u,v) \\ &= a_k(\hat{T})^k(u,v) + \dots + a_1\hat{T}(u,v) + a_0I(u,v) \\ &= a_k(T^k(u), T^k(v)) + \dots + a_1(T(u), T(v)) + a_0(u,v) \\ &= (a_kT^k(u), a_kT^k(v)) + \dots + (a_1T(u), a_1T(v)) + (a_0u, a_0v) \\ &= (a_kT^k(u) + \dots + a_1T(u) + a_0u, a_kT^k(v) + \dots + a_1T(v) + a_0v) \\ &= (f(T)(u), f(T)(v)) \\ &= \widehat{f(T)}(u,v).\end{aligned}$$

Então, se $f(T) = \mathcal{O}$, temos que $\widehat{f(\hat{T})} = \mathcal{O}$, ou seja, $f(\hat{T}) = \mathcal{O}$.

Logo, $\widehat{p_m}$ divide p_m em $\mathbb{C}[x]$.

Seja, agora, $g = (a_k + ib_k)x^k + \dots + (a_1 + ib_1)x + a_0 + ib_0$ um polinômio com coeficientes complexos. Então, $g = f_1 + if_2$, onde

$$f_1 = a_k x^k + \dots + a_1 x + a_0 \quad \text{e} \quad f_2 = b_k x^k + \dots + b_1 x + b_0$$

são polinômios com coeficientes reais.

Suponha que $g(\hat{T}) = \mathcal{O}$. Então,

$$\begin{aligned} g(\hat{T})(u, v) &= (a_k + ib_k)(\hat{T})^k(u, v) + \dots + (a_1 + ib_1)\hat{T}(u, v) + (a_0 + ib_0)(u, v) \\ &= (a_k + ib_k)(T^k(u), T^k(v)) + \dots + (a_1 + ib_1)(T(u), T(v)) + (a_0 + ib_0)(u, v) \\ &= (a_k T^k(u) - b_k T^k(v), b_k T^k(u) + a_k T^k(v)) + \dots + (a_1 T(u) - b_1 T(v), b_1 T(u) \\ &\quad + a_1 T(v)) + (a_0 u - b_0 v, b_0 u + a_0 v) \\ &= ((a_k T^k(u) + \dots + a_1 T(u) + a_0 u) - (b_k T^k(v) + \dots + b_1 T(v) + b_0 v), \\ &\quad (b_k T^k(u) + \dots + b_1 T(u) + b_0 u) + (a_k T^k(v) + \dots + a_1 T(v) + a_0 v)). \end{aligned}$$

Ou seja, $g(\hat{T})(u, v) = (f_1(T)(u) - f_2(T)(v), f_2(T)(u) + f_1(T)(v))$.

Como $g(\hat{T})(u, v) = \mathbf{0}$, temos que

$$f_1(T)u - f_2(T)(v) = \mathbf{0} \quad \text{e} \quad f_2(T)(u) + f_1(T)(v) = \mathbf{0},$$

para todo par $u, v \in V$.

Fazendo $v = \mathbf{0}$, obtemos que $f_1(T)(u) = f_2(T)(u) = \mathbf{0}$, para todo $u \in V$. Então, $g(\hat{T}) = \mathcal{O}$ se, e somente se $f_1(T) = f_2(T) = \mathcal{O}$.

Seja $\widehat{p_m} = p_1 + ip_2$, onde $p_1, p_2 \in \mathbb{R}[x]$.

Como $\widehat{p_m}(\hat{T}) = \mathcal{O}$, temos que $p_1(T) = p_2(T) = \mathcal{O}$. Logo, p_m divide os polinômios p_1 e p_2 , ou seja, p_m divide $\widehat{p_m}$ em $\mathbb{C}[x]$.

Sendo p_m e $\widehat{p_m}$ polinômios unitários, obtemos que $\widehat{p_m} = p_m$.

Mostrando assim a afirmação.

- Seja $\sigma : \hat{V} \longrightarrow \hat{V}$ a aplicação definida por $\sigma(u, v) = (u, -v)$.

Então σ é um isomorfismo linear sobre $\mathbb{R}$ (quando os espaços são considerados reais), $\sigma(\mu(u, v)) = \bar{\mu}\sigma(u, v)$ e $\sigma^2 = \text{id}$ (aplicação identidade).

De fato:

- $\sigma((u, v) + (u', v')) = \sigma(u + u', v + v') = (u + u', -(v + v'))$
 $= (u, -v) + (u', -v') = \sigma(u, v) + \sigma(u', v')$;
- $\sigma((a + ib)(u, v)) = \sigma(au - bv, bu + av) = (au - bv, -bu - av)$
 $= (a - ib)(u, -v) = (a - ib)\sigma(u, v) = \overline{(a + ib)}\sigma(u, v)$;
- $\sigma^2(u, v) = \sigma(u, -v) = (u, v) = \text{id}(u, v)$.

Observação 4.2

Se W é um subconjunto de $\hat{V}$, usaremos a notação: $\overline{W} = \sigma(W)$.

Seja $T : V \rightarrow V$ um operador linear sobre um espaço vetorial real V de dimensão finita e seja $\hat{T} : \hat{V} \rightarrow \hat{V}$ o operador linear sobre $\hat{V}$ definido acima.

Como $p_c \in \mathbb{R}[x]$, temos que se $\mu = a + ib$ é uma raiz complexa de p_c , então $\bar{\mu} = a - ib$ também o é.

Sendo $\hat{p}_c = p_c$, temos que μ e $\bar{\mu}$ são autovalores de $\hat{T}$.

Sejam

$$\widehat{W}_\mu = \{(u, v) \in \hat{V} \mid \hat{T}(u, v) = \mu(u, v)\},$$

e

$$\widehat{W}_{\bar{\mu}} = \{(u, v) \in \hat{V} \mid \hat{T}(u, v) = \bar{\mu}(u, v)\}.$$

Afirmção: $\sigma(\widehat{W}_\mu) = \widehat{W}_{\bar{\mu}}$.

Seja $(u, v) \in \widehat{W}_\mu$. Então, $\hat{T}(u, v) = \mu(u, v)$.

Logo, $\sigma(\hat{T}(u, v)) = \bar{\mu}\sigma(u, v)$.

Como

$$\sigma(\hat{T}(u, v)) = \sigma(T(u), T(v)) = (T(u), -T(v)) = \hat{T}(u, -v) = \hat{T}(\sigma(u, v))$$

ou seja, $\hat{T} \circ \sigma = \sigma \circ \hat{T}$, temos que

$$\hat{T}(\sigma(u, v)) = \bar{\mu}\sigma(u, v).$$

Logo, $\sigma(\widehat{W}_\mu) \subset \widehat{W}_{\bar{\mu}}$.

De modo análogo, temos que $\sigma(\widehat{W}_{\bar{\mu}}) \subset \widehat{W}_\mu$.

Então, $\sigma^2(\widehat{W}_{\bar{\mu}}) \subset \sigma(\widehat{W}_\mu)$, ou seja, $\widehat{W}_{\bar{\mu}} \subset \sigma(\widehat{W}_\mu)$. Logo, $\sigma(\widehat{W}_\mu) = \widehat{W}_{\bar{\mu}}$.

Afirmção: Se $\{(u_1, v_1), \dots, (u_k, v_k)\}$ é base de um subespaço W de $\hat{V}$, então $\{(u_1, -v_1), \dots, (u_k, -v_k)\}$ é base do subespaço $\overline{W}$ de $\hat{V}$.

Antes temos que provar que se W é um subespaço de $\hat{V}$, então $\overline{W} = \sigma(W)$ é um subespaço de $\hat{V}$.

Se $(u, v), (u', v') \in \overline{W} = \sigma(W)$, então $(u, -v), (u', -v') \in W$.

Logo, $(a - ib)(u, -v) + (u', -v') \in W$.

Então, $\sigma((a + ib)(u, v) + (u', v')) = (a - ib)(u, -v) + (u', -v') \in W$, ou seja, $(a + ib)(u, v) + (u', v') \in \overline{W}$.

Agora provemos a afirmação.

Seja $\sigma(u, v) = (u, -v) \in W$. Como $(u, v) \in W$, existem $a_j + ib_j \in \mathbb{C}$, $j = 1, \dots, k$, tais que

$$(u, v) = (a_1 + ib_1)(u_1, v_1) + \dots + (a_k + ib_k)(u_k, v_k).$$

Logo,

$$(u, -v) = \sigma(u, v) = (a_1 - ib_1)(u_1, -v_1) + \dots + (a_k - ib_k)(u_k, -v_k),$$

ou seja, $\{(u_1, -v_1), \dots, (u_k, -v_k)\}$ gera $\overline{W}$.

Seja $(a_1 - ib_1)(u_1, -v_1) + \dots + (a_k - ib_k)(u_k, -v_k) = \mathbf{0}$. Então,

$$\sigma((a_1 - ib_1)(u_1, -v_1) + \dots + (a_k - ib_k)(u_k, -v_k)) = \mathbf{0},$$

ou seja,

$$(a_1 + ib_1)(u_1, v_1) + \dots + (a_k + ib_k)(u_k, v_k) = \mathbf{0}.$$

Como $\{(u_1, v_1), \dots, (u_k, v_k)\}$ é LI, temos que $a_j + ib_j = 0$, $j = 1, \dots, k$.

Logo, $a_j - ib_j = 0$, $j = 1, \dots, k$, provando que $\{(u_1, -v_1), \dots, (u_k, -v_k)\}$ é LI, e portanto uma base, como afirmado.

Seja $T : V \longrightarrow V$ um operador sobre um espaço vetorial V de dimensão finita sobre $\mathbb{R}$ e seja $\hat{T} : \hat{V} \longrightarrow \hat{V}$ o seu **complexificado**.

Seja

$$p_c = p_1^{d_1} \dots p_k^{d_k} (x - \lambda_1)^{q_1} \dots (x - \lambda_\ell)^{q_\ell},$$

o polinômio característico de T , onde $\lambda_1, \dots, \lambda_\ell \in \mathbb{R}$ e $p_j = (x - \mu_j)(x - \overline{\mu_j})$, $\mu_j = a_j + ib_j$, $b_j > 0$, $j = 1, \dots, k$, e

$$p_m = p_1^{r_1} \dots p_k^{r_k} (x - \lambda_1)^{s_1} \dots (x - \lambda_\ell)^{s_\ell},$$

o polinômio minimal de T .

Como T e $\hat{T}$ tem o mesmo polinômio característico e o mesmo polinômio minimal temos que, em $\mathbb{C}$,

$$p_c = (x - \mu_1)^{d_1} (x - \overline{\mu_1})^{d_1} \dots (x - \mu_k)^{d_k} (x - \overline{\mu_k})^{d_k} (x - \lambda_1)^{q_1} \dots (x - \lambda_\ell)^{q_\ell},$$

$$p_m = (x - \mu_1)^{r_1} (x - \overline{\mu_1})^{r_1} \dots (x - \mu_k)^{r_k} (x - \overline{\mu_k})^{r_k} (x - \lambda_1)^{s_1} \dots (x - \lambda_\ell)^{s_\ell},$$

ou seja, as multiplicidades de μ_j e $\overline{\mu_j}$ em p_c e p_m são iguais, $j = 1, \dots, k$.

Afirmção: $\text{Ker}(\hat{T} - \lambda_i I)^s$ é o complexificado de $\text{Ker}(T - \lambda_i I)$.

De fato, como

$$\begin{aligned} (\hat{T} - \lambda_i I)(u, v) &= \hat{T}(u, v) - \lambda_i(u, v) = (T(u), T(v)) - (\lambda_i u, \lambda_i v) \\ &= ((T - \lambda_i I)(u), (T - \lambda_i I)(v)), \end{aligned}$$

podemos provar, por indução sobre s , que

$$(\hat{T} - \lambda_i I)^s(u, v) = ((T - \lambda_i I)^s(u), (T - \lambda_i I)^s(v)),$$

ou seja,

$$((T - \lambda_i I)^s)^\wedge = (\hat{T} - \lambda_i I)^s.$$

Logo, $(u, v) \in \text{Ker}(\hat{T} - \lambda_i I)^s$ se, e somente se, $u, v \in \text{Ker}(T - \lambda_i I)$, ou seja,

$$(\text{Ker}(T - \lambda_i I)^s)^\wedge = \{(u, v) \mid u, v \in \text{Ker}(T - \lambda_i I)^s\} = \text{Ker}(\hat{T} - \lambda_i I)^s.$$

Então, $\dim(\text{Ker}(T - \lambda_i I)^s) = \dim(\text{Ker}(\hat{T} - \lambda_i I)^s)$, para todo $s \in \mathbb{N}$.

Em particular, se $\widetilde{W}_i = \text{Ker}(\hat{T} - \lambda_i I)^{s_i}$ e $W_i = \text{Ker}(T - \lambda_i I)^{s_i}$, temos que $\widetilde{W}_i = \widehat{W}_i$ e $\dim(\widetilde{W}_i) = \dim(W_i)$.

Seja $\widetilde{\mathcal{B}}_i$ uma base de $\widetilde{W}_i$ tal que $[\hat{T}|_{\widetilde{W}_i}]_{\widetilde{\mathcal{B}}_i}$ está na forma canônica de Jordan. Então, $[\hat{T}|_{\widetilde{W}_i}]_{\widetilde{\mathcal{B}}_i}$ possui $\dim(\text{Ker}(\hat{T} - \lambda_i I)) = \dim(\text{Ker}(T - \lambda_i I))$ blocos elementares de Jordan associados ao autovalor λ_i , sendo as quantidades dos blocos de cada ordem dadas em termos das dimensões

$$\dim(\text{Ker}(\hat{T} - \lambda_i I)^s) = \dim(\text{Ker}(T - \lambda_i I)^s), \quad 1 \leq s \leq s_i.$$

Assim, a matriz na forma canônica de Jordan de $T|_{W_i}$ é igual à matriz na forma canônica de Jordan de $\hat{T}|_{\widetilde{W}_i}$.

Além disso, se $\mathcal{B}_i = \{v_1, \dots, v_{q_i}\}$ é uma base de W_i tal que $[T|_{W_i}]_{\mathcal{B}_i}$ está na forma canônica de Jordan, então $\widetilde{\mathcal{B}}_i = \{(v_1, \mathbf{0}), \dots, (v_{q_i}, \mathbf{0})\}$ é uma base de $\widetilde{W}_i$ tal que $[\hat{T}|_{\widetilde{W}_i}]_{\widetilde{\mathcal{B}}_i} = [T|_{W_i}]_{\mathcal{B}_i}$ está na forma canônica de Jordan.

Afirmção: $\text{Ker}(\hat{T} - \bar{\mu}I)^s = \sigma(\text{Ker}(\hat{T} - \mu I)^s)$, para $\mu = a + ib$, $b > 0$.

De fato, $(\hat{T} - \mu I)^s = p_\mu(\hat{T})$, onde $p_\mu = (x - \mu)^s$ é um polinômio cujos coeficientes são polinômios em μ com coeficientes inteiros, pois:

$$p_\mu = x^s + p_{s-1}(\mu)x^{s-1} + \dots + p_1(\mu)x + p_0(\mu),$$

com

$$p_{s-j}(\mu) = (-1)^j \binom{s}{j} \mu^j = (-1)^j \frac{s!}{j!(s-j)!} \mu^j.$$

Logo, $(\hat{T} - \bar{\mu}I)^s = (\hat{T})^s + p_{s-1}(\bar{\mu})(\hat{T})^{s-1} + \dots + p_1(\bar{\mu})\hat{T} + p_0(\bar{\mu}) = \bar{p}_\mu(\hat{T})$,
pois

$$\begin{aligned} \bar{p}_\mu &= x^s + \overline{p_{s-1}(\mu)}x^{s-1} + \dots + \overline{p_1(\mu)}x + \overline{p_0(\mu)} \\ &= x^s + p_{s-1}(\bar{\mu})x^{s-1} + \dots + p_1(\bar{\mu})x + p_0(\bar{\mu}), \end{aligned}$$

já que os coeficientes de p_j são inteiros.

Então, se $p_\mu = g_\mu + ih_\mu$, $g_\mu, h_\mu \in \mathbb{R}[x]$, temos que

$$(\hat{T} - \mu I)^s = p_\mu(\hat{T}) = g_\mu(\hat{T}) + ih_\mu(\hat{T}), \quad \text{e} \quad (\hat{T} - \bar{\mu}I)^s = \bar{p}_\mu(\hat{T}) = g_\mu(\hat{T}) - ih_\mu(\hat{T}).$$

Assim, como g_μ e h_μ têm coeficientes reais,

$$\begin{aligned} p_\mu(\hat{T})(u, v) &= \widehat{g_\mu(T)}(u, v) + i\widehat{h_\mu(T)}(u, v) \\ &= (g_\mu(T)(u), g_\mu(T)(v)) + i(h_\mu(T)(u), h_\mu(T)(v)) \\ &= (g_\mu(T)(u), g_\mu(T)(v)) + (-h_\mu(T)(v), h_\mu(T)(u)) \\ &= (g_\mu(T)(u) - h_\mu(T)(v), g_\mu(T)(v) + h_\mu(T)(u)). \end{aligned}$$

Por outro lado,

$$\begin{aligned} (\hat{T} - \bar{\mu}I)^s(u, -v) &= g_\mu(\hat{T})(u, -v) - ih_\mu(\hat{T})(u, -v) \\ &= (g_\mu(T)(u), -g_\mu(T)(v)) - i(h_\mu(T)(u), -h_\mu(T)(v)) \\ &= (g_\mu(T)(u), -g_\mu(T)(v)) + (-h_\mu(T)(v), -h_\mu(T)(u)) \\ &= (g_\mu(T)(u) - h_\mu(T)(v), -g_\mu(T)(v) - h_\mu(T)(u)), \end{aligned}$$

ou seja,

$$(\hat{T} - \bar{\mu}I)^s(\sigma(u, v)) = \sigma((\hat{T} - \mu I)^s(u, v)).$$

Logo, se $(u, v) \in \text{Ker}(\hat{T} - \mu I)^s$ então $\sigma(u, v) \in \text{Ker}(\hat{T} - \bar{\mu}I)^s$. Assim,
 $\sigma(\text{Ker}(\hat{T} - \mu I)^s) \subset \text{Ker}(\hat{T} - \bar{\mu}I)^s$.

De modo análogo, temos que

$$\sigma(\text{Ker}(\hat{T} - \bar{\mu}I)^s) \subset \text{Ker}(\hat{T} - \mu I)^s.$$

Logo,

$$\text{Ker}(\hat{T} - \bar{\mu}I)^s = \sigma^2(\text{Ker}(\hat{T} - \bar{\mu}I)^s) \subset \sigma(\text{Ker}(\hat{T} - \mu I)^s).$$

Portanto,

$$\text{Ker}(\hat{T} - \bar{\mu}I)^s = \sigma(\text{Ker}(\hat{T} - \mu I)^s)$$

e

$$\dim(\text{Ker}(\hat{T} - \bar{\mu}I)^s) = \dim(\text{Ker}(\hat{T} - \mu I)^s),$$

para todo $s \in \mathbb{N}$.

Se $\mu_j, \bar{\mu}_j, j = 1, \dots, k$, são as raízes complexas de p_c , temos que $\dim(\text{Ker}(\hat{T} - \bar{\mu}_j I)^s) = \dim(\text{Ker}(\hat{T} - \mu_j I)^s)$, para todo $s \in \mathbb{N}$. Logo, $A_{\bar{\mu}_j} = \overline{A_{\mu_j}}$, onde A_{μ_j} é a matriz na forma canônica de $\hat{T}|_{V_j}$, $A_{\bar{\mu}_j}$ é a matriz na forma canônica de $\hat{T}|_{\bar{V}_j}$, $V_j = \text{Ker}(\hat{T} - \mu_j I)^{r_j}$ e $\bar{V}_j = \text{Ker}(\hat{T} - \bar{\mu}_j I)^{r_j}$.

Afirmção: Se $\mathcal{B}_j$ é uma base de V_j tal que $[\hat{T}|_{V_j}]_{\mathcal{B}_j} = A_{\mu_j}$, então $\bar{\mathcal{B}}_j = \sigma(\mathcal{B}_j)$ é base de $\bar{V}_j$ e $[\hat{T}|_{\bar{V}_j}]_{\bar{\mathcal{B}}_j} = A_{\bar{\mu}_j}$.

Seja

$$\mathcal{B}_j = \mathcal{B}_j^1 \cup \dots \cup \mathcal{B}_j^\ell, \quad \ell = \dim(\text{Ker}(\hat{T} - \mu_j I)),$$

onde $\mathcal{B}_j^i = \{(v_1, w_1), \dots, (v_k, w_k)\}$ é um conjunto LI de vetores de $\hat{V}$ que dá origem a um bloco de ordem k de A_{μ_j} .

Então,

$$(T(v_p), T(w_p)) = \hat{T}(v_p, w_p) = \mu_j(v_p, w_p) + (v_{p+1}, w_{p+1}), \quad 1 \leq p \leq k-1,$$

e

$$(T(v_k), T(w_k)) = \hat{T}(v_k, w_k) = \mu_j(v_k, w_k).$$

Como $\hat{T}\sigma = \sigma\hat{T}$, temos que

$$\begin{aligned} \hat{T}(v_p, -w_p) &= (T(v_p), -T(w_p)) = \sigma(T(v_p), T(w_p)) \\ &= \sigma(\hat{T}(v_p, w_p)) = \sigma(\mu_j(v_p, w_p) + (v_{p+1}, w_{p+1})) \\ &= \bar{\mu}_j(v_p, -w_p) + (v_{p+1}, -w_{p+1}), \end{aligned}$$

e

$$\begin{aligned} \hat{T}(v_k, -w_k) &= (T(v_k), -T(w_k)) = \sigma(T(v_k), T(w_k)) \\ &= \sigma(\hat{T}(v_k, w_k)) = \sigma(\mu_j(v_k, w_k)) = \bar{\mu}_j(v_k, -w_k). \end{aligned}$$

Observação 4.3

Sejam $\mathcal{B}_j = \{(v_1, w_1), \dots, (v_{d_j}, w_{d_j})\}$ e $\overline{\mathcal{B}}_j = \{(v_1, -w_1), \dots, (v_{d_j}, -w_{d_j})\}$ bases de $V_j = \text{Ker}(\hat{T} - \mu_j I)^{r_j}$ e $\overline{V}_j = \text{Ker}(\hat{T} - \overline{\mu}_j I)^{r_j}$, respectivamente, onde $d_j = \dim(\text{Ker}(\hat{T} - \mu_j I)^{r_j}) = \dim(\text{Ker}(\hat{T} - \overline{\mu}_j I)^{r_j})$, tais que $[\hat{T}|_{V_j}]_{\mathcal{B}_j} = A_{\mu_j}$ e $[\hat{T}|_{\overline{V}_j}]_{\overline{\mathcal{B}}_j} = A_{\overline{\mu}_j}$ estão na forma canônica.

Como V_j e $\overline{V}_j$ são subespaços LI, pois estão associados a autovalores μ_j e $\overline{\mu}_j$, distintos, temos que

$$\{(v_1, w_1), \dots, (v_{d_j}, w_{d_j}), (v_1, -w_1), \dots, (v_{d_j}, -w_{d_j})\}$$

é uma base de $V_j \oplus \overline{V}_j$.

Afirmção: Seja $p_j = (x - \mu_j)(x - \overline{\mu}_j) \in \mathbb{R}[x]$, $j = 1, \dots, k$.

Então,

$$V_j \oplus \overline{V}_j = (\text{Ker}(p_j(T))^{r_j})^\wedge$$

e $\{v_1 + w_1, v_1 - w_1, \dots, v_{d_j} + w_{d_j}, v_{d_j} - w_{d_j}\}$ é uma base de $\text{Ker}(p_j(T))^{r_j}$.

Pelo Teorema Generalizado de Cayley-Hamilton, temos que

$$\dim(\text{Ker}(p_j(T))^{r_j}) = d_j \times \text{grau}(p_j) = 2d_j.$$

Sejam $p_{\mu_j} = (x - \mu_j)^{r_j}$ e $p_{\overline{\mu}_j} = (x - \overline{\mu}_j)^{r_j}$.

Como $p_{\mu_j} = g_{\mu_j} + ih_{\mu_j}$ e $p_{\overline{\mu}_j} = g_{\mu_j} - ih_{\mu_j}$, onde $g_{\mu_j}, h_{\mu_j} \in \mathbb{R}[x]$, temos:

$$p_j^{r_j} = p_{\mu_j} p_{\overline{\mu}_j} = (g_{\mu_j} + ih_{\mu_j})(g_{\mu_j} - ih_{\mu_j}) = g_{\mu_j}^2 + h_{\mu_j}^2.$$

Sendo

$$(\hat{T} - \mu_j I)^{r_j} = p_{\mu_j}(\hat{T}) = (g_{\mu_j} + ih_{\mu_j})(\hat{T}),$$

temos que $(u, v) \in V_j = \text{Ker}(\hat{T} - \mu_j I)^{r_j} = \text{Ker}(g_{\mu_j} + ih_{\mu_j})(\hat{T})$ se, e somente se,

$$\begin{aligned} (\mathbf{0}, \mathbf{0}) &= (g_{\mu_j} + ih_{\mu_j})(\hat{T})(u, v) = (g_{\mu_j}(\hat{T}) + ih_{\mu_j}(\hat{T}))(u, v) \\ &= g_{\mu_j}(\hat{T})(u, v) + ih_{\mu_j}(\hat{T})(u, v) \\ &= (g_{\mu_j}(T)(u), g_{\mu_j}(T)(v)) + i(h_{\mu_j}(T)(u), h_{\mu_j}(T)(v)) \\ &= (g_{\mu_j}(T)(u), g_{\mu_j}(T)(v)) + (-h_{\mu_j}(T)(v), h_{\mu_j}(T)(u)) \\ &= (g_{\mu_j}(T)(u) - h_{\mu_j}(T)(v), g_{\mu_j}(T)(v) + h_{\mu_j}(T)(u)), \end{aligned}$$

ou seja $g_{\mu_j}(T)(u) = h_{\mu_j}(T)(v)$ e $g_{\mu_j}(T)(v) = -h_{\mu_j}(T)(u)$.

Logo,

$$p_j^{r_j}(T)(u) = g_{\mu_j}^2(T)(u) + h_{\mu_j}^2(u) = g_{\mu_j}(T)h_{\mu_j}(T)(v) - h_{\mu_j}(T)g_{\mu_j}(T)(v) = \mathbf{0},$$

e

$$p_j^{r_j}(T)(v) = g_{\mu_j}^2(T)(v) + h_{\mu_j}^2(v) = -g_{\mu_j}(T)h_{\mu_j}(T)(u) + h_{\mu_j}(T)g_{\mu_j}(T)(u) = \mathbf{0}.$$

Como

$$\begin{aligned} (\text{Ker}(p_j(T)^{r_j}))^\wedge &= \{(u, v) \mid p_j(T)^{r_j}(u) = p_j(T)^{r_j}(v) = \mathbf{0}\} \\ &= \{(u, v) \mid g_{\mu_j}^2(T)(u) + h_{\mu_j}^2(T)(u) = g_{\mu_j}^2(T)(v) + h_{\mu_j}^2(T)(v) = \mathbf{0}\}, \end{aligned}$$

temos que $V_j \subset (\text{Ker}(p_j(T)^{r_j}))^\wedge$.

Sendo $(\hat{T} - \overline{\mu_j}I)^{r_j} = p_{\overline{\mu_j}}(\hat{T}) = (g_{\mu_j} - ih_{\mu_j})(\hat{T})$, podemos provar, de modo análogo, que se

$$(u, v) \in \overline{V_j} = \text{Ker}(\hat{T} - \overline{\mu_j}I)^{r_j} = \text{Ker}(p_{\overline{\mu_j}}(\hat{T})) = \text{Ker}((g_{\mu_j} - ih_{\mu_j})(\hat{T})),$$

então $(u, v) \in (\text{Ker}(p_j(T)^{r_j}))^\wedge$.

Logo, $V_j \oplus \overline{V_j} \subset (\text{Ker}(p_j(T)^{r_j}))^\wedge$. Como

$$\dim(V_j \oplus \overline{V_j}) = 2d_j = \dim(\text{Ker}(p_j(T)^{r_j}))^\wedge,$$

temos que

$$V_j \oplus \overline{V_j} = (\text{Ker}(p_j(T)^{r_j}))^\wedge.$$

Vamos, agora, provar que

$$\{v_1 + w_1, v_1 - w_1, \dots, v_{d_j} + w_{d_j}, v_{d_j} - w_{d_j}\}$$

é uma base de $\text{Ker}(p_j(T)^{r_j})$.

De fato, como $(v_i, w_i) \in V_j$ e $(v_i, -w_i) \in \overline{V_j}$, $i = 1, \dots, d_j$, temos que:

- $(v_i, w_i) + i(v_i, -w_i) = (v_i + w_i, v_i + w_i) \in V_j \oplus \overline{V_j} = (\text{Ker}(p_j(T)^{r_j}))^\wedge$;
- $(v_i, w_i) - i(v_i, -w_i) = (v_i - w_i, -(v_i - w_i)) \in V_j \oplus \overline{V_j} = (\text{Ker}(p_j(T)^{r_j}))^\wedge$.

Logo, $v_i + w_i, v_i - w_i \in \text{Ker}(p_j(T)^{r_j})$, $i = 1, \dots, d_j$.

• Os vetores

$$\begin{aligned} &(v_1 + w_1, v_1 + w_1), \dots, (v_{d_j} + w_{d_j}, v_{d_j} + w_{d_j}), \dots, \\ &(v_1 - w_1, -(v_1 - w_1)), \dots, (v_{d_j} - w_{d_j}, -(v_{d_j} - w_{d_j})) \end{aligned}$$

formam uma base de $V_j \oplus \overline{V_j} = (\text{Ker}(p_j(T)^{r_j}))^\wedge$.

Como $\dim(\text{Ker}(p_j(T)^{r_j}))^\wedge = 2d_j$, basta mostrar que esses vetores são LI.

Sejam $\lambda_1, \dots, \lambda_{d_j}, \delta_1, \dots, \delta_{d_j} \in \mathbb{C}$, tais que

$$\lambda_1(v_1 + w_1, v_1 + w_1) + \dots + \lambda_{d_j}(v_{d_j} + w_{d_j}, v_{d_j} + w_{d_j}) \\ + \delta_1(v_1 - w_1, -(v_1 - w_1)) + \dots + \delta_{d_j}(v_{d_j} - w_{d_j}, -(v_{d_j} - w_{d_j})) = (\mathbf{0}, \mathbf{0}).$$

Como

$$\begin{aligned} e \quad (v_i + w_i, v_i + w_i) &= (v_i, w_i) + i(v_i, -w_i) \\ (v_i - w_i, -(v_i - w_i)) &= (v_i, w_i) - i(v_i, -w_i), \end{aligned}$$

para $i = 1, \dots, d_j$, temos que

$$\begin{aligned} (\lambda_1 + \delta_1)(v_1, w_1) + \dots + (\lambda_{d_j} + \delta_{d_j})(v_{d_j}, w_{d_j}) \\ + (\lambda_1 - \delta_1)(v_1, -w_1) + \dots + (\lambda_{d_j} - \delta_{d_j})(v_{d_j}, -w_{d_j}) = (\mathbf{0}, \mathbf{0}). \end{aligned}$$

Logo, $\lambda_i + \delta_i = \lambda_i - \delta_i = 0$, $i = 1, \dots, d_j$, pois

$$(v_1, w_1), \dots, (v_{d_j}, w_{d_j}), (v_1, -w_1), \dots, (v_{d_j}, -w_{d_j})\}$$

é uma base de $V_j \oplus \overline{V_j}$.

Então, $\lambda_i = \delta_i = 0$, $i = 1, \dots, d_j$.

Seja

$$\begin{aligned} (a_1 + ib_1)(v_1 + w_1, v_1 + w_1) + \dots + (a_{d_j} + ib_{d_j})(v_{d_j} + w_{d_j}, v_{d_j} + w_{d_j}) \\ + (c_1 + ie_1)(v_1 - w_1, -(v_1 - w_1)) + \dots + (c_{d_j} + ie_{d_j})(v_{d_j} - w_{d_j}, -(v_{d_j} - w_{d_j})) \end{aligned} \quad (I)$$

uma combinação linear dos vetores da base

$$\begin{aligned} \{(v_1 + w_1, v_1 + w_1), \dots, (v_{d_j} + w_{d_j}, v_{d_j} + w_{d_j}), \\ (v_1 - w_1, -(v_1 - w_1)), \dots, (v_{d_j} - w_{d_j}, -(v_{d_j} - w_{d_j}))\} \end{aligned}$$

de $(\text{Ker}(p_j(T)^{r_j}))^\wedge$, onde $a_i, b_i, c_i, e_i \in \mathbb{R}$, $i = 1, \dots, d_j$.

Como para todo $i = 1, \dots, d_j$:

$$(a_i + ib_i)(v_i + w_i, v_i + w_i) = ((a_i - b_i)(v_i + w_i), (a_i + b_i)(v_i + w_i))$$

e

$$(c_i + ie_i)(v_i - w_i, -(v_i - w_i)) = ((c_i + e_i)(v_i - w_i), (e_i - c_i)(v_i - w_i)),$$

temos que a combinação linear (I) pode ser reescrita na forma:

$$\begin{aligned} ((a_1 - b_1)(v_1 + w_1), (a_1 + b_1)(v_1 + w_1)) + \dots \\ + ((a_{d_j} - b_{d_j})(v_{d_j} + w_{d_j}), (a_{d_j} + b_{d_j})(v_{d_j} + w_{d_j})) \\ + ((c_1 + e_1)(v_1 - w_1), (e_1 - c_1)(v_1 - w_1)) + \dots \\ + ((c_{d_j} + e_{d_j})(v_{d_j} - w_{d_j}), (e_{d_j} - c_{d_j})(v_{d_j} - w_{d_j})). \end{aligned}$$

Logo, todo vetor pertencente a $\text{Ker}(p_j(T)^{r_j})$ pode ser escrito como uma combinação linear dos vetores do conjunto

$$\{v_1 + w_1, \dots, v_{d_j} + w_{d_j}, v_1 - w_1, \dots, v_{d_j} - w_{d_j}\}.$$

Como $\dim(\text{Ker}(p_j(T)^{r_j})) = 2d_j$, temos que esse conjunto é uma base de $\text{Ker}(p_j(T)^{r_j})$. Como queríamos provar.

- Seja $\mu_j = a_j + ib_j$, $b_j > 0$, $j = 1, \dots, k$ um autovalor complexo de $\hat{T}$.

Seja $\mathcal{B}_j = \{(v_1, w_1), \dots, (v_{d_j}, w_{d_j})\}$ uma base de $V_j = \text{Ker}(\hat{T} - \mu_j I)^{r_j}$ tal que $A_{\mu_j} = [\hat{T}|_{V_j}]_{\mathcal{B}_j}$ está na forma canônica de Jordan. Então, $\overline{\mathcal{B}}_j = \{(v_1, -w_1), \dots, (v_{d_j}, -w_{d_j})\}$ é uma base de $\overline{V}_j = \text{Ker}((\hat{T} - \overline{\mu_j} I)^{r_j})$ tal que $A_{\overline{\mu_j}} = [\hat{T}|_{\overline{V}_j}]_{\overline{\mathcal{B}}_j}$ está na forma canônica de Jordan e $A_{\overline{\mu_j}} = \overline{A_{\mu_j}}$.

Seja $\tilde{\mathcal{B}}_j = \{v_1 + w_1, v_1 - w_1, \dots, v_{d_j} + w_{d_j}, v_{d_j} - w_{d_j}\}$ a base de $\text{Ker}(p_j(T)^{r_j})$ obtida a partir das bases $\mathcal{B}_j$ e $\overline{\mathcal{B}}_j$.

Como

$$(T(v_i), T(w_i)) = \hat{T}(v_i, w_i) = (a_j + ib_j)(v_i, w_i) + (v_{i+1}, w_{i+1}),$$

ou

$$(T(v_i), T(w_i)) = \hat{T}(v_i, w_i) = (a_j + ib_j)(v_i, w_i),$$

temos que

$$\begin{cases} T(v_i) = a_j v_i - b_j w_i + v_{i+1} \\ T(w_i) = b_j v_i + a_j w_i + w_{i+1} \end{cases} \quad \text{ou} \quad \begin{cases} T(v_i) = a_j v_i - b_j w_i \\ T(w_i) = b_j v_i + a_j w_i. \end{cases}$$

Logo,

$$T(v_i + w_i) = a_j(v_i + w_i) + b_j(v_i - w_i) + (v_{i+1} + w_{i+1})$$

$$T(v_i - w_i) = -b_j(v_i + w_i) + a_j(v_i - w_i) + (v_{i+1} - w_{i+1}),$$

ou

$$T(v_i + w_i) = a_j(v_i + w_i) + b_j(v_i - w_i)$$

$$T(v_i - w_i) = -b_j(v_i + w_i) + a_j(v_i - w_i).$$

Então, se $\{(v_{i_1}, w_{i_1}), \dots, (v_{i_k}, w_{i_k})\}$ determinam um bloco elementar de Jordan de A_{μ_j} de ordem k ,

$$\{(v_{i_1} + w_{i_1}), (v_{i_1} - w_{i_1}), \dots, (v_{i_k} + w_{i_k}), (v_{i_k} - w_{i_k})\},$$

determinam um bloco de ordem $2k$ na forma:

$$\left(\begin{array}{cc|cc|cc|cc|cc} a_j & -b_j & & & & & & & & \\ b_j & a_j & & & & & & & & \\ \hline 1 & 0 & a_j & -b_j & & & & & & \\ 0 & 1 & b_j & a_j & & & & & & \\ \hline & & 1 & 0 & a_j & -b_j & & & & \\ & & 0 & 1 & b_j & a_j & & & & \\ \hline & & & & \ddots & & \ddots & & & \\ & & & & & \ddots & & \ddots & & \\ \hline & & & & & & 1 & 0 & a_j & -b_j \\ & & & & & & 0 & 1 & b_j & a_j \end{array} \right)$$

Sejam $U_i = \text{Ker}(p_i(T)^{r_i})$, $i = 1, \dots, k$ e $W_i = \text{Ker}((T - \lambda_i I)^{s_i})$, $i = 1, \dots, \ell$. Sejam $\tilde{\mathcal{B}}_i$ a base de U_i construída acima e $\tilde{\mathcal{B}}_{i+k}$ uma base de W_i tal que $[T|_{W_i}]_{\tilde{\mathcal{B}}_{i+k}}$ está na forma canônica de Jordan.

