Parte 2

Autovalores e autovetores

Introduziremos os espaços vetoriais sobre $\mathbb{C}$ e generalizaremos os resultados obtidos para espaços vetoriais reais finitamente gerados. Vamos trabalhar, daqui por diante, com espaços vetorias reais ou complexos finitamente gerados. Trabalharemos com K-espaços vetoriais, onde $\mathsf{K} = \mathbb{R}$ ou $\mathsf{K} = \mathbb{C}$.

Apresentaremos os conceitos de subespaço invariante por meio de um operador K-linear e daremos ênfase a um tipo especial de subespaço invariante, a saber, os subespaços característicos, que são subespaços gerados por autovetores associados a um autovalor de um operador K-linear.

Ensinaremos como determinar, caso existam, os autovalores e os autovetores de um operador K-linear em um espaço vetorial de dimensão finita. Apresentaremos os conceitos de polinômio característico e polinômio mínimo de um operador K-linear, multiplicidade algébrica e multiplicidade geométrica de um autovalor. Introduziremos o conceito de operadores diagonalizáveis e daremos condições necessárias e suficientes para um operador K-linear em um espaço vetorial de dimensão $n \geq 1$ ser diagonalizável, em termos do polinômio característico e dos subespaços característicos, mas também do polinômio mínimo.

Espaços vetoriais sobre $\mathbb{C}$

Na primeira disciplina de Álgebra Linear vocês estudaram espaços vetoriais sobre $\mathbb{R}$ finitamente gerados. Vamos generalizar os conceitos lá aprendidos para espaços vetoriais sobre $\mathbb{C}$.

Sabemos que $\mathbb{C} = \{a + bi \; ; \; a, b \in \mathbb{R} \text{ e } i^2 = -1\}$ está munido com as operações de adição e multiplicação, respectivamente:

$(a + bi) + (c + di) = (a + c) + (b + d)i$ e

$(a + bi) \cdot (c + di) = (ac - bd) + (ad + bc)i.$

Com essas operações $\mathbb{C}$ é um corpo.

Definição 1 (Espaço vetorial sobre $\mathbb{C}$)

Um $\mathbb{C}$-espaço vetorial é um conjunto não vazio V, munido com as operações de adição e multiplicação por escalar:

$$\begin{array}{rccc} + : & V \times V & \longrightarrow & V \\ & (u, v) & \longmapsto & u + v \end{array} \quad \text{e} \quad \begin{array}{rccc} \cdot : & \mathbb{C} \times V & \longrightarrow & V \\ & (a, v) & \longmapsto & a \cdot v \end{array}$$

tendo as seguintes propriedades, para quaisquer $u, v, w \in V$ e $a, b \in \mathbb{C}$:

(1) Comutativa: $u + v = v + u$.

(2) Associativa: $(u + v) + w = u + (v + w)$.

(3) Existência de elemento neutro: Existe 0_V, tal que $v + 0_V = v$, para todo $v \in V$.

(4) Existência de simétrico: Para cada $v \in V$, existe $u \in V$ tal que $u + v = 0_V$.

(5) $1 \cdot v = v$.

(6) Associativa: $(a \cdot b) \cdot v = a \cdot (b \cdot v)$.

(7) Distributiva: $(a + b) \cdot v = a \cdot v + b \cdot v$.

(8) Distributiva: $a \cdot (u + v) = a \cdot u + a \cdot v$.

Exemplo 1

$\mathbb{C}$ é um $\mathbb{C}$-espaço vetorial com as operações de adição e multiplicação usuais de números complexos.

Verifique!

Exemplo 2

Seja $n \geq 1$ um número natural. Definimos

$$\mathbb{C}^n = \{(x_1, \ldots, x_n) \; ; \; x_j \in \mathbb{C}, \text{ para todo } j = 1, \ldots, n\}.$$

$\mathbb{C}^n$ é um $\mathbb{C}$-espaço vetorial com as seguintes operações de adição e multiplicação por escalar definidas a seguir, chamadas de *operações usuais*:

$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$, onde $x_j, y_j \in \mathbb{C}$, para $j = 1, \ldots, n$ e

$a \cdot (x_1, \ldots, x_n) = (a \cdot x_1, \ldots, a \cdot x_n)$, onde $a \in \mathbb{C}$ e $x_j \in \mathbb{C}$, para $j = 1, \ldots, n$.

Observamos que $(0, 0, \ldots, 0)$ é o elemento neutro de $\mathbb{C}^n$ e o simétrico de $v = (x_1, \ldots, x_n) \in \mathbb{C}^n$ é $u = (-x_1, \ldots, -x_n)$.

Verifique as outras propriedades. Todas elas são consequência das propriedades das operações do corpo dos números complexos.

Exemplo 3

Sejam $m \geq 1, n \geq 1$ números naturais. O conjunto $M_{m\times n}(\mathbb{C})$ é um $\mathbb{C}$-espaço vetorial com as operações usuais de adição de matrizes e multiplicação de um escalar por uma matriz. A saber, sejam $A = (a_{ij}), B = (b_{ij}) \in M_{m\times n}(\mathbb{C})$ e $a \in \mathbb{C}$. Definimos:

$$A + B = (a_{ij} + b_{ij}) \text{ e } a \cdot B = (a \cdot b_{ij}),$$

para quaisquer $i = 1, \ldots, m$ e $j = 1, \ldots, n$.

Observamos que $0 = (a_{ij})$, onde $a_{ij} = 0$, para quaisquer $i = 1, \ldots, m$ e $j = 1, \ldots, n$, é o elemento neutro e o simétrico de $A = (a_{ij})$ é $B = (-a_{ij})$.

Verifique as outras propriedades.

Definição 2 (Subespaço vetorial)

Um *subespaço vetorial* de um $\mathbb{C}$-espaço vetorial V é um subconjunto não vazio W de V, que com as operações de adição e multiplicação por escalares de V é um $\mathbb{C}$-espaço vetorial.

Equivalentemente, um subconjunto W de um $\mathbb{C}$-espaço vetorial V é um subespaço de V se, e somente se, tem as seguintes propriedades:

(a) $0_V \in W$;

(b) se $u, w \in W$, então $u + w \in W$;

(c) se $a \in \mathbb{C}$ e $w \in W$, então $a \cdot w \in W$.

Exemplo 4

Seja V um $\mathbb{C}$-espaço vetorial. Então, $\{0_V\}$ e V são subespaços de V.

Exemplo 5

Consideremos $W = \{(x, y) \in \mathbb{C}^2 \ ; \ x + 2y = 0\}$. W é um subespaço vetorial de $\mathbb{C}^2$. De fato,

(a) $0 + 2 \cdot 0 = 0 \Longrightarrow (0, 0) \in \mathbb{C}^2$;

(b) se $u = (x, y), v = (x', y') \in W$, então $x + 2y = 0$, $x' + 2y' = 0$, $u + v = (x + x', y + y')$ e $(x + x') + 2(y + y') \stackrel{(1)}{=} (x + 2y) + (x' + 2y') \stackrel{(2)}{=} 0 + 2 \cdot 0 = 0$, logo $u + v \in W$;

Em (1) usamos a associatividade e a comutatividade da adição em $\mathbb{C}$ e em (2), que $u, v \in W$.

(c) se $w = (x, y) \in W$ e $a \in C$, então $x + 2y = 0$, $a \cdot w = (a \cdot x, a \cdot y)$ e $a \cdot x + 2(a \cdot y) \overset{(3)}{=} a \cdot (x + 2y) \overset{(4)}{=} a \cdot 0$, logo $a \cdot w \in W$.

Em (3) usamos a comutatividade da multiplicação e a distributividade em $\mathbb{C}$ e em (4), que $w \in W$.

As noções de dependência e independência linear em um $\mathbb{C}$-espaço vetorial são análogas às de um $\mathbb{R}$-espaço vetorial.

Definição 3 (Dependência e independência linear)
Seja V um $\mathbb{C}$-espaço vetorial. Os vetores $v_1, \ldots, v_n \in V$ são *linearmente independentes* se, e somente se,

se $a_1, \ldots, a_n \in \mathbb{C}$ e $a_1v_1 + \cdots + a_nv_n = 0$, então $a_1 = \cdots = a_n = 0$.

Caso contrário, $v_1, \ldots, v_n$ são ditos *linearmente dependentes*, ou seja,

existem $a_1, \ldots, a_n \in \mathbb{C}$, nem todos nulos, tais que $a_1v_1 + \cdots + a_nv_n = 0$.

Exemplo 6
Os vetores $v_1 = (1, 0), v_2 = (i, 0)$ são linearmente dependentes no $\mathbb{C}$-espaço vetorial $\mathbb{C}^2$, pois $v_2 = iv_1$ é equivalente a $iv_1 - v_2 = (0, 0)$, que é uma combinação linear nula de v_1 e v_2, com escalares em $\mathbb{C}$ nem todos nulos.

Entretanto, com $a, b \in \mathbb{R}$, temos:

$(0, 0) = av_1 + bv_2 = a(1, 0) + b(i, 0) = (a + bi, 0)$ se, e somente se, $a + bi = 0$ se, e somente se, $a = b = 0$.

Portanto, v_1 e v_2 são linearmente independentes no $\mathbb{R}$-espaço vetorial $\mathbb{C}^2$.

Exemplo 7
Os vetores $u_1 = (1, 0), u_2 = (0, 1)$ são linearmente independentes no $\mathbb{C}$-espaço vetorial $\mathbb{C}^2$, pois se $a_1, a_2 \in \mathbb{C}$ e $(0, 0) = a_1(1, 0) + a_2(0, 1) = (a_1, a_2)$, então $a_1 = a_2 = 0$.

Definição 4 (Conjunto gerador)
Seja V um $\mathbb{C}$-espaço vetorial. Dizemos que V é um $\mathbb{C}$-*espaço vetorial finitamente gerado* se, e somente se, existe $\{v_1, \ldots, v_n\} \subset V$, tal que para todo $v \in V$, existem $a_1, \ldots, a_n \in \mathbb{C}$ tais que

$$v = a_1v_1 + \cdots + a_nv_n.$$

Nesse caso, dizemos que $\{v_1, \ldots, v_n\}$ é um *conjunto gerador* de V, V é o $\mathbb{C}$-espaço vetorial gerado por $\{v_1, \ldots, v_n\}$ e escrevemos $V = [v_1, \ldots, v_n]$.

Definição 5 (Base e dimensão)
Seja $V \neq \{0\}$ um $\mathbb{C}$-espaço vetorial finitamente gerado. Dizemos que o subconjunto $\beta = \{v_1, \ldots, v_n\}$ de V é uma *base* de V se, e somente se, β gera V e é linearmente independente sobre $\mathbb{C}$. Nesse caso, dizemos que a *dimensão* de V é n e escrevemos $\dim_{\mathbb{C}} V = n$.

Observação:

(1) A definição de dimensão está bem posta, pois todas as bases têm o mesmo número de elementos. O número de elementos de uma base é o número máximo de vetores linearmente independentes e o número mínimo de geradores.

(2) Quando $V = \{0_V\}$ definimos $\dim_{\mathbb{C}} V = 0$.

Exemplo 8

Temos que $\dim_{\mathbb{C}} \mathbb{C}^n = n$. De fato, se $v \in \mathbb{C}^n$, então existem $x_1, \dots, x_n \in \mathbb{C}$ tais que

$$\begin{aligned} v = (x_1, \dots, x_n) &= (x_1, 0, \dots, 0) + (0, x_2, \dots, 0) + \dots + (0, 0, \dots, x_n) \\ &= x_1(1, 0, \dots, 0) + x_2(0, 1, 0, \dots, 0) + \dots + x_n(0, 0, \dots, 1), \end{aligned}$$

logo $\beta = \{e_1 = (1, 0, \dots, 0), e_2 = (0, 1, \dots, 0), \dots, e_n = (0, 0, \dots, 1)\}$ gera $\mathbb{C}^n$.

Como $(0, 0, \dots, 0) = x_1e_1 + x_2e_2 + \dots + x_ne_n = (x_1, x_2, \dots, x_n)$ se, e somente se, $x_1 = x_2 = \dots = x_n = 0$, então β é $\mathbb{C}$-linearmente independente.

Logo, β é uma base de $\mathbb{C}^n$ como $\mathbb{C}$-espaço vetorial.

Exemplo 9

Temos que $\dim_{\mathbb{C}} M_{2\times 2}(\mathbb{C}) = 4$.

De fato, tomando $E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, temos que se $x, y, z, w \in \mathbb{C}$, então

$$\begin{aligned} \begin{pmatrix} x & y \\ z & w \end{pmatrix} &= \begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & y \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ z & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & w \end{pmatrix} \\ &= x\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + y\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + z\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + w\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ &= xE_{11} + yE_{12} + zE_{21} + wE_{22}, \end{aligned}$$

logo $\{E_{11}, E_{12}, E_{21}, E_{22}\}$ gera $M_{2\times 2}(\mathbb{C})$.

Além disso,

$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = xE_{11} + yE_{12} + zE_{21} + wE_{22} = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$ se, e somente se, $x = y = z = w = 0$, mostrando que $\{E_{11}, E_{12}, E_{21}, E_{22}\}$ é $\mathbb{C}$-linearmente independente.

Definição 6 (Transformação $\mathbb{C}$-linear)

Sejam V e W $\mathbb{C}$-espaços vetoriais. Uma função $T : V \longrightarrow W$ é chamada uma *transformação $\mathbb{C}$-linear* se, e somente se,

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

(ii) $T(av) = aT(v)$, para quaisquer $a \in \mathbb{C}$ e $v \in V$.

Exemplo 10

Seja $T : \mathbb{C}^2 \longrightarrow \mathbb{C}$ defnida por $T(x, y) = 2ix + (1+i)y$.

Se $u = (x, y)$ e $v = (x', y')$, então $u + v = (x + x', y + y')$ e

$$\begin{aligned} T(u+v) &\overset{(1)}{=} T(x+x', y+y') \\ &\overset{(2)}{=} 2i(x+x') + (1+i)(y+y') \\ &\overset{(3)}{=} 2ix + 2ix' + (1+i)y + (1+i)y' \\ &\overset{(4)}{=} (2ix + (1+i)y) + (2ix' + (1+i)y') \\ &\overset{(5)}{=} T(u) + T(v) \end{aligned}$$

Em (1) usamos a definição da adição em $\mathbb{C}^2$; em (2), a definição de T; em (3), a distributividade em $\mathbb{C}$; em (4), a comutatividade e associatividade da adição em $\mathbb{C}$; em (5), novamente, a definição de T.

Observação: Continuam válidas as seguintes propriedades, onde V e W são $\mathbb{C}$-espaços vetoriais, que deixamos como exercício.

(1) Toda $T : V \longrightarrow W$ transformação $\mathbb{C}$-linear está perfeitamente determinada se é conhecida numa base de V.

(2) Se $T : V \longrightarrow W$ é $\mathbb{C}$-linear, então

$$\text{Núcleo}(T) = \{v \in V \; ; \; T(v) = 0_W\}$$

é um subespaço de V.

Além disso, T é injetora se, e somente se, $\text{Núcleo}(T) = \{0_V\}$.

