Cultura Acadêmica

Anízio Perissinotto Junior
João Peres Vieira
Carina Alves

FORMAS ELEMENTARES: DIAGONAL, TRIANGULAR E DE JORDAN

CULTURA ACADÊMICA Editora

unesp

Pró-reitoria de Graduação / UNESP
prograd

FORMAS ELEMENTARES: DIAGONAL, TRIANGULAR E DE JORDAN

Cultura Acadêmica

Anízio Perissinotto Junior
João Peres Vieira
Carina Alves

FORMAS ELEMENTARES: DIAGONAL, TRIANGULAR E DE JORDAN

CULTURA ACADÊMICA
Editora

unesp

Pró-reitoria de Graduação / UNESP
prograd

São Paulo
2014

Ficha catalográfica elaborada pela Coordenadoria Geral de Bibliotecas da Unesp

P446f

Perissinotto Junior, Anízio
Formas elementares: diagonal, triangular e de Jordan / Anízio Perissinotto Junior, João Peres Vieira, Carina Alves. – São Paulo : Cultura Acadêmica : Universidade Estadual Paulista, Pró-Reitoria de Graduação, 2014.

96 p.
Bibliografia
ISBN 978-85-7983-524-7

1. Álgebra linear. 2. Matemática. 3. Jordan, Forma de. I. Vieira, João Peres. II. Alves, Carina. III. Universidade Estadual Paulista. Pró-Reitoria de Graduação.

CDD 512.5

PROGRAMA DE APOIO
À PRODUÇÃO DE MATERIAL DIDÁTICO

Considerando a importância da produção de material didático-pedagógico dedicado ao ensino de graduação e de pós-graduação, a Reitoria da UNESP, por meio da Pró-Reitoria de Graduação (PROGRAD) e em parceria com a Fundação Editora UNESP (FEU), mantém o Programa de Apoio à Produção de Material Didático de Docentes da UNESP, que contempla textos de apoio às aulas, material audiovisual, *homepages*, *softwares*, material artístico e outras mídias, sob o selo CULTURA ACADÊMICA da Editora da UNESP, disponibilizando aos alunos material didático de qualidade com baixo custo e editado sob demanda.

Assim, é com satisfação que colocamos à disposição da comunidade acadêmica mais esta obra, "Formas Elementares: Diagonal, Triangular e de Jordan", de autoria dos professores Dr. Anízio Perissinotto Junior, Dr. João Peres Vieira e Dr. Carina Alves, do Instituto de Geociências e Ciências Exatas do Câmpus de Rio Claro, esperando que ela traga contribuição não apenas para estudantes da UNESP, mas para todos aqueles interessados no assunto abordado.

SUMÁRIO

INTRODUÇÃO

As formas elementares são parte integrante de um curso de Álgebra Linear para licenciandos, bacharelandos e pós-graduandos em Matemática. Trata-se de um tema extremamente importante não apenas na Matemática como também em aplicações na Física e Engenharia. O objetivo central deste livro são as formas elementares de um operador linear, isto é, dado $T \in L(V)$, encontrar uma base de V na qual a matriz de T assume uma forma particularmente agradável. Essas matrizes serão denominadas *formas elementares* e as formas que veremos são a *forma diagonal*, a *forma triangular* e a *forma de Jordan*. Para um estudo mais completo, abordamos inicialmente: autovalor, polinômio minimal, subespaço invariante e espaço quociente. Daremos ênfase aos exemplos, mas sempre manteremos o rigor na parte teórica. Para fixarmos a notação, no decorrer de todo o texto, denotaremos por K o corpo dos números reais ou complexos, por V um espaço vetorial de dimensão finita sobre o corpo K, por $L(V)$ o espaço dos operadores lineares sobre V e por M_m o espaço das matrizes quadradas de ordem m. Para $T \in L(V)$, $\mathrm{Ker}(T) = \{x \in V \mid T(x) = 0\}$ e $\mathrm{Im}(T) = \{T(x) \mid x \in V\}$ são o *núcleo* e a *imagem* de T, respectivamente.

Os autores agradecem ao parecerista pelas sugestões que muito contribuíram para a melhoria do texto. Agradecem também ao Prof. Dr. Thiago de Melo pela diagramação do texto.

1

AUTOVALOR. POLINÔMIO MINIMAL. SUBESPAÇO INVARIANTE. ESPAÇO QUOCIENTE.

Neste capítulo vamos abordar autovalor, cuja importância surgiu a partir de estudos da Física e no estudo de formas quadráticas e equações diferenciais. Além disso vamos abordar polinômio minimal e subespaço invariante, que nos permitirá obter caracterizações de operadores diagonalizáveis (e trianguláveis) em termos de seus polinômios minimais. Por último faremos um estudo de espaço quociente.

Definição 1.1. Seja $T \in L(V)$. Se existirem $v \in V$, $v \neq 0$ e $\lambda \in K$ tais que $T(v) = \lambda v$, dizemos que λ é *autovalor* de T e v é *autovetor* de T associado a λ.

Nos três exemplos abaixo verificamos a existência ou não de autovalor.

Exemplo 1.2. Seja $T_1 \in L(\mathbb{R}^2)$ dado por $T_1(x,y) = (x, 2y)$. Temos $T_1(1,0) = 1(1,0)$ e portanto 1 é autovalor de T_1 e $(1,0)$ é autovetor de T_1 associado a 1. Também, $T_1(0,1) = (0,2) = 2(0,1)$ e portanto 2 é autovalor de T_1 e $(0,1)$ é autovetor de T_1 associado a 2. Assim, este operador possui dois autovalores distintos.

Exemplo 1.3. Seja $T_2 \in L(\mathbb{R}^2)$ dado por $T_2(x,y) = (3x + y, 3y)$. Temos $T_2(1,0) = (3,0) = 3(1,0)$ e portanto 3 é autovalor de T_2 e $(1,0)$ é autovetor de T_2 associado a 3. Também, $T_2(0,1) = (1,3) \neq \lambda(0,1)$, para todo $\lambda \in \mathbb{R}$ e portanto $(0,1)$ não é autovetor de T_2. Este operador possui só um autovalor.

Exemplo 1.4. Seja $T_3 \in L(\mathbb{R}^2)$ dado por $T_3(x,y) = (-y, x)$. Observamos que T_3 não tem autovalor em $\mathbb{R}$.

Sugestão 1.5. Mostre que se o operador T_3 do Exemplo 1.4 é tal que $T_3 \in L(\mathbb{C}^2)$ então T_3 tem autovalor em $\mathbb{C}$.

Teorema 1.6. Sejam $T \in L(V)$ e $\lambda \in K$. Considere o seguinte conjunto $V(\lambda) = \{v \in V \mid T(v) = \lambda v\}$. Então $V(\lambda)$ é um subespaço de V, chamado de auto-espaço.

Demonstração.

i) Temos que $0 \in V(\lambda)$ pois $T(0) = 0 = \lambda 0$.

ii) Se $v, u \in V(\lambda)$ então $T(v+u) = T(v) + T(u) = \lambda v + \lambda u = \lambda(v+u)$ e portanto $v + u \in V(\lambda)$.

iii) Se $a \in K$, $v \in V(\lambda)$ então $T(av) = aT(v) = a(\lambda v) = \lambda(av)$ e portanto $av \in V(\lambda)$. □

Exemplo 1.7. Do Exemplo 1.2, temos que $\lambda = 1$ e $\lambda = 2$ são autovalores de $T_1(x,y) = (x, 2y)$. Então,

$$V(1) = \left\{(x,y) \in \mathbb{R}^2 \mid T_1(x,y) = 1(x,y)\right\} = \left\{(x,0) \in \mathbb{R}^2\right\}$$

e

$$V(2) = \left\{(x,y) \in \mathbb{R}^2 \mid T_1(x,y) = 2(x,y)\right\} = \left\{(0,y) \in \mathbb{R}^2\right\}.$$

Teorema 1.8. Sejam $T \in L(V)$ e $\lambda \in K$. São equivalentes:

i) λ é um autovalor de T.

ii) O operador $T - \lambda I$ não é injetor.

Demonstração. Se λ é um autovalor de T, existe $v \neq 0$ tal que $Tv = \lambda v$. Então $Tv - \lambda v = 0$, ou equivalentemente, $(T - \lambda I)v = 0$. Assim $0 \neq v \in \operatorname{Ker}(T - \lambda I)$ e portanto $T - \lambda I$ não é injetor. Reciprocamente, se $\operatorname{Ker}(T - \lambda I) \neq 0$ então existe $0 \neq v \in V$ tal que $(T - \lambda I)(v) = 0$, ou seja, $T(v) - \lambda v = 0$, ou ainda, $T(v) = \lambda v$ e portanto λ é autovalor de T. □

Exemplo 1.9. Do Exemplo 1.3, temos que $\lambda = 3$ é autovalor de $T_2(x,y) = (3x + y, 3y)$. Então, o operador $T_2 - 3I$ é dado por $(T_2 - 3I)(x,y) = (y, 0)$. Portanto, $(T_2 - 3I)(1,0) = (0,0)$. Logo não é injetor.

Teorema 1.10. Autovetores associados a autovalores distintos são linearmente independentes.

Demonstração. Sejam $v_1, \ldots, v_n$ autovetores associados aos autovalores distintos $\lambda_1, \ldots, \lambda_n$. Vamos fazer a demonstração por indução em n.

i) Se $n = 1$ e tomamos o autovetor $v_1 \neq 0$ associado ao autovalor λ_1 temos que $\{v_1\}$ é linearmente independente.

ii) Suponhamos válido para n e mostremos que o teorema vale para $n+1$. Consideremos então $v_1, \ldots, v_{n+1}$ autovetores associados aos autovalores distintos $\lambda_1, \ldots, \lambda_{n+1}$ e façamos

$$a_1 v_1 + \cdots + a_n v_n + a_{n+1} v_{n+1} = 0. \tag{1.1}$$

Então, aplicando T a ambos os lados de (1.1) obtemos:

$$a_1 T(v_1) + \cdots + a_n T(v_n) + a_{n+1} T(v_{n+1}) = 0$$

e portanto

$$a_1 \lambda_1 v_1 + \cdots + a_n \lambda_n v_n + a_{n+1} \lambda_{n+1} v_{n+1} = 0. \tag{1.2}$$

Agora, multiplicando (1.1) por λ_{n+1}, obtemos:

$$a_1 \lambda_{n+1} v_1 + \cdots + a_n \lambda_{n+1} v_n + a_{n+1} \lambda_{n+1} v_{n+1} = 0. \tag{1.3}$$

Subtraindo (1.3) de (1.2) obtemos:

$$a_1(\lambda_1 - \lambda_{n+1}) v_1 + \cdots + a_n(\lambda_n - \lambda_{n+1}) v_n = 0.$$

Agora, por indução, cada um dos coeficientes acima é 0 e como $\lambda_i \neq \lambda_j$, segue que

$$a_i = 0, \text{ para } i = 1, \ldots, n. \tag{1.4}$$

Substituindo (1.4) em (1.1) temos que $a_{n+1} v_{n+1} = 0$ e portanto $a_{n+1} = 0$, o que demonstra o teorema. □

Definição 1.11. Seja A uma matriz quadrada sobre K. Um *autovalor* de A em K é um escalar $\lambda \in K$ tal que a matriz $A - \lambda I$ não é inversível.

Observação 1.12. Já sabemos que podemos associar um $T \in L(V)$ a uma matriz A, em relação a uma base. Assim, podemos escrever, λ é um autovalor de T se, e somente se, $\det(A - \lambda I) = \det(T - \lambda I) = 0$.

Definição 1.13. O polinômio $p_A(\lambda) = \det(A - \lambda I)$ é chamado de *polinômio característico* de A. Observe então que λ é autovalor de T se, e somente se, $p_A(\lambda) = \det(T - \lambda I) = 0$.

Exemplo 1.14. Para cada matriz

$$A = \begin{pmatrix} 0 & 7 & -6 \\ -1 & 4 & 0 \\ 0 & 2 & -2 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}, \qquad C = \begin{pmatrix} -3 & 1 & -1 \\ -7 & 5 & -1 \\ -6 & 6 & -2 \end{pmatrix},$$

encontramos todos os autovalores e uma base de cada auto-espaço.

1. $p_A(\lambda) = \det\begin{pmatrix} -\lambda & 7 & -6 \\ -1 & 4-\lambda & 0 \\ 0 & 2 & -2-\lambda \end{pmatrix} = -(\lambda - 1)(\lambda + 1)(\lambda - 2)$. Logo, os autovalores de A são 1, -1 e 2.

Para $\lambda = 1$ temos

$$\begin{pmatrix} 0 & 7 & -6 \\ -1 & 4 & 0 \\ 0 & 2 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

cujo espaço solução é gerado pelo autovetor $v_1 = (9, 3, 2)$ e v_1 é base de $V(1)$.

Para $\lambda = -1$ temos

$$\begin{pmatrix} 0 & 7 & -6 \\ -1 & 4 & 0 \\ 0 & 2 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (-1) \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

cujo espaço solução é gerado pelo autovetor $v_2 = (5, 1, 2)$ e v_2 é base de $V(-1)$.

Para $\lambda = 2$ temos

$$\begin{pmatrix} 0 & 7 & -6 \\ -1 & 4 & 0 \\ 0 & 2 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 2 \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

cujo espaço solução é gerado pelo autovetor $v_3 = (4, 2, 1)$ e v_3 é base de $V(2)$.

2. $p_B(\lambda) = \det\begin{pmatrix} 1-\lambda & -3 & 3 \\ 3 & -5-\lambda & 3 \\ 6 & -6 & 4-\lambda \end{pmatrix} = -(\lambda+2)^2(\lambda-4)$. Portanto, os autovalores de B são -2 e 4.

Para $\lambda = -2$ temos

$$\begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = (-2)\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Então $V(-2) = \{(x,y,z) \mid x - y + z = 0\} = [(1,1,0),(1,0,-1)]$ e portanto $\{(1,1,0),(1,0,-1)\}$ é base de $V(-2)$.

Para $\lambda = 4$ temos

$$\begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 4\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Então $V(4) = \left\{\left(\frac{z}{2},\frac{z}{2},z\right) \mid z \in \mathbb{R}\right\} = [(1,1,2)]$ e portanto $\{(1,1,2)\}$ é base de $V(4)$.

3. $p_C(\lambda) = \det\begin{pmatrix} -3-\lambda & 1 & -1 \\ -7 & 5-\lambda & -1 \\ -6 & 6 & -2-\lambda \end{pmatrix} = -(\lambda+2)^2(\lambda-4)$. Portanto, os autovalores de C são -2 e 4.

Para $\lambda = -2$ temos

$$\begin{pmatrix} -3 & 1 & -1 \\ -7 & 5 & -1 \\ -6 & 6 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = (-2)\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Então $V(-2) = \{(x,x,0) \mid x \in \mathbb{R}\} = [(1,1,0)]$ e portanto $\{(1,1,0)\}$ é base de $V(-2)$.

Para $\lambda = 4$ temos

$$\begin{pmatrix} -3 & 1 & -1 \\ -7 & 5 & -1 \\ -6 & 6 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 4\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Então $V(4) = \{(0,y,y) \mid y \in \mathbb{R}\} = [(0,1,1)]$ e portanto $\{(0,1,1)\}$ é base de $V(4)$.

Proposição 1.15. Se A e B são matrizes $n \times n$ sobre K, então AB e BA têm exatamente os mesmos autovalores.

Demonstração. Seja λ autovalor de AB com autovetor $v \neq 0$, isto é, $AB(v) = \lambda v$. Então $BA(B(v)) = B(AB(v)) = B(\lambda v) = \lambda B(v)$ e portanto λ é também autovalor de BA com autovetor $B(v)$ se $B(v) \neq 0$. Se $B(v) = 0$ então temos que 0 é autovalor de AB com autovetor $v \neq 0$. Assim $0 = p_{AB}(0) = \det(AB) = \det(BA) = p_{BA}(0)$ e portanto 0 é também autovalor de BA. □

Proposição 1.16. Matrizes semelhantes têm o mesmo polinômio característico.

Demonstração. Suponhamos B semelhante a A, isto é, existe uma matriz inversível P tal que $B = P^{-1}AP$. Então, $\det(B-\lambda I) = \det(P^{-1}AP-\lambda I) = \det(P^{-1}AP-\lambda P^{-1}P) = \det(P^{-1}(A-\lambda I)P) = \det(P^{-1})\det(A-\lambda I)\det(P) = \det(A-\lambda I)$. □

Exemplo 1.17. São semelhantes as matrizes

$$A = \begin{pmatrix} 0 & 7 & -6 \\ -1 & 4 & 0 \\ 0 & 2 & -2 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

A matriz que dá a semelhança é $P = \left(\begin{smallmatrix} 9 & 5 & 4 \\ 3 & 1 & 2 \\ 2 & 2 & 1 \end{smallmatrix}\right)$. Temos

$$\det(A - \lambda I) = -(\lambda - 1)(\lambda + 1)(\lambda - 2),$$
$$\det(B - \lambda I) = -(\lambda - 1)(\lambda + 1)(\lambda - 2).$$

Sugestão 1.18. Matrizes semelhantes possuem o mesmo traço.

Definição 1.19. Sejam $T \in L(V)$, $A \in M_m$ e $p(t) = a_0 + \cdots + a_n t^n$. Então definimos: $p(T) = a_0 I + a_1 T + \cdots + a_n T^n \in L(V)$ e $p(A) = a_0 I + a_1 A + \cdots + a_n A^n \in M_m$.

Exemplo 1.20. Sejam $p(t) = 1 + 2t + t^2$, $T(x, y) = (x + y, x)$ e $A = \left(\begin{smallmatrix} 1 & 0 \\ 2 & 3 \end{smallmatrix}\right)$. Então $p(T) = I + 2T + T^2$, isto é, $p(T)(x, y) = (5x + 3y, 3x + 2y)$ e $p(A) = I + 2A + A^2 = \left(\begin{smallmatrix} 4 & 0 \\ 12 & 16 \end{smallmatrix}\right)$.

Observação 1.21. Denotamos por $\sigma(A) = \{\lambda \in K \mid \det(A - \lambda I) = 0\}$. Este conjunto recebe o nome de *espectro* de A. Vale que $p_A(\sigma(A)) = \sigma(p_A(A)) = \{0\}$, onde $p_A(\lambda)$ é o polinômio característico de A.

Definição 1.22. Seja $T \in L(V)$ $[A \in M_m]$. O *polinômio minimal* de T $[A]$ é um polinômio $m(t)$ tal que:

i) $m(t)$ é o polinômio de menor grau entre os que anulam T $[A]$;

ii) $m(t)$ é um polinômio mônico (o coeficiente da maior potência de t é 1);

iii) o polinômio característico e minimal de T $[A]$ têm as mesmas raízes, exceto possivelmente, por multiplicidade.

Exemplo 1.23. Encontre o polinômio minimal das seguintes matrizes:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}; \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}; \quad C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \quad D = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix};$$

$$E = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}; \quad F = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}; \quad G = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Temos:

$$\begin{aligned}
p_A(\lambda) &= (1-\lambda)(2-\lambda), & m_A(\lambda) &= p_A(\lambda);\\
p_B(\lambda) &= (1-\lambda)^2, & m_B(\lambda) &= p_B(\lambda);\\
p_C(\lambda) &= (1-\lambda)^2, & m_C(\lambda) &= \lambda - 1;\\
p_D(\lambda) &= (1-\lambda)(2-\lambda), & m_D(\lambda) &= p_D(\lambda);\\
p_E(\lambda) &= (2-\lambda)^3, & m_E(\lambda) &= \lambda - 2;\\
p_F(\lambda) &= (2-\lambda)^3, & m_F(\lambda) &= (\lambda - 2)^2;\\
p_G(\lambda) &= (2-\lambda)^3, & m_G(\lambda) &= (\lambda - 2)^3.
\end{aligned}$$

Exemplo 1.24. Sejam A e B matrizes $n \times n$ sobre K. Pela Proposição 1.15, AB e BA têm os mesmos autovalores. Eles possuem o mesmo polinômio minimal?

Não. Vejamos um contraexemplo. Sejam $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ e $B = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. Então, $AB = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ e $BA = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$. Assim,

$$p_{AB}(\lambda) = \lambda^2, \quad m_{AB}(\lambda) = \lambda^2, \quad p_{BA}(\lambda) = \lambda^2, \quad m_{BA}(\lambda) = \lambda.$$

Exemplo 1.25. Sejam a, b, c elementos de K e $A = \left(\begin{smallmatrix} 0 & 0 & c \\ 1 & 0 & b \\ 0 & 1 & a \end{smallmatrix}\right)$. Então, o polinômio característico de A é $-x^3 + ax^2 + bx + c$ e este polinômio é também o polinômio minimal de A.

De fato, temos

$$p(x) = \det\begin{pmatrix} x & 0 & -c \\ -1 & x & -b \\ 0 & -1 & x-a \end{pmatrix} = x^2(a-x) + c + bx = -x^3 + ax^2 + bx + c.$$

Portanto, os candidatos a polinômios minimais são o próprio $p(x)$, $m_1(x) = x^2 + b_1x + b_0$ e $m_2(x) = x + a_0$. Temos,

$$m_1(A) = \begin{pmatrix} 0 & c & ac \\ 0 & b & c+ab \\ 1 & a & b+a^2 \end{pmatrix} + \begin{pmatrix} 0 & 0 & b_1c \\ b_1 & 0 & b_1b \\ 0 & b_1 & b_1a \end{pmatrix} + \begin{pmatrix} b_0 & 0 & 0 \\ 0 & b_0 & 0 \\ 0 & 0 & b_0 \end{pmatrix}$$

$$= \begin{pmatrix} b_0 & c & ac + b_1c \\ b_1 & b + b_0 & c + ab + b_1b \\ 1 & a + b_1 & b + a^2 + b_1a + b_0 \end{pmatrix},$$

que obviamente não pode ser a matriz nula e portanto $m_1(x)$ não pode ser minimal.

Passemos agora para o cálculo de $m_2(A)$. Temos,

$$m_2(A) = \begin{pmatrix} 0 & 0 & c \\ 1 & 0 & b \\ 0 & 1 & a \end{pmatrix} + \begin{pmatrix} a_0 & 0 & 0 \\ 0 & a_0 & 0 \\ 0 & 0 & a_0 \end{pmatrix} = \begin{pmatrix} a_0 & 0 & c \\ 1 & a_0 & b \\ 0 & 1 & a+a_0 \end{pmatrix},$$

que obviamente não pode ser a matriz nula e portanto $m_2(x)$ também não pode ser o minimal. Logo o polinômio minimal é o próprio polinômio característico.

Exemplo 1.26. Seja A a matriz real 4×4 dada por

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

Então o polinômio característico de A é $x^2(x-1)^2$ e esse polinômio é também o polinômio minimal.

De fato, temos

$$\begin{aligned} p(x) &= \det \begin{pmatrix} 1-x & 1 & 0 & 0 \\ -1 & -1-x & 0 & 0 \\ -2 & -2 & 2-x & 1 \\ 1 & 1 & -1 & -x \end{pmatrix} \\ &= \det \begin{pmatrix} 1-x & 1 \\ -1 & -1-x \end{pmatrix} \det \begin{pmatrix} 2-x & 1 \\ -1 & -x \end{pmatrix} \\ &= x^2(x-1)^2. \end{aligned}$$

Portanto, os candidatos a polinômios minimais, além de $p(x)$, são:

$$m_1(x) = x(x-1) = x^2 - x, \qquad m_2(x) = x^2(x-1) = x^3 - x^2,$$
$$m_3(x) = x(x-1)^2 = x^3 - 2x^2 + x.$$

Como,

$$A^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -3 & -3 & 3 & 2 \\ 2 & 2 & -2 & -1 \end{pmatrix} \quad \text{e} \quad A^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -4 & -4 & 4 & 3 \\ 3 & 3 & -3 & -2 \end{pmatrix},$$

temos $m_1(A) = A^2 - A \neq 0$; $m_2(A) = A^3 - A^2 \neq 0$ e $m_3(A) = A^3 - 2A^2 + A \neq 0$. Portanto o polinômio minimal é o polinômio característico.

Exemplo 1.27. Seja $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 0 \end{smallmatrix}\right)$ e seja $T : M_2 \to M_2$ definida por $T(B) = AB$, $B \in M_2$. Então o polinômio minimal de T é o polinômio minimal de A.

De fato, temos

$$p_A(x) = \det\begin{pmatrix} 1-x & 2 \\ 3 & -x \end{pmatrix} = x^2 - x - 6 = (x+2)(x-3) = m_A(x),$$

$$\begin{aligned}
T\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right) &= \left(\begin{smallmatrix} 1 & 0 \\ 3 & 0 \end{smallmatrix}\right) = 1e_1 + 0e_2 + 3e_3 + 0e_4,\\
T\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) &= \left(\begin{smallmatrix} 0 & 1 \\ 0 & 3 \end{smallmatrix}\right) = 0e_1 + 1e_2 + 0e_3 + 3e_4,\\
T\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right) &= \left(\begin{smallmatrix} 2 & 0 \\ 0 & 0 \end{smallmatrix}\right) = 2e_1 + 0e_2 + 0e_3 + 0e_4,\\
T\left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right) &= \left(\begin{smallmatrix} 0 & 2 \\ 0 & 0 \end{smallmatrix}\right) = 0e_1 + 2e_2 + 0e_3 + 0e_4,
\end{aligned}$$

onde e_1, e_2, e_3 e e_4 denotam a base canônica de M_2; a saber

$$e_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad e_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Logo,

$$[T] = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \\ 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \end{pmatrix}$$

e o polinômio característico é

$$\begin{aligned}
p_T(x) &= \det\begin{pmatrix} 1-x & 0 & 2 & 0 \\ 0 & 1-x & 0 & 2 \\ 3 & 0 & -x & 0 \\ 0 & 3 & 0 & -x \end{pmatrix}\\
&= (1-x)\det\begin{pmatrix} 1-x & 0 & 2 \\ 0 & -x & 0 \\ 3 & 0 & -x \end{pmatrix} + 2\det\begin{pmatrix} 0 & 1-x & 2 \\ 3 & 0 & 0 \\ 0 & 3 & -x \end{pmatrix}\\
&= x^4 - 2x^3 - 11x^2 + 12x + 36 = (x+2)^2(x-3)^2.
\end{aligned}$$

Vejamos se $m_T(x) = (x+2)(x-3) = x^2 - x - 6$. Com efeito, temos

$$m_T(T) = \begin{pmatrix} 7 & 0 & 2 & 0 \\ 0 & 7 & 0 & 2 \\ 3 & 0 & 6 & 0 \\ 0 & 3 & 0 & 6 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \\ 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 6 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 6 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Sugestão 1.28. Seja A uma matriz real simétrica de ordem três. Dê o seu polinômio característico e o minimal.