Como $V = U_1 \oplus \dots \oplus U_k \oplus W_1 \oplus \dots \oplus W_\ell$, temos que

$$\tilde{\mathcal{B}} = \tilde{\mathcal{B}}_1 \cup \dots \cup \tilde{\mathcal{B}}_k \cup \tilde{\mathcal{B}}_{k+1} \cup \dots \cup \tilde{\mathcal{B}}_{k+\ell}$$

é uma base de V tal que

$$[T]_{\tilde{\mathcal{B}}} = \begin{pmatrix} A_1 & & & & \\ & \ddots & & & \\ & & A_k & & \\ & & & B_1 & \\ & & & & \ddots \\ & & & & & B_\ell \end{pmatrix}_{n \times n}$$

onde:

(I)

$$\bullet A_j = \begin{pmatrix} R_1^j & & \\ & \ddots & \\ & & R_{n_j}^j \end{pmatrix}_{2d_j \times 2d_j};$$

$$\bullet R_i^j = \begin{pmatrix} a_j & -b_j & & & & \\ b_j & a_j & & & & \\ 1 & 0 & a_j & -b_j & & \\ 0 & 1 & b_j & a_j & & \\ & & \ddots & & \ddots & \\ & & & \ddots & & \ddots \\ & & & & 1 & 0 & a_j & -b_j \\ & & & & 0 & 1 & b_j & a_j \end{pmatrix}_{2k_i \times 2k_i}, \quad i = 1, \dots, n_j;$$

- $2n_j = \dim(\text{Ker}(p_j(T))) = 2 \dim(\text{Ker}(\hat{T} - \mu_j I))$;
- $k_i \geq k_{i+1}$, $i = 1, \dots, n_j - 1$;
- $2k_1 = 2r_j$, ou seja, $k_1 = r_j$;
- $\nu_k(\mu_j) = -\delta_{k-1}(\mu_j) + 2\delta_k(\mu_j) - \delta_{k+1}(\mu_j)$, sendo $\nu_k(\mu_j)$ o número de blocos de tamanho $2k \times 2k$ associado ao autovalor $\mu_j = \alpha_j + ib_j$, $b_j > 0$, e $2\delta_k(\mu_j) = \dim(\text{Ker}(p_j(T)^k)) = 2 \dim(\text{Ker}(\hat{T} - \mu_j I)^k)$.

(II)

- $B_i = \begin{pmatrix} J_1^i & & \\ & \ddots & \\ & & J_{m_i}^i \end{pmatrix}_{q_i \times q_i}$;
- $J_j^i = \begin{pmatrix} \lambda_i & & & \\ 1 & \lambda_i & & \\ & 1 & \ddots & \\ & & \ddots & \ddots \\ & & & 1 & \lambda_i \end{pmatrix}_{f_j \times f_j}$, $j = 1, \dots, m_i$;
- $q_i = \text{multiplicidade de } \lambda_i \text{ como raiz do polinômio característico de } T$;
- $m_i = \dim(\text{Ker}(T - \lambda_i I))$;
- $f_j \geq f_{j+1}$, $j = 1, \dots, m_i - 1$;
- $f_1 = s_1$ (multiplicidade de λ_i como raiz do polinômio minimal de T);
- $\nu_k(\lambda_i) = -\delta_{k-1}(\lambda_i) + 2\delta_k(\lambda_i) - \delta_{k+1}(\lambda_i)$, sendo $\nu_k(\lambda_i)$ o número de blocos de tamanho $k \times k$ associado ao autovalor λ_i e $\delta_k(\lambda_i) = \dim(\text{Ker}(T - \lambda_i I)^k)$.

A matriz

$$[T]_{\tilde{B}} = \begin{pmatrix} A_1 & & & \\ & \ddots & & \\ & & A_k & \\ & & & B_1 & \\ & & & & \ddots & \\ & & & & & B_\ell \end{pmatrix}_{n \times n}$$

é a forma canônica de Jordan real do operador T .

Observação 4.4

Se as raízes do polinômio característico de um operador T sobre um espaço vetorial real são todas reais, a forma canônica de Jordan real de T é igual à sua forma canônica de Jordan.

Em particular, se T é nilpotente, sua forma canônica de Jordan real é igual à sua forma canônica de Jordan, que, por sua vez, é igual à sua forma racional.

Unicidade: Seja $\tilde{B}'$ uma base de V tal que $[T]_{\tilde{B}'}$ está na forma canônica de Jordan real. Então, $[T]_{\tilde{B}}$ e $[T]_{\tilde{B}'}$ podem diferir apenas pela ordem em que aparecem os blocos associados a uma raiz do polinômio característico de T .

De fato, seja $\tilde{B}'$ uma base de V tal que

$$[T]_{\tilde{B}'} = \begin{pmatrix} C_1 & & & & \\ & \ddots & & & \\ & & C_s & & \\ & & & D_1 & \\ & & & & \ddots \\ & & & & & D_q \end{pmatrix}_{n \times n},$$

onde:

$$\bullet C_i = \begin{pmatrix} C_1^i & & \\ & \ddots & \\ & & C_{p_i}^i \end{pmatrix}_{e_i \times e_i};$$

$$\bullet C_j^i = \begin{pmatrix} \alpha_i & -\beta_i & & & \\ \beta_i & \alpha_i & & & \\ 1 & 0 & \alpha_i & -\beta_i & \\ 0 & 1 & \beta_i & \alpha_i & \\ & & \ddots & & \ddots \\ & & & \ddots & \ddots \\ & & & & 1 & 0 & \alpha_i & -\beta_i \\ & & & & 0 & 1 & \beta_i & \alpha_i \end{pmatrix}_{2h_j^i \times 2h_j^i};$$

$$\bullet h_j^i \geq h_{j+1}^i, j = 1, \dots, p_i - 1;$$

$$\bullet D_i = \begin{pmatrix} D_1^i & & \\ & \ddots & \\ & & D_{t_i}^i \end{pmatrix}_{u_i \times u_i};$$

$$\bullet D_j^i = \begin{pmatrix} \delta_i & & & \\ 1 & \delta_i & & \\ & 1 & \ddots & \\ & & \ddots & \ddots \\ & & & 1 & \delta_i \end{pmatrix}_{\ell_j^i \times \ell_j^i} ;$$

$$\bullet \ell_j^i \geq \ell_{j+1}^i, j = 1, \dots, t_i - 1.$$

Como $[T]_{\tilde{B}'}$ está na forma de blocos, temos que

$$\begin{aligned} p_c &= \det(xI - [T]_{\tilde{B}'}) \\ &= ((x - c_1)^2 + d_1^2)^{e_1} \dots ((x - c_s)^2 + d_s^2)^{e_s} (x - \delta_1)^{u_1} \dots (x - \delta_q)^{u_q}, \end{aligned}$$

ou seja, $\alpha_1 + i\beta_1, \alpha_1 - i\beta_1, \dots, \alpha_s + i\beta_s, \alpha_s - i\beta_s$ são as raízes complexas distintas de p_c e $\delta_1, \dots, \delta_q$ são as raízes reais distintas de p_c .

Logo: $k = s$; $\ell = q$; $\{\alpha_1 + i\beta_1, \dots, \alpha_s + i\beta_s\} = \{a_1 + ib_1, \dots, a_k + ib_k\}$ e $\{\delta_1, \dots, \delta_q\} = \{\lambda_1, \dots, \lambda_q\}$. Além disso,

$$e_i = d_j, \text{ se } \alpha_i + i\beta_i = a_j + ib_j \quad \text{e} \quad u_i = q_j, \text{ se } \delta_i = \lambda_j.$$

• Seja $\tilde{B}' = \{\tilde{v}_1, \tilde{w}_1, \dots, \tilde{v}_m, \tilde{w}_m, \tilde{u}_1, \dots, \tilde{u}_p\}$, onde $m = d_1 + \dots + d_k$ e $p = q_1 + \dots + q_\ell$.

Vamos provar que

$$B' = \{(\tilde{v}_1, \tilde{w}_1), (\tilde{v}_1, -\tilde{w}_1), \dots, (\tilde{v}_m, \tilde{w}_m), (\tilde{v}_m, -\tilde{w}_m), (\tilde{u}_1, \mathbf{0}), \dots, (\tilde{u}_p, \mathbf{0})\}$$

é uma base de $\hat{V}$.

Como $\dim(\hat{V}) = \dim(V) = 2m + p$, basta provar que B' é LI.

De fato, seja

$$\begin{aligned} &(\lambda_1^1 + i\delta_1^1)(\tilde{v}_1, \tilde{w}_1) + (\lambda_1^2 + i\delta_1^2)(\tilde{v}_1, -\tilde{w}_1) + \dots + (\lambda_m^1 + i\delta_m^1)(\tilde{v}_m, \tilde{w}_m) \\ &+ (\lambda_m^2 + i\delta_m^2)(\tilde{v}_m, -\tilde{w}_m) + \lambda_1^3(\tilde{u}_1, \mathbf{0}) + \dots + \lambda_p^3(\tilde{u}_p, \mathbf{0}) = (\mathbf{0}, \mathbf{0}), \end{aligned}$$

onde $\lambda_i^1, \lambda_j^2, \lambda_k^3, \delta_i^1, \delta_i^2 \in \mathbb{R}$, $i, j = 1, \dots, m$, $k = 1, \dots, p$.

Então,

$$\begin{aligned} &(\lambda_1^1 \tilde{v}_1 - \delta_1^1 \tilde{w}_1, \delta_1^1 \tilde{v}_1 + \lambda_1^1 \tilde{w}_1) + (\lambda_1^2 \tilde{v}_1 + \delta_1^2 \tilde{w}_1, \delta_1^2 \tilde{v}_1 - \lambda_1^2 \tilde{w}_1) + \dots \\ &+ (\lambda_m^1 \tilde{v}_m - \delta_m^1 \tilde{w}_m, \delta_m^1 \tilde{v}_m + \lambda_m^1 \tilde{w}_m) + (\lambda_m^2 \tilde{v}_m + \delta_m^2 \tilde{w}_m, \delta_m^2 \tilde{v}_m - \lambda_m^2 \tilde{w}_m) \\ &+ (\lambda_1^3 \tilde{u}_1, \mathbf{0}) + \dots + (\lambda_p^3 \tilde{u}_p, \mathbf{0}) = (\mathbf{0}, \mathbf{0}) \\ \implies &(\lambda_1^1 + \lambda_1^2) \tilde{v}_1 + (-\delta_1^1 + \delta_1^2) \tilde{w}_1 + \dots + (\lambda_m^1 + \lambda_m^2) \tilde{v}_m + (-\delta_m^1 + \delta_m^2) \tilde{w}_m \\ &+ \lambda_1^3 \tilde{u}_1 + \dots + \lambda_p^3 \tilde{u}_p = \mathbf{0} \end{aligned}$$

e

$$(\delta_1^1 + \delta_1^2)\widetilde{v}_1 + (\lambda_1^1 - \lambda_1^2)\widetilde{w}_1 + \dots + (\delta_m^1 + \delta_m^2)\widetilde{v}_m + (\lambda_m^1 - \lambda_m^2)\widetilde{w}_m = \mathbf{0}.$$

Como $\{\widetilde{v}_1, \widetilde{w}_1, \dots, \widetilde{v}_m, \widetilde{w}_m, \widetilde{u}_1, \dots, \widetilde{u}_p\}$ é uma base de V , temos que

$$\begin{aligned} \lambda_j^1 + \lambda_j^2 &= \lambda_j^1 - \lambda_j^2 = 0, \\ \delta_j^1 + \delta_j^2 &= -\delta_j^1 + \delta_j^2 = 0, \end{aligned} \quad j = 1, \dots, m,$$

e $\lambda_i^3 = 0, i = 1, \dots, p$.

Logo, $\lambda_j^1 = \lambda_j^2 = \delta_j^1 = \delta_j^2 = 0, j = 1, \dots, m$, e $\lambda_i^3 = 0, i = 1, \dots, p$.

Observe, também, que se

$$\begin{cases} T(\widetilde{v}_i) &= a_j \widetilde{v}_i + b_j \widetilde{w}_i + \widetilde{v}_{i+1} \\ T(\widetilde{w}_i) &= -b_j \widetilde{v}_i + a_j \widetilde{w}_i + \widetilde{w}_{i+1}, \end{cases}$$

ou

$$\begin{cases} T(\widetilde{v}_i) &= a_j \widetilde{v}_i + b_j \widetilde{w}_i \\ T(\widetilde{w}_i) &= -b_j \widetilde{v}_i + a_j \widetilde{w}_i, \end{cases}$$

então:

$$\begin{aligned} \widehat{T}(\widetilde{v}_i, \widetilde{w}_i) &= (T(\widetilde{v}_i), T(\widetilde{w}_i)) \\ &= (a_j \widetilde{v}_i + b_j \widetilde{w}_i + \widetilde{v}_{i+1}, -b_j \widetilde{v}_i + a_j \widetilde{w}_i + \widetilde{w}_{i+1}) \\ &= (a_j - ib_j)(\widetilde{v}_i, \widetilde{w}_i) + (\widetilde{v}_{i+1}, \widetilde{w}_{i+1}) \\ &= \overline{\mu}_j(\widetilde{v}_i, \widetilde{w}_i) + (\widetilde{v}_{i+1}, \widetilde{w}_{i+1}), \end{aligned}$$

e

$$\begin{aligned} \widehat{T}(\widetilde{v}_i, -\widetilde{w}_i) &= (T(\widetilde{v}_i), -T(\widetilde{w}_i)) \\ &= (a_j \widetilde{v}_i + b_j \widetilde{w}_i + \widetilde{v}_{i+1}, b_j \widetilde{v}_i - a_j \widetilde{w}_i - \widetilde{w}_{i+1}) \\ &= (a_j + ib_j)(\widetilde{v}_i, -\widetilde{w}_i) + (\widetilde{v}_{i+1}, -\widetilde{w}_{i+1}) \\ &= \mu_j(\widetilde{v}_i, -\widetilde{w}_i) + (\widetilde{v}_{i+1}, -\widetilde{w}_{i+1}), \end{aligned}$$

ou

$$\begin{aligned} \widehat{T}(\widetilde{v}_i, \widetilde{w}_i) &= (T(\widetilde{v}_i), T(\widetilde{w}_i)) \\ &= (a_j \widetilde{v}_i + b_j \widetilde{w}_i, -b_j \widetilde{v}_i + a_j \widetilde{w}_i) \\ &= (a_j - ib_j)(\widetilde{v}_i, \widetilde{w}_i) \\ &= \overline{\mu}_j(\widetilde{v}_i, \widetilde{w}_i), \end{aligned}$$

e

$$\begin{aligned} \widehat{T}(\widetilde{v}_i, -\widetilde{w}_i) &= (T(\widetilde{v}_i), -T(\widetilde{w}_i)) \\ &= (a_j \widetilde{v}_i + b_j \widetilde{w}_i, b_j \widetilde{v}_i - a_j \widetilde{w}_i) \\ &= (a_j + ib_j)(\widetilde{v}_i, -\widetilde{w}_i) \\ &= \mu_j(\widetilde{v}_i, -\widetilde{w}_i), \end{aligned}$$

Reordenando a base $\mathcal{B}'$ de $\hat{V}$, obtemos a base $\mathcal{B}'' = \{(\tilde{v}_1, \tilde{w}_1), \dots, (\tilde{v}_m, \tilde{w}_m), (\tilde{v}_1, -\tilde{w}_1), \dots, (\tilde{v}_m, -\tilde{w}_m), (\tilde{u}_1, \mathbf{0}), \dots, (\tilde{u}_p, \mathbf{0})\}$ de $\hat{V}$ tal que $[\hat{T}]_{\mathcal{B}''}$ está na forma canônica de Jordan. Como no caso complexo já provamos a unicidade, obtemos a unicidade no caso real.

Exemplo 4.1

Seja T o operador linear sobre $\mathbb{R}^3$ que é representado em relação à base canônica pela matriz

$$A = \begin{pmatrix} 0 & 0 & 0 & -8 \\ 1 & 0 & 0 & 16 \\ 0 & 1 & 0 & -14 \\ 0 & 0 & 1 & 6 \end{pmatrix}.$$

Logo,

$$\begin{aligned} p_c &= \det(xI - A) = \det \begin{pmatrix} x & 0 & 0 & 8 \\ -1 & x & 0 & -16 \\ 0 & -1 & x & 14 \\ 0 & 0 & -1 & x-6 \end{pmatrix} \\ &= x \det \begin{pmatrix} x & 0 & -16 \\ -1 & x & 14 \\ 0 & -1 & x-6 \end{pmatrix} + \det \begin{pmatrix} 0 & 0 & 8 \\ -1 & x & 14 \\ 0 & -1 & x-6 \end{pmatrix} \\ &= x^2 \det \begin{pmatrix} x & 14 \\ -1 & x-6 \end{pmatrix} + x \det \begin{pmatrix} 0 & -16 \\ -1 & x-6 \end{pmatrix} + \det \begin{pmatrix} 0 & 8 \\ -1 & x-6 \end{pmatrix} \\ &= x^2(x(x-6) + 14) + x(-16) + 8 = x^2(x^2 - 6x + 14) - 16x + 8 \\ &= x^4 - 6x^3 + 14x^2 - 16x + 8 = (x^2 - 2x + 2)(x - 2)^2 \\ &= (x - (1 + i))(x - (1 - i))(x - 2)^2. \end{aligned}$$

Como

$$A - 2I = \begin{pmatrix} -2 & 0 & 0 & -8 \\ 1 & -2 & 0 & 16 \\ 0 & 1 & -2 & -14 \\ 0 & 0 & 1 & 4 \end{pmatrix}$$

tem posto igual a 3, temos que $\dim(\text{Ker}(A - 2I)) = 1$, ou seja, a forma canônica de Jordan real de T tem apenas um bloco 2×2 elementar associado ao autovalor 2 e $p_m = p_c$.

Sendo que $x^2 - 2x + 2$ é um polinômio primo em $\mathbb{R}[x]$, temos que existe apenas um bloco 2×2 associado à raiz complexa $1 + i$. Ou seja,

$$B = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix},$$

é a forma canônica de Jordan real do operador T .

Para determinar uma base $\mathcal{B} = \{v_1, v_2, v_3, v_4\}$ de $\mathbb{R}^4$ tal que $[T]_{\mathcal{B}} = B$, devemos determinar primeiro um autovetor associado ao autovalor complexo $1 + i$ da matriz A .

Seja $X = (x, y, z, w)$ tal que $((1 + i)I - A)X = \mathbf{0}$, isto é,

$$\begin{pmatrix} 1+i & 0 & 0 & 8 \\ -1 & 1+i & 0 & -16 \\ 0 & -1 & 1+i & 14 \\ 0 & 0 & -1 & -5+i \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Então,

$$\begin{cases} (1+i)x & + & 8w & = & 0 \\ -x & + & (1+i)y & - & 16w & = & 0 \\ & - & y & + & (1+i)z & + & 14w & = & 0 \\ & & & - & z & + & (-5+i)w & = & 0. \end{cases}$$

ou seja,

$$w = -\frac{1}{8}(1+i)x, \quad z = (-5+i)w \quad \text{e} \quad y = (1+i)(-5+i)w + 14w = (8-4i)w.$$

Fazendo $x = 8$, obtemos $w = -(1+i)$, $z = 6+4i$ e $y = -(12+4i)$. Logo, $X = (8, -12, 6, -1) + i(0, -4, 4, -1)$.

Como $u = (8, -12, 6, -1)$ é a parte real e $v = (0, -4, 4, -1)$ é a parte imaginária do autovetor X , os primeiros vetores da base $\mathcal{B}$ são

$$v_1 = u + v = (8, -16, 10, -2) \quad \text{e} \quad v_2 = u - v = (8, -8, 2, 0).$$

Os vetores v_3 e v_4 da base $\mathcal{B}$ devem ser tais que

$$v_4 = (A - 2I)v_3 \quad \text{e} \quad (A - 2I)^2 v_3 = \mathbf{0}.$$

Como

$$(A - 2I)^2 = \begin{pmatrix} -2 & 0 & 0 & -8 \\ 1 & -2 & 0 & 16 \\ 0 & 1 & -2 & -14 \\ 0 & 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} -2 & 0 & 0 & -8 \\ 1 & -2 & 0 & 16 \\ 0 & 1 & -2 & -14 \\ 0 & 0 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 4 & 0 & -8 & -16 \\ -4 & 4 & 16 & 24 \\ 1 & -4 & -10 & -12 \\ 0 & 1 & 2 & 2 \end{pmatrix}$$

temos que $v_3 = (a, b, c, d) \in \text{Ker}((A - 2I)^2)$ se, e somente se, $4a = 8c + 16d$ e $b = -2c - 2d$. Fazendo $c = d = 1$, obtemos $v_3 = (6, -4, 1, 1)$ e $v_4 = (A - 2I)v_3 = (-20, 30, -20, 5)$, pois

$$\begin{pmatrix} -2 & 0 & 0 & -8 \\ 1 & -2 & 0 & 16 \\ 0 & 1 & -2 & -14 \\ 0 & 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} 6 \\ -4 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -20 \\ 30 \\ -20 \\ 5 \end{pmatrix}.$$

□

Exemplo 4.2

Seja T um operador linear sobre um espaço vetorial real V tal que

$p_c = p_1^3 p_2^3 (x - 4)^4$ e $p_m = p_1^2 p_2^3 (x - 4)^3$, onde

$$p_1 = x^2 - 2x + 5 = (x - (1 + 2i))(x - (1 - 2i)),$$

e

$$p_2 = x^2 + 4 = (x - 2i)(x + 2i).$$

Então, a matriz do operador T na forma canônica tem: um bloco 4×4 elementar de Jordan real associado ao autovalor complexo $1 + 2i$; um bloco 2×2 elementar de Jordan real associado ao autovalor complexo $1 + 2i$; um bloco 6×6 elementar de Jordan real associado ao autovalor complexo $2i$; um bloco 3×3 elementar de Jordan associado ao autovalor 4 e um bloco 1×1 elementar de Jordan associado ao autovalor 4 . Ou seja,

$$\begin{pmatrix} \begin{matrix} 1 & -2 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & 0 & 1 & -2 \\ 0 & 1 & 2 & 1 \end{matrix} & \begin{matrix} 1 & -2 \\ 2 & 1 \end{matrix} & \begin{matrix} 0 & -2 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -2 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -2 \\ 0 & 0 & 0 & 1 & 2 & 0 \end{matrix} & \begin{matrix} 4 & 0 & 0 \\ 1 & 4 & 0 \\ 0 & 1 & 4 \end{matrix} & 4 \end{pmatrix}_{16 \times 16}$$

é a forma canônica de Jordan real do operador T . □

Exemplo 4.3

Seja A uma matriz 7×7 real cujo polinômio minimal é $p_m = (x-1)^2(x+1)^2x$.

Vamos determinar todas as matrizes $\tilde{A}$ na forma canônica de Jordan real que podem ser semelhantes à matriz A .

Seja $p_c = (x-1)^a(x+1)^b x^c$ o polinômio característico de A .

Como $a + b + c = 7$, $a \geq 2$, $b \geq 2$ e $c \geq 1$, temos as seguintes 8 possibilidades:

- $a = 2$, $b = 2$ e $c = 3$.

Nesse caso:

$$\tilde{A} = \begin{pmatrix} 1 & 0 & & & & & \\ 1 & 1 & & & & & \\ & & -1 & 0 & & & \\ & & 1 & -1 & & & \\ & & & & 0 & & \\ & & & & & 0 & \\ & & & & & & 0 \end{pmatrix}.$$

- $a = 2$, $b = 3$, $c = 2$.

Nesse caso:

$$\tilde{A} = \begin{pmatrix} 1 & 0 & & & & & \\ 1 & 1 & & & & & \\ & & -1 & 0 & & & \\ & & 1 & -1 & & & \\ & & & & -1 & & \\ & & & & & 0 & \\ & & & & & & 0 \end{pmatrix}.$$

- $a = 2$, $b = 4$, $c = 1$.

Nesse caso:

$$\tilde{A} = \begin{pmatrix} 1 & 0 & & & & & \\ 1 & 1 & & & & & \\ & & -1 & 0 & & & \\ & & 1 & -1 & & & \\ & & & & -1 & 0 & \\ & & & & 1 & -1 & \\ & & & & & & 0 \end{pmatrix} \quad \text{ou} \quad \tilde{A} = \begin{pmatrix} 1 & 0 & & & & & \\ 1 & 1 & & & & & \\ & & -1 & 0 & & & \\ & & 1 & -1 & & & \\ & & & & -1 & & \\ & & & & & -1 & \\ & & & & & & 0 \end{pmatrix}.$$

- $a = 3, b = 2, c = 2.$

Nesse caso:

$$\tilde{A} = \begin{pmatrix} 1 & 0 & & & & \\ 1 & 1 & & & & \\ & & 1 & & & \\ & & & -1 & 0 & \\ & & & 1 & -1 & \\ & & & & & 0 \\ & & & & & & 0 \end{pmatrix}.$$

- $a = 3, b = 3, c = 1.$

Nesse caso:

$$\tilde{A} = \begin{pmatrix} 1 & 0 & & & & \\ 1 & 1 & & & & \\ & & 1 & & & \\ & & & -1 & 0 & \\ & & & 1 & -1 & \\ & & & & & -1 \\ & & & & & & 0 \end{pmatrix}.$$

- $a = 4, b = 2, c = 1.$

Nesse caso:

$$\tilde{A} = \begin{pmatrix} 1 & 0 & & & & \\ 1 & 1 & & & & \\ & & 1 & 0 & & \\ & & 1 & 1 & & \\ & & & & -1 & 0 \\ & & & & 1 & -1 \\ & & & & & & 0 \end{pmatrix} \quad \text{ou} \quad \tilde{A} = \begin{pmatrix} 1 & 0 & & & & \\ 1 & 1 & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & 1 & \\ & & & & & -1 & 0 \\ & & & & & 1 & -1 \\ & & & & & & & 0 \end{pmatrix}.$$

□

5. Operadores Semi-Simples

Definição 5.1

Seja V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{R}$ dos números reais. Dizemos que um operador linear $L : V \longrightarrow V$ é *semi-simples* se $\hat{L} : \hat{V} \longrightarrow \hat{V}$ é um operador linear diagonalizável.

Lema 5.1

Seja $\widehat{V}$ o complexificado do espaço vetorial real V . Seja $T : \widehat{V} \longrightarrow \widehat{V}$ um operador linear. Então existe um operador linear $L : V \longrightarrow V$ tal que $\widehat{L} = T$ se, e somente se $\sigma T = T\sigma$.

Prova.

($\implies$) Suponhamos que $T = \widehat{L}$. Então, como $T(u, v) = (L(u), L(v))$, temos que

$$\sigma T(u, v) = \sigma(L(u), L(v)) = (L(u), -L(v)) = (L(u), L(-v)) = T\sigma(u, v).$$

($\impliedby$) Seja $\pi : \widehat{V} \longrightarrow V$ a aplicação definida por $\pi(u, v) = u$. Como $\pi((u_1, v_1) + (u_2, v_2)) = \pi(u_1 + u_2, v_1 + v_2) = u_1 + u_2 = \pi(u_1, v_1) + \pi(u_2, v_2)$, e $\pi(\lambda(u, v)) = \pi(\lambda u, \lambda v) = \lambda u$, se $\lambda \in \mathbb{R}$, temos que π é um operador linear real.

Considere a aplicação $L : V \longrightarrow V$ dada por $L(u) = \pi \circ T(u, \mathbf{0})$.

Como, para todo $\lambda \in \mathbb{R}$ vale que

$$\begin{aligned} L(\lambda u + w) &= (\pi \circ T)(\lambda u + w, \mathbf{0}) \\ &= (\pi \circ T)(\lambda(u, \mathbf{0}) + (w, \mathbf{0})) \\ &= \pi(\lambda T(u, \mathbf{0}) + T(w, \mathbf{0})) \\ &= \lambda \pi \circ T(u, \mathbf{0}) + \pi \circ T(w, \mathbf{0}) \\ &= \lambda L(u) + L(w), \end{aligned}$$

temos que L é um operador linear.

Resta mostrar que $T(u, v) = (L(u), L(v))$ para todo par $u, v \in V$.

De fato, como $L(u) = \pi(T(u, \mathbf{0}))$, temos que $T(u, \mathbf{0}) = (L(u), x)$ para algum $x \in V$.

Mas, como $\sigma T = T\sigma$, temos que

$$\sigma \circ T(u, \mathbf{0}) = (L(u), -x) = T\sigma(u, \mathbf{0}) = T(u, \mathbf{0}) = (L(u), x).$$

Logo, $x = 0$, ou seja, $T(u, \mathbf{0}) = (L(u), \mathbf{0})$.

Assim,

$$\begin{aligned} T(u, v) &= T((u, \mathbf{0}) + i(v, \mathbf{0})) = T(u, \mathbf{0}) + iT(v, \mathbf{0}) \\ &= (L(u), \mathbf{0}) + i(L(v), \mathbf{0}) = (L(u), L(v)) = \widehat{L}(u, v), \end{aligned}$$

ou seja, $\widehat{L} = T$. ■

Teorema 5.1

Seja V um espaço vetorial real de dimensão finita e seja $L \in \mathcal{L}(V, V)$.

Então existe um único operador linear semi-simples S e um único operador linear nilpotente N , tais que:

$$(1) L = S + N;$$

$$(2) SN = NS.$$

Além disso, S e N são polinômios em L .

Prova.

Existência.

Seja $\hat{L} : \hat{V} \rightarrow \hat{V}$ o operador linear complexificado do operador L . Como $\hat{V}$ é um espaço vetorial de dimensão finita sobre o corpo $\mathbb{C}$ dos números complexos, existe um único operador linear $D_0 : \hat{V} \rightarrow \hat{V}$ e um único operador nilpotente $N_0 : \hat{V} \rightarrow \hat{V}$ tais que: $\hat{L} = D_0 + N_0$ e $D_0 N_0 = N_0 D_0$, sendo D_0 e N_0 polinômios em $\hat{L}$.

Sejam $D_1 = \sigma D_0 \sigma^{-1}$ e $N_1 = \sigma N_0 \sigma^{-1}$. Então D_1 é diagonalizável e N_1 é nilpotente.

De fato, se $N_0^k = \mathcal{O}$, temos que

$$N_1^k = (\sigma N_0 \sigma^{-1})^k = \sigma N_0^k \sigma^{-1} = \mathcal{O},$$

ou seja, N_1 é nilpotente.

Se $\{(u_1, v_1), \dots, (u_n, v_n)\}$ é uma base de autovetores de D_0 , temos $D_0(u_i, v_i) = c_i(u_i, v_i)$, $i = 1, \dots, n$, e, então

$$\begin{aligned} D_1(u_i, -v_i) &= \sigma \circ D_0 \circ \sigma^{-1}(u_i, -v_i) \\ &= \sigma \circ D_0(u_i, v_i) \\ &= \sigma(c_i(u_i, v_i)) \\ &= \overline{c_i}(u_i, v_i), \end{aligned}$$

ou seja, $\{(u_1, -v_1), \dots, (u_n, -v_n)\}$ é uma base de autovetores de D_1 . Logo, D_1 é diagonalizável.

Além disso,

$$\begin{aligned} D_1 N_1 &= \sigma D_0 \sigma^{-1} \sigma N_0 \sigma^{-1} = \sigma D_0 N_0 \sigma^{-1} \\ &= \sigma N_0 D_0 \sigma^{-1} = \sigma N_0 \sigma^{-1} \sigma D_0 \sigma^{-1} = N_1 D_1, \end{aligned}$$

ou seja, D_1 e N_1 comutam.

Como $\sigma\hat{L} = \hat{L}\sigma$, temos que

$$\hat{L} = \sigma\hat{L}\sigma^{-1} = \sigma(D_0 + N_0)\sigma^{-1} = \sigma D_0 \sigma^{-1} + \sigma N_0 \sigma^{-1} = D_1 + N_1.$$

Então, pela unicidade, $D_0 = \sigma D_0 \sigma^{-1}$ e $N_0 = \sigma N_0 \sigma^{-1}$, ou seja, $\sigma D_0 = D_0 \sigma$ e $\sigma N_0 = N_0 \sigma$.

Pelo lema anterior, existem $S : V \rightarrow V$ e $N : V \rightarrow V$ operadores lineares tais que $D_0 = \hat{S}$ e $N_0 = \hat{N}$.

Logo, S é semi-simples e N é nilpotente, pois se $N_0^k = \mathcal{O}$, temos

$$N_0^k(u, \mathbf{0}) = (N^k(u), \mathbf{0}) = (\mathbf{0}, \mathbf{0}),$$

para todo $u \in V$, ou seja, $N^k = \mathcal{O}$.

Temos, também, que

- $SN = NS$, pois:

$$\begin{aligned} (SN - NS)^{\wedge}(u, v) &= ((SN - NS)(u), (SN - NS)(v)) \\ &= (SN(u), SN(v)) - (NS(u), NS(v)) \\ &= \hat{S}\hat{N}(u, v) - \hat{N}\hat{S}(u, v) \\ &= (\hat{S}\hat{N} - \hat{N}\hat{S})(u, v) \\ &= (D_0 N_0 - N_0 D_0)(u, v) \\ &= (\mathbf{0}, \mathbf{0}). \end{aligned}$$

- $L = S + N$, pois, como $\hat{L} = D_0 + N_0 = \hat{S} + \hat{N}$, temos que $(L - S - N)^{\wedge} = \mathcal{O}$.

Unicidade.

Sejam $S_1 : V \rightarrow V$ um operador semi-simples e $N_1 : V \rightarrow V$ um operador nilpotente tais que: $L = S_1 + N_1$ e $S_1 N_1 = N_1 S_1$.

Então, $\hat{L} = \hat{S}_1 + \hat{N}_1$, onde $\hat{S}_1$ é diagonalizável e $\hat{N}_1$ é nilpotente, e

$$\hat{S}_1 \hat{N}_1 = \hat{S}_1 \hat{N}_1 = \hat{N}_1 \hat{S}_1 = \hat{N}_1 \hat{S}_1.$$

Pela unicidade da decomposição de $\hat{L}$ como uma soma de um operador diagonalizável com um operador nilpotente que comutam, temos que $\hat{S}_1 = D_0$ e $\hat{N}_1 = N_0$. Mas, como $D_0 = \hat{S}$ e $N_0 = \hat{N}$, temos que $\hat{S}_1 = \hat{S}$ e $\hat{N}_1 = \hat{N}$, ou seja, $S_1 = S$ e $N_1 = N$.

- Vamos provar agora que S e N são polinômios em L .

Seja

$$p_m = (x - \mu_1)^{r_1} (x - \overline{\mu_1})^{r_1} \dots (x - \mu_k)^{r_k} (x - \overline{\mu_k})^{r_k} (x - c_1)^{s_1} \dots (x - c_\ell)^{s_\ell}$$

o polinômio minimal de L e $\widehat{L}$, onde μ_i e $\overline{\mu_i}$, $i = 1, \dots, k$, são as raízes complexas e c_j , $j = 1, \dots, \ell$ são as raízes reais.

Sejam

$$f_{c_j} = \frac{p_m}{(x - c_j)^{s_j}}, \quad f_{\mu_i} = \frac{p_m}{(x - \mu_i)^{r_i}}, \quad e \quad f_{\overline{\mu_i}} = \frac{p_m}{(x - \overline{\mu_i})^{r_i}}.$$

Então, $f_{c_j} \in \mathbb{R}[x]$ e $f_{\overline{\mu_i}} = \overline{f_{\mu_i}}$, $j = 1, \dots, \ell$ e $i = 1, \dots, k$.