(3) Se $T : V \longrightarrow W$ é $\mathbb{C}$-linear com $\dim_{\mathbb{C}} V = n$ e $\dim_{\mathbb{C}} W = m$, então $\dim_{\mathbb{C}} V = \dim_{\mathbb{C}} \text{Núcleo}(T) + \dim_{\mathbb{C}} \text{Imagem}(T)$.

(4) Seja $T : V \longrightarrow W$ $\mathbb{C}$-linear bijetora. Então, a função $T^{-1} : W \longrightarrow V$ é $\mathbb{C}$-linear.

(5) Seja $T : V \longrightarrow W$ $\mathbb{C}$-linear com $\dim_{\mathbb{C}} V = n$ e $\dim_{\mathbb{C}} W = m$. Sejam $\alpha = \{v_1, \ldots, v_n\}$ e $\beta = \{w_1, \ldots, w_m\}$ bases de V e W, respectivamente. Seja $A \in M_{m\times n}(\mathbb{C})$ definida por $A = T]^{\alpha}_{\beta} = (T(v_1)]_\beta \, T(v_2)]_\beta \cdots T(v_n)]_\beta)$. Se $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$, então $v]_\alpha = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ e $T(v)]_\beta = A\, v]_\alpha = T]^{\alpha}_{\beta} v]_\alpha$.

Exercícios

1. Seja V um $\mathbb{C}$-espaço vetorial. Mostre que:

   (a) O elemento neutro de V é único.

   (b) $a \cdot 0_V = 0_V$, para todo $a \in \mathbb{C}$.

(c) $0 \cdot v = 0_V$, para todo $v \in V$.

(d) Para cada $v \in V$, o simétrico de v é único.

(e) Para cada $v \in V$, $(-1) \cdot v$ é o simétrico de v.

2. Consideremos o $\mathbb{C}$-espaço vetorial $\mathbb{C}^3$ e seja
$$W = \{(x, y, z) \in \mathbb{C}^3 \,;\ x + y = 0, x + z = 0\}.$$

   (a) Mostre que W é um subespaço de $\mathbb{C}^3$.

   (b) Determine uma base e a dimensão de W.

3. Descreva todos os subespaços do $\mathbb{C}$-espaço vetorial $\mathbb{C}^2$.

4. Mostre que $\dim_{\mathbb{C}} M_{m\times n}(\mathbb{C}) = m \cdot n$.

5. Considere $V = M_{n\times n}(\mathbb{C})$ como $\mathbb{C}$-espaço vetorial.

   Seja $W = \{A \in V \,;\ A = A^t\}$.

   (a) Mostre que W é um subespaço de V.

   (b) Determine a dimensão de W.

6. Considere $\mathbb{C}$ como $\mathbb{R}$-espaço vetorial.

   (a) Mostre que $\dim_{\mathbb{R}} \mathbb{C} = 2$.

   (b) Mostre que $\dim_{\mathbb{R}} \mathbb{C}^n = 2 \cdot n$.

   (c) Mostre que $\dim_{\mathbb{R}} M_{m\times n}(\mathbb{C}) = 2 \cdot m \cdot n$.

7. Verifique que a função T é uma transformação $\mathbb{C}$-linear:

   (a) $T : \mathbb{C}^2 \longrightarrow \mathbb{C}^3$ definida por
   $T(x, y) = (x + iy, (1 + i)x - iy, (2 - i)x - 2y)$.

   (b) $T : \mathbb{C}^2 \longrightarrow \mathbb{C}$ definida por $T(x, y) = (1 + 2i)x + (1 - i)y$.

   (c) $T : \mathbb{C}^2 \longrightarrow \mathbb{C}^2$ definida por $T(x, y) = (x + iy, 2ix - 2y)$

8. No Exercício anterior, determine o núcleo e a imagem de cada T. Verifique o teorema do núcleo e da imagem.

9. Sejam U, V e W $\mathbb{C}$-espaços vetoriais. Mostre que:

   (a) Se $T : V \longrightarrow W$ é $\mathbb{C}$-linear, então

      i. Núcleo(T) é um subespaço de V.

ii. T é injetora se, e somente se, Núcleo$(T) = \{0_V\}$.

(b) Se $S : U \longrightarrow V$ e $T : V \longrightarrow W$ são $\mathbb{C}$-lineares, então a função $T \circ S : U \longrightarrow W$ é $\mathbb{C}$-linear.

(c) Se $T : V \longrightarrow W$ é $\mathbb{C}$-linear e bijetora, então a função inversa de T, $T^{-1} : W \longrightarrow V$, é $\mathbb{C}$-linear.

(d) Se $\{v_1, \ldots, v_n\}$ é uma base de V, $\{w_1, \ldots, w_n\} \subset W$ e $v = a_1v_1 + \cdots + a_nv_n$, definindo $T(v) = a_1w_1 + \cdots + a_nw_n$, então temos que $T(v_j) = w_j$, para cada $j = 1, \ldots, n$, e $T : V \longrightarrow W$ é $\mathbb{C}$-linear.

Subespaços invariantes e autovetores

Vamos, daqui por diante, trabalhar com K-espaços vetoriais, onde $K = \mathbb{R}$ ou $K = \mathbb{C}$.

Introduzimos agora um tipo especial de subespaço que será muito importante para entendermos, geometricamente, um operador linear num espaço vetorial.

Definição 7 (Subespaço invariante)
Sejam V um K-espaço vetorial e $T : V \longrightarrow V$ um operador K-linear. Um subespaço W de V é chamado *subespaço invariante por* T se, e somente se, $T(W) \subset W$.

Exemplo 11
Sejam V um K-espaço vetorial e $T : V \longrightarrow V$ um operador K-linear. Então, $\{0_V\}$ e V são invariantes por T.

Exemplo 12
Seja $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definida por $T(x, y, z) = \left(x, \frac{3}{5}y - \frac{4}{5}z, \frac{4}{5}y - \frac{3}{5}z\right)$.

Consideremos o plano Π com equação $x = 0$.

Planos que passam pela origem são subespaços do $\mathbb{R}^3$.

Esse plano é um subespaço invariante por T. De fato, se $v = (0, y, z) \in \Pi$, então

$$T(v) = T(0, y, z) = \left(0, \tfrac{3}{5}y - \tfrac{4}{5}z, \tfrac{4}{5}y - \tfrac{3}{5}z\right) \in \Pi,$$

logo $T(\Pi) \subset \Pi$.

Consideremos agora o subespaço $W = [(1, 0, 0)]$. Esse subespaço do $\mathbb{R}^3$ também é invariante por T. De fato, se $w \in W$, então existe $a \in \mathbb{R}$ tal que $w = a(1, 0, 0) = (a, 0, 0)$. Logo, $T(w) = T(a, 0, 0) = (a, 0, 0) = w \in W$.

Há um tipo de subespaço invariante muito interessante, conforme veremos a seguir.

Definição 8 (Autovalor e autovetor)
Sejam V um K-espaço vetorial e $T : V \longrightarrow V$ um operador K-linear. Um elemento $\lambda \in K$ é chamado um *autovalor* de T e $v \in V$, $v \neq 0_V$, é chamado um *autovetor associado ao autovalor* λ se, e somente se, $T(v) = \lambda v$.

Exemplo 13
No Exemplo anterior, $v = (a, 0, 0)$, com $a \neq 0$, $a \in \mathbb{R}$, é um autovetor de T associado ao autovalor 1.

Observação: O vetor 0_V foi excluído da definição de autovetor. Por quê? Como T é K-linear, temos que $T(0_V) = 0_V = \lambda \cdot 0_V$, *para qualquer* $\lambda \in K$.

Exemplo 14
Se $V \neq 0_V$ é um K-espaço vetorial e $I_V : V \longrightarrow V$ é definido por $I_V(v) = v$, para qualquer $v \in V$, então $\lambda = 1$ é um autovalor do operador linear I_V e qualquer $v \in V$, $v \neq 0_V$, é um autovetor associado ao autovalor 1.

t é uma indeterminada e $\mathbb{R}[t]$ é o conjunto dos polinômios com coeficientes reais com as operações usuais de adição de polinômios e multiplicação de um número real por um polinômio. Temos que $\dim_{\mathbb{R}} \mathbb{R}[t] = \infty$ e $\{1, t, t^2, \ldots\}$ é uma base de $\mathbb{R}[t]$.

Exemplo 15
Sejam $n \geq 0$ um natural e $P_n(\mathbb{R})$, subespaço vetorial do espaço vetorial real $\mathbb{R}[t]$, definido por

$$P_n(\mathbb{R}) = \{a_0 + a_1 t + \cdots + a_n t^n \,;\; a_j \in \mathbb{R}, \text{ para todo } j = 0, \ldots, n\}.$$

Seja $D : P_n(\mathbb{R}) \longrightarrow P_n(\mathbb{R})$ o operador $\mathbb{R}$-linear derivação. Para cada $a \in \mathbb{R}$ temos que $D(a) = 0 = 0 \cdot a$. Logo, todo $a \in \mathbb{R}$, $a \neq 0$, é um autovetor associado ao autovalor $\lambda = 0$.

Exemplo 16
Seja $C^\infty = \{f : \mathbb{R} \longrightarrow \mathbb{R}$ funções com derivadas de todas as ordens $\}$.

Consideremos os operadores lineares $D : C^\infty \longrightarrow C^\infty$, operador linear derivação, e $D^2 : C^\infty \longrightarrow C^\infty$, operador linear derivação de segunda ordem.

Como $D\big(e^{at}\big) = ae^{at}$, então todo $a \in \mathbb{R}$ é autovalor de D e $f(t) = e^{at}$ é autovetor de D associado ao autovalor a.

Como $D^2(\cos t) = -\cos t$ e $D^2(\operatorname{sen} t) = -\operatorname{sen} t$, então -1 é um autovalor de D^2 e $f(t) = \cos t$ e $g(t) = \operatorname{sen} t$ são autovetores de D^2 associados ao autovalor -1.

Mais ainda, $D^2(e^t) = e^t$ e $D^2(e^{-t}) = e^{-t}$, logo e^t e e^{-t} são autovetores de D^2 associados ao autovalor 1.

Para todo real $a > 0$, temos que $e^{\sqrt{a}t}$ e $e^{-\sqrt{a}t}$ são autovetores de D^2 associados ao autovalor a.

Exemplo 17
Seja $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ o operador linear definido por $T(x, y) = (y, x)$.

Observamos que

$T(1, 1) = (1, 1) = 1 \cdot (1, 1)$ e

$T(-1, 1) = (1, -1) = (-1) \cdot (-1, 1)$.

Portanto, 1 e -1 são autovalores de T e $(1, 1)$ é autovetor associado ao autovalor 1, enquanto $(-1, 1)$ é autovetor associado ao autovalor -1.

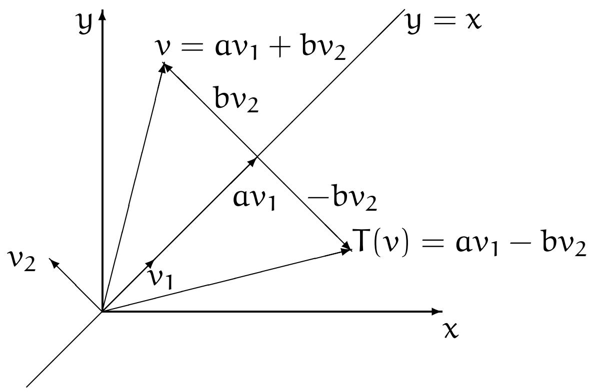
Como $\beta = \{v_1 = (1, 1), v_2 = (-1, 1)\}$ é uma base de $\mathbb{R}^2$, todo $v \in \mathbb{R}^2$ se escreve como uma combinação linear única dessa base, assim, existem $a, b \in \mathbb{R}$, unicamente determinados, tais que

$$v = av_1 + bv_2.$$

Como qualquer transformação linear está perfeitamente determinada pelos seus valores numa base, temos que

$$\begin{aligned} T(v) &= aT(v_1) + bT(v_2) \\ &= av_1 + b(-v_2) \\ &= av_1 - bv_2. \end{aligned}$$

T tem a propriedade de fixar a componente de v na direção de v_1 e mandar no seu simétrico a componente de v na direção de v_2. Geometricamente, T é a simetria com respeito à reta $y = x$, que é a reta que passa pela origem gerada por v_1.



Nesse caso, $T]_\beta^\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Exemplo 18

Seja $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definida por $T(x, y, z) = \left(\frac{2x-y-z}{3}, \frac{-x+2y-z}{3}, \frac{-x-y+2z}{3} \right)$.

Quem é esse operador $\mathbb{R}$-linear? A fórmula acima não dá nenhuma informação, apenas permite determinar os seus valores em cada ponto.

Sejam $v_1 = (1, 1, 1)$, $v_2 = (1, -1, 0)$ e $v_3 = (1, 1, -2)$.

Observamos que $T(v_1) = (0, 0, 0) = 0 \cdot v_1$, $T(v_2) = (1, -1, 0) = v_2$ e $T(v_3) = (1, 1, -2) = v_3$. Portanto, v_1 é autovetor de T associado ao autovalor 0 e v_2 e v_3 são autovetores de T associados ao autovalor 1.

Geometricamente, T é a projeção sobre o plano gerado por v_2 e v_3 segundo a direção da reta gerada por v_1. Por quê?

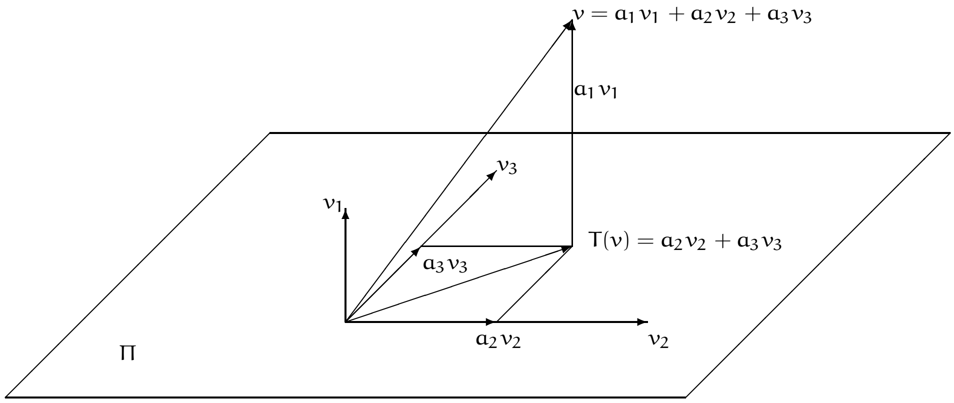
O conjunto $\beta = \{v_1, v_2, v_3\}$ é uma base do $\mathbb{R}^3$. Cada $v \in \mathbb{R}^3$ se escreve de uma única maneira como combinação linear de β. Assim, existem $a_1, a_2, a_3 \in \mathbb{R}$, unicamente determinados tais que

$$v = a_1v_1 + a_2v_2 + a_3v_3.$$

Portanto,

Faça um desenho ilustrativo com v e $T(v)$, como no Exemplo anterior.

$$\begin{aligned} T(v) &= a_1T(v_1) + a_2T(v_2) + a_3T(v_3) \\ &= a_1.(0,0,0) + a_2v_2 + a_3v_3 \\ &= a_2v_2 + a_3v_3 \end{aligned}$$



Lembramos que se $u = (x_1, x_2, x_3)$ e $v = (y_1, y_2, y_3)$, então o *produto interno usual* do $\mathbb{R}^3$ é definido por $u \cdot v = \sum_{j=1}^{3} x_j \cdot y_j$. Vale que $u \perp v$, se e somente se, $u \cdot v = 0$.