Definição 1.29. Um polinômio $p(x) \in P(K)$ de grau maior ou igual a 1 é *irredutível* se ele não pode ser escrito como o produto $p(x) = q(x)r(x)$ onde $q(x)$ e $r(x)$ são polinômios de grau maior ou igual a 1.

Proposição 1.30. Suponha que $f(t)$ é um polinômio mônico irredutível, para o qual $f(T) = 0$, onde $T \in L(V)$. Então $f(t)$ é o polinômio minimal de T.

Demonstração. Suponhamos $m(\lambda)$ o minimal. Então $\partial m < \partial f$ e também $f = mq + r$ com $r = 0$ ou $\partial r < \partial m$. Se $\partial r < \partial m$, como $f(T) = m(T)q(T) + r(T)$ e $f(T) = m(T) = 0$ segue que $r(T) = 0$ e daí m não seria o minimal. Logo $f = mq$, o que também é um absurdo, pois f é irredutível. □

Exemplo 1.31. Seja $T(x, y) = (x - y, 2y + 2x)$. Então $p(\lambda) = \lambda^2 - 3\lambda + 4$ e daí $p(T)(x, y) = (-x - 3y, 6x + 2y) - (3x - 3y, 6x + 6y) + (4x, 4y) = (0, 0)$. Logo $p(\lambda)$ é o polinômio minimal de T.

Teorema 1.32 (Teorema de Cayley–Hamilton). Cada matriz é um zero de seu polinômio característico, ou equivalentemente: "Seja $T \in L(V)$. Se $p(x)$ é o polinômio característico de T então $p(T) = 0$".

Demonstração. Consideremos $\beta = \{x_1, \dots, x_n\}$ uma base de V e chamemos de

$$A = [T]_\beta = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix}.$$

Então

$$T(x_j) = \sum_{i=1}^{n} a_{ij} x_i, \quad j = 1, \dots, n,$$

ou equivalentemente,

$$\sum_{i=1}^{n} (a_{ij} I - \delta_{ij} T)(x_i) = 0, \quad j = 1, \dots, n, \tag{1.5}$$

onde

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

Definamos agora $C \in M_n$ tal que $C_{ij} = a_{ij}I - \delta_{ij}T$. Afirmamos que $p(T) = \det(C)$. Com efeito, $p(x) = \det(A - xI)$ e o elemento ij da matriz $[A - xI]$ é dado por $a_{ij} - \delta_{ij}x$. Assim,

$$p(T) = \det(A - TI) = \det(C).$$

Mostremos agora que $p(T) = 0$. Para isto basta observarmos que $p(T) = 0$ se, e somente se, $\det(C)x_k = 0$ para todo $k = 1, \ldots, n$. Primeiramente, denotemos por $\bar{C}$ a matriz adjunta de C e lembremos que $C\bar{C} = \bar{C}C = \det(C)I$. Então por (1.5) temos

$$\sum_{i=1}^{n} C_{ij}x_i = 0, \quad j = 1, \ldots, n$$

e portanto

$$\sum_{i=1}^{n} C_{ij}\bar{C}_{jk}x_i = 0, \quad \text{para cada par } (k, j).$$

Logo,

$$\sum_{j=1}^{n}\sum_{i=1}^{n} C_{ij}\bar{C}_{jk}x_i = 0,$$

ou equivalentemente,

$$\sum_{i=1}^{n}\sum_{j=1}^{n} C_{ij}\bar{C}_{jk}x_i = 0,$$

ou ainda,

$$\sum_{i=1}^{n} \delta_{ik} \det(C)x_i = 0.$$

Portanto, $\det(C)x_k = 0$ o que demonstra que $p(T) = 0$. □

Exemplo 1.33. Verificamos o teorema de Cayley–Hamilton para a matriz $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 0 \end{smallmatrix}\right)$ e o operador linear $T(x, y) = (3x + y, 3y)$. Neste caso temos

$$p_A(\lambda) = \det\begin{pmatrix} 1 - \lambda & 2 \\ 3 & -\lambda \end{pmatrix} = \lambda^2 - \lambda - 6$$

e portanto

$$p_A(A) = A^2 - A - 6I = \begin{pmatrix} 7 & 2 \\ 3 & 6 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix} - \begin{pmatrix} 6 & 0 \\ 0 & 6 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

O polinômio característico de T é $p_T(\lambda) = (3-\lambda)^2$. Então $p_T(T) = (3I - T)^2$ e daí $p_T(T)(x,y) = (9x + 6y, 9y) - (18x + 6y, 18y) + (9x, 9y) = (0,0)$.

Definição 1.34. Sejam $T \in L(V)$ e W subespaço de V. Diz-se que W é *invariante por* T ou T-*invariante* se $T(W) \subset W$, isto é, para todo $w \in W$ temos $T(w) \in W$.

Exemplo 1.35. Seja $T \in L(V)$. Então, os subespaços $\{0\}$, V, $\mathrm{Im}(T)$, $\mathrm{Ker}(T)$ e $V(\lambda)$ são invariantes por T.

Sugestão 1.36. Seja $v \neq 0$ autovetor de T associado a λ. Então $[v]$ é T-invariante. Reciprocamente, se $U = [u]$, $u \neq 0$, é um subespaço T-invariante, então u é autovetor de T.

Exemplo 1.37. Seja $T \in L(V)$ e $\lambda \in K$ um autovalor de T. Se $S \in L(V)$ comuta com T, então o auto-espaço de λ, $V(\lambda)$, é invariante sob S. De fato, temos $V(\lambda) = \{v \in V \mid T(v) = \lambda v\}$. Devemos mostrar que se $v \in V(\lambda)$ então $S(v) \in V(\lambda)$. Agora, $S(v) \in V(\lambda)$ se, e somente se, $T(S(v)) = \lambda(S(v))$. Mas, $T(S(v)) = S(T(v)) = S(\lambda v) = \lambda S(v)$ e portanto o resultado segue.

Sugestão 1.38. Se W_1 e W_2 são subespaços T-invariantes, então $W_1 \cap W_2$ é também T-invariante.

Sugestão 1.39. Sejam $M = \left\{\left(\begin{smallmatrix} a \\ 0 \end{smallmatrix}\right),\ a \in \mathbb{R}\right\}$ e $N = \left\{\left(\begin{smallmatrix} 0 \\ b \end{smallmatrix}\right),\ b \in \mathbb{R}\right\}$. Então, $A = \left(\begin{smallmatrix} \alpha & \beta \\ 0 & \gamma \end{smallmatrix}\right)$ com $\alpha, \beta, \gamma \in \mathbb{R}$ mantém M invariante; $B = \left(\begin{smallmatrix} \alpha & 0 \\ \beta & \gamma \end{smallmatrix}\right)$ com $\alpha, \beta, \gamma \in \mathbb{R}$ mantém N invariante e M e N são invariantes com relação a C, se $C = \left(\begin{smallmatrix} \alpha & 0 \\ 0 & \gamma \end{smallmatrix}\right)$ com $\alpha, \gamma \in \mathbb{R}$.

Observação 1.40. Seja $T \in L(V)$ e W um subespaço T-invariante. Então podemos definir uma transformação linear $\hat{T} : W \to W$ por $\hat{T}(w) = T(w)$, isto é, $\hat{T}$ é a restrição de T a W, denotada por T_W.

Exemplo 1.41. Seja λ um autovalor de T e seja $V(\lambda)$ o auto-espaço associado a λ. Pergunta-se: qual é o operador $T_{V(\lambda)}$? Temos que $V(\lambda) = \{v \in V \mid T(v) = \lambda v\}$ é T-invariante. Assim, $T_{V(\lambda)} : V(\lambda) \to V(\lambda)$ é dada por $T_{V(\lambda)}(w) = T(w) = \lambda w$ e portanto $T_{V(\lambda)} = \lambda I_{V(\lambda)}$ ou seja $T_{V(\lambda)}$ é um múltiplo da identidade de $V(\lambda)$.

Proposição 1.42. Para qualquer polinômio $f(t)$ temos:

i) $f(\hat{T})(w) = f(T)(w)$, para todo $w \in W$, com W T-invariante.

ii) $\hat{m}(\lambda) \mid m(\lambda)$, isto é, o polinômio minimal de $\hat{T}$ divide o polinômio minimal de T.

Demonstração. Primeiramente, façamos a demonstração do item (i) por indução no grau de f. O caso em que o grau de f é 0 é imediato. Se $f(t) = a_0 + a_1 t$ então $f(\hat{T})(w) = (a_0 I + a_1 \hat{T})(w) = a_0 w + a_1 \hat{T}(w) = a_0 w + a_1 T(w) = (a_0 I + a_1 T)(w) = f(T)(w)$. Suponhamos agora que o grau de f seja maior que 1 e que o resultado vale para polinômios de grau menor do que n. Seja $f(t) = a_0 + a_1 t + \cdots + a_{n-1} t^{n-1} + a_n t^n$. Então, para todo $w \in W$,

$$
\begin{aligned}
f(\hat{T})(w) &= (a_0 I + a_1 \hat{T} + \cdots + a_{n-1} \hat{T}^{n-1})(w) + a_n \hat{T}^{n-1}(\hat{T}(w)) = \\
&= (a_0 I + a_1 T + \cdots + a_{n-1} T^{n-1})(w) + a_n T^{n-1}(T(w)) = f(T)(w).
\end{aligned}
$$

Passemos agora a demonstração do item (ii). Por (i), $m(\hat{T})(w) = m(T)(w) = 0w = 0$, para todo $w \in W$. Portanto $\hat{T}$ é raiz do polinômio minimal de T e daí segue que o polinômio minimal de $\hat{T}$ divide o de T. □

Exemplo 1.43. Verificamos a Proposição 1.42 para $T \in L(\mathbb{R}^2)$, definido por $T(x, y) = (x + y, y)$, $W = [(1, 0)]$ e $f(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3$. Com efeito, seja $w \in W$. Então $w = a(1, 0)$ e $T(w) = aT(1, 0) = a(1, 0)$ e portanto W é T-invariante. Defina $\hat{T} : W \to W$ por $\hat{T}(w) = \hat{T}(a, 0) = T(a, 0) = (a, 0)$. Seja $f(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3$. Para $w \in W$, $f(\hat{T})(w) = (a_0 I + a_1 \hat{T} + a_2 \hat{T}^2 + a_3 \hat{T}^3)(w) = a_0 w + a_1 \hat{T}(w) + a_2 \hat{T}(\hat{T}(w)) + a_3 \hat{T}(\hat{T}(\hat{T}(w))) = a_0(a, 0) + a_1(a, 0) + a_2(a, 0) + a_3(a, 0)$. Por outro lado, $f(T)(w) = (a_0 I + a_1 T + a_2 T^2 + a_3 T^3)(w) = a_0 w + a_1 T(w) + a_2 T(T(w)) + a_3 T(T(T(w))) =$

$a_0(a,0) + a_1(a,0) + a_2(a,0) + a_3(a,0)$. Temos $\{(1,0),(0,1)\}$ base de V e $\{(1,0)\}$ base de W. Ainda,

$$p_T(\lambda) = \det\begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix} = (1-\lambda)^2 = m_T(\lambda)$$

e

$$p_{\hat{T}}(\lambda) = \det(1-\lambda) = (1-\lambda) = m_{\hat{T}}(\lambda).$$

Portanto, $m_{\hat{T}} \mid m_T$.

Teorema 1.44. Sejam $T \in L(V)$ e U, W subespaços T-invariantes com $V = U \oplus W$. Definamos $\hat{T} = T_W$ e $\tilde{T} = T_U$. Então o polinômio minimal de T é o menor múltiplo comum dos polinômios minimais de $\hat{T}$ e $\tilde{T}$.

Demonstração. Denotemos por $m(\lambda)$, $\hat{m}(\lambda)$ e $\tilde{m}(\lambda)$ os polinômios minimais de $T, \hat{T}$ e $\tilde{T}$, respectivamente. Pela Proposição 1.42, $\hat{m}(\lambda) \mid m(\lambda)$ e $\tilde{m}(\lambda) \mid m(\lambda)$. Seja $f(\lambda)$ um múltiplo comum de $\hat{m}(\lambda)$ e $\tilde{m}(\lambda)$. Logo $f(\hat{T})(W) = \{0\}$ e $f(\tilde{T})(U) = \{0\}$. Seja $v \in V$ com $v = u + w$. Então $f(T)(v) = f(T)(u) + f(T)(w) = f(\tilde{T})(u) + f(\hat{T})(w) = 0 + 0 = 0$. Portanto, $f(T) = 0$ e daí $m(\lambda) \mid f(\lambda)$. Logo $m(\lambda)$ é o menor múltiplo comum de $\hat{m}(\lambda)$ e $\tilde{m}(\lambda)$. □

Exemplo 1.45. Verificamos o Teorema 1.44 para $T \in L(\mathbb{R}^2)$ definido por $T(x,y) = (2x+y, y+2x)$, $W = [(1,-2)]$ e $U = [(1,1)]$. Com efeito, o polinômio minimal de T é $m(\lambda) = \lambda^2 - 3\lambda$. Como $T(1,-2) = (0,0) \in W$ então W é T-invariante. Também, $T(1,1) = (3,3) = 3(1,1) \in U$ e portanto U é T-invariante. Ainda $\mathbb{R}^2 = U \oplus W$. Temos que $\hat{T} : W \to W$ é dada por $\hat{T}(x,-2x) = (0,0) = T(x,-2x)$ e $\tilde{T} : U \to U$ é dada por $\tilde{T}(x,x) = (3x,3x) = T(x,x)$. Logo $[\hat{T}] = [0]$ e $[\tilde{T}] = [3]$ donde segue que $\hat{m}(\lambda) = \lambda$ e $\tilde{m}(\lambda) = (\lambda - 3)$. Portanto, $m(\lambda) = \text{mmc}\{\hat{m}(\lambda), \tilde{m}(\lambda)\}$.

Exemplo 1.46. Seja $A = \left(\begin{smallmatrix} 1 & -1 \\ 2 & 2 \end{smallmatrix}\right)$.

i) Se T é o operador linear sobre $\mathbb{R}^2$, cuja matriz em relação à base canônica é A, então os únicos subespaços T-invariantes são $\mathbb{R}^2$ e o subespaço nulo.

ii) Se S é o operador linear sobre $\mathbb{C}$, cuja matriz em relação à base canônica é A, então $\mathbb{C}^2$ possui subespaços S-invariantes unidimensionais.

Verifiquemos inicialmente o item (i). Qualquer outro subespaço T-invariante teria necessariamente dimensão 1, isto é, $W = [v]$, com $0 \neq v \in \mathbb{R}^2$. Logo $T(v) = xv$, pois W é T-invariante. E daí x seria autovalor de T. Absurdo, pois T não possui autovalores reais. Para o item (ii), observemos que $p(x) = x^2 - 3x + 4$ e como o discriminante de $p(x)$ é $\Delta = 9 - 16 = -7$, temos dois autovalores distintos e daí cada auto-espaço é S-invariante e unidimensional.

Sugestão 1.47. Sejam $\dim V = n$, $\dim M = m$, $A \in L(V)$ e M invariante com relação a A. Seja $\beta = \{v_1, \ldots, v_m, v_{m+1}, \ldots, v_n\}$ base de V onde $\gamma = \{v_1, \ldots, v_m\}$ é base de M. Seja $A_1 = A_M$. Então

$$[A]_\beta = \begin{pmatrix} [A_1]_\gamma & [B_0]_{m\times(n-m)} \\ [0]_{(n-m)\times m} & [A_2]_{(n-m)\times(n-m)} \end{pmatrix}.$$

Exemplo 1.48. Seja $T \in L(\mathbb{R}^3)$ dada por $T(x, y, z) = (x + y - z, 2y - z, z)$. Então $W = [(1, 1, 0), (0, 1, 1)]$ é T-invariante e a matriz de T em relação à base $\{(1, 0, 0), (1, 1, 0), (0, 1, 1)\}$ é $\left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$. De fato, $T((a, a, 0) + (0, b, b)) = T(a, a + b, b) = (2a, 2a + b, b) = 2a(1, 1, 0) + b(0, 1, 1) \in W$ e portanto, W é T-invariante.

Ainda, como

$$\begin{aligned} T(1,0,0) &= (1,0,0) = 1(1,0,0) + 0(1,1,0) + 0(0,1,1), \\ T(1,1,0) &= (2,2,0) = 0(1,0,0) + 2(1,1,0) + 0(0,1,1), \\ T(0,1,1) &= (0,1,1) = 0(1,0,0) + 0(1,1,0) + 1(0,1,1) \end{aligned}$$

segue que

$$[T] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Proposição 1.49. Se $V = V_1 \oplus V_2 \oplus \cdots \oplus V_r$, onde cada subespaço V_i é de dimensão n_i e é invariante sob $T \in L(V)$, então pode-se determinar uma base de V tal que a matriz de T em relação a esta base seja da forma

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_r \end{pmatrix},$$

onde cada A_i é uma matriz $n_i \times n_i$, a matriz da transformação linear induzida por T sobre V_i.

Demonstração. Sejam $\{v_{11}, \ldots, v_{1n_1}\}$ base de V_1, ..., $\{v_{r1}, \ldots, v_{rn_r}\}$ base de V_r. Então, $\beta = \{v_{11}, \ldots, v_{1n_1}\} \cup \cdots \cup \{v_{r1}, \ldots, v_{rn_r}\}$ é base de V. Lembrando que V_i é T-invariante, temos

$$\begin{aligned}
V_1 \ni Tv_{11} &= a_{11}v_{11} + \cdots + a_{n_11}v_{1n_1} + \cdots + 0v_{r1} + \cdots + 0v_{rn_r},\\
&\vdots\\
V_1 \ni Tv_{1n_1} &= a_{1n_1}v_{11} + \cdots + a_{n_1n_1}v_{1n_1} + \cdots + 0v_{r1} + \cdots + 0v_{rn_r},\\
&\vdots\\
V_r \ni Tv_{r1} &= 0v_{11} + \cdots + 0v_{1n_1} + \cdots + c_{11}v_{r1} + \cdots + c_{n_r1}v_{rn_r},\\
&\vdots\\
V_r \ni Tv_{rn_r} &= 0v_{11} + \cdots + 0v_{1n_1} + \cdots + c_{1n_r}v_{r1} + \cdots + c_{n_rn_r}v_{rn_r}.
\end{aligned}$$

Assim,

$$[T]_\beta = \begin{pmatrix} \begin{matrix} a_{11} & \cdots & a_{1n_1} \\ \vdots & \ddots & \vdots \\ a_{n_11} & \cdots & a_{n_1n_1} \end{matrix} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \begin{matrix} c_{11} & \cdots & c_{1n_r} \\ \vdots & \ddots & \vdots \\ c_{n_r1} & \cdots & c_{n_rn_r} \end{matrix} \end{pmatrix} = \begin{pmatrix} [T_1] & 0 & \cdots & 0 \\ 0 & [T_2] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & [T_r] \end{pmatrix}. \qquad \square$$

Exemplo 1.50. Verificamos a Proposição 1.49 para os subespaços do $\mathbb{R}^4$:

$$V_1 = [(1,1,0,0), (0,0,1,1)] \quad \text{e} \quad V_2 = [(0,0,0,1), (0,1,1,0)],$$

onde $T(x, y, z, w) = (2x, 4x - 2y, x - y - z, -4x + 4y - 4z + 3w)$. É claro que $\mathbb{R}^4 = V_1 \oplus V_2$ e que $T \in L(\mathbb{R}^4)$. Ainda,

$$T(a, a, b, b) = (2a, 2a, -b, -b) = 2a(1,1,0,0) + (-b)(0,0,1,1) \in V_1$$

e

$$T(0, b, b, a) = (0, -2b, -2b, 3a) = 3a(0,0,0,1) + (-2b)(0,1,1,0) \in V_2.$$

Assim V_1 e V_2 são T-invariantes. Como $\{(1,1,0,0),(0,0,1,1)\}$ é base de V_1 e $\{(0,0,0,1),(0,1,1,0)\}$ é base de V_2, então

$$\{(1,1,0,0),(0,0,1,1),(0,0,0,1),(0,1,1,0)\}$$

é base do $\mathbb{R}^4$. Temos

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

Logo,

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

Mostremos agora que as matrizes $A_1 = \left(\begin{smallmatrix} 2 & 0\\ 0 & -1 \end{smallmatrix}\right)$ e $A_2 = \left(\begin{smallmatrix} 3 & 0\\ 0 & -2 \end{smallmatrix}\right)$ são exatamente as matrizes das transformações lineares T_1 e T_2, induzidas por T sobre V_1 e V_2 respectivamente. Temos que $T_1 : V_1 \to V_1$ e $T_2 : V_2 \to V_2$ são dadas por

$$\begin{aligned}
T_1(a,a,b,b) &= T(a,a,b,b) = (2a,2a,-b,-b),\\
T_2(0,b,b,a) &= T(0,b,b,a) = (0,-2b,-2b,3a).
\end{aligned}$$

Calculemos agora as matrizes de T_1 e T_2, respectivamente. Temos,

$$\begin{aligned}
T_1(1,1,0,0) &= (2,2,0,0) = 2(1,1,0,0)+0(0,0,1,1),\\
T_1(0,0,1,1) &= (0,0,-1,-1) = 0(1,1,0,0)+(-1)(0,0,1,1),\\
T_2(0,0,0,1) &= (0,0,0,3) = 3(0,0,0,1)+0(0,1,1,0),\\
T_2(0,1,1,0) &= (0,-2,-2,0) = 0(0,0,0,1)+(-2)(0,1,1,0),
\end{aligned}$$

e portanto, $[T_1] = A_1$ e $[T_2] = A_2$, como queríamos mostrar.

Seja V um espaço vetorial sobre o corpo K e seja W um subespaço de V. Construiremos a partir de V e W um espaço vetorial V/W sobre o corpo K chamado de espaço quociente. Este espaço quociente não é um subespaço de V, ele é um espaço vetorial definido apenas em termos de V e W que tem a seguinte propriedade: se W' é outro subespaço de V tal que $V = W \oplus W'$ então V/W é isomorfo a W'. Se W é um subespaço de V e $v \in V$, então definimos $v + W = \{v + w \mid w \in W\}$. Esse conjunto recebe o nome de *classe lateral* de W em V. Do fato da operação de adição num espaço vetorial ser comutativa segue que $v + W = W + v$. Considere agora, para $v \in V$, o conjunto $[v]_W = \{u \in V \mid u - v \in W\}$.

Lema 1.51. Para todo $v \in V$, $v + W = [v]_W$.

Demonstração. Primeiramente mostremos que $v + W \subset [v]_W$. Se $w \in W$, então $(v+w) - v = w$ também é elemento de W pois W é subespaço. Da definição de $[v]_W$ temos que $v + w \in [v]_W$ para cada $w \in W$. Logo, $v + W \subset [v]_W$. Suponha agora que $u \in [v]_W$. Então $u - v \in W$. Logo $u - v = w$ para algum $w \in W$. Assim $u = v + w$ e então $u \in v + W$. Portanto, $[v]_W \subset v + W$. Assim, concluímos a prova. □

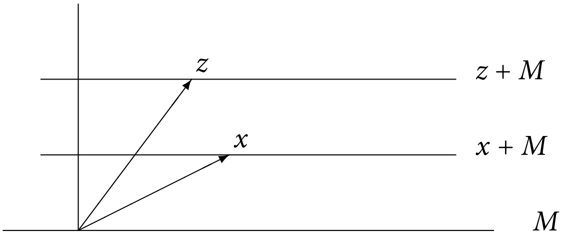
Lema 1.52. Se $v_1 + W$ e $v_2 + W$ são duas classes laterais de W em V, então ou elas são iguais ou não têm ponto em comum.

Demonstração. Suponhamos que estas duas classes laterais têm um ponto em comum, isto é, existe $v_3 \in V$ tal que $v_3 \in v_1 + W$ e $v_3 \in v_2 + W$. Assim, $v_3 = v_1 + w_1$ e $v_3 = v_2 + w_2$ para alguns w_1, w_2 em W. Como W é subespaço, temos que $v_2 - v_3 = -w_2 \in W$. Portanto, $v_2 - v_1 = v_2 - v_3 + v_3 - v_1 = -w_2 + w_1 \in W$. Logo, $v_2 \in v_1 + W$. Analogamente, $v_1 \in v_2 + W$. Assuma $u \in v_2 + W$. Como $u - v_2 \in W$ e $v_2 - v_1 \in W$ temos que $u - v_2 + v_2 - v_1 = u - v_1 \in W$. Logo $u \in v_1 + W$ e assim $v_2 + W \subset v_1 + W$. Analogamente $v_1 + W \subset v_2 + W$. □

Sugestão 1.53. Se $v_1 + W$ e $v_2 + W$ são duas classes laterais de W em V, então $v_1 + W = v_2 + W$ se, e somente se, $v_1 - v_2 \in W$.

Definição 1.54. A coleção de todas as classes laterais de W em V será indicada por V/W, isto e, $V/W = \{v + W \mid v \in V\}$.

Geometricamente:



Se $x \in M$, então $x + M = M$. Como um exemplo, consideremos $W = \{(a, a) \mid a \in \mathbb{R}\}$ e $v = (2, 3)$. Então $v + W = \{(2, 3) + (a, a) \mid a \in \mathbb{R}\} = \{(2 + a, 3 + a) \mid a \in \mathbb{R}\}$. Portanto, $v + W$ é uma reta passando pelo ponto $(2, 3)$ e paralela ao vetor $(1, 1)$.

Teorema 1.55. Seja V um espaço vetorial sobre o corpo K. Então V/W é um espaço vetorial sobre K, com as seguintes operações de adição e multiplicação por escalar:

i) $(u + W) + (v + W) = (u + v) + W$, para $u, v \in V$;

ii) $a(v + W) = av + W$, para $a \in K, v \in V$.

Demonstração. É necessário, primeiramente, mostrar que as operações estão bem definidas, isto é, sempre que $u + W = u' + W$ e $v + W = v' + W$, então:

i) $(u + v) + W = (u' + v') + W$, para quaisquer $u, v, u', v' \in V$;

ii) $ku + W = ku' + W$, para quaisquer $k \in K, u, u' \in V$.