Como $f_{\mu_1}, f_{\overline{\mu_1}}, \dots, f_{\mu_k}, f_{\overline{\mu_k}}, f_{c_1}, f_{c_\ell}$ são polinômios primos entre si, existem polinômios $h_{\mu_i}, h_{\overline{\mu_i}}, h_{c_j} \in \mathbb{C}[x]$, $i = 1, \dots, k$ e $j = 1, \dots, \ell$, tais que:

$$h_{\mu_1} f_{\mu_1} + h_{\overline{\mu_1}} \overline{f_{\mu_1}} + \dots + h_{\mu_k} f_{\mu_k} + h_{\overline{\mu_k}} \overline{f_{\mu_k}} + h_{c_1} f_{c_1} + \dots + h_{c_\ell} f_{c_\ell} = 1. \quad (I)$$

Então:

$$\overline{h_{\mu_1}} \overline{f_{\mu_1}} + \overline{h_{\overline{\mu_1}}} f_{\mu_1} + \dots + \overline{h_{\mu_k}} \overline{f_{\mu_k}} + \overline{h_{\overline{\mu_k}}} f_{\mu_k} + \overline{h_{c_1}} f_{c_1} + \dots + \overline{h_{c_\ell}} f_{c_\ell} = 1. \quad (II)$$

Logo, por (I) e (II), obtemos que:

$$\begin{aligned} \frac{h_{\mu_1} + \overline{h_{\mu_1}}}{2} f_{\mu_1} + \frac{(\overline{h_{\mu_1}} + h_{\mu_1})}{2} f_{\overline{\mu_1}} + \dots + \frac{h_{\mu_k} + \overline{h_{\mu_k}}}{2} f_{\mu_k} + \frac{(\overline{h_{\mu_k}} + h_{\mu_k})}{2} f_{\overline{\mu_k}} \\ + \frac{h_{c_1} + \overline{h_{c_1}}}{2} f_{c_1} + \dots + \frac{h_{c_\ell} + \overline{h_{c_\ell}}}{2} f_{c_\ell} = 1, \end{aligned}$$

Fazendo $h_i = \frac{h_{\mu_i} + \overline{h_{\mu_i}}}{2} \in \mathbb{C}[x]$, $i = 1, \dots, k$ e $g_j = \frac{h_{c_j} + \overline{h_{c_j}}}{2} \in \mathbb{R}[x]$, temos que

$$h_1 f_{\mu_1} + \overline{h_1} \overline{f_{\mu_1}} + \dots + h_k f_{\mu_k} + \overline{h_k} \overline{f_{\mu_k}} + g_1 f_{c_1} + \dots + g_\ell f_{c_\ell} = 1.$$

Logo, como foi provado no Teorema 5.1, $D_0 = p(\widehat{L})$, onde

$$\begin{aligned} p = \mu_1 h_1 f_{\mu_1} + \overline{\mu_1} \overline{h_1} \overline{f_{\mu_1}} + \dots + \mu_k h_k f_{\mu_k} + \overline{\mu_k} \overline{h_k} \overline{f_{\mu_k}} \\ + c_1 g_1 f_{c_1} + \dots + c_\ell g_\ell f_{c_\ell} \in \mathbb{R}[x]. \end{aligned}$$

Então, como $p \in \mathbb{R}[x]$ e $\widehat{S} = D_0$, temos que $\widehat{S} = D_0 = p(\widehat{L}) = \widehat{p(L)}$, ou seja, $S = p(L)$ e $N = L - S = L - p(L) = q(L)$, onde $q = x - p \in \mathbb{R}[x]$. ■



Espaços Vetoriais com Produto Interno

1. Produto Interno - Definições básicas

Definição 1.1

Seja V um espaço vetorial sobre o corpo $\mathbb{K}$, sendo $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

Um *produto interno sobre V* é uma aplicação

$$\begin{aligned} \langle \cdot, \cdot \rangle: V \times V &\longrightarrow \mathbb{K} \\ (u, v) &\longmapsto \langle u, v \rangle, \end{aligned}$$

que satisfaz as seguintes condições:

(a) Para cada $v \in V$, a aplicação $u \mapsto \langle u, v \rangle$ é linear, ou seja,

$$\langle \lambda u + w, v \rangle = \lambda \langle u, v \rangle + \langle w, v \rangle.$$

quaisquer que sejam $u, w \in V$ e $\lambda \in \mathbb{K}$.

(b) $\langle u, v \rangle = \overline{\langle v, u \rangle}$, onde a barra indica conjugação complexa.

(c) $\langle u, u \rangle \geq 0$ e $\langle u, u \rangle = 0$ se, e só se, $u = \mathbf{0}$.

Observação 1.1

$$\begin{aligned} \text{(1)} \quad \langle u, \lambda v + w \rangle &= \overline{\langle \lambda v + w, u \rangle} = \overline{\lambda \langle v, u \rangle + \langle w, u \rangle} = \overline{\lambda} \overline{\langle v, u \rangle} + \overline{\langle w, u \rangle} \\ &= \overline{\lambda} \langle u, v \rangle + \langle u, w \rangle. \end{aligned}$$

(2) No caso $\mathbb{K} = \mathbb{R}$, temos que $\langle u, \lambda v + w \rangle = \lambda \langle u, v \rangle + \langle u, w \rangle$.

(3) No caso $\mathbb{K} = \mathbb{C}$, a conjugação complexa em (b) é necessária, pois, sem ela, teríamos que $\langle u, u \rangle > 0$ e $\langle iu, iu \rangle = -\langle u, u \rangle > 0$, se $u > 0$.

Exemplo 1.1

Sejam $X = (x_1, \dots, x_n), Y = (y_1, \dots, y_n) \in \mathbb{K}^n$, onde $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

Então, $\langle X, Y \rangle = \sum_{i=1}^n x_i \overline{y_i}$ é um produto interno sobre $\mathbb{K}^n$, denominado *produto interno canônico*. $\square$

Exemplo 1.2

Seja $V = \mathbb{K}^{n \times n}$, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$, o espaço vetorial das matrizes $n \times n$ sobre $\mathbb{K}$. Como $\mathbb{K}^{n \times n}$ é isomorfo a $\mathbb{K}^{n^2}$ de uma maneira natural, temos:

$$\langle A, B \rangle = \sum_{i,j=1}^n A_{ij} \overline{B_{ij}}$$

define um produto interno sobre V .

Seja B^* a conjugada da matriz transposta de B , ou seja, $B_{ij}^* = \overline{B_{ji}}$. Então,

$$\langle A, B \rangle = \text{Tr}(AB^*) = \text{Tr}(B^*A),$$

onde $\text{Tr}(\cdot)$ é o traço da matriz.

De fato,

$$\text{Tr}(AB^*) = \sum_{j=1}^n (AB^*)_{jj} = \sum_{j=1}^n \sum_{k=1}^n A_{jk} B_{kj}^* = \sum_{j=1}^n \sum_{k=1}^n A_{jk} \overline{B_{jk}}.$$

$\square$

Exemplo 1.3

Seja $\mathbb{K}^{n \times 1}$ o espaço vetorial das matrizes $n \times 1$, sendo $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

Sejam $X = (x_1, \dots, x_n)$ e $Y = (y_1, \dots, y_n)$ em $\mathbb{K}^{n \times 1}$. Então,

$$\langle X, Y \rangle = \sum_{i=1}^n x_i \overline{y_i} = Y^* X$$

é um produto interno sobre $\mathbb{K}^{n \times 1}$.

Generalizando, temos que

$$\langle X, Y \rangle_Q = Y^* Q^* Q X = (QY)^* Q X = \langle QX, QY \rangle$$

é um produto interno sobre $\mathbb{K}^{n \times 1}$, onde Q é uma matriz $n \times n$ invertível sobre $\mathbb{K}$.

De fato:

- $\langle \lambda X + Z, Y \rangle_Q = \langle Q(\lambda X + Z), Y \rangle$
 $= \langle \lambda QX + QZ, QY \rangle$
 $= \lambda \langle QX, QY \rangle + \langle QZ, QY \rangle$
 $= \lambda \langle X, Y \rangle_Q + \langle Z, Y \rangle_Q$
- $\langle X, Y \rangle_Q = \langle QX, QY \rangle = \overline{\langle QY, QX \rangle} = \overline{\langle Y, X \rangle_Q}$
- $\langle X, X \rangle_Q = \langle QX, QX \rangle \geq 0$
- $\langle X, X \rangle_Q = \langle QX, QX \rangle = 0 \iff QX = \mathbf{0}$
 $\iff X = \mathbf{0}$ pois Q é invertível.

□

Exemplo 1.4

Seja $V = \mathbb{C}^0([0, 1], \mathbb{C})$ o espaço vetorial das funções contínuas definidas no intervalo $[0, 1]$ com valores complexos. Então,

$$\langle f, g \rangle = \int_0^1 f(t) \overline{g(t)} dt$$

é um produto interno sobre V . □

Observação 1.2

Sejam V e W espaços vetoriais sobre $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) e suponhamos que $\langle \cdot, \cdot \rangle$ seja um produto interno sobre W . Se $T : V \longrightarrow W$ é uma transformação linear injetora, então

$$\langle u, v \rangle_T = \langle T(u), T(v) \rangle$$

define um produto interno sobre V . (Verifique como exercício).

O produto interno sobre $\mathbb{K}^{n \times 1}$ definido no exemplo 1.3 é um caso particular dessa observação. De fato, tomando $T : \mathbb{K}^{n \times 1} \longrightarrow \mathbb{K}^{n \times 1}$, $T(X) = QX$, temos que T é injetora e que

$$\langle X, Y \rangle_T = \langle T(X), T(Y) \rangle = \langle QX, QY \rangle = \langle X, Y \rangle_Q.$$

Exemplo 1.5

Sejam V um espaço vetorial de dimensão finita sobre $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) e seja $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base de V . Seja $\{e_1, \dots, e_n\}$ a base canônica de $\mathbb{K}^n$ e seja $T : V \longrightarrow \mathbb{K}^n$ a transformação linear de V em $\mathbb{K}^n$ tal que $T(v_j) = e_j$, $j = 1, \dots, n$. Se tomarmos o produto interno canônico sobre $\mathbb{K}^n$, temos:

$$\langle u, v \rangle_T = \left\langle \sum_{j=1}^n x_j v_j, \sum_{k=1}^n y_k v_k \right\rangle_T = \sum_{j=1}^n x_j \bar{y}_j$$

é um produto interno sobre V , onde $u = \sum_{j=1}^n x_j v_j$ e $v = \sum_{k=1}^n y_k v_k$.

Assim, para toda base $\mathcal{B}$ de V existe um produto interno sobre V tal que $\langle v_i, v_j \rangle_{\mathcal{B}} = \delta_{ij}$, $i, j = 1, \dots, n$. Na verdade, é fácil mostrar que existe exatamente um tal produto interno.

- Mostraremos depois que todo produto interno sobre V é determinado por alguma base $\mathcal{B}$ da maneira acima. $\square$

Observação 1.3

Seja V um espaço vetorial complexo com um produto interno $\langle \cdot, \cdot \rangle$. Então, para $v, w \in V$,

$$\langle v, w \rangle = \Re \langle v, w \rangle + i \Im \langle v, w \rangle,$$

onde $\Re \langle v, w \rangle$ e $\Im \langle v, w \rangle$ são as partes real e imaginária do número complexo $\langle v, w \rangle$.

Como, para todo $z \in \mathbb{C}$, $\Im(z) = \Re(-iz)$, temos

$$\langle u, v \rangle = \Re \langle v, w \rangle + i \Re(-i \langle v, w \rangle),$$

ou seja

$$\langle u, v \rangle = \Re \langle v, w \rangle + i \Re(\langle v, iw \rangle).$$

Assim, o produto interno é completamente determinado por sua parte real.

Definição 1.2

Seja $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$) um produto interno sobre V .

- Para $v \in V$, $\|v\| = \sqrt{\langle v, v \rangle}$ é chamada a **norma de v** em relação ao produto interno $\langle \cdot, \cdot \rangle$.
- A **forma quadrática** determinada pelo produto interno $\langle \cdot, \cdot \rangle$ é a função

$$\begin{aligned} V &\longrightarrow \mathbb{K} \\ v &\longmapsto \|v\|^2 = \langle v, v \rangle \end{aligned}$$

Então,

$$\|v \pm w\|^2 = \langle v \pm w, v \pm w \rangle = \|v\|^2 \pm 2\Re \langle v, w \rangle + \|w\|^2.$$

Assim, no caso real,

$$\|v + w\|^2 = \|v\|^2 + 2\langle v, w \rangle + \|w\|^2,$$

e

$$\|v - w\|^2 = \|v\|^2 - 2\langle v, w \rangle + \|w\|^2,$$

implica que:

$$\langle v, w \rangle = \frac{1}{4} (\|v + w\|^2 - \|v - w\|^2) \quad (I)$$

No caso complexo,

$$\Re \langle v, w \rangle = \frac{1}{4} (\|v + w\|^2 - \|v - w\|^2)$$

e

$$\Re \langle v, iw \rangle = \frac{1}{4} (\|v + iw\|^2 - \|v - iw\|^2)$$

Logo, como $\langle v, w \rangle = \Re \langle v, w \rangle + i \Re \langle v, iw \rangle$, temos

$$\langle v, w \rangle = \frac{1}{4} (\|v + w\|^2 - \|v - w\|^2) + \frac{i}{4} (\|v + iw\|^2 - \|v - iw\|^2) \quad (II)$$

As identidades (I) e (II) são denominadas as *identidades de polarização*.

- Sejam V um espaço vetorial de dimensão finita, $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base ordenada de V e $\langle \cdot, \cdot \rangle$ um produto interno sobre V . Mostraremos que *o produto interno é completamente determinado pelos números:*

$$G_{jk} = \langle v_k, v_j \rangle, \quad k, j = 1, \dots, n.$$

calculados sobre todos os pares de vetores em $\mathcal{B}$.

De fato, se $u = \sum_{k=1}^n x_k v_k$ e $w = \sum_{j=1}^n y_j v_j$, então

$$\langle u, v \rangle = \left\langle \sum_{k=1}^n x_k v_k, \sum_{j=1}^n y_j v_j \right\rangle = \sum_{j,k=1}^n x_k \overline{y_j} \langle v_k, v_j \rangle = \sum_{j,k=1}^n \overline{y_j} G_{jk} x_k = Y^* G X,$$

onde $X = [u]_{\mathcal{B}}$, $Y = [v]_{\mathcal{B}}$ e G é a matriz com entradas $G_{jk} = \langle v_k, v_j \rangle$.

Denominamos *G a matriz do produto interno em relação à base $\mathcal{B}$* .

Como $G_{jk} = \langle v_k, v_j \rangle = \overline{\langle v_j, v_k \rangle} = \overline{G_{kj}}$, temos que G é *hermitiana*, ou seja, $G^* = G$.

Além disso, G é invertível, pois, como $X^* G X > 0$ para todo $X \in \mathbb{K}^{n \times 1} - \{0\}$, temos que $G X \neq 0$, para $X \neq 0$.

Fazendo $X = [v_i]_{\mathcal{B}} = e_i$, e_i sendo o i -ésimo termo da base canônica de $\mathbb{K}^{n \times 1}$, temos que

Uma matriz é chamada *hermitiana* quando é igual ao conjugado da sua própria transposta.

$$G_{ii} = \sum_{k,j=1}^n \overline{x_j} G_{jk} x_k = X^* G X > 0,$$

para todo $i = 1, \dots, n$, ou seja, todas as entradas da diagonal de G são positivas.

Porém, nem toda matriz hermitiana A que possui as entradas da diagonal positivas satisfaz a condição

$$X^* A X > 0,$$

para todo $X \neq \mathbf{0}$. Basta tomar, por exemplo,

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

Observação 1.4

Seja G uma matriz $n \times n$ sobre o corpo $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) tal que $G^* = G$ e $X^* G X > 0$ para todo $X \in \mathbb{K}^{n \times 1} - \{\mathbf{0}\}$. Então G é a matriz de algum produto interno em relação à base $\mathcal{B}$.

De fato, definindo

$$\langle u, w \rangle = Y^* G X,$$

onde $X = [u]_{\mathcal{B}}$ e $Y = [w]_{\mathcal{B}}$, temos que $\langle \cdot, \cdot \rangle$ é um produto interno sobre V e G é a sua matriz em relação à base $\mathcal{B}$, pois:

- $\overline{\langle w, u \rangle} = \overline{X^* G Y} = (X^* G Y)^* = Y^* G^* X = Y^* G X = \langle u, w \rangle.$
- $\langle v_j, v_k \rangle = e_k^* G e_j = G_{kj}, k, j = 1, \dots, n.$
- $\langle \lambda u_1 + u_2, w \rangle = Y^* G [\lambda u_1 + u_2]_{\mathcal{B}} = Y^* G (\lambda [u_1]_{\mathcal{B}} + [u_2]_{\mathcal{B}})$
 $= \lambda Y^* G [u_1]_{\mathcal{B}} + Y^* G [u_2]_{\mathcal{B}} = \lambda \langle u_1, w \rangle + \langle u_2, w \rangle.$

Estabeleceremos agora as propriedades básicas dos conceitos de *comprimento* e *ortogonalidade* que são impostas ao espaço pelo produto interno.

Proposição 1.1

Seja V um espaço vetorial com produto interno. Então, para quaisquer vetores $u, v \in V$ e todo escalar $\lambda \in \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$),

(a) $\|\lambda u\| = |\lambda| \|u\|;$

(b) $\|u\| \geq 0$ e $\|u\| = 0 \iff u = \mathbf{0};$

(c) $|\langle u, v \rangle| \leq \|u\| \|v\|$ (*Desigualdade de Cauchy-Schwarz*);

(d) $|\langle u, v \rangle| = \|u\| \|v\| \iff u$ e v são LD;

(e) $\|u + v\| \leq \|u\| + \|v\|$ (*Desigualdade Triangular*)

Prova.

Como $\|\lambda u\|^2 = \langle \lambda u, \lambda u \rangle = \lambda \bar{\lambda} \langle u, u \rangle = |\lambda|^2 \|u\|^2$, temos que $\|\lambda u\| = |\lambda| \|u\|$.

O item (b) segue imediatamente das propriedades de um produto interno.

Para provarmos o item (c), faremos separadamente o caso real ($\mathbb{K} = \mathbb{R}$) e o caso complexo ($\mathbb{K} = \mathbb{C}$).

• **Caso real.**

Como $\langle u + \lambda v, u + \lambda v \rangle \geq 0$, para todo $\lambda \in \mathbb{R}$, temos

$$\lambda^2 \|v\|^2 + 2\lambda \langle u, v \rangle + \|u\|^2 \geq 0$$

para todo $\lambda \in \mathbb{R}$. Logo, o discriminante Δ da equação do segundo grau $\lambda^2 \|v\|^2 + 2\lambda \langle u, v \rangle + \|u\|^2 = 0$ (na variável λ) é menor ou igual do que zero, isto é,

$$\Delta = 4\langle u, v \rangle^2 - 4\|u\|^2 \|v\|^2 \leq 0.$$

Então, $|\langle u, v \rangle| \leq \|u\| \|v\|$. Além disso, $|\langle u, v \rangle| = \|u\| \|v\|$ se, e somente se, $\Delta = 0$, ou seja, se e somente se, existe um único $\lambda_0 \in \mathbb{R}$ tal que $\langle u + \lambda_0 v, u + \lambda_0 v \rangle = 0$.

Como $\langle u + \lambda_0 v, u + \lambda_0 v \rangle = 0$ se, e somente se, $u + \lambda_0 v = \mathbf{0}$, temos que $|\langle u, v \rangle| = \|u\| \|v\|$ se, e somente se, u e v são LD.

• **Caso complexo.**

Como $\langle u + \lambda v, u + \lambda v \rangle \geq 0$ para todo $\lambda \in \mathbb{C}$, temos que

$$|\lambda|^2 \|v\|^2 + \bar{\lambda} \langle u, v \rangle + \lambda \langle v, u \rangle + \|u\|^2 \geq 0,$$

para todo $\lambda \in \mathbb{C}$.

Logo, $|\lambda|^2 \|v\|^2 + 2\Re(\lambda \langle v, u \rangle) + \|u\|^2 \geq 0$, para todo $\lambda \in \mathbb{C}$, pois $\bar{\lambda} \langle u, v \rangle = \overline{\lambda \langle v, u \rangle}$.

Em particular, $|\lambda|^2 \|v\|^2 + 2\Re(\lambda \langle v, u \rangle) + \|u\|^2 \geq 0$, para todo $\lambda \in \mathbb{R}$. Então,

$$\lambda^2 \|v\|^2 + 2\lambda \Re(\langle v, u \rangle) + \|u\|^2 \geq 0,$$

para todo $\lambda \in \mathbb{R}$.

Assim,

$$\Delta = 4 (\Re \langle u, v \rangle)^2 - 4 \|u\|^2 \|v\|^2 \leq 0,$$

onde Δ é o discriminante da equação (na variável λ)

$$\lambda^2 \|v\|^2 + 2\lambda \Re(\langle v, u \rangle) + \|u\|^2 = 0.$$

Logo,

$$|\Re \langle v, u \rangle| \leq \|u\| \|v\|. \quad (I)$$

Suponhamos que $\langle u, v \rangle \neq 0$. Seja $z = \frac{\langle u, v \rangle}{|\langle u, v \rangle|}$.

Então, como

$$\Re \langle zv, u \rangle = \Re (z \langle v, u \rangle) = \Re \left(\frac{\langle u, v \rangle}{|\langle u, v \rangle|} \langle v, u \rangle \right) = \Re \left(\frac{|\langle u, v \rangle|^2}{|\langle u, v \rangle|} \right) = |\langle u, v \rangle|,$$

temos, por (I), que

$$|\langle u, v \rangle| = \Re \langle zv, u \rangle \leq \|zv\| \|u\| = \|u\| \|v\|,$$

pois $|z| = 1$.

Se u e v são LD, é fácil verificar que $|\langle u, v \rangle| = \|u\| \|v\|$.

Se $|\langle u, v \rangle| = \|u\| \|v\|$, então $\Re \langle zv, u \rangle = \|zv\| \|u\|$. Logo, a equação

$$\langle u + \lambda zv, u + \lambda zv \rangle = \lambda \|v\|^2 + 2\lambda \Re \langle zv, u \rangle + \|u\|^2 = 0,$$

com $\lambda \in \mathbb{R}$, tem discriminante zero. Então, existe $\lambda_0 \in \mathbb{R}$ tal que

$$\langle u + \lambda_0 zv, u + \lambda_0 zv \rangle = 0,$$

ou seja, $u + \lambda_0 zv = \mathbf{0}$. Portanto, u e v são LD.

Vamos, agora, provar a desigualdade triangular.

Como

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle = \|u\|^2 + \langle u, v \rangle + \langle v, u \rangle + \|v\|^2 \\ &= \|u\|^2 + 2\Re \langle u, v \rangle + \|v\|^2, \end{aligned}$$

temos, pela desigualdade de Cauchy-Schwarz, que

$$\begin{aligned} \|u + v\|^2 &= \|u\|^2 + \|v\|^2 + 2\Re \langle u, v \rangle \leq \|u\|^2 + \|v\|^2 + 2\|u\| \|v\| \\ &= (\|u\| + \|v\|)^2, \end{aligned}$$

ou seja, $\|u + v\| \leq \|u\| + \|v\|$, pois

$$\Re \langle u, v \rangle \leq |\Re \langle u, v \rangle| \leq |\langle u, v \rangle| \leq \|u\| \|v\|.$$

Além disso, temos que $\|u + v\| = \|u\| + \|v\|$ se, e somente se,

$$\Re \langle u, v \rangle = |\Re \langle u, v \rangle| = |\langle u, v \rangle| = \|u\| \|v\|.$$

Logo, existe $c \in \mathbb{C}$ tal que $u = cv$.

Como $|\langle u, v \rangle| = |\Re \langle u, v \rangle|$, temos que $\langle u, v \rangle \in \mathbb{R}^+$. Sendo,

$$\langle u, v \rangle = \langle cv, v \rangle = c |\langle u, v \rangle| = c \|v\|^2,$$

temos que $c > 0$, se $\langle u, v \rangle \neq 0$, ou seja, u é um múltiplo positivo de v .

Se $\langle u, v \rangle = 0$, temos que $u = \mathbf{0}$ ou $v = \mathbf{0}$. Nesse caso, temos, também, que um dos vetores u ou v é múltiplo positivo do outro.

Reciprocamente, se $u = cv$, onde $c \in \mathbb{R}$, $c > 0$, então,

$$\begin{aligned} \|u + v\| &= \|cv + v\| = |c + 1| \|v\| = (c + 1) \|v\| = c \|v\| + \|v\| \\ &= \|cv\| + \|v\| = \|u\| + \|v\|, \end{aligned}$$



Observação 1.5

Aplicando a desigualdade de Cauchy-Schwarz aos produtos internos dos exemplos 1.1, 1.2 e 1.4, obtemos, respectivamente, que

- $\left| \sum_{k=1}^n x_k \overline{y_k} \right| \leq \left(\sum_{k=1}^n |x_k|^2 \right)^{1/2} \left(\sum_{k=1}^n |y_k|^2 \right)^{1/2};$
- $|\text{Tr}(AB^*)| \leq (\text{Tr}(AA^*))^{1/2} (\text{Tr}(BB^*))^{1/2};$
- $\left| \int_0^1 f(x) \overline{g(x)} dx \right| \leq \left(\int_0^1 |f(x)|^2 dx \right)^{1/2} \left(\int_0^1 |g(x)|^2 dx \right)^{1/2}.$

Definição 1.3

Sejam u e v vetores num espaço vetorial V com produto interno. Dizemos que u e v são *ortogonais* se $\langle u, v \rangle = 0$.

Definição 1.4

seja S um subconjunto de V . Dizemos que S é um *conjunto ortogonal* se dois vetores distintos quaisquer em S são ortogonais. Se, além disso, todos os vetores em S são unitários, dizemos que S é um *conjunto ortonormal*.

Observação 1.6

O vetor nulo é ortogonal a todo vetor em V e é o único vetor com essa propriedade.

Exemplo 1.6

A base canônica de $\mathbb{R}^n$ ou de $\mathbb{C}^n$ é um conjunto ortonormal em relação ao produto interno canônico. $\square$

Exemplo 1.7

Seja $V = \mathbb{C}^{n \times n}$ o espaço vetorial de todas as matrizes complexas e seja E^{pq} a matriz cuja única entrada não-nula é 1 localizada na linha p e na coluna q .

Então, o conjunto das matrizes E^{pq} , $p, q = 1, \dots, n$, é ortonormal em relação ao produto interno definido no exemplo 1.2. De fato,

$$\langle E^{pq}, E^{rs} \rangle = \text{Tr}(E^{pq}E^{sr}) = \delta_{qs} \text{Tr}(E^{pr}) = \delta_{qs}\delta_{pr}.$$

$\square$

Exemplo 1.8

Seja $V = \mathcal{C}^0([0, 1], \mathbb{C})$ o espaço vetorial das funções contínuas definidas no intervalo $[0, 1]$ que tomam valores em $\mathbb{C}$, com o produto interno definido no exemplo 1.4.

Sejam $f_n(x) = \sqrt{2} \cos(2\pi nx)$ e $g_n = \sqrt{2} \sin(2\pi nx)$, $n \in \mathbb{N}$.

Então, $\{1, f_1, g_1, f_2, g_2, \dots, f_n, g_n, \dots\}$ é um subconjunto de V que é ortonormal infinito.

Temos, também, que o conjunto de todas as funções da forma

$$h_n(x) = e^{2\pi i n x}, \quad n \in \mathbb{Z},$$

é um subconjunto de V que é ortonormal e infinito. $\square$

Proposição 1.2

Todo conjunto S ortonormal de vetores não nulos é LI.

Prova.

Sejam $v_1, \dots, v_m$ vetores distintos em S e seja

$$w = \lambda_1 v_1 + \dots + \lambda_m v_m.$$

Então,

$$\langle w, v_k \rangle = \sum_{j=1}^m \lambda_j \langle v_j, v_k \rangle = \lambda_k \langle v_k, v_k \rangle.$$

Como $\langle v_k, v_k \rangle \neq 0$, temos que $\lambda_k = \frac{\langle w, v_k \rangle}{\langle v_k, v_k \rangle}$, $k = 1, \dots, m$.

Assim, $\lambda_1 = \dots = \lambda_m = 0$, se $w = \mathbf{0}$. Logo, S é um conjunto LI. ■

Observação 1.7

Se $w = \lambda_1 v_1 + \dots + \lambda_m v_m$ é uma combinação linear de vetores não-nulos ortogonais, então

$$w = \sum_{k=1}^m \frac{\langle w, v_k \rangle}{\|v_k\|^2} v_k.$$

Observação 1.8

Se $\{v_1, \dots, v_m\}$ é um conjunto ortogonal de vetores não-nulos em V , então $m \leq \dim(V)$.

Teorema 1.1

Seja V um espaço vetorial com produto interno e sejam $v_1, \dots, v_n$ vetores LI em V . Então, existem vetores ortogonais $w_1, \dots, w_n$ em V tais que, para cada $k = 1, \dots, n$, o conjunto $\{w_1, \dots, w_k\}$ é uma base do subespaço gerado pelos vetores $v_1, \dots, v_k$.

Processo de ortogonalização de Gram-Schmidt

Prova.

Tome $w_1 = v_1$. Suponhamos que $w_1, \dots, w_m$, $1 \leq m < n$, tenham sido escolhidos de modo que, para cada k ,

$$\{w_1, \dots, w_k\}, 1 \leq k \leq m,$$

seja uma base ortonormal do subespaço gerado por $v_1, \dots, v_k$.

Tome

$$w_{m+1} = v_{m+1} - \sum_{k=1}^m \frac{\langle v_{m+1}, w_k \rangle}{\|w_k\|^2} w_k.$$

Então, $w_{m+1} \neq \mathbf{0}$, pois, caso contrário, v_{m+1} seria combinação linear de $w_1, \dots, w_m$ e, portanto, uma combinação linear de $v_1, \dots, v_m$.

Além disso, se $1 \leq j \leq m$,

$$\begin{aligned} \langle w_{m+1}, w_j \rangle &= \langle v_{m+1}, w_j \rangle - \sum_{k=1}^m \frac{\langle v_{m+1}, w_k \rangle}{\|w_k\|^2} \langle w_k, w_j \rangle \\ &= \langle w_{m+1}, w_j \rangle - \langle w_{m+1}, w_j \rangle = 0. \end{aligned}$$

Portanto, $\{w_1, \dots, w_{m+1}\}$ é um conjunto ortogonal de $m + 1$ vetores não-nulos do subespaço gerado pelos vetores $v_1, \dots, v_{m+1}$. Como, pela proposição anterior, $\{w_1, \dots, w_{m+1}\}$ é um conjunto LI, temos que $\{w_1, \dots, w_{m+1}\}$ é uma base do subespaço gerado pelos vetores $v_1, \dots, v_{m+1}$.

Prosseguindo desta maneira, podemos obter n vetores $w_1, \dots, w_n$ ortogonais tais que $\{w_1, \dots, w_n\}$ é uma base do subespaço gerado pelos vetores $\{v_1, \dots, v_n\}$. ■

Corolário 1.1

Todo espaço vetorial V de dimensão finita com produto interno possui uma base ortonormal.

Prova.

Seja $\{v_1, \dots, v_n\}$ uma base de V . Pelo processo de ortogonalização de Gram-Schmidt, podemos construir uma base ortogonal $\{w_1, \dots, w_n\}$.

Então $\{w'_1, \dots, w'_n\}$ é uma base ortonormal de V , onde

$$w'_1 = \frac{w_1}{\|w_1\|} = \frac{v_1}{\|v_1\|};$$

$$w'_2 = \frac{w_2}{\|w_2\|} = \frac{v_2 - \langle v_2, w'_1 \rangle w'_1}{\|v_2 - \langle v_2, w'_1 \rangle w'_1\|};$$

$$w'_3 = \frac{w_3}{\|w_3\|} = \frac{v_3 - \langle v_3, w'_1 \rangle w'_1 - \langle v_3, w'_2 \rangle w'_2}{\|v_3 - \langle v_3, w'_1 \rangle w'_1 - \langle v_3, w'_2 \rangle w'_2\|};$$

...

$$w'_j = \frac{w_j}{\|w_j\|} = \frac{v_j - \sum_{i=1}^{j-1} \langle v_j, w'_i \rangle w'_i}{\left\| v_j - \sum_{i=1}^{j-1} \langle v_j, w'_i \rangle w'_i \right\|};$$

...

$$w'_n = \frac{w_n}{\|w_n\|} = \frac{v_n - \sum_{i=1}^{n-1} \langle v_n, w'_i \rangle w'_i}{\left\| v_n - \sum_{i=1}^{n-1} \langle v_n, w'_i \rangle w'_i \right\|}. \quad \blacksquare$$

Observação 1.9

Seja V um espaço vetorial com produto interno e seja G a matriz do produto interno em relação a uma base ortonormal $\mathcal{B} = \{v_1, \dots, v_n\}$. Então, G é a matriz identidade e

$$\langle u, w \rangle = Y^* X = \sum_{j=1}^n x_j \overline{y_j},$$

onde $X = [u]_{\mathcal{B}}$ e $Y = [w]_{\mathcal{B}}$.

Observação 1.10

O processo de Gram-Schmidt também pode ser usado para testar dependência linear. De fato, suponhamos que $v_1, \dots, v_n$ sejam vetores LD em um espaço vetorial com produto interno e que $v_1 \neq \mathbf{0}$. Seja m o maior inteiro para o qual os vetores $v_1, \dots, v_m$ são LI. Então $1 \leq m < n$. Sejam $w_1, \dots, w_m$ os vetores obtidos aplicando o processo de ortogonalização aos vetores $v_1, \dots, v_m$. Então, o vetor

$$w_{m+1} = v_{m+1} - \sum_{k=1}^m \frac{\langle v_{m+1}, w_k \rangle}{\|w_k\|^2} w_k,$$

é nulo, pois como w_{m+1} está no subespaço gerado por $w_1, \dots, w_m$, já que v_{m+1} pertence ao subespaço gerado por $v_1, \dots, v_m$ que é igual ao subespaço gerado por $w_1, \dots, w_m$ e é ortogonal a cada um desses vetores, temos

$$w_{m+1} = \sum_{k=1}^m \frac{\langle w_{m+1}, w_k \rangle}{\|w_k\|^2} w_k = \mathbf{0}.$$

Reciprocamente, se $w_1, \dots, w_m$ são não-nulos e $w_{m+1} = \mathbf{0}$, então $v_1, \dots, v_n$ são LI, mas $v_1, \dots, v_{m+1}$ são LD.

Exemplo 1.9

Considere os vetores $v_1 = (3, 0, 4)$, $v_2 = (-1, 0, 7)$ e $v_3 = (2, 9, 11)$ em $\mathbb{R}^3$ munido do produto interno canônico. Aplicando o processo de Gram-Schmidt aos vetores v_1 , v_2 e v_3 , obtemos os seguintes vetores:

- $w_1 = (3, 0, 4)$;
- $w_2 = (-1, 0, 7) - \frac{\langle (-1, 0, 7), (3, 0, 4) \rangle}{25} (3, 0, 4)$
 $= (-1, 0, 7) - (3, 0, 4) = (-4, 0, 3)$;

$$\begin{aligned} \bullet w_3 &= (2, 9, 11) - \frac{\langle (2, 9, 11), (3, 0, 4) \rangle}{25} (3, 0, 4) - \frac{\langle (2, 9, 11), (-4, 0, 3) \rangle}{25} (-4, 0, 3) \\ &= (2, 9, 11) - 2(3, 0, 4) - (-4, 0, 3) = (0, 9, 0). \end{aligned}$$

Como w_1 , w_2 e w_3 são não-nulos e ortogonais, $\{w_1, w_2, w_3\}$ é uma base ortogonal de $\mathbb{R}^3$ e, portanto,

$$\left\{ w'_1 = \frac{w_1}{\|w_1\|} = \frac{(3, 0, 4)}{5}, w'_2 = \frac{w_2}{\|w_2\|} = \frac{(-4, 0, 3)}{5}, w'_3 = \frac{w_3}{\|w_3\|} = (0, 1, 0) \right\}$$

é uma base ortonormal de $\mathbb{R}^3$.

Seja $X = (x_1, x_2, x_3) \in \mathbb{R}^3$. Então,

$$\begin{aligned} X = (x_1, x_2, x_3) &= \langle X, w'_1 \rangle w'_1 + \langle X, w'_2 \rangle w'_2 + \langle X, w'_3 \rangle w'_3 \\ &= \frac{3x_1 + 4x_3}{5} w'_1 + \frac{-4x_1 + 3x_3}{5} w'_2 + x_2 w'_3. \end{aligned}$$

E a base $\{f_1, f_2, f_3\}$ de $(\mathbb{R}^3)^*$ dual da base $\{w'_1, w'_2, w'_3\}$ é dada por:

$$\begin{aligned} f_1(x_1, x_2, x_3) &= \frac{3x_1 + 4x_3}{5} \\ f_2(x_1, x_2, x_3) &= \frac{-4x_1 + 3x_3}{5} \\ f_3(x_1, x_2, x_3) &= x_2. \end{aligned}$$

□

Definição 1.5

Seja W um subespaço de um espaço vetorial V com produto interno e seja $v \in V$. Uma **melhor aproximação** de v por vetores de W é um vetor $w \in W$ tal que

$$\|v - w\| \leq \|v - u\|, \forall u \in W.$$

Teorema 1.2

Seja W um subespaço de um espaço vetorial V com produto interno e seja $v \in V$.

- (a) O vetor $w \in W$ é uma melhor aproximação de v por vetores de W se, e somente se, $v - w$ é ortogonal a todo vetor de W .
- (b) Se uma melhor aproximação de v por vetores de W existir, ela é única.
- (c) Se W é de dimensão finita e $\{w_1, \dots, w_n\}$ é uma base ortonormal de W , então

$$w = \sum_{k=1}^n \langle v, w_k \rangle w_k,$$

e a (única) melhor aproximação de v por vetores de W .

Prova.

Se u é um vetor arbitrário de V , então $v - u = (v - w) + (w - u)$ e

$$\|v - u\|^2 = \|v - w\|^2 + 2\Re\langle v - w, w - u \rangle + \|w - u\|^2 \quad (I)$$

Suponhamos que $v - w$ seja ortogonal a todo vetor em W , $u \in W$ e $u \neq w$.

Como $w - u \in W$, temos que

$$\|v - u\|^2 = \|v - w\|^2 + \|w - u\|^2 > \|v - w\|^2.$$

Reciprocamente, suponhamos que $\|v - u\| \geq \|v - w\|$ para todo $u \in W$.

Então, por (I), temos que

$$2\Re\langle v - w, w - u \rangle + \|w - u\|^2 \geq 0,$$

para todo vetor $u \in W$.

Como todo vetor em W pode ser expresso na forma $w - u$, com $u \in W$, temos que

$$2\Re\langle v - w, x \rangle + \|x\|^2 \geq 0 \quad (II)$$

para todo vetor $x \in W$. Em particular, se $u \in W$ e $u \neq w$, podemos tomar

$$x = -\frac{\langle w - u, v - w \rangle}{\|w - u\|^2} (w - u).$$

Então, a desigualdade (II) se reduz a

$$-2 \frac{|\langle v - w, w - u \rangle|^2}{\|w - u\|^2} + \frac{|\langle v - w, w - u \rangle|^2}{\|w - u\|^2} \geq 0,$$

ou seja $-\langle v - w, w - u \rangle^2 \geq 0$. Logo, $\langle v - w, w - u \rangle = 0$.