Geometricamente, T fixa as componentes de v nas direções de v_2 e v_3, logo fixa a componente de v sobre o plano gerado por v_2 e v_3, e manda em $(0,0,0)$ a componente de v na direção de v_1. Portanto, T projeta cada $v \in \mathbb{R}^3$ sobre o plano gerado por v_2 e v_3, segundo a direção de v_1. Como v_1 é ortogonal a v_2 e a v_3, então T é a projeção ortogonal sobre o plano $x+y+z=0$. Nesse caso, $T]_\beta^\beta = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

O Exemplo anterior motiva a seguinte Proposição.

Proposição 1

Sejam V um K-espaço vetorial e $v_1, \dots, v_s$ autovetores associados ao autovalor λ do operador K-linear T em V. Então, $W = [v_1, \dots, v_s]$ é invariante por T.

Demonstração: Seja $w \in W$. Então, existem $a_1, \dots, a_s \in K$ tais que

$$w = a_1v_1 + \cdots + a_sv_s. \qquad (\star)$$

Logo,

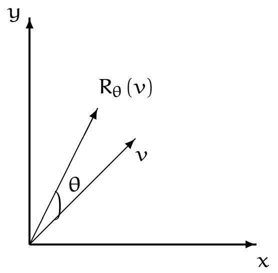
$$\begin{aligned} T(w) &\overset{(1)}{=} T(a_1v_1 + \cdots + a_sv_s) \\ &\overset{(2)}{=} a_1T(v_1) + \cdots + a_sT(v_s) \\ &\overset{(3)}{=} a_1(\lambda v_1) + \cdots + a_s(\lambda v_s) \\ &\overset{(4)}{=} \lambda(a_1v_1 + \cdots + a_sv_s) \\ &\overset{(5)}{=} \lambda w \in W. \quad \blacksquare \end{aligned}$$

Em (1) usamos ($\star$); em (2), que T é linear; em (3), que $T(v_j) = \lambda v_j$; em (4), que a multiplicação por escalar é comutativa e distributiva; em (5), ($\star$).

Exemplo 19

Seja $R_\theta : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ o operador $\mathbb{R}$-linear rotação de θ radianos, no sentido positivo das rotações, onde $0 \leq \theta < 2\pi$. Lembramos que

$$R_\theta(x, y) = (x\cos\theta - y\,\mathrm{sen}\,\theta, x\,\mathrm{sen}\,\theta + y\cos\theta).$$



No momento essa fórmula não tem interesse. Geometricamente, a rotação muda a direção do vetor $v \neq (0, 0)$, sempre que $\theta \neq 0$ e $\theta \neq \pi$. Portanto, $R_\theta(v) = av$, com $v \neq (0, 0)$, se e somente se, $\theta = 0$ rd ou $\theta = \pi$ rd, isto é, $R_0 = I_{\mathbb{R}^2}$ ou $R_\pi = -I_{\mathbb{R}^2}$. Em ambos os casos, todo $v \neq (0, 0)$ é autovetor, sendo 1 o autovalor de R_0 e -1 o autovalor de R_π.

Além disso, R_θ, com $\theta \neq 0$ e $\theta \neq \pi$, é um operador $\mathbb{R}$-linear sem autovalores e autovetores.

Os Exemplos 17 e 18 mostram que o conceito de autovalores e autovetores permite compreender geometricamente esses operadores lineares, enquanto o Exemplo 19 mostra que pode não haver autovalores e autovetores. A pergunta natural é: como determinar, caso existam, os autovalores e autovetores de um operador linear?

Vamos a seguir desenvolver um método para responder à questão no caso dos espaços vetoriais de dimensão finita.

Para os nossos propósitos precisamos do conceito a seguir.

Definição 9 (Determinante de um operador)
Sejam V um K-espaço vetorial, $\dim_K(V) = n \geq 1$ e T um operador K-linear em V. Definimos $\det(T) = \det\left(T]^\alpha_\alpha\right)$, onde α é qualquer base de V.

Devemos mostrar que a definição acima está bem posta. De fato, sejam α e β bases de V, então

$$\begin{aligned} T]^\beta_\beta &= I]^\alpha_\beta \cdot T]^\alpha_\alpha \cdot I]^\beta_\alpha \\ &= \left(I]^\beta_\alpha\right)^{-1} \cdot T]^\alpha_\alpha \cdot I]^\beta_\alpha \end{aligned}$$

Logo,

$$\begin{aligned} \det\left(T]^\beta_\beta\right) &= \det\left(\left(I]^\beta_\alpha\right)^{-1} \cdot T]^\alpha_\alpha \cdot I]^\beta_\alpha\right) \\ &= \det\left(\left(I]^\beta_\alpha\right)^{-1}\right) \cdot \det\left(T]^\alpha_\alpha\right) \cdot \det\left(I]^\beta_\alpha\right) \\ &= \left(\det\left(I]^\beta_\alpha\right)\right)^{-1} \det \cdot \left(T]^\alpha_\alpha\right) \cdot \det\left(I]^\beta_\alpha\right) \\ &= \det\left(T]^\alpha_\alpha\right), \end{aligned}$$

$\det(A^{-1}) = (\det(A))^{-1}$ e a multiplicação em K é comutativa.

mostrando que o valor $\det(T)$ não depende da base escolhida.

Proposição 2

Sejam V um K-espaço vetorial, $\dim_K V = n \geq 1$ e T um operador K-linear em V. T é invertível se, e somente se, $\det(T) \neq 0$.

Demonstração: Seja $T : V \longrightarrow V$ um operador K-linear. Pelo Teorema do núcleo e da imagem, T é injetor se, e somente se, T é sobrejetor. Assim,

$$\begin{array}{lll} T \text{ é invertível} & \text{se, e somente se,} & T \text{ é bijetor} \\ & \text{se, e somente se,} & \text{existe } S : V \longrightarrow V, K\text{-linear,} \\ & & \text{tal que } I_V = S \circ T \text{ e } I_V = T \circ S \\ & \text{se, e somente se,} & \text{para qualquer base } \beta \text{ de } V, \\ & & I = S]^{\beta}_{\beta} \cdot T]^{\beta}_{\beta} \text{ e } I = T]^{\beta}_{\beta} \cdot S]^{\beta}_{\beta}, \\ & & \text{onde } I = I_V]^{\beta}_{\beta} \\ & \text{se, e somente se,} & 1 = \det(S) \cdot \det(T) \\ & \text{se, e somente se,} & \det(T) \neq 0_K. \quad \blacksquare \end{array}$$

Lembramos que $I_V : V \longrightarrow V$ é definida por $I_V(v) = v$, $\forall\, v \in V$.

Nesse caso, a expressão *para qualquer base* é equivalente a *para alguma base*.

Agora estamos prontos para desenvolver o método para determinação de autovalores e autovetores de operadores lineares em espaços vetoriais de dimensão finita, caso existam.

Teorema 1

Sejam V um K-espaço vetorial, $\dim_K V = n \geq 1$ e T um operador K-linear em V. As seguintes condições são equivalentes:

(a) $\lambda \in K$ é autovalor de T.

(b) O operador K-linear $\lambda I - T$ não é invertível.

(c) $\det(\lambda I - T) = 0$.

Demonstração:

$$\begin{array}{lll} \lambda \in K \text{ é autovalor de } T & \text{se, e somente se,} & \text{existe } v \in V,\ v \neq 0_V, \text{ tal que} \\ & & T(v) = \lambda v = \lambda I_V(v) \\ & \text{se, e somente se,} & \text{existe } v \in V,\ v \neq 0_V, \text{ tal que} \\ & & (\lambda I_V - T)(v) = 0_V \\ & \text{se, e somente se,} & \lambda I_V - T \text{ não é invertível} \\ & \text{se, e somente se,} & \det(\lambda I_V - T) = 0. \quad \blacksquare \end{array}$$

Com as notações do Teorema acima, seja α uma base qualquer de V, seja $A = T]^{\alpha}_{\alpha} \in M_{n\times n}(K)$. Então, $(\lambda I_V - T)]^{\alpha}_{\alpha} = (\lambda I_V)]^{\alpha}_{\alpha} - T]^{\alpha}_{\alpha} = \lambda I - A$.

Logo,

$\lambda \in K$ é autovalor de T se, e somente se, $\det(\lambda I_V - T) = \det(\lambda I - A) = 0$.

Escrevendo $A = (a_{ij})$, com $i, j = 1, \ldots, n$, temos que

$$\lambda I - A = \begin{pmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{pmatrix}.$$

Assim, $p(\lambda) = \det(\lambda I - A)$ é um polinômio mônico de grau n com coeficientes em K, chamado de *polinômio característico* de T.

Um polinômio não nulo cujo coeficiente líder (coeficiente do termo de mais alto grau) é 1 é chamado de *polinômio mônico.*

$\lambda \in K$ é autovalor de T se, e somente se, $p(\lambda) = 0$
se, e somente se, $\lambda \in K$ é raiz do polinômio característico de T.

Quem são os autovetores de T, caso existam autovalores?

Escrevendo $\alpha = \{v_1, \ldots, v_n\}$, temos que para cada $v \in V$, existem $x_1, \ldots, x_n$ em K, tais que $v = x_1v_1 + \cdots + x_nv_n$.

Seja $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = v]_\alpha$.

Lembramos que $v \neq 0_V$ se, e somente se, $X = v]_\alpha \neq 0$.

Denotaremos $v \in V$, $v \neq 0_V$, autovetor de T associado ao autovalor λ por $v \leftrightarrow \lambda$. Portanto,

$v \leftrightarrow \lambda$ se, e somente se, $(\lambda I_V - T)(v) = 0_V$, com $v \neq 0_V$
se, e somente se, $(\lambda I_V - T)]_\alpha^\alpha v]_\alpha = 0_V]_\alpha = 0$, com $v]_\alpha \neq 0$
se, e somente se, $(\lambda I - A)X = 0$, com $X \neq 0$.

Para determinar os autovetores de T, resolvemos o sistema linear homogêneo, cuja matriz associada é $\lambda I - A$. Cada solução $X \neq 0$ desse sistema corresponde ao autovetor $v = x_1v_1 + \cdots + x_nv_n$ de T, pois $v]_\alpha = X$.

Exemplo 20

Vamos determinar, caso existam, os autovalores e autovetores do operador $\mathbb{R}$-linear $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definido por $T(x, y) = (-y, x)$.

Tomamos $\alpha = \{(1, 0), (0, 1)\}$, a base canônica do $\mathbb{R}^2$.

Então, $A = T]_\alpha^\alpha = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ e $\lambda I - A = \begin{pmatrix} \lambda & 1 \\ -1 & \lambda \end{pmatrix}$. Logo, o polinômio caraterístico de T é $p(\lambda) = \det(\lambda I - A) = \lambda^2 + 1 \in \mathbb{R}[\lambda]$. Como esse polinômio não tem raízes reais, T não tem autovalores, consequentemente, não tem autovetores. Observamos que $T = R_{\frac{\pi}{2}}$.

Antes de mais um exemplo, introduzimos uma terminologia.

Verifique que V_λ é um subespaço de V.

Definição 10 (Subespaço característico)
Sejam V um K-espaço vetorial de dimensão $n \geq 1$, T um operador K-linear em V e $\lambda \in K$. O conjunto

$$V_\lambda = \{v \in V\ ;\ T(v) = \lambda v\}$$

é chamado de *subespaço característico* de T.

Observação:

(1) $V_\lambda \neq \{0_V\}$ se, e somente se, λ é autovalor de T.

(2) $V_\lambda = \{0_V\} \cup \{v \in V\,;\, v$ é autovetor associado ao autovalor $\lambda\}$.

Exemplo 21
Consideremos $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definida por

$$T(x, y, z) = (5x - 6y - 6z, -x + 4y + 2z, 3x - 6y - 4z).$$

Tomamos $\alpha = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ a base canônica do $\mathbb{R}^3$.

Então, $T]^\alpha_\alpha = \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix}$ e $\lambda I - A = \begin{pmatrix} \lambda - 5 & 6 & 6 \\ 1 & \lambda - 4 & -2 \\ -3 & 6 & \lambda + 4 \end{pmatrix}$.

Para fatorar $p(\lambda)$ pesquisamos, primeiramente, as possíveis raízes racionais. Lembramos que se $p(\lambda)$ tem coeficientes inteiros e a fração irredutível $\frac{a}{b}$ é sua raiz, então a divide o termo constante e b divide o coeficiente líder.

O polinômio característico de T é

$$p(\lambda) = \det(\lambda I - A) = \lambda^3 - 5\lambda^2 + 8\lambda - 4 = (\lambda - 1)(\lambda - 2)^2 \in \mathbb{R}[\lambda].$$

Temos os autovalores reais 1 e 2.

Vamos determinar os autovetores, resolvendo os sistemas lineares homogêneos correspondentes.

$\boxed{\lambda = 1}$

$$I - A = \begin{pmatrix} -4 & 6 & 6 \\ 1 & -3 & -2 \\ -3 & 6 & 5 \end{pmatrix} \sim_1 \begin{pmatrix} 0 & -6 & -2 \\ 1 & -3 & -2 \\ 0 & -3 & -1 \end{pmatrix} \sim_2 \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ 0 & -3 & -1 \end{pmatrix}$$

$$\sim_3 \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & \frac{1}{3} \end{pmatrix}$$

Usamos a seguinte sequência de operações elementares:

em $\sim_1$: $L_1 \to L_1 + 4L_2$ e $L_3 \to L_3 + 3L_2$;

em $\sim_2$: $L_1 \to L_1 - 2L_3$ e $L_2 \to L_2 - L_3$;

em $\sim_3$: $L_3 \leftrightarrow -\frac{1}{3}L_3$.

A solução do sistema é o subespaço característico

$$V_{\lambda=1} = \{(x, y, z) \in \mathbb{R}^3\ ;\ x - z = 0, y + \tfrac{1}{3}z = 0\} = \{\left(z, -\tfrac{1}{3}z, z\right)\ ;\ z \in \mathbb{R}\}.$$

Logo, $\lambda = 1 \leftrightarrow v = z\left(1, -\frac{1}{3}, 1\right)$, com $z \neq 0$.

$\boxed{\lambda = 2}$

$$2I - A = \begin{pmatrix} -3 & 6 & 6 \\ 1 & -2 & -2 \\ -3 & 6 & 6 \end{pmatrix} \sim_1 \begin{pmatrix} 0 & 0 & 0 \\ 1 & -2 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

Usamos a seguinte sequência de operações elementares em $\sim_1$: $L_1 \to L_1 + 3L_2$ e $L_3 \to L_3 + 3L_2$.