Provemos inicialmente o item (i). Primeiramente observemos que $u + W = u' + W$ se, e somente se, $u - u' \in W$, pois $u + w_0 = u' + w_0'$ se, e somente se, $u - u' = w_0 - w_0' \in W$. Similarmente, de $v + W = v' + W$ temos $v - v' \in W$. Mas então $(u + v) - (u' + v') = (u - u') + (v - v') \in W$. Portanto, $(u + v) + W = (u' + v') + W$. Agora para o item (ii), observemos que, como $u - u' \in W$ implica $k(u - u') \in W$, então $ku - ku' = k(u - u') \in W$; portanto, $ku + W = ku' + W$. Agora, as propriedades de espaço vetorial são facilmente verificadas e deixadas como exercício para o leitor. $\square$

Proposição 1.56. Seja W um subespaço de V. Então, a função $\phi : V \to V/W$, definida por $\phi(v) = v + W$, é linear.

Demonstração. Com efeito, temos: $\phi(u+v) = (u+v)+W = (u+W)+(v+W) = \phi(u) + \phi(v)$, para todos $u, v \in V$ e $\phi(au) = au + W = a(u + W) = a\phi(u)$, para todos $a \in K, u \in V$. □

Teorema 1.57. Sejam $T \in L(V)$ e W um subespaço de V, T-invariante. Então, T induz um operador linear $\bar{T}$ em V/W, definido por $\bar{T}(v + W) = Tv + W$. Além disso, se T é zero de algum polinômio, $\bar{T}$ também o é. Assim, o polinômio minimal de $\bar{T}$ divide o polinômio minimal de T.

Demonstração. Mostremos inicialmente que $\bar{T}$ é bem definido. De fato, se $u + W = v+W$, então $u-v \in W$ e como W é T-invariante, $T(u-v) = T(u)-T(v) \in W$. Assim $\bar{T}(u + W) = T(u) + W = T(v) + W = \bar{T}(v + W)$. Mostremos agora que $\bar{T}$ é linear.

(L1) $\bar{T}(u+W+v+W) = \bar{T}(u+v+W) = T(u+v)+W = T(u)+T(v)+W = T(u) + W + T(v) + W = \bar{T}(u + W) + \bar{T}(v + W)$.

(L2) $\bar{T}(a(u+W)) = \bar{T}(au+W) = T(au)+W = aT(u)+W = a(Tu+W) = a\bar{T}(u + W)$.

Afirmamos que $\bar{T}^n = \overline{T^n}$, para qualquer n. Faremos a prova por indução em n. O caso $n = 1$ é imediato. Suponhamos válido para $n - 1$ e provemos que o resultado vale para n. Com efeito, $\bar{T}^n(u+W) = \bar{T}(\bar{T}^{n-1}(u+W)) = \bar{T}(\overline{T^{n-1}}(u+W)) = \bar{T}(T^{n-1}(u) + W) = T(T^{n-1}(u)) + W = T^n(u) + W = \overline{T^n}(u + W)$. Portanto, $\bar{T}^n = \overline{T^n}$. Finalmente, seja $f(t) = a_n t^n + \cdots + a_1 t + a_0$. Então, $f(T) = a_n T^n + \cdots + a_1 T + a_0 I$ e $f(\bar{T}) = a_n \bar{T}^n + \cdots + a_1 \bar{T} + a_0 \bar{I} = a_n \overline{T^n} + \cdots + a_1 \bar{T} + a_0 \bar{I}$. Portanto, $\overline{f(T)}(u+W) = f(T)(u)+W = (a_n T^n(u)+\cdots+a_1 T(u)+a_0 I(u)) + W = a_n T^n(u) + W + \cdots + a_1 T(u) + W + a_0 I(u) + W = a_n(T^n(u) + W) + \cdots + a_1(T(u) + W) + a_0(I(u) + W) = a_n \overline{T^n}(u + W) + \cdots + a_1 \bar{T}(u + W) + a_0 \bar{I}(u + W) = f(\bar{T})(u + W)$ e portanto $\overline{f(T)} = f(\bar{T})$. Assim, se $f(T) = 0$, então $\overline{f(T)} = \bar{0} = W = f(\bar{T})$ e daí $\bar{T}$ também é raiz de f. □

Exemplo 1.58. Aplicamos o Teorema 1.57 ao operador linear do $\mathbb{R}^2$ definido por $T(x,y) = (-y, x-2y)$ e ao subespaço $W = [(1,1)]$ de $\mathbb{R}^2$, mostrando inicialmente que W é T-invariante. Para isto, seja $(a,a) \in W$. Então $T(a,a) = (-a,-a) = -a(1,1) \in W$. Temos $\mathbb{R}^2/W = \{(a,b) + [(1,1)] \mid (a,b) \in \mathbb{R}^2\}$ e $\bar{T} : \mathbb{R}^2/W \to \mathbb{R}^2/W$ definida por $\bar{T}((a,b)+W) = T(a,b)+W = (-b, a-2b) + [(1,1)]$. Seja $f(t) = t^3 + 2t^2 + t$. Então $f(T) = T^3 + 2T^2 + T$ e $f(T)(x,y) = T(T(T(x,y))) + 2T(T(x,y)) + T(x,y) = (2x-3y, 3x-4y) + (-2x+4y, -4x+6y) + (-y, x-2y) = (0,0)$. Logo T é um zero desse polinômio. Ainda, $f(\bar{T}) = \bar{T}^3 + 2\bar{T}^2 + \bar{T}$ e portanto $f(\bar{T})((a,b)+W) = \bar{T}(\bar{T}(\bar{T}((a,b)+W))) + 2\bar{T}(\bar{T}((a,b)+W)) + \bar{T}((a,b)+W) = (2a-3b, 3a-4b) + W + (-2a+4b, -4a+6b) + W + (-b, a-2b) + W = (0,0) + W = W$. Portanto, $\bar{T}$ também é um zero desse polinômio. Temos,

$$p(\lambda) = \det\begin{pmatrix} -\lambda & -1 \\ 1 & -2-\lambda \end{pmatrix} = \lambda^2 + 2\lambda + 1 = (\lambda+1)^2 = m(\lambda).$$

Afirmamos que $\{(-1,1)+W\}$ é base de $\mathbb{R}^2/W$. Este conjunto é linearmente independente pois $(-1,1) + W \neq W$. Seja $(a,b) + W \in \mathbb{R}^2/W$. Queremos mostrar que $(a,b) + W = k((-1,1) + W) = (-k,k) + W$ para algum $k \in \mathbb{R}$, ou equivalentemente, que $(a,b) - (-k,-k) \in W$ para algum $k \in \mathbb{R}$. Com efeito, tome $k = \frac{b-a}{2}$. Então $(a,b) + W = (\frac{-b+a}{2}, \frac{b-a}{2}) + W = \frac{b-a}{2}[(-1,1) + W]$. Logo gera e portanto é base. $\bar{T}((-1,1) + W) = T(-1,1) + W = (-1,-3) + W = (-1)[(-1,1) + W]$. Portanto, $[\bar{T}] = [-1]$ e então $\bar{p}(\lambda) = -1 - \lambda$ e $\bar{m}(\lambda) = \lambda + 1$.

Teorema 1.59. Suponha que $\{w_1, \ldots, w_r\}$ é base do subespaço W de V e que $\{v_1 + W, \ldots, v_s + W\}$ é base do espaço quociente V/W. Então $\{v_1, \ldots, v_s, w_1, \ldots, w_r\}$ é base de V. Assim,

$$\dim V = \dim W + \dim V/W.$$

Demonstração. Seja $u \in V$. Então, $u + W = a_1(v_1 + W) + \cdots + a_s(v_s + W)$. Logo,

$$u = (a_1v_1 + \cdots + a_sv_s) + w, \tag{1.6}$$

onde $w \in W$. Como $\{w_1, \ldots, w_r\}$ é base de W, então, $w = b_1w_1 + \cdots + b_rw_r$ e daí, substituindo w em (1.6), $u = a_1v_1 + \cdots + a_sv_s + b_1w_1 + \cdots + b_rw_r$. Logo, esses vetores geram V. Mostremos, agora, que são linearmente independentes. De fato, sejam $c_1v_1+\cdots+c_sv_s+d_1w_1+\cdots+d_rw_r = 0$. Então, $c_1v_1+\cdots+c_sv_s+W = 0+W$ pois $d_1w_1+\cdots+d_rw_r \in W$ e portanto, $c_1(v_1+W)+\cdots+c_s(v_s+W) = 0+W = W$. Como $\{v_1+W,\ldots,v_s+W\}$ é linearmente independente então $c_i = 0$, para todo $i = 1,\ldots,s$. Logo, $d_1w_1+\cdots+d_rw_r = 0$, o que implica que os d_i's são todos nulos e o teorema está demonstrado. □

Exemplo 1.60. Verificamos o Teorema 1.59 para $W = [(1,1,0),(0,1,1)]$, subespaço do $\mathbb{R}^3$, observando que $\{(1,1,0),(0,1,1)\}$ é base de W e $\mathbb{R}^3/W = \{v + [(1,1,0),(0,1,1)] \mid v \in \mathbb{R}^3\}$. Afirmamos que $\{(1,0,0) + [(1,1,0),(0,1,1)]\}$ é base de $\mathbb{R}^3/W$. Esse vetor é não nulo pois $(1,0,0) \notin [(1,1,0),(0,1,1)]$. Mostremos que ele gera $\mathbb{R}^3/W$. Com efeito, seja $v + W \in \mathbb{R}^3/W$; $v = (x,y,z)$. Queremos mostrar que $v + W = k((1,0,0) + W) = k(1,0,0) + W$. Sejam $(x,y,z) + a(1,1,0) + b(0,1,1)$ e $(k,0,0) + c(1,1,0) + d(0,1,1)$ elementos de $v + W$ e $(k,0,0) + W$, respectivamente. Então, $(x,y,z) + (a,a,0) + (0,b,b) = (k,0,0) + (c,c,0) + (0,d,d)$. Resolvendo, encontramos: $(x,y,z) + a(1,1,0) + b(0,1,1) = (x-y+z)(1,0,0) + (y-z+a)(1,1,0) + (z+b)(0,1,1)$. Assim gera e logo é base. É claro que $\{(1,0,0),(1,1,0),(0,1,1)\}$ é base do $\mathbb{R}^3$ e ainda que $\dim W = 2$ e $\dim \mathbb{R}^3/W = 1$.

Teorema 1.61. Sejam V um espaço vetorial e M, N subespaços vetoriais de V tais que $V = M \oplus N$. Então V/M é isomorfo a N.

Demonstração. Definamos $\phi : N \to V/M$ por $\phi(x) = x + M$. Observamos que ϕ é linear. Com efeito, $\phi(x+y) = (x+y)+M = (x+M)+(y+M) = \phi(x)+\phi(y)$ e $\phi(ax) = (ax) + M = a(x + M) = a\phi(x)$. Também, ϕ é sobrejetora. De fato, seja $x_0 + M \in V/M$. Então $x_0 = u_0 + v_0 \in V = M \oplus N$. Assim, $x_0 + M = v_0 + u_0 + M = v_0 + M$ e portanto $\phi(v_0) = v_0 + M = x_0 + M$. Ainda, ϕ é injetora. Temos $\operatorname{Ker}\phi = \{x \in N \mid \phi(x) = 0 + M\}$. Mas $x + M = 0 + M$ implica que $x \in M$ e portanto $x = 0$ desde que $M \cap N = \{0\}$ e daí $\operatorname{Ker}\phi = \{0\}$. □

Com este teorema a prova do Teorema 1.59 fica reduzida assim:

> "Seja N subespaço de V tal que $V = M \oplus N$. Então $\dim V = \dim M + \dim N = \dim M + \dim V/M$".

Proposição 1.62. Suponha que o subconjunto $\{v_1 + W, \ldots, v_r + W\}$ de V/W é linearmente independente. Então $\{v_1, \ldots, v_r\}$ em V é também linearmente independente.

Demonstração. Com efeito, se $a_1v_1 + \cdots + a_rv_r = 0$ então $a_1(v_1 + W) + \cdots + a_r(v_r + W) = 0 + W = W$ e como o subconjunto $\{v_1 + W, \ldots, v_r + W\}$ de V/W é linearmente independente segue que os a_i's são todos nulos. □

Proposição 1.63. Suponha que $V = U \oplus W$ e que $\{u_1, \ldots, u_r\}$ é base de U. Então $\{u_1 + W, \ldots, u_r + W\}$ é base de V/W.

Demonstração. Com efeito, suponhamos que $a_1(u_1+W)+\cdots+a_r(u_r+W) = W$. Então, $(a_1u_1+\cdots+a_ru_r)+W = W$ e portanto $(a_1u_1+\cdots+a_ru_r) = 0$ de U. Como $\{u_1, \ldots, u_r\}$ é base de U segue que os a_i's são todos nulos. □

Exemplo 1.64. Seja $A \in L(V)$. Defina $T : V/\operatorname{Ker}(A) \to \operatorname{Im}(A)$ por $T(x + \operatorname{Ker}(A)) = A(x)$. Então T é um isomorfismo. De fato, verifiquemos primeiramente que T está bem definida. Com efeito, se $x_1 + \operatorname{Ker}(A) = x_2 + \operatorname{Ker}(A)$ então $(x_1 - x_2) \in \operatorname{Ker}(A)$ e portanto $A(x_1 - x_2) = 0$ o que implica $A(x_1) = A(x_2)$. Logo está bem definida. Também, T é linear, pois $T(a(x + \operatorname{Ker}(A)) + b(y + \operatorname{Ker}(A))) = T((ax + \operatorname{Ker}(A)) + (by + \operatorname{Ker}(A))) = T((ax + by) + \operatorname{Ker}(A)) = A(ax+by) = aA(x)+bA(y) = aT(x+\operatorname{Ker}(A))+bT(y+\operatorname{Ker}(A))$. Ainda, T é sobre pois se $A(v) \in \operatorname{Im}(A)$ tome $v+\operatorname{Ker}(A) \in V/\operatorname{Ker}(A)$ e daí $T(v+\operatorname{Ker}(A)) = A(v)$ e mais, T é injetora desde que se $A(x) = 0$ então $x \in \operatorname{Ker}(A)$ e portanto $\operatorname{Ker}(T) = \operatorname{Ker}(A)$ e daí T é um isomorfismo.

Sugestão 1.65. Suponha que W e U são subespaços de V. Então $(W + U)/W$ é isomorfo a $U/(W \cap U)$.

Sugestão 1.66. Sejam U e W subespaços de V com $W \subset U \subset V$. Mostre que:

i) U/W é subespaço de V/W.

ii) $(V/W)/(U/W)$ é isomorfo a V/U.

iii) $\dim V/W = \dim V/U + \dim U/W$.

2

FORMA DIAGONAL

Neste capítulo pretendemos encontrar condições necessárias e/ou suficientes para um operador linear ser diagonalizável. Veremos como autovalores e autovetores podem ajudar no processo de diagonalização.

Definição 2.1. Seja $T \in L(V)$. Dizemos que T é *diagonalizável* se existe uma base de V formada de autovetores.

Nos próximos quatro exemplos veremos a diagonalização ou não de certos operadores lineares.

Exemplo 2.2. Seja $T \in L(\mathbb{R}^2)$ definido por $T(x, y) = (x, 2y)$. Então T é diagonalizável pois o conjunto $\{(1, 0), (0, 1)\}$ é uma base de $\mathbb{R}^2$ formada de autovetores. Como fica a matriz de T em relação à essa base? Temos

$$T(1, 0) = (1, 0) = 1(1, 0) + 0(0, 1), \qquad T(0, 1) = (0, 2) = 0(1, 0) + 2(0, 1).$$

Portanto, $[T] = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right)$.

Exemplo 2.3. Seja $T \in L(\mathbb{R}^3)$ definido por $T(x, y, z) = (7y - 6z, -x + 4y, 2y - 2z)$. Obtemos os autovalores $1, -1, 2$ e autovetores associados $(9, 3, 2)$, $(5, 1, 2)$ e $(4, 2, 1)$, respectivamente, que formam uma base de $\mathbb{R}^3$ e portanto T é diagonalizável. Qual a matriz de T em relação à essa base? Temos

$$\begin{aligned}
T(9, 3, 2) &= (9, 3, 2) = 1(9, 3, 2) + 0(5, 1, 2) + 0(4, 2, 1),\\
T(5, 1, 2) &= (-5, -1, -2) = 0(9, 3, 2) + (-1)(5, 1, 2) + 0(4, 2, 1),\\
T(4, 2, 1) &= (8, 4, 2) = 0(9, 3, 2) + 0(5, 1, 2) + 2(4, 2, 1).
\end{aligned}$$

Portanto, $[T] = \left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{smallmatrix}\right)$.

Exemplo 2.4. Seja $T \in L(\mathbb{R}^3)$ definido por $T(x, y, z) = (x - 3y + 3z, 3x - 5y + 3z, 6x-6y+4z)$. Obtemos os autovalores -2 e 4 e autovetores $(1,1,0)$ e $(1,0,-1)$ associados a -2 e $(1,1,2)$ associado a 4, que formam uma base de $\mathbb{R}^3$ e portanto T é diagonalizável e a matriz de T em relação a essa base é

$$[T] = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

Exemplo 2.5. Seja $T \in L(\mathbb{R}^3)$; $T(x, y, z) = (-3x+y-z, -7x+5y-z, -6x+6y-2z)$. Então T não é diagonalizável, pois não possui uma base de autovetores.

Teorema 2.6. Sejam $T \in L(V)$ e $\dim V = n$. Suponha que T possua n autovalores distintos $\lambda_1, \dots, \lambda_n$. Então T é diagonalizável.

Demonstração. Sejam $v_1, \dots, v_n$ autovetores não nulos associados aos autovalores distintos $\lambda_1, \dots, \lambda_n$. Pelo Teorema 1.10 sabemos que eles são linearmente independentes. Logo formam uma base, pois $\dim V = n$. Então T é diagonalizável. Como temos

$$\begin{array}{ccccc} T(v_1) = \lambda_1 v_1 & = \lambda_1 v_1 & + 0v_2 & + \cdots & + 0v_n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ T(v_n) = \lambda_n v_n & = 0v_1 & + 0v_2 & + \cdots & + \lambda_n v_n \end{array}$$

segue que

$$[T] = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$ □

Definição 2.7. Chamamos de *multiplicidade algébrica* de um autovalor λ a quantidade de vezes que ele aparece como raiz do polinômio característico. Chamamos de *multiplicidade geométrica* de um autovalor λ a dimensão do autoespaço $V(\lambda)$.

Sugestão 2.8. Verifique, no exemplo 1.14, as multiplicidades geométricas.

Teorema 2.9. O operador linear $T \in L(V)$ é diagonalizável se, e somente se,

i) o polinômio característico de T tem todas as raízes em K;

ii) a multiplicidade algébrica de cada autovalor λ_i é igual à multiplicidade geométrica de λ_i.

Demonstração. Vamos fazer a demonstração para um caso particular. Inicialmente, vamos supor que $\{v_1, v_2, u_1, u_2, u_3, w\}$ seja uma base de autovetores de V onde $\{v_1, v_2\}$, $\{u_1, u_2, u_3\}$ e $\{w\}$ são os autovetores associados aos autovalores distintos λ_1, λ_2 e λ_3, respectivamente. A matriz de T em relação a essa base é

$$[T] = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_3 \end{pmatrix}.$$

Logo $p_T(\lambda) = (\lambda_1 - \lambda)^2(\lambda_2 - \lambda)^3(\lambda_3 - \lambda)$ cujas raízes estão em K. Mostremos agora que $V(\lambda_1) = [v_1, v_2]$, $V(\lambda_2) = [u_1, u_2, u_3]$ e $V(\lambda_3) = [w]$. É claro que $[v_1, v_2] \subset V(\lambda_1)$, $[u_1, u_2, u_3] \subset V(\lambda_2)$ e $[w] \subset V(\lambda_3)$.

Seja $v \in V(\lambda_1)$. Então $T(v) = \lambda_1 v$ e por outro lado temos

$$v = a_1 v_1 + a_2 v_2 + b_1 u_1 + b_2 u_2 + b_3 u_3 + cw \tag{2.1}$$

e multiplicando (2.1) por λ_1 obtemos

$$\lambda_1 v = a_1\lambda_1 v_1 + a_2\lambda_1 v_2 + b_1\lambda_1 u_1 + b_2\lambda_1 u_2 + b_3\lambda_1 u_3 + c\lambda_1 w. \tag{2.2}$$

Aplicando T em (2.1), obtemos

$$\lambda_1 v = a_1 T(v_1) + a_2 T(v_2) + b_1 T(u_1) + b_2 T(u_2) + b_3 T(u_3) + cT(w)$$

e portanto

$$\lambda_1 v = a_1\lambda_1 v_1 + a_2\lambda_1 v_2 + b_1\lambda_2 u_1 + b_2\lambda_2 u_2 + b_3\lambda_2 u_3 + c\lambda_3 w. \tag{2.3}$$

Igualando (2.2) e (2.3), obtemos

$$b_1\lambda_1 = b_1\lambda_2, \qquad b_2\lambda_1 = b_2\lambda_2, \qquad b_3\lambda_1 = b_3\lambda_2, \qquad c\lambda_1 = c\lambda_3,$$

donde segue que $b_1 = b_2 = b_3 = c = 0$ e portanto $v = a_1v_1 + a_2v_2 \in [v_1, v_2]$. Portanto, $V(\lambda_1) = [v_1, v_2]$ e $\dim V(\lambda_1) = 2$. Logo, temos

$$\text{multiplicidade algébrica} = \text{multiplicidade geométrica.}$$

Analogamente, concluímos que $V(\lambda_2) = [u_1, u_2, u_3]$, donde $\dim V(\lambda_2) = 3$ e $V(\lambda_3) = [w]$, donde $\dim V(\lambda_3) = 1$.

Reciprocamente, assumamos por hipótese, que o polinômio característico de T possa ser fatorado sobre K. Suponhamos $p(\lambda) = (\lambda_1-\lambda)^2(\lambda_2-\lambda)^3(\lambda_3-\lambda)$ onde $\lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1$ e $2+3+1 = \text{grau}\, p(\lambda) = \dim V$. Também $\dim V(\lambda_1) = 2$; $\dim V(\lambda_2) = 3$ e $\dim V(\lambda_3) = 1$. Seja $H = V(\lambda_1) + V(\lambda_2) + V(\lambda_3)$. Mostremos que $V(\lambda_1) \cap V(\lambda_2) = V(\lambda_1) \cap V(\lambda_3) = V(\lambda_2) \cap V(\lambda_3) = \{0\}$. Com efeito, seja $u \in V(\lambda_i) \cap V(\lambda_j)$. Então $T(u) = \lambda_i u = \lambda_j u$ e daí $(\lambda_i - \lambda_j)u = 0$ e portanto $u = 0$ desde que $\lambda_i \neq \lambda_j$. Assim, $V(\lambda_2) \cap V(\lambda_1) = \{0\}$ e $V(\lambda_3) \cap (V(\lambda_1) + V(\lambda_2)) = \{0\}$, então $H = V(\lambda_1) \oplus V(\lambda_2) \oplus V(\lambda_3)$.

Sendo H um subespaço de V e de mesma dimensão que V segue que $H = V$. Então $\beta = \{v_1, v_2, u_1, u_2, u_3, w\}$, onde $\{v_1, v_2\}$ é base de $V(\lambda_1)$, $\{u_1, u_2, u_3\}$ é base de $V(\lambda_2)$ e $\{w\}$ é base de $V(\lambda_3)$, será base de V formada por autovetores. Logo, T é diagonalizável. □

Proposição 2.10. Sejam $T \in L(V)$ e $m(\lambda) = \lambda - \lambda_1$ o polinômio minimal de T. Então $V = \text{Ker}(T - \lambda_1 I)$.

Demonstração. Como $m(\lambda) = \lambda - \lambda_1$ então $m(T) = T - \lambda_1 I = 0$. Logo, para $v \in V$, temos $m(T)(v) = (T - \lambda_1 I)(v) = 0$. □

Proposição 2.11. Sejam $T \in L(V)$ e $m(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) = m_1(\lambda)m_2(\lambda)$ o polinômio minimal de T. Então:

i) $V = \text{Ker}(\lambda_1 I - T) \oplus \text{Ker}(\lambda_2 I - T)$;

ii) O polinômio minimal da restrição de T a $\text{Ker}(\lambda_i I - T)$ é $m_i(\lambda)$, $i = 1, 2$.

Demonstração. Provemos inicialmente o item (i). Como $m_1(\lambda)$ e $m_2(\lambda)$ são primos entre si, existem polinômios $r(\lambda)$ e $s(\lambda)$ tais que $r(\lambda)m_1(\lambda) + s(\lambda)m_2(\lambda) = 1$. Portanto, para o operador T,

$$r(T)m_1(T) + s(T)m_2(T) = I. \tag{2.4}$$

Seja $v \in V$. Aplicando (2.4) temos

$$r(T)m_1(T)(v) + s(T)m_2(T)(v) = v.$$

Mas, $m_2(T)r(T)m_1(T)(v) = r(T)m_1(T)m_2(T)(v) = r(T)0(v) = 0$. Logo, $r(T)m_1(T)(v) \in \operatorname{Ker}(\lambda_2 I - T)$.

Semelhantemente, $s(T)m_2(T)(v) \in \operatorname{Ker}(\lambda_1 I - T)$. Portanto, $V = \operatorname{Ker}(\lambda_1 I - T) + \operatorname{Ker}(\lambda_2 I - T)$.

Vamos agora provar que é de maneira única. Suponha $v = u + w = u_1 + w_1$, para $u, u_1 \in \operatorname{Ker}(\lambda_1 I - T)$ e $w, w_1 \in \operatorname{Ker}(\lambda_2 I - T)$. Temos

$$r(T)m_1(T)(v) = r(T)m_1(T)(u) + r(T)m_1(T)(w) = r(T)m_1(T)(w)$$

e

$$r(T)m_1(T)(v) = r(T)m_1(T)(u_1) + r(T)m_1(T)(w_1) = r(T)m_1(T)(w_1)$$

Aplicando (2.4) em w e w_1 obtemos

$$w = r(T)m_1(T)(w) + s(T)m_2(T)(w) = r(T)m_1(T)(w)$$

e

$$w_1 = r(T)m_1(T)(w_1) + s(T)m_2(T)(w_1) = r(T)m_1(T)(w_1).$$

Portanto,

$$w = r(T)m_1(T)(w) = r(T)m_1(T)(v) = r(T)m_1(T)(w_1) = w_1.$$

Analogamente, $u = u_1$. Portanto, $V = \operatorname{Ker}(\lambda_1 I - T) \oplus \operatorname{Ker}(\lambda_2 I - T)$.