Como todo vetor em W pode ser escrito na forma $w - u$, $u \in W$, temos que $v - w$ é ortogonal a todo vetor em W .

Seja w' uma outra melhor aproximação de v por vetores de W . Então,

$$\langle v - w, w - w' \rangle = 0, \quad \text{e} \quad \langle v - w', w - w' \rangle = 0,$$

ou seja,

$$\langle v, w - w' \rangle = \langle w, w - w' \rangle, \quad \text{e} \quad \langle v, w - w' \rangle = \langle w', w - w' \rangle.$$

Logo, $\langle w - w', w - w' \rangle = 0$, isto é, $w' = w$.

Suponhamos agora que W tem dimensão finita e que $\{w_1, \dots, w_n\}$ é uma

base ortonormal de W . Seja $w = \sum_{k=1}^n \langle v, w_k \rangle w_k$. Então,

$$\begin{aligned} \langle v - w, w_j \rangle &= \langle v, w_j \rangle - \langle w, w_j \rangle \\ &= \langle v, w_j \rangle - \sum_{k=1}^n \langle v, w_k \rangle \langle w_k, w_j \rangle \\ &= \langle v, w_j \rangle - \langle v, w_j \rangle = 0, \end{aligned}$$

para todo $j = 1, \dots, n$.

Assim, $v - w$ é ortogonal a toda combinação linear dos vetores $w_1, \dots, w_n$, isto é, a todo vetor de W . Então, por (a), w é a melhor aproximação de v por vetores de W . ■

Definição 1.6

Seja V um espaço vetorial com produto interno e seja S um subconjunto de V . O **suplementar ortogonal de S** é o conjunto $S^\perp$ de todos os vetores em V que são ortogonais a todo vetor em S , ou seja,

$$S^\perp = \{v \in V \mid \langle v, w \rangle = 0 \quad \forall w \in S\}.$$

Observação 1.11

$S^\perp$ é um subespaço vetorial. De fato, $S^\perp$ é não-vazio pois $\mathbf{0} \in S^\perp$ e se $v_1, v_2 \in S^\perp$ e $c \in \mathbb{K}$, então

$$\langle cv_1 + v_2, w \rangle = c\langle v_1, w \rangle + \langle v_2, w \rangle = 0,$$

para todo $w \in S$. Logo, $cv_1 + v_2 \in S^\perp$.

Exemplo 1.10

Tem-se que $\{\mathbf{0}\}^\perp = V$ e $V^\perp = \{\mathbf{0}\}$. □

Observação 1.12

- Se W é um subespaço de V e $w \in W$ é a melhor aproximação de $v \in V$ por vetores de W , então w é o único vetor de W tal que $v - w \in W^\perp$.

Nesse caso, w é chamado a **projeção ortogonal de v sobre W** .

- Se todo vetor de V possuir uma projeção ortogonal sobre W , a aplicação que a cada vetor de V associa a sua projeção ortogonal sobre W é chamada a **projeção ortogonal de V sobre W** .

- Pelo teorema anterior, sempre existe a projeção ortogonal de um espaço com produto interno sobre um subespaço de dimensão finita.

Corolário 1.2

Sejam V um espaço vetorial com produto interno, W um subespaço de dimensão finita e π a projeção ortogonal de V sobre W . Então a aplicação

$$v \mapsto v - \pi(v)$$

é a projeção ortogonal de V sobre $W^\perp$.

Prova.

Seja $v \in V$. Então, $v - \pi(v) \in W^\perp$ e $v - u = \pi(v) + (v - \pi(v) - u)$ para todo $u \in W^\perp$.

Como $\pi(v) \in W$ e $v - \pi(v) - u \in W^\perp$, temos que

$$\|v - u\|^2 = \|\pi(v)\|^2 + \|v - \pi(v) - u\|^2 \geq \|v - (v - \pi(v))\|^2,$$

valendo a desigualdade estrita se $u \neq v - \pi(v)$.

Portanto, $v - \pi(v)$ é a melhor aproximação de v por vetores em $W^\perp$. ■

Teorema 1.3

Seja W um subespaço vetorial de dimensão finita de um espaço V com produto interno. Seja π a projeção ortogonal de V sobre W . Então π é uma transformação linear idempotente de V sobre W . Além disso, $W^\perp = \text{Ker}(\pi)$ e $V = W \oplus W^\perp$.

Prova.

Seja $v \in V$. Como $\pi(v)$ é a melhor aproximação de v por vetores de W , temos que $\pi(v) = v$ se $v \in W$.

Portanto, $\pi(\pi(v)) = \pi(v)$ para todo $v \in V$, isto é, π é idempotente ($\pi^2 = \pi$).

Vamos, agora, provar que π é uma transformação linear. Sejam $u, v \in V$ e $c \in \mathbb{K}$. Como $u - \pi(u)$ e $v - \pi(v)$ são ortogonais a todos os vetores de W , temos que $u - \pi(u)$ e $v - \pi(v)$ pertencem a $W^\perp$. Portanto, o vetor

$$c(u - \pi(u)) + (v - \pi(v)) = (cu + v) - (c\pi(u) + \pi(v))$$

também pertence a $W^\perp$. Como $c\pi(u) + \pi(v) \in W$, temos que $c\pi(u) + \pi(v)$ é a melhor aproximação de $cu + v$ por vetores de W , ou seja,

$$\pi(cu + v) = c\pi(u) + \pi(v).$$

Seja $v \in V$. Como $\pi(v)$ é o único vetor em W tal que $v - \pi(v) \in W^\perp$, temos que $\pi(v) = \mathbf{0}$ se $v \in W^\perp$.

Reciprocamente, se $\pi(v) = \mathbf{0}$ então $v \in W^\perp$. Logo, $W^\perp$ é o núcleo de π .

Como $v = \pi(v) + (v - \pi(v))$ para todo vetor $v \in V$ e $\pi(v) \in W$, $v - \pi(v) \in W^\perp$, temos que $V = W + W^\perp$.

Além disso, se $v \in W \cap W^\perp$, então $\langle v, v \rangle = 0$, ou seja, $v = 0$. Logo, $W \cap W^\perp = \{\mathbf{0}\}$ e, portanto, $V = W \oplus W^\perp$. ■

Observação 1.13

A projeção ortogonal $I - \pi$ de V sobre $W^\perp$, é uma transformação linear idempotente de V sobre $W^\perp$ com núcleo W .

De fato, como π é linear e $\pi^2 = \pi$, temos que $I - \pi$ é linear e

$$(I - \pi)(I - \pi) = (I - 2\pi) + \pi^2 = I - 2\pi + \pi = I - \pi.$$

Além disso, $(I - \pi)(v) = \mathbf{0}$ se, e somente se, $\pi(v) = v$, ou seja, $(I - \pi)(v) = \mathbf{0}$ se, e só se, $v \in W$. Logo, W é o núcleo de $I - \pi$.

Exemplo 1.11

Seja $\mathbb{R}^3$ munido do produto interno canônico e seja π a projeção ortogonal de $\mathbb{R}^3$ sobre o subespaço W gerado pelo vetor $w = (3, 12, -1)$.

Então, se $x = (x_1, x_2, x_3) \in \mathbb{R}^3$,

$$\pi(x) = \frac{\langle x, w \rangle}{\|w\|^2} w = \frac{3x_1 + 12x_2 - x_3}{154} (3, 12, -1).$$

Portanto, $W^\perp = \text{Ker}(\pi) = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid 3x_1 + 12x_2 - x_3 = 0\}$ tem dimensão dois e

$$(I - \pi)(x) = \frac{1}{154} (145x_1 - 36x_2 + 3x_3, -36x_1 + 10x_2 + 12x_3, 3x_1 + 12x_2 + 153x_3)$$

é a projeção ortogonal de V sobre $W^\perp$. □

Observação 1.14

O processo de Gram-Schmidt pode ser, agora, descrito geometricamente.

Seja V um espaço vetorial com produto interno e $v_1, \dots, v_n$ vetores LI em V . Seja P_k , $k > 1$ a projeção ortogonal de V sobre o suplementar ortogonal do subespaço gerado por $v_1, \dots, v_{k-1}$, e seja $P_1 = I$.

Então, os vetores que se obtém aplicando o processo de ortogonalização aos vetores $v_1, \dots, v_n$ são definidos por:

$$w_k = P_k(v_k), \quad 1 \leq k \leq n.$$

Corolário 1.3

Seja $\{v_1, \dots, v_n\}$ um conjunto ortogonal de vetores não-nulos em um espaço vetorial V com produto interno.

Desigualdade de Bessel

Se $v \in V$, então

$$\sum_{k=1}^n \frac{|\langle v, v_k \rangle|^2}{\|v_k\|^2} \leq \|v\|^2$$

e a igualdade ocorre se, e somente se,

$$v = \sum_{k=1}^n \frac{\langle v, v_k \rangle}{\|v_k\|^2} v_k.$$

Prova.

Seja $w = \sum_{k=1}^n \frac{\langle v, v_k \rangle}{\|v_k\|^2} v_k$. Então, $v = w + u$, onde $\langle w, u \rangle = 0$.

Observe que $w = \pi(v)$, onde π é a projeção ortogonal de V sobre o subespaço W gerado pelos vetores $v_1, \dots, v_n$. Logo,

$$\|v\|^2 = \|w\|^2 + \|u\|^2,$$

onde $\|w\|^2 = \sum_{k=1}^n \frac{|\langle v, v_k \rangle|^2}{\|v_k\|^2}.$

Então, $\|v\|^2 \geq \|w\|^2 = \sum_{k=1}^n \frac{|\langle v, v_k \rangle|^2}{\|v_k\|^2}$ e a igualdade ocorre se, e só se, $u = \mathbf{0}$,

ou seja, se, e somente se, $v = \sum_{k=1}^n \frac{\langle v, v_k \rangle}{\|v_k\|^2} v_k$.

Logo, v pertence ao subespaço W gerado pelos vetores $v_1, \dots, v_k$ se, e somente se,

$$v = \sum_{k=1}^n \frac{\langle v, v_k \rangle}{\|v_k\|^2} v_k,$$

ou seja, se, e só se, a desigualdade de Bessel é, na verdade, uma igualdade. ■

Observação 1.15

No caso especial em que $\{v_1, \dots, v_n\}$ é um conjunto ortonormal, a desigualdade de Bessel afirma que

$$\sum_{k=1}^n |\langle v, v_k \rangle|^2 \leq \|v\|^2.$$

Exemplo 1.12

Aplicando a desigualdade de Bessel aos conjuntos ortonormais do exemplo 1.8, temos que:

- $\sum_{k=-n}^n \left| \int_0^1 f(t) e^{-2\pi n i t} dt \right|^2 \leq \int_0^1 |f(t)|^2 dt;$
- $\int_0^1 \left| \sum_{k=-n}^n c_k e^{2\pi n i t} \right|^2 dt = \sum_{k=-n}^n |c_k|^2;$
- $\int_0^1 \left(\sqrt{2} \cos(2\pi t) + \sqrt{2} \sin(4\pi t) \right)^2 dt = 1 + 1 = 2. \quad \square$

2. Funcionais Lineares e Adjuntos

Seja V um espaço vetorial com produto interno e seja $w \in V$. Definimos a aplicação

$$\begin{aligned} f_w &\longrightarrow \mathbb{K} \\ v &\longmapsto f_w(v) = \langle v, w \rangle. \end{aligned}$$

Sendo o produto interno $\langle v, w \rangle$ linear como função de v , f_w é um funcional linear sobre V .

Teorema 2.1

Seja V um espaço vetorial de dimensão finita com produto interno e seja f um funcional linear sobre V . Então existe um único vetor $w \in V$ tal que $f(v) = \langle v, w \rangle$, para todo $v \in V$.

Prova.

Seja $\{v_1, \dots, v_n\}$ uma base ortonormal de V e seja

$$w = \sum_{k=1}^n \overline{f(v_k)} v_k.$$

Seja $f_w : V \rightarrow \mathbb{K}$ o funcional linear definido por $f_w(v) = \langle v, w \rangle$.

Como

$$\begin{aligned} f_w(v_j) &= \langle v_j, w \rangle = \left\langle v_j, \sum_{k=1}^n \overline{f(v_k)} v_k \right\rangle \\ &= \sum_{k=1}^n f(v_k) \langle v_j, v_k \rangle = f(v_j), \end{aligned}$$

para todo $j = 1, \dots, n$, temos que $f = f_w$.

Suponhamos que existe outro vetor $w' \in V$ tal que $f(v) = \langle v, w' \rangle$ para todo $v \in V$. Então,

$$\langle v, w \rangle = \langle v, w' \rangle,$$

ou seja, $\langle v, w - w' \rangle = 0$ para todo $v \in V$. Fazendo $v = w - w'$, obtemos que $\langle w - w', w - w' \rangle = 0$.

Logo, $w' = w$. ■

Observação 2.1

O vetor w pertence a $(\text{Ker}(f))^\perp$. De fato, seja $W = \text{Ker}(f)$.

Então, $V = W \oplus W^\perp$. Se π é a projeção ortogonal de V sobre $W^\perp$, então

$$f(v) = f(\pi(v)),$$

para todo $v \in V$.

Suponhamos $f \neq 0$. Então, $\text{posto}(f) = 1$ e $\dim W^\perp = 1$.

Se w_0 é um vetor não-nulo de $W^\perp$, temos que

$$\pi(v) = \frac{\langle v, w_0 \rangle}{\|w_0\|^2} w_0$$

para todo $v \in V$.

Logo,

$$f(v) = \frac{\langle v, w_0 \rangle}{\|w_0\|^2} f(w_0) = \left\langle v, \frac{\overline{f(w_0)}}{\|w_0\|^2} w_0 \right\rangle = 0$$

para todo $v \in V$. Assim, $w = \frac{\overline{f(w_0)}}{\|w_0\|^2} w_0$.

Exemplo 2.1

Mostraremos, agora, através de um exemplo, que o teorema anterior não é válido sem a hipótese de V ser de dimensão finita.

Seja V o espaço vetorial dos polinômios com coeficientes complexos com o produto interno

$$\langle f, g \rangle = \int_0^1 f(t) \overline{g(t)} dt.$$

Este produto interno também pode ser definido algebricamente.

Se $f = \sum a_k x^k$ e $g = \sum b_j x^j$, então

$$\langle f, g \rangle = \sum_{j,k} \frac{1}{j+k+1} a_k \overline{b_j}.$$

Seja $z \in \mathbb{C}$ e seja L o funcional linear definido por $L(f) = f(z)$.

Suponhamos que existe um polinômio $g \in V$ tal que

$$L(f) = f(z) = \langle f, g \rangle = \int_0^1 f(t) \overline{g(t)} dt,$$

para todo $f \in V$.

Seja $h = x - z$. Então $(hf)(z) = 0$ para todo $f \in V$. Assim,

$$0 = \int_0^1 h(t) f(t) \overline{g(t)} dt,$$

para todo $f \in V$. Tome $f = \overline{h}g$. Então,

$$0 = \int_0^1 |h(t)|^2 |g(t)|^2 dt.$$

Logo, $hg = 0$. Como $h \neq 0$, devemos ter $g = 0$, o que é uma contradição, já que L não é o funcional nulo. $\square$

Teorema 2.2

Seja T um operador linear sobre um espaço V de dimensão finita com produto interno. Então, existe um único operador linear T^* sobre V tal que

$$\langle T(v), w \rangle = \langle v, T^*(w) \rangle,$$

para todo $v, w \in V$.

Prova.

Seja $w \in V$. Então, a aplicação $v \mapsto \langle T(v), w \rangle$ é um funcional linear

sobre V . Pelo teorema anterior, existe um único vetor $w' \in V$ tal que

$$\langle T(v), w \rangle = \langle v, w' \rangle,$$

para todo $v \in V$. Faça $T^*(w) = w'$.

Devemos verificar que $T^* : V \rightarrow V$ é uma transformação linear.

Sejam $w, u \in V$ e $c \in \mathbb{K}$. Então, para qualquer $v \in V$, temos

$$\begin{aligned} \langle v, T^*(cw + u) \rangle &= \langle T(v), cw + u \rangle \\ &= \bar{c} \langle T(v), w \rangle + \langle T(v), u \rangle \\ &= \bar{c} \langle v, T^*(w) \rangle + \langle v, T^*(u) \rangle \\ &= \langle v, cT^*(w) + T^*(u) \rangle. \end{aligned}$$

Logo, $T^*(cw + u) = cT^*(w) + T^*(u)$, ou seja, T^* é linear. ■

Proposição 2.1

Seja V um espaço vetorial de dimensão finita com produto interno e seja $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base ortonormal de V . Seja T um operador linear sobre V e seja A a matriz de T em relação à base ordenada $\mathcal{B}$. Então, $A_{ij} = \langle T(v_j), v_i \rangle$.

Prova.

Como $\mathcal{B}$ é uma base ortonormal, temos que

$$T(v_j) = \sum_{i=1}^n \langle T(v_j), v_i \rangle v_i.$$

Logo, $A_{ij} = \langle T(v_j), v_i \rangle$. ■

Teorema 2.3

Seja V um espaço vetorial de dimensão finita com produto interno e seja T um operador linear sobre V . Então, em relação a qualquer base ortonormal de V , a matriz de T^* é a transposta conjugada da matriz de T .

Prova.

Sejam $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base ortonormal de V , $A = [T]_{\mathcal{B}}$ e $B = [T^*]_{\mathcal{B}}$.

Pelo teorema anterior, $A_{kj} = \langle T(v_j), v_k \rangle$ e $B_{kj} = \langle T^*(v_j), v_k \rangle$.

Logo,

$$B_{kj} = \langle T^*(v_j), v_k \rangle = \overline{\langle v_k, T^*(v_j) \rangle} = \overline{\langle T(v_k), v_j \rangle} = \overline{A_{jk}}.$$

■

Exemplo 2.2

Seja V um espaço vetorial de dimensão finita com produto interno e seja π a projeção ortogonal de V sobre um subespaço W . Então, para quaisquer vetores $v, w \in V$,

$$\begin{aligned}\langle \pi(v), w \rangle &= \langle \pi(v), \pi(w) + (w - \pi(w)) \rangle \\ &= \langle \pi(v), \pi(w) \rangle \\ &= \langle \pi(v) + (v - \pi(v)), \pi(w) \rangle \\ &= \langle v, \pi(w) \rangle.\end{aligned}$$

Logo, $\pi^* = \pi$. $\square$

Exemplo 2.3

Seja $\pi : \mathbb{R}^3 \longrightarrow W$ dada por

$$\pi(x_1, x_2, x_3) = \frac{3x_1 + 12x_2 - x_3}{154} (3, 12, -1),$$

a projeção ortogonal de $\mathbb{R}^3$ sobre o subespaço W gerado pelo vetor $(3, 12, -1)$, do exemplo 1.11.

Então,

$$A = \frac{1}{154} \begin{pmatrix} 9 & 36 & -3 \\ 36 & 144 & -12 \\ -3 & -12 & 1 \end{pmatrix}$$

é a matriz de π em relação à base ortonormal canônica. Como $\pi^* = \pi$, A é também a matriz de π^* em relação à base canônica. E, como $A^* = A$, isso não contradiz o teorema acima.

Considere, agora a base $\mathcal{B}' = \{v_1, v_2, v_3\}$, onde

$$v_1 = (154, 0, 0), \quad v_2 = (145, -36, 3), \quad v_3 = (-36, 10, 12).$$

Como

$$\pi(v_1) = (9, 36, -3) = (154, 0, 0) - (145, -36, 3), \quad \pi(v_2) = (0, 0, 0), \quad \pi(v_3) = (0, 0, 0),$$

temos que

$$[\pi]_{\mathcal{B}'} = B = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Nesse caso, $B \neq B^*$ e B^* não é a matriz de $\pi^* = \pi$ em relação à base $\mathcal{B}'$. Isso ocorre porque a base $\mathcal{B}'$ não é ortonormal. $\square$

Definição 2.1

Seja T um operador linear sobre um espaço V com produto interno. Dizemos que T possui um *adjunto* sobre V se existe um operador linear T^* sobre V tal que $\langle T(v), w \rangle = \langle v, T^*(w) \rangle$ para todos $v, w \in V$.

Observação 2.2

O Teorema 2.2 afirma que todo operador linear sobre um espaço V de dimensão finita com produto interno possui um adjunto. Mas, no caso de dimensão infinita isso nem sempre ocorre. Em qualquer caso, existe no máximo um operador T^* , que é denominado *adjunto de T* .

Observação 2.3

O adjunto de T depende não só de T , como também do produto interno de V .

Exemplo 2.4

Seja $V = \mathbb{C}^{n \times 1}$ o espaço das matrizes complexas $n \times 1$ com o produto interno $\langle X, Y \rangle = Y^*X$. Se A é uma matriz complexa $n \times n$, o adjunto do operador linear $X \mapsto AX$ é o operador $X \mapsto A^*X$. De fato,

$$\langle AX, Y \rangle = Y^*AX = (A^*Y)^*X = \langle X, A^*Y \rangle.$$

□

Exemplo 2.5

Seja $V = \mathbb{C}^{n \times n}$ o espaço das matrizes complexas $n \times n$ com o produto interno $\langle A, B \rangle = \text{Tr}(B^*A)$. Seja $M \in \mathbb{C}^{n \times n}$. O adjunto do operador $L_M(A) = MA$ é o operador $L_{M^*}(A) = M^*A$.

De fato,

$$\begin{aligned} \langle L_M(A), B \rangle &= \text{Tr}(B^*(MA)) \\ &= \text{Tr}(MAB^*) \\ &= \text{Tr}(AB^*M) \\ &= \text{Tr}(A(M^*B)^*) \\ &= \text{Tr}((M^*B)^*A) \\ &= \langle A, M^*B \rangle. \end{aligned}$$

Logo, $(L_M)^*(B) = M^*B = L_{M^*}(B)$, ou seja, $(L_M)^* = L_{M^*}$. □

Exemplo 2.6

Seja V o espaço vetorial dos polinômios sobre $\mathbb{C}$, com o produto interno

$$\langle f, g \rangle = \int_0^1 f(t) \overline{g(t)} dt.$$

Se $f = \sum a_k x^k$, seja $\bar{f} = \sum \overline{a_k} x^k$. Então,

$$\bar{f}(t) = \overline{f(t)}, \quad \text{para todo } t \in \mathbb{R}.$$

Consideremos o operador $M_f(g) = fg$, isto é, M_f é o operador multiplicação por f . Então, o adjunto de M_f é o operador $M_{\bar{f}}$.

De fato,

$$\begin{aligned} \langle M_f(g), h \rangle &= \int_0^1 f(t) g(t) \overline{h(t)} dt \\ &= \int_0^1 g(t) \overline{\overline{f(t)} h(t)} dt \\ &= \langle g, M_{\bar{f}}(h) \rangle. \end{aligned}$$

Logo, $(M_f)^* = M_{\bar{f}}$. $\square$

Exemplo 2.7

Daremos, agora, um exemplo para mostrar que nem todo operador sobre um espaço vetorial de dimensão infinita com produto interno possui um adjunto.

Seja V o espaço dos polinômios sobre $\mathbb{C}$ com o produto interno

$$\langle f, g \rangle = \int_0^1 f(t) \overline{g(t)} dt.$$

Seja D o operador derivação sobre V . Então,

$$\begin{aligned} \langle D(f), g \rangle &= \int_0^1 f'(t) \overline{g(t)} dt = f(1) \overline{g(1)} - f(0) \overline{g(0)} - \int_0^1 f(t) \overline{g(t)}' dt \\ &= f(1) \overline{g(1)} - f(0) \overline{g(0)} - \langle f, D(g) \rangle. \end{aligned}$$

Suponhamos que existe um polinômio $D^*(g)$ tal que

$$\langle D(f), g \rangle = \langle f, D^*(g) \rangle, \quad \text{para todo } f \in V.$$

Ou seja,

$$\langle f, D^*(g) \rangle = f(1) \overline{g(1)} - f(0) \overline{g(0)} - \langle f, D(g) \rangle,$$

ou ainda,

$$\langle f, D^*(g) + D(g) \rangle = f(1)\bar{g}(1) - f(0)\bar{g}(0),$$

para todo $f \in V$.

Seja $L(f) = f(1)\bar{g}(1) - f(0)\bar{g}(0)$, onde $g \in \mathbb{C}[x]$ é fixo. Então, L é um funcional linear e

$$L(f) = \langle f, h \rangle, \quad \text{para todo } f \in \mathbb{C}[x],$$

onde $h = D^*(g) + D(g)$.

Tome $\ell = (x - 1)x$. Então,

$$0 = L(\ell f) = \ell f(1)\bar{g}(1) - \ell f(0)\bar{g}(0) = \int_0^1 \ell(t)f(t)\bar{h}(t) dt,$$

para todo $f \in \mathbb{C}[x]$.

Fazendo $f = \bar{\ell}h$, temos que

$$\int_0^1 |\ell(t)|^2 |h(t)|^2 dt = 0,$$

o que implica que $h(t) = 0 \forall t \in \mathbb{R}$.

Assim,

$$L(f) = \langle f, h \rangle = f(1)\bar{g}(1) - f(0)\bar{g}(0) = 0,$$

para todo $f \in \mathbb{C}[x]$.

Tomando $f = x$ e $f = x - 1$, obtemos que

$$g(1) = g(0) = 0.$$

Logo, se $g \in \mathbb{C}[x]$ é tal que $g(1) \neq 0$ ou $g(0) \neq 0$, então não existe $D^*(g)$ tal que $\langle D(f), g \rangle = \langle f, D^*(g) \rangle$, para todo f . $\square$

- Veremos agora como a operação que leva T em T^* se assemelha com a conjugação sobre os números complexos.

Teorema 2.4

Seja V um espaço vetorial de dimensão finita com produto interno. Se T e U são operadores lineares sobre V e c é um escalar, então:

- (a) $(T + U)^* = T^* + U^*$;
- (b) $(cT)^* = \bar{c}T^*$;
- (c) $(TU)^* = U^*T^*$;

(d) $(T^*)^* = T$.

Prova.

Como, para todos $v, w \in V$ temos

$$\begin{aligned}\langle (T + U)(v), w \rangle &= \langle T(v) + U(v), w \rangle \\ &= \langle T(v), w \rangle + \langle U(v), w \rangle \\ &= \langle v, T^*(w) \rangle + \langle v, U^*(w) \rangle \\ &= \langle v, T^*(w) + U^*(w) \rangle \\ &= \langle v, (T^* + U^*)(w) \rangle,\end{aligned}$$

temos, pela unicidade do adjunto, que $(T + U)^* = T^* + U^*$.

De modo análogo, para todos $v, w \in W$, temos

- $\langle (cT)(v), w \rangle = \langle cT(v), w \rangle = c\langle T(v), w \rangle = c\langle v, T^*(w) \rangle = \langle v, (\bar{c}T^*)(w) \rangle$;
- $\langle T U(v), w \rangle = \langle U(v), T^*(w) \rangle = \langle v, U^*T^*(w) \rangle$;
- $\langle T^*(v), w \rangle = \overline{\langle w, T^*(v) \rangle} = \overline{\langle T(w), v \rangle} = \langle v, T(w) \rangle$.

Logo, pela unicidade do adjunto, temos que

$$(cT)^* = \bar{c}T^*, \quad (TU)^* = U^*T^*, \quad \text{e} \quad (T^*)^* = T.$$



Observação 2.4

Se T é um operador linear sobre um espaço vetorial complexo de dimensão finita com produto interno, então T pode ser expresso de modo único na forma:

$$T = U_1 + iU_2,$$

onde $U_1^* = U_1$ e $U_2^* = U_2$.

De fato, se $T = U_1 + iU_2$, onde $U_1^* = U_1$ e $U_2^* = U_2$, então

$$T^* = (U_1 + iU_2)^* = U_1^* - iU_2^* = U_1 - iU_2.$$

Logo,

$$U_1 = \frac{T + T^*}{2} \quad \text{e} \quad U_2 = \frac{T - T^*}{2i}.$$

Assim, T possui, de certa forma, uma parte real, U_1 , e uma parte imaginária, U_2 .

(Lembre-se: $z \in \mathbb{C}$ é real se, e somente se, $z = \bar{z}$)

Definição 2.2

Um operador linear T tal que $T^* = T$ é dito *auto-adjunto* ou *hermitiano*.

Então, T é auto-adjunto se, e somente se,

$$[T]_{\mathcal{B}} = [T^*]_{\mathcal{B}} = [T]_{\mathcal{B}}^*,$$

para toda base ortonormal $\mathcal{B}$ de V , ou seja, se, e somente se, $[T]_{\mathcal{B}}$ é uma matriz auto-adjunta (ou hermitiana) para toda base ortonormal $\mathcal{B}$ de V .

Observação 2.5

Seja a transformação

$$\begin{aligned} i_V : V &\longrightarrow V^* \\ y &\longmapsto i_V(y)(x) = \langle x, y \rangle. \end{aligned}$$

Então:

(a) $i_V(y_1 + y_2) = i_V(y_1) + i_V(y_2)$, pois

$$\begin{aligned} i_V(y_1 + y_2)(x) &= \langle x, y_1 + y_2 \rangle \\ &= \langle x, y_1 \rangle + \langle x, y_2 \rangle \\ &= i_V(y_1)(x) + i_V(y_2)(x), \end{aligned}$$

para todo $x \in V$.

(b) $i_V(\lambda y) = \bar{\lambda} i_V(y)$, pois

$$i_V(\lambda y)(x) = \langle x, \lambda y \rangle = \bar{\lambda} \langle x, y \rangle = \bar{\lambda} i_V(y)(x),$$

para todo $x \in V$.

(c) i_V é injetora, pois

$$\begin{aligned} i_V(y_1) = i_V(y_2) &\iff i_V(y_1)(x) = i_V(y_2)(x), \quad \forall x \in V \\ &\iff \langle x, y_1 \rangle = \langle x, y_2 \rangle, \quad \forall x \in V \\ &\iff \langle x, y_1 - y_2 \rangle = 0, \quad \forall x \in V. \end{aligned}$$

Fazendo $x = y_1 - y_2$, obtemos que $\langle y_1 - y_2, y_1 - y_2 \rangle = 0$, logo $y_1 - y_2 = \mathbf{0}$.

Isto é, $y_1 = y_2$.

(d) Se V tem dimensão finita, i_V é sobrejetora, pois se $f \in V^*$ já provamos que existe um (único) vetor $y \in V$ tal que $f(x) = \langle x, y \rangle$, para todo $x \in V$. Logo, $f = i_V(y)$.

Seja $T : V \longrightarrow V$ um operador linear sobre um espaço vetorial V de dimensão finita com produto interno e seja

$$\begin{aligned} T_1^* : V^* &\longrightarrow V^* \\ f &\longmapsto T_1^*(f) = f \circ T, \end{aligned}$$

o adjunto de T anteriormente definido. Então, o seguinte diagrama é comutativo

$$\begin{array}{ccc} V^* & \xleftarrow{T_1^*} & V^* \\ i_V \uparrow & & \uparrow i_V \\ V & \xleftarrow{T^*} & V \end{array}$$

ou seja, $i_V \circ T^* = T_1^* \circ i_V$.

De fato,

$$\begin{aligned} T_1^* \circ i_V(y)(x) &= (i_V(y) \circ T)(x) = i_V(y)(T(x)) \\ &= \langle T(x), y \rangle = \langle x, T^*(y) \rangle = i_V(T^*(y))(x) \\ &= (i_V \circ T^*)(y)(x), \end{aligned}$$

para todos $x, y \in V$. Logo, $T^* \circ i_V = i_V \circ T^*$.

3. Operadores Unitários

Definição 3.1

Sejam V e W espaços vetoriais com produto interno sobre o mesmo corpo $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) e seja T uma transformação linear de V em W . Dizemos que T **preserva produto interno** se

$$\langle T(v), T(w) \rangle = \langle v, w \rangle,$$

para todos $v, w \in V$.

Uma **isometria** de V em W é um isomorfismo de V em W que preserva o produto interno. Uma isometria de um espaço vetorial com produto interno sobre si mesmo é um **operador unitário**.

Observação 3.1

Se $T \in \mathcal{L}(V, W)$ preserva produto interno, então $\|T(v)\| = \|v\|$ para todo $v \in V$. Isto é, T **preserva a norma**.

Em particular T é uma transformação injetora.

Reciprocamente, se $T \in \mathcal{L}(V, W)$ preserva a norma, ou seja $\|T(v)\| = \|v\|$

para todo $v \in V$, então T preserva produto interno.

De fato, para $\mathbb{K} = \mathbb{R}$, usando a identidade de polarização, temos que

$$\begin{aligned}\langle T(v), T(w) \rangle &= \frac{1}{4} (\|T(v) + T(w)\|^2 - \|T(v) - T(w)\|^2) \\ &= \frac{1}{4} (\|T(v + w)\|^2 - \|T(v - w)\|^2) \\ &= \frac{1}{4} (\|v + w\|^2 - \|v - w\|^2) \\ &= \langle v, w \rangle,\end{aligned}$$

e para $\mathbb{K} = \mathbb{C}$,

$$\begin{aligned}\langle T(v), T(w) \rangle &= \frac{1}{4} (\|T(v) + T(w)\|^2 - \|T(v) - T(w)\|^2) \\ &\quad + \frac{i}{4} (\|T(v) + iT(w)\|^2 - \|T(v) - iT(w)\|^2) \\ &= \frac{1}{4} (\|T(v + w)\|^2 - \|T(v - w)\|^2) \\ &\quad + \frac{i}{4} (\|T(v + iw)\|^2 - \|T(v - iw)\|^2) \\ &= \frac{1}{4} (\|v + w\|^2 - \|v - w\|^2) + \frac{i}{4} (\|v + iw\|^2 - \|v - iw\|^2) \\ &= \langle v, w \rangle,\end{aligned}$$

Proposição 3.1

Sejam V e W espaços vetoriais com produto interno de mesma dimensão finita n sobre o corpo $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$). Se $T \in \mathcal{L}(V, W)$, as seguintes afirmações são equivalentes:

- (a)** T preserva produto interno.
- (b)** T é um isomorfismo que preserva produto interno (ou seja T é uma isometria).
- (c)** T leva toda base ortonormal de V em uma base ortonormal de W .
- (d)** T leva alguma base ortonormal de V em uma base ortonormal de W .

Prova.

(a) $\implies$ (b) Como T é injetora e $\dim(V) = \dim(W) = n$, temos que T é sobrejetora. Logo, T é um isomorfismo.

(b) $\implies$ (c) Seja $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base ortonormal de V . Como T é um

isomorfismo, temos que $\mathcal{B}' = \{T(v_1), \dots, T(v_n)\}$ é uma base de W tal que

$$\langle T(v_i), T(v_j) \rangle = \langle v_i, v_j \rangle = \delta_{ij},$$

ou seja, $\mathcal{B}'$ é uma base ortonormal de W .

(c) $\implies$ (d) É de fácil verificação.

(d) $\implies$ (a) Seja $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base ortonormal de V tal que $\mathcal{B}' = \{T(v_1), \dots, T(v_n)\}$ é uma base ortonormal de W .

Então, se $v = \sum_{j=1}^n x_j v_j$ e $w = \sum_{k=1}^n y_k v_k$, temos que

$$\begin{aligned} \langle T(v), T(w) \rangle &= \left\langle \sum_{j=1}^n x_j T(v_j), \sum_{k=1}^n y_k T(v_k) \right\rangle = \sum_{j,k=1}^n x_j \overline{y_k} \langle T(v_j), T(v_k) \rangle \\ &= \sum_{j,k=1}^n x_j \overline{y_k} \delta_{jk} = \sum_{j,k=1}^n x_j \overline{y_k} \langle v_j, v_k \rangle \\ &= \left\langle \sum_{j=1}^n x_j v_j, \sum_{k=1}^n y_k v_k \right\rangle = \langle v, w \rangle, \end{aligned}$$

ou seja, T preserva produto interno. ■

Corolário 3.1

Sejam V e W espaços vetoriais de dimensão finita com produto interno sobre o mesmo corpo $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$). Existe um isomorfismo $T : V \rightarrow W$ que preserva produto interno se, e somente se, $\dim(V) = \dim(W)$.

Prova.

($\implies$) trivial.

($\impliedby$) Se $\{v_1, \dots, v_n\}$ é uma base ortonormal de V e $\{w_1, \dots, w_n\}$ é uma base ortonormal de W , seja T a transformação linear de V em W definida por $T(v_i) = w_i$, $i = 1, \dots, n$. Então, T é um isomorfismo de V em W que preserva produto interno. ■

Exemplo 3.1

Se V é um espaço n -dimensional com produto interno sobre o corpo $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$), então toda base ortonormal $\mathcal{B} = \{v_1, \dots, v_n\}$ determina uma isometria de V em $\mathbb{K}^n$ com o produto interno canônico. De fato, a transformação

$$T(v) = T\left(\sum_{j=1}^n x_j v_j\right) = (x_1, \dots, x_n),$$

onde $v = \sum_{j=1}^n x_j v_j$, define uma isometria de V em $\mathbb{K}^n$, pois $\{T(v_1) = e_1, \dots, T(v_n) = e_n\}$ é a base canônica, que é ortonormal em relação ao produto interno canônico.

Se consideramos o espaço $\mathbb{K}^{n \times 1}$ com o produto interno $\langle X, Y \rangle = Y^* X$, temos que

$$v \longmapsto [v]_{\mathcal{B}}$$

é uma isometria de V em $\mathbb{K}^{n \times 1}$ se, e somente se, $\mathcal{B}$ é uma base ortonormal de V . $\square$

Exemplo 3.2

Seja $P \in \mathbb{C}^{n \times n}$ uma matriz complexa invertível e seja $G = P^* P$. Sejam $\langle X, Y \rangle = Y^* X$ e $[X, Y] = Y^* G X$ produtos internos em $\mathbb{C}^{n \times 1}$. Então,

$$\begin{aligned} T : (\mathbb{C}^{n \times 1}, [\ , \]) &\longrightarrow (\mathbb{C}^{n \times 1}, \langle \ , \ \rangle) \\ X &\longmapsto PX \end{aligned}$$

é uma isometria. De fato,

$$\begin{aligned} \langle T(X), T(Y) \rangle &= \langle PX, PY \rangle = (PY)^* PX = Y^* P^* P X \\ &= Y^* G X = [X, Y]. \end{aligned}$$

$\square$

Definição 3.2

Um **operador unitário** sobre um espaço vetorial V com produto interno é um isomorfismo de V em V que preserva o produto interno.