A solução do sistema é o subespaço característico

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

Logo, todo vetor não nulo do plano acima é autovetor de T associado ao autovalor $\lambda = 2$.

Nesse caso, existe uma base β do $\mathbb{R}^3$ formada por autovetores de T, por exemplo $\beta = \{v_1 = (3, -1, 3), v_2 = (0, 1, -1), v_3 = (2, 0, 1)\}$, onde v_1 está na reta $V_{\lambda=1}$ e v_2 e v_3 foram escolhidos no plano $V_{\lambda=2}$.

Observamos que $T]^{\beta}_{\beta} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

Agora, entendemos geometricamente o operador linear T. Quem é T?

Exemplo 22

Consideremos $V = \left\{ \begin{pmatrix} x & y \\ z & 0 \end{pmatrix}\ ;\ x, y, z \in \mathbb{R} \right\}$ e seja $T : V \longrightarrow V$ definida por $T \begin{pmatrix} x & y \\ z & 0 \end{pmatrix} = \begin{pmatrix} 2x - y + 2z & 3x \\ -2x + 4y + z & 0 \end{pmatrix}$. Vamos determinar os autovalores e autovetores de T.

Primeiramente, V é um espaço vetorial real, pois é um subespaço de $M_{2\times 2}(\mathbb{R})$ e T é linear. Precisamos de uma base de V. Afirmamos que $\dim_{\mathbb{R}} V = 3$. De fato, dado $v \in V$ temos que

$$\begin{aligned} v &= \begin{pmatrix} x & y \\ z & 0 \end{pmatrix} \\ &= \begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & y \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ z & 0 \end{pmatrix} \\ &= x \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + y \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + z \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \in [v_1, v_2, v_3], \end{aligned}$$

onde $v_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ e $v_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.

Como esses vetores são linearmente independentes, então $\alpha = \{v_1, v_2, v_3\}$ é uma base de V e $v]_\alpha = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

Seja $A = T]_\alpha^\alpha = \begin{pmatrix} 2 & -1 & 2 \\ 3 & 0 & 0 \\ -2 & 4 & 1 \end{pmatrix}$. Então, $\lambda I - A = \begin{pmatrix} \lambda-2 & 1 & -2 \\ -3 & \lambda & 0 \\ 2 & -4 & \lambda-1 \end{pmatrix}$.

Temos $p(\lambda) = \det(\lambda I - A) = \lambda^3 - 3\lambda^2 + 9\lambda - 27 = \lambda^2(\lambda-3) + 9(\lambda-3) = (\lambda^2+9)(\lambda-3)$.

Portanto, $\lambda = 3$ é o único autovalor de T. Para determinar os autovetores associados devemos resolver o sistema linear $(3I - A)X = 0$, onde $X = v]_\alpha$.

Como $3I - A = \begin{pmatrix} 1 & 1 & -2 \\ -3 & 3 & 0 \\ 2 & -4 & 2 \end{pmatrix}$, reduzindo por linhas, temos

$$3I - A \sim_1 \begin{pmatrix} 1 & 1 & -2 \\ 0 & 6 & -6 \\ 0 & -6 & 6 \end{pmatrix} \sim_2 \begin{pmatrix} 1 & 1 & -2 \\ 0 & 1 & -1 \\ 0 & -6 & 6 \end{pmatrix} \sim_3 \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix},$$

com a seguinte sequência de operações elementares:

em $\sim_1$: $L_2 \longrightarrow L_2 + 3L_1, L_3 \longrightarrow L_3 - 2L_1$;

em $\sim_2$: $L_2 \longleftrightarrow \frac{1}{6}L_2$ e

em $\sim_3$: $L_1 \longrightarrow L_1 - L_2$ e $L_3 \longrightarrow L_3 + 6L_2$.

O conjunto solução do sistema é

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

Logo, o subespaço característico é

$$V_{\lambda=3} = \left\{ v \in V \,;\; v = zv_1 + zv_2 + zv_3 = \begin{pmatrix} z & z \\ z & 0 \end{pmatrix} = z \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \right\}.$$

Assim, os autovetores de T associados a $\lambda = 3$ são $v = z\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, com $z \neq 0$.

Nesse caso, temos $\dim_{\mathbb{R}} V = 3$ e o subespaço característico $V_{\lambda=3}$ tem dimensão 1. Não é possível construir uma base de V formada por autovetores de T.

Exemplo 23

Vamos determinar os autovalores e autovetores de $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definida por $T(x,y,z) = (2z, x+z, y-2z)$.

Seja $\alpha = \{(1,0,0),(0,1,0),(0,0,1)\}$ e seja $A = T]_\alpha^\alpha = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & -2 \end{pmatrix}$. Então,

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

T tem três autovalores distintos: 1, -1 e -2.

$\boxed{\lambda = -1}$

$-I - A = \begin{pmatrix} -1 & 0 & -2 \\ -1 & -1 & -1 \\ 0 & -1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$. Logo, o subespaço característico é

$$V_{\lambda=-1} = \{(x,y,z) \in \mathbb{R}^3 \,;\; x + 2z = 0 \text{ e } y - z = 0\} = \{(-2z,z,z) \,;\; z \in \mathbb{R}\}$$

e

$v \leftrightarrow \lambda = -1$ se, e somente se, $v = z(-2,1,1)$, com $z \in \mathbb{R}$ e $z \neq 0$.

$\boxed{\lambda = -2}$

$-2I - A = \begin{pmatrix} -2 & 0 & -2 \\ -1 & -2 & -1 \\ 0 & -1 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Logo, o subespaço característico é

$$V_{\lambda=-2} = \{(x,y,z) \in \mathbb{R}^3 \,;\; x + z = 0 \text{ e } y = 0\} = \{(-z,0,z) \,;\; z \in \mathbb{R}\}$$

e

$v \leftrightarrow \lambda = -2$ se, e somente se, $v = z(-1,0,1)$, com $z \in \mathbb{R}$ e $z \neq 0$.

$\boxed{\lambda = 1}$

$I - A = \begin{pmatrix} 1 & 0 & -2 \\ -1 & 1 & -1 \\ 0 & -1 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{pmatrix}$. Logo, o subespaço característico é

$$V_{\lambda=1} = \{(x,y,z) \in \mathbb{R}^3 \,;\; x - 2z = 0 \text{ e } y - 3z = 0\} = \{(2z,3z,z) \,;\; z \in \mathbb{R}\}$$

e

$v \leftrightarrow \lambda = 1$ se, e somente se, $v = z(2,3,1)$, com $z \in \mathbb{R}$ e $z \neq 0$.

Nesse caso, existe uma base β do $\mathbb{R}^3$ formada por autovetores de T, digamos $\beta = \{v_1 = (-2,1,1), v_2 = (-1,0,1), v_3 = (2,3,1)\}$, onde escolhemos os

Por que $\{v_1, v_2, v_3\}$ é linearmente independente sobre $\mathbb{R}$?

vetores $v_1 \in V_{\lambda=1}$, $v_2 \in V_{\lambda=-2}$ e $v_3 \in V_{\lambda=-1}$. Temos $T]_\beta^\beta = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Exercícios

1. Seja V um K-espaço vetorial com $\dim_K V \geq 1$, onde $K = \mathbb{R}$ ou $K = \mathbb{C}$.

   (a) Seja I_V o operador identidade em V. Mostre que todo subespaço W de V é invariante por I_V.

   (b) Mostre que $W = \{0_V\}$ é um subespaço invariante por T, para todo operador K-linear T.

   (c) Seja $T : V \longrightarrow V$ um operador K-linear.

      i. Mostre que Núcleo(T) é invariante por T.

      ii. Sejam $f(x) = a_0 + a_1x + \cdots + a_nx^n$ em $K[x]$ e o operador definido por $f(T) = a_0I_V + a_1T + \cdots + a_nT^n$. Mostre que Núcleo$(f(T))$ é invariante por T.

      iii. Se T não é injetor mostre $\lambda = 0$ é um autovalor de T.

2. Sejam V um K-espaço vetorial, $K = \mathbb{R}$ ou $K = \mathbb{C}$, $v_1, \ldots, v_s$ autovetores do operador K-linear T em V associados ao autovalor λ e $W = [v_1, \ldots, v_s]$. Mostre que todo $w \in W$, $w \neq 0$, é autovetor de T associado ao autovalor λ.

3. Sejam V um K-espaço vetorial de dimensão $n \geq 1$, T um operador linear em V e $\lambda \in K$. Mostre que $V_\lambda = \{v \in V\,;\, T(v) = \lambda v\}$ é um subespaço de V.

4. Seja $T : \mathbb{C}^2 \longrightarrow \mathbb{C}^2$ definida por $T(x, y) = (-y, x)$.

   (a) Mostre que T tem dois autovalores distintos.

   (b) Determine os autovetores associados a cada um dos autovalores.

   (c) Construa uma base β de $\mathbb{C}^2$ formada por autovetores de T e dê $T]_\beta^\beta$.

5. Interprete geometricamente o operador $\mathbb{R}$-linear do Exemplo 21.

6. Seja $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ a transformação linear definida por
$$T(x, y, z) = (x - 3y + 3z, 3x - 5y + 3z, 6x - 6y + 4z).$$
Mostre que $W = \{(x, y, z) \; ; \; x - y + z = 0\}$ é um subespaço invariante por T.

7. Sejam $A \in M_{n\times n}(\mathbb{R})$ e $T_A : \mathbb{R}^n \to \mathbb{R}^n$ a transformação linear, tal que $T_A]^{\alpha}_{\alpha} = A$, onde α é a base canônica do $\mathbb{R}^n$.

   Mostre que os escalares indicados são autovalores de T_A, justificando sua resposta, e determine uma base para o subespaço característico do $\mathbb{R}^n$ associado a cada autovalor.

   (a) $A = \begin{pmatrix} 10 & -9 \\ 4 & -2 \end{pmatrix}$, $\lambda = 4$ (b) $A = \begin{pmatrix} 4 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & 3 & 4 \end{pmatrix}$, $\lambda \in \{1, 4\}$

   (c) $A = \begin{pmatrix} 4 & 0 & 1 \\ -2 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$ e $\lambda \in \{1, 2, 3\}$

   (d) $A = \begin{pmatrix} 3 & 0 & 2 & 0 \\ 1 & 3 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$ e $\lambda = 4$

Os autovalores e autovetores de $A \in M_{n\times n}(K)$ são os autovalores e autovetores de $T_A : K^n \longrightarrow K^n$, tal que $A = T]^{\alpha}_{\alpha}$, onde α é a base canônica de K^n.

8. Seja $A \in M_{n\times n}(K)$. Determine os autovalores de A, para $K = \mathbb{R}$ e $K = \mathbb{C}$.

   (a) $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ (b) $A = \begin{pmatrix} -1 & 0 & 0 \\ 2 & 3 & 0 \\ 1 & -1 & -2 \end{pmatrix}$

   (c) $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ (d) $A = \begin{pmatrix} 2 & -1 & 2 \\ 3 & 0 & 0 \\ -2 & 4 & 1 \end{pmatrix}$

9. Sejam $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$ e $B = \begin{pmatrix} 5 & 5 & 5 \\ 5 & 5 & 5 \\ 5 & 5 & 5 \end{pmatrix}$ matrizes em $M_{3\times 3}(\mathbb{R})$.

   Determine, sem fazer cálculos e justifique sua resposta:

   (a) Um autovalor de A.

   (b) Um autovalor e dois autovetores de B linearmente independentes associados a esse autovalor.

10. Seja $A \in M_{n\times n}(K)$, onde $K = \mathbb{R}$ ou $K = \mathbb{C}$. Mostre que:

    (a) Se A é matriz diagonal, então os autovalores de A são os elementos da sua diagonal principal.

    (b) $\lambda = 0$ é autovalor de A se, e somente se, A não é invertível.

    (c) A e A^t têm os mesmos autovalores.

11. Seja $T : V \longrightarrow V$ uma transformação K-linear, onde $K = \mathbb{R}$ ou $K = \mathbb{C}$ e $\dim_K V = n \geq 1$. Mostre que:

    (a) Se T é invertível e λ é autovalor de T, então $\lambda \neq 0$ e λ^{-1} é autovalor de T^{-1}.

    (b) Se T^2 é o operador nulo, então λ é autovalor de T se, e somente se, $\lambda = 0$.

12. Seja $T : V \longrightarrow V$ um operador K-linear, onde $K = \mathbb{R}$ ou $K = \mathbb{C}$.

    (a) Mostre que se $\lambda \in K$ é um autovalor de T, então λ^m é um autovalor de T^m, para todo inteiro $m \geq 1$.

    (b) Seja $f(t) = a_n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$ um polinômio com coeficientes em K. Mostre que se $\lambda \in K$ é um autovalor de T, então $f(\lambda)$ é um autovalor de $S = a_n T^n + a_{n-1}T^{n-1} + \cdots + a_1 T + a_0 I_V$, onde $I_V(v) = v$, para todo $v \in V$.

13. Sejam S e T operadores K-lineares e W um subespaço de V invariante por T e por S. Mostre que W é invariante por $S + T$ e por $S \circ T$.

14. Sejam S e T operadores K-lineares, tais que $S \circ T = T \circ S$. Sejam $\lambda \in K$ um autovalor de T e W o subespaço característico associado a λ.

    Mostre que W é um subespaço invariante por S.

15. Seja V um espaço vetorial real de dimensão ímpar $n \geq 3$. Mostre que para todo operador linear $T : V \longrightarrow V$ existe W subespaço de V, tal que $W \neq V$, $W \neq \{0\}$ e W é invariante por T.

16. Seja V um espaço vetorial complexo de dimensão $n \geq 2$. Mostre que para todo operador linear $T : V \longrightarrow V$ existe W subespaço de V, tal que $W \neq V$, $W \neq \{0\}$ e W é invariante por T.

Operadores diagonalizáveis

Os diversos exemplos da Seção anterior mostram que nem sempre existe uma base do espaço vetorial V formada por autovetores do operador linear em V. Vamos dar condições necessárias e suficientes para a existência de uma base de autovetores.

Definição 11 (Operador diagonalizável)
Sejam V um K-espaço vetorial com $\dim_K V = n \geq 1$. Dizemos que o operador K-linear T em V é *diagonalizável* se, e somente se, existe β base de V formada por autovetores de T.

A definição acima equivale à existência de uma base β de V tal que $T]_\beta^\beta$ é uma matriz diagonal. De fato, digamos que $\beta = \{v_1, \dots, v_n\}$ é uma base de V formada por autovetores de T e v_j é autovetor associado ao autovalor λ_j, para cada $j = 1, \dots, n$. Então, $T]_\beta^\beta = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$ é matriz diagonal.

$\lambda_1, \dots, \lambda_n$ não são necessariamente distintos.

Reciprocamente, se existe uma base $\beta = \{v_1, \dots, v_n\}$ de V tal que $T]_\beta^\beta = (a_{ij})$ é uma matriz diagonal, então $a_{ij} = 0$, para todo $1 \leq i \neq j \leq n$. Portanto, $T(v_j)]_\beta = \begin{pmatrix} 0 \\ \vdots \\ a_{jj} \\ \vdots \\ 0 \end{pmatrix}$, $T(v_j) = a_{jj}v_j$, para cada $j = 1, \dots, n$, e β é uma base de V formada por autovetores de T.