Passemos agora ao item (ii). Sejam $\hat{m}_1(\lambda)$ o polinômio minimal da restrição T_1 de T ao $\operatorname{Ker}(\lambda_1 I - T)$ e $\hat{m}_2(\lambda)$ o polinômio minimal da restrição T_2 de T ao

$\text{Ker}(\lambda_2 I - T)$. Temos, $m_1(T_1)(u) = (\lambda_1 I - T_1)(u) = \lambda_1 u - T_1(u) = \lambda_1 u - T(u) = (\lambda_1 I - T)(u) = 0$, para todo $u \in \text{Ker}(\lambda_1 I - T)$. Logo, $m_1(T_1) = 0$. Assim, também, $m_2(T_2) = 0$. Portanto, $\hat{m}_1(\lambda) \mid m_1(\lambda)$ e $\hat{m}_2(\lambda) \mid m_2(\lambda)$. Logo, $\hat{m}_1(\lambda) = m_1(\lambda) = (\lambda_1 - \lambda)$ e $\hat{m}_2(\lambda) = m_2(\lambda) = (\lambda_2 - \lambda)$. □

Teorema 2.12. Sejam $T \in L(V)$ e $m(\lambda) = (\lambda_1 - \lambda)\cdots(\lambda_r - \lambda) = m_1(\lambda)\cdots m_r(\lambda)$ o polinômio minimal de T. Então:

i) $V = \text{Ker}(\lambda_1 I - T) \oplus \cdots \oplus \text{Ker}(\lambda_r I - T)$;

ii) O polinômio minimal da restrição de T a $\text{Ker}(\lambda_j I - T)$ é $m_j(\lambda)$, $j = 1, \dots, r$.

Demonstração. Faremos a prova por indução em r. Os casos $r = 1$ e $r = 2$ seguem das Proposições 2.10 e 2.11, respectivamente. Suponhamos válido para $r - 1$ e provemos para r. Façamos inicialmente o item (i). Como $\bar{m}(\lambda) = (\lambda_1 - \lambda)\cdots(\lambda_{r-1} - \lambda)$ e $m_r(\lambda)$ são primos entre si, existem polinômios $r(\lambda)$ e $s(\lambda)$ tais que $r(\lambda)\bar{m}(\lambda) + s(\lambda)m_r(\lambda) = 1$. Portanto, para o operador linear T,

$$r(T)\bar{m}(T) + s(T)m_r(T) = I. \tag{2.5}$$

Seja $v \in V$. Então aplicando (2.5), temos $r(T)\bar{m}(T)(v) + s(T)m_r(T)(v) = v$.

Mas,

$$m_r(T)r(T)\bar{m}(T)(v) = r(T)m_r(T)\bar{m}(T)(v) = r(T)m(T)(v) = r(t)0(v) = 0.$$

Logo, $r(T)\bar{m}(T)(v) \in \text{Ker}(\lambda_r I - T)$. Semelhantemente, $s(T)m_r(T)(v) \in \text{Ker}(\bar{m}(T))$. Portanto, $V = \text{Ker}(\bar{m}(T)) + \text{Ker}(\lambda_r I - T)$.

Vamos agora provar que é de maneira única. Suponha $v = w + u = w_1 + u_1$, $u, u_1 \in \text{Ker}(\lambda_r I - T)$ e $w, w_1 \in \text{Ker}(\bar{m}(T))$. Temos

$$r(T)\bar{m}(T)(v) = r(T)\bar{m}(T)(w) + r(T)\bar{m}(T)(u) = r(T)\bar{m}(T)(u)$$

e

$$r(T)\bar{m}(T)(v) = r(T)\bar{m}(T)(w_1) + r(T)\bar{m}(T)(u_1) = r(T)\bar{m}(T)(u_1).$$

Aplicando (2.5) em u e u_1 obtemos

$$u = r(T)\bar{m}(T)(u) + s(T)m_r(T)(u) = r(T)\bar{m}(T)(u)$$

e

$$u_1 = r(T)\bar{m}(T)(u_1) + s(T)m_r(T)(u_1) = r(T)\bar{m}(T)(u_1).$$

Portanto, $u = r(T)\bar{m}(T)(u) = r(T)\bar{m}(T)(u_1) = u_1$.

Analogamente, $w = w_1$ e assim $V = \mathrm{Ker}(\bar{m}(T)) \oplus \mathrm{Ker}(\lambda_r I - T)$.

Passemos agora ao item (ii). Sejam $\hat{m}(\lambda)$ o polinômio minimal da restrição $\hat{T}$ de T ao $\mathrm{Ker}(\bar{m}(T))$ e $\tilde{m}(\lambda)$ o polinômio minimal da restrição $\tilde{T}$ de T ao $\mathrm{Ker}(\lambda_r I - T)$. Temos, $m_r(\tilde{T})(u) = (\lambda_r I - \tilde{T})(u) = \lambda_r I(u) - \tilde{T}(u) = \lambda_r u - T(u) = (\lambda_r I - T)(u) = 0$, para todo $u \in \mathrm{Ker}(\lambda_r I - T)$. Logo, $m_r(\tilde{T}) = 0$. Também temos $\tilde{m}(\tilde{T}) = 0$. Portanto, $\tilde{m}(\lambda) \mid m_r(\lambda) = \lambda_r - \lambda$ e daí $\tilde{m}(\lambda) = m_r(\lambda)$. Como $\bar{m}(\lambda)$ e $m_r(\lambda)$ são primos entre si então o $\mathrm{mmc}\{\bar{m}(\lambda), m_r(\lambda)\} = \bar{m}(\lambda)m_r(\lambda) = m(\lambda) = \mathrm{mmc}\{\hat{m}(\lambda), \tilde{m}(\lambda)\}$ e portanto, $\bar{m}(\lambda) = \hat{m}(\lambda)$. Aplicando a hipótese de indução a $\hat{T}$ e $\hat{m}(\lambda) = (\lambda_1 - \lambda)\cdots(\lambda_{r-1} - \lambda)$ obtemos

i) $\mathrm{Ker}(\bar{m}(T)) = \mathrm{Ker}(\lambda_1 I - \hat{T}) \oplus \cdots \oplus \mathrm{Ker}(\lambda_{r-1} I - \hat{T})$;

ii) O polinômio minimal da restrição de $\hat{T}$ a $\mathrm{Ker}(\lambda_j I - \hat{T})$ é $m_j(\lambda)$, $j = 1, \ldots, r-1$.

Logo, $V = \mathrm{Ker}(\bar{m}(T)) \oplus \mathrm{Ker}(\lambda_r I - T) = \mathrm{Ker}(\lambda_1 I - \hat{T}) \oplus \cdots \oplus \mathrm{Ker}(\lambda_{r-1} I - \hat{T}) \oplus \mathrm{Ker}(\lambda_r I - \tilde{T})$. Pela Proposição 2.10, temos que $\mathrm{Ker}(\lambda_r I - \tilde{T}) = \mathrm{Ker}(\lambda_r I - T)$. Precisamos mostrar que $\mathrm{Ker}(\lambda_j I - \hat{T}) = \mathrm{Ker}(\lambda_j I - T)$, para $j = 1, \ldots, r-1$. De fato, se $z \in \mathrm{Ker}(\lambda_j I - \hat{T})$, $j = 1, \ldots, r-1$, então $z \in \mathrm{Ker}(\bar{m}(T))$ e $T(z) = \hat{T}(z) = \lambda_j z$. Logo $z \in \mathrm{Ker}(\lambda_j I - T)$. Reciprocamente, se $z \in \mathrm{Ker}(\lambda_j I - T)$ então $z \in V$ e $T(z) = \lambda_j z$. Mas $z = w + u$, $u \in \mathrm{Ker}(\lambda_r I - \tilde{T})$ e $w \in \mathrm{Ker}(\bar{m}(T))$. Assim, $T(z) = T(w) + T(u) = \hat{T}(w) + \tilde{T}(u)$ e como, por outro lado, $T(z) = \lambda_j z = \lambda_j(w + u) = \lambda_j w + \lambda_j u$ segue que $\hat{T}(w) = \lambda_j w$ e $\lambda_r u = \tilde{T}(u) = \lambda_j u$. Assim, $(\lambda_r - \lambda_j)u = 0$ e portanto $u = 0$. Logo, $z = w$ e $\hat{T}(z) = T(z) = \lambda_j z$ e daí $z \in \mathrm{Ker}(\lambda_j I - \hat{T})$. □

Teorema 2.13. Sejam $T \in L(V)$ e $m(\lambda) = (\lambda_1 - \lambda)\cdots(\lambda_r - \lambda)$ o polinômio minimal de T. Então T é diagonalizável.

Demonstração. Pelo Teorema 2.12 temos $V = \text{Ker}(\lambda_1 I - T) \oplus \cdots \oplus \text{Ker}(\lambda_r I - T)$. Seja $0 \neq w \in \text{Ker}(\lambda_i I - T)$. Então $(\lambda_i I - T)(w) = 0$ e daí $T(w) = \lambda_i w$. Portanto todo vetor não nulo de $\text{Ker}(\lambda_i I - T)$ é autovetor associado a λ_i. Como a união das bases de $\text{Ker}(\lambda_1 I - T), \ldots, \text{Ker}(\lambda_r I - T)$ é base de V, segue que temos uma base de autovetores para V e portanto T é diagonalizável. □

Teorema 2.14. Seja T um operador linear diagonalizável e sejam $\lambda_1, \ldots, \lambda_r$ autovalores distintos de T. Então o polinômio minimal de T é o polinômio $m(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_r)$.

Demonstração. Precisamos apenas mostrar que $m(T) = 0$. Se v é um autovetor, então um dos operadores $T - \lambda_1 I, T - \lambda_2 I, \ldots, T - \lambda_r I$ aplica v em zero. Portanto $m(T)(v) = (T - \lambda_1 I)(T - \lambda_2 I)\cdots(T - \lambda_r I)(v) = 0$, para todo autovetor v, e como existe uma base de autovetores de T, segue que $m(T) = 0$, e logo é o minimal. □

Exemplo 2.15. Verificamos se $T \in L(\mathbb{R}^3)$ é diagonalizável onde $T(x, y, z) = (7y - 6z, -x + 4y, 2y - 2z)$. Temos $A = [T]_C = \begin{pmatrix} 0 & 7 & -6 \\ -1 & 4 & 0 \\ 0 & 2 & -2 \end{pmatrix}$, onde C denota a base canônica de $\mathbb{R}^3$ e

$$p(\lambda) = \det \begin{pmatrix} -\lambda & 7 & -6 \\ -1 & 4 - \lambda & 0 \\ 0 & 2 & -2 - \lambda \end{pmatrix} = -(\lambda - 1)(\lambda + 1)(\lambda - 2).$$

Logo, $m(\lambda) = (\lambda - 1)(\lambda + 1)(\lambda - 2)$. Portanto esse operador é diagonalizável. Assim existe uma base $\beta = \{v_1, v_2, v_3\}$ de autovetores de T de $\mathbb{R}^3$ tal que $[T]_\beta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, a saber $v_1 = (9, 3, 2)$ associado ao autovalor $\lambda = 1$, $v_2 = (5, 1, 2)$ associado ao autovalor $\lambda = -1$ e $v_3 = (4, 2, 1)$ associado ao autovalor $\lambda = 2$.

Exemplo 2.16. Verificamos se é diagonalizável o operador $T \in L(\mathbb{R}^4)$ cuja matriz em relação à base canônica é $A = \begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$. Temos

$$p(\lambda) = \det \begin{pmatrix} 3 - \lambda & 1 & 0 & 0 \\ 0 & 3 - \lambda & 0 & 0 \\ 0 & 0 & 3 - \lambda & 0 \\ 0 & 0 & 0 & 4 - \lambda \end{pmatrix} = (3 - \lambda)^3(4 - \lambda).$$

Logo, $m(\lambda) = (\lambda - 3)^2(\lambda - 4)$. Portanto esse operador não é diagonalizável.

Exemplo 2.17. Sejam $A = \left(\begin{smallmatrix} 3 & -1 \\ 1 & 1 \end{smallmatrix}\right)$ e $B = \left(\begin{smallmatrix} 1 & -1 \\ 2 & -1 \end{smallmatrix}\right)$. Vejamos se elas são diagonalizáveis, vistas como matrizes sobre $\mathbb{R}$ e também sobre $\mathbb{C}$.

Caso real: $p_A(\lambda) = \det\left(\begin{smallmatrix} 3-\lambda & -1 \\ 1 & 1-\lambda \end{smallmatrix}\right) = (3-\lambda)(1-\lambda) + 1 = (\lambda - 2)^2$. Então $m_A(\lambda) = (\lambda - 2)^2$ ou $m_A(\lambda) = (\lambda - 2)$. Agora, $m_A(A) = A - 2I = \left(\begin{smallmatrix} 1 & -1 \\ 1 & -1 \end{smallmatrix}\right)$. Logo, $m_A(\lambda) = (\lambda - 2)^2$ e A não é diagonalizável em $\mathbb{R}$.

Vejamos de outra forma: para $\lambda = 2$ temos

$$\begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 2\begin{pmatrix} x \\ y \end{pmatrix}.$$

Então $\{3x - y = 2x, x + y = 2y\}$. Portanto, $V(2) = [(1,1)]$ e

$$\dim V(2) = 1 = \text{multiplicidade geométrica} \neq 2 = \text{multiplicidade algébrica.}$$

Ainda, $p_B(\lambda) = \det\left(\begin{smallmatrix} 1-\lambda & -1 \\ 2 & -1-\lambda \end{smallmatrix}\right) = (1-\lambda)(-1-\lambda) + 2 = \lambda^2 + 1$. Portanto B não tem autovalor em $\mathbb{R}$ e daí não é diagonalizável em $\mathbb{R}$.

Caso complexo: $p_A(\lambda) = (\lambda-2)^2$ e $m_A(\lambda) = (\lambda-2)^2$. Ainda, $V(2) = [(1,1)]$ e portanto A não é diagonalizável em $\mathbb{C}$. Também, $p_B(\lambda) = \lambda^2+1 = (\lambda+i)(\lambda-i) = m_B(\lambda)$ e portanto B é diagonalizável em $\mathbb{C}$. De outra forma, para $\lambda = i$ temos

$$\begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = i\begin{pmatrix} x \\ y \end{pmatrix}.$$

Então $\{(1-i)x - y = 0, 2x - (1+i)y = 0\}$, ou equivalentemente, $(1-i)x - y = 0$. Portanto, $V(i) = \{(x,(1-i)x)\} = [(1,1-i)]$ e $\dim V(i) = 1$, donde segue que

$$\text{multiplicidade geométrica} = \text{multiplicidade algébrica.}$$

Para $\lambda = -i$ temos

$$\begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = (-i)\begin{pmatrix} x \\ y \end{pmatrix}.$$

Então $\{(1+i)x - y = 0, 2x + (i-1)y = 0\}$, ou equivalentemente, $(1+i)x - y = 0$. Portanto, $V(-i) = \{(x,(1+i)x)\} = [(1, 1+i)]$ e $\dim V(-i) = 1$, donde segue que

$$\text{multiplicidade geométrica} = \text{multiplicidade algébrica.}$$

Definição 2.18. Dizemos que uma matriz está na *forma diagonal* se é do tipo

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Sugestão 2.19. Dada a matriz $A = \left(\begin{smallmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{smallmatrix}\right)$, encontrar uma matriz inversível P tal que $P^{-1}AP$ é diagonal.

Sugestão 2.20. Para cada matriz

$$A = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & 2 \\ 3 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} 5 & -1 \\ 1 & 3 \end{pmatrix}$$

encontrar, quando possível, matrizes inversíveis P_1, P_2 e P_3 tais que $P_1{}^{-1}AP_1$, $P_2{}^{-1}BP_2$ e $P_3{}^{-1}CP_3$ são diagonais.

Exemplo 2.21. Suponhamos que A seja uma matriz 2×2 com elementos reais e simétrica $(A^t = A)$. Então A é semelhante sobre $\mathbb{R}$ a uma matriz diagonal. De fato, se $A = \left(\begin{smallmatrix} a_{11} & a \\ a & a_{22} \end{smallmatrix}\right)$ então o polinômio característico de A é $p(\lambda) = \lambda^2 + (-a_{11} - a_{22})\lambda + a_{11}a_{22} - a^2$. Se $A = 0$ ou $A = \left(\begin{smallmatrix} a_{11} & 0 \\ 0 & a_{11} \end{smallmatrix}\right)$ o resultado é imediato. Suponhamos A diferente dessas duas matrizes. Calculando o discriminante de $p(\lambda)$ obtemos $\Delta = a_{11}^2 + 2a_{11}a_{22} + a_{22}^2 - 4(a_{11}a_{22} - a^2) = (a_{11} - a_{22})^2 + 4a^2$. Portanto $\Delta > 0$ e assim $p(\lambda)$ tem duas raízes reais e distintas, o que implica que A tem dois autovalores reais e distintos e portanto A é diagonalizável.

Exemplo 2.22. Seja N uma matriz complexa 2×2 tal que $N^2 = 0$. Então $N = 0$ ou N é semelhante sobre $\mathbb{C}$ à matriz $\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$. De fato, se $N \neq 0$ então existe $v \neq 0$ tal que $N(v) \neq 0$ e $N^2(v) = 0$, onde $N : \mathbb{C}^2 \to \mathbb{C}^2$. Afirmamos que $\beta = \{v, N(v)\}$ é base de $\mathbb{C}^2$. Mostremos inicialmente que β é linearmente independente. Com efeito, se $av + bN(v) = 0$, então aplicando N obtemos $aN(v) + bN^2(v) = 0$, ou equivalentemente $aNv = 0$ e portanto $a = 0$, donde segue que $bN(v) = 0$ e consequentemente $b = 0$. Portanto β é de fato linearmente independente. Agora, como $\dim_{\mathbb{C}} \mathbb{C}^2 = 2$, segue que β é base e como $N(v) = 0v + 1N(v)$ e $N(N(v)) = 0 = 0v + 0N(v)$, temos imediatamente que $[N]_\beta = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$.

Exemplo 2.23. Seja $T \in L(\mathbb{R}^4)$ representado, em relação à base ordenada canônica, pela matriz

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \end{pmatrix}.$$

Pergunta-se: em que condições sobre a, b e c , T é diagonalizável?

Temos $p(\lambda) = \lambda^4$. Para ser diagonalizável, $m(\lambda) = \lambda$ e daí deveríamos ter $m(T) = 0$. Logo, $a = b = c = 0$.

Exemplo 2.24. Toda matriz A tal que $A^2 = A$ é semelhante a uma matriz diagonal. De fato, inicialmente observemos que as únicas matrizes diagonais D tais que $D^2 = D$ são matrizes em que os elementos da diagonal são 0 ou 1. Se $A = 0$ ou $A = I$, o resultado é imediato. Suponhamos então que $A \neq 0$ e $A \neq I$. Notemos que $q(x) = x^2 - x$ é um polinômio que anula a matriz A, desde que $q(A) = A^2 - A = 0$. Assim, os possíveis polinômios minimais para A são da forma $m(x) = x + a$ ou $m(x) = x^2 + bx + c$ para alguns a, b e c. Observemos agora que não podemos ter $m(x) = x + a$. De fato, se $m(x) = x + a$, como $m(A) = 0$ então deveríamos ter $A + aI = 0$ e portanto $A = -aI$ donde seguiria que $A = 0$ ou $A = I$, o que é um absurdo. Suponhamos então que $m(x) = x^2 + bx + c$. Logo, como $m(A) = 0$ então deveríamos ter $A^2 + bA + cI = 0$ e como $A^2 = A$, então teríamos $(b+1)A + cI = 0$. Logo, $b + 1 = 0$, senão teríamos $A = \left(\frac{-c}{b+1}\right)I$ e portanto $A = 0$ ou $A = I$, o que é um absurdo. Daí $b = -1$ e $c = 0$, donde segue que $m(x) = x^2 - x$ e portanto A é diagonalizável.

Sugestão 2.25. Mostre que se A é uma matriz tal que $A^2 = A$, então o polinômio característico de A é da forma $p_A(\lambda) = \lambda^m(\lambda - 1)^p$, onde $m + p$ é a ordem da matriz A.

Proposição 2.26. Seja $T \in L(V)$ diagonalizável e W um subespaço T-invariante. Então T_W é diagonalizável.

Demonstração. Temos que T é diagonalizável. Então o polinômio minimal de T se fatora em fatores lineares distintos. Pela Proposição 1.42, o polinômio minimal de T_W divide o de T. E daí também se fatora em fatores lineares distintos. Logo T_W é diagonalizável. □

Exemplo 2.27. É falsa a afirmação abaixo:

> "Se a matriz triangular A for semelhante a uma matriz diagonal, então A já é diagonal".

De fato, vejamos um contraexemplo: seja $A = \left(\begin{smallmatrix} 1 & 2 \\ 0 & 2 \end{smallmatrix}\right)$. Então $p(x) = (1 - x)(2 - x) = m(x)$ e portanto A é diagonalizável e $B = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right) = P^{-1}AP$.

Sugestão 2.28. Seja $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ uma matriz sobre o corpo real $\mathbb{R}$. Encontre condições necessárias e suficientes em a, b, c e d para que A seja diagonalizável.

3

FORMA TRIANGULAR

Neste capítulo veremos como polinômios característicos, minimais e autovalores podem dar condições para um operador linear ser triangulável. Também faremos um estudo sobre operadores nilpotentes.

Definição 3.1. Dizemos que uma matriz está na *forma triangular* se ela tem a forma

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} \quad \text{ou} \quad \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

No primeiro caso dizemos que a matriz é *triangular superior* e no segundo caso, *triangular inferior*.

Propriedades de uma matriz triangular

a) Se nenhum elemento da diagonal principal é 0, então A é inversível.

b) Se um elemento da diagonal principal é 0, então A é não inversível.

c) Os autovalores de A são exatamente os elementos de sua diagonal principal.

d) Se A é uma matriz quadrada de ordem n e todos os elementos de sua diagonal principal são nulos, então $A^n = 0$.

e) O determinante de A, $\det A$, é dado pelo produto dos elementos de sua diagonal principal (no determinante consideramos sempre um produto de n elementos tal que um e só um elemento provém de cada linha e um e só um elemento provém de cada coluna).

Definição 3.2. O operador linear T se diz *triangulável* se existir uma base em relação à qual T seja representado por uma matriz triangular.

Teorema 3.3. Se T é triangulável, então o polinômio característico de T tem a forma: $p(x) = (x - c_1)^{d_1} \cdots (x - c_r)^{d_r}$, $c_i \in K$. Assim o polinômio minimal de T tem a forma: $m(x) = (x - c_1)^{e_1} \cdots (x - c_r)^{e_r}$, com $c_i \in K$, $e_i \leq d_i$, $i = 1, \ldots, r$.

Demonstração. Como T é triangulável então é semelhante a uma matriz triangular B. Como matrizes semelhantes possuem o mesmo polinômio característico, então $\det(T - xI) = \det(B - xI)$ e das propriedades de matriz triangular sabemos que $\det(B - xI)$ é igual ao produto dos elementos da diagonal da matriz $(B - xI)$. Logo, $p(x) = (x - c_1)^{d_1} \cdots (x - c_r)^{d_r}$, com $c_i \in K$. Como o polinômio minimal divide o polinômio característico, então $m(x) = (x - c_1)^{e_1} \cdots (x - c_r)^{e_r}$, $c_i \in K$, $e_i \leq d_i$, $i = 1, \ldots, r$. □

Teorema 3.4. Seja $T \in L(V)$ com todos os seus autovalores distintos $c_1, \ldots, c_r$ em K e seu polinômio característico $p(x) = (x - c_1)^{d_1} \cdots (x - c_r)^{d_r}$. Então T é triangulável.

Demonstração. A demonstração será feita por indução na dimensão de V. Faremos inicialmente para os casos de dimensão 1 e 2, e depois para dimensão n.

Caso $\dim V = 1$**.** Neste caso, cada representação matricial de T é uma matriz 1×1, que é triangular.

Caso $\dim V = 2$**.** Neste caso, temos $p(x) = (x - c_1)(x - c_2)$, $c_1 \neq c_2$ ou $p(x) = (x - c_1)^2$. Se $p(x) = (x - c_1)(x - c_2)$, $c_1 \neq c_2$, então T é diagonalizável e daí triangulável. Se $p(x) = (x - c_1)^2$, então temos dois candidatos a polinômio minimal: $m_1(x) = (x - c_1)$ e $m_2(x) = p(x)$. Se for $m_1(x)$, então é diagonalizável e daí triangular. Suponhamos agora que seja $m_2(x)$. Existe $0 \neq v \in V$ tal que $T(v) = c_1 v$. Seja $\bar{V} = V/[v]$ com $\dim \bar{V} = \dim V - \dim[v] = 1$. Seja $\{v_1 + [v]\}$ base de $\bar{V}$. Assim $\{v, v_1\}$ é base de V. Temos $\bar{T} : \bar{V} \to \bar{V}$ dada por $\bar{T}(u + [v]) = T(u) + [v]$. Vejamos qual a matriz de $\bar{T}$ em relação a esta base. Temos $\bar{T}(v_1 + [v]) = T(v_1) + [v]$. Um representante seria: $T(v_1) + av = bv + dv_1 + av = dv_1 + (b + a)v$. Logo $\bar{T}(v_1 + [v]) = dv_1 + [v] = d(v_1 + [v])$ e portanto $[\bar{T}] = [d]$. Assim, d é autovalor de $\bar{T}$ e daí $\bar{p}(x) = (x - d)$. Como $\bar{p} \mid p$, então $d = c_1$ e assim $\bar{T}(v_1 + [v]) = c_1 v_1 + [v]$ e portanto, $T(v_1) = c_1 v_1 + \alpha v$. Logo, $[T] = \begin{pmatrix} c_1 & \alpha \\ 0 & c_1 \end{pmatrix}$ e T é triangular.