Observação 3.2

O conjunto dos operadores unitários sobre um espaço V com produto interno é um **grupo** com a operação de composição.

De fato, se U_1 e U_2 são unitários, então $U_2 U_1$ é um isomorfismo e

$$\|U_2 U_1(v)\| = \|U_1(v)\| = \|v\|, \quad \text{para todo } v \in V.$$

Além disso, o inverso U^{-1} de um operador unitário U é unitário, pois

$$\|U^{-1}(v)\| = \|U(U^{-1}(v))\| = \|v\|, \quad \text{para todo } v \in V.$$

Temos, também, que o operador identidade é unitário.

Observação 3.3

Se V é um espaço de dimensão finita com produto interno e U é um operador linear sobre V , então, pela proposição 3.1, temos que U é unitário se, e somente se, U preserva produto interno, ou ainda, se, e só se, U leva alguma (toda) base ortonormal de V em outra base ortonormal de V .

Proposição 3.2

Seja U um operador linear sobre um espaço vetorial V com produto interno. Então, U é unitário se, e só se, o adjunto U^* de U existe e $UU^* = U^*U = I$.

Prova.

Suponhamos que U é unitário. Então U é invertível e

$$\langle U(v), w \rangle = \langle Uv, UU^{-1}(w) \rangle = \langle v, U^{-1}(w) \rangle,$$

para todos $v, w \in V$. Logo, $U^* = U^{-1}$ é o adjunto de U .

Suponhamos, agora, que U^* existe e $UU^* = U^*U = I$. Então U é invertível, $U^{-1} = U^*$ e

$$\langle U(v), U(w) \rangle = \langle v, U^*U(w) \rangle = \langle v, w \rangle,$$

para todos $v, w \in W$. Logo, U é um isomorfismo que preserva produto interno, ou seja, U é unitário. ■

Observação 3.4

Seja V um espaço vetorial com produto interno e seja T um operador linear sobre V . Então, T é unitário se, e só se, T^* é unitário.

Exemplo 3.3

Seja $V = \mathbb{C}^{n \times 1}$ com o produto interno $\langle X, Y \rangle = Y^*X$, e seja $A \in \mathbb{C}^{n \times n}$. Seja o operador $U : \mathbb{C}^{n \times 1} \rightarrow \mathbb{C}^{n \times 1}$ o operador definido por $U(X) = AX$.

Como

$$\langle U(X), U(Y) \rangle = \langle AX, AY \rangle = (AY)^*AX = Y^*A^*AX,$$

para todo X , temos que U é unitário se, e somente se, $A^*A = I$. □

Definição 3.3

Uma matriz $n \times n$ real ou complexa A é dita **unitária** se $A^*A = I$.

Proposição 3.3

Seja V um espaço vetorial de dimensão finita com produto interno e seja U um operador linear sobre V . Então U é unitário se, e só se, a matriz de U em relação a alguma (toda) base ortonormal ordenada de V é uma matriz unitária.

Prova.

Suponhamos que U é unitário. Seja $\mathcal{B}$ uma base ortonormal de V . Como $U^*U = I$, temos que

$$I = [I]_{\mathcal{B}} = [U^*U]_{\mathcal{B}} = [U^*]_{\mathcal{B}}[U]_{\mathcal{B}} = [U]_{\mathcal{B}}^*[U]_{\mathcal{B}},$$

ou seja $[U]_{\mathcal{B}}$ é uma matriz unitária.

Suponhamos, agora, que $[U]_{\mathcal{B}}$ é uma matriz unitária para alguma base ortonormal $\mathcal{B}$. Então,

$$I = [U]_{\mathcal{B}}^*[U]_{\mathcal{B}} = [U^*]_{\mathcal{B}}[U]_{\mathcal{B}} = [U^*U]_{\mathcal{B}}.$$

Logo, $U^*U = I$, ou seja, U é um operador unitário. ■

Observação 3.5

Seja A uma matriz $n \times n$ complexa. Então,

$$\begin{aligned} A \text{ é unitária} &\iff A^*A = I \\ &\iff (A^*A)_{jk} = \delta_{jk}, \quad j, k = 1, \dots, n \\ &\iff \sum_{r=1}^n \overline{A_{rj}} A_{rk} = \delta_{jk}, \quad j, k = 1, \dots, n \end{aligned}$$

$\iff$ as colunas da matriz A formam um conjunto ortonormal de vetores de $\mathbb{C}^{n \times 1}$ com respeito ao produto interno canônico $\langle X, Y \rangle = Y^*X$.

Como $A^*A = I \iff AA^* = I$, temos que A é unitária $\iff$ as linhas de A formam um conjunto ortonormal de vetores de $\mathbb{C}^n$ com respeito ao produto interno canônico.

Então, A é unitária $\iff$ as colunas de A formam uma base ortonormal de $\mathbb{C}^{n \times 1}$ com respeito ao produto interno canônico $\iff$ as linhas de A formam uma base ortonormal de $\mathbb{C}^n$ com respeito ao produto interno canônico.

Definição 3.4

- Uma matriz $n \times n$ A , real ou complexa, é dita **ortogonal** se $A^t A = I$.
- Uma matriz unitária é ortogonal se, e somente se, todas as suas entradas são reais.

De fato, $A^* = A^t \iff \overline{a_{ij}} = a_{ij}, i, j = 1, \dots, n$.

Exemplo 3.4

Uma matriz 1×1 $[c]$ é ortogonal se, e somente se, $c = \pm 1$, e é unitária se, e somente se, $\overline{c}c = 1$, ou seja, se, e somente se, $|c| = 1$, ou ainda, se, e só se, $c = e^{i\theta}$, para algum $\theta \in \mathbb{R}$. $\square$

Exemplo 3.5

Seja a matriz 2×2 sobre $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$)

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Então A é ortogonal se, e só se, $A^{-1} = A^t$, ou seja, se, e só se,

$$A^t = \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = A^{-1}.$$

Como $\det(A^t) = \det(A)$ e $A^t A = I$, temos que $(\det(A))^2 = 1$. Logo, $\det A = \pm 1$.

Assim, A é ortogonal se, e só se,

$$\bullet \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \text{ e } ad - bc = 1,$$

ou seja, $d = a$, $c = -b$ e $ad - bc = 1$;

ou

$$\bullet \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} -d & b \\ c & -a \end{pmatrix}, \text{ e } ad - bc = -1,$$

ou seja, $d = -a$, $c = b$ e $ad - bc = -1$.

Então, A é ortogonal se, e só se,

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \quad \text{ou} \quad A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix},$$

onde $a^2 + b^2 = 1$.

- Em particular, se $\theta \in \mathbb{R}$, a matriz

$$A_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix},$$

é uma matriz ortogonal real, logo unitária.

Como A_θ é a matriz em relação à base canônica de $\mathbb{R}^2$ do operador linear U_θ , rotação de ângulo θ em torno da origem, temos que U_θ é um operador unitário, ou seja, preserva o produto interno. $\square$

Exemplo 3.6

Seja a matriz 2×2 sobre $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$)

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Então, A é unitária se, e só se,

$$A^* = \begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = A^{-1}.$$

Como $\det(A^*) = \overline{\det(A)}$ e $A^*A = I$, temos que $\det(A) = e^{i\theta}$, para algum $\theta \in \mathbb{R}$.

Assim, A é unitária se, e só se,

$$A^* = \begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{pmatrix} = e^{-i\theta} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = A^{-1}, \quad \text{e} \quad ad - bc = e^{i\theta},$$

ou seja, $d = e^{i\theta}\bar{a}$, $c = -e^{i\theta}\bar{b}$ e $ad - bc = e^{i\theta}$ para algum $\theta \in \mathbb{R}$.

Logo, A é unitária se, e só se,

$$A = \begin{pmatrix} a & b \\ -e^{i\theta}\bar{b} & e^{i\theta}\bar{a} \end{pmatrix},$$

onde $\theta \in \mathbb{R}$ e $|a|^2 + |b|^2 = 1$. $\square$

Observação 3.6

O conjunto $U(n)$ das matrizes unitárias $n \times n$ é um grupo.

De fato, se $A \in U(n)$, então $(A^{-1})^{-1} = A = (A^*)^{-1} = (A^{-1})^*$. E, se A e B são unitárias, então

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

Observação 3.7

Sejam V um espaço vetorial de dimensão finita com produto interno, $\mathcal{B} = \{v_1, \dots, v_n\}$ e $\mathcal{B}' = \{v'_1, \dots, v'_n\}$ bases ortonormais de V . Seja $P = [I]_{\mathcal{B}\mathcal{B}'}$ a matriz de mudança da base $\mathcal{B}'$ para a base $\mathcal{B}$, onde

$$v'_k = \sum_{j=1}^n P_{jk} v_j.$$

Seja $U : V \longrightarrow V$ o operador linear definido por $U(v_j) = v'_j$, $j = 1, \dots, n$. Como U leva a base ortonormal $\mathcal{B}$ na base ortonormal $\mathcal{B}'$ e $[U]_{\mathcal{B}} = P$, temos que U é um operador unitário e P é uma matriz unitária.

Logo, se $T : V \longrightarrow V$ é um operador linear, temos que

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

onde P é uma matriz unitária e $\mathcal{B}, \mathcal{B}'$ são bases ortonormais.

Definição 3.5

Sejam A e B matrizes complexas $n \times n$. Dizemos que B é *unitariamente equivalente* a A se existe uma matriz unitária P $n \times n$ tal que

$$B = P^{-1}AP = P^*AP.$$

E dizemos que B é *ortogonalmente equivalente* a A se existe uma matriz ortogonal P $n \times n$, tal que

$$B = P^{-1}AP = P^tAP.$$

Observação 3.8

Seja $T : V \longrightarrow V$ um operador linear sobre um espaço vetorial V de dimensão finita com produto interno, e sejam $\mathcal{B}$ e $\mathcal{B}'$ bases ortonormais de V .

Então, $[T]_{\mathcal{B}'}$ e $[T]_{\mathcal{B}}$ são unitariamente equivalentes, e se V é real, $[T]_{\mathcal{B}'}$ e $[T]_{\mathcal{B}}$ são ortogonalmente equivalentes, através de uma matriz ortogonal real.

4. Operadores Normais

Seja $T : V \longrightarrow V$ um operador linear sobre um espaço V de dimensão finita e seja $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base ortonormal de V , sendo $T(v_i) = c_i v_i$, $i = 1, \dots, n$.

Então,

$$[T]_{\mathcal{B}} = \begin{pmatrix} c_1 & & \\ & \ddots & \\ & & c_n \end{pmatrix} \quad \text{e} \quad [T^*]_{\mathcal{B}} = ([T]_{\mathcal{B}})^* = \begin{pmatrix} \overline{c_1} & & \\ & \ddots & \\ & & \overline{c_n} \end{pmatrix}.$$

Se V é *real*, $c_i = \overline{c_i}$, $i = 1, \dots, n$ e $[T^*]_{\mathcal{B}} = [T]_{\mathcal{B}}$, ou seja, $T = T^*$. Então, T é auto-adjunto.

Se V é *complexo*, podemos apenas afirmar que $[T]_{\mathcal{B}}[T^*]_{\mathcal{B}} = [T^*]_{\mathcal{B}}[T]_{\mathcal{B}}$, pois $[T]_{\mathcal{B}}$ e $[T^*]_{\mathcal{B}}$ são matrizes diagonais. Logo, $TT^* = T^*T$.

Definição 4.1

Seja V um espaço vetorial de dimensão finita com produto interno e T um operador linear sobre V . Dizemos que T é *normal* se $TT^* = T^*T$, ou seja, se T e T^* comutam.

Observação 4.1

- Todo operador auto-adjunto é normal.
- Todo operador unitário é normal

Observação 4.2

Todo múltiplo de um operador normal é um operador normal, mas a soma e o produto de operadores normais nem sempre são normais.

Proposição 4.1

Seja V um espaço vetorial com produto interno e T um operador linear auto-adjunto sobre V . Então todo autovalor de T é real.

Prova.

Seja c um autovalor de T e seja v não-nulo tal que $T(v) = cv$. Então,

$$\begin{aligned} c\langle v, v \rangle &= \langle cv, v \rangle = \langle T(v), v \rangle \\ &= \langle v, T^*(v) \rangle = \langle v, T(v) \rangle \\ &= \langle v, cv \rangle = \bar{c} \langle v, v \rangle, \end{aligned}$$

ou seja, $(c - \bar{c})\langle v, v \rangle = 0$. Como $\langle v, v \rangle \neq 0$, temos que $c = \bar{c}$. Logo, c é real. ■

Observação 4.3

Se T é normal, então $\|T^*(v)\| = \|T(v)\|$ para todo $v \in V$.

De fato

$$\langle T^*T(v), v \rangle = \langle T(v), (T^*)^*(v) \rangle = \langle T(v), T(v) \rangle = \|T(v)\|^2,$$

e

$$\langle TT^*(v), v \rangle = \langle T^*(v), T^*(v) \rangle = \|T^*(v)\|^2.$$

Como $TT^* = T^*T$, temos que $\|T(v)\| = \|T^*(v)\|$.

Observação 4.4

Se T é normal, então $T - \lambda I$ é normal para todo $\lambda \in \mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$).

De fato,

$$\begin{aligned} (T - \lambda I)(T - \lambda I)^* &= (T - \lambda I)(T^* - \bar{\lambda} I) = TT^* - \lambda T^* - \bar{\lambda} T + |\lambda|^2 I \\ &= T^*T - \lambda T^* - \bar{\lambda} T + |\lambda|^2 I = (T^* - \bar{\lambda} I)(T - \lambda I) \\ &= (T - \lambda I)^*(T - \lambda I). \end{aligned}$$

Proposição 4.2

Seja V um espaço vetorial com produto interno e seja T um operador normal sobre V . Seja $v \in V$. Então v é um autovetor de T associado a um autovalor c se, e só se, v é um autovalor de T^* associado ao autovalor $\bar{c}$.

Prova.

Pela observação 4.4, $T - cI$ é normal, e pela observação 4.3, temos que

$$0 = \|(T - cI)(v)\| = \|(T - cI)^*(v)\| = \|(T^* - \bar{c}I)(v)\|,$$

ou seja, $T^*(v) = \bar{c}v$. ■

Proposição 4.3

Seja V um espaço vetorial com produto interno e T um operador normal sobre V . Então, autovetores associados a autovalores distintos são ortogonais.

Prova.

Sejam λ e μ autovalores distintos de T , e v, w autovetores de T associados aos autovalores λ e μ , respectivamente. Então,

$$\begin{aligned}\lambda \langle v, w \rangle &= \langle \lambda v, w \rangle = \langle T(v), w \rangle = \langle v, T^*(w) \rangle \\ &= \langle v, \bar{\mu} w \rangle = \mu \langle v, w \rangle.\end{aligned}$$

Logo, $(\lambda - \mu) \langle v, w \rangle = 0$. Como $\lambda \neq \mu$, temos que $\langle v, w \rangle = 0$. ■

Proposição 4.4

Seja V um espaço vetorial com produto interno e T um operador linear sobre V . Seja W um subespaço de V invariante por T . Então, $W^\perp$ é invariante por T^* .

Prova.

Sejam $w \in W^\perp$ e $v \in W$. Então,

$$\langle v, T^*(w) \rangle = \langle T(v), w \rangle = 0,$$

pois $T(v) \in W$ e $w \in W^\perp$.

Como $\langle v, T^*(w) \rangle = 0$ para todo $v \in W$, temos que $T^*(w) \in W^\perp$. ■

Proposição 4.5

Seja V um espaço vetorial sobre $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) e seja T um operador auto-adjunto sobre V . Então, T possui um autovetor não-nulo.

Prova.

Seja $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base ortonormal de V e seja $A = [T]_{\mathcal{B}}$.

Como $p_c = \det(xI - A)$ é um polinômio com coeficientes complexos, p_c possui uma raiz $c \in \mathbb{C}$.

Então, $\det(cI - A) = 0$, ou seja $cI - A$ não é invertível.

Logo, existe $X \in \mathbb{C}^{n \times 1} - \{\mathbf{0}\}$ tal que $AX = cX$.

Se V é um espaço vetorial real, temos que $c \in \mathbb{R}$, pois, como $A^* = A$, o operador $U : \mathbb{C}^{n \times 1} \rightarrow \mathbb{C}^{n \times 1}$ dado por $U(X) = AX$ é auto-adjunto com o produto interno canônico e c é um autovalor de U .

Logo, como $cI - A$ é uma matriz com entradas reais e $\det(cI - A) = 0$, existe $X \in \mathbb{R}^{n \times 1} - \{\mathbf{0}\}$ tal que $AX = cX$.

Seja $v = \sum_{j=1}^n x_j v_j$, onde $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \neq \mathbf{0}$.

Então, $v \neq \mathbf{0}$ e

$$[T(v)]_{\mathcal{B}} = [T]_{\mathcal{B}}[v]_{\mathcal{B}} = AX = cX = c[v]_{\mathcal{B}}.$$

Logo, $T(v) = cv$. ■

Observação 4.5

A hipótese de T ser um operador auto-adjunto só foi usada no caso real para garantir que toda raiz do polinômio característico de T é real. No caso complexo, essa hipótese é desnecessária.

Observação 4.6

O polinômio característico de um operador auto-adjunto sobre um espaço vetorial complexo ou real tem todos os coeficientes reais, pois todas as suas raízes são reais.

Observação 4.7

Um operador auto-adjunto sobre um espaço vetorial de dimensão infinita com produto interno pode não ter autovetores não-nulos.

De fato, seja $V = \mathbb{C}^0([0, 1], \mathbb{C})$ com o produto interno

$$\langle f, g \rangle = \int_0^1 f(t) \overline{g(t)} dt,$$

e seja $T : V \rightarrow V$ o operador linear definido por $T(f)(t) = t f(t)$, para todo $t \in [0, 1]$.

Como

$$\langle T(f), g \rangle = \int_0^1 t f(t) \overline{g(t)} dt = \int_0^1 f(t) \overline{t g(t)} dt = \langle f, T(g) \rangle, \quad \forall f, g \in V,$$

temos que $T^* = T$, ou seja, T é auto-adjunto.

Suponha que existe $c \in \mathbb{C}$ e $f \in V - \{\mathbf{0}\}$ tal que $T(f) = cf$.

Então,

$$t f(t) = c f(t), \quad \forall t \in [0, 1],$$

ou seja, $(t - c) f(t) = 0$ para todo $t \in [0, 1]$.

Logo, $f(t) = 0$ para todo $t \in [0, 1]$, o qual é uma contradição.

Teorema 4.1

Seja V um espaço vetorial de dimensão finita com produto interno sobre $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) e seja T um operador auto-adjunto sobre V .

Então, V possui uma base ortonormal de autovetores de T .

Prova.

Pela proposição 4.5, T possui um autovetor v não-nulo. Tome $v_1 = \frac{v}{\|v\|}$, que também é um autovetor de T de norma 1.

Se $\dim(V) = 1$ o teorema fica provado.

Se $\dim(V) > 1$, vamos proceder por indução sobre a dimensão de V para provar o resultado.

Suponhamos que o teorema seja válido para espaços com produto interno de dimensão menor que $\dim(V)$.

Seja W o subespaço gerado pelo vetor v_1 .

Como v_1 é um autovetor de T , temos que $T(W) \subset W$, ou seja, W é invariante por T . Pela proposição 4.4, $W^\perp$ é invariante por $T^* = T$.

Seja $U = T|_{W^\perp}$. Como

$$\langle U(v), w \rangle = \langle T(v), w \rangle = \langle v, T(w) \rangle = \langle v, U(w) \rangle, \quad \forall v, w \in W^\perp,$$

temos que U é um operador auto-adjunto sobre o espaço $W^\perp$.

Como $\dim(W^\perp) = \dim(V) - 1$ temos, pela hipótese de indução, que $W^\perp$ possui uma base ortonormal $\{v_2, \dots, v_n\}$ de autovetores de U .

Então, $\{v_1, v_2, \dots, v_n\}$ é uma base ortonormal de autovetores de T , já que $v_2, \dots, v_n$ também são autovetores de T e $V = W \oplus W^\perp$. ■

Corolário 4.1

Seja A uma matriz $n \times n$ hermitiana (auto-adjunta). Então A é unitariamente equivalente a uma matriz diagonal, ou seja, existe uma matriz P unitária tal que $P^{-1}AP$ é diagonal.

Se A é uma matriz simétrica real ($A^t = A$), existe uma matriz ortogonal real P tal que $P^{-1}AP$ é diagonal.

Prova.

Seja V o espaço $\mathbb{C}^{n \times 1}$ com o produto interno canônico e seja T o ope-

rador sobre V que é representado por A em relação à base canônica $\mathcal{B}_0$ de $\mathbb{C}^{n \times 1}$, ou seja, $T(X) = AX$ para todo $X \in V$.

Como $A = A^*$, temos que $T = T^*$, ou seja, T é auto-adjunto.

Então, pelo teorema anterior, V possui uma base ortonormal $\mathcal{B} = \{v_1, \dots, v_n\}$ de autovetores de T , ou seja, $[T]_{\mathcal{B}}$ é diagonal.

Seja P a matriz cujos vetores colunas são os vetores $v_1, \dots, v_n$. Então, P é unitária e $P = [I]_{\mathcal{B}_0 \mathcal{B}}$.

Logo, $[T]_{\mathcal{B}} = P^{-1}AP$, ou seja, A é unitariamente equivalente a uma matriz diagonal.

Caso todas as entradas de A sejam reais, tomamos $V = \mathbb{R}^{n \times 1}$ com o produto interno canônico e repetimos o argumento. Neste caso, a matriz P é uma matriz unitária com entradas reais, ou seja, P é uma matriz ortogonal real. ■

Observação 4.8

Seja V um espaço vetorial *real* de dimensão finita com produto interno e seja T um operador linear sobre V .

Pelas observações feitas no início da seção e pelo teorema anterior, temos que:

- Um operador T é auto-adjunto $\iff$ o espaço V possui uma base ortonormal de autovetores de T .
- Equivalentemente, se A é uma matriz $n \times n$ com entradas *reais*, temos que:

A é simétrica $\iff$ existe uma matriz $n \times n$ P ortogonal real tal que $P^tAP = P^{-1}AP$ é uma matriz diagonal.

De fato, a implicação $(\implies)$ já foi provada no corolário acima. Para provar a recíproca, basta observar que se $P^tAP = D$ é uma matriz diagonal, então

$$D = D^t = (P^tAP)^t = P^tA^tP,$$

ou seja, $P^tAP = P^tA^tP$. E como $P^tP = PP^t = I$, temos

$$A = PP^tAPP^t = PP^tA^tPP^t = A^t,$$

ou seja, A é uma matriz simétrica.

Observação 4.9

Seja V um espaço vetorial sobre $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) de dimensão finita com produto interno e seja T um operador auto-adjunto sobre V .

Se $\lambda_1, \dots, \lambda_k$ são os autovalores distintos de T , W_i o autoespaço associado ao autovalor λ_i e $\pi_i : V \rightarrow W_i$ a projeção de V sobre W_i , $i = 1, \dots, k$, então:

- $\lambda_i \in \mathbb{R}$, $i = 1, \dots, k$;
- $p_c = (x - \lambda_1)^{d_1} \dots (x - \lambda_k)^{d_k}$;
- $p_m = (x - \lambda_1) \dots (x - \lambda_k)$;
- $W_i \perp W_j$, $i \neq j$;
- $V = W_1 \oplus \dots \oplus W_k$ (soma direta ortogonal);
- π_i é a projeção ortogonal de V sobre W_i ;
- $\dim(W_i) = d_i$, $i = 1, \dots, k$;
- $I = \pi_1 + \dots + \pi_k$;
- $\pi_i^2 = \pi_i$, $i = 1, \dots, k$;
- $\pi_i \pi_j = \mathcal{O}$, se $i \neq j$;
- $T = \lambda_1 \pi_1 + \dots + \lambda_k \pi_k$;
- $\pi_i = p_i(T)$, onde $p_i = \frac{\prod_{j \neq i} (x - \lambda_j)}{\prod_{j \neq i} (\lambda_i - \lambda_j)}$, $i = 1, \dots, k$.

Seja, agora, um espaço vetorial V de dimensão finita com produto interno e T um operador sobre V .

Já provamos, no início da seção, que se V possui uma base ortonormal de autovetores de T , então T é normal.

O teorema abaixo mostra que a recíproca também é verdadeira para espaços vetoriais **complexos**.

O teorema só é válido para espaços vetoriais complexos, pois um operador normal sobre um espaço real pode não ter autovalores reais. Isto é verdadeiro, por exemplo, para todas as rotações de $\mathbb{R}^2$, salvo duas.

De fato, o operador rotação

$$\begin{aligned} T_\theta : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ X &\longrightarrow AX \end{aligned} \quad \text{onde} \quad A = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

é normal, pois $A^*A = A^tA = AA^t = AA^* = I$, e o seu polinômio característico é

$$p_c = (x - \cos \theta)^2 + \sin^2 \theta.$$

Como $e^{i\theta}$ e $e^{-i\theta}$ são as raízes deste polinômio, que não são reais, se $\theta \neq 0, \pi$, temos que não existe uma base ortonormal de $\mathbb{R}^2$ formada por autovetores de T_θ , se $\theta \neq 0, \pi$.

Teorema 4.2

Seja V um espaço vetorial complexo de dimensão finita com produto interno e seja T um operador normal sobre V . Então, V possui uma base ortonormal de autovetores de T .

Prova.

Sejam $\lambda_1, \dots, \lambda_m$ os autovalores distintos de T , $n = \dim(V)$, e W_{λ_i} o autoespaço associado ao autovalor λ_i , $i = 1, \dots, m$.

Já sabemos que $W_{\lambda_i} \perp W_{\lambda_j}$, se $i \neq j$.

Vamos provar, por indução sobre n , que

$$V = W_{\lambda_1} \oplus \dots \oplus W_{\lambda_m}.$$

Se $n = \dim(V) = 1$, não há nada a provar. Suponhamos que o resultado seja verdadeiro para operadores normais sobre um espaço vetorial complexo de dimensão menor ou igual a $n - 1$.

Seja W_{λ_1} o autoespaço associado ao autovalor λ_1 .

Como $T(W_{\lambda_1}) \subset W_{\lambda_1}$, temos, pela proposição 4.4, que $T^*(W_{\lambda_1}^\perp) \subset W_{\lambda_1}^\perp$.

Além disso, como

$$T(T^*(v)) = T^*(T(v)) = \lambda_1 T^*(v), \quad \forall v \in W_{\lambda_1},$$

temos que $T^*(W_{\lambda_1}) \subset W_{\lambda_1}$.

Logo, pela proposição 4.4, $T(W_{\lambda_1}^\perp) \subset W_{\lambda_1}^\perp$.

Assim, $T|_{W_{\lambda_1}^\perp} : W_{\lambda_1}^\perp \longrightarrow W_{\lambda_1}^\perp$ é um operador normal sobre $W_{\lambda_1}^\perp$, pois $[T|_{W_{\lambda_1}^\perp}]^* = [T^*|_{W_{\lambda_1}^\perp}]$, já que $T^*(W_{\lambda_1}^\perp) \subset W_{\lambda_1}^\perp$ e $\langle T(v), w \rangle = \langle v, T^*(w) \rangle$ para quaisquer $v, w \in W_{\lambda_1}^\perp$.

Como $\dim(W_{\lambda_1}^\perp) \leq n - 1$, temos, pela hipótese de indução, que

$$W_{\lambda_1}^\perp = W'_{\mu_1} \oplus \dots \oplus W'_{\mu_k},$$

onde $\mu_1, \dots, \mu_k$ são os autovalores distintos de $T|_{W_{\lambda_1}^\perp}$ e W'_{μ_i} é o autoespaço correspondente ao autovalor μ_i .

Como todo autovetor de $T|_{W_{\lambda_1}^\perp}$ é um autovetor de T , temos que, para cada $i = 1, \dots, k$, existe $j_i \in \{1, \dots, m\}$, tal que $\mu_i = \lambda_{j_i}$.

Logo,

$$W'_{\mu_i} = \{v \in W_{\lambda_1}^\perp \mid T(v) = \mu_i v\} = \{v \in W_{\lambda_1}^\perp \mid T(v) = \lambda_{j_i} v\} = W_{\lambda_1}^\perp \cap W'_{\lambda_{j_i}}.$$

Observe, também, que $\lambda_{j_i} \neq \lambda_1$, para todo $i = 1, \dots, k$, pois se $\lambda_{j_i} = \lambda_1$, existiria $v \in W_{\lambda_1}^\perp - \{0\}$ tal que $T(v) = \lambda_1 v$, o que é um absurdo, já que $W_{\lambda_1} \cap W_{\lambda_1}^\perp = \{0\}$.

Como $W_{\lambda_{j_i}} \perp W_{\lambda_1}$, para todo $i = 1, \dots, k$, temos que $W_{\lambda_{j_i}} \subset W_{\lambda_1}^\perp$, para todo $i = 1, \dots, k$.

Logo, $W'_{\mu_i} = W_{\lambda_1}^\perp \cap W_{\lambda_{j_i}} = W_{\lambda_{j_i}}$, $i = 1, \dots, k$.

Então,

$$V = W_{\lambda_1} \oplus W_{\lambda_1}^\perp = W_{\lambda_1} \oplus W_{\lambda_{j_1}} \oplus \dots \oplus W_{\lambda_{j_k}}.$$

Basta, agora, mostrar que $\lambda_1, \lambda_{j_1}, \dots, \lambda_{j_k}$ são todos os autovalores de T .

Seja λ um autovalor de T e seja $v \neq 0$ tal que $T(v) = \lambda v$. Então, existem $v_1 \in W_{\lambda_1}$ e $w_i \in W_{\lambda_{j_i}}$, $i = 1, \dots, k$, tais que $v = v_1 + w_1 + \dots + w_k$.

Como

$$\begin{aligned} T(v) &= \lambda v = \lambda(v_1 + w_1 + \dots + w_k) = T(v_1) + T(w_1) + \dots + T(w_k) \\ &= \lambda_1 v_1 + \lambda_{j_1} w_1 + \dots + \lambda_{j_k} w_k, \end{aligned}$$

temos que $(\lambda - \lambda_1)v_1 + (\lambda - \lambda_{j_1})w_1 + \dots + (\lambda - \lambda_{j_k})w_k = 0$.

Logo, $(\lambda - \lambda_1)v_1 = 0$ e $(\lambda - \lambda_{j_i})w_i = 0$, $i = 1, \dots, k$. Sendo $v \neq 0$, devemos ter que, $\lambda = \lambda_1$ ou $\lambda = \lambda_{j_i}$, para algum $i = 1, \dots, k$.

Provamos, assim, que $V = W_{\lambda_1} \oplus \dots \oplus W_{\lambda_m}$, onde $\lambda_1, \dots, \lambda_m$ são os autovalores distintos de T .

Então, se $\mathcal{B}_i$ é uma base ortonormal de W_{λ_i} , $i = 1, \dots, m$, $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_m$ é uma base ortonormal de V formada por autovetores de T . ■

Corolário 4.2

Seja A uma matriz $n \times n$ complexa normal ($AA^* = A^*A$). Então, existe uma matriz $n \times n$ P unitária tal que $P^{-1}AP = P^*AP$ é diagonal.

Teorema 4.3

Seja V um espaço vetorial real de dimensão finita com produto interno e seja T um operador normal sobre V .

Então, se $p_c = p_1^{d_1} p_2^{d_2} \dots p_k^{d_k} (x - \lambda_1)^{f_1} \dots (x - \lambda_\ell)^{f_\ell}$ é o polinômio característico de T , onde $\lambda_1, \dots, \lambda_\ell$ são os autovalores reais distintos de T e $p_i = (x - \mu_j)(x - \overline{\mu_j})$, $\mu_j = \alpha_j e^{i\theta_j}$, $\theta_j \in (0, \pi)$, $j = 1, \dots, k$, temos que:

- (1) $p_m = p_1 \dots p_k (x - \lambda_1) \dots (x - \lambda_\ell)$ é o polinômio minimal de T .
- (2) $V = \text{Ker}(p_1(T)) \oplus \dots \oplus \text{Ker}(p_k(T)) \oplus W_{\lambda_1} \oplus \dots \oplus W_{\lambda_\ell}$ é uma decomposição de V numa soma direta de subespaços invariantes ortogonais, onde:
 - $W_{\lambda_i} = \text{Ker}(T - \lambda_i I)$, $T|_{W_{\lambda_i}} = \lambda_i I$, $\dim(W_{\lambda_i}) = f_i$ e $x - \lambda_i$ é o polinômio minimal de $T|_{W_{\lambda_i}}$, $i = 1, \dots, \ell$.
 - $\dim(\text{Ker}(p_j(T))) = 2d_j$ e p_j é o polinômio minimal de $T|_{\text{Ker}(p_j(T))}$, $j = 1, \dots, k$.
- (3) Existe uma base $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k \cup \mathcal{B}_{k+1} \cup \dots \cup \mathcal{B}_{k+\ell}$ ortonormal de V tal que:
 - $\mathcal{B}_i$ é uma base ortonormal de $\text{Ker}(p_i(T))$ tal que

$$[T|_{\text{Ker}(p_i(T))}]_{\mathcal{B}_i} = A_i = \begin{pmatrix} R_1^i & & & \\ & R_2^i & & \\ & & \ddots & \\ & & & R_{d_i}^i \end{pmatrix}_{2d_i \times 2d_i},$$

onde $R_j^i = \alpha_i \begin{pmatrix} \cos \theta_i & -\sin \theta_i \\ \sin \theta_i & \cos \theta_i \end{pmatrix}$, para todo $j = 1, \dots, d_i$ e $i = 1, \dots, k$.

- $\mathcal{B}_{k+j}$ é uma base ortonormal de W_{λ_j} tal que

$$[T|_{W_{\lambda_j}}]_{\mathcal{B}_{k+j}} = B_j = \begin{pmatrix} \lambda_j & & \\ & \ddots & \\ & & \lambda_j \end{pmatrix}_{f_j \times f_j}, \quad j = 1, \dots, \ell.$$

Ou seja,

$$[T]_B = \begin{pmatrix} A_1 & & & & \\ & \ddots & & & \\ & & A_k & & \\ & & & B_1 & \\ & & & & \ddots \\ & & & & & B_\ell \end{pmatrix}_{n \times n}$$

está na forma canônica de Jordan real.

Prova.

Seja $\widehat{V} = \{(u, v) \mid u, v \in V\}$ o complexificado do espaço V .

Afirmção: $\langle \cdot, \cdot \rangle_{\mathbb{C}} : \widehat{V} \times \widehat{V} \longrightarrow \mathbb{C}$, definido por

$$\langle (u, v), (u', v') \rangle_{\mathbb{C}} = \langle u, u' \rangle + \langle v, v' \rangle - i\langle u, v' \rangle + \langle v, u' \rangle$$

é um produto interno sobre $\widehat{V}$. De fato,

- $\begin{aligned} & \langle (a + ib)(u_1, v_1) + (u_2, v_2), (u', v') \rangle_{\mathbb{C}} \\ &= \langle (au_1 - bv_1 + u_2, bu_1 + av_1 + v_2), (u', v') \rangle_{\mathbb{C}} \\ &= \langle au_1 - bv_1 + u_2, u' \rangle + \langle bu_1 + av_1 + v_2, v' \rangle \\ &\quad - i\langle au_1 - bv_1 + u_2, v' \rangle + i\langle bu_1 + av_1 + v_2, u' \rangle \\ &= a\langle u_1, u' \rangle - b\langle v_1, u' \rangle + \langle u_2, u' \rangle + b\langle u_1, v' \rangle + a\langle v_1, v' \rangle + \langle v_2, v' \rangle \\ &\quad - ia\langle u_1, v' \rangle + ib\langle v_1, v' \rangle - i\langle u_2, v' \rangle + ib\langle u_1, u' \rangle + ia\langle v_1, u' \rangle + i\langle v_2, u' \rangle \\ &= a\langle u_1, u' \rangle + a\langle v_1, v' \rangle + ib\langle v_1, v' \rangle + ib\langle u_1, u' \rangle \\ &\quad - ia\langle u_1, v' \rangle + ia\langle v_1, u' \rangle + b\langle u_1, v' \rangle - b\langle v_1, u' \rangle \\ &\quad + \langle u_2, u' \rangle + \langle v_2, v' \rangle - i\langle u_2, v' \rangle + i\langle v_2, u' \rangle \\ &= (a + ib)\langle u_1, u' \rangle + (a + ib)\langle v_1, v' \rangle - i(a + ib)\langle u_1, v' \rangle + i(a + ib)\langle v_1, u' \rangle \\ &\quad + \langle (u_2, v_2), (u', v') \rangle_{\mathbb{C}} \\ &= (a + ib)(\langle u_1, u' \rangle + \langle v_1, v' \rangle - i\langle u_1, v' \rangle + i\langle v_1, u' \rangle) + \langle (u_2, v_2), (u', v') \rangle_{\mathbb{C}} \\ &= (a + ib)\langle (u_1, v_1), (u', v') \rangle_{\mathbb{C}} + \langle (u_2, v_2), (u', v') \rangle_{\mathbb{C}}. \end{aligned}$
- $\begin{aligned} & \langle (u', v'), (u, v) \rangle_{\mathbb{C}} = \langle u', u \rangle + \langle v', v \rangle - i\langle u', v \rangle + i\langle v', u \rangle \\ &= \langle u, u' \rangle + \langle v, v' \rangle + i\langle u, v' \rangle - i\langle v, u' \rangle = \overline{\langle (u, v), (u', v') \rangle_{\mathbb{C}}}. \end{aligned}$
- $\langle (u, v), (u, v) \rangle_{\mathbb{C}} = \langle u, u \rangle + \langle v, v \rangle - i\langle u, v \rangle + i\langle v, u \rangle = \langle u, u \rangle + \langle v, v \rangle \geq 0.$
- $\langle (u, v), (u, v) \rangle_{\mathbb{C}} = \|u\|^2 + \|v\|^2 = 0 \iff u = v = \mathbf{0} \iff (u, v) = (\mathbf{0}, \mathbf{0}). \blacksquare$

Afirmção: $(\widehat{L})^* = (L^*)^\wedge$.