Se α é qualquer base de V, tomando a matriz de mudança de base $P = I]_\alpha^\beta$, temos $P^{-1} = I]_\beta^\alpha$ e

$$T]_\beta^\beta = I]_\beta^\alpha \cdot T]_\alpha^\alpha \cdot I]_\alpha^\beta = P^{-1}\, T]_\alpha^\alpha P.$$

Logo, para cada base α de V, existe uma matriz invertível $P \in M_{n\times n}(K)$ tal que $P^{-1}\, T]_\alpha^\alpha P$ é uma matriz diagonal. Dizemos que P *diagonaliza* T.

Exemplo 24
O operador $\mathbb{R}$-linear do Exemplo 23 é diagonalizável, assim como, os operadores $\mathbb{R}$-lineares dos Exemplos 17, 18 e 21. Enquanto, a rotação R_θ no plano, com $\theta \neq 0$ e $\theta \neq \pi$ não é diagonalizável.

Proposição 3

Sejam V um K-espaço vetorial, T um operador K-linear em V, $\lambda_1, \ldots, \lambda_s$ autovalores distintos de T e $v_1, \ldots, v_s$ autovetores associados, respectivamente, a $\lambda_1, \ldots, \lambda_s$. Então, $\{v_1, \ldots, v_s\}$ é K-linearmente independente.

Demonstração: Faremos indução sobre s. Seja v_1 autovetor de T associado ao autovalor λ_1. Como $v_1 \neq 0_V$, então $\{v_1\}$ é linearmente independente e a afirmação vale para $s = 1$.

Seja $s \geq 1$ e suponhamos a afirmação válida para s.

Sejam $v_1, \ldots, v_s, v_{s+1}$ autovetores de T associados, respectivamente, aos autovalores distintos $\lambda_1, \ldots, \lambda_s, \lambda_{s+1}$. Consideremos $a_1, \ldots, a_{s+1} \in K$ tais que

$$0_V = a_1 v_1 + \cdots + a_s v_s + a_{s+1} v_{s+1}. \qquad (\star)$$

Aplicando T, obtemos

$$\begin{aligned} 0_V = T(0_V) &= a_1 T(v_1) + \cdots + a_s T(v_s) + a_{s+1} T(v_{s+1}) \\ &= a_1 \lambda_1 v_1 + \cdots + a_s \lambda_s v_s + a_{s+1} \lambda_{s+1} v_{s+1}. \quad (\star\star) \end{aligned}$$

Multiplicando a igualdade $(\star)$ por λ_{s+1}, obtemos

$$0_V = a_1 \lambda_{s+1} v_1 + \cdots + a_s \lambda_{s+1} v_s + a_{s+1} \lambda_{s+1} v_{s+1}. \qquad (\star\star\star)$$

Subtraindo $(\star\star\star)$ de $(\star\star)$, obtemos

$$0_V = a_1(\lambda_1 - \lambda_{s+1}) v_1 + \cdots + a_s(\lambda_s - \lambda_{s+1}) v_s.$$

Pela hipótese de indução, temos que $\{v_1, \ldots, v_s\}$ é K-linearmente independente, logo $a_j(\lambda_j - \lambda_{s+1}) = 0$, com $\lambda_j \neq \lambda_{s+1}$, para $j = 1, \ldots, s$. Portanto, $a_j = 0$, para $j = 1, \ldots, s$.

Substituindo em $(\star)$, temos $a_{s+1} v_{s+1} = 0_v$. Como $v_{s+1} \neq 0_V$, então $a_{s+1} = 0$. Assim, $\{v_1, \ldots, v_{s+1}\}$ é K-linearmente independente. ■

Exemplo 25

Seja $T : P_2(\mathbb{R}) \longrightarrow P_2(\mathbb{R})$ definido por

$$T(a + bt + ct^2) = (a + 2b + 3c) + (2b + 3c)t + 3ct^2.$$

Esse operador é diagonalizável. De fato, tomando $\alpha = \{1, t, t^2\}$ temos $A = T]_\alpha^\alpha = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{pmatrix}$, $\lambda I - A = \begin{pmatrix} \lambda - 1 & -2 & -3 \\ 0 & \lambda - 2 & -3 \\ 0 & 0 & \lambda - 3 \end{pmatrix}$ e o polinômio característico é $p(\lambda) = (\lambda - 1)(\lambda - 2)(\lambda - 3)$.

Como $\dim_{\mathbb{R}} P_2(\mathbb{R}) = 3$ e T tem três autovalores distintos, pela Proposição anterior, antes de determinarmos os autovetores, já sabemos que é possível construir uma base de $P_2(\mathbb{R})$ formada por autovetores de T. Por exemplo, escolhendo em $P_2(\mathbb{R})$, v_1 autovetor associado a $\lambda_1 = 1$, v_2 autovetor associado

Determine uma base β de $P_2(\mathbb{R})$ formada por autovetores de T e uma matriz $P = I]_{\alpha}^{\beta}$ que diagonaliza T .

a $\lambda_2 = 2$ e v_3 autovetor associado a $\lambda_3 = 3$ então, $\{v_1, v_2, v_3\}$ é $\mathbb{R}$-linearmente independente e é uma base de $P_2(\mathbb{R})$ formada por autovetores de T.

Quando $a \in K$ é autovalor de um operador K-linear T em V, então a é uma raiz do polinômio característico $p(\lambda) \in K[\lambda]$. Então, $\lambda - a$ divide $p(\lambda)$.

Definição 12 (Multiplicidade algébrica)

Seja V um K-espaço vetorial com $\dim_K V = n \geq 1$ e seja $a \in K$ um autovalor do operador K-linear T em V. Dizemos que o natural $m \geq 1$ é a *multiplicidade algébrica* do autovalor a se, e somente se, $(\lambda - a)^m$ divide $p(\lambda)$, o polinômio característico de T, e $(\lambda - a)^{m+1}$ não divide $p(\lambda)$.

Exemplo 26

No Exemplo anterior temos $p(\lambda) = (\lambda - 1)(\lambda - 2)(\lambda - 3)$ e a multiplicidade algébrica de cada autovalor é 1.

No Exemplo 17 temos $p(\lambda) = (\lambda - 1)(\lambda + 1)$ e a multiplicidade algébrica de cada autovalor é 1.

Determine o polinômio característico dos operadores $\mathbb{R}$-lineares dos Exemplos 17 e 18.

No Exemplo 18 temos $p(\lambda) = \lambda(\lambda - 1)^2$ e a multiplicidade algébrica do autovalor $\lambda = 0$ é 1 e do autovalor $\lambda = 1$ é 2.

No Exemplo 21 temos $p(\lambda) = (\lambda - 1)(\lambda - 2)^2$ e a multiplicidade algébrica do autovalor $\lambda = 1$ é 1 e do autovalor $\lambda = 2$ é 2.

No Exemplo 22 temos $p(\lambda) = (\lambda^2 + 9)(\lambda - 3)$ e a multiplicidade algébrica do único autovalor é 1. Nesse caso, o polinômio característico não se decompõe em produto de fatores lineares em $\mathbb{R}[\lambda]$.

Definição 13 (Multiplicidade geométrica)

Seja V um K-espaço vetorial com $\dim_K V = n \geq 1$ seja $a \in K$ um autovalor do operador K-linear $T : V \longrightarrow V$. Dizemos que o natural $m \geq 1$ é a *multiplicidade geométrica* do autovalor $a \in K$ se, e somente se, $m = \dim_K V_{\lambda=a}$.

Exemplo 27

No Exemplo 17 temos $p(\lambda) = (\lambda-1)(\lambda+1)$, $\dim_{\mathbb{R}} V_{\lambda=1} = 1$ e $\dim_{\mathbb{R}} V_{\lambda=-1} = 1$. Nesse caso, a multiplicidade geométrica de cada autovalor é igual à multiplicidade algébrica.

No Exemplo 18 temos $p(\lambda) = \lambda(\lambda - 1)^2$, $\dim_{\mathbb{R}} V_{\lambda=0} = 1$ e $\dim_{\mathbb{R}} V_{\lambda=1} = 2$. Nesse caso, a multiplicidade geométrica de cada autovalor é igual à multiplicidade algébrica de cada autovalor.

No Exemplo 21 temos $p(\lambda) = (\lambda-1)(\lambda-2)^2$, $\dim_{\mathbb{R}} V_{\lambda=1} = 1$ e $\dim_{\mathbb{R}} V_{\lambda=2} = 2$. Nesse caso, a multiplicidade geométrica de cada autovalor é igual à multiplicidade algébrica.

No Exemplo 22 temos $p(\lambda) = (\lambda^2 + 9)(\lambda - 3)$ e a multiplicidade geométrica do único autovalor coincide com a sua multiplicidade algébrica.

Proposição 4

Seja V um K-espaço vetorial com $\dim_K V = n \geq 1$ e seja $a \in K$ um autovalor do operador K-linear $T : V \longrightarrow V$ com multiplicidade algébrica m. Então, $\dim_K V_{\lambda=a} \leq m$.

Demonstração: Seja r a multiplicidade geométrica do autovalor a de T, isto é, $r = \dim_K V_{\lambda=a}$. Seja $\{v_1, \dots, v_r\}$ uma base do subespaço característico $V_{\lambda=a}$. Sejam $u_1, \dots, u_s$ vetores de V, tais que $\alpha = \{v_1, \dots, v_r, u_1, \dots, u_s\}$ seja uma base de V. Seja $A = T]^{\alpha}_{\alpha}$.

Então, $n = r + s$ e existem matrizes $B \in M_{r\times s}(K)$ e $C \in M_{s\times s}(K)$ tais que $A = \begin{pmatrix} aI_r & B \\ 0_{s\times r} & C \end{pmatrix}$, onde I_r é a matriz identidade de ordem r e $0_{s\times r}$ é a matriz nula s por r.

Podemos escrever a matriz identidade de ordem n como $\begin{pmatrix} I_r & 0_{r\times s} \\ 0_{s\times r} & I_s \end{pmatrix}$, onde $0_{r\times s}$, $0_{s\times r}$ são matrizes nulas e I_r e I_s são as matrizes identidades de ordens r e s, respectivamente.

O polinômio característico de T é

$$\begin{aligned} p(\lambda) &= \det(\lambda I - A) \\ &= \det \begin{pmatrix} \lambda I_r - aI_r & -B \\ 0_{s\times r} & \lambda I_s - C \end{pmatrix} \\ &= \det((\lambda - a)I_r)\det(\lambda I_s - C) \\ &= (\lambda - a)^r \det(\lambda I_s - C). \end{aligned}$$

Portanto, $(\lambda - a)^r$ divide $p(\lambda)$. Assim, $r \leq m$. ■

Vamos agora dar condições necessárias e suficientes para um operador K-linear em V ser diagonalizável.

Teorema 2

Seja V um K-espaço vetorial com $\dim_K V = n \geq 1$. O operador K-linear $T : V \longrightarrow V$ é diagonalizável se, e somente se,

(a) o polinômio característico de T tem todas as suas raízes em K;

(b) a multiplicidade algébrica de cada autovalor λ de T é igual a $\dim_K V_\lambda$.

A multiplicidade geométrica de cada autovalor coincide com a multiplicidade algébrica.

Demonstração:

($\Longrightarrow$:) Suponhamos que o operador K-linear T em V seja diagonalizável. Sejam $\lambda_1, \dots, \lambda_s$ em K os autovalores distintos de T e seja β uma base de V formada por autovetores de T, ordenada de modo que $\beta = \beta_1 \cup \ldots \cup \beta_s$, onde β_j é o subconjunto da base dos n_j autovetores de T associados ao autovalor λ_j, para cada $j = 1, \dots, s$. Então, é claro que $n = \dim_K V = \sharp\beta = n_1 + \cdots + n_s$.

Seja $B_j = \begin{pmatrix} \lambda_j & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_j \end{pmatrix} \in M_{n_j \times n_j}(K)$ matriz diagonal.

Então, $T]_\beta^\beta = \begin{pmatrix} B_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & B_s \end{pmatrix}$ é matriz diagonal em blocos de ordem n.

Escrevendo a matriz identidade de ordem n como uma matriz em blocos $I = \begin{pmatrix} I_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & I_s \end{pmatrix}$, onde I_j é a matriz identidade de ordem n_j, para cada $j = 1, \dots, s$, temos que $B_j = \lambda_j I_j$, para cada $j = 1, \dots, s$, e

$$\begin{aligned}
(\lambda I_V - T)]_\beta^\beta = (\lambda I_V)]_\beta^\beta - T]_\beta^\beta &= \begin{pmatrix} \lambda I_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda I_s \end{pmatrix} - \begin{pmatrix} B_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & B_s \end{pmatrix} \\
&= \begin{pmatrix} \lambda I_1 - \lambda_1 I_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda I_s - \lambda_s I_s \end{pmatrix} \\
&= \begin{pmatrix} (\lambda - \lambda_1) I_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & (\lambda - \lambda_s) I_s \end{pmatrix}
\end{aligned}$$

é matriz diagonal em blocos. Portanto,

$$\begin{aligned}
p(\lambda) &= \det\left((\lambda I - T)]_\beta^\beta\right) \\
&= \det((\lambda - \lambda_1)I_1) \cdot \ldots \cdot \det((\lambda - \lambda_s)I_s) \\
&= (\lambda - \lambda_1)^{n_1} \cdot \ldots \cdot (\lambda - \lambda_s)^{n_s},
\end{aligned}$$

mostrando o item (a) e que a multiplicidade algébrica do autovalor λ_j é n_j, para cada $j = 1, \dots, s$.

Consideremos W_j o subespaço de V de dimensão n_j gerado pelo subconjunto linearmente independente $\beta_j = \{v_{j1}, \dots, v_{jn_j}\}$. Para mostramos o item (b) basta mostrarmos que $W_j = V_{\lambda_j}$, para cada $j = 1, \dots, s$.

De fato, se $v \in W_j$, então $v = a_{j1}v_{j1} + \cdots + a_{jn_j}v_{jn_j}$, com $a_{j\ell} \in K$, para cada $\ell = 1, \dots, n_j$ e

$$\begin{aligned}
T(v) &= a_{j1}T(v_{j1}) + \cdots + a_{jn_j}T(v_{jn_j}) \\
&= a_{j1}\lambda_j v_{j1} + \cdots + a_{jn_j}\lambda_j v_{jn_j} \\
&= \lambda_j(a_{j1}v_{j1} + \cdots + a_{jn_j}v_{jn_j}) \\
&= \lambda_j v,
\end{aligned}$$

Logo, $v \in V_{\lambda_j}$, mostrando que $W_j \subset V_{\lambda_j}$, para cada $j = 1, \dots, s$.

Consideremos agora $v \in V_{\lambda_j}$. Então, $T(v) = \lambda_j v$. Escrevendo v como combinação linear da base $\beta = \beta_1 \cup \cdots \cup \beta_s$, temos que $v = \displaystyle\sum_{k=1}^{s} \left(\sum_{\ell=1}^{n_k} a_{k\ell} v_{k\ell} \right)$ e $T(v) = \displaystyle\sum_{k=1}^{s} \left(\sum_{\ell=1}^{n_k} a_{k\ell} T(v_{k\ell}) \right) = \sum_{k=1}^{s} \left(\sum_{\ell=1}^{n_k} a_{k\ell} \lambda_k v_{k\ell} \right)$.