Caso $\dim V = n > 2$. Suponha que o teorema vale para espaços de dimensão menor que n. Como o polinômio característico de T se fatora em polinômios lineares, T tem pelo menos um autovalor e, portanto, pelo menos um autovetor não nulo v, digamos, ou seja, $T(v) = a_{11}v$. Seja $W = [v]$. Então W é T-invariante. Faça $\bar{V} = V/W$. Então $\dim \bar{V} = n - 1$. Seja $\bar{T}(u + W) = T(u) + W$. Sabemos que $\bar{p} \mid p$ e também $\bar{m} \mid m$. Assim $\bar{V}$ e $\bar{T}$ satisfazem as hipóteses do teorema. Portanto, por indução, existe uma base $\{v_2 + W, \ldots, v_n + W\}$ de $\bar{V}$ tal que $\bar{T}(v_2 + W) = a_{22}(v_2 + W)$, $\bar{T}(v_3 + W) = a_{23}(v_2 + W) + a_{33}(v_3 + W)$, ..., $\bar{T}(v_n + W) = a_{2n}(v_2 + W) + \cdots + a_{nn}(v_n + W)$. Então $\{v, v_2, \ldots, v_n\}$ é base de V. Como $\bar{T}(v_2 + W) = a_{22}(v_2 + W)$, temos $\bar{T}(v_2 + W) - a_{22}(v_2 + W) = 0 + W$ e portanto $T(v_2) + W - a_{22}v_2 + W = 0 + W$ e daí $T(v_2) - a_{22}v_2 \in W$. Logo, $T(v_2) - a_{22}v_2 = a_{12}v$, ou equivalentemente, $T(v_2) = a_{12}v + a_{22}v_2$. Analogamente, para $i = 3, \ldots, n$, $T(v_i) - a_{2i}v_2 - \cdots - a_{ii}v_i \in W$, logo $T(v_i) = a_{1i}v + a_{2i}v_2 + \cdots + a_{ii}v_i$. Assim, $T(v) = a_{11}v$, $T(v_2) = a_{12}v + a_{22}v_2$, ..., $T(v_n) = a_{1n}v + a_{2n}v_2 + \cdots + a_{nn}v_n$ e portanto a matriz de T nessa base é triangular. E observando essa matriz triangular, os elementos da diagonal $a_{11}, \ldots, a_{nn}$ são os autovalores c_i repetidos d_i vezes. □

Corolário 3.5. Se $T \in L(V)$ tem todos os autovalores em K então T é triangulável.

Exemplo 3.6. Verificamos se $A = \begin{pmatrix} -3 & 1 & -1 \\ -7 & 5 & -1 \\ -6 & 6 & -2 \end{pmatrix}$ é triangulável. De fato, temos $p_A(x) = (x - 4)(x + 2)^2$ e portanto A tem todos os seus autovalores em $\mathbb{R}$, donde segue que A é triangulável. Temos $m_A(x) = (x - 4)(x + 2)^2$ e portanto A não é diagonalizável. Para o autovalor $x = 4$, $v = (0, 1, 1)$ é autovetor. Consideremos $W = [(0, 1, 1)]$ e $\bar{V} = \mathbb{R}^3/W$. Então $\dim \bar{V} = 2$. Afirmamos que $\{(1, 1, 0) + W, (0, 1, 0) + W\}$ é base de $\bar{V}$. Precisamos mostrar somente que os elementos desse conjunto são linearmente independentes. Com efeito, se $a((1, 1, 0) + W) + b((0, 1, 0) + W) = W$ então $a(1, 1, 0) + W + b(0, 1, 0) + W = W$ e portanto $(a, a + b, 0) + W = W$. Assim, $(a, a + b, 0) = x(0, 1, 1)$, ou equivalentemente, $\{a = 0, a + b = x, 0 = x\}$. Logo, $a = b = 0$ e daí são linearmente independentes.

Consideremos então a seguinte base para $\mathbb{R}^3$: $\{(0,1,1),(1,1,0),(0,1,0)\}$. Então,

$$
\begin{aligned}
A(0,1,1) &= (0,4,4) = 4(0,1,1) + 0(1,1,0) + 0(0,1,0),\\
A(1,1,0) &= (-2,-2,0) = 0(0,1,1) - 2(1,1,0) + 0(0,1,0),\\
A(0,1,0) &= (1,5,6) = 6(0,1,1) + 1(1,1,0) - 2(0,1,0),
\end{aligned}
$$

e logo $B = \begin{pmatrix} 4 & 0 & 6 \\ 0 & -2 & 1 \\ 0 & 0 & -2 \end{pmatrix}$. Se $P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, $P^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ -1 & 1 & -1 \end{pmatrix}$ e daí obtemos $P^{-1}AP = B$.

Sugestão 3.7. Achar uma base triangular para as aplicações de $\mathbb{C}^2$ representadas pelas matrizes

$$
A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}; \qquad B = \begin{pmatrix} 1 & i \\ 1 & i \end{pmatrix}; \qquad C = \begin{pmatrix} 1 & 2 \\ i & i \end{pmatrix}.
$$

3.1. OPERADORES NILPOTENTES

Definição 3.8. Dizemos que $T \in L(V)$ é *nilpotente* se $T^m = 0$ para algum inteiro $m \geq 1$. Uma matriz quadrada A é dita nilpotente se existir um inteiro $m \geq 1$ tal que $A^m = 0$.

Exemplo 3.9. Seja $T \in L(\mathbb{R}^2)$ dado por $T(x,y) = (2y,0)$. Então, T é nilpotente. De fato, temos $T^2(x,y) = T(T(x,y)) = T(2y,0) = (0,0)$.

Exemplo 3.10. A matriz $A = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}$ é nilpotente. De fato, temos $A^2 = \begin{pmatrix} 0 & 0 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ e portanto $A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

Sugestão 3.11. Suponha que S e T são operadores nilpotentes que comutam entre si. Então $S + T$ e ST também são nilpotentes.

Proposição 3.12. Se $T \in L(V)$ é nilpotente, então todos os autovalores de T são iguais a zero.

Demonstração. Suponha que $T^m = 0$ e que existe $0 \neq v \in V$ tal que $T(v) = \lambda v$. Então $T^2(v) = T(T(v)) = T(\lambda v) = \lambda T(v) = \lambda(\lambda v) = \lambda^2 v$. Suponhamos agora que $T^{m-1}(v) = \lambda^{m-1} v$ e mostremos que $T^m(v) = \lambda^m v$. Com efeito, temos

$T^m(v) = T(T^{m-1}(v)) = T(\lambda^{m-1}v) = \lambda^{m-1}T(v) = \lambda^{m-1}(\lambda v) = \lambda^m v$. Portanto, $\lambda^m v = 0$ e daí $\lambda = 0$. □

Proposição 3.13. Sejam $\dim V = n$ e $T \in L(V)$. Se todos os autovalores de T são iguais a zero, então T é nilpotente.

Demonstração. Temos que o polinômio característico de T é $p(x) = x^n$ e portanto, pelo teorema de Cayley–Hamilton, $p(T) = T^n = 0$. Logo, T é nilpotente. □

Definição 3.14. Se $T \in L(V)$ (ou $A \in M_n$) é nilpotente, então o número k é dito *índice de nilpotência* de T (ou de A) se $T^k = 0$ (ou $A^k = 0$) mas $T^{k-1} \neq 0$ (ou $A^{k-1} \neq 0$).

Nos Exemplos 3.9 e 3.10, respectivamente, vimos que T tem índice de nilpotência 2 e A tem índice de nilpotência 3.

Proposição 3.15. Se $T \in L(V)$ é nilpotente de índice k, então $x = 0$ é autovalor de T.

Demonstração. Temos que existe $0 \neq v \in V$ tal que $T^{k-1}(v) \neq 0$. Portanto, $T(T^{k-1}(v)) = T^k(v) = 0 = 0T^{k-1}(v)$ e assim 0 é autovalor de T. □

Proposição 3.16. Sejam $\dim V = n$ e $T \in L(V)$ nilpotente. Então o polinômio característico de T é x^n.

Demonstração. Suponha que $T^m = 0$ e considere o seguinte polinômio: $f(x) = x^m$. Então $f(T) = T^m = 0$ e portanto T é raiz de f. Logo $m(x) \mid f(x)$, de onde segue que $m(x) = x^r$ e daí $p_T(x) = x^{\dim V} = x^n$. □

Teorema 3.17. Se $T \in L(V)$ é nilpotente então $f(T) = a_0 I + a_1 T + \cdots + a_r T^r$ é inversível, desde que $a_0 \neq 0$.

Demonstração. Suponha que $T^m = 0$ e considere $S = a_1 T + a_2 T^2 + \cdots + a_r T^r$. Então $S^m = 0$ e

$$(a_0I+S)\left(\frac{I}{a_0}-\frac{S}{a_0^2}+\frac{S^2}{a_0^3}+\cdots+(-1)^{m-2}\frac{S^{m-2}}{a_0^{m-1}}+(-1)^{m-1}\frac{S^{m-1}}{a_0^m}\right)=$$

$$I-\frac{S}{a_0}+\frac{S^2}{a_0^2}+\cdots+(-1)^{m-2}\frac{S^{m-2}}{a_0^{m-2}}+(-1)^{m-1}\frac{S^{m-1}}{a_0^{m-1}}+$$

$$+\frac{S}{a_0}-\frac{S^2}{a_0^2}+\cdots+(-1)^{m-3}\frac{S^{m-2}}{a_0^{m-2}}+(-1)^{m-2}\frac{S^{m-1}}{a_0^{m-1}}+(-1)^{m-1}\frac{S^m}{a_0^m}=I.$$

Assim, para $a_0 \neq 0$, $f(T) = a_0I + S$ é inversível. □

Teorema 3.18. Se $T \in L(V)$ tem índice de nilpotência k e $T^{k-1}(v) \neq 0$ então $\{T^{k-1}(v), \dots, T^2(v), T(v), v\}$ é linearmente independente.

Demonstração. Considere $f(T) = a_{k-1}T^{k-1} + \cdots + a_1T + a_0I$. Seja $v \neq 0$ e $f(T)(v) = 0$. Então $f(T)$ não é inversível e portanto $a_0 = 0$. Reescrevendo, obtemos $(a_1I+a_2T+\cdots+a_{k-1}T^{k-2})(T(v)) = 0$ e daí $a_1I+a_2T+\cdots+a_{k-1}T^{k-2}$ não é inversível e portanto $a_1 = 0$. Semelhantemente, obtemos $a_2 = \cdots = a_{k-1} = 0$ e o teorema segue imediatamente. □

Teorema 3.19. Sejam $\dim V = n$ e $T \in L(V)$ de índice de nilpotência n. Então, se $0 \neq T^{n-1}(v) \in V$, $\{T^{n-1}(v), \dots, T^2(v), T(v), v\}$ é base de V e a matriz de T em relação a essa base tem a forma

$$\begin{pmatrix} 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Demonstração. Pelo Teorema 3.18, $\{T^{n-1}(v), \dots, T^2(v), T(v), v\}$ é linearmente independente e, como tem n elementos, é uma base de V.

Ainda, temos:

$$T(T^{n-1}(v)) = T^n(v) = 0 = 0T^{n-1}(v) + 0T^{n-2}(v) + \cdots + 0T(v) + 0v;$$

$$\begin{aligned} T(T^{n-2}(v)) &= T^{n-1}(v) = 1T^{n-1}(v) + 0T^{n-2}(v) + \cdots + 0T(v) + 0v;\\ &\vdots\\ T(T(v)) &= T^2(v) = 0T^{n-1}(v) + \cdots + 1T^2(v) + 0T(v) + 0v;\\ T(v) &= 0T^{n-1}(v) + \cdots + 0T^2(v) + 1T(v) + 0v, \end{aligned}$$

e daí a matriz de T em relação à base acima é, de fato, a matriz enunciada. □

Teorema 3.20. Se $T \in L(V)$ é nilpotente de índice de nilpotência k, então pode-se determinar uma base de V tal que a matriz de T em relação a essa base tenha a forma $\begin{pmatrix} M_k & A_1 \\ 0 & A_2 \end{pmatrix}$, com

$$M_k = \begin{pmatrix} 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}_{k\times k}.$$

Demonstração. Seja $T^{k-1}(v) \neq 0$. Pelo Teorema 3.18, $\{T^{k-1}(v), \ldots, T(v), v\}$ é linearmente independente. Completando-o para uma base de V,

$$\{T^{k-1}(v), \ldots, T(v), v, v_{k+1}, \ldots, v_n\},$$

a matriz de T tem a forma acima. □

Teorema 3.21. Se $T \in L(V)$ é nilpotente de índice de nilpotência k então $W = [T^{k-1}(v), \ldots, T(v), v]$ é T-invariante.

Demonstração. Seja $w \in W$. Então $w = a_{k-1}T^{k-1}(v) + \cdots + a_1T(v) + a_0v$. Portanto, $T(w) = a_{k-1}T^k(v) + a_{k-2}T^{k-1}(v) + \cdots + a_1T^2(v) + a_0T(v) = a_{k-2}T^{k-1}(v) + \cdots + a_1T^2(v) + a_0T(v)$ e daí $T(w) \in W$. □

Lema 3.22. Se $T \in L(V)$ é nilpotente de índice de nilpotência k e $u \in W = [T^{k-1}(v), \ldots, T(v), v]$ é tal que $T^{k-r}(u) = 0$, onde $0 < r \leq k$, então $T^r(u_0) = u$, para algum $u_0 \in W$.

Demonstração. O caso $r = k$ é imediato. Agora, se $0 < r < k$ e $u \in W$, então

$$u = a_{k-1}T^{k-1}(v) + \cdots + a_r T^r(v) + a_{r-1}T^{r-1}(v) + \cdots + a_1 T(v) + a_0 v.$$

Assim, aplicando T^{k-r} em u obtemos $0 = T^{k-r}(u) = a_{r-1}T^{k-1}(v) + \cdots + a_0 T^{k-r}(v)$. No entanto, $\{T^{k-1}(v), \ldots, T^{k-r}(v)\}$ é linearmente independente, donde segue que $a_0 = \cdots = a_{r-1} = 0$ e então, $u = a_{k-1}T^{k-1}(v) + \cdots + a_r T^r(v)$. Seja $u_0 = a_{k-1}T^{k-r-1}(v) + \cdots + a_r v \in W$. Então $T^r(u_0) = u$. □

Teorema 3.23. Se $T \in L(V)$ é nilpotente de índice de nilpotência k, então existe um subespaço W_1 de V, invariante sob T, tal que $V = W \oplus W_1$, onde $W = \left[T^{k-1}(v), \ldots, T(v), v\right]$ com $T^{k-1}(v) \neq 0$ e $T^k = 0$.

Demonstração. Consideremos W_1 o maior subespaço de V tal que $W \cap W_1 = \{0\}$ e W_1 é invariante sob T. Observemos que existe pelo menos um subespaço de V satisfazendo as propriedades acima, a saber, o subespaço nulo. Afirmamos que $V = W + W_1$.

De fato, suponha que $V \neq W + W_1$. Então existe $z \in V$ tal que $z \notin W + W_1$. Desde que $T^k = 0$ então existe um inteiro r, $0 < r \leq k$, tal que $T^r(z) \in W + W_1$ e $T^i(z) \notin W + W_1$, para $i < r$. Portanto, $T^r(z) = u + w_1$ onde $u \in W$ e $w_1 \in W_1$.

Temos que $0 = T^k(z) = T^{k-r}(T^r(z)) = T^{k-r}(u) + T^{k-r}(w_1)$. Como W e W_1 são T-invariantes então $T^{k-r}(u) \in W$ e $T^{k-r}(w_1) \in W_1$. Portanto, $T^{k-r}(u) = -T^{k-r}(w_1) \in W \cap W_1 = \{0\}$ e daí $T^{k-r}(u) = 0$. Logo, pelo Lema 3.22, $T^r(u_0) = u$, para algum $u_0 \in W$. Portanto, $T^r(z) = u + w_1 = T^r(u_0) + w_1$.

Seja $z_1 = z - u_0$. Então $T^r(z_1) = T^r(z) - T^r(u_0) = w_1 \in W_1$. Agora, como W_1 é T-invariante então $T^m(z_1) \in W_1$, para todo $m \geq r$. Afirmamos agora que se $i < r$ então $T^i(z_1) = T^i(z) - T^i(u_0) \notin W + W_1$, pois caso contrário $T^i(z) \in W + W_1$, contradizendo a escolha de r.

Seja $\tilde{W} = \left[W_1 \cup \{z_1, T(z_1), \ldots, T^{r-1}(z_1)\}\right]$. Desde que $z_1 \notin W_1$ e $\tilde{W} \supset W_1$, $\dim \tilde{W} > \dim W_1$. Ainda, como $T^r(z_1) \in W_1$ e W_1 é T-invariante temos que $\tilde{W}$ é T-invariante. Pela natureza maximal de W_1 deve existir um elemento da forma $w_0 + \alpha_1 z_1 + \alpha_2 T(z_1) + \cdots + \alpha_r T^{r-1}(z_1) \neq 0 \in \tilde{W} \cap W$ onde $w_0 \in W_1$.

Afirmamos que $\alpha_1, \ldots, \alpha_r$ não podem ser todos nulos, senão $0 \neq w_0 \in W_1 \cap W = \{0\}$, uma contradição.

Seja α_s o primeiro α não nulo; então

$$w_0 + \alpha_s T^{s-1}(z_1) + \alpha_{s+1} T^s(z_1) + \cdots + \alpha_r T^{r-s+s-1}(z_1) =$$
$$w_0 + (\alpha_s I + \alpha_{s+1} T + \cdots + \alpha_r T^{r-s})(T^{s-1}(z_1)) \in W.$$

Desde que $\alpha_s \neq 0$, pelo Teorema 3.17, $\alpha_s I + \alpha_{s+1} T + \cdots + \alpha_r T^{r-s}$ é inversível e seu inverso, R, é um polinômio em T. Portanto, W_1 e W são invariantes por R; logo, $R(w_0) + T^{s-1}(z_1) \in R(W) \subset W$ e daí $T^{s-1}(z_1) \in W + R(W_1) \subset W + W_1$.

Desde que $s - 1 < r$, isto é impossível e portanto $W + W_1 = V$. Como $W \cap W_1 = \{0\}$ então $V = W \oplus W_1$. □

Teorema 3.24. Se $T \in L(V)$, então existem subespaços W e U invariantes por T tais que $V = W \oplus U$, com $T_W : W \to W$ nilpotente e $T_U : U \to U$ inversível.

Demonstração. Temos que $\mathrm{Ker}(T) \subset \mathrm{Ker}(T^2) \subset \cdots \subset V$. Como V é de dimensão finita e $\mathrm{Ker}(T^j)$ é subespaço, estas inclusões não podem ser todas próprias, logo existe um inteiro k tal que $\mathrm{Ker}(T^k) = \mathrm{Ker}(T^{k+1})$. Usando indução sobre j, concluímos que $\mathrm{Ker}(T^k) = \mathrm{Ker}(T^{k+j})$, $j = 1, 2, \ldots$. De fato, já observamos a validade para $j = 1$. Suponhamos por indução, que valha para $j - 1$, isto é, que $\mathrm{Ker}(T^k) = \mathrm{Ker}(T^{k+j-1})$, e provemos para j. Para isto, basta mostrarmos que $\mathrm{Ker}(T^{k+j}) \subset \mathrm{Ker}(T^k)$. Com efeito, seja $y \in \mathrm{Ker}(T^{k+j})$. Então $T^{k+j}(y) = 0$ e portanto $T^{k+j-1}(T(y)) = 0$, donde segue que $T(y) \in \mathrm{Ker}(T^{k+j-1}) = \mathrm{Ker}(T^k)$. Logo, $T^{k+1}(y) = 0$ e portanto $y \in \mathrm{Ker}(T^{k+1}) = \mathrm{Ker}(T^k)$. Afirmamos que $W = \mathrm{Ker}(T^k)$ é um dos subespaços procurados. De fato, se $w \in W = \mathrm{Ker}(T^k)$, então $T^k(w) = 0$ e daí $T^k(T(w)) = T(T^k(w)) = 0$, donde $T(w) \in W$. Logo W é T-invariante. Assim podemos definir $T_W : W \to W$ por $T_W(w) = T(w)$. Para $w \in W$, $T_W^k(w) = T^k(w) = 0$. Logo T_W é nilpotente e de índice k. Seja $U = \mathrm{Im}(T^k)$. Então $W \cap U = \{0\}$, pois se $w \in W \cap U$, temos que $T^k(w) = 0$ e existe $v \in V$ tal que $T^k(v) = w$. Logo $T^k(T^k(v)) = 0$. Assim, $v \in \mathrm{Ker}(T^{2k}) = \mathrm{Ker}(T^k)$ e daí $T^k(v) = 0$, o que implica $w = 0$. Como $\dim V = \dim \mathrm{Ker}(T^k) + \dim \mathrm{Im}(T^k)$, temos que $V = W \oplus U$. Que W é T-invariante e T_W é nilpotente já vimos acima. É claro que U é T-invariante. Vejamos que $T_U : U \to U$ é inversível. Seja $u \in U$ tal que $T_U(u) = T(u) = 0$. Então $T^{k-1}(T(u)) = 0$ e daí $u \in \mathrm{Ker}(T^k)$. Logo $u = 0$ e portanto T_U é injetiva.

Que T_U é sobrejetora decorre da própria definição. $\square$

Exemplo 3.25. Verificamos o Teorema 3.24 para o operador linear $T(x, y, z) = (-x + y - z, -7x + 7y - z, -6x + 6y)$. Com efeito, temos:

$$\begin{aligned}\operatorname{Ker}(T) &= \{(x, y, z) \mid (-x + y - z, -7x + 7y - z, -6x + 6y) = (0, 0, 0)\} \\ &= \{(x, x, 0) \mid x \in \mathbb{R}\} = [(1, 1, 0)].\end{aligned}$$

$$\begin{aligned}\operatorname{Ker}(T^2) &= \{(x, y, z) \mid (0, -36x + 36y, -36x + 36y) = (0, 0, 0)\} \\ &= \{(x, x, z) \mid x, z \in \mathbb{R}\} = [(1, 1, 0), (0, 0, 1)].\end{aligned}$$

$$\begin{aligned}\operatorname{Ker}(T^3) &= \{(x, y, z) \mid (0, -216x + 216y, -216x + 216y) = (0, 0, 0)\} \\ &= \{(x, x, z) \mid x, z \in \mathbb{R}\} = [(1, 1, 0), (0, 0, 1)].\end{aligned}$$

Assim, $\operatorname{Ker}(T^2) = \operatorname{Ker}(T^3) = \operatorname{Ker}(T^4) = \cdots$. Seja $W = \operatorname{Ker}(T^2)$. É claro que é T-invariante. Temos que $T_W : W \to W$ é dada por $T_W(x, x, z) = T(x, x, z) = (-z, -z, 0)$ e $T_W^2(x, x, z) = T_W(-z, -z, 0) = T(-z, -z, 0) = (0, 0, 0)$. Logo T_W é nilpotente de índice 2. Como $W = \operatorname{Ker}(T^2) = [(1, 1, 0), (0, 0, 1)]$ e $U = \operatorname{Im}(T^2) = \{(0, -36x + 36y, -36x + 36y) \mid x, y \in \mathbb{R}\} = [(0, 1, 1)]$, temos que $W \cap U = \{(0, 0, 0)\}$. Ainda, como $\dim W = 2$ e $\dim U = 1$, temos que $W \oplus U = \mathbb{R}^3$. Considere agora $T_U : U \to U$. Assim, $T_U(0, x, x) = (0, 6x, 6x) = (0, 0, 0)$ se, e somente se, $x = 0$. Logo, $\operatorname{Ker}(T_U) = \{(0, 0, 0)\}$ e daí é injetora. É sobrejetora pois $T_U\left(0, \frac{x}{6}, \frac{x}{6}\right) = (0, x, x)$. Logo é inversível.

Sugestão 3.26. Verifique o Teorema 3.24 para o operador linear $T(x, y, z) = (y, z, 0)$.

Exemplo 3.27. Para o operador linear $T(x, y, z) = (-3x + y - z, -7x + 5y - z, -6x + 6y - 2z)$, vamos:

i) Calcular os polinômios característico e minimal;

ii) Calcular as multiplicidades algébrica e geométrica de cada autovalor λ_i;

iii) Para cada autovalor λ_i, encontrar os subespaços $H_i = \operatorname{Ker}((T - \lambda_i I)^{k_i})$, onde k_i é o primeiro inteiro positivo tal que $\operatorname{Ker}((T - \lambda_i I)^{k_i}) = \operatorname{Ker}((T - \lambda_i I)^{k_i + 1})$ e $\dim H_i = k_i$;

iv) H_i é invariante por $(T - \lambda_i I)$?

v) Provar que $(T - \lambda_i I)_{H_i}$ é nilpotente;

vi) Determinar a matriz de $(T - \lambda_i I)_{H_i}$ em relação à base

$$\left\{(T - \lambda_i I)^{k_i - 1}(u), \ldots, (T - \lambda_i I)(u), u\right\},$$

onde u é escolhido tal que $(T - \lambda_i I)^{k_i}(u) = 0$ e $(T - \lambda_i I)^{k_i - 1}(u) \neq 0$.

O polinômio característico é $p_T(\lambda) = -\lambda^3 + 12\lambda + 16 = (\lambda - 4)(\lambda + 2)^2$. Assim, os autovalores são $\lambda_1 = 4$ com multiplicidade algébrica igual a 1 e $\lambda_2 = -2$ com multiplicidade algébrica igual a 2.

O polinômio minimal é $m_T(\lambda) = (\lambda - 4)(\lambda + 2)^2 = p_T(\lambda)$. Temos $\mathrm{Ker}(T - 4I) = \{(0, y, y) \mid y \in \mathbb{R}\}$ e $\mathrm{Ker}(T + 2I) = \{(x, x, 0) \mid x \in \mathbb{R}\}$. Então a multiplicidade geométrica de $\lambda_1 = 4$ é $1 = \dim \mathrm{Ker}(T - 4I)$ e a de $\lambda_2 = -2$ é $1 = \dim \mathrm{Ker}(T + 2I)$. Ainda, $\mathrm{Ker}(T - 4I) = \mathrm{Ker}((T - 4I)^2)$ e assim $k_1 = 1$ e como $\mathrm{Ker}((T + 2I)^2) = \mathrm{Ker}((T + 2I)^3)$ temos $k_2 = 2$. Logo, $H_1 = \mathrm{Ker}(T - 4I)$ e $H_2 = \mathrm{Ker}((T + 2I)^2)$.

Agora, $(T - 4I)(0, y, y) = (0, 0, 0)$, e portanto $(T - 4I)_{H_1}$ é nilpotente de índice 1 e como $(T + 2I)(x, x, z) = (-z, -z, 0)$, segue que $(T + 2I)^2(x, x, z) = (0, 0, 0)$ e daí $(T + 2I)_{H_2}$ é nilpotente de índice 2.

Esses dois subespaços são invariantes. Seja $0 \neq w \in H_1$, assim $\epsilon_1 = \{w\}$ é base de H_1. Como $(T - 4I)(w) = 0$ temos que $[T - 4I]_{\epsilon_1} = [0]$. Seja $u \in H_2$ tal que $(T + 2I)(u) \neq 0$. Então $\epsilon_2 = \{(T + 2I)(u), u\}$ é base de H_2. Como $u \in H_2$ temos $(T + 2I)^2(u) = 0$. Seja $u_1 = (T + 2I)(u)$, assim $(T + 2I)(u_1) = 0 = 0u_1 + 0u$ e $(T + 2I)(u) = u_1 = 1u_1 + 0u$ e daí

$$[T + 2I]_{\epsilon_2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Responderemos agora a mais três questões: Quais seriam as matrizes $[T_{H_1}]_{\epsilon_1}$, $[T_{H_2}]_{\epsilon_2}$ e $[T]_{\epsilon_1 \cup \epsilon_2}$?