De fato, como $\langle L(u), v \rangle = \langle u, L(v) \rangle$, para quaisquer $u, v \in V$, temos:

$$\begin{aligned} \langle \widehat{L}(u, v), (u', v') \rangle_{\mathbb{C}} &= \langle (L(u), L(v)), (u', v') \rangle_{\mathbb{C}} \\ &= \langle L(u), u' \rangle + \langle L(v), v' \rangle - i \langle L(u), v' \rangle + i \langle L(v), u' \rangle \\ &= \langle u, L^*(u') \rangle + \langle v, L^*(v') \rangle - i \langle u, L^*(v') \rangle + i \langle v, L^*(u') \rangle \\ &= \langle (u, v), (L^*(u'), L^*(v')) \rangle_{\mathbb{C}} = \langle (u, v), \widehat{L}^*(u', v') \rangle_{\mathbb{C}}. \end{aligned}$$

Logo, $(\widehat{L})^*(u', v') = \widehat{L}^*(u', v')$ para todo $(u', v') \in \widehat{V}$.

Afirmção: $\widehat{L}$ é normal.

De fato, como $LL^* = L^*L$, temos que

$$\widehat{L}(\widehat{L})^* = \widehat{L}\widehat{L}^* = (LL^*)^\wedge = (L^*L)^\wedge = (\widehat{L}^*)\widehat{L} = (\widehat{L})^*\widehat{L}.$$

• Como $\widehat{L} : \widehat{V} \rightarrow \widehat{V}$ é um operador normal sobre o espaço complexo $\widehat{V}$, temos, pelo teorema 4.2, que $\widehat{L}$ é diagonalizável. Então:

$$\rightarrow p_m = (x - \mu_1)(x - \overline{\mu_1}) \dots (x - \mu_k)(x - \overline{\mu_k})(x - \lambda_1) \dots (x - \lambda_\ell)$$

é o polinômio minimal de $\widehat{L}$ (que é igual ao polinômio minimal de L).

$$\rightarrow \sigma(\widetilde{W_{\mu_r}}) = \widetilde{W_{\overline{\mu_r}}}, \quad r = 1, \dots, k.$$

$$\rightarrow \dim(\widetilde{W_{\mu_r}}) = \dim(\widetilde{W_{\overline{\mu_r}}}) = d_r, \quad r = 1, \dots, k, \text{ onde}$$

$$\widetilde{W_{\mu}} = \{(u, v) \in \widehat{V} \mid \widehat{T}(u, v) = \mu(u, v)\}.$$

$$\rightarrow (\text{Ker}(p_r(T)))^\wedge = \widetilde{W_{\mu_r}} \oplus \widetilde{W_{\overline{\mu_r}}}, \quad r = 1, \dots, k.$$

$$\rightarrow \text{Ker}(\widehat{T} - \lambda_j I) = \widehat{W_{\lambda_j}}, \quad \dim(W_{\lambda_j}) = \dim(\widehat{W_{\lambda_j}}) = f_j, \quad j = 1, \dots, \ell.$$

$$\rightarrow \widehat{V} = \widetilde{W_{\mu_1}} \oplus \widetilde{W_{\overline{\mu_1}}} \oplus \dots \oplus \widetilde{W_{\mu_k}} \oplus \widetilde{W_{\overline{\mu_k}}} \oplus \widehat{W_{\lambda_1}} \oplus \dots \oplus \widehat{W_{\lambda_\ell}},$$

soma direta de subespaços invariantes e ortogonais.

• Já vimos, também, que se $\mathcal{B}_r = \{(v_1^r, w_1^r), \dots, (v_{d_r}^r, w_{d_r}^r)\}$ é uma base de $\widehat{W_{\mu_r}}$, então $\sigma(\mathcal{B}_r) = \overline{\mathcal{B}_r} = \{(v_1^r, -w_1^r), \dots, (v_{d_r}^r, -w_{d_r}^r)\}$ é uma base de $\widetilde{W_{\mu_r}}$.

Afirmção: $\mathcal{B}_r$ ortonormal $\implies \overline{\mathcal{B}_r}$ ortonormal.

De fato, como

$$\begin{aligned} \langle (v_k^r, w_k^r), (v_j^r, w_j^r) \rangle_{\mathbb{C}} &= \langle v_k^r, v_j^r \rangle + \langle w_k^r, w_j^r \rangle - i \langle v_k^r, w_j^r \rangle + i \langle w_k^r, v_j^r \rangle \\ &= \begin{cases} 0, & \text{se } k \neq j \\ 1, & \text{se } k = j, \end{cases} \end{aligned}$$

temos que

$$\begin{aligned}\langle (v_k^r, -w_k^r), (v_j^r, -w_j^r) \rangle_{\mathbb{C}} &= \langle v_k^r, v_j^r \rangle + \langle w_k^r, w_j^r \rangle + i\langle v_k^r, w_j^r \rangle - i\langle w_k^r, v_j^r \rangle \\ &= \overline{\langle (v_k^r, w_k^r), (v_j^r, w_j^r) \rangle_{\mathbb{C}}} \\ &= \begin{cases} 0, & \text{se } k \neq j \\ 1, & \text{se } k = j, \end{cases}\end{aligned}$$

• Já sabemos também que

$$\{v_1 + w_1, v_1 - w_1, \dots, v_{d_r} + w_{d_r}, v_{d_r} - w_{d_r}\}$$

é uma base de $\text{Ker}(p_r(T))$ tal que

$$T(v_j^r + w_j^r) = a_r \cos \theta_r (v_j^r + w_j^r) + a_r \sin \theta_r (v_j^r - w_j^r),$$

e

$$T(v_j^r - w_j^r) = -a_r \sin \theta_r (v_j^r + w_j^r) + a_r \cos \theta_r (v_j^r - w_j^r),$$

para $j = 1, \dots, d_r$, onde $\mu_r = a_r \cos \theta_r + ia_r \sin \theta_r$.

Afirmção: $\langle v_j^r, v_j^r \rangle = \langle w_j^r, w_j^r \rangle = \frac{1}{2}$ e $\langle v_j^r, w_j^r \rangle = 0, j = 1, \dots, d_r$.

De fato,

$$1 = \langle (v_j^r, w_j^r), (v_j^r, w_j^r) \rangle_{\mathbb{C}} = \langle v_j^r, v_j^r \rangle + \langle w_j^r, w_j^r \rangle,$$

e

$$0 = \langle (v_j^r, w_j^r), (v_j^r, -w_j^r) \rangle_{\mathbb{C}} = \langle v_j^r, v_j^r \rangle - \langle w_j^r, w_j^r \rangle + i\langle v_j^r, w_j^r \rangle + i\langle w_j^r, v_j^r \rangle,$$

pois

$$(v_j^r, w_j^r) \in \widetilde{W_{\mu_r}}, (v_j^r, -w_j^r) \in \widetilde{W_{\mu_r}} \text{ e } \widetilde{W_{\mu_r}} \perp \widetilde{W_{\mu_r}}.$$

Logo, $\langle v_j^r, w_j^r \rangle = 0$ e

$$\begin{cases} \langle v_j^r, v_j^r \rangle + \langle w_j^r, w_j^r \rangle = 1 \\ \langle v_j^r, v_j^r \rangle - \langle w_j^r, w_j^r \rangle = 0. \end{cases}$$

Então, $\langle v_j^r, w_j^r \rangle = 0$ e $\langle v_j^r, v_j^r \rangle = \langle w_j^r, w_j^r \rangle = \frac{1}{2}, r = 1, \dots, d_r$.

Afirmção: $\mathcal{B}_r = \{v_1^r + w_1^r, v_1^r - w_1^r, \dots, v_{d_r}^r + w_{d_r}^r, v_{d_r}^r - w_{d_r}^r\}$ é uma base ortonormal de $\text{Ker}(p_r(T))$.

De fato,

- $\langle v_j^r + w_j^r, v_j^r + w_j^r \rangle = \langle v_j^r, v_j^r \rangle + \langle w_j^r, w_j^r \rangle = 1$
- $\langle v_j^r - w_j^r, v_j^r - w_j^r \rangle = \langle v_j^r, v_j^r \rangle + \langle w_j^r, w_j^r \rangle = 1$
- $\langle v_j^r + w_j^r, v_j^r - w_j^r \rangle = \langle v_j^r, v_j^r \rangle - \langle w_j^r, w_j^r \rangle = 0$

Além disso, como os vetores $(v_j^r, w_j^r), (v_j^r, -w_j^r), (v_t^r, w_t^r)$ e $(v_t^r, -w_t^r)$ são dois a dois ortogonais, se $t \neq j$, temos que

- o vetor $(v_j^r, \mathbf{0}) = \frac{1}{2}(v_j^r, w_j^r) + \frac{1}{2}(v_j^r, -w_j^r)$ é ortogonal ao vetor $(w_t^r, \mathbf{0}) = \frac{i}{2}(v_t^r, -w_t^r) - \frac{i}{2}(v_t^r, w_t^r)$, ou seja,

$$\langle (v_j^r, \mathbf{0}), (w_t^r, \mathbf{0}) \rangle_{\mathbb{C}} = \langle v_j^r, w_t^r \rangle = 0.$$

- o vetor $(v_j^r, \mathbf{0}) = \frac{1}{2}(v_j^r, w_j^r) + \frac{1}{2}(v_j^r, -w_j^r)$ é ortogonal ao vetor $(v_t^r, \mathbf{0}) = \frac{1}{2}(v_t^r, w_t^r) + \frac{1}{2}(v_t^r, -w_t^r)$, ou seja,

$$\langle (v_j^r, \mathbf{0}), (v_t^r, \mathbf{0}) \rangle_{\mathbb{C}} = \langle v_j^r, v_t^r \rangle = 0.$$

- o vetor $(w_j^r, \mathbf{0}) = \frac{i}{2}(v_j^r, -w_j^r) - \frac{i}{2}(v_j^r, w_j^r)$ é ortogonal ao vetor $(w_t^r, \mathbf{0}) = \frac{i}{2}(v_t^r, -w_t^r) - \frac{i}{2}(v_t^r, w_t^r)$, ou seja,

$$\langle (w_j^r, \mathbf{0}), (w_t^r, \mathbf{0}) \rangle_{\mathbb{C}} = \langle w_j^r, w_t^r \rangle = 0.$$

Logo,

- $\langle v_j^r + w_j^r, v_t^r + w_t^r \rangle = \langle v_j^r, v_t^r \rangle + \langle w_j^r, w_t^r \rangle + \langle w_j^r, v_t^r \rangle + \langle v_j^r, w_t^r \rangle = 0$, se $t \neq j$.
- $\langle v_j^r + w_j^r, v_t^r - w_t^r \rangle = \langle v_j^r, v_t^r \rangle - \langle w_j^r, w_t^r \rangle - \langle v_j^r, w_t^r \rangle + \langle w_j^r, v_t^r \rangle = 0$, se $t \neq j$.
- Seja F_j^r o subespaço de V gerado por $\mathcal{B}_j^r = \{v_j^r + w_j^r, v_j^r - w_j^r\}$, $j = 1, \dots, d_r$. Então, $F_j^r \perp F_t^r$, se $j \neq t$, $\dim F_j^r = 2$, $\text{Ker}(p_r(T)) = F_1^r \oplus \dots \oplus F_{d_r}^r$ e

$$[T|_{F_j^r}]_{\mathcal{B}_j^r} = \begin{pmatrix} a_r \cos \theta_r & -a_r \sin \theta_r \\ a_r \sin \theta_r & a_r \cos \theta_r \end{pmatrix}_{2 \times 2}$$

Afirmção: $\text{Ker}(p_r(T)) \perp \text{Ker}(p_s(T))$, se $r \neq s$.

De fato, se $v \in \text{Ker}(p_r(T))$ e $w \in \text{Ker}(p_s(T))$, então $(v, \mathbf{0}) \in \widetilde{W}_{\mu_r} \oplus \widetilde{W}_{\mu_r}$ e $(w, \mathbf{0}) \in \widetilde{W}_{\mu_s} \oplus \widetilde{W}_{\mu_s}$.

Como $\widetilde{W}_{\mu_r} \oplus \widetilde{W}_{\mu_r} \perp \widetilde{W}_{\mu_s} \oplus \widetilde{W}_{\mu_s}$, se $s \neq r$, temos que

$$\langle v, w \rangle = \langle (v, \mathbf{0}), (w, \mathbf{0}) \rangle_{\mathbb{C}} = 0.$$

Afirmção: $\text{Ker}(p_r(T)) \perp W_{\lambda_j}$, $r = 1, \dots, k$, $j = 1, \dots, \ell$.

De fato, como

$$\widehat{W}_{\lambda_j} = \text{Ker}(\widehat{T} - \lambda_j I), \quad \text{Ker}(p_r(T))^\wedge = \widetilde{W}_{\mu_r} \oplus \widetilde{W}_{\mu_r} \text{ e } \widehat{W}_{\lambda_j} \perp \widetilde{W}_{\mu_r} \oplus \widetilde{W}_{\mu_r},$$

temos que se $v \in W_{\lambda_j}$ e $w \in \text{Ker}(p_r(T))$, então

$$(v, \mathbf{0}) \in \widehat{W_{\lambda_j}} \quad \text{e} \quad (w, \mathbf{0}) \in \widehat{W_{\mu_r}} \oplus \widehat{W_{\mu_r}}.$$

Logo,

$$\langle v, w \rangle = \langle (v, 0), (w, 0) \rangle_{\mathbb{C}} = 0.$$

Sendo também $W_{\lambda_j} \perp W_{\lambda_t}$, se $t \neq j$, obtemos que

$$V = \text{Ker}(p_1(T)) \oplus \dots \oplus \text{Ker}(p_k(T)) \oplus W_{\lambda_1} \oplus \dots \oplus W_{\lambda_\ell}$$

é uma decomposição em soma direta por subespaços invariantes dois a dois ortogonais.

Como $\dim(W_{\lambda_j}) = f_j$ e $T|_{W_{\lambda_j}} = \lambda_j I$, existe uma base ortonormal $\mathcal{B}_{k+j}$ de W_{λ_j} tal que

$$[T|_{W_{\lambda_j}}]_{\mathcal{B}_{k+j}} = \begin{pmatrix} \lambda_j & & \\ & \ddots & \\ & & \lambda_j \end{pmatrix}_{f_j \times f_j}.$$

Logo, $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k \cup \Gamma_1 \cup \dots \cup \Gamma_\ell$ é uma base ortonormal de V tal que $[T]_{\mathcal{B}}$ está na forma canônica de Jordan real.

Corolário 4.3

Seja A uma matriz normal real.

Então, existe uma matriz ortogonal real P tal que $P^{-1}AP = P^tAP$ está na forma canônica de Jordan real.

Observação 4.10

Se V é real e $T : V \rightarrow V$ é um operador normal, então T é semi-simples, pois $\hat{T} : \hat{V} \rightarrow \hat{V}$ é normal, e, portanto, diagonalizável.

Proposição 4.6

Seja V um espaço vetorial de dimensão finita sobre $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) com produto interno e seja T um operador normal sobre V .

Então:

- (a) T é auto-adjunto $\iff$ todas as raízes de p_c são reais.
- (b) T é unitário $\iff$ todas as raízes de p_c tem norma igual a 1.

Prova.

(a) ($\implies$) Se T é auto-adjunto, todos os autovalores de T são reais e existe uma base ortonormal $\mathcal{B}$ de V tal que $[T]_{\mathcal{B}}$ é diagonal.

Logo, $p_c = (x - \lambda_1)^{d_1} \dots (x - \lambda_k)^{d_k}$, onde $\lambda_1, \dots, \lambda_k$ são os autovalores distintos de T . Então, todas as raízes de p_c são reais.

($\impliedby$) Se T é normal e todas as raízes de p_c são reais, existe, V possui uma base ortonormal $\mathcal{B}$ de autovetores, pelo teorema 4.2 ($\mathbb{K} = \mathbb{C}$), e pelo teorema 4.3 ($\mathbb{K} = \mathbb{R}$). Como

$$[T]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \text{ onde } \lambda_i \in \mathbb{R}, i = 1, \dots, n,$$

temos que $[T^*]_{\mathcal{B}} = [T]_{\mathcal{B}}^* = [T]_{\mathcal{B}}$, ou seja, $T^* = T$.

(b) ($\implies$) Seja T um operador unitário e seja v um autovetor não-nulo tal que $T(v) = \lambda v$. Então,

$$|\lambda|^2 \langle v, v \rangle = \lambda \bar{\lambda} \langle v, v \rangle = \langle \lambda v, \lambda v \rangle = \langle T(v), T(v) \rangle = \langle v, v \rangle.$$

Como $\langle v, v \rangle \neq 0$, temos que $|\lambda| = 1$, ou seja, $\lambda = e^{i\theta}$, $\theta \in (0, \pi)$, se $\lambda \in \mathbb{C}$ e $\lambda = \pm 1$, se $\lambda \in \mathbb{R}$.

Caso $\mathbb{K} = \mathbb{C}$: Como todo operador unitário T é normal, V possui uma base ortonormal $\mathcal{B}$ de autovetores de T . Logo, todas as raízes de p_c têm norma igual a 1.

Caso $\mathbb{K} = \mathbb{R}$: Seja $\mathcal{B}$ uma base ortonormal de V e seja $A = [T]_{\mathcal{B}}$. Como $[T^*]_{\mathcal{B}} = [T]_{\mathcal{B}}^* = A^*$, temos que

$$AA^* = [T]_{\mathcal{B}}[T]_{\mathcal{B}}^* = [T]_{\mathcal{B}}[T^*]_{\mathcal{B}} = [TT^*]_{\mathcal{B}} = I = [T^*T]_{\mathcal{B}} = [T]_{\mathcal{B}}^*[T]_{\mathcal{B}} = A^*A.$$

Logo, o operador $L : \mathbb{C}^{n \times 1} \rightarrow \mathbb{C}^{n \times 1}$ definido por $L(X) = AX$ é unitário sobre o espaço complexo $\mathbb{C}^{n \times 1}$. Como o polinômio característico p_c de T é igual ao polinômio característico de L , isto é, $p_c = \det(xI - A)$, temos que todas as raízes de p_c têm norma igual a 1.

($\impliedby$) Caso $\mathbb{K} = \mathbb{C}$: Se T é normal e todas as raízes de p_c têm norma igual a 1, existe uma base ortonormal de autovetores $\mathcal{B} = \{v_1, \dots, v_n\}$ tais que $T(v_i) = \lambda_i v_i$, sendo $|\lambda_i| = 1$, $i = 1, \dots, n$.

Como $\mathcal{B}' = \{T(v_1), \dots, T(v_n)\} = \{\lambda_1 v_1, \dots, \lambda_n v_n\}$ é uma base ortonormal de V , temos que T é um operador unitário, pois leva a base ortonormal $\mathcal{B}$ na base ortonormal $\mathcal{B}'$.

Caso $\mathbb{K} = \mathbb{R}$: Seja $\mathcal{B}$ uma base ortonormal de V e seja $A = [T]_{\mathcal{B}}$. Como $AA^* = A^*A$, pois T é normal, temos que o operador $L : \mathbb{C}^{n \times 1} \rightarrow \mathbb{C}^{n \times 1}$ definido por $L(X) = AX$ é normal.

Como $p_c = p_c^L = \det(xI - A)$ têm, por hipótese, todas suas raízes de norma 1, temos que L é unitário, ou seja, $AA^* = A^*A = I$.

Logo, T é unitário, pois

$$[T^*T]_{\mathcal{B}} = [T^*]_{\mathcal{B}}[T]_{\mathcal{B}} = [T]_{\mathcal{B}}^*[T]_{\mathcal{B}} = I = [T]_{\mathcal{B}}[T]_{\mathcal{B}}^* = [T]_{\mathcal{B}}[T^*]_{\mathcal{B}} = [TT^*]_{\mathcal{B}},$$

ou seja, $[T^*T]_{\mathcal{B}} = [TT^*]_{\mathcal{B}} = I$. Assim, $T^*T = TT^* = I$. ■

Como consequência da proposição anterior e do teorema 4.3, obtemos o seguinte teorema sobre operadores unitários.

Teorema 4.4

Seja V um espaço vetorial sobre $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) de dimensão finita com produto interno e seja T um operador unitário sobre V . Então:

(1) Se $\mathbb{K} = \mathbb{C}$, V possui uma base ortonormal de autovetores de T com autovalores de módulo 1.

(2) Se $\mathbb{K} = \mathbb{R}$,

- $p_c = p_1^{d_1} \dots p_k^{d_k} (x-1)^{f_1} (x+1)^{f_2}$, onde $p_j = (x - e^{i\theta_j})(x - e^{-i\theta_j})$, $\theta_j \in (0, \pi)$, $j = 1, \dots, k$.
- $p_m = p_1 \dots p_k (x-1)(x+1)$.
- $V = \text{Ker}(p_1(T)) \oplus \dots \oplus \text{Ker}(p_k(T)) \oplus W_1 \oplus W_{-1}$ é soma direta de subespaços invariantes dois a dois ortogonais, onde $W_1 = \text{Ker}(T - I)$ e $W_{-1} = \text{Ker}(T + I)$, $\dim(\text{Ker}(p_j(T))) = 2d_j$, $j = 1, \dots, k$, $\dim(W_1) = f_1$ e $\dim(W_{-1}) = f_2$.
- existe uma base ortonormal $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k \cup \gamma_1 \cup \gamma_{-1}$ de V tal que:

- γ_1 é uma base de W_1 tal que

$$[T|_{W_1}]_{\gamma_1} = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}_{f_1 \times f_1} = B_1.$$

- γ_{-1} é uma base de W_{-1} tal que

$$[T|_{W_{-1}}]_{\gamma_{-1}} = \begin{pmatrix} -1 & & \\ & \ddots & \\ & & -1 \end{pmatrix}_{f_2 \times f_2} = B_{-1}.$$

◦ $\mathcal{B}_j$ é uma base de $\text{Ker}(p_j(T))$ tal que

$$[T|_{\text{Ker}(p_j(T))}]_{\mathcal{B}_j} = \begin{pmatrix} R_1^j & & \\ & \ddots & \\ & & R_{d_j}^j \end{pmatrix}_{2d_j \times 2d_j} = A_j,$$

sendo $R_i^j = \begin{pmatrix} \cos \theta_j & -\sin \theta_j \\ \sin \theta_j & \cos \theta_j \end{pmatrix}$, $i = 1, \dots, d_j$, $j = 1, \dots, k$.

Ou seja,

$$[T]_{\mathcal{B}} = \begin{pmatrix} A_1 & & & & \\ & \ddots & & & \\ & & A_k & & \\ & & & B_1 & \\ & & & & B_{-1} \end{pmatrix}$$

está na forma canônica de Jordan real.

Veremos, agora, algumas propriedades importantes de um operador normal.

Proposição 4.7

Seja T um operador normal sobre um espaço vetorial V sobre $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) com produto interno. Então, $(T(V))^\perp = \text{Ker}(T)$.

Prova.

Suponhamos que $w \in (T(V))^\perp$, ou seja, $\langle w, T(v) \rangle = 0$ para todo $v \in V$. Então, $\langle T^*(w), v \rangle = 0$ para todo $v \in V$. Logo, $T^*(w) = \mathbf{0}$.

Como $\|T(w)\| = \|T^*(w)\|$, temos que $T(w) = \mathbf{0}$, ou seja $w \in \text{Ker}(T)$.

Suponhamos, agora, que $w \in \text{Ker}(T)$, ou seja, $T(w) = \mathbf{0}$.

Então, $T^*(w) = \mathbf{0}$ e $\langle T^*(w), v \rangle = \langle w, T(v) \rangle = 0$, para todo $v \in V$.

Logo, $w \in (T(V))^\perp$. ■

Proposição 4.8

Se T é um operador normal e v um vetor tal que $T^2(v) = \mathbf{0}$, então $T(v) = \mathbf{0}$.

Prova.

Como $T(T(v)) = \mathbf{0}$, temos que $T(v) \in T(V) \cap \text{Ker}(T)$. Logo, $T(v) \in T(V) \cap (T(V))^\perp = \{\mathbf{0}\}$. Ou seja, $T(v) = \mathbf{0}$. ■

Proposição 4.9

Seja T um operador normal e f um polinômio com coeficientes no corpo de escalares. Então, $f(T)$ é um operador normal.

Prova.

Seja $f = a_0 + a_1x + \dots + a_nx^n$. Então, $f(T) = a_0I + a_1T + \dots + a_nT^n$ e $f(T)^* = \overline{a_0}I + \overline{a_1}T^* + \dots + \overline{a_n}(T^n)^*$.

Como $T^*T = TT^*$, segue que $f(T)$ e $f(T)^*$ comutam. ■

Proposição 4.10

Seja T um operador normal sobre um espaço V de dimensão finita com produto interno e sejam $W_1, \dots, W_k$ as componentes primárias de V sob T . Se W é um subespaço de V invariante por T , então

$$W = \bigoplus_{i=1}^k W_i \cap W$$

Prova.

Obviamente, $\sum_{i=1}^k (W_i \cap W) \subset W$. Seja $v \in W$ e seja $\pi_j : V \rightarrow W_j$ a projeção ortogonal de V sobre W_j , $j = 1, \dots, k$.

Como cada π_j é um polinômio em T e W é invariante por T , temos que $\pi_j(v) \in W \cap W_j$, $j = 1, \dots, k$. Além disso, como

$$I = \pi_1 + \dots + \pi_k,$$

temos que $v = \pi_1(v) + \dots + \pi_k(v)$.

Logo, $v \in \sum_{j=1}^k W_j \cap W$, ou seja, $W \subset \sum_{j=1}^k W_j \cap W$.

Assim, $W = \sum_{j=1}^k W_j \cap W$, ou melhor ainda, $W = \bigoplus_{j=1}^k W_j \cap W$, pois os subespaços $W_j \cap W$, $j = 1, \dots, k$, são linearmente independentes. ■

Teorema 4.5

Seja V um espaço vetorial real de dimensão finita com produto interno e seja T um operador normal sobre V .

Se

$$p = (x - a)^2 + b^2, \quad a, b \in \mathbb{R}, \quad b \neq 0,$$

é o polinômio minimal de T , então T é invertível e $T^* = (a^2 + b^2)T^{-1}$.

Prova.

Pelo teorema 4.3, temos que $V = \text{Ker}(p(T))$, $p_c = p^r$, onde $2r = \dim(V)$ e V possui uma base ortonormal $\mathcal{B}$ tal que

$$[T]_{\mathcal{B}} = \begin{pmatrix} a & -b & & & \\ b & a & & & \\ & & a & -b & \\ & & b & a & \\ & & & \ddots & \\ & & & & a & -b \\ & & & & b & a \end{pmatrix}_{2r \times 2r}$$

Então, $\det[T]_{\mathcal{B}} = (a^2 + b^2)^r \neq 0$. Logo, T é invertível.

Como

$$[T^*]_{\mathcal{B}} = [T]_{\mathcal{B}}^* = \begin{pmatrix} a & b & & & \\ -b & a & & & \\ & & a & b & \\ & & -b & a & \\ & & & \ddots & \\ & & & & a & b \\ & & & & -b & a \end{pmatrix}_{2r \times 2r},$$

temos que

$$[TT^*]_{\mathcal{B}} = \begin{pmatrix} a^2 + b^2 & & & & \\ & a^2 + b^2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & a^2 + b^2 \end{pmatrix}_{2r \times 2r},$$

ou seja, $TT^* = (a^2 + b^2)I$. Logo, $T^* = (a^2 + b^2)T^{-1}$. ■

Teorema 4.6

Seja T um operador normal sobre um espaço V de dimensão finita com produto interno.

Então, qualquer operador linear que comuta com T , também comuta com T^* . Além disso, todo subespaço T -invariante é também T^* -invariante.

Prova.

Suponhamos que U é um operador linear que comuta com T .

Seja $V = W_1 \oplus \dots \oplus W_k$ a decomposição primária de V com respeito a T , e seja $\pi_j : V \longrightarrow W_j$ a projeção ortogonal de V sobre W_j , $j = 1, \dots, k$.

Como T comuta com U e π_j é um polinômio em T , temos que π_j e U comutam, $j = 1, \dots, k$.

Assim, se $v \in W_j$,

$$U(v) = U(\pi_j(v)) = \pi_j(U(v)),$$

ou seja, $U(v) \in W_j$. Logo, W_j é invariante por U , $j = 1, \dots, k$.

Sejam $U_j = U|_{W_j}$ e $T_j = T|_{W_j}$.

Se $W_j = \text{Ker}(T - c_j I)$, então $T_j = c_j I$. Logo, U_j comuta com $T_j^* = \overline{c_j} I$.

Se V é real e $W_j = \text{Ker}(p_j(T))$, onde $p_j = (x - \alpha_j)^2 + b_j^2$, $b_j > 0$, temos que p_j é o polinômio minimal de T_j . Logo, pelo teorema anterior,

$$T_j^* = (\alpha_j^2 + b_j^2)T_j^{-1}.$$

Como $T_j U_j = U_j T_j$, temos que $U_j T_j^{-1} = T_j^{-1} U_j$. Então, U_j comuta com T_j^* .

Temos, também, que T^* comuta com π_j , pois π_j é um polinômio em T e $T T^* = T^* T$. Logo, W_j é invariante por T^* e $T_j^* = T^*|_{W_j}$.

Assim, se $v \in W_j$,

$$U T^*(v) = U_j T_j^*(v) = T_j^* U_j(v) = T_j^*(U(v)) = T^* U(v),$$

pois $U(v) \in W_j$.

Como $V = W_1 \oplus \dots \oplus W_k$, temos que $U T^*(v) = T^* U(v)$ para todo $v \in V$, pois se $v = v_1 + \dots + v_k$, $v_j \in W_j$, $j = 1, \dots, k$, então

$$\begin{aligned} U T^*(v) &= U T^*(v_1 + \dots + v_k) = U T^*(v_1) + \dots + U T^*(v_k) \\ &= T^* U(v_1) + \dots + T^* U(v_k) = T^* U(v_1 + \dots + v_k) \\ &= T^* U(v). \end{aligned}$$

Logo, U comuta com T^* .

Seja, agora, W um subespaço de V invariante por T e seja $Z_j = W \cap W_j$, $j = 1, \dots, k$.

Pela proposição 4.10, $W = Z_1 \oplus \dots \oplus Z_k$.

Como W e W_j são invariantes por T , temos que $Z_j = W \cap W_j$ é invariante por T e, portanto, por $T_j = T|_{W_j}$.

Vamos, agora, mostrar que Z_j é também invariante por T_j^* .

Se $T_j = c_j I$, temos que Z_j é invariante por T_j^* , pois $T_j^* = \bar{c}_j I$.

Se não for o caso, sabemos que T_j é invertível e $T_j(Z_j) \subset Z_j$.

Logo, $T_j(Z_j) = Z_j$. Então, $T_j^{-1}(Z_j) = Z_j$.

Como $T_j^* = (a_j^2 + b_j^2)T_j^{-1}$, temos que $T_j^*(Z_j) = Z_j$.

Se $w \in W$, existem $w_j \in Z_j$, $j = 1, \dots, k$, tais que $w = w_1 + \dots + w_k$. Logo,

$$\begin{aligned} T^*(w) &= T^*(w_1 + \dots + w_k) = T^*(w_1) + \dots + T^*(w_k) \\ &= T_1^*(w_1) + \dots + T_k^*(w_k) \in Z_1 \oplus \dots \oplus Z_k = W. \end{aligned}$$

Assim, $T^*(W) \subset W$. ■

Observação 4.11

Seja V um espaço vetorial de dimensão finita com produto interno e T um operador normal sobre V .

Já sabemos que se $T(W) \subset W$, então $T^*(W) \subset W$, pelo teorema anterior, e $T^*(W^\perp) \subset W^\perp$, pela proposição 4.4. Logo, $T(W^\perp) \subset W^\perp$, pois $T(W^\perp) = (T^*)^*(W^\perp) \subset W^\perp$, já que $T^*(W^\perp) \subset W^\perp$. Ou seja, se W é invariante por T , então W é invariante por T^* e $W^\perp$ é invariante por T e T^* .

Observação 4.12

Se $T : V \rightarrow V$ é normal e $T(W) \subset W$, então $(T|_W)^* = (T^*)|_W$ e $T|_W$ é normal.

De fato, como

$$\langle v, T^*w \rangle = \langle T(v), w \rangle = \langle T|_W(v), w \rangle = \langle v, (T|_W)^*(w) \rangle,$$

para quaisquer $v, w \in W$, e $T^*(w) \in W$, temos que $T^*(w) = (T|_W)^*(w)$, já que

$$\langle T^*(w) - (T|_W)^*(w), T^* - (T|_W)^*(w) \rangle = 0.$$

Logo, $T^*|_W = (T|_W)^*$ e $(T|_W)(T|_W)^* = (T|_W)^*(T|_W)$.

Com base nas observações acima, podemos demonstrar a seguinte versão mais forte do teorema da decomposição cíclica para operadores normais.

Teorema 4.7

Seja T um operador linear normal sobre um espaço V de dimensão finita com produto interno. Então existem r vetores não-nulos $v_1, \dots, v_r$, com respectivos T -anuladores $p_1, \dots, p_r$, tais que:

(a) $V = Z(v_1, T) \oplus \dots \oplus Z(v_r, T)$.

(b) $p_{k+1} | p_k$, $k = 1, \dots, r-1$.

(c) $Z(v_i, T) \perp Z(v_j, T)$, se $i \neq j$.

Além do mais, o número r e os T -anuladores $p_1, \dots, p_r$ são determinados de modo único pelas condições (a) e (b) e pelo fato de que nenhum v_k é nulo.

Prova.

Vamos provar o resultado por indução sobre a dimensão do espaço V .

Se $\dim(V) = 1$, $V = Z(v, T)$, para todo $v \in V - \{0\}$.

Suponhamos que o teorema é válido para todo operador normal sobre um espaço de dimensão menor ou igual a $n-1$.

Seja V um espaço de dimensão n e T um operador normal sobre V .

Pelo Teorema da Decomposição Cíclica, existem r vetores $w_1, \dots, w_r$ não-nulos tais que

$$V = Z(w_1, T) \oplus \dots \oplus Z(w_r, T),$$

tais que q_{j+1} divide q_j , para todo $j = 1, \dots, r-1$, onde q_j é o T -anulador de w_j .

Seja $W = Z(w_1, T)$. Então, $\dim(W) \geq 1$, W é T -invariante, e q_1 é o polinômio minimal de T . Além disso, $W^\perp$ é T -invariante e $V = W \oplus W^\perp$.

Logo, $T|_{W^\perp}$ é normal e $\dim(W^\perp) \leq n-1$.

Pela hipótese de indução, existem s vetores não-nulos $u_1, \dots, u_s$ em $W^\perp$ tais que

$$W^\perp = Z(u_1, T|_W) \oplus \dots \oplus Z(u_s, T|_W),$$

tais que $Z(u_i, T|_W) \perp Z(u_j, T|_W)$, se $i \neq j$ e f_{i+1} divide f_i , $i = 1, \dots, s-1$, onde f_i é o $T|_W$ -anulador de u_i .

Como $Z(u_i, T|_W) = Z(u_i, T)$, $i = 1, \dots, s$, temos que

$$W^\perp = Z(u_1, T) \oplus \dots \oplus Z(u_s, T).$$

Além disso, como f_1 é o polinômio minimal de $T|_{W^\perp}$, temos que f_1 divide p_1 .

Logo, $V = W \oplus W^\perp = Z(w_1, T) \oplus Z(u_1, T) \oplus \dots \oplus Z(u_s, T)$ e $p_1 \mid f_1 \mid f_2 \mid \dots \mid f_s$.

Ou seja

$$V = Z(w_1, T) \oplus Z(u_1, T) \oplus \dots \oplus Z(u_s, T),$$

é uma decomposição cíclica ortogonal que satisfaz (b), $w_1 \neq \mathbf{0}$ e $u_j \neq \mathbf{0}$, $j = 1, \dots, s$.

Logo, $s = r - 1$. ■

Observação 4.13

Pode não existir uma base $\mathcal{B}$ ortonormal tal que $[T]_{\mathcal{B}}$ está na forma canônica racional.

Com efeito, considere o operador $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ dado por $T(X) = AX$, onde $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$.

Como $A^t = A$, T é auto-adjunto e, portanto, normal.

Além disso, $p_m = p_c = (x - 1)(x - 2) = x^2 - 3x + 2$. Logo, T possui um vetor cíclico.

Tome $v = e_1 + e_2 = (1, 1)$. Então, $T(v) = e_1 + 2e_2 = (1, 2)$.

Logo, $V = Z(v, T)$.

Suponha que existe uma base $\mathcal{B}$ ortonormal tal que

$$[T]_{\mathcal{B}} = \begin{pmatrix} 0 & -2 \\ 1 & 3 \end{pmatrix}.$$

Então, existe $w \in V - \{\mathbf{0}\}$ tal que $\mathcal{B} = \{w, T(w)\}$.

Suponha que $w = ae_1 + be_2$, com $a^2 + b^2 = 1$.

Assim, $T(w) = aT(e_1) + bT(e_2) = ae_1 + 2be_2$, com $a^2 + 4b^2 = 1$, e $\langle w, T(w) \rangle = a^2 + 2b^2 = 0$. Logo, $a = b = 0$, o que é uma contradição, já que $w \neq \mathbf{0}$.