Portanto, $\lambda_j v = \displaystyle\sum_{k=1}^{s} \left(\sum_{\ell=1}^{n_k} \lambda_j\, a_{k\ell} v_{k\ell} \right) = \sum_{k=1}^{s} \left(\sum_{\ell=1}^{n_k} \lambda_k a_{k\ell} v_{k\ell} \right)$.

Assim, $\lambda_j a_{k\ell} = \lambda_k a_{k\ell}$, para quaisquer $k = 1, \dots, s$ e $\ell = 1, \dots, n_k$.

Logo, $(\lambda_j - \lambda_k) a_{k\ell} = 0$, para quaisquer $k = 1, \dots, s$ e $\ell = 1, \dots, n_k$. Então, se $k \neq j$, temos $a_{k\ell} = 0$, para todo $\ell = 1, \dots, n_k$.

Portanto, $v = \displaystyle\sum_{\ell=1}^{n_j} a_{j\ell} v_{j\ell} \in W_j$, mostrando que $V_{\lambda_j} \subset W_j$. Concluímos então que $W_j = V_{\lambda_j}$, para todo $k = 1, \dots, s$.

($\Longleftarrow$:) Suponhamos que sejam válidas as propriedades (a) e (b). Então, existem $\lambda_1, \dots, \lambda_s \in K$ distintos tais que o polinômio característico de T é da forma

$$p(\lambda) = (\lambda - \lambda_1)^{n_1} \cdot \ldots \cdot (\lambda - \lambda_s)^{n_s},$$

onde $n_j = \dim_K V_{\lambda_j}$.

Consideremos o subespaço de V definido por $W_s = V_{\lambda_1} + \cdots + V_{\lambda_s}$.

Afirmamos que a soma é uma soma direta.

Sabendo que $W_s = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s}$, temos que

$$\dim_K W_s = \sum_{j=1}^{s} \dim_K V_{\lambda_j} = \sum_{j=1}^{s} n_j = \text{grau}(p(\lambda)) = \dim_K V.$$

Como $W_s \subset V$ concluímos que $W_s = V$. Tomando β_j uma base de V_{λ_j}, para cada $j = 1, \dots, s$, temos que $\beta = \beta_1 \cup \cdots \cup \beta_s$ é uma base de V formada por autovetores de T. Logo, T é diagonalizável.

A demonstração da afirmação é por indução sobre s.

Para $j \neq k$, temos que se $v \in V_{\lambda_j} \cap V_{\lambda_k}$, então $T(v) = \lambda_j v = \lambda_k v$, que é equivalente a $(\lambda_j - \lambda_k) v = 0_V$, com $\lambda_j - \lambda_k \neq 0$, logo $v = 0_V$. Assim, $V_{\lambda_j} \cap V_{\lambda_k} = \{0_V\}$.

Portanto, se $s = 2$, então $W_2 = V_{\lambda_1} \oplus V_{\lambda_2}$.

Seja $s \geq 2$ e suponhamos que $W_s = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s}$. Seja $v \in V_{\lambda_{s+1}} \cap W_s$. Então, existem $v_j \in V_{\lambda_j}$, para $j = 1, \dots, s$, unicamente determinados, tais que $v = v_1 + \cdots + v_s$. Logo,

$$T(v) = T(v_1) + \cdots + T(v_s) = \lambda_1 v_1 + \cdots + \lambda_s v_s. \qquad (1)$$

Por outro lado,

$$T(v) = \lambda_{s+1} v = \lambda_{s+1}(v_1 + \cdots + v_s) = \lambda_{s+1} v_1 + \cdots + \lambda_{s+1} v_s. \qquad (2)$$

Subtraindo (2) de (1), obtemos

$$(\lambda_1 - \lambda_{s+1})v_1 + \cdots + (\lambda_s - \lambda_{s+1})v_s = 0_V.$$

Como em W_s a soma é uma soma direta, temos que $(\lambda_j - \lambda_{s+1})v_j = 0_V$, para $j = 1, \ldots, s$, com $\lambda_j - \lambda_{s+1} \neq 0$. Portanto, $v_j = 0_V$, para $j = 1, \ldots, s$, e $v = 0_V$. Isto mostra que a soma $W_{s+1} = W_s + V_{\lambda_{s+1}}$ é uma soma direta. Logo, $W_{s+1} = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s} \oplus V_{\lambda_{s+1}}$ e o resultado é válido para $s+1$. ■

Exemplo 28

Seja $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definido por $T(x, y, z) = (4x + 2y, -x + y, y + 2z)$. Afirmamos que esse operador $\mathbb{R}$-linear não é diagonalizável.

De fato, tomando α a base canônica do $\mathbb{R}^3$ temos $A = T]^{\alpha}_{\alpha} = \begin{pmatrix} 4 & 2 & 0 \\ -1 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}$,

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

Assim, o autovalor $\lambda = 2$ tem multiplicidade algébrica 2. Vamos determinar o subespaço característico $V_{\lambda=2}$.

$\boxed{\lambda = 2}$

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

Logo, $V_{\lambda=2} = \{(x, y, z) \in \mathbb{R}^3 \,;\, x = 0, y = 0\} = \{(0, 0, z) \,;\, z \in \mathbb{R}\}$.

Nesse caso, $\dim_{\mathbb{R}} V_{\lambda=2} = 1 < 2 =$multiplicidade algébrica do autovalor $\lambda = 2$. Portanto, T não é diagonalizável.

Exemplo 29

Volte aos Exemplos 17, 18 e 21 e verifique a validade dos itens (a) e (b) do Teorema anterior.

Exemplo 30

O operador $\mathbb{C}$-linear $T : \mathbb{C}^2 \longrightarrow \mathbb{C}^2$ definido por $T(x, y) = (-y, x)$ é diagonalizável. De fato, $p(\lambda) = \lambda^2 + 1 = (\lambda - i)(\lambda + i)$. Nesse caso,

$1 \leq \dim_{\mathbb{C}} V_{\lambda=i} \leq 1$ =multiplicidade algébrica do autovalor $\lambda = i$, assim como,

$1 \leq \dim_{\mathbb{C}} V_{\lambda=-i} \leq 1$ =multiplicidade algébrica do autovalor $\lambda = -i$.

Portanto, $\dim_{\mathbb{C}} V_{\lambda=i} = 1$ e $\dim_{\mathbb{C}} V_{\lambda=-i} = 1$.

Podemos usar o conceito de diagonalização de operadores para determinar potências de números naturais de operadores ou matrizes.

Aplicação da diagonalização: Cálculo de potências

Seja $T : K^n \longrightarrow K^n$ um operador K-linear diagonalizável. Sejam α a base canônica do K^n e β uma base do K^n formada por autovetores de T.

Sejam $A = T]^{\alpha}_{\alpha}$ e a matriz diagonal $D = T]^{\beta}_{\beta}$. Tomando a matriz de mudança de base $P = I]^{\beta}_{\alpha}$ e $P^{-1} = I]^{\alpha}_{\beta}$ temos

$$D = T]^{\beta}_{\beta} = I]^{\alpha}_{\beta} T]^{\alpha}_{\alpha} I]^{\beta}_{\alpha} = P^{-1}AP.$$

Para cada natural $m \geq 1$, temos que

$$D^m = (P^{-1}AP)^m = P^{-1}A^mP.$$

Faça a demonstração da última igualdade por indução sobre m.

Portanto, $T^m]^{\alpha}_{\alpha} = (T]^{\alpha}_{\alpha})^m = A^m = PD^mP^{-1}$.

Dessa maneira, podemos determinar o operador T^m ou a matriz A^m, para todo natural $m \geq 1$. ■

Vejamos agora um exemplo.

Exemplo 31

Vamos determinar A^{10}, onde $A = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$.

Consideramos o operador $\mathbb{R}$-linear do $\mathbb{R}^2$, cuja matriz com respeito à base canônica do $\mathbb{R}^2$ é $T]^{\alpha}_{\alpha} = A$.

Então, $p(\lambda) = \det \begin{pmatrix} \lambda - 1 & 0 \\ 1 & \lambda - 2 \end{pmatrix} = (\lambda - 1)(\lambda - 2)$ é o polinômio característico de T. O operador T é diagonalizável, pois $2 = \dim_{\mathbb{R}} \mathbb{R}^2$ é igual ao número de autovalores distintos.

$\boxed{\lambda = 1}$

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

Logo, $V_{\lambda=1} = \{(x, y) \in \mathbb{R}^2 \,;\, x - y = 0\} = \{(x, x) \,;\, x \in \mathbb{R}\}$ e

$\lambda = 1 \leftrightarrow v = x(1, 1), x \neq 0$.

$\boxed{\lambda = 2}$

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

Logo, $V_{\lambda=2} = \{(x, y) \in \mathbb{R}^2 \,;\, x = 0\} = \{(0, y) \,;\, y \in \mathbb{R}\}$ e

$\lambda = 2 \leftrightarrow v = y(0, 1), y \neq 0$.

Portanto, $\beta = \{(1, 1), (0, 1)\}$ é uma base do $\mathbb{R}^2$ formada por autovetores do operador T.

Temos que $P = I]_{\alpha}^{\beta} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ e $P^{-1} = I]_{\beta}^{\alpha} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$.

Como $D = T]_{\beta}^{\beta} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = P^{-1}AP$, então $D^{10} = \begin{pmatrix} 1^{10} & 0 \\ 0 & 2^{10} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1024 \end{pmatrix} = P^{-1}A^{10}P$. Assim,

$$\begin{aligned} A^{10} = PD^{10}P^{-1} &= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1024 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 1 & 1024 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ -1023 & 1024 \end{pmatrix}. \end{aligned}$$

Podemos determinar agora o operador T^{10}. Temos que $A^{10} = T^{10}]_{\alpha}^{\alpha}$ e

$$T^{10}(x, y) = (x, -1023x + 1024y).$$

Exercícios

1. Seja $T : V \longrightarrow V$ um operador K-linear, onde $K = \mathbb{R}$ ou $K = \mathbb{C}$ e $\dim_K V = n \geq 1$. Diga quais das afirmações são falsas ou verdadeiras, justificando a sua resposta.

   (a) Se $T(v) = \lambda v$, para algum $v \in V$, então λ é autovetor de T.

   (b) T é operador linear invertível se, e somente se, zero não é autovalor de T.

   (c) Zero é autovalor de T se, e somente se, núcleo de T é não nulo.

   (d) $c \in K$ é autovalor de T se, e somente se, existe $v \in V$, $v \neq 0_V$, tal que $(T - cI)(v) = 0_V$.

   (e) Se $T(v) = \lambda v$, para algum escalar $\lambda \in K$, então v é autovetor de T.

   (f) Se v_1 e v_2 são autovetores de T linearmente independentes, então correspondem a autovalores distintos.

   (g) Se a dimensão de V é 2, então T pode ter 3 autovalores.

   (h) O número máximo de autovalores de T é a dimensão de V.

   (i) Todo operador linear T tem autovalores e autovetores.

2. Seja $V = \left\{ \begin{pmatrix} x & y \\ z & 0 \end{pmatrix} ; x, y, z \in \mathbb{C} \right\}$ e seja $T : V \longrightarrow V$ definida por $T \begin{pmatrix} x & y \\ z & 0 \end{pmatrix} = \begin{pmatrix} 2x - y + 2z & 3x \\ -2x + 4y + z & 0 \end{pmatrix}$. Mostre que T é diagonalizável. Construa uma base β de V formada por autovetores de T e dê $T]^{\beta}_{\beta}$.

3. Determine, caso exista, um autovalor do operador linear T e o subespaço característico, sem descrever T explicitamente.

Não encontre uma fórmula para T. Faça geometricamente.

   (a) $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ é a simetria com relação a uma reta pela origem. T é diagonalizável?
   (b) $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ é a projeção ortogonal sobre uma reta passando na origem. T é diagonalizável?
   (c) $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ é a rotação de $\theta \in [0, 2\pi)$. T é diagonalizável?
   (d) $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ é a rotação de θ em torno de uma reta pela origem. T é diagonalizável?
   (e) $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ é a projeção ortogonal sobre um plano passando pela origem. T é diagonalizável?
   (f) $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ é a projeção ortogonal sobre uma reta passando pela origem. T é diagonalizável?
   (g) $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ é a simetria com respeito a um plano passando pela origem. T é diagonalizável?
   (h) $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ é a simetria com respeito a um reta passando pela origem. T é diagonalizável?

4. Determine, caso existam, uma matriz invertível P e uma matriz diagonal D em $M_{n\times n}(\mathbb{R})$, tais que $D = P^{-1}AP$, para cada $A \in M_{n\times n}(\mathbb{R})$:

   (a) $A = \begin{pmatrix} 19 & -9 & -6 \\ 25 & -11 & -9 \\ 17 & -9 & -4 \end{pmatrix}$
   (b) $A = \begin{pmatrix} -1 & 4 & -2 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{pmatrix}$
   (c) $A = \begin{pmatrix} 5 & 0 & 0 \\ 1 & 5 & 0 \\ 0 & 1 & 5 \end{pmatrix}$
   (d) $A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 3 & 0 & 1 \end{pmatrix}$
   (e) $A = \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 1 & 3 \end{pmatrix}$
   (f) $A = \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 5 & -5 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$

5. Determine, caso existam, uma matriz invertível P e uma matriz diagonal D em $M_{n\times n}(\mathbb{C})$, tais que $D = P^{-1}AP$, para cada $A \in M_{n\times n}(\mathbb{C})$:

   (a) $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ (b) $A = \begin{pmatrix} 2 & -1 & 2 \\ 3 & 0 & 0 \\ -2 & 4 & 1 \end{pmatrix}$

6. Sejam $P_2(\mathbb{R}) = \{a + bt + ct^2 \;;\; a, b, c \in \mathbb{R}\}$ e $T : P_2(\mathbb{R}) \longrightarrow P_2(\mathbb{R})$ definida por
   $$T(a + bt + ct^2) = (5a + 6b + 2c) - (b + 8c)t + (a - 2c)t^2.$$

   (a) Determine o polinômio característico, os autovalores e os subespaços característicos de T.

   (b) T é diagonalizável?

7. Para cada $T : V \longrightarrow V$ $\mathbb{R}$-linear, determine uma base β de V, tal que $T]^{\beta}_{\beta}$ seja matriz diagonal D. Dê a matriz diagonal D.

   (a) $V = \mathbb{R}^2$ e $T(x, y) = (3x + 4y, 2x + y)$.

   (b) $V = P_1(\mathbb{R}) = \{a + bt \;;\; a, b \in \mathbb{R}\}$ e $T(a + bt) = a + (6a - b)t$.

   (c) $V = M_{2\times 2}(\mathbb{R})$ e $T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 2c & a + c \\ b - 2c & d \end{pmatrix}$.