Sabemos que $(T - 4I)(w) = 0$, ou seja, que $T(w) = 4w$. Logo, $[T_{H_1}]_{\epsilon_1} = [4]$. Por outro lado, de $(T + 2I)(u_1) = 0$ segue que $T(u_1) = -2u_1 + 0u$ e de

$(T+2I)(u) = u_1$ segue que $T(u) = u_1 - 2u$. Portanto,

$$[T_{H_2}]_{\epsilon_2} = \begin{pmatrix} -2 & 1 \\ 0 & -2 \end{pmatrix}.$$

Temos que $\epsilon_1 \cup \epsilon_2 = \{w, (T+2I)(u), u\}$. Então, de $(T-4I)(w) = 0$ segue que $T(w) = 4w = 4w + 0u_1 + 0u$; de $(T+2I)(u_1) = 0$ segue que $T(u_1) = -2u_1 = 0w - 2u_1 + 0u$; e de $(T+2I)(u) = u_1$ segue que $T(u) = u_1 - 2u = 0w + u_1 - 2u$.

Assim

$$[T]_{\epsilon_1 \cup \epsilon_2} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & -2 \end{pmatrix}.$$

Teorema 3.28. Se $T \in L(V)$ é nilpotente de índice de nilpotência n_1 então pode-se determinar uma base de V tal que a matriz de T em relação a esta base tenha a forma abaixo, com $n_1 \geq n_2 \geq \cdots \geq n_r$ e $n_1 + n_2 + \cdots + n_r = \dim V$.

$$\begin{pmatrix} \underbrace{\begin{matrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{matrix}}_{n_1 \times n_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \underbrace{\begin{matrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{matrix}}_{n_2 \times n_2} \quad \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \underbrace{\begin{matrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{matrix}}_{n_r \times n_r} \end{pmatrix}$$

Demonstração. Pelo Teorema 3.23, $V = W \oplus W_1$, onde W_1 é invariante por T. Associando o Teorema 4.9, pode-se determinar uma base na qual a matriz de T será $\begin{pmatrix} M_{n_1} & 0 \\ 0 & A_2 \end{pmatrix}$, onde A_2 é a matriz de T_2, a transformação linear induzida sobre W_1. Assim, temos $T_2 : W_1 \to W_1$ linear com $T_2(W_1) \subset W_1$ e $T_2(w_1) = T(w_1)$.

Como $T^{n_1} = 0$, $T_2^r = 0$, para algum $r \leq n_1$. Seja $n_2 \leq r$ tal que $T_2^{n_2} = 0$ e $T_2^{n_2-1} \neq 0$. Logo T_2 é nilpotente, de índice de nilpotência n_2. Repetindo para T_2 sobre W_1 o argumento usado para T sobre V, podemos decompor W_1 da mesma forma que fizemos com V (ou recorrer a uma indução sobre a dimensão do espaço vetorial considerado). Continuando desta forma, obtemos uma base de V em relação a qual a matriz de T é da forma

$$\begin{pmatrix} M_{n_1} & 0 & \cdots & 0 \\ 0 & M_{n_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & M_{n_r} \end{pmatrix}$$

e é evidente que $n_1 + n_2 + \cdots + n_r = \dim V$, pois a dimensão da matriz é $n \times n$, onde $n = \dim V$. □

Exemplo 3.29. Verificamos o Teorema 3.28 para o operador linear $T \in L(\mathbb{R}^6)$ definido por $T(x_1, x_2, x_3, x_4, x_5, x_6) = (0, 3x_1, 8x_1 + 3x_2, 5x_5, 0, 0)$.

Temos

$$\begin{aligned} T^2(x_1, x_2, x_3, x_4, x_5, x_6) &= T(T(x_1, x_2, x_3, x_4, x_5, x_6)) \\ &= T(0, 3x_1, 8x_1 + 3x_2, 5x_5, 0, 0) \\ &= (0, 0, 9x_1, 0, 0, 0), \\ T^3(x_1, x_2, x_3, x_4, x_5, x_6) &= (0, 0, 0, 0, 0, 0), \end{aligned}$$

donde segue que T é nilpotente, de índice $n_1 = 3$.

Seja $v = (1, 0, 0, 0, 0, 0) \in \mathbb{R}^6$. Então, $\{T^2(v), T(v), v\}$ é base de $W = [T^2(v), T(v), v]$, onde temos $T(v) = (0, 3, 8, 0, 0, 0)$ e $T^2(v) = (0, 0, 9, 0, 0, 0)$.

Como $T_1 : W \to W$ é dada por

$$T_1(c, 3b, 9a + 8b, 0, 0, 0) = T(c, 3b, 9a + 8b, 0, 0, 0) = (0, 3c, 8b + 9a, 0, 0, 0),$$

segue que

$$T_1(0, 0, 9, 0, 0, 0) = (0, 0, 0, 0, 0, 0) = 0T^2(v) + 0T(v) + 0v,$$

$$T_1(0,3,8,0,0,0) = (0,0,9,0,0,0) = 1T^2(v) + 0T(v) + 0v,$$
$$T_1(1,0,0,0,0,0) = (0,3,8,0,0,0) = 0T^2(v) + 1T(v) + 0v.$$

Portanto $[T_1] = \left(\begin{smallmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{smallmatrix}\right)$. Assim, $\mathbb{R}^6 = W \oplus W_1$, com

$$W = [(0,0,9,0,0,0),(0,3,8,0,0,0),(1,0,0,0,0,0)],$$
$$W_1 = [(0,0,0,1,0,1),(0,0,0,0,1,0),(0,0,0,0,0,1)].$$

Notemos que W_1 é T-invariante pois

$$\begin{aligned} T(0,0,0,1,0,1) &= (0,0,0,0,0,0) \in W_1, \\ T(0,0,0,0,1,0) &= (0,0,0,5,0,0) \\ &= 5(0,0,0,1,0,1) + 0(0,0,0,0,1,0) - 5(0,0,0,0,0,1) \in W_1, \\ T(0,0,0,0,0,1) &= (0,0,0,0,0,0) \in W_1. \end{aligned}$$

Agora, seja $T_2 : W_1 \to W_1$ dada por

$$T_2(0,0,0,a,b,a+c) = T(0,0,0,a,b,a+c) = (0,0,0,5b,0,0).$$

Então, $T_2^2(0,0,0,a,b,a+c) = T(0,0,0,5b,0,0) = (0,0,0,0,0,0)$ e portanto $T_2 \in L(W_1)$ é nilpotente, de índice de nilpotência 2.

Sejam $v_1 = (0,0,0,0,1,0) \in W_1$ e $T_2(v_1) = (0,0,0,5,0,0)$. Então $\{T_2(v_1), v_1\}$ é base de $W_2 = [(0,0,0,5,0,0),(0,0,0,0,1,0)]$.

Também, seja $T_3 : W_2 \to W_2$ dada por

$$T_3(0,0,0,5a,b,0) = T_2(0,0,0,5a,b,0) = T(0,0,0,5a,b,0) = (0,0,0,5b,0,0),$$

de modo que

$$T_3(0,0,0,5,0,0) = (0,0,0,0,0,0) = 0T_2v_1 + 0v_1,$$
$$T_3(0,0,0,0,1,0) = (0,0,0,5,0,0) = 1T_2v_1 + 0v_1.$$

Logo, $[T_3] = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$.

Temos que $W_1 = W_2 \oplus W_3$, onde $W_2 = [(0,0,0,5,0,0),(0,0,0,0,1,0)]$ é T_2-invariante e $W_3 = [(0,0,0,0,0,1)]$ é T_2-invariante, pois $T_2(0,0,0,0,0,1) = (0,0,0,0,0,0) \in W_3$.

Por fim, seja $T_4 : W_3 \to W_3$ dada por

$$T_4(0,0,0,0,0,a) = T_2(0,0,0,0,0,a) = (0,0,0,0,0,0).$$

Então, $T_4(0,0,0,0,0,1) = (0,0,0,0,0,0)$ e portanto, $[T_4] = [0]$.

Vejamos agora qual a matriz de T em relação à base

$$\{(0,0,9,0,0,0),(0,3,8,0,0,0),(1,0,0,0,0,0),\\ (0,0,0,5,0,0),(0,0,0,0,1,0),(0,0,0,0,0,1).\}$$

Temos:

$$\begin{aligned}
T(0,0,9,0,0,0) &= (0,0,0,0,0,0) = 0v_1 + 0v_2 + 0v_3 + 0v_4 + 0v_5 + 0v_6,\\
T(0,3,8,0,0,0) &= (0,0,9,0,0,0) = 1v_1 + 0v_2 + 0v_3 + 0v_4 + 0v_5 + 0v_6,\\
T(1,0,0,0,0,0) &= (0,3,8,0,0,0) = 0v_1 + 1v_2 + 0v_3 + 0v_4 + 0v_5 + 0v_6,\\
T(0,0,0,5,0,0) &= (0,0,0,0,0,0) = 0v_1 + 0v_2 + 0v_3 + 0v_4 + 0v_5 + 0v_6,\\
T(0,0,0,0,1,0) &= (0,0,0,5,0,0) = 0v_1 + 0v_2 + 0v_3 + 1v_4 + 0v_5 + 0v_6,\\
T(0,0,0,0,0,1) &= (0,0,0,0,0,0) = 0v_1 + 0v_2 + 0v_3 + 0v_4 + 0v_5 + 0v_6,
\end{aligned}$$

e portanto,

$$[T] = \begin{pmatrix} 0 & 1 & 0 & & & \\ 0 & 0 & 1 & & & \\ 0 & 0 & 0 & & & \\ & & & 0 & 1 & \\ & & & 0 & 0 & \\ & & & & & 0 \end{pmatrix} = \begin{pmatrix} [T_1] & 0 & 0 \\ 0 & [T_3] & 0 \\ 0 & 0 & [T_4] \end{pmatrix}.$$

Sugestão 3.30. Para cada um dos operadores:

1. $T(x,y,z) = (3x + y, 3y + z, 3z)$;

2. $T(x,y) = (2x + y, 2y)$;

3. $T(u,v,x,y,z) = (2u + x + y, 2v + 3y + z, 2x + 4z, 2y + 5z, 3z)$;

i) Calcule o polinômio característico;

ii) Calcule o polinômio minimal;

iii) Calcule a multiplicidade algébrica de cada autovalor λ_i;

iv) Calcule a multiplicidade geométrica de cada autovalor λ_i;

v) Para cada autovalor λ_i, encontre o subespaço $\operatorname{Ker}((A-\lambda_i I)^{k_i})$, onde k_i é o primeiro inteiro positivo tal que $\operatorname{Ker}((A-\lambda_i I)^{k_i}) = \operatorname{Ker}((A-\lambda_i I)^{k_i+1})$;

vi) Calcule $\dim \operatorname{Ker}((A-\lambda_i I)^{k_i})$;

vii) Se $H_i = \operatorname{Ker}((A-\lambda_i I)^{k_i})$, mostre que $(A-\lambda_i I)_{H_i}$ é nilpotente e calcule o seu índice de nilpotência;

viii) Determine a matriz de $(A-\lambda_i I)_{H_i}$ em relação à base

$$\left\{(A-\lambda_i I)^{k_i-1}(u), \ldots, (A-\lambda_i I)(u), u\right\},$$

onde u é escolhido tal que $(A-\lambda_i I)^{k_i}(u) = 0$ e $(A-\lambda_i I)^{k_i-1}(u) \neq 0$.

4

FORMA DE JORDAN

Já sabemos que nem toda matriz quadrada (todo operador linear) é diagonalizável. Neste caso, uma forma especial de deixar uma matriz quadrada (um operador linear) mais simples é obter a sua forma de Jordan, que será estudada neste capítulo.

Definição 4.1. A matriz quadrada

$$\begin{pmatrix} x & 1 & 0 & \dots & 0 & 0 \\ 0 & x & 1 & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 1 & 0 \\ 0 & \ddots & \ddots & \ddots & x & 1 \\ 0 & 0 & 0 & \cdots & 0 & x \end{pmatrix}$$

com x na diagonal principal, 1 na diagonal imediatamente acima desta e 0 no resto é chamada *bloco básico de Jordan associado a x*.

Observação 4.2. Um $(m \times m)$-bloco básico de Jordan associado a x é simplesmente $xI_m + M_m$, onde M_m é a matriz de um operador nilpotente em relação a uma base, dada por

$$M_m = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}_{m \times m}$$

Definição 4.3. Dizemos que uma matriz A de ordem n está sob a forma de Jordan *(forma canônica de Jordan)* se A for triangular superior da forma

$$A = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_r \end{pmatrix},$$

onde os blocos de Jordan J_i, $i = 1, \ldots, r$, são formados de blocos básicos de Jordan associado a x_i, isto é,

$$J_i = \begin{pmatrix} \underbrace{\begin{matrix} x_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & x_i & 1 & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & x_i & 1 \\ 0 & 0 & \cdots & \cdots & 0 & x_i \end{matrix}}_{n_1 \times n_1} & \mathbf{0} & & \mathbf{0} \\ \mathbf{0} & \underbrace{\begin{matrix} x_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & x_i & 1 & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & x_i & 1 \\ 0 & 0 & \cdots & \cdots & 0 & x_i \end{matrix}}_{n_2 \times n_2} & & \mathbf{0} \\ & & \ddots & \\ \mathbf{0} & \mathbf{0} & & \underbrace{\begin{matrix} x_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & x_i & 1 & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & x_i & 1 \\ 0 & 0 & \cdots & \cdots & 0 & x_i \end{matrix}}_{n_{r_i} \times n_{r_i}} \end{pmatrix}$$

$$= \begin{pmatrix} B_{i1} & & & \\ & B_{i2} & & \\ & & \ddots & \\ & & & B_{ir_i} \end{pmatrix},$$

com os blocos da diagonal de ordem decrescente, isto é, $n_1 \geq n_2 \geq \cdots \geq n_{r_i}$.

Exemplo 4.4. A matriz de Jordan 1×1 é da forma $[x]$.

Exemplo 4.5. As matrizes de Jordan 2×2 são:

$$\begin{pmatrix} x_1 & 0 \\ 0 & x_2 \end{pmatrix}, \qquad \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix}, \qquad \begin{pmatrix} x & 1 \\ 0 & x \end{pmatrix}.$$

Exemplo 4.6. As matrizes de Jordan 3×3 são:

$$\begin{pmatrix} x_1 & 0 & 0 \\ 0 & x_2 & 0 \\ 0 & 0 & x_3 \end{pmatrix}, \quad \begin{pmatrix} x_1 & 0 & 0 \\ 0 & x_1 & 0 \\ 0 & 0 & x_2 \end{pmatrix}, \quad \begin{pmatrix} x_1 & 1 & 0 \\ 0 & x_1 & 0 \\ 0 & 0 & x_2 \end{pmatrix}, \quad \begin{pmatrix} x_1 & 0 & 0 \\ 0 & x_2 & 0 \\ 0 & 0 & x_2 \end{pmatrix},$$

$$\begin{pmatrix} x_1 & 0 & 0 \\ 0 & x_2 & 1 \\ 0 & 0 & x_2 \end{pmatrix}, \quad \begin{pmatrix} x & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x \end{pmatrix}, \quad \begin{pmatrix} x & 1 & 0 \\ 0 & x & 0 \\ 0 & 0 & x \end{pmatrix}, \quad \begin{pmatrix} x & 1 & 0 \\ 0 & x & 1 \\ 0 & 0 & x \end{pmatrix}.$$

Teorema 4.7. Seja T um elemento de $L(V)$, com todos os seus autovalores distintos $x_1, \dots, x_r$ em K. Então V pode ser escrito como $V = V_1 \oplus \cdots \oplus V_r$, onde $V_i = \mathrm{Ker}((T - x_i I)^{n_i}) = \{v \in V \mid (T - x_i I)^{n_i}(v) = 0\}$. Além disso, o polinômio minimal de T_i, o operador linear induzido por T sobre V_i, é $m_i(x) = (x - x_i)^{n_i}$.

Demonstração. Se $T \in L(V)$ possui todos os seus autovalores em K, então o polinômio minimal de T tem a forma $m(x) = (x - x_1)^{n_1}(x - x_2)^{n_2} \cdots (x - x_r)^{n_r}$, onde $x_1, x_2, \dots, x_r$ são autovalores distintos de T.

Seja $m_i(x) = (x - x_i)^{n_i}$, $i = 1, 2, \dots, r$. Consideremos agora os r polinômios

$$\begin{aligned} h_1(x) &= m_2(x) m_3(x) \cdots m_{r-1}(x) m_r(x); \\ h_2(x) &= m_1(x) m_3(x) \cdots m_{r-1}(x) m_r(x); \\ &\vdots \\ h_{r-1}(x) &= m_1(x) m_2(x) \cdots m_{r-2}(x) m_r(x); \\ h_r(x) &= m_1(x) m_2(x) \cdots m_{r-2}(x) m_{r-1}(x). \end{aligned}$$

Seja $v \in V$. Daí se $w_i = h_i(T)(v)$, então $m_i(T)(w_i) = m_i(T) h_i(T)(v) = m(T)(v) = 0$. Portanto, $w_i = h_i(T)(v) \in V_i$, $i = 1, 2, \dots, r$. Assim, $h_i(T)(V) \subset V_i$, $i = 1, 2, \dots, r$. Como os polinômios $h_1(x), h_2(x), \dots, h_r(x)$ são primos entre si, podemos determinar polinômios $s_1(x), s_2(x), \dots, s_r(x)$ tais que

$$s_1(x) h_1(x) + s_2(x) h_2(x) + \cdots + s_r(x) h_r(x) = 1.$$

Então, $s_1(T) h_1(T) + s_2(T) h_2(T) + \cdots + s_r(T) h_r(T) = I$. Logo, se $v \in V$ então $v = s_1(T) h_1(T)(v) + s_2(T) h_2(T)(v) + \cdots + s_r(T) h_r(T)(v) = v_1 + v_2 + \cdots + v_r$, onde cada $v_i = s_i(T)h_i(T)(v) \in V_i$. Logo, $V = V_1 + \cdots + V_r$.

Precisamos agora verificar que esta soma é uma soma direta, isto é, se $u_1 + u_2 + \cdots + u_r = 0$, $u_i \in V_i$, então $u_i = 0$, $i = 1, 2, \ldots, r$. Aplicando $h_i(T)$ a $u_1 + u_2 + \cdots + u_r = 0$, obtemos

$$h_i(T)(u_1) + h_i(T)(u_2) + \cdots + h_i(T)(u_r) = h_i(T)(0) = 0$$

e portanto $h_i(T)(u_i) = 0$. Mas também, $m_i(T)(u_i) = 0$. Porém $h_i(T)$ e $m_i(T)$ são primos entre si, e daí, existem polinômios $s(x)$ e $r(x)$ tais que $s(x)h_i(x) + r(x)m_i(x) = 1$. Portanto, $s(T)h_i(T) + r(T)m_i(T) = I$ e aplicando essa expressão em u_i, obtemos $s(T)h_i(T)(u_i) + r(T)m_i(T)(u_i) = u_i$, donde segue que $u_i = 0$ e portanto a soma é direta.

Mostremos agora que o polinômio minimal de T_i, o operador linear induzido por T sobre V_i, é $m_i(x) = (x - x_i)^{n_i}$. Com efeito, pela definição de V_i, temos que $m_i(T)(V_i) = 0$ e daí $m_i(T_i) = 0$, donde o polinômio minimal de T_i é, necessariamente, um divisor de $m_i(x)$, sendo assim da forma $(x - x_i)^{s_i}$ com $s_i \leq n_i$. Mas o polinômio minimal de T sobre K é

$$\operatorname{mmc}\left\{(x - x_1)^{s_1}, (x - x_2)^{s_2}, \ldots, (x - x_r)^{s_r}\right\} = (x - x_1)^{s_1}(x - x_2)^{s_2} \cdots (x - x_r)^{s_r}.$$

Como o minimal é $(x - x_1)^{n_1}(x - x_2)^{n_2} \cdots (x - x_r)^{n_r}$ temos, necessariamente, $n_1 \leq s_1, n_2 \leq s_2, \ldots, n_r \leq s_r$. E daí $n_i = s_i$, para $i = 1, 2, \ldots, r$. □

Exemplo 4.8. Verificamos o Teorema 4.7 para o operador $T \in L(\mathbb{R}^3)$ dado por

$$T(x, y, z) = (x - 3y + 3z, 3x - 5y + 3z, 6x - 6y + 4z).$$

Temos que $p(\lambda) = (\lambda + 2)^2(\lambda - 4)$ e $m(\lambda) = (\lambda + 2)(\lambda - 4)$. Ainda $V_1 = \operatorname{Ker}(T + 2I)$ e $V_2 = \operatorname{Ker}(T - 4I)$. É claro que V_1 e V_2 são subespaços T-invariantes. Ainda mais, $V_1 = [(1,1,0), (-1,0,1)]$ e portanto $\dim V_1 = 2$, $V_2 = [(1,1,2)]$ e portanto $\dim V_2 = 1$ e $V_1 \cap V_2 = \{0\}$. Como

$$(x, y, z) = (y - x)(-1, 0, 1) + \left(\frac{-x + 3y - z}{2}\right)(1, 1, 0) + \left(\frac{x - y + z}{2}\right)(1, 1, 2)$$

segue que $\mathbb{R}^3 = V_1 \oplus V_2$.

Temos que $T_1 : V_1 \to V_1$ e $T_2 : V_2 \to V_2$ são dados por

$$T_1(a-b,a,b) = T(a-b,a,b) = (-2(a-b),-2a,-2b),$$
$$T_2(c,c,2c) = T(c,c,2c) = (4c,4c,8c).$$

Então $[T_1] = \left(\begin{smallmatrix} -2 & 0 \\ 0 & -2 \end{smallmatrix}\right)$ e $[T_2] = [4]$. Assim, $p_1(\lambda) = (\lambda+2)^2$, $m_1(\lambda) = (\lambda+2)$ e $p_2(\lambda) = (\lambda-4)$, $m_2(\lambda) = \lambda - 4$.

Teorema 4.9. Seja $T \in L(V)$ com todos os seus autovalores distintos $x_1, \ldots, x_r$ em K. Então, pode-se determinar uma base de V em relação à qual a matriz de T é da forma

$$\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_r \end{pmatrix}, \text{ onde } J_i = \begin{pmatrix} B_{i_1} & & & \\ & B_{i_2} & & \\ & & \ddots & \\ & & & B_{i_{k_i}} \end{pmatrix}$$

e $B_{i_1}, \ldots, B_{i_{k_i}}$ são blocos básicos de Jordan associados a x_i.

Demonstração. Pelo Teorema 4.7, $V = V_1 \oplus \cdots \oplus V_r$, onde $V_i = \mathrm{Ker}((T - x_i I)^{n_i})$, $i = 1, 2, \ldots r$. Assim, podemos determinar uma base de V tal que, em relação a esta base, a matriz de T é da forma

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_r \end{pmatrix},$$

onde cada A_i é uma matriz $d_i \times d_i$ que é, de fato, a matriz de T_i, induzida de T sobre V_i. Observemos que cada operador T_i possui apenas um autovalor x_i e que $T_i - x_i I$ é nilpotente pois $(T_i - x_i I)^{n_i} = 0$. Podemos escrever $T_i = x_i I + (T_i - x_i I)$. Como $T_i - x_i I$ é nilpotente de índice de nilpotência n_i, existe uma base em relação à qual sua matriz é da forma

$$\begin{pmatrix} M_{n_{i_1}} & & & \\ & M_{n_{i_2}} & & \\ & & \ddots & \\ & & & M_{n_{i_{k_i}}} \end{pmatrix},$$

com $n_{i_1} + n_{i_2} + \cdots + n_{i_{k_i}} = d_i$. Portanto a matriz de T_i é da forma

$$\begin{pmatrix} x_i I_{n_{i_1}} & & & \\ & x_i I_{n_{i_2}} & & \\ & & \ddots & \\ & & & x_i I_{n_{i_{k_i}}} \end{pmatrix} + \begin{pmatrix} M_{n_{i_1}} & & & \\ & M_{n_{i_2}} & & \\ & & \ddots & \\ & & & M_{n_{i_{k_i}}} \end{pmatrix} = \begin{pmatrix} B_{i_1} & & & \\ & B_{i_2} & & \\ & & \ddots & \\ & & & B_{i_{k_i}} \end{pmatrix}.$$

Observemos que, pelo Teorema 3.28, poderíamos tomar a base de V de modo que ordem $B_{i_1} \geq$ ordem $B_{i_2} \geq \cdots \geq$ ordem $B_{i_{k_i}}$. □

Teorema 4.10. Sejam T em $L(V)$, $p(x) = (x - c_1)^{d_1} \cdots (x - c_r)^{d_r}$, $c_i \in K$, polinômio característico e $m(x) = (x - c_1)^{n_1} \cdots (x - c_r)^{n_r}$ polinômio minimal de T, com $n_i \leq d_i$, $i = 1, 2, \ldots, r$ e $d_1 + \cdots + d_r = \dim V$. Então:

i) $V = V_1 \oplus \cdots \oplus V_r$ onde $V_i = \operatorname{Ker}((T - c_i I)^{n_i})$;

ii) Os subespaços V_i são T-invariantes;

iii) O polinômio minimal de T_i, restrição de T a V_i, é $m_i(x) = (x - c_i)^{n_i}$;

iv) $\dim V_i = d_i$.

Demonstração. Os itens (i) e (iii) seguem do Teorema 4.7. Para a demonstração do item (ii), coloquemos $m_i(x) = (x - c_i)^{n_i}$ e assim $m_i(T) = (T - c_i I)^{n_i}$. Logo $V_i = \operatorname{Ker}(m_i(T))$. Seja $v \in \operatorname{Ker}(m_i(T))$. Então $m_i(T)(v) = 0$. Queremos mostrar que $T(v) \in \operatorname{Ker}(m_i(T))$, ou seja, que $m_i(T)(T(v)) = 0$. Como $f(t)t = tf(t)$, então $m_i(T) \circ T = T \circ m_i(T)$. Assim $m_i(T)(T(v)) = T(m_i(T)(v)) = T(0) = 0$ e portanto V_i é T-invariante. Para a demonstração do item iv), consideremos $\phi_i = T_i - c_i I$. Então ϕ_i é nilpotente pois ${\phi_i}^{n_i}(v_i) = (T_i - c_i I)^{n_i}(v_i) = 0$.