Definição 4.2

Um operador T sobre um espaço vetorial V de dimensão finita com produto interno é dito:

- **não-negativo** se $T^* = T$ e $\langle T(v), v \rangle \geq 0$, para todo $v \in V$;
- **positivo** se $T^* = T$ e $\langle T(v), v \rangle > 0$ para todo $v \in V - \{0\}$.

Proposição 4.11

Seja T um operador normal sobre um espaço vetorial V de dimensão finita com produto interno sobre $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$). Então T é não-negativo (ou positivo) se, e somente se, todas as raízes do seu polinômio característico são números reais não-negativos (resp. positivos).

Prova.

Se T é não-negativo (ou positivo), todas as raízes de p_c são reais pela proposição 4.6, pois T é auto-adjunto. Assim, todas as raízes de p_c são autovalores de T .

Então, se λ é uma raiz de p_c , existe $v \in V - \{0\}$ tal que $T(v) = \lambda v$. Logo, $\lambda \langle v, v \rangle = \langle T(v), v \rangle \geq 0$ (ou > 0).

Como $\langle v, v \rangle > 0$, temos que $\lambda \geq 0$ (ou > 0).

Suponhamos, agora, que todas as raízes de p_c são números reais não-negativos (ou positivos). Então, pela proposição 4.6, T é auto-adjunto e, portanto, diagonalizável. Seja $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base ortonormal de autovetores de T , com $T(v_i) = \lambda_i v_i$, $\lambda_i \geq 0$ (ou > 0), $i = 1, \dots, n$.

Seja $v \in V - \{0\}$. Então existem $a_1, \dots, a_n \in \mathbb{K}$ não todos nulos, tais que $v = a_1 v_1 + \dots + a_n v_n$.

Logo,

$$\begin{aligned} \langle T(v), v \rangle &= \langle a_1 \lambda_1 v_1 + \dots + a_n \lambda_n v_n, a_1 v_1 + \dots + a_n v_n \rangle \\ &= \lambda_1 a_1^2 + \dots + \lambda_n a_n^2 \geq 0 \text{ (ou } > 0 \text{)}. \end{aligned}$$

Assim, T é não-negativo (ou positivo). ■

Já observamos antes que podemos fazer uma analogia entre a operação de se tomar adjunto de um operador e a operação de se tomar o conjugado de um número complexo. Um número complexo z é real ou de módulo 1 conforme $z = \bar{z}$ ou $z\bar{z} = 1$. E um operador T é auto-adjunto ou unitário conforme $T = T^*$ ou $TT^* = I$.

Vamos, agora demonstrar dois teoremas que são os análogos para operadores destas duas afirmações sobre números complexos:

(a) Todo número não-negativo possui uma única raiz quadrada não-negativa.

(b) Todo número complexo pode ser expresso sob a forma $z = ru$, onde r é não-negativo e $|u| = 1$, ou seja, $u = e^{i\theta}$, para algum $\theta \in \mathbb{R}$, que é a decomposição polar para números complexos.

Teorema 4.8

Seja T um operador não-negativo sobre um espaço vetorial V de dimensão finita com produto interno. Então T possui uma única raiz quadrada não-negativa, ou seja, existe um, e somente um, operador N não-negativo tal que $N^2 = T$.

Prova.

Existência: Sejam $\lambda_1, \dots, \lambda_k$ os autovalores distintos de T . Como T é diagonalizável, pois T é auto-adjunto, temos que:

- $V = W_{\lambda_1} \oplus \dots \oplus W_{\lambda_k}$, onde W_{λ_i} é o autoespaço de T correspondente ao autovalor λ_i , $i = 1, \dots, k$.
- $I = \pi_1 + \dots + \pi_k$, onde $\pi_i : V \rightarrow W_{\lambda_i}$ é a projeção ortogonal de V sobre W_{λ_i} , $i = 1, \dots, k$.
- $\pi_i^2 = \pi_i$, $i = 1, \dots, k$.
- $\pi_i \pi_j = \mathcal{O}$, se $i \neq j$.
- $T = \lambda_1 \pi_1 + \dots + \lambda_k \pi_k$ (resolução espectral de T).

Como $\lambda_i \geq 0$, $i = 1, \dots, k$, podemos considerar o operador

$$N = \sqrt{\lambda_1} \pi_1 + \dots + \sqrt{\lambda_k} \pi_k.$$

Como os escalares $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_k}$ são distintos e os operadores π_i , $i = 1, \dots, k$, não-nulos, satisfazem as condições acima, temos que o operador N é diagonalizável e $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_k}$ são os seus autovalores distintos.

Além disso, como $\pi_i^* = \pi_i$ e $\sqrt{\lambda_i} \geq 0$, $i = 1, \dots, k$, temos que N é auto-adjunto com todos os autovalores não-negativos, ou seja N é um operador não-negativo tal que

$$\begin{aligned} N^2 &= (\sqrt{\lambda_1} \pi_1 + \dots + \sqrt{\lambda_k} \pi_k)(\sqrt{\lambda_1} \pi_1 + \dots + \sqrt{\lambda_k} \pi_k) \\ &= \lambda_1 \pi_1 + \dots + \lambda_k \pi_k = T, \end{aligned}$$

Unicidade: Seja U um operador não-negativo tal que $U^2 = T$. Como

U é auto-adjunto e todos os seus autovalores são não-negativos, U é diagonalizável e

$$U = d_1 \widetilde{\pi}_1 + \dots + d_s \widetilde{\pi}_s,$$

onde $d_1, \dots, d_s$ são os autovalores distintos de U e $\widetilde{\pi}_i : V \longrightarrow \widetilde{W}_{d_i}$ é a projeção ortogonal de V sobre $\widetilde{W}_{d_i}$, autoespaço de U correspondente ao autovalor d_i , $i = 1, \dots, s$.

Como $T = U^2 = d_1^2 \widetilde{\pi}_1 + \dots + d_s^2 \widetilde{\pi}_s$, $d_i^2 \neq d_j^2$, para $i \neq j$, $I = \widetilde{\pi}_1 + \dots + \widetilde{\pi}_s$, $\widetilde{\pi}_i \widetilde{\pi}_j = \mathcal{O}$, para $i \neq j$, $\widetilde{\pi}_i^2 = I$, $i = 1, \dots, s$, temos que $d_1^2, \dots, d_s^2$ são os autovalores distintos de T e $\widetilde{\pi}_i(V) = \widetilde{W}_{d_i^2}$ é o autoespaço de T associado a d_i^2 , $i = 1, \dots, s$.

Logo, $s = k$ e, para cada $j = 1, \dots, k$, existe $r_j \in \{1, \dots, k\}$ tal que $d_j^2 = \lambda_{r_j}$ e $\widetilde{\pi}_j = \pi_{r_j}$.

Então, $U = \sqrt{\lambda_{r_1}} \pi_{r_1} + \dots + \sqrt{\lambda_{r_k}} \pi_{r_k} = N$. ■

Teorema 4.9

Teorema da Decomposição Polar

Seja V um espaço de dimensão finita com produto interno e seja T um operador linear arbitrário sobre V . Então, existe um operador unitário U sobre V e um operador não-negativo N sobre V tais que $T = UN$.

O operador não-negativo N é único.

Se T é invertível, o operador unitário U também é único.

Prova.

Suponhamos que $T = UN$, onde U é unitário e N é não-negativo. Então, $T^* = N^* U^* = NU^*$.

Assim, $T^*T = NU^*UN = NIN = N^2$.

Como T^*T é não-negativo, pois $(T^*T)^* = T^*(T^*)^* = T^*T$ e $\langle T^*T(v), v \rangle = \langle T(v), T(v) \rangle \geq 0$, para todo $v \in V$, temos que N é determinado de modo único como sendo a única raiz quadrada não-negativa de T^*T .

Então, para iniciarmos a demonstração da existência de U e N , tomamos N como sendo a única raiz quadrada não negativa de T^*T .

Se T é *invertível*, então N também o é, porque, nesse caso, N é positivo, já que

$$\langle N(v), N(v) \rangle = \langle N^2(v), v \rangle = \langle T^*T(v), v \rangle = \langle T(v), T(v) \rangle > 0,$$

para todo $v \in V - \{0\}$.

Seja $U = TN^{-1}$. Então, $U^* = (N^{-1})^*T^* = (N^*)^{-1}T^* = N^{-1}T^*$. Logo,

$$\begin{aligned} UU^* &= TN^{-1}N^{-1}T^* = T(N^{-1})^2T^* = T(N^2)^{-1}T^* \\ &= T(T^*T)^{-1}T^* = TT^{-1}(T^*)^{-1}T^* = I, \end{aligned}$$

ou seja, U é unitário.

Suponhamos, agora, que T não é invertível.

Definamos primeiro U sobre a imagem de N .

Seja v um vetor na imagem de N , digamos $N(w) = v$.

Definamos $U(v) = T(w)$, porque queremos que $UN(w) = T(w)$.

Precisamos verificar que U está bem definido sobre a imagem de N , ou seja, se $N(w) = N(w')$ então $T(w) = T(w')$.

Verificamos acima que $\|N(u)\| = \|T(u)\|$ para todo $u \in V$. Assim, fazendo $u = w - w'$, temos que $\|N(w - w')\| = 0$ se, e somente se, $\|T(w - w')\| = 0$. Portanto U está bem definida sobre a imagem de N e é linear.

Vamos, agora, definir U sobre $W^\perp$, onde $W = N(V)$.

Como $\text{Ker}(N) = \text{Ker}(T)$, temos que $\dim(N(V)) = \dim(T(V))$ e, portanto, $\dim(N(V)^\perp) = \dim(T(V)^\perp)$.

Logo, existe um isomorfismo $U_0 : W^\perp \rightarrow T(V)^\perp$ que preserva produto interno.

Como $V = W \oplus W^\perp$, todo vetor v pode ser escrito de modo único sob a forma $v = N(w) + u$, onde $N(w) \in W$ e $u \in W^\perp$. Definamos

$$U(v) = T(w) + U_0(u).$$

Então, U está bem definida e é linear.

Além disso,

$$\begin{aligned} \langle U(v), U(v) \rangle &= \langle T(w) + U_0(u), T(w) + U_0(u) \rangle \\ &= \langle T(w), T(w) \rangle + \langle U_0(u), U_0(u) \rangle, \end{aligned}$$

pois $T(w) \in T(V)$ e $U_0(u) \in T(V)^\perp$.

Logo,

$$\langle U(v), U(v) \rangle = \langle N(w), N(w) \rangle + \langle u, u \rangle = \langle N(w) + u, N(w) + u \rangle = \langle v, v \rangle,$$

pois $N(w) \in W$ e $u \in W^\perp$.

Então U é unitário e $UN(w) = T(w)$, para todo $w \in V$. ■

Observação 4.14

Seja $T = UN$, onde U é unitário e N é não-negativo. Temos que U e N comutam se, e somente se, N é normal.

De fato, se $T = UN = NU$, então

$$TT^* = (NU)(NU)^* = NUU^*N^* = NUU^*N = NIN = N^2 = T^*T.$$

Reciprocamente, se $TT^* = T^*T$, temos que $UN^2U^* = N^2$, pois

$$TT^* = UNN^*U^* = UN^2U^*,$$

e

$$T^*T = N^*U^*UN = N^2.$$

Como UNU^* é auto-adjunto e

$$\langle UNU^*(v), v \rangle = \langle N(U^*(v)), U^*(v) \rangle \geq 0,$$

para todo $v \in V$, temos que UNU^* é não-negativo e

$$(UNU^*)^2 = (UNU^*)(UNU^*) = UNU^*UNU^* = UNINU^* = UN^2U^* = N^2.$$

Pela unicidade da raiz quadrada de um operador não-negativo,

$$N = UNU^* = UNU^{-1}, \text{ ou seja, } NU = UN.$$

Observação 4.15

Todo operador T pode ser escrito na forma $T = N_1U_1$, onde N_1 é não-negativo e U_1 é unitário.

De fato, pelo teorema anterior, $T^* = UN$, onde U é unitário e N é não-negativo.

Logo,

$$T = (T^*)^* = N^*U^* = NU^*,$$

onde N é não-negativo e U^* é unitário.

5. Diagonalização simultânea de uma família comutativa de operadores normais

Definição 5.1

Seja $\mathcal{F}$ uma família de operadores lineares sobre um espaço V com produto interno sobre o corpo $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$).

Uma função $r : \mathcal{F} \rightarrow \mathbb{K}$ é chamada **uma raiz da família $\mathcal{F}$** se existir um vetor $v \in V - \{0\}$ tal que

$$T(v) = r(T)(v),$$

para todo $T \in \mathcal{F}$.

Para uma função arbitrária $r : \mathcal{F} \rightarrow \mathbb{K}$ indicaremos

$$V(r) = \{\alpha \in V \mid T(\alpha) = r(T)(\alpha), \quad \forall T \in \mathcal{F}\}$$

Observação 5.1

- $V(r)$ é um subespaço de V , pois $0 \in V(r)$ e se $v, w \in V(r)$ e $\lambda \in \mathbb{K}$, então

$$\begin{aligned} T(\lambda v + w) &= \lambda T(v) + T(w) = \lambda r(T)(v) + r(T)(w) \\ &= r(T)(\lambda v + w), \end{aligned}$$

para todo $T \in \mathcal{F}$.

- $V(r) \neq \{0\}$ se, e somente se, r é uma raiz de $\mathcal{F}$.
- Todo vetor v não-nulo em $V(r)$ é um autovetor para todo $T \in \mathcal{F}$.

Teorema 5.1

Seja $\mathcal{F}$ uma família comutativa de operadores normais diagonalizáveis sobre um espaço V de dimensão finita com produto interno. Então $\mathcal{F}$ possui apenas um número finito de raízes. Se $r_1, \dots, r_k$ são as raízes distintas de $\mathcal{F}$, então

$$(1) \quad V(r_i) \perp V(r_j), \text{ se } i \neq j.$$

$$(2) \quad V = V(r_1) \oplus \dots \oplus V(r_k).$$

Prova.

Suponhamos que r e s são raízes distintas de $\mathcal{F}$. Então existe um operador T em $\mathcal{F}$ tal que $r(T) \neq s(T)$.

Sejam $v \in V(r) - \{0\}$ e $w \in V(s) - \{0\}$.

Como $T(v) = r(T)(v)$, $T(w) = s(T)(w)$ e $r(T) \neq s(T)$, temos que v e w são autovetores de T associados a autovalores distintos. Logo, $\langle v, w \rangle = 0$, ou seja, $V(r) \perp V(s)$.

Como V tem dimensão finita, $V(r) \perp V(s)$, se $r \neq s$, e $\dim(V(r)) \geq 1$, temos que $\mathcal{F}$ possui apenas um número finito $r_1, \dots, r_k$ de raízes distintas.

Seja $\{T_1, \dots, T_m\}$ uma base do subespaço de $\mathcal{L}(V, V)$ gerado por $\mathcal{F}$.

• Para cada $i = 1, \dots, m$, sejam $\pi_1^i, \dots, \pi_{k_i}^i$ as projeções que determinam a resolução espectral do operador identidade determinado por T_i , ou seja

◦ $V = W_{\lambda_1^i} \oplus \dots \oplus W_{\lambda_{k_i}^i}$, onde λ_j^i , $j = 1, \dots, k_i$, são os autovalores distintos de T_i e $W_{\lambda_j^i}^i$ é o autoespaço de T_i correspondente ao autovalor λ_j^i .

◦ $\pi_j^i : V \longrightarrow W_{\lambda_j^i}^i$ é a projeção ortogonal de V sobre $W_{\lambda_j^i}^i$, $j = 1, \dots, k_i$.

◦ $I = \pi_1^i + \dots + \pi_{k_i}^i$.

◦ $T = \lambda_1^i \pi_1^i + \dots + \lambda_{k_i}^i \pi_{k_i}^i$.

◦ $\pi_j^i \pi_\ell^i = 0$, se $j \neq \ell$.

◦ $(\pi_j^i)^2 = \pi_j^i$, $j = 1, \dots, k_i$.

As projeções π_j^i , $i = 1, \dots, m$, $j = 1, \dots, k_i$, comutam entre si, pois π_j^i é um polinômio em T_i e os operadores $T_1, \dots, T_m$ comutam.

Como

$$I = \left(\sum_{j_1=1}^{k_1} \pi_{j_1}^1 \right) \dots \left(\sum_{j_m=1}^{k_m} \pi_{j_m}^m \right),$$

temos que

$$v = \sum_{j_1, \dots, j_m} \pi_{j_1}^1 \dots \pi_{j_m}^m(v), \quad (I)$$

para todo $v \in V$.

Suponhamos que $j_1, \dots, j_m$ são índices tais que $w = \pi_{j_1}^1 \dots \pi_{j_m}^m(v) \neq 0$.

Seja $w_i = \prod_{n \neq i} \pi_{j_n}^n(v)$.

Então, $w = \pi_{j_i}^i(w_i)$, pois as projeções comutam.

Logo, w é um autovetor de T_i associado ao autovalor $\lambda_{j_i}^i$, ou seja,

$$T_i(w) = \lambda_{j_i}^i w.$$

Seja $T \in \mathcal{F}$. Então existem escalares $b_1, \dots, b_m$ tais que $T = \sum_{i=1}^m b_i T_i$.

Assim,

$$T(w) = \sum_{i=1}^m b_i T_i(w) = \left(\sum_{i=1}^m b_i \lambda_{j_i}^i \right) (w).$$

Logo, $r : \mathcal{F} \rightarrow \mathbb{K}$, definida por

$$r(T) = \sum_{i=1}^m b_i \lambda_{j_i}^i,$$

é uma raiz de $\mathcal{F}$.

Então, $r = r_t$, para algum $t \in \{1, \dots, k\}$, e $w \in V(r_t) - \{0\}$.

Provamos, assim, que todo termo não-nulo de (I) pertence a um dos subespaços $V(r_1), \dots, V(r_k)$. Logo, $V = V(r_1) \oplus \dots \oplus V(r_k)$, pois os subespaços são ortogonais, logo, LI. ■

Observação 5.2

Seja $P_j : V \rightarrow V(r_j)$, $j = 1, \dots, k$, a projeção ortogonal sobre $V(r_j)$.

Então, $P_i P_j = 0$, se $i \neq j$,

$$I = P_1 + \dots + P_k,$$

e todo $T \in \mathcal{F}$ pode ser escrito na forma

$$T = \sum_{j=1}^k r_j(T) P_j. \quad (**)$$

Definição 5.2

A família de projeções ortogonais $\{P_1, \dots, P_k\}$ é denominada a *resolução espectral do operador identidade determinada por $\mathcal{F}$* e, $(**)$ é a *resolução espectral de T em termos desta família*.

Corolário 5.1

Se $\mathcal{F}$ é uma família comutativa de operadores normais diagonalizáveis sobre um espaço V de dimensão finita, então V possui uma base ortonormal $\mathcal{B}$ tal que $[T]_{\mathcal{B}}$ é diagonal para todo $T \in \mathcal{F}$.

Prova.

Seja $\mathcal{B}_i$ uma base ortonormal de $V(r_i)$, $i = 1, \dots, k$. Então $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ é uma base ortonormal de V de autovetores de T , para todo $T \in \mathcal{F}$, pois todos os vetores de $V(r_i)$, $i = 1, \dots, k$, são autovetores de T , para todo $T \in \mathcal{F}$. ■

Definição 5.3

Uma *álgebra auto-adjunta de operadores* sobre um espaço vetorial V com produto interno é uma sub-álgebra de $\mathcal{L}(V, V)$ que contém o adjunto de cada um de seus elementos.

Observação 5.3

A interseção de uma coleção qualquer de álgebras auto-adjuntas é uma álgebra auto-adjunta.

Definição 5.4

Se $\mathcal{F}$ é uma família de operadores lineares sobre um espaço de dimensão finita com produto interno, a *álgebra auto-adjunta gerada por $\mathcal{F}$* é a menor álgebra auto-adjunta que contém $\mathcal{F}$.

Teorema 5.2

Seja $\mathcal{F}$ uma família comutativa de operadores normais diagonalizáveis sobre um espaço V de dimensão finita com produto interno e seja $\mathcal{A}$ a álgebra auto-adjunta gerada por $\mathcal{F}$ e I (o operador identidade). Seja $\{P_1, \dots, P_k\}$ a resolução do operador identidade definida por $\mathcal{F}$. Então $\mathcal{A}$ é o conjunto de todos os operadores sobre V da forma

$$T = \sum_{j=1}^k c_j P_j, \quad (I)$$

onde $c_1, \dots, c_k$ são escalares arbitrários.

Prova.

Seja $\mathcal{C}$ o conjunto de todos os operadores sobre V da forma (I). Então $\mathcal{C}$ contém o operador identidade e o adjunto

$$T^* = \sum_{j=1}^k \bar{c}_j P_j,$$

de cada um de seus elementos.

Se $T = \sum_{j=1}^k c_j P_j$ e $U = \sum_{j=1}^k d_j P_j$, então, para λ escalar,

$$\lambda T + u = \sum_{j=1}^k (\lambda c_j + d_j) P_j,$$

e

$$T U = \sum_{i,j=1}^k c_j d_i P_j P_i = \sum_{j=1}^k c_j d_j P_j = U T.$$

Assim, $\mathbb{C}$ é uma álgebra comutativa auto-adjunta que contém $\mathcal{F}$ e o operador identidade. Logo, $\mathcal{A} \subset \mathcal{C}$.

Sejam $r_1, \dots, r_k$ as raízes distintas de $\mathcal{F}$. Então, para cada par de índices (i, n) , com $i \neq n$, existe um operador $T_{in} \in \mathcal{F}$ tal que $r_i(T_{in}) \neq r_n(T_{in})$.

Sejam $a_{in} = r_i(T_{in}) - r_n(T_{in})$ e $b_{in} = r_n(T_{in})$.

Então, o operador linear

$$Q_i = \prod_{n \neq i} a_{in}^{-1} (T_{in} - b_{in} I) \in \mathcal{A}.$$

Vamos mostrar que $Q_i = P_i$, $i = 1, \dots, k$.

Seja $j \neq i$ e $v \in V(r_j)$. Então

$$T_{ij}(v) = r_j(T_{ij})(v) = b_{ij}v,$$

ou seja, $(T_{ij} - b_{ij}I)(v) = \mathbf{0}$. Como todos os fatores de Q_i comutam, temos que $Q_i(v) = \mathbf{0}$.

Logo, $Q_i(v) = P_i(v)$, para todo $v \in V(r_j)$, $j \neq i$.

Seja, agora, $v \in V(r_i)$. Então, $T_{in}(v) = r_i(T_{in})(v)$ e

$$\begin{aligned} a_{in}^{-1} (T_{in} - b_{in} I)(v) &= a_{in}^{-1} (r_i(T_{in}) - b_{in})(v) \\ &= a_{in}^{-1} (r_i(T_{in}) - r_n(T_{in}))(v) = v. \end{aligned}$$

Assim, $Q_i(v) = v = P_i(v)$ para todo $v \in V(r_i)$. Logo, $Q_i = P_i$, $i = 1, \dots, k$.

Então, $\mathcal{C} \subset \mathcal{A}$. ■

Observação 5.4

O teorema mostra que a álgebra $\mathcal{A}$ é comutativa e que cada elemento de $\mathcal{A}$ é um operador normal diagonalizável.

De fato, se $\mathcal{B}_i$ é uma base ortonormal de $V(r_i)$, $i = 1, \dots, k$, então $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ é uma base ortonormal de V tal que $[T]_{\mathcal{B}}$ é diagonal para todo $T \in \mathcal{A}$, pois se $T = \sum_{j=1}^k c_j P_j$, então $T(v) = c_j v$ para todo $v \in V(r_j)$, $j = 1, \dots, k$, já que $P_i(v) = \mathbf{0}$, se $i \neq j$, e $P_j(v) = v$.

6. Formas

Definição 6.1

Uma **forma** (*sesquilinear*) sobre um espaço vetorial V sobre o corpo $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) é uma função $f : V \times V \longrightarrow \mathbb{K}$ tal que:

- $f(\lambda v + v', w) = \lambda f(v, w) + f(v', w)$;
- $f(v, \lambda w + w') = \bar{\lambda} f(v, w) + f(v, w')$;

quaisquer que sejam $v, w, v', w' \in V$ e $\lambda \in \mathbb{K}$.

Assim, f é uma função linear com respeito à primeira variável, e linear-conjugada com respeito à segunda variável.

Observação 6.1

- Se $\mathbb{K} = \mathbb{R}$, f é **bilinear**, ou seja, f é linear com respeito às duas variáveis.
- Se $\mathbb{K} = \mathbb{C}$, temos que f é bilinear se, e somente se, $f = \mathcal{O}$. De fato, se f é bilinear, $f(v, iw) = -if(v, w)$ e $f(v, iw) = if(v, w)$. Logo, $2if(v, w) = 0$, ou seja, $f(v, w) = 0$.

Observação 6.2

O conjunto das formas sobre V é um subespaço vetorial do espaço de todas as funções de $V \times V$ em $\mathbb{K}$.

Teorema 6.1

Seja V um espaço vetorial de dimensão finita com produto interno e f uma forma sobre V . Então existe um único operador linear T_f sobre V tal que

$$f(v, w) = \langle T_f(v), w \rangle,$$

para todos $v, w \in V$. Além disso, a aplicação $f \longmapsto T_f$ é um isomorfismo do espaço das formas em $\mathcal{L}(V, V)$.

Prova.

Seja $w \in V$ fixo. Como $v \mapsto f(v, w)$ é um funcional linear sobre V , existe um único vetor $w' \in V$ tal que $f(v, w) = \langle v, w' \rangle$ para todo $v \in V$.

Definimos a função $L : V \rightarrow V$, colocando $L(w) = w'$.

O operador L é linear, pois se $w_1, w_2 \in V$ e $\lambda \in \mathbb{K}$,

$$\begin{aligned}\langle v, L(\lambda w_1 + w_2) \rangle &= f(v, \lambda w_1 + w_2) = \bar{\lambda} f(v, w_1) + f(v, w_2) \\ &= \bar{\lambda} \langle v, L(w_1) \rangle + \langle v, L(w_2) \rangle \\ &= \langle v, \lambda L(w_1) + L(w_2) \rangle\end{aligned}$$

para todo $v \in V$.

Seja $T_f = L^*$. Então,

$$f(v, w) = \langle v, L(w) \rangle = \langle L^*(v), w \rangle = \langle T_f(v), w \rangle,$$

para todos $v, w \in V$.

Suponhamos que existe um outro operador T' tal que

$$f(v, w) = \langle T'(v), w \rangle$$

para todos $v, w \in V$. Então,

$$\langle T_f(v), w \rangle = \langle T'(v), w \rangle,$$

ou seja,

$$\langle T_f(v) - T'(v), w \rangle = 0,$$

para todos $v, w \in V$. Logo, $T'(v) = T_f(v)$ para todo $v \in V$.

Vamos, agora, provar que a função $f \mapsto T_f$ é um isomorfismo linear.

Sejam f e g formas sobre V e $\lambda \in \mathbb{K}$.

Então,

$$(\lambda f + g)(v, w) = \langle T_{\lambda f + g}(v), w \rangle,$$

e

$$\begin{aligned}(\lambda f + g)(v, w) &= \lambda f(v, w) + g(v, w) = \lambda \langle T_f(v), w \rangle + \langle T_g(v), w \rangle \\ &= \langle \lambda T_f(v) + T_g(v), w \rangle,\end{aligned}$$

para todos $v, w \in V$. Logo, $T_{\lambda f + g} = \lambda T_f + T_g$.

Além disso, se $T \in \mathcal{L}(V, V)$,

$$f(v, w) = \langle T(v), w \rangle$$

é uma forma sobre V tal que $T_f = T$. Logo, a aplicação linear $f \mapsto T_f$ é sobrejetora.

Temos, também, que $f = \mathcal{O}$ se, e somente se, $\langle T_f(v), w \rangle = 0$, para todos $v, w \in V$. Logo, $f = \mathcal{O}$ se, e somente se, $T_f = \mathcal{O}$.

Assim, a aplicação $f \mapsto T_f$ é um isomorfismo linear. ■

Observação 6.3

Sabemos que a função

$$\begin{aligned} \mathbb{K}^{n \times n} \times \mathbb{K}^{n \times n} &\longrightarrow \mathbb{K} \\ (A, B) &\longmapsto \text{Tr}(AB^*) \end{aligned}$$

define um produto interno sobre $\mathbb{K}^{n \times n}$, e que, se $\mathcal{B}$ é uma base ortonormal de V ,

$$\begin{aligned} \mathcal{L}(V, V) &\longrightarrow \mathbb{K}^{n \times n} \\ T &\longmapsto [T]_{\mathcal{B}} \end{aligned}$$

é um isomorfismo linear. Então,

$$\langle T, U \rangle = \text{Tr}([T]_{\mathcal{B}} [U]_{\mathcal{B}}^*) = \text{Tr}([T U^*]_{\mathcal{B}}),$$

é um produto interno sobre $\mathcal{L}(V, V)$, que independe da base ortonormal $\mathcal{B}$, pois se $\mathcal{B}'$ é outra base ortonormal de V , existe $P \in \mathbb{K}^{n \times n}$ unitária tal que $[T]_{\mathcal{B}'} = P^{-1} [T]_{\mathcal{B}} P$ para todo $T \in \mathcal{L}(V, V)$. Logo,

$$\begin{aligned} \text{Tr}([T U^*]_{\mathcal{B}'}) &= \text{Tr}([T]_{\mathcal{B}'} [U]_{\mathcal{B}'}^*) = \text{Tr}(P^{-1} [T]_{\mathcal{B}} P P^* [U]_{\mathcal{B}}^* (P^{-1})^*) \\ &= \text{Tr}(P^{-1} [T U^*]_{\mathcal{B}} P) = \text{Tr}([T U^*]_{\mathcal{B}}), \end{aligned}$$

já que $P^* = P^{-1}$.

Assim, como $f \mapsto T_f$ é um isomorfismo do espaço das formas sobre $\mathcal{L}(V, V)$, temos que

$$\langle f, g \rangle = \text{Tr}(T_f T_g^*),$$

é um produto interno sobre o espaço das formas sobre V .

Além disso, se $\mathcal{B}$ é uma base ortonormal de V , então $A_{ij} = \langle T_f(v_j), v_i \rangle$ e $B_{ij} = \langle T_g(v_j), v_i \rangle$, onde $A = [T_f]_{\mathcal{B}}$ e $B = [T_g]_{\mathcal{B}}$.

Logo,

$$\begin{aligned} \langle f, g \rangle &= \text{Tr}(T_f T_g^*) = \text{Tr}(AB^*) = \sum_{i,j=1}^n A_{ij} \overline{B_{ij}} \\ &= \sum_{i,j=1}^n \langle T_f(v_j), v_i \rangle \overline{\langle T_g(v_j), v_i \rangle} = \sum_{i,j=1}^n f(v_j, v_i) \overline{g(v_j, v_i)} \end{aligned}$$

Definição 6.2

Sejam f uma forma sobre V e $\mathcal{B}$ uma base de V . A matriz A com entradas

$$A_{ij} = f(v_j, v_i)$$

é denominada a **matriz de f em relação à base ordenada $\mathcal{B}$** .

Observação 6.4

Se $\mathcal{B}$ é uma base ortogonal, $[T_f]_{\mathcal{B}} = A$, mas para uma base qualquer, isso nem sempre é verdade.

Observação 6.5

Se A é a matriz de f em relação à base $\mathcal{B} = \{v_1, \dots, v_n\}$, então

$$\begin{aligned} f(v, w) &= f\left(\sum_{k=1}^n x_k v_k, \sum_{j=1}^n y_j v_j\right) = \sum_{k,j=1}^n x_k f(v_k, v_j) \overline{y_j} \\ &= \sum_{j,k=1}^n \overline{y_j} A_{jk} x_k = Y^* A X = \langle A X, Y \rangle_C \end{aligned}$$

onde $X = [v]_{\mathcal{B}}$, $Y = [w]_{\mathcal{B}}$ e $\langle \cdot, \cdot \rangle_C$ é o produto interno canônico de $\mathbb{K}^{n \times 1}$.

Observação 6.6

Se $\mathcal{B}' = \{v'_1, \dots, v'_n\}$ é outra base de V e

$$v'_j = \sum_{i=1}^n P_{ij} v_i, \quad i, j = 1, \dots, n.$$

então a matriz A' de f na base $\mathcal{B}'$ é dada por **$A' = P^* A P$** .

De fato,

$$\begin{aligned} A'_{jk} &= f(v'_k, v'_j) = f\left(\sum_{s=1}^n P_{sk} v_s, \sum_{r=1}^n P_{rj} v_r\right) \\ &= \sum_{s,r=1}^n P_{sk} \overline{P_{rj}} f(v_s, v_r) = \sum_{s,r=1}^n \overline{P_{rj}} A_{rs} P_{sk} \\ &= (P^* A P)_{jk}, \quad j, k = 1, \dots, n. \end{aligned}$$

Teorema 6.2

Seja T um operador linear sobre um espaço vetorial complexo V de dimensão finita com produto interno. Então V possui uma base ortonormal $\mathcal{B}$ tal que $[T]_{\mathcal{B}}$ é triangular superior.

Prova.

Vamos provar o resultado por indução sobre a dimensão n de V .

Se $n = 1$, não há nada a provar. Suponhamos que o resultado seja válido para operadores sobre espaços de dimensão menor ou igual a $n - 1$.

Seja $T : V \longrightarrow V$ um operador linear, onde $\dim(V) = n$.

Como V é complexo, o operador T^* possui um autovetor v não-nulo. Seja W o subespaço de dimensão 1 gerado por v . Como $T^*(W) \subset W$ temos, pela proposição 4.4, que $T(W^\perp) \subset W^\perp$, ou seja, $W^\perp$ é T -invariante.

Então, pela hipótese de indução, $\dim(W^\perp) = n - 1$, $W^\perp$ possui uma base $\mathcal{B}' = \{v_2, \dots, v_n\}$ ortonormal tal que $[T|_{W^\perp}]_{\mathcal{B}'}$ é triangular superior.

Logo, $\mathcal{B} = \{v_1, \dots, v_n\}$ é uma base ortonormal de V tal que $[T]_{\mathcal{B}}$ é triangular superior, pois

$$[T]_{\mathcal{B}} = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & A \end{array} \right),$$

onde $A = [T|_{W^\perp}]_{\mathcal{B}'}$. ■

Corolário 6.1

Seja f uma forma sobre um espaço vetorial complexo V de dimensão finita com produto interno. Então, V possui uma base ortonormal $\mathcal{B}$ tal que a matriz de f em relação à base $\mathcal{B}$ é triangular superior.

Prova.

Seja $T \in \mathcal{L}(V, V)$ tal que $f(v, w) = \langle T(v), w \rangle$ para $v, w \in V$.

Pelo teorema anterior, existe uma base $\mathcal{B}$ ortonormal tal que $[T]_{\mathcal{B}}$ é triangular superior. Mas, como $\mathcal{B}$ é ortonormal, $[T]_{\mathcal{B}}$ é a matriz A da forma f em relação à base $\mathcal{B}$.

Logo, A é triangular superior. ■

Definição 6.3

Seja V um espaço vetorial real ou complexo. Uma forma f sobre V é dita **hermitiana** se

$$f(v, w) = \overline{f(w, v)},$$

para todos $v, w \in V$.

Observação 6.7

Se T é um operador linear sobre um espaço vetorial V de dimensão finita com produto interno e f é a forma

$$f(v, w) = \langle T(v), w \rangle,$$

então, f é hermitiana se, e somente se, T é auto-adjunto.

De fato,

$$f(v, w) = \overline{f(w, v)} \iff \langle T(v), w \rangle = \overline{\langle T(w), v \rangle} = \langle v, T(w) \rangle,$$

para todos $v, w \in V$. Logo, $T^* = T$.

Observação 6.8

Se f é hermitiana, então $f(v, v) \in \mathbb{R}$ para todo $v \in V$.

Teorema 6.3

Seja V um espaço vetorial **complexo** e seja f uma forma sobre V tal que $f(v, v) \in \mathbb{R}$ para todo $v \in V$. Então, f é hermitiana.

Prova.

Sejam $v, w \in V$. Como

$$f(v + w, v + w) = f(v, v) + f(v, w) + f(w, v) + f(w, w)$$

e

$$f(v + iw, v + iw) = f(v, v) + if(w, v) - if(v, w) + f(w, w),$$

temos que

$$f(v, w) + f(w, v) \in \mathbb{R} \quad \text{e} \quad if(w, v) - if(v, w) \in \mathbb{R},$$

pois $f(v + w, v + w)$, $f(v, v)$, $f(w, w)$ e $f(v + iw, v + iw)$ são, por hipótese, números reais.

Então,

$$f(v, w) + f(w, v) = \overline{f(v, w)} + \overline{f(w, v)}$$

e

$$if(w, v) - if(v, w) = -i \overline{f(w, v)} + i \overline{f(v, w)}.$$

Multiplicando a segunda identidade por i , obtemos que

$$\begin{aligned} f(v, w) + f(w, v) &= \overline{f(v, w)} + \overline{f(w, v)} \\ -f(w, v) + f(v, w) &= \overline{f(w, v)} - \overline{f(v, w)}. \end{aligned}$$

Logo, $2f(v, w) = 2\overline{f(w, v)}$, ou seja, $f(v, w) = \overline{f(w, v)}$. ■

Corolário 6.2

Seja T um operador linear sobre um espaço vetorial **complexo** V de dimensão finita com produto interno.

Então, T é auto-adjunto se, e somente se, $\langle T(v), v \rangle \in \mathbb{R}$, para todo $v \in V$.

Teorema 6.4

Seja V um espaço vetorial de dimensão finita com produto interno e seja f uma forma hermitiana sobre V . Então V possui uma base ortonormal $\mathcal{B}$ em relação à qual f é representada por uma matriz diagonal com entradas reais.