8. Sejam $K = \mathbb{R}$ ou $K = \mathbb{C}$, $A, P \in M_{n\times n}(K)$ com P invertível.

   Mostre que $(P^{-1}AP)^m = P^{-1}A^mP$, para todo inteiro $m \geq 1$.

9. Usando o exercício anterior, calcule:

   (a) A^{25}, onde $A = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}$.

   (b) A^{2009}, onde $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

10. O traço de uma matriz $A = (a_{ij}) \in M_{n\times n}(K)$, onde $K = \mathbb{R}$ ou $K = \mathbb{C}$, é definido por
    $$\mathrm{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}.$$

    Mostre que se $n = 2$, então o polinômio característico de A é $p(\lambda) = \lambda^2 - \mathrm{tr}(A)\lambda + \det(A)$.

11. Seja $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$. Mostre que:

    (a) Se $(a-d)^2 + 4bc > 0$, então A é diagonalizável.

    (b) Se $(a-d)^2 + 4bc < 0$, então A não é diagonalizável.

Teorema de Hamilton-Cayley e polinômio mínimo

Relembramos algumas propriedades relevantes da álgebra dos operadores lineares antes de apresentar o Teorema de Hamilton-Cayley.

Seja V um K-espaço vetorial com $\dim_K V = n \geq 1$. Lembramos que $\mathcal{L}(V, V) = \{T : V \longrightarrow V\,;\, T \text{ é } K\text{-linear }\}$ é um K-espaço vetorial com $\dim_K \mathcal{L}(V, V) = n^2$. De fato, fixada α uma base de V a função

$$\begin{array}{rcl} \varphi_\alpha : \mathcal{L}(V, V) & \longrightarrow & M_{n\times n}(K) \\ T & \longmapsto & T]^\alpha_\alpha \end{array}$$

é um isomorfismo de K-espaços vetoriais.

Logo, $\dim_K \mathcal{L}(V, V) = \dim_K M_{n\times n}(K) = n^2$.

Também, $T = 0$ se, e somente se, $T]^\alpha_\alpha = 0 \in M_{n\times n}(K)$.

Seja V um K-espaço vetorial com $\dim_K V = n \geq 1$ e seja $T : V \longrightarrow V$ um operador K-linear.

Temos $T^1 = T$ e, para cada número natural $m > 1$, definimos o operador $T^m = T^{m-1} \circ T$. Então, $T^m : V \longrightarrow V$ é um operador K-linear.

Para cada polinômio $f(x) = b_0 + b_1x + \cdots + b_mx^m$ com coeficientes em K definimos o operador K-linear $f(T) = b_0I_V + b_1T + \cdots + b_mT^m$.

Fixemos α uma base de V. Seja $A = T]^\alpha_\alpha \in M_{n\times n}(K)$.

A matriz da composição de operadores lineares é o produto das matrizes.

Para cada número natural $\ell \geq 1$, temos que $A^\ell = (T]^\alpha_\alpha)^\ell = T^\ell]^\alpha_\alpha$ e assim,

$$\begin{array}{rcl} f(T)]^\alpha_\alpha & = & (b_0I_V + b_1T + \cdots + b_mT^m)]^\alpha_\alpha \\ & = & b_0(I_V]^\alpha_\alpha) + b_1(T]^\alpha_\alpha) + \cdots + b_m(T^m]^\alpha_\alpha) \\ & = & b_0I + b_1A + \cdots + b_mA^m \\ & = & f(A). \end{array}$$

Observamos que fixado um operador K-linear T em V, tal que $\dim_K V = n \geq 1$, então os $n^2 + 1$ operadores lineares $I_V, T, \ldots, T^{n^2}$ de $\mathcal{L}(V, V)$ são linearmente dependentes sobre K. Portanto, existem $a_0, \ldots, a_{n^2}$ em K, nem todos nulos, tais que

$$a_0I_V + a_1T + \cdots + a_{n^2}T^{n^2} = 0,$$

onde 0 é o operador identicamente nulo de $\mathcal{L}(V, V)$.

$\mathcal{L}(V, V)$ é isomorfo a $M_{n\times n}(K)$ e em qualquer espaço vetorial de dimensão finita, todo subconjunto com mais elementos do que a dimensão do espaço vetorial é linearmente dependente.

Logo, existe um polinômio não nulo $g(x) = a_0 + a_1x + \cdots + a_{n^2}x^{n^2}$ com coeficientes em K, tal que $g(T) = 0$.

O Teorema de Hamilton-Cayley afirma que é possível construir um polinômio de grau igual à dimensão de V, a saber, $p(x)$, o polinômio característico de T, tal que $p(T) = 0$.

Exemplo 32

Consideremos o operador linear T em $\mathbb{R}^2$ definido por $T(x, y) = (x + y, y)$.

Tomando $\alpha = \{(1, 0), (0, 1)\}$, temos que $A = T]_\alpha^\alpha = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

O polinômio característico de T é

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

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

Isto é equivalente a $p(T) = (T - I_V)^2 = T^2 - 2T + I_V = 0$.

Como consequência temos que $I_V = -T^2 + 2T = T \circ (-T + 2I_V)$. Portanto, T é um operador invertível e o seu inverso é o operador $T^{-1} = -T + 2I_V$. Logo,

$$T^{-1}(x, y) = -T(x, y) + 2I_V(x, y) = (-x - y, -y) + (2x, 2y) = (x - y, y).$$

Para a compreensão da demonstração do Teorema de Hamilton-Cayley a seguinte observação é muito importante.

Observação: Toda matriz quadrada de ordem n cujos coeficientes são polinômios em $K[\lambda]$ de grau no máximo m ou o polinômio nulo pode ser escrita como

$$A = A_0 + A_1\lambda + \cdots + A_m\lambda^m,$$

onde $A_0, A_1, \ldots, A_m \in M_{n\times n}(K)$.

Vejamos o procedimento com um exemplo.

Exemplo 33

Consideremos $A = \begin{pmatrix} 4 + 3\lambda^2 - \lambda^3 & 2 + \lambda - \lambda^2 + \lambda^3 \\ 2 + 3\lambda + 2\lambda^2 & \sqrt{3} + 2\lambda^2 + 5\lambda^3 \end{pmatrix} \in M_{2\times 2}(\mathbb{R}[\lambda])$.

Temos

$$\begin{aligned} A &= \begin{pmatrix} 4 & 2 \\ 2 & \sqrt{3} \end{pmatrix} + \begin{pmatrix} 0\lambda & \lambda \\ 3\lambda & 0\lambda \end{pmatrix} + \begin{pmatrix} 3\lambda^2 & -\lambda^2 \\ 2\lambda^2 & 2\lambda^2 \end{pmatrix} + \begin{pmatrix} -\lambda^3 & \lambda^3 \\ 0\lambda^3 & 5\lambda^3 \end{pmatrix} \\ &= \begin{pmatrix} 4 & 2 \\ 2 & \sqrt{3} \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 3 & 0 \end{pmatrix}\lambda + \begin{pmatrix} 3 & -1 \\ 2 & 2 \end{pmatrix}\lambda^2 + \begin{pmatrix} -1 & 1 \\ 0 & 5 \end{pmatrix}\lambda^3 \end{aligned}$$

Teorema 3 (Hamilton-Cayley)

Sejam V um K-espaço vetorial com $\dim_K V = n \geq 1$, T um operador linear em V e $p(\lambda)$ o polinômio característico de T. Então, $p(T) = 0$.

Demonstração: Seja α uma base qualquer de V e seja $A = T]^{\alpha}_{\alpha}$. Mostraremos que $p(T)]^{\alpha}_{\alpha} = p(A) = 0 \in M_{n\times n}(K)$, concluindo que $p(T) = 0$.

Seja $p(\lambda) = \det(\lambda I - A)$. Seja $B = \text{adj}(\lambda I - A)$. Temos que B é uma matriz n por n cujos coeficientes são polinômios na indeterminada λ com coeficientes em K e de grau no máximo $n-1$.

Escrevemos $B = B_0 + B_1\lambda + \cdots + B_{n-1}\lambda^{n-1}$, com $B_j \in M_{n\times n}(K)$, (I)

Da propriedade da adjunta clássica, segue que:

$$B(\lambda I - A) = \det(\lambda I - A)I = p(\lambda)I \qquad \text{(II)}$$

Escrevemos

$$p(\lambda) = a_0 + a_1\lambda + \cdots + a_{n-1}\lambda^{n-1} + \lambda^n, \qquad \text{(III)}$$

com $a_j \in K$, para $j = 1, \ldots, n-1$.

Substituindo (I) e (III) em (II), temos:

$$(B_0 + B_1\lambda + \cdots + B_{n-1}\lambda^{n-1})(\lambda I - A) = (a_0 + a_1\lambda + \cdots + a_{n-1}\lambda^{n-1} + \lambda^n)I.$$

Esta é uma igualdade de matrizes polinomiais. Comparando as matrizes dos coeficientes dos termos de mesmo grau, temos que:

$$\begin{array}{crcll}
(0) & -B_0A & = & a_0I & \text{(coeficiente de } \lambda^0) \\
(1) & B_0 - B_1A & = & a_1I & \text{(coeficiente de } \lambda) \\
(2) & B_1 - B_2A & = & a_2I & \text{(coeficiente de } \lambda^2) \\
\vdots & & \vdots & & \\
(j) & B_{j-1} - B_jA & = & a_jI & \text{(coeficiente de } \lambda^j) \\
\vdots & & \vdots & & \\
(n-1) & B_{n-2} - B_{n-1}A & = & a_{n-1}I & \text{(coeficiente de } \lambda^{n-1}) \\
(n) & B_{n-1} & = & I & \text{(coeficiente de } \lambda^n)
\end{array}$$

Para cada $j = 1, \ldots, n$ multiplicamos a equação (j) por A^j, obtendo:

$$\begin{array}{crcl}
(0) & -B_0A & = & a_0I \\
(1)' & B_0A - B_1A^2 & = & a_1A \\
(2)' & B_1A^2 - B_2A^3 & = & a_2A^2 \\
\vdots & & \vdots & \\
(j)' & B_{j-1}A^j - B_jA^{j+1} & = & a_jA^j \\
\vdots & & \vdots & \\
(n-1)' & B_{n-2}A^{n-1} - B_{n-1}A^n & = & a_{n-1}A^{n-1} \\
(n)' & B_{n-1}A^n & = & A^n.
\end{array}$$

Somando membro a membro, obtemos:

$$0 = a_0I + a_1A + a_2A^2 + \cdots + a_{n-1}A^{n-1} + A^n = p(A). \blacksquare$$

Quem são os polinômios $f(x) \in K[x]$ tais que $f(T) = 0$, onde T é um operador linear em um K-espaço vetorial de dimensão $n \geq 1$? A Proposição a seguir responde essa questão.

Proposição 5

Seja V um K-espaço vetorial de dimensão $n \geq 1$ e seja $T : V \longrightarrow V$ um operador linear. Seja $I = \{f(x) \in K[x] \ ; \ f(T) = 0\}$. Então, existe um único polinômio mônico $m(x) \in K[x] \backslash K$ tal que $I = \{g(x)m(x) \ ; \ g(x) \in K[x]\}$. Em particular, $m(x)$ divide $f(x)$ e $m(x)$ é o polinômio mônico de menor grau que se anula em T.

Demonstração: Vamos mostrar, primeiramente, que existe um polinômio $m(x) \in K[x] \backslash K$ tal que $I = \{g(x)m(x) \ ; \ g(x) \in K[x]\}$.

Observamos que $I \neq \{0\}$ pois, pelo Teorema de Hamilton-Cayley, o polinômio característico de T $p(x) \in I$.

Consideremos $S = \{\text{grau}(f(x)) \ ; \ f(x) \in I \text{ e } f(x) \neq 0\}$. Temos que $S \neq \emptyset$ e $S \subset \mathbb{N}$. Pelo princípio da boa ordenação, S tem menor elemento, digamos s. Então, existe polinômio $m(x) \in K[x]$, $m(x) \neq 0$, com $\text{grau}(m(x)) = s$ e $m(T) = 0$.

Afirmamos que $I = \{ \ g(x)m(x) \ ; \ g(x) \in K[x] \ \}$.

De fato, sejam $J = \{ \ g(x)m(x) \ ; \ g(x) \in K[x] \ \}$, α uma base de V e $A = T]^{\alpha}_{\alpha}$.

Seja $f(x) \in J$. Então existe $g(x) \in K[x]$ tal que $f(x) = g(x)m(x)$. Como $m(x) \in I$, então $m(A) = 0$ e $f(A) = g(A)m(A) = g(A) \cdot 0 = 0$, que é equivalente, a $f(T) = g(T)m(T) = 0$, logo $f(x) = g(x)m(x) \in I$ e $J \subset I$.

Consideremos agora $f(x) \in I$. Pela divisão euclidiana de $f(x)$ por $m(x)$, existem polinômios unicamente determinados $q(x)$ e $r(x)$ tais que

$$f(x) = q(x)m(x) + r(x),$$

onde $r(x) = 0$ ou $0 \leq \text{grau}(r(x)) < \text{grau}(m(x)) = s$.

Assim,

$$r(x) = f(x) - q(x)m(x) \text{ e } r(A) = f(A) - q(A)m(A) = 0 - q(A) \cdot 0 = 0.$$

Portanto, $r(T) = 0$. O caso $r(x) \neq 0$ não pode ocorrer, em virtude de contradizer a escolha de $m(x)$, logo concluímos que $r(x) = 0$. Portanto, $f(x) = q(x)m(x) \in J$, mostrando que $I \subset J$.

Seja $a_s \neq 0$ o coeficiente líder de $m(x)$. Escrevemos

$$\begin{aligned} m(x) &= a_s x^s + a_{s-1}x^{s-1} + \cdots + a_1 x + a_0 \\ &= a_s \underbrace{\left(x^s + a_s^{-1}a_{s-1}x^{s-1} + \cdots + a_s^{-1}a_1 x + a_s^{-1}a_0\right)}_{m_1(x) \text{ é mônico}} \\ &= a_s m_1(x), a_s \neq 0 \text{ e } m_1(x) \text{ mônico.} \end{aligned}$$

Como $f(A) = 0$ se, e somente se, $af(A) = 0$, para todo $a \in K$, $a \neq 0$. Então, $m_1(A) = 0$, com grau$(m_1(x)) = s$ e $m_1(x)$ mônico. Logo, existe um único polinômio mônico de menor grau que se anula em T. ∎

Definição 14 (Polinômio mínimo de um operador)
Seja V um K-espaço vetorial de dimensão $n \geq 1$ e seja $T : V \longrightarrow V$ um operador linear. O *polinômio mínimo* de T é $m(x)$, o polinômio mônico de menor grau no conjunto

$$I = \{\ f(x) \in K[x]\ ;\ f(T) = 0\ \} = \{\ g(x)m(x)\ ;\ g(x) \in K[x]\ \}.$$

Em particular,

$$f(x) \in K[x] \text{ e } f(T) = 0 \text{ se, e somente se, } m(x) \text{ divide } f(x).$$

Corolário 1
Sejam V um K-espaço vetorial de dimensão $n \geq 1$, $T : V \longrightarrow V$ um operador linear, $p(x)$ e $m(x)$, respectivamente, os polinômios característico e mínimo de T. Então, $m(x)$ divide $p(x)$.