Portanto, o polinômio característico de ϕ_i é x^{e_i}, onde $e_i = \dim V_i$. Logo, o polinômio característico de T_i é $(x - c_i)^{e_i}$ pois

$$\begin{aligned} p_{T_i}(x) &= \det(T_i - xI) = \det(\phi_i + c_i I - xI) = \det(\phi_i - (x - c_i)I) \\ &= p_{\phi_i}(x - c_i) = (x - c_i)^{e_i}. \end{aligned}$$

Mas, por hipótese, o polinômio característico de T_i é $(x - c_i)^{d_i}$ e portanto $e_i = d_i$, ou equivalentemente, $\dim V_i = d_i$. □

Observação 4.11. Desejamos agora fazer mais algumas observações sobre o operador T e a matriz de Jordan J que representa T em relação a uma certa base.

Sejam $p(x) = (x - x_1)^{d_1} \cdots (x - x_r)^{d_r}$ o polinômio característico e $m(x) = (x - x_1)^{n_1} \cdots (x - x_r)^{n_r}$ o polinômio minimal, com $d_1 + \cdots + d_r = n = \dim V$ e $n_i \leq d_i$, $i = 1, \ldots, r$. Assim, J e J_i são como no enunciado do Teorema 4.9 e valem:

i) Todo elemento de J que não esteja na diagonal principal ou imediatamente acima dela é nulo.

ii) Na diagonal principal de J aparecem os r autovalores $x_1, \ldots, x_r$ de T. Além disso, x_i se repete d_i vezes (multiplicidade algébrica) e assim J_i é um bloco $d_i \times d_i$.

iii) Para cada i, o primeiro bloco básico de Jordan B_{i1} no bloco de Jordan J_i é uma matriz $n_i \times n_i$, e todos os outros blocos básicos de Jordan B_{ij} são de ordem menor ou igual a n_i.

iv) O número de blocos básicos de Jordan B_{ij} é igual à multiplicidade geométrica dos x_i.

Exemplo 4.12. Suponha que os polinômios característico e minimal de T são, respectivamente, $p(x) = (x - 2)^4 (x - 3)^3$ e $m(x) = (x - 2)^2 (x - 3)^2$. Como é a forma canônica de Jordan de T?

A forma de Jordan de T é uma das seguintes matrizes, onde a primeira (resp. segunda) matriz ocorre se a multiplicidade geométrica do autovalor 2 for 2 (resp. 3).

$$\begin{pmatrix} 2 & 1 & & & & & \\ 0 & 2 & & & & & \\ & & 2 & 1 & & & \\ & & 0 & 2 & & & \\ & & & & 3 & 1 & \\ & & & & 0 & 3 & \\ & & & & & & 3 \end{pmatrix} \quad \text{ou} \quad \begin{pmatrix} 2 & 1 & & & & & \\ 0 & 2 & & & & & \\ & & 2 & & & & \\ & & & 2 & & & \\ & & & & 3 & 1 & \\ & & & & 0 & 3 & \\ & & & & & & 3 \end{pmatrix}.$$

Exemplo 4.13. Se A é uma matriz com polinômio característico $p(x) = (x-2)^3(x+7)^2$ e minimal $m(x) = (x-2)^2(x+7)$, então a forma de Jordan de A é

$$J_A = \begin{pmatrix} 2 & 1 & & & \\ 0 & 2 & & & \\ & & 2 & & \\ & & & -7 & \\ & & & & -7 \end{pmatrix}.$$

Exemplo 4.14. Quantas formas de Jordan são possíveis para uma matriz quadrada, cujo polinômio característico é $p(x) = (x+2)^4(x-1)^2$?

As possibilidades para o polinômio minimal são:

$$\begin{aligned} m_1(x) &= (x+2)^4(x-1)^2; & m_2(x) &= (x+2)^4(x-1); \\ m_3(x) &= (x+2)^3(x-1)^2; & m_4(x) &= (x+2)^3(x-1); \\ m_5(x) &= (x+2)^2(x-1)^2; & m_6(x) &= (x+2)^2(x-1); \\ m_7(x) &= (x+2)(x-1)^2; & m_8(x) &= (x+2)(x-1). \end{aligned}$$

Logo, são possíveis 10 formas de Jordan para esta matriz, pois $m_5(x)$ e $m_6(x)$ contribuem com duas, cada um. São elas:

$$J_1 = \begin{pmatrix} -2 & 1 & 0 & 0 & & \\ 0 & -2 & 1 & 0 & & \\ 0 & 0 & -2 & 1 & & \\ 0 & 0 & 0 & -2 & & \\ & & & & 1 & 1 \\ & & & & 0 & 1 \end{pmatrix} \qquad J_2 = \begin{pmatrix} -2 & 1 & 0 & 0 & & \\ 0 & -2 & 1 & 0 & & \\ 0 & 0 & -2 & 1 & & \\ 0 & 0 & 0 & -2 & & \\ & & & & 1 & \\ & & & & & 1 \end{pmatrix}$$

$$J_3 = \begin{pmatrix} -2 & 1 & 0 & & & \\ 0 & -2 & 1 & & & \\ 0 & 0 & -2 & & & \\ & & & -2 & & \\ & & & & 1 & 1 \\ & & & & 0 & 1 \end{pmatrix} \qquad J_4 = \begin{pmatrix} -2 & 1 & 0 & & & \\ 0 & -2 & 1 & & & \\ 0 & 0 & -2 & & & \\ & & & -2 & & \\ & & & & 1 & \\ & & & & & 1 \end{pmatrix}$$

$$J_5 = \begin{pmatrix} -2 & 1 & & & & \\ 0 & -2 & & & & \\ & & -2 & 1 & & \\ & & 0 & -2 & & \\ & & & & 1 & 1 \\ & & & & 0 & 1 \end{pmatrix} \qquad J_5' = \begin{pmatrix} -2 & 1 & & & & \\ 0 & -2 & & & & \\ & & -2 & & & \\ & & & -2 & & \\ & & & & 1 & 1 \\ & & & & 0 & 1 \end{pmatrix}$$

$$J_6 = \begin{pmatrix} -2 & 1 & & & & \\ 0 & -2 & & & & \\ & & -2 & 1 & & \\ & & 0 & -2 & & \\ & & & & 1 & \\ & & & & & 1 \end{pmatrix} \qquad J_6' = \begin{pmatrix} -2 & 1 & & & & \\ 0 & -2 & & & & \\ & & -2 & & & \\ & & & -2 & & \\ & & & & 1 & \\ & & & & & 1 \end{pmatrix}$$

$$J_7 = \begin{pmatrix} -2 & & & & & \\ & -2 & & & & \\ & & -2 & & & \\ & & & -2 & & \\ & & & & 1 & 1 \\ & & & & 0 & 1 \end{pmatrix} \qquad J_8 = \begin{pmatrix} -2 & & & & & \\ & -2 & & & & \\ & & -2 & & & \\ & & & -2 & & \\ & & & & 1 & \\ & & & & & 1 \end{pmatrix}$$

Acontecerá J_5 (resp. J_5') se a multiplicidade geométrica do autovalor -2 for 2 (resp. 3). Da mesma forma para J_6 e J_6'.

Exemplo 4.15. Seja a matriz

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

Temos que $p_A(x) = (x-2)^5(x+1)$ e $m_A(x) = (x-2)^4(x+1)$. Logo, a forma de Jordan de A é

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

Continuaremos buscando a forma de Jordan de um operador linear T.

Exemplo 4.16. Suponhamos que $\lambda = 3$ seja autovalor de um operador linear T e que satisfaça às seguintes condições:

i) multiplicidade algébrica de $\lambda = 3$ em $p(\lambda)$ igual a 10;

ii) multiplicidade algébrica de $\lambda = 3$ em $m(\lambda)$ igual a 3;

iii) multiplicidade geométrica de $\lambda = 3$ em $p(\lambda)$ igual a 6.

Das hipóteses acima, podemos tirar as seguintes conclusões:

1. Na diagonal principal do bloco de Jordan aparece o número 3 (3 é o autovalor do operador linear T).

2. A ordem do bloco de Jordan é 10 (10 é a multiplicidade algébrica de $\lambda = 3$ em $p(\lambda)$).

3. O primeiro bloco básico de Jordan tem ordem 3 (3 é a multiplicidade algébrica de $\lambda = 3$ em $m(\lambda)$).

4. Os demais blocos básicos de Jordan têm ordem menor ou igual a 3.

5. O bloco de Jordan possui 6 blocos básicos de Jordan (6 é a multiplicidade geométrica de $\lambda = 3$ em $p(\lambda)$).

Assim, as possíveis formas de Jordan são:

$$\begin{pmatrix} 3 & 1 & 0 & & & & & & & \\ 0 & 3 & 1 & & & & & & & \\ 0 & 0 & 3 & & & & & & & \\ & & & 3 & 1 & 0 & & & & \\ & & & 0 & 3 & 1 & & & & \\ & & & 0 & 0 & 3 & & & & \\ & & & & & & 3 & & & \\ & & & & & & & 3 & & \\ & & & & & & & & 3 & \\ & & & & & & & & & 3 \end{pmatrix}, \begin{pmatrix} 3 & 1 & 0 & & & & & & & \\ 0 & 3 & 1 & & & & & & & \\ 0 & 0 & 3 & & & & & & & \\ & & & 3 & 1 & & & & & \\ & & & 0 & 3 & & & & & \\ & & & & & 3 & 1 & & & \\ & & & & & 0 & 3 & & & \\ & & & & & & & 3 & & \\ & & & & & & & & 3 & \\ & & & & & & & & & 3 \end{pmatrix}.$$

Temos portanto informações sobre a ordem do maior bloco de Jordan e sobre o número de blocos existentes para cada λ_i. Falta somente informações sobre a ordem desses blocos. Para isto, sejam $N = \dim V$, $d_j = \dim \mathrm{Ker}((T - \lambda_i I)^j)$ e n_j o número de blocos básicos de Jordan de ordem j.

Observe que devemos calcular as dimensões d_j até obtermos o primeiro inteiro k tal que $d_k = d_{k+1}$, que é o índice de nilpotência do operador $T - \lambda_i I$. A partir desse índice temos $d_j = d_k$, $j \geq k$.

Observe ainda que

$$d_0 = \dim \mathrm{Ker}(I) = 0.$$

Sabemos que o número de blocos básicos de Jordan é igual a multiplicidade geométrica, logo

$$d_1 = n_1 + n_2 + \cdots + n_N,$$

que são todas as ordens possíveis.

Agora, quando calculamos $(T - \lambda_i I)^2$, a "diagonal" de 1's escorrega para a "diagonal" imediatamente acima e isto significa que, nos blocos básicos de Jordan de ordens maiores ou iguais a 2, aumenta uma coluna de zeros em cada um. Logo

$$d_2 = n_1 + 2n_2 + \cdots + 2n_N.$$

Com o mesmo raciocínio concluímos para os subsequentes, isto é,

$$\begin{aligned}
d_3 &= n_1 + 2n_2 + 3n_3 + \cdots + 3n_N = n_1 + 2n_2 + 3(n_3 + \cdots + n_N),\\
&\vdots\\
d_{N-1} &= n_1 + 2n_2 + \cdots + (N-2)n_{N-2} + (N-1)(n_{N-1} + n_N),\\
d_N &= n_1 + 2n_2 + \cdots + Nn_N,\\
d_{N+1} &= d_N.
\end{aligned}$$

Os d_i's são conhecidos (já foram calculados). Vamos resolver para os n_i's. Subtraindo cada equação da anterior obtemos:

$$\begin{aligned}
d_1 - d_0 &= n_1 + \cdots + n_N,\\
d_2 - d_1 &= n_2 + \cdots + n_N,\\
&\vdots\\
d_N - d_{N-1} &= n_N,\\
d_{N+1} - d_N &= 0.
\end{aligned}$$

Subtraindo cada equação da subsequente, vem

$$-d_0 + 2d_1 - d_2 = n_1, \ \ldots, \ -d_{N-1} + 2d_N - d_{N+1} = n_N.$$

Obtemos portanto a relação

$$n_j = -d_{j-1} + 2d_j - d_{j+1}, \quad 1 \le j \le N \tag{4.1}$$

que fornece o número de blocos básicos de Jordan de ordem j correspondentes ao autovalor λ_i.

Voltemos ao exemplo anterior, supondo $d_2 = 9$. Como a multiplicidade geométrica de $\lambda = 3$ em $p(\lambda)$ é igual a 6, temos que $d_1 = \dim \operatorname{Ker}(T - 3I) = 6$ e como a multiplicidade algébrica de $\lambda = 3$ em $m(\lambda)$ é igual a 3, temos que $d_j = d_3 = 10$, para $j \ge 3$.

Assim, faltam somente as ordens dos blocos básicos de Jordan e a quantidade de cada um deles. Para isto, usamos a fórmula (4.1) e obtemos

$$n_1 = -d_0 + 2d_1 - d_2 = 0 + 12 - 9 = 3,$$
$$n_2 = -d_1 + 2d_2 - d_3 = -6 + 18 - 10 = 2,$$
$$n_3 = -d_2 + 2d_3 - d_4 = -9 + 20 - 10 = 1.$$

Logo, o bloco de Jordan correspondente ao autovalor 3 é

$$\begin{pmatrix}
3 & 1 & 0 & & & & & & & \\
0 & 3 & 1 & & & & & & & \\
0 & 0 & 3 & & & & & & & \\
 & & & 3 & 1 & & & & & \\
 & & & 0 & 3 & & & & & \\
 & & & & & 3 & 1 & & & \\
 & & & & & 0 & 3 & & & \\
 & & & & & & & 3 & & \\
 & & & & & & & & 3 & \\
 & & & & & & & & & 3
\end{pmatrix}$$

Teorema 4.17. Duas transformações lineares em $L(V)$ que têm todos os seus autovalores em K são semelhantes se, e somente se, elas podem ser reduzidas à mesma forma de Jordan (a menos da ordem dos seus autovalores).

Demonstração. Sejam A e B duas transformações lineares em $L(V)$ que têm todos os seus autovalores em K e suponhamos inicialmente que elas possam ser reduzidas à mesma forma de Jordan, isto é, existam P e Q inversíveis tais que $P^{-1}AP = J_A = J_B = Q^{-1}BQ$.

Logo, $QP^{-1}APQ^{-1} = B$ e colocando $R = PQ^{-1}$ temos que R é inversível e $R^{-1} = QP^{-1}$, donde segue que $R^{-1}AR = B$ e portanto A é semelhante a B.

Reciprocamente, suponhamos que A é semelhante a B e mostremos então que elas podem ser reduzidas a mesma forma de Jordan. De fato:

1) A e B possuem o mesmo polinômio característico pois são semelhantes (vide Proposição 1.16).

2) Seja $f(t) = a_0 + a_1t + \cdots + a_nt^n$ um polinômio de grau n. Mostremos que$f(B) = 0$ se, e somente se, $f(A) = 0$.

Com efeito, como A e B são semelhantes existe uma matriz inversível P tal que $P^{-1}AP = B$. Assim, temos que $f(B) = 0$ se, e somente se, $a_0I + a_1B + \cdots + a_nB^n = 0$, o que é equivalente a

$$a_0P^{-1}IP + a_1P^{-1}AP + \cdots + a_n(P^{-1}AP)^n = 0,$$

ou ainda, $P^{-1}[a_0I + a_1A + \cdots + a_nA^n]P = 0$. Portanto $f(B) = 0$ se, e somente se, $P^{-1}f(A)P = 0$, ou equivalentemente, $f(A) = 0$.

Se denotarmos por m_A e m_B os polinômios minimais de A e B respectivamente, temos que $m_B(B) = 0$ se, e somente se, $m_B(A) = 0$ e $m_A(A) = 0$ se, e somente se, $m_A(B) = 0$.

Logo, $m_A \mid m_B$ e $m_B \mid m_A$, donde segue que os polinômios minimais m_A e m_B são iguais.

3) Para o mesmo autovalor λ de A e B temos que $\mathrm{Ker}(A - \lambda I)$ é isomorfo à $\mathrm{Ker}(B - \lambda I)$. De fato, como A e B são semelhantes existe uma matriz inversível P tal que $P^{-1}AP = B$. Defina a transformação linear

$$S : \mathrm{Ker}(A - \lambda I) \to \mathrm{Ker}(B - \lambda I)$$

por $S(v) = P^{-1}(v)$. Observemos inicialmente que se $v \in \mathrm{Ker}(A - \lambda I)$ então $P^{-1}(v) \in \mathrm{Ker}(B - \lambda I)$ pois

$$\begin{aligned}(B - \lambda I)(P^{-1}(v)) &= (P^{-1}AP - \lambda P^{-1}IP)(P^{-1}(v)) = (P^{-1}(A - \lambda I)P)(P^{-1}(v)) \\ &= (P^{-1}(A - \lambda I))(v) = P^{-1}((A - \lambda I)(v)) = P^{-1}(0) = 0\end{aligned}$$

e portanto S é bem definida.

Note que a transformação linear inversa de S,

$$S^{-1} : \mathrm{Ker}(B - \lambda I) \to \mathrm{Ker}(A - \lambda I),$$

é dada por $S^{-1}(w) = P(w)$. Basta observarmos que se $w \in \mathrm{Ker}(B - \lambda I)$ então $(A - \lambda I)(P(w)) = (PBP^{-1} - \lambda PIP^{-1})(P(w)) = (P(B - \lambda I)P^{-1})(P(w)) = (P(B - \lambda I))(w) = P(0) = 0$ e portanto S^{-1} é bem definida. Logo, $\mathrm{Ker}(A - \lambda I)$

é isomorfo a $\text{Ker}(B - \lambda I)$ e portanto a multiplicidade geométrica de λ, visto como autovalor de A, é a mesma que a multiplicidade geométrica de λ, visto como autovalor de B.

4) Sabemos que $n_j = -d_{j+1} + 2d_j - d_{j-1}$, para $j = 1, \ldots, N = \dim V$, onde n_j denota o número de blocos básicos de Jordan de ordem j e $d_j = \dim \text{Ker}((A - \lambda_i I)^j)$, com λ_i autovalor de A (e portanto de B). Se mostrarmos que $\text{Ker}((A - \lambda_i I)^j)$ é isomorfo à $\text{Ker}((B-\lambda_i I)^j)$ concluímos que o número de blocos básicos de Jordan de ordem j em J_A e o número de blocos básicos de Jordan de ordem j em J_B são os mesmos.

Mas $(B-\lambda_i I)^j = (P^{-1}AP-\lambda_i P^{-1}IP)^j = (P^{-1}(A-\lambda_i I)P)^j = P^{-1}(A-\lambda_i I)^j P$. Assim, a transformação linear

$$S : \text{Ker}((A - \lambda_i I)^j) \to \text{Ker}((B - \lambda_i I)^j)$$

dada por $S(v) = P^{-1}(v)$ é bem definida e dá um isomorfismo entre $\text{Ker}((A - \lambda_i I)^j)$ e $\text{Ker}((B-\lambda_i I)^j)$, com inversa $S^{-1} : \text{Ker}((B-\lambda_i I)^j) \to \text{Ker}((A-\lambda_i I)^j)$, dada por $S^{-1}(w) = P(w)$.

Pelos itens de 1) a 4) concluímos que $J_A = J_B$ (a menos da ordem dos seus autovalores). $\square$

Exemplo 4.18. Toda matriz complexa é semelhante a sua transposta. De fato, temos:

1) A e A^t possuem o mesmo polinômio característico, pois $\det(A - \lambda I) = \det(A - \lambda I)^t = \det(A^t - (\lambda I)^t) = \det(A^t - \lambda I^t) = \det(A^t - \lambda I)$. Logo, possuem os mesmos autovalores.

2) Se $f(t) = a_0 + a_1 t + \cdots + a_n t^n$ e $f(A) = 0$, então

$$\begin{aligned} f(A^t) &= a_0 I + a_1 A^t + \cdots + a_n (A^t)^n = a_0 I^t + a_1 A^t + \cdots + a_n (A^n)^t = \\ &= (a_0 I + a_1 A + \cdots + a_n A^n)^t = f(A)^t = 0^t = 0. \end{aligned}$$

Logo as matrizes A e A^t possuem o mesmo polinômio minimal.

3) Apesar de terem os mesmos autovalores, podem ter autovetores diferentes, como sugere o exemplo a seguir: Seja $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. Então $\lambda = 1$ é o único autovalor e $V_A(1) = [(1,0)]$. $A^t = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$, $\lambda = 1$ é ainda o único autovalor e $V_{A^t}(1) = [(0,1)]$.

4) As multiplicidades geométricas são iguais, pois

$$\dim \operatorname{Ker}(A - \lambda I) = \dim \operatorname{Ker}((A - \lambda I)^t) = \dim \operatorname{Ker}(A^t - \lambda I).$$

5) A e A^t possuem o mesmo número de blocos básicos de Jordan de ordem j pois $\operatorname{Ker}((A - \lambda I)^j)$ é isomorfo a $\operatorname{Ker}(((A - \lambda I)^j)^t) = \operatorname{Ker}(((A - \lambda I)^t)^j) = \operatorname{Ker}((A^t - \lambda I)^j)$. Portanto A e A^t possuem a mesma forma canônica de Jordan e daí são semelhantes.

Sugestão 4.19. As matrizes A e B abaixo são semelhantes.

$$A = \begin{pmatrix} 0 & 1 & & & \\ & 0 & \ddots & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & & & & \\ 1 & 0 & & & \\ & \ddots & \ddots & & \\ & & \ddots & 0 & \\ & & & 1 & 0 \end{pmatrix}$$

Exemplo 4.20. Duas matrizes nilpotentes de ordem três são semelhantes se, e somente se, elas possuem o mesmo índice de nilpotência. A afirmação não é verdadeira para matrizes nilpotentes de ordem 4.

De fato, sejam A e B matrizes de ordem três nilpotentes. Suponhamos que A é semelhante a B. Então A e B possuem a mesma forma de Jordan que é a matriz do Teorema 3.28. Temos três casos a considerar, a saber:

i) Se o índice de nilpotência de A for 3, então $J_A = J_B = \left(\begin{smallmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{smallmatrix}\right)$. Logo, B tem índice de nilpotência 3.

ii) Se o índice de nilpotência de A for 2, então $J_A = J_B = \left(\begin{smallmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right)$. Logo, B tem índice de nilpotência 2.

iii) Se o índice de nilpotência de A for 1, então $J_A = J_B = \left(\begin{smallmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right)$. Logo, B tem índice de nilpotência 1.

Por outro lado, se A e B são matrizes de ordem três nilpotentes, com o mesmo índice de nilpotência, então elas possuem a mesma forma de Jordan (matriz do Teorema 3.28). Logo elas são semelhantes.

Para vermos que a afirmação não é verdadeira para matrizes nilpotentes de ordem 4, consideremos as matrizes

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

Então A e B são nilpotentes de índice 2 mas não são semelhantes.

Exemplo 4.21. Sejam A e B matrizes complexas de ordem n, para as quais $A^n = I$ e $B^n = I$. Serão A e B semelhantes?

Não! Consideremos $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ e $B = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Então temos que $A^2 = I$, $p_A(x) = (1-x)^2$, $m_A(x) = (1-x)$, $B^2 = I$, $p_B(x) = (1-x)(-1-x)$ e $m_B(x) = (1-x)(-1-x)$. Logo, $J_A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ e $J_B = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Portanto, $J_A \neq J_B$ e daí A e B não são semelhantes.

Exemplo 4.22. Sejam N_1 e N_2 matrizes de ordem três nilpotentes sobre K. Então N_1 e N_2 são semelhantes se, e somente se, possuem o mesmo polinômio minimal.

De fato, se N_1 é semelhante a N_2, então $J_{N_1} = J_{N_2}$. Agora, matrizes semelhantes possuem o mesmo polinômio característico. Por serem 3×3 e $J_{N_1} = J_{N_2}$ então possuem o mesmo minimal (ver Exemplo 4.20). Se são matrizes de ordem três nilpotentes então o polinômio característico é x^3. E como o minimal é o mesmo, então possuem a mesma forma de Jordan. Logo são semelhantes.

Exemplo 4.23. Se A é uma matriz $n \times n$ sobre K com polinômio característico $p(x) = (x - c_1)^{d_1} \cdots (x - c_r)^{d_r}$, $c_i \in K$, qual o traço de A?

Sabemos que A é semelhante a uma matriz de Jordan J, onde na diagonal aparecem os c_i. Assim,

$$\operatorname{tr}(A) = \operatorname{tr}(P^{-1}JP) = \operatorname{tr}(JP^{-1}P) = \operatorname{tr}(J) = d_1c_1 + \cdots + d_rc_r.$$

Exemplo 4.24. Se A é uma matriz $n \times n$ nilpotente elementar, isto é, $A^n = 0$ e $A^{n-1} \neq 0$, então A e A^t são semelhantes.

De fato, temos

$$(A^t)^n = A^t A^t \cdots A^t = (AA\cdots A)^t = (A^n)^t = 0^t = 0.$$

Assim também $(A^t)^{n-1} \neq 0$. Logo, possuem a mesma forma de Jordan (ver Teorema 3.19). Então são semelhantes.

Agora, dado um operador linear $T \in L(V)$ veremos, através de exemplos, como encontrar uma matriz inversível P tal que $P^{-1}[T]P$ é uma matriz de Jordan, ou equivalentemente, como encontrar uma base de V em relação a qual a matriz de T é uma matriz de Jordan.

Exemplo 4.25. Suponhamos que a matriz do operador linear $T \in L(\mathbb{R}^5)$ seja

$$A = \begin{pmatrix} 2 & 0 & 1 & 1 & 0 \\ & 2 & 2 & 0 & 1 \\ & & 2 & 2 & 0 \\ & & & 2 & 0 \\ & & & & 3 \end{pmatrix}.$$

Então temos $p_A(\lambda) = (3 - \lambda)(\lambda - 2)^4$. Logo, os autovalores de A são 3 e 2.

Temos $(A - 3I)(v) = 0$ se, e somente se,

$$\begin{pmatrix} -1 & 0 & 1 & 1 & 0 \\ & -1 & 2 & 0 & 1 \\ & & -1 & 2 & 0 \\ & & & -1 & 0 \\ & & & & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

equivalente a

$$-x_1 + x_3 + x_4 = 0, \qquad -x_2 + 2x_3 + x_5 = 0, \qquad -x_3 + 2x_4 = 0,$$
$$-1x_4 = 0, \qquad 0x_5 = 0,$$

ou ainda, $x_1 = x_3 = x_4 = 0$ e $x_2 = x_5$.

Logo $\operatorname{Ker}(A-3I) = \{(0,x,0,0,x) \mid x \in \mathbb{R}\} = [(0,1,0,0,1)]$.