Teorema do Eixo Principal

Prova.

Seja T o operador linear tal que $f(v, w) = \langle T(v), w \rangle$ para todos $v, w \in V$. Como f é hermitiana, T é um operador auto-adjunto.

Então, pelo teorema 4.1, V possui uma base ortonormal $\mathcal{B} = \{v_1, \dots, v_n\}$ formada por autovetores de T .

Seja $T(v_i) = c_i v_i$, $i = 1, \dots, n$. Como T é auto-adjunto, seus autovalores c_i , $i = 1, \dots, n$, são todos reais.

Logo,

$$A_{ij} = f(v_j, v_i) = \langle T(v_j), v_i \rangle = c_j \delta_{ij},$$

onde A é a matriz de f com respeito à base $\mathcal{B}$.

Assim, A é diagonal com entradas reais. ■

Observação 6.9

Nas condições acima,

$$f(v, w) = Y^* A X = \sum_{j,k=1}^n \overline{y_k} A_{kj} x_j = \sum_{j=1}^n c_j x_j \overline{y_j},$$

e

$$f(v, v) = \sum_{j=1}^n c_j |x_j|^2,$$

onde $X = [v]_{\mathcal{B}}$ e $Y = [w]_{\mathcal{B}}$.

7. Formas Positivas

Definição 7.1

Uma forma f sobre um espaço vetorial V real ou complexo é dita

- **não-negativa**, se f é hermitiana e $f(v, v) \geq 0$, para todo $v \in V$.
- **positiva**, se f é hermitiana e $f(v, v) > 0$, para todo $v \in V - \{0\}$.

Observação 7.1

Uma forma positiva sobre V é um produto interno sobre V .

Observação 7.2

Seja f uma forma sobre o espaço V de dimensão finita. Seja $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base de V e seja A a matriz de f em relação à base $\mathcal{B}$, ou seja,

$$A_{jk} = f(v_k, v_j).$$

Se $v = x_1 v_1 + \dots + x_n v_n$, então

$$f(v, v) = \sum_{k,j=1}^n A_{kj} x_j \overline{x_k} = X^* A X.$$

Assim, f é não-negativa se, e somente se, $A = A^*$ e $\sum_{j,k=1}^n A_{kj} x_j \overline{x_k} \geq 0$, para todo $(x_1, \dots, x_n) \in \mathbb{K}^n$.

De fato, f é hermitiana se, e só se, $f(v, w) = \overline{f(w, v)}$ para todo $v, w \in V$, ou seja, se, e só se, $\langle AX, Y \rangle_C = \overline{\langle AY, X \rangle_C} = \langle X, AY \rangle_C$, para todos $X, Y \in \mathbb{K}^{n \times 1}$, onde $\langle \cdot, \cdot \rangle_C$ é o produto interno canônico de $\mathbb{K}^{n \times 1}$.

Logo, f é hermitiana se, e só se, $A = A^*$.

E f é positiva se, e só se, $A = A^*$ e

$$\sum_{j,k=1}^n A_{kj} x_j \overline{x_k} > 0, \quad \forall (x_1, \dots, x_n) \in \mathbb{K}^n - \{0\},$$

ou seja, se, e só se, $g(X, Y) = Y^* A X$ é um produto interno sobre $\mathbb{K}^{n \times 1}$.

Mas, para $\mathbb{K} = \mathbb{C}$, temos que:

- f é não-negativa $\iff \sum_{j,k} A_{kj} x_j \overline{x_k} \geq 0$ para todo $(x_1, \dots, x_n) \in \mathbb{C}^n$.

- f é positiva $\iff \sum_{j,k} A_{kj} x_j \overline{x_k} > 0$ para todo $(x_1, \dots, x_n) \in \mathbb{C}^n - \{\mathbf{0}\}$.

De fato, pelo teorema 6.3, se $\sum_{j,k} A_{kj} x_j \overline{x_k} \in \mathbb{R}$ para todo $(x_1, \dots, x_n) \in \mathbb{C}^n$, então f é hermitiana.

Teorema 7.1

Seja A uma matriz $n \times n$ sobre $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$). A função g definida por

$$g(X, Y) = Y^* A X$$

é uma forma positiva sobre $\mathbb{K}^{n \times 1}$ se, e só se, existe uma matriz $P \in \mathbb{K}^{n \times n}$ invertível, tal que $A = P^* P$.

Prova.

Suponhamos que $A = P^* P$, sendo P invertível.

Como $A^* = A$, temos que g é hermitiana. Além disso,

$$g(X, X) = X^* A X = X^* P^* P X = (P X)^* P X > 0,$$

para todo $X \in \mathbb{K}^{n \times 1} - \{\mathbf{0}\}$, já que $P X = \mathbf{0}$ se, e só se, $X = \mathbf{0}$.

Suponhamos, agora, que g é uma forma positiva sobre $\mathbb{K}^{n \times 1}$, ou seja, g é um produto interno sobre $\mathbb{K}^{n \times 1}$.

Então, $\mathbb{K}^{n \times 1}$ possui uma base $\{v_1, \dots, v_n\}$ ortonormal em relação ao produto interno g .

Seja Q a matriz $n \times n$ cuja j -ésima coluna é v_j . Como $\mathcal{B}$ é uma base ortonormal, temos que

$$\delta_{jk} = g(v_j, v_k) = v_k^* A v_j,$$

ou seja, $Q^* A Q = I$. Além disso, Q é invertível, pois $\mathcal{B}$ é uma base de $\mathbb{K}^{n \times 1}$.

Fazendo $P = Q^{-1}$, temos que $A = (Q^*)^{-1} Q^{-1} = (Q^{-1})^* Q^{-1} = P^* P$, sendo P invertível. ■

Observação 7.3

Se g é positiva, então $\det(A) > 0$, pois

$$\det(A) = \det(P^* P) = \det(P^*) \det(P) = \overline{\det(P)} \det(P) = |\det(P)|^2 > 0.$$

A condição de que $\det(A) > 0$ é necessária, mas não suficiente, como veremos abaixo.

Definição 7.2

Seja A uma matriz $n \times n$ sobre o corpo $\mathbb{K}$. Os *menores principais* de A são os escalares

$$\Delta_k(A) = \det \begin{pmatrix} A_{11} & \cdots & A_{1k} \\ \vdots & \ddots & \vdots \\ A_{k1} & \cdots & A_{kk} \end{pmatrix}, \quad k = 1, \dots, n.$$

Lema 7.1

Seja A uma matriz $n \times n$ invertível sobre um corpo $\mathbb{K}$. As seguintes afirmações são equivalentes:

- (a) Existe uma matriz triangular superior P , com $P_{kk} = 1$, $1 \leq k \leq n$, tal que a matriz $B = AP$ é triangular inferior.
- (b) Os menores principais de A são todos diferentes de zero.

Prova.

(b) $\implies$ (a) Seja $P \in \mathbb{K}^{n \times n}$ e façamos $B = AP$. Então,

$$B_{jk} = \sum_{r=1}^n A_{jr} P_{rk}, \quad j, k = 1, \dots, n.$$

Se P é triangular superior, ou seja, $P_{rk} = 0$, se $r > k$, e $P_{kk} = 1$, para todo $k = 1, \dots, n$, então

$$\sum_{r=1}^{k-1} A_{jr} P_{rk} + A_{jk} = B_{jk}, \quad \text{se } k \geq 2.$$

Assim, B será triangular inferior se, e só se, $B_{jk} = 0$, $j < k$, ou seja, se, e só se,

$$\sum_{r=1}^{k-1} A_{jr} P_{rk} = -A_{jk}, \quad 1 \leq j \leq k-1 \quad \text{e} \quad 2 \leq k \leq n. \quad (*)$$

Então, a afirmação (a) do lema é equivalente à afirmação de que existem escalares P_{rk} , $1 \leq r \leq k$ e $1 \leq k \leq n$, que satisfazem (*) e $P_{kk} = 1$, $1 \leq k \leq n$.

Para cada $2 \leq k \leq n$, (*) é um sistema de $k-1$ equações lineares nas incógnitas $P_{1,k}, \dots, P_{k-1,k}$. A matriz dos coeficientes do sistema é

$$\begin{pmatrix} A_{11} & \cdots & A_{1,k-1} \\ \vdots & \ddots & \vdots \\ A_{k-1,1} & \cdots & A_{k-1,k-1} \end{pmatrix},$$

cujo determinante é o menor principal $\Delta_{k-1}(A)$.

Como $\Delta_{k-1}(A) \neq 0$, para todo $2 \leq k \leq n+1$, os sistemas possuem soluções únicas. Mostramos, assim, que (b) $\implies$ (a) e que a matriz P é única.

(a) $\implies$ (b) Suponhamos que $B = AP$, onde P é triangular superior, com $P_{kk} = 1$, $k = 1, \dots, n$, e B é triangular inferior.

Afirmção: $\Delta_k(A) = \Delta_k(B) = B_{11} \dots B_{kk}$, $k = 1, \dots, n$.

Sejam $A_1, \dots, A_n$ e $B_1, \dots, B_n$ as colunas de A e B , respectivamente. Então, $B_1 = A_1$ e

$$B_r = \sum_{j=1}^{r-1} P_{jr} A_j + A_r, \quad 2 \leq r \leq n \quad (I)$$

Seja $k \in \{1, \dots, n\}$ fixo. Então, por (I), a r -ésima coluna da matriz

$$\begin{pmatrix} B_{11} & \cdots & B_{1k} \\ \vdots & \ddots & \vdots \\ B_{k1} & \cdots & B_{kk} \end{pmatrix}$$

é obtida, adicionando-se à r -ésima coluna da matriz

$$\begin{pmatrix} A_{11} & \cdots & A_{1k} \\ \vdots & \ddots & \vdots \\ A_{k1} & \cdots & A_{kk} \end{pmatrix}$$

uma combinação linear de suas outras colunas. Então, $\Delta_k(A) = \Delta_k(B)$.

Como A e P são invertíveis, temos que B é invertível e, portanto, $\Delta_n(B) = B_{11} \dots B_{nn} \neq 0$.

Logo, $\Delta_k(A) = \Delta_k(B) = B_{11} \dots B_{kk} \neq 0$, $k = 1, \dots, n$. ■

Teorema 7.2

Seja f uma forma sobre um espaço vetorial V de dimensão finita e seja A a matriz de f em relação a uma base $\mathcal{B}$ de V .

Então, f é positiva se, e só se, $A^* = A$ e $\Delta_k(A) > 0$ para todo $k = 1, \dots, n$.

Prova.

• Suponhamos que $A = A^*$ e $\Delta_k(A) > 0$, $k = 1, \dots, n$. Pelo lema anterior, existe uma (única) matriz P , com $P_{kk} = 1$, $k = 1, \dots, n$, triangular superior tal que $B = AP$ é triangular inferior.

Logo, P^* é triangular inferior e, portanto, $P^*B = P^*AP$ também é triangular inferior.

Como A é auto-adjunta, a matriz $D = P^*AP$ é auto-adjunta.

Logo, $D = P^*B = P^*AP$ é diagonal, pois D é triangular e auto-adjunta.

Além disso, $\Delta_k(D) = \Delta_k(P^*B) = \Delta_k(B)$, $k = 1, \dots, n$.

De fato, se $B^1, \dots, B^n$ e $D^1, \dots, D^n$ são as linhas de B e D , respectivamente, então $D^1 = B^1$ e

$$D^r = \sum_{j=1}^{r-1} P_{rj}^* B^j + B^r, \quad r = 2, \dots, n.$$

Assim, a r -ésima linha da matriz

$$\begin{pmatrix} D_{11} & \cdots & D_{1k} \\ \vdots & \ddots & \vdots \\ D_{k1} & \cdots & D_{kk} \end{pmatrix}$$

é obtida, adicionando-se à r -ésima linha da matriz

$$\begin{pmatrix} B_{11} & \cdots & B_{1k} \\ \vdots & \ddots & \vdots \\ B_{k1} & \cdots & B_{kk} \end{pmatrix}$$

uma combinação linear de suas outras linhas.

Logo, $\Delta_k(D) = \Delta_k(B)$, $k = 1, \dots, n$.

Como $\Delta_k(B) = \Delta_k(A) > 0$, $k = 1, \dots, n$ e D é diagonal, temos que

$$\Delta_k(D) = D_{11} \dots D_{kk} > 0, \quad \forall k = 1, \dots, n.$$

Logo, $D_{kk} > 0$, para todo $k = 1, \dots, n$.

Seja $\mathcal{B} = \{v_1, \dots, v_n\}$.

Como A é a matriz de f com respeito à base $\mathcal{B}$, então, $D = P^*AP$ é a matriz de f com respeito à base $\mathcal{B}' = \{v'_1, \dots, v'_n\}$, onde

$$v'_j = \sum_{i=1}^n P_{ij} v_i.$$

Logo,

$$f(v, v) = X^*DX = \sum_{k=1}^n D_{kk}|x_k|^2 > 0,$$

onde $X = [v]_{\mathcal{B}'} \in \mathbb{K}^{n \times 1} - \{\mathbf{0}\}$. Ou seja, f é uma forma positiva.

- Suponhamos, agora, que f é uma forma positiva.

Seja $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base de V e seja A a matriz de f na base $\mathcal{B}$. Já sabemos que $A^* = A$.

Para cada $k = 1, \dots, n$, seja V_k o subespaço de V gerado pelos vetores $v_1, \dots, v_k$ e seja f_k a forma $f|_{V_k \times V_k}$.

Como f_k é positiva e, em relação à base $\{v_1, \dots, v_k\}$ de V_k , f_k é representada pela matriz

$$A_k = \begin{pmatrix} A_{11} & \cdots & A_{1k} \\ \vdots & \ddots & \vdots \\ A_{k1} & \cdots & A_{kk} \end{pmatrix},$$

temos que $\Delta_k(A) = \det(A_k) > 0$, $k = 1, \dots, n$, pois já provamos, anteriormente, que se uma forma é positiva, o determinante de qualquer matriz que a represente é positivo. ■

Observação 7.4

Seja $f : V \times V \longrightarrow \mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) uma forma. Já provamos que:

- se $\mathbb{K} = \mathbb{C}$, f é positiva $\iff A = A^*$ e $X^*AX > 0$, $\forall X \in \mathbb{C}^{n \times 1} - \{\mathbf{0}\} \iff X^*AX > 0$, $\forall X \in \mathbb{C}^{n \times 1} - \{\mathbf{0}\}$.
- se $\mathbb{K} = \mathbb{R}$, f é positiva $\iff A = A^t$ e $X^tAX > 0$, $\forall X \in \mathbb{R}^{n \times 1} - \{\mathbf{0}\}$.

A hipótese de que $A = A^t$ no caso real é necessária, pois uma matriz real que satisfaz $X^tAX > 0$, $\forall X \in \mathbb{R}^{n \times 1} - \{\mathbf{0}\}$ não é necessariamente simétrica.

Exemplo 7.1

Seja $f : \mathbb{R}^{2 \times 1} \longrightarrow \mathbb{R}^{2 \times 1}$ a forma definida por $f(X, Y) = \langle AX, Y \rangle$, onde

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

Então, para todo $X = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{2 \times 1} - \{\mathbf{0}\}$, temos

$$\begin{aligned} f(X, X) &= X^*AX = (x, y) \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (x, y) \begin{pmatrix} x + 2y \\ -2x + y \end{pmatrix} \\ &= x^2 + 2yx - 2xy + y^2 = x^2 + y^2 > 0. \end{aligned}$$

Mas $A^t \neq A$. □

Observação 7.5

Um fato verdadeiro é que, se $A \in \mathbb{R}^{n \times n}$, então:

$$A = A^t \text{ e } X^t A X > 0, \forall X \in \mathbb{R}^{n \times 1} - \{\mathbf{0}\} \iff X^* A X > 0, \forall X \in \mathbb{C}^{n \times 1} - \{\mathbf{0}\}.$$

Com efeito:

- $X^* A X > 0, \forall X \in \mathbb{C}^{n \times 1} - \{\mathbf{0}\} \implies A = A^* = A^t \text{ e } X^t A X > 0, \forall X \in \mathbb{R}^{n \times 1} - \{\mathbf{0}\}.$
- $(X + iY)^* A (X + iY) = (X^t - iY^t) A (X + iY)$
 $= X^t A X + Y^t A Y + iX^t A Y - iY^t A X$
 $= X^t A X + Y^t A Y + i(X^t A Y - Y^t A X)$
 $= X^t A X + Y^t A Y > 0, \text{ se } X + iY \neq \mathbf{0}, X, Y \in \mathbb{R}^{n \times 1},$

pois $Y^t A X = Y^t A^t X = (X^t A Y)^t = X^t A Y$.

Logo, uma forma $f : V \times V \longrightarrow \mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) é positiva se, e só se, $X^* A X > 0$, para todo $X \in \mathbb{C}^{n \times 1} - \{\mathbf{0}\}$, onde A é a matriz de f com respeito a alguma base de V .

Observação 7.6

Seja f uma forma positiva sobre um espaço V de dimensão finita **com produto interno** e seja T o único operador sobre V tal que $f(v, w) = \langle T(v), w \rangle$ para todos $v, w \in W$.

Então, f é uma forma positiva se, e só se, T é um operador positivo, ou seja, $T = T^*$ e $\langle T(v), v \rangle > 0$ para todo $v \in V$.

Existe, portanto, uma correspondência biunívoca entre os operadores positivos e as formas positivas.

8. Resumo sobre matrizes positivas

- Se $A \in \mathbb{C}^{n \times n}$, as seguintes afirmações são equivalentes:

$$(1) \sum_{j,k=1}^n A_{kj} x_j \overline{x_k} > 0, \forall X \in \mathbb{C}^{n \times 1} - \{\mathbf{0}\}.$$

$$(2) [X, Y] = Y^* A X \text{ é um produto interno sobre } \mathbb{C}^{n \times 1}.$$

(3) Em relação ao produto interno canônico $\langle X, Y \rangle = Y^*X$ de $\mathbb{C}^{n \times 1}$, o operador $X \mapsto AX$ sobre $\mathbb{C}^{n \times 1}$ é positivo.

(4) Existe uma matriz $P \in \mathbb{C}^{n \times n}$ invertível tal que $A = P^*P$.

(5) $A = A^*$ e os menores principais de A são positivos.

• Se $A \in \mathbb{R}^{n \times n}$, as afirmações acima também são equivalentes às seguintes:

(6) $A = A^t$ e $X^tAX > 0$, $\forall X \in \mathbb{R}^{n \times 1} - \{0\}$.

(7) $(X, Y) = X^tAY$ é um produto interno sobre $\mathbb{R}^{n \times 1}$.

(8) Em relação ao produto interno canônico $\langle X, Y \rangle = Y^tX$ de $\mathbb{R}^{n \times 1}$, o operador $X \mapsto AX$ sobre $\mathbb{R}^{n \times 1}$ é positivo.

(9) Existe uma matriz $P \in \mathbb{R}^{n \times n}$ invertível tal que $A = P^tP$.

Exemplo 8.1

A forma $f: \mathbb{C}^3 \times \mathbb{C}^3 \rightarrow \mathbb{C}$ definida por

$$f((z_1, z_2, z_3), (w_1, w_2, w_3)) = z_1 \overline{w_1} - iz_2 \overline{w_1} + iz_1 \overline{w_2} + 2z_2 \overline{w_2} + 2z_3 \overline{w_3},$$

é positiva.

De fato, se $\mathcal{B} = \{e_1, e_2, e_3\}$ é a base canônica de $\mathbb{C}^3$, a matriz A de f em relação a $\mathcal{B}$ tem entradas:

$$\begin{aligned} A_{11} &= f(e_1, e_1) = 1 & ; & & A_{12} &= f(e_2, e_1) = -i & ; & & A_{13} &= f(e_3, e_1) = 0 & ; \\ A_{21} &= f(e_1, e_2) = i & ; & & A_{22} &= f(e_2, e_2) = 2 & ; & & A_{23} &= f(e_3, e_2) = 0 & ; \\ A_{31} &= f(e_1, e_3) = 0 & ; & & A_{32} &= f(e_2, e_3) = 0 & ; & & A_{33} &= f(e_3, e_3) = 2 & . \end{aligned}$$

Assim,

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

Então, $A^* = A$, $\Delta_1(A) = 1$, $\Delta_2(A) = 2 - 1 = 1 > 0$ e $\Delta_3(A) = 2 > 0$.

Logo, por (5), f é uma forma positiva. $\square$

9. Funções de operadores

De maneira análoga ao que foi feito para obter a raiz quadrada não-negativa de um operador não-negativo, podemos também definir, sob certas condições, $f(T)$, onde f é uma função definida sobre um subconjunto do corpo $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) e T é um operador linear.

Definição 9.1

Seja T um operador diagonalizável sobre um espaço V de dimensão finita. O **espectro de T** é o conjunto $S = \{c_1, \dots, c_k\}$, cujos elementos são os autovalores distintos de T .

Definição 9.2

Seja T um operador normal diagonalizável sobre um espaço V de dimensão finita com produto interno e seja

$$T = c_1\pi_1 + \dots + c_k\pi_k,$$

sua resolução espectral. Seja f uma função com valores no corpo dos escalares, cujo domínio contém o espectro de T . Então o operador $f(T)$ é definido por:

$$f(T) = f(c_1)\pi_1 + \dots + f(c_k)\pi_k.$$

Teorema 9.1

Seja T um operador normal diagonalizável com espectro S sobre um espaço V de dimensão finita com produto interno. Suponhamos que f seja uma função com valores no corpo dos escalares, cujo domínio contém S . Então, $f(T)$ é um operador normal diagonalizável com espectro $f(S)$. Se U é uma aplicação unitária de V sobre V' e $T' = UTU^{-1}$, então S é o espectro de T' e

$$f(T') = Uf(T)U^{-1}.$$

Prova.

Seja $T = \sum_{j=1}^k c_j\pi_j$ a resolução espectral de T e seja

$$f(T) = \sum_{j=1}^k f(c_j)\pi_j.$$

Como

$$f(T)^* = \sum_{j=1}^k \overline{f(c_j)} \pi_j^* = \sum_{j=1}^k f(\overline{c_j}) \pi_j,$$

já que π_j é auto-adjunta, para $j = 1, \dots, k$, temos que

$$f(T)^* f(T) = f(T) f(T)^*,$$

pois cada π_j é um polinômio em T . Logo, $f(T)$ é um operador normal.

Afirmção: $f(S)$ é o espectro de $f(T)$.

De fato, se $v \in \pi_j(V) - \{0\}$, então $f(T)(v) = f(c_j)(v)$, ou seja, $f(c_j)$ é um autovalor de $f(T)$.

Reciprocamente, suponhamos que $v \neq 0$ e

$$f(T)(v) = bv.$$

Como $v = \pi_1(v) + \dots + \pi_k(v)$, temos que

$$f(T)(v) = \sum_{j=1}^k f(T)\pi_j(v) = \sum_{j=1}^k f(c_j)\pi_j(v) = \sum_{j=1}^k b\pi_j(v).$$

Logo,

$$0 = \left\| \sum_{j=1}^k (f(c_j) - b)\pi_j(v) \right\|^2 = \sum_{j=1}^k |f(c_j) - b|^2 \|\pi_j(v)\|^2,$$

já que $\pi_j(v) \perp \pi_i(v)$, se $i \neq j$. Como $v \neq 0$, existe $i \in \{1, \dots, k\}$ tal que $\pi_i(v) \neq 0$. Então, $f(c_i) = b$, já que $|f(c_j) - b|^2 \|\pi_j(v)\|^2 = 0$, $j = 1, \dots, k$.

Provamos, assim, que $f(S)$ é o espectro de $f(T)$.

Seja $f(T) = \{b_1, \dots, b_r\}$, com $b_m \neq b_n$, se $m \neq n$. Seja J_m o conjunto dos índices $i \in \{1, \dots, k\}$, tais que $f(c_i) = b_m$ e seja

$$P_m = \sum_{j \in J_m} \pi_j, \quad m = 1, \dots, r.$$

Então:

- $P_m^2 = P_m$, ou seja, P_m é uma projeção.
- $P_m P_n = 0$, se $m \neq n$.
- $I = P_1 + \dots + P_r$.
- $f(T) = \sum_{m=1}^r b_m P_m$, onde $\{b_1, \dots, b_r\}$ é o espectro de $f(T)$.

- $P_m(V)$ é o autoespaço de $f(T)$ associado ao autovalor b_m .
- $P_m(V) \perp P_n(V)$, se $m \neq n$.
- P_m é a projeção ortogonal de V sobre $P_m(V)$.
- $V = P_1(V) \oplus \dots \oplus P_r(V)$.

Ou seja,

$$f(T) = \sum_{m=1}^r b_m P_m,$$

é a resolução espectral do operador $f(T)$.

Suponhamos, agora, que U é uma transformação unitária de V sobre V' e que $T' = UTU^{-1}$.

Sendo $T' = UTU^{-1}$ e $\pi'_j = U\pi_j U^{-1}$, $j = 1, \dots, k$, temos que

$$T' = \sum_{j=1}^k c_j U\pi_j U^{-1} = \sum_{j=1}^k c_j \pi'_j,$$

é a resolução espectral de T' , pois:

- $c_1, \dots, c_k$ são escalares distintos.
- $\pi'_j \pi'_i = U\pi_j U^{-1} U\pi_i U^{-1} = U\pi_j \pi_i U^{-1} = \mathcal{O}$, se $i \neq j$.
- $\pi'_j \pi'_j = U\pi_j U^{-1} U\pi_j U^{-1} = U\pi_j \pi_j U^{-1} = \pi_j$, $j = 1, \dots, k$.
- π'_j é um operador não-nulo, $j = 1, \dots, k$.
- $\pi'_1 + \dots + \pi'_k = U\pi_1 U^{-1} + \dots + U\pi_k U^{-1} = U(\pi_1 + \dots + \pi_k) U^{-1} = UIU^{-1} = I$.

Logo, $\{c_1, \dots, c_k\}$ é o espectro de T' e $\pi'_j(V') = U\pi_j U^{-1}(V') = U\pi_j(V)$ é o autoespaço de T' associado ao autovalor c_j , $j = 1, \dots, k$.

Assim,

$$f(T') = \sum_{j=1}^k f(c_j) \pi'_j = \sum_{j=1}^k f(c_j) U\pi_j U^{-1} = U \left(\sum_{j=1}^k f(c_j) \pi_j \right) U^{-1} = Uf(T)U^{-1}.$$



Corolário 9.1

Com as mesmas hipóteses do teorema 9.1, suponhamos que T seja representado, em relação à base ordenada $\mathcal{B} = \{v_1, \dots, v_n\}$, pela matriz

diagonal D com entradas $D_{ii} = d_i$, $i = 1, \dots, n$ na diagonal. Então, em relação à base $\mathcal{B}$, $f(T)$ é representado pela matriz $f(D)$ com entradas $f(d_1), \dots, f(d_n)$ na diagonal.

Além disso, se $\mathcal{B}' = \{v'_1, \dots, v'_n\}$ é outra base de V e P é a matriz tal que

$$v'_j = \sum_{i=1}^n P_{ij} v_i,$$

então $P^{-1}f(D)P$ é a matriz de $f(T)$ em relação à base $\mathcal{B}'$.

Prova.

Para cada $i \in \{1, \dots, n\}$, existe um único $j \in \{1, \dots, k\}$, tal que $v_i \in \pi_j(V)$ e $d_i = c_j$. Portanto, $f(T)(v_i) = f(d_i)(v_i)$ para todo $i = 1, \dots, n$, e

$$\begin{aligned} f(T)(v'_j) &= \sum_{i=1}^n P_{ij} f(T)(v_i) = \sum_{i=1}^n f(d_i) P_{ij} v_i = \sum_{i=1}^n (f(D)P)_{ij} v_i \\ &= \sum_{i=1}^n (f(D)P)_{ij} \sum_{k=1}^n P_{ki}^{-1} v'_k = \sum_{k=1}^n \sum_{i=1}^n P_{ki}^{-1} (f(D)P)_{ij} v'_k \\ &= \sum_{k=1}^n (P^{-1}f(D)P)_{kj} v'_k. \end{aligned}$$

Logo, $[f(T)]_{\mathcal{B}} = f(D)$ e $[f(T)]_{\mathcal{B}'} = P^{-1}f(D)P$. ■

Observação 9.1

Seja A uma matriz $n \times n$ sobre $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) normal e diagonalizável. Sejam P e Q matrizes $n \times n$ invertíveis tais que $D = PAP^{-1}$ e $D' = QAQ^{-1}$ são matrizes diagonais.

Considere o espaço $\mathbb{K}^{n \times 1}$ com o produto interno canônico e o operador $T : \mathbb{K}^{n \times 1} \rightarrow \mathbb{K}^{n \times 1}$ definido por $T(X) = AX$.

Então, T é um operador normal diagonalizável, pois $[T]_{\mathcal{B}_C} = A$ e $A^*A = AA^*$, e

- $[T]_{\mathcal{B}} = [I]_{\mathcal{B}\mathcal{B}_C} [T]_{\mathcal{B}_C} [I]_{\mathcal{B}_C\mathcal{B}} = PAP^{-1} = D$,
- $[T]_{\mathcal{B}'} = [I]_{\mathcal{B}'\mathcal{B}_C} [T]_{\mathcal{B}_C} [I]_{\mathcal{B}_C\mathcal{B}'} = QAQ^{-1} = D'$,

onde $\mathcal{B}_C$ é a base canônica de $\mathbb{K}^{n \times 1}$, $\mathcal{B}$ é a base de $\mathbb{K}^{n \times 1}$ formada pelos vetores colunas de P^{-1} e $\mathcal{B}'$ é a base de $\mathbb{K}^{n \times 1}$ formada pelos vetores colunas de Q^{-1} .

Como as entradas $d_1, \dots, d_n$ e $d'_1, \dots, d'_n$ das matrizes diagonais de D e D' , respectivamente, são os autovalores de T , temos que $\{d_1, \dots, d_n\} = \{d'_1, \dots, d'_n\}$.

Seja f uma função com valores no corpo $\mathbb{K}$, cujo domínio contém o conjunto $S = \{d_1, \dots, d_n\}$.

• Definimos $f(A) = P^{-1}f(D)P$.

A matriz $f(A)$ está bem definida, pois, pelo corolário anterior, temos $f(D) = [f(T)]_{\mathcal{B}}$, $f(D') = [f(T)]_{\mathcal{B}'}$ e

$$\begin{aligned} [f(T)]_{\mathcal{B}'} &= [I]_{\mathcal{B}'\mathcal{B}}[f(T)]_{\mathcal{B}}[I]_{\mathcal{B}\mathcal{B}'} \\ &= [I]_{\mathcal{B}'\mathcal{B}_C}[I]_{\mathcal{B}_C\mathcal{B}}f(D)[I]_{\mathcal{B}\mathcal{B}_C}[I]_{\mathcal{B}_C\mathcal{B}'} \\ &= QP^{-1}f(D)PQ^{-1}, \end{aligned}$$

ou seja, $Q^{-1}f(D')Q = P^{-1}f(D)P$.

Exemplo 9.1

Seja A uma matriz $n \times n$ complexa normal e seja P uma matriz unitária tal

que $D = PAP^* = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ é uma matriz diagonal.

Então, $e^A = P^*e^DP = P^* \begin{pmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n} \end{pmatrix} P$.

No caso particular em que $A = \mathcal{O}$ é a matriz nula, $e^A = I$, pois

$$\mathcal{O} = P \begin{pmatrix} 0 & & \\ & \ddots & \\ & & 0 \end{pmatrix} P^* \quad \text{e} \quad e^A = P^* \begin{pmatrix} e^0 & & \\ & \ddots & \\ & & e^0 \end{pmatrix} P = P^*IP = I.$$

Seja B outra matriz normal complexa $n \times n$ que comuta com A .

Considere $\mathbb{C}^{n \times 1}$ com o produto interno canônico e os operadores

$$T_A: \mathbb{C}^{n \times 1} \longrightarrow \mathbb{C}^{n \times 1} \quad \text{e} \quad T_B: \mathbb{C}^{n \times 1} \longrightarrow \mathbb{C}^{n \times 1},$$

definidos por

$$T_A(X) = AX \quad \text{e} \quad T_B(X) = BX.$$

Como $A = [T_A]_{\mathcal{B}_C}$ e $B = [T_B]_{\mathcal{B}_C}$, onde $\mathcal{B}_C$ é a base canônica de $\mathbb{C}^{n \times 1}$, temos que T_A e T_B são operadores normais que comutam.

Seja $\mathbb{C}^{n \times 1} = W_1 \oplus \dots \oplus W_k$ a decomposição primária de $\mathbb{C}^{n \times 1}$ com res-

peito ao operador T_A , ou seja W_j , $j = 1, \dots, k$, são os autoespaços de T_A associados aos seus autovalores distintos.

Seja $v \in W_j$. Então,

$$T_A(T_B(v)) = AB(v) = BA(v) = T_B(T_A(v)) = T_B(c_j v) = c_j T_B(v).$$

Logo, $T_B(W_j) \subset W_j$.

Como $T_B|_{W_j}$ é um operador normal sobre um espaço complexo, W_j possui uma base ortonormal $\mathcal{B}_j$ tal que $[T_B|_{W_j}]_{\mathcal{B}_j}$ é diagonal.

Logo, $\mathcal{B} = \{\mathcal{B}_1, \dots, \mathcal{B}_k\}$ é uma base ortonormal de $\mathbb{C}^{n \times 1}$, já que $W_j \perp W_i$, $j \neq i$, tal que $[T_A]_{\mathcal{B}}$ e $[T_B]_{\mathcal{B}}$ são matrizes diagonais.

Então,

$$D_A = [T_A]_{\mathcal{B}} = PAP^* = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \quad \text{e} \quad D_B = [T_B]_{\mathcal{B}} = PBP^* = \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix}$$

onde P^* é a matriz cujas colunas são os vetores da base $\mathcal{B}$.

Assim,

$$D_A + D_B = P(A + B)P^* = \begin{pmatrix} \lambda_1 + \mu_1 & & \\ & \ddots & \\ & & \lambda_n + \mu_n \end{pmatrix},$$

$$e^A = P^* \begin{pmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n} \end{pmatrix} P \quad \text{e} \quad e^B = P^* \begin{pmatrix} e^{\mu_1} & & \\ & \ddots & \\ & & e^{\mu_n} \end{pmatrix} P.$$

Logo,

$$\begin{aligned} \bullet e^A e^B &= P^* \begin{pmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n} \end{pmatrix} P P^* \begin{pmatrix} e^{\mu_1} & & \\ & \ddots & \\ & & e^{\mu_n} \end{pmatrix} P \\ &= P^* \begin{pmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n} \end{pmatrix} \begin{pmatrix} e^{\mu_1} & & \\ & \ddots & \\ & & e^{\mu_n} \end{pmatrix} P \\ &= P^* \begin{pmatrix} e^{\lambda_1 + \mu_1} & & \\ & \ddots & \\ & & e^{\lambda_n + \mu_n} \end{pmatrix} P \\ &= e^{A+B}. \end{aligned}$$

$$\bullet e^{-A} = P^* \begin{pmatrix} e^{-\lambda_1} & & \\ & \ddots & \\ & & e^{-\lambda_n} \end{pmatrix} P.$$

- $e^A e^{-A} = e^O = I.$

No caso particular em que $A^* = -A$, temos que A é uma matriz normal cujos autovalores são imaginários puros.

De fato, se $v \neq 0$ e $Av = \lambda v$, então $A^*v = \bar{\lambda}v$. Logo, $\bar{\lambda}v = -\lambda v$, ou seja, $\bar{\lambda} + \lambda = 0$.

Assim, $D = PAP^* = \begin{pmatrix} e^{i\theta_1} & & \\ & \ddots & \\ & & e^{i\theta_n} \end{pmatrix}$, para alguma matriz $n \times n$ invertível P , onde $i\theta_1, \dots, i\theta_n$, $\theta_j \in \mathbb{R}$, $j = 1, \dots, n$, são os autovalores de A .

Logo

$$e^A = P^* \begin{pmatrix} e^{i\theta_1} & & \\ & \ddots & \\ & & e^{i\theta_n} \end{pmatrix} P.$$

Então,

- $\det(e^A) = \det \begin{pmatrix} e^{i\theta_1} & & \\ & \ddots & \\ & & e^{i\theta_n} \end{pmatrix} = e^{i\theta_1 + \dots + i\theta_n} = e^{\text{Tr}(A)}.$

- $(e^A)^* = P^* \begin{pmatrix} e^{-i\theta_1} & & \\ & \ddots & \\ & & e^{-i\theta_n} \end{pmatrix} P = e^{-A}.$

- $e^A (e^A)^* = e^A e^{-A} = I$, ou seja A é uma matriz unitária.

No caso geral, se C é uma matriz $n \times n$ complexa normal e P é uma matriz unitária tal que

$$D = PCP^* = \begin{pmatrix} a_1 + ib_1 & & \\ & \ddots & \\ & & a_n + ib_n \end{pmatrix},$$

é uma matriz diagonal, onde $a_j, b_j \in \mathbb{R}$, $j = 1, \dots, n$, temos que

$$C = P^*DP = C_1 + C_2,$$

onde

$$C_1 = P^* \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix} P$$

é uma matriz auto-adjunta e

$$C_2 = P^* \begin{pmatrix} ib_1 & & \\ & \ddots & \\ & & ib_n \end{pmatrix} P$$

é uma matriz tal que $C_2^* = -C_2$ que comuta com C_1 .

Logo,

$$e^C = e^{C_1+C_2} = e^{C_1} e^{C_2},$$

onde

$$e^{C_1} = P^* \begin{pmatrix} e^{a_1} & & \\ & \ddots & \\ & & e^{a_n} \end{pmatrix} P$$

é uma matriz positiva e

$$e^{C_2} = P^* \begin{pmatrix} e^{b_1} & & \\ & \ddots & \\ & & e^{b_n} \end{pmatrix} P$$

é uma matriz unitária. $\square$