Também, $(A-3I)^2(v) = 0$ se, e somente se,

$$\begin{pmatrix} 1 & 0 & -2 & 0 & 0 \\ & 1 & -4 & 4 & -1 \\ & & 1 & -4 & 0 \\ & & & 1 & 0 \\ & & & & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

que é equivalente a

$$x_1 - 2x_3 = 0, \qquad x_2 - 4x_3 + 4x_4 - x_5 = 0, \qquad x_3 - 4x_4 = 0,$$
$$1x_4 = 0, \qquad 0x_5 = 0,$$

ou ainda, $x_1 = x_3 = x_4 = 0$ e $x_2 = x_5$. Logo

$$\operatorname{Ker}((A-3I)^2) = \{(0,x,0,0,x) \mid x \in \mathbb{R}\} = [(0,1,0,0,1)]$$

e portanto $\operatorname{Ker}(A-3I) = \operatorname{Ker}((A-3I)^2)$. Ainda,

$$A - 2I = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ & 0 & 2 & 0 & 1 \\ & & 0 & 2 & 0 \\ & & & 0 & 0 \\ & & & & 1 \end{pmatrix}, \quad (A-2I)^2 = \begin{pmatrix} 0 & 0 & 0 & 2 & 0 \\ & 0 & 0 & 4 & 1 \\ & & 0 & 0 & 0 \\ & & & 0 & 0 \\ & & & & 1 \end{pmatrix},$$

$$(A-2I)^3 = (A-2I)^4 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 & 1 \\ & & 0 & 0 & 0 \\ & & & 0 & 0 \\ & & & & 1 \end{pmatrix}.$$

Portanto

$$\operatorname{Ker}((A-2I)^3) = \operatorname{Ker}((A-2I)^4) = \{(x,y,z,w,0) \mid x,y,z,w \in \mathbb{R}\}$$

e segue que $\mathbb{R}^5 = \operatorname{Ker}(A-3I) \oplus \operatorname{Ker}((A-2I)^3)$.

Vamos agora encontrar uma base especial para $\mathbb{R}^5$. Seja $u = (0,1,0,0,1)$. Então $\{u\}$ é base de $\operatorname{Ker}(A-3I)$. Procuremos agora uma base para o núcleo $\operatorname{Ker}((A-2I)^3)$. Para isto, tomemos v_0 tal que $(A-2I)^3(v_0) = 0$ mas $(A-2I)^2(v_0) \neq 0$. Por exemplo, tomemos $v_0 = (0,0,0,1,0)$. Então temos $v_1 = (A-2I)(v_0) = (1,0,2,0,0)$ e $v_2 = (A-2I)^2(v_0) = (2,4,0,0,0)$. Como $\dim\operatorname{Ker}((A-2I)^3) = 4$, precisamos de mais um vetor. Escolhemos esse quarto vetor de tal forma que $v_3 \in \operatorname{Ker}(A-2I)$ e $\{v_2, v_1, v_0, v_3\}$ seja linearmente independente. Por exemplo, tomemos $v_3 = (1,1,0,0,0)$. Assim, esse conjunto é base de $\operatorname{Ker}((A-2I)^3)$. Então, $\{v_2, v_1, v_0, v_3, u\}$ é a base especial de $\mathbb{R}^5$ procurada.

Com relação a essa base temos:

$$(A-2I)(v_2) = (A-2I)((A-2I)^2(v_0)) = 0 \Rightarrow A(v_2) = 2v_2,$$

$$(A-2I)(v_1) = (A-2I)((A-2I)(v_0)) = v_2 \Rightarrow A(v_1) = v_2 + 2v_1,$$

$$(A-2I)(v_0) = v_1 \Rightarrow A(v_0) = v_1 + 2v_0,$$

$$A(v_3) = 2v_3, \quad A(u) = 3u.$$

Logo, a matriz de T em relação a essa base especial é

$$[T] = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ & 2 & 1 & 0 & 0 \\ & & 2 & 0 & 0 \\ & & & 2 & 0 \\ & & & & 3 \end{pmatrix}$$

e assim

$$P = \begin{pmatrix} 2 & 1 & 0 & 1 & 0 \\ 4 & 0 & 0 & 1 & 1 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \qquad P^{-1} = \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} & \frac{1}{4} & 0 & -\frac{1}{2} \\ 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 2 & -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$P^{-1}AP = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ & 2 & 1 & 0 & 0 \\ & & 2 & 0 & 0 \\ & & & 2 & 0 \\ & & & & 3 \end{pmatrix}.$$

Exemplo 4.26. Suponhamos que a matriz do operador linear $T \in L(\mathbb{R}^5)$ seja

$$A = \begin{pmatrix} 2 & 1 & 1 & 0 & 0 \\ & 2 & 0 & 0 & 0 \\ & & 2 & 0 & 0 \\ & & & 2 & 1 \\ & & & & 3 \end{pmatrix}.$$

Então, temos $p_A(\lambda) = (\lambda - 3)(\lambda - 2)^4$ e assim os autovalores de A são 3 e 2.

Temos $(A - 3I)(v) = 0$ se, e somente se,

$$\begin{pmatrix} -1 & 1 & 1 & 0 & 0 \\ & -1 & 0 & 0 & 0 \\ & & -1 & 0 & 0 \\ & & & -1 & 1 \\ & & & & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

que é equivalente a

$$-x_1 + x_2 + x_3 = 0, \qquad -x_2 = 0, \qquad -x_3 = 0,$$
$$-x_4 + x_5 = 0, \qquad 0x_5 = 0,$$

ou ainda, $x_1 = x_2 = x_3 = 0$ e $x_4 = x_5$.

Logo, temos $\mathrm{Ker}(A - 3I) = \{(0,0,0,x,x) \mid x \in \mathbb{R}\} = [(0,0,0,1,1)] = \mathrm{Ker}((A-3I)^2)$. Ainda,

$$(A - 2I) = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 0 \\ & & & 0 & 1 \\ & & & & 1 \end{pmatrix}, \quad (A-2I)^2 = (A-2I)^3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 0 \\ & & & 0 & 1 \\ & & & & 1 \end{pmatrix}.$$

Logo, $\mathrm{Ker}((A-2I)^2) = \mathrm{Ker}((A-2I)^3) = \{(x,y,z,w,0) \mid x,y,z,w \in \mathbb{R}\}$. Portanto, temos que $\mathbb{R}^5 = \mathrm{Ker}(A-3I) \oplus \mathrm{Ker}((A-2I)^2)$.

Vamos agora encontrar uma base especial para $\mathbb{R}^5$. Se $u = (0,0,0,1,1)$, então $\{u\}$ é base de $\mathrm{Ker}(A-3I)$. Para uma base de $\mathrm{Ker}((A-2I)^2)$, tomemos v_0 tal que $(A-2I)^2(v_0) = 0$ mas $(A-2I)(v_0) \neq 0$. Por exemplo, tomemos

$v_0 = (0,1,0,0,0)$. Então temos $v_1 = (A-2I)(v_0) = (1,0,0,0,0)$. Como $\dim \operatorname{Ker}((A-2I)^2) = 4$, precisamos de mais dois vetores. Escolhemos esses dois vetores da seguinte forma: $v_2, v_3 \in \operatorname{Ker}(A-2I)$ tais que $\{v_1, v_0, v_2, v_3\}$ seja linearmente independente. Por exemplo, tomemos $v_2 = (0,1,-1,0,0)$ e $v_3 = (0,0,0,1,0)$.

Assim, o conjunto $\{v_1, v_0, v_2, v_3, u\}$ é a base especial de $\mathbb{R}^5$ procurada e então temos:

$$(A-2I)(v_1) = (A-2I)((A-2I)(v_0)) = 0 \Rightarrow A(v_1) = 2v_1,$$

$$(A-2I)(v_0) = v_1 \Rightarrow A(v_0) = v_1 + 2v_0,$$

$$A(v_2) = 2v_2, \quad A(v_3) = 2v_3, \quad A(u) = 3u.$$

Logo, a matriz de T em relação a essa base especial é

$$[T] = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ & 2 & 0 & 0 & 0 \\ & & 2 & 0 & 0 \\ & & & 2 & 0 \\ & & & & 3 \end{pmatrix}.$$

Como

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ & 1 & 1 & 0 & 0 \\ & & -1 & 0 & 0 \\ & & & 1 & 1 \\ & & & & 1 \end{pmatrix} \quad \text{e} \quad P^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ & 1 & 1 & 0 & 0 \\ & & -1 & 0 & 0 \\ & & & 1 & -1 \\ & & & & 1 \end{pmatrix}$$

então

$$P^{-1}AP = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ & 2 & 0 & 0 & 0 \\ & & 2 & 0 & 0 \\ & & & 2 & 0 \\ & & & & 3 \end{pmatrix}.$$

Exemplo 4.27. Suponhamos que a matriz do operador linear $T \in L(\mathbb{R}^5)$ seja

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 & 1 \\ & 2 & 0 & 0 & 2 \\ & & 2 & 0 & 1 \\ & & & 2 & 2 \\ & & & & 3 \end{pmatrix}.$$

Então temos $p_A(\lambda) = (\lambda - 3)(\lambda - 2)^4$ e os autovalores de A são 3 e 2.

Temos $(A - 3I)(v) = 0$ se, e somente se,

$$\begin{pmatrix} -1 & 0 & 0 & 0 & 1 \\ & -1 & 0 & 0 & 2 \\ & & -1 & 0 & 1 \\ & & & -1 & 2 \\ & & & & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

que é equivalente a

$$-x_1 + x_5 = 0, \qquad -x_2 + 2x_5 = 0, \qquad -x_3 + x_5 = 0,$$
$$-x_4 + 2x_5 = 0, \qquad 0x_5 = 0,$$

ou ainda, $x_1 = x_3 = x_5$ e $x_2 = x_4 = 2x_5$. Logo,

$$\text{Ker}(A - 3I) = \{(x, 2x, x, 2x, x) \mid x \in \mathbb{R}\} = [(1, 2, 1, 2, 1)] = \text{Ker}((A - 3I)^2).$$

Ainda,

$$A - 2I = (A - 2I)^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ & 0 & 0 & 0 & 2 \\ & & 0 & 0 & 1 \\ & & & 0 & 2 \\ & & & & 1 \end{pmatrix}.$$

Mas $\text{Ker}(A-2I) = \{(x, y, z, w, 0) \mid x, y, z, w \in \mathbb{R}\}$. Logo, $\dim \text{Ker}(A-2I) = 4$ e assim temos que $\mathbb{R}^5 = \text{Ker}(A-3I) \oplus \text{Ker}(A-2I)$ e portanto possuímos uma base de autovetores donde segue que A é diagonalizável. Sendo

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ & 1 & 0 & 0 & 2 \\ & & 1 & 0 & 1 \\ & & & 1 & 2 \\ & & & & 1 \end{pmatrix} \quad \text{e} \quad P^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ & 1 & 0 & 0 & -2 \\ & & 1 & 0 & -1 \\ & & & 1 & -2 \\ & & & & 1 \end{pmatrix}$$

segue que

$$P^{-1}AP = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ & 2 & 0 & 0 & 0 \\ & & 2 & 0 & 0 \\ & & & 2 & 0 \\ & & & & 3 \end{pmatrix}.$$

Quando o operador T possui autovalores complexos $\lambda = a + bi$, a forma de Jordan obtida pelo processo acima é complexa e os autovalores são complexos conjugados. Nesse caso o operador T deve ser considerado sobre $\mathbb{C}^N$ e a forma de Jordan pode ser utilizada normalmente com os mesmos objetivos. Desejamos obter também nesse caso uma matriz real.

Por exemplo, se $[T] = A = \left(\begin{smallmatrix} -1 & -5 \\ 2 & 5 \end{smallmatrix}\right)$, então $p_A(\lambda) = \lambda^2 - 4\lambda + 5$ e os autovalores de A são $\lambda = 2 + i$ e $\bar{\lambda} = 2 - i$. Um autovetor associado a λ é $v = (-3 + i, 2)$ e a $\bar{\lambda}$ é $\bar{v} = (-3 - i, 2)$. Logo T é diagonalizável e a matriz de T com relação a base $\{v, \bar{v}\}$ de $\mathbb{C}^2$ é

$$\begin{pmatrix} 2 + i & 0 \\ 0 & 2 - i \end{pmatrix}.$$

Sejam $v_1 = (v + \bar{v})/2 = (-3, 2)$ e $v_2 = (\bar{v} - v)/2i = (-1, 0)$. Então $\{v_1, v_2\}$ é uma base de $\mathbb{R}^2$ e como

$$A(v_1) = (-7, 4) = 2(-3, 2) + 1(-1, 0),$$
$$A(v_2) = (1, -2) = (-1)(-3, 2) + 2(-1, 0),$$

a matriz de T em relação a essa base é

$$\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} \operatorname{Re} \lambda & -\operatorname{Im} \lambda \\ \operatorname{Im} \lambda & \operatorname{Re} \lambda \end{pmatrix}.$$

Essa matriz encontrada é chamada *forma canônica real* associada a T. Ainda mais, existe uma matriz D tal que

$$D^{-1}\begin{pmatrix} 2+i & 0 \\ 0 & 2-i \end{pmatrix} D = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}.$$

Basta tomar $D = \begin{pmatrix} 1/2 & i/2 \\ 1/2 & -i/2 \end{pmatrix}$.

Assim, de uma matriz real A de ordem 2 com autovalores $\lambda = a + bi$, onde $b > 0$ e $\bar{\lambda} = a - bi$, existe uma base tal que a matriz de T em relação a essa base é

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

chamada *forma canônica real* associada a A.

De fato, suponha que v é autovetor associado a λ e $\bar{v}$ é autovetor associado a $\bar{\lambda}$. Defina $v_1 = (v + \bar{v})/2$ e $v_2 = (\bar{v} - v)/2i$. Então v_1 e v_2 são reais. Como $b > 0$, temos que $\lambda \neq \bar{\lambda}$ e daí $\{v, \bar{v}\}$ é linearmente independente e então $\beta = \{v_1, v_2\}$ também é; logo são bases.

Observe que $\bar{v} = v_1 + iv_2$. Então $A(v_1 + iv_2) = A(\bar{v}) = \bar{\lambda}\bar{v} = (a - bi)(v_1 + iv_2)$ e portanto $A(v_1) + iA(v_2) = (av_1 + bv_2) + i(-bv_1 + av_2)$. Logo $A(v_1) = av_1 + bv_2$ e $A(v_2) = -bv_1 + av_2$ e daí

$$[T] = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Em resumo, quando o operador T possui autovalores complexos

$$\lambda = a + bi \ (b > 0) \quad \text{e} \quad \bar{\lambda} = a - bi,$$

utilizando o processo acima, podemos obter suas formas canônicas complexa e real, como nos tipos (I)–(III) da Tabela 4.1.

Já vimos uma justificativa para (I). Agora vamos ver uma justificativa para (III). O caso (II) deixamos como exercício para o leitor.

	Forma Canônica Complexa	Forma Canônica Real
I.	$\begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix}$	$D = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$
II.	$\begin{pmatrix} \lambda & 0 & \cdots & 0 & & & & \\ 0 & \lambda & \ddots & \vdots & & & & \\ \vdots & \ddots & \ddots & 0 & & & & \\ 0 & \cdots & 0 & \lambda & & & & \\ & & & & \bar{\lambda} & 0 & \cdots & 0 \\ & & & & 0 & \bar{\lambda} & \ddots & \vdots \\ & & & & \vdots & \ddots & \ddots & 0 \\ & & & & 0 & \cdots & 0 & \bar{\lambda} \end{pmatrix}$	$\begin{pmatrix} D & & & \\ & D & & \\ & & \ddots & \\ & & & D \end{pmatrix}$
III.	$\begin{pmatrix} \lambda & 1 & \cdots & 0 & & & & \\ 0 & \lambda & \ddots & \vdots & & & & \\ \vdots & \ddots & \ddots & 1 & & & & \\ 0 & \cdots & 0 & \lambda & & & & \\ & & & & \bar{\lambda} & 1 & \cdots & 0 \\ & & & & 0 & \bar{\lambda} & \ddots & \vdots \\ & & & & \vdots & \ddots & \ddots & 1 \\ & & & & 0 & \cdots & 0 & \bar{\lambda} \end{pmatrix}$	$\begin{pmatrix} D & I & & \\ & D & \ddots & \\ & & \ddots & I \\ & & & D \end{pmatrix}$

Tabela 4.1: Formas canônicas complexa e real

Justificativa de (III). Vamos denotar por λ o autovalor com parte imaginária positiva e por $\bar{\lambda}$, seu conjugado.

Seja $\{v_1, \ldots, v_k\}$, onde $v_j = x_j + iy_j$, uma base do autoespaço generalizado correspondente a λ, isto é, uma base para $\mathrm{Ker}((T - \lambda I)^k)$, onde v_k é tal que $(T - \lambda I)^k(v_k) = 0$ e $(T - \lambda I)^{k-1}(v_k) \neq 0$, e

$$\begin{aligned} v_{k-1} &= (T - \lambda I)(v_k), \\ v_{k-2} &= (T - \lambda I)^2(v_k), \\ &\vdots \\ v_2 &= (T - \lambda I)^{k-2}(v_k), \\ v_1 &= (T - \lambda I)^{k-1}(v_k). \end{aligned}$$

Observe que $v_{k-1} = T(v_k) - \lambda v_k$ e portanto $T(v_k) = \lambda v_k + v_{k-1}$. Também, $v_{k-2} = (T-\lambda I)^2(v_k) = (T-\lambda I)((T-\lambda I)(v_k)) = (T-\lambda I)(v_{k-1}) = T(v_{k-1}) - \lambda v_{k-1}$ de modo que $T(v_{k-1}) = \lambda v_{k-1} + v_{k-2}$, e assim por diante, ou seja, $T(v_j) = \lambda v_j + v_{j-1}$ para $1 < j \leq k$.

Também, $\{\bar{v_1}, \ldots, \bar{v_k}\}$ (onde $\bar{\ }$ denota o conjugado) é a base correspondente a $\bar{\lambda}$; logo esses $2k$ vetores são linearmente independentes sobre o corpo $\mathbb{C}$ dos números complexos, e portanto $\{y_1, x_1, \ldots, y_k, x_k\}$ são $2k$ vetores linearmente independentes sobre o corpo $\mathbb{R}$. Isto segue das seguintes igualdades:

$$a_1x_1 + b_1y_1 + \cdots + a_kx_k + b_ky_k =$$

$$\frac{a_1}{2}(v_1 + \bar{v_1}) + \frac{b_1}{2i}(v_1 - \bar{v_1}) + \cdots + \frac{a_k}{2}(v_k + \bar{v_k}) + \frac{b_k}{2i}(v_k - \bar{v_k}) =$$

$$\left(\frac{a_1}{2} + \frac{b_1}{2i}\right)v_1 + \cdots + \left(\frac{a_k}{2} + \frac{b_k}{2i}\right)v_k + \left(\frac{a_1}{2} - \frac{b_1}{2i}\right)\bar{v_1} + \cdots + \left(\frac{a_k}{2} - \frac{b_k}{2i}\right)\bar{v_k}.$$

Agora, como $\{v_1, \ldots, v_k\}$ é base do autoespaço generalizado na qual T está na forma de Jordan, temos que, para $1 < j \leq k$,

$$\begin{aligned} T(v_j) &= T(x_j) + iT(y_j) = \lambda v_j + v_{j-1} \\ &= (ax_j - by_j + x_{j-1}) + i(bx_j + ay_j + y_{j-1}) \end{aligned}$$

e portanto

$$T(y_j) = ay_j + bx_j + 1y_{j-1} + 0x_{j-1} \quad \text{e} \quad T(x_j) = -by_j + ax_j + 0y_{j-1} + 1x_{j-1}.$$

Logo, podemos concluir que a matriz de T na forma de Jordan real (na base formada por blocos de vetores do tipo $\{y_1, x_1, \ldots, y_k, x_k\}$, nessa ordem) é uma matriz (verifique) formada por blocos em diagonal da forma

$$\begin{pmatrix} D & I & & \\ & D & \ddots & \\ & & \ddots & I \\ & & & D \end{pmatrix}, \text{ onde } D = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \text{ e } I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Aplicação. Consideremos $T \in L(\mathbb{C}^4)$ tal que $p_T(x) = (x-\lambda)^2(x-\bar{\lambda})^2$. Há duas possibilidades para o polinômio minimal de T:

i) $m_T(x) = p_T(x)$, ii) $m_T(x) = (x-\lambda)(x-\bar{\lambda})$.

Se ocorre (i), $\mathbb{C}^4 = W_1 \oplus W_2$, onde $W_1 = \operatorname{Ker}((T-\lambda I)^2)$ e $W_2 = \operatorname{Ker}((T-\bar{\lambda} I)^2)$, e $(T-\lambda I)_{W_1}$ e $(T-\bar{\lambda}I)_{W_2}$ são nilpotentes de índice 2. Então

$$W_1 = \left[\{(T-\lambda I)(v_2), v_2\}\right] \quad \text{e} \quad W_2 = \left[\{(T-\bar{\lambda} I)(\bar{v_2}), \bar{v_2}\}\right]$$

e relativamente à base $\beta = \left\{(T-\lambda I)(v_2), v_2, (T-\bar{\lambda}I)(\bar{v_2}), \bar{v_2}\right\}$ a forma de Jordan complexa de T é

$$[T]_\beta = \begin{pmatrix} \lambda & 1 & & \\ 0 & \lambda & & \\ & & \bar{\lambda} & 1 \\ & & 0 & \bar{\lambda} \end{pmatrix}.$$

Reordenando β como $\theta = \left\{(T-\lambda I)(v_2), (T-\bar{\lambda}I)(\bar{v_2}), v_2, \bar{v_2}\right\}$ obtemos

$$[T]_\theta = \begin{pmatrix} \lambda & 0 & 1 & 0 \\ 0 & \bar{\lambda} & 0 & 1 \\ & & \lambda & 0 \\ & & 0 & \bar{\lambda} \end{pmatrix}.$$

Olhando para as matrizes de T_{K_1} e T_{K_2}, onde

$$K_1 = \left[\left\{(T-\lambda I)(v_2), (T-\bar{\lambda}I)(\bar{v_2})\right\}\right] \quad \text{e} \quad K_2 = \left[\left\{v_2, \bar{v_2}\right\}\right]$$

temos

$$[T_{K_1}] = \begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix} \quad \text{e} \quad [T_{K_2}] = \begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix}$$

e a forma canônica real para ambas é $A = \left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)$. Portanto a forma canônica real para $[T]_\theta$ é

$$A' = \begin{pmatrix} a & -b & 1 & 0 \\ b & a & 0 & 1 \\ & & a & -b \\ & & b & a \end{pmatrix}.$$

Se ocorre (ii), $m_T(x) = (x-\lambda)(x-\bar{\lambda})$ e portanto o operador T é diagonalizável, ou seja, exite uma base de autovetores $\gamma = \{v_1, v_2, v_3, v_4\}$ tal que

$$[T]_\gamma = \begin{pmatrix} \lambda & 0 & & \\ 0 & \bar{\lambda} & & \\ & & \lambda & 0 \\ & & 0 & \bar{\lambda} \end{pmatrix}.$$

Como, para ambos os (2×2)-blocos $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{smallmatrix}\right)$ da matriz acima, temos que a forma canônica real associada é

$$B = \begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

segue que a forma canônica real associada a $[T]_\gamma$ é

$$B' = \begin{pmatrix} a & -b & & \\ b & a & & \\ & & a & -b \\ & & b & a \end{pmatrix}.$$

Sugestão 4.28. Seja $T \in L(\mathbb{C}^5)$ tal que o polinômio minimal de T é

$$m_T(x) = (x-i)^2(x+i)(x-1).$$

Encontre as formas canônicas complexa e real associadas a T.

REFERÊNCIAS BIBLIOGRÁFICAS

[1] HERSTEIN, I.N. *Tópicos de Álgebra*. São Paulo, Editora da Universidade de São Paulo e Polígono, 1970.

[2] HOFFMAN, K. e KUNZE, R. *Álgebra Linear*. Livros Técnicos e Científicos, Rio de Janeiro, 1979.

[3] NEVES, A.F. *Forma de Jordan e Equações Diferenciais*. Notas de aula.

[4] BROWN, W.C. *A Second Course in Linear Agebra*. Wiley–Interscience, 1987.

ÍNDICE REMISSIVO

S

T

As formas elementares são parte integrante de um curso de Álgebra Linear para licenciandos, bacharelandos e pós-graduandos em Matemática. Trata-se de um tema extremamente importante não apenas na Matemática como também em aplicações na Física e Engenharia.

O objetivo central deste livro é abordar de forma clara e objetiva a forma diagonal, a forma triangular e a forma de Jordan, tópicos estes em via de regra explorados de forma bem superficial nos livros didáticos de Álgebra Linear em nível de graduação.

Este texto é fruto de nossa experiência como professores e/ou ex-professores do Instituto de Geociências e Ciências Exatas, câmpus de Rio Claro, Unesp – Universidade Estadual Paulista, onde ministramos a disciplina Álgebra Linear para o cursos de Graduação em Matemática e de Pós-Graduação em Matemática.

Anízio Perissinotto Junior possui doutorado em Ciências da Computação e Matemática Computacional pela Universidade de São Paulo e pós-doutorado pela School of Mathematics – Center for Dynamical Systems and Non-linear Studies – Georgia Institute of Technology. Atualmente é professor aposentado, como professor assistente da PP do QDUNESP, efetivo, com função de professor adjunto, lotado no Departamento de Matemática do Instituto de Geociências e Ciências Exatas da Universidade Estadual Paulista. Tem experiência na área de Equações Diferenciais e Análise.

João Peres Vieira possui doutorado e pós-doutorado em Matemática pela Universidade de São Paulo. Atualmente é professor adjunto do Instituto de Geociências e Ciências Exatas, câmpus de Rio Claro, Unesp – Universidade Estadual Paulista, onde atua desde 1986. É revisor do Zentralblatt Math e líder do grupo de pesquisa: Topologia Algébrica, Diferencial e Geométrica. Pesquisa na área de Topologia Algébrica, atuando principalmente nos seguintes temas: pontos fixos e coincidência de aplicações fibradas.

Carina Alves é doutora em Matemática pela Universidade Estadual de Campinas e possui pós-doutorado pela Telecom Paristech – Paris, França. Atualmente é professora assistente doutora do Instituto de Geociências e Ciências Exatas, câmpus de Rio Claro, Unesp – Universidade Estadual Paulista, onde atua desde 2009. Tem experiência na área de Álgebra, atuando principalmente nos seguintes temas: teoria de códigos, reticulados e teoria algébrica dos números.

ISBN 978-85-7983-524-7

9 788579 835247