Demonstração: Como $p(T) = 0$, pela Proposição anterior, temos que existe $g(x) \in K[x]$ tal que $p(x) = g(x)m(x)$. Logo, $m(x)$ divide $p(x)$. ∎

Podemos dar uma caracterização dos operadores diagonalizáveis em termos do polinômio mínimo. Para isto, precisamos do seguinte resultado.

Lema 1
Sejam V um K-espaço vetorial de dimensão $n \geq 1$, $T : V \longrightarrow V$ um operador linear, $p(x)$ e $m(x)$, respectivamente, os polinômios característico e mínimo de T. Então, $p(x)$ divide $m(x)^n$.

Demonstração: Seja α uma base de V e $A = T]^{\alpha}_{\alpha}$.

Suponhamos que $m(x) = t^s + c_1x^{s-1} + \cdots + c_{s-1}x + c_s$, com $c_j \in K$, para $j = 1, \ldots, s$. Consideremos as matrizes

$$\begin{aligned}
B_0 &= I \\
B_1 &= A + c_1 I \\
B_2 &= A^2 + c_1 A + c_2 I \\
&\vdots \\
B_{s-2} &= A^{s-2} + c_1 A^{s-3} + \cdots + c_{s-3} A + c_{s-2} I \\
B_{s-1} &= A^{s-1} + c_1 A^{s-2} + \cdots + c_{s-2} A + c_{s-1} I.
\end{aligned}$$

Então,

$$\begin{aligned}
B_0 &= I \\
B_1 - AB_0 &= c_1 I \\
B_2 - AB_1 &= c_2 I \\
&\vdots \\
B_{s-1} - AB_{s-2} &= c_{s-1} I
\end{aligned}$$

Além disso,

$$\begin{aligned}
-AB_{s-1} &= -(A^s + c_1A^{s-1} + \cdots + c_{s-1}A) \\
&= -(A^s + c_1A^{s-1} + \cdots + c_{s-1}A + c_sI) + c_sI \\
&= -m(A) + c_sI \\
&= c_sI.
\end{aligned}$$

Definimos a matriz polinomial

$$B(x) = B_0x^{s-1} + B_1x^{s-2} + \cdots + B_{s-2}x + B_{s-1}.$$

Então,

$$\begin{aligned}
(xI - A)B(x) &= (xI - A)(B_0x^{s-1} + B_1x^{s-2} + \cdots + B_{s-2}x + B_{s-1}) \\
&= (B_0x^s + B_1x^{s-1} + \cdots + B_{s-1}x) + \\
&\quad -(AB_0x^{s-1} + AB_1x^{s-2} + \cdots + AB_{s-2}x + AB_{s-1}) \\
&= B_0x^s + (B_1 - AB_0)x^{s-1} + \cdots + (B_{s-1} - AB_{s-2})x + \\
&\quad -AB_{s-1} \\
&= Ix^s + c_1Ix^{s-1} + c_2Ix^{s-2} + \cdots + c_{s-1}Ix + c_sI \\
&= (x^s + c_1x^{s-1} + c_2x^{s-2} + \cdots + c_{s-1}x + c_s)I \\
&= m(x)I
\end{aligned}$$

Calculando o determinante em ambos os lados da igualdade acima, obtemos:

$$\det(xI - A)\det(B(x)) = \det(m(x)I) = m(x)^n.$$

Como $\det(B(x))$ é um polinômio com coeficientes em K e o polinômio característico de T é $p(x) = \det(xI - A)$, então $p(x) \cdot \det(B(x)) = m(x)^n$, que é equivalente a $p(x)$ divide $m(x)^n$. ■

Proposição 6

Sejam V um K-espaço vetorial de dimensão $n \geq 1$, $T : V \longrightarrow V$ um operador linear, $p(x)$ e $m(x)$, respectivamente, os polinômios característico e mínimo de T. Então, $m(x)$ e $p(x)$ têm os mesmos fatores irredutíveis em $K[x]$.

Demonstração: Seja $q(x)$ um fator irredutível de $m(x)$. Como $q(x)$ divide $m(x)$ e $m(x)$ divide $p(x)$, então existem $f(x)$ e $g(x)$ em $K[x]$ tais que

$$m(x) = q(x)f(x) \text{ e } p(x) = m(x)g(x),$$

logo $p(x) = (q(x)f(x))g(x)$. Assim, $q(x)$ divide $p(x)$.

Reciprocamente, suponhamos que $q(x)$ seja um fator irredutível de

$p(x)$, então $q(x)$ divide $p(x)$. Pelo Lema anterior, $p(x)$ divide $m(x)^n$. Logo, existem $f(x)$ e $g(x)$ em $K[x]$ tais que

$$p(x) = f(x)q(x) \text{ e } m(x)^n = g(x)p(x),$$

logo $m(x)^n = g(x)(f(x)q(x))$. Assim, $q(x)$ divide $m(x)^n$. Como $q(x)$ é irredutível, então $q(x)$ divide $m(x)$. ■

Lembramos que se $q(x)$ é irredutível em $K[x]$ e divide um produto de fatores em $K[x]$, então $q(x)$ divide um dos fatores.

Exemplo 34

Vamos determinar os polinômios característicos e mínimos das seguintes matrizes: $A = \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $C = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix}$.

Como $\lambda I - A = \begin{pmatrix} \lambda-2 & -3 \\ 0 & \lambda-2 \end{pmatrix}$, então $p(\lambda) = (\lambda-2)^2$. Pela Proposição 6 e pelo Corolário 1, as possibilidades para o polinômio mínimo são: $f(\lambda) = \lambda-2$ ou $g(\lambda) = (\lambda-2)^2$. Em virtude de $f(A) = A-2I \neq 0$, temos $m(\lambda) = g(\lambda) = (\lambda-2)^2$.

Como $\lambda I - B = \begin{pmatrix} \lambda-1 & -1 & 0 \\ 0 & \lambda-2 & 0 \\ 0 & 0 & \lambda-1 \end{pmatrix}$, então $p(\lambda) = (\lambda-1)^2(\lambda-2)$. Pela Proposição 6 e pelo Corolário 1, as possibilidades para o polinômio mínimo são: $f(t) = (\lambda-1)(\lambda-2)$ ou $g(T) = (\lambda-1)^2(\lambda-2)$. Temos $f(B) = (B-I)(B-2I) = 0$, logo $m(\lambda) = f(\lambda) = (\lambda-1)(\lambda-2)$.

Determine $B-I$ e $B-2I$ e calcule $(B-I)(B-2I)$.

Como $\lambda I - C = \begin{pmatrix} \lambda-2 & 0 & 0 \\ 0 & \lambda-2 & -2 \\ 0 & 0 & \lambda-1 \end{pmatrix}$, então $p(\lambda) = (\lambda-1)(\lambda-2)^2$. Pela Proposição 6 e pelo Corolário 1, as possibilidades para o polinômio mínimo são: $f(\lambda) = (\lambda-1)(\lambda-2)$ ou $g(\lambda) = (\lambda-1)(\lambda-2)^2$. Verificamos que $f(C) = (C-I)(C-2I) = 0$, logo $m(\lambda) = f(\lambda) = (\lambda-1)(\lambda-2)$.

Determine $C-I$ e $C-2I$ e calcule $(C-I)(C-2I)$.

Vamos dar condições necessárias e suficientes para um operador linear em um espaço vetorial de dimensão finita ser diagonalizável em termos do polinômio mínimo. Para isto, precisamos do seguinte resultado.

Lema 2

Seja V um K-espaço vetorial de dimensão $n \geq 1$ e seja $T : V \longrightarrow V$ um operador linear. Se $f(x), g(x), h(x) \in K[x]$ são polinômios mônicos tais que $f(x) = g(x)h(x)$, $\text{mdc}_{K[x]}(g(x), h(x)) = 1$ e $f(T) = 0$, então

(a) $V = U \oplus W$, onde $U = \text{Núcleo}(g(T))$ e $W = \text{Núcleo}(h(T))$ e os subespaços U e W são invariantes por T.

(b) Se $f(x)$ é o polinômio mínimo de T, então $g(x)$ é o polinômio mínimo de $T_1 = T|_U$, a restrição de T a U, e $h(x)$ é o polinômio mínimo de $T_2 = T|_W$, a restrição de T a W.

Vale a recíproca do item (b).

Demonstração:

(a) Como $\text{mdc}_{K[x]}(g(x), h(x)) = 1$, existem $a(x), b(x) \in K[x]$ tais que

$$1 = a(x)h(x) + b(x)g(x).$$

Então, $I_V = a(T)h(T) + b(T)g(T)$. Aplicando em $v \in V$ temos

$$v = I_V(v) = \big(a(T)h(T)\big)(v) + \big(b(T)g(T)\big)(v). \qquad (\star)$$

Sejam $u = \big(a(T)h(T)\big)(v)$ e $w = \big(b(T)g(T)\big)(v)$.

Afirmamos que $u \in U = \text{Núcleo}(g(T))$ e $w \in W = \text{Núcleo}(h(T))$.

De fato,

$$\begin{array}{rcl} g(T)(u) &=& g(T)\big(a(T)h(T)(v)\big) \\ &=& \big(g(T)a(T)h(T)\big)(v) \\ &=& \big(a(T)g(T)h(T)\big)(v) \\ &=& a(T)\big(g(T)h(T)(v)\big) \\ &=& a(T)\big(f(T)(v)\big) \\ &=& a(T)(0_V) \\ &=& 0_V \end{array} \quad \text{e} \quad \begin{array}{rcl} h(T)(w) &=& h(T)\big(b(T)g(T)(v)\big) \\ &=& \big(h(T)b(T)g(T)\big)(v) \\ &=& \big(b(T)g(T)h(T)\big)(v) \\ &=& b(T)\big(g(T)h(T)(v)\big) \\ &=& b(T)\big(f(T)(v)\big) \\ &=& b(T)(0_V) \\ &=& 0_V. \end{array}$$

Usamos a definição de composição de funções; que operadores lineares polinomiais em T comutam; que $f(T) = g(T)h(T)$ e $f(T) = 0$, isto é, $f(T)(v) = 0_V$, para todo $v \in V$; e que $a(T)$ e $b(T)$ são lineares.

Portanto, $v = u + w$, onde $u \in U$ e $w \in W$. Logo, $V = U + W$.

Para que a soma seja uma soma direta, falta apenas mostrarmos que $U \cap W = \{0_V\}$.

Seja $v \in \text{Núcleo}(g(T)) \cap \text{Núcleo}(h(T))$. De $(\star)$, temos

$$\begin{array}{rcl} v &=& \big(a(T)h(T)\big)(v) + \big(b(T)g(T)\big)(v) \\ &=& a(T)\big(h(T)(v)\big) + b(T)\big(g(T)(v)\big) \\ &=& a(T)(0_V) + b(T)(0_V) \\ &=& 0_V + 0_V \\ &=& 0_V. \end{array}$$

Vamos mostrar que $U = \text{Núcleo}(g(T))$ e $W = \text{Núcleo}(h(T))$ são invariantes por T.

Seja $u \in U = \text{Núcleo}(g(T))$. Temos $g(T)(u) = 0_V$ e

$$g(T)\big(T(u)\big) = \big(g(T) \circ T\big)(u) = \big(T \circ g(T)\big)(u) = T\big(g(T)(u)\big) = T(0_V) = 0_V.$$

Portanto, $T(u) \in U = \text{Núcleo}(g(T))$ e U é invariante por T.

Seja $w \in W = \text{Núcleo}(h(T))$. Temos $h(T)(w) = 0_V$ e

$$h(T)\big(T(w)\big) = \big(h(T) \circ T\big)(w) = \big(T \circ h(T)\big)(w) = T\big(h(T)(w)\big) = T(0_V) = 0_V.$$

Portanto, $T(w) \in W = \text{Núcleo}(h(T))$ e W é invariante por T.

(b) Sabemos que $V = U \oplus W$. Como U e W são invariantes por T, então $T_1 = T|_U$ e $T_2 = T|_W$ definem operadores lineares em U e W, a saber,

$$\begin{array}{rcl} T_1 : U & \longrightarrow & U \\ u & \longmapsto & T(u) \end{array} \quad \text{e} \quad \begin{array}{rcl} T_2 : W & \longrightarrow & W \\ w & \longmapsto & T(w). \end{array}$$

Sejam β e γ bases, respectivamente, de U e W. Então, $\alpha = \beta \cup \gamma$ é uma base de V. Sejam $A = T]^{\alpha}_{\alpha}$, $B = T_1]^{\beta}_{\beta}$ e $C = T_2]^{\gamma}_{\gamma}$. Então, $A = \begin{pmatrix} B & 0_{r\times s} \\ 0_{s\times r} & C \end{pmatrix}$, onde B e C são matrizes quadradas de ordens, respectivamente, $r = \dim_K U$ e $s = \dim_K W$, com $r + s = \dim_K V = n$, e as matrizes $0_{r\times s}$, $0_{s\times r}$ são nulas.

Observamos, primeiramente, que como A é uma matriz diagonal em blocos, então $A^j = \begin{pmatrix} B^j & 0 \\ 0 & C^j \end{pmatrix}$, para todo $j \geq 1$, também é uma matriz diagonal em blocos; assim como $a_j A^j$, $a_j \in K$, e $\sum_{j=0}^{m} a_j A^j = \ell(A)$, onde $\ell(x) = \sum_{j=0}^{m} a_j x^j$, com blocos de ordens r e s, as mesmas ordens dos blocos de A.

Logo, para qualquer polinômio $\ell(x) \in K[x]$ temos $\ell(A) = \begin{pmatrix} \ell(B) & 0 \\ 0 & \ell(C) \end{pmatrix}$.

Verifique essa propriedade. Faça o Exercício 10.

Suponhamos que $f(x)$ seja o polinômio mínimo de T e sejam $m_1(x)$ e $m_2(x)$ os polinômios mínimos de T_1 e T_2, respectivamente.

Como $m_1(B) = 0$ e $m_2(C) = 0$, então

$$\begin{aligned} m_1(A)m_2(A) &= \begin{pmatrix} m_1(B) & 0 \\ 0 & m_1(C) \end{pmatrix} \begin{pmatrix} m_2(B) & 0 \\ 0 & m_2(C) \end{pmatrix} \\ &= \begin{pmatrix} m_1(B)m_2(B) & 0 \\ 0 & m_1(C)m_2(C) \end{pmatrix} \\ &= \begin{pmatrix} 0 \cdot m_2(B) & 0 \\ 0 & m_1(C)\cdot 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

é a matriz nula de ordem n. Logo, $f(x)$, o polinômio mínimo de T, divide $m_1(x)m_2(x)$.

Observamos que como o subespaço U é invariante por T e $T_1 = T|_U$, então $(aT^j)|_U = aT_1^j$, para todo $j \geq 1$ e para todo $a \in K$. Assim, dado $\ell(x) = \sum_{j=0}^{m} a_j x^j$ temos que $\ell(T)|_U = \ell(T_1)$. Analogamente, W é invariante por T, $T_2 = T|_W$ e $(aT^j)|_W = aT_2^j$, para todo $j \geq 1$ e para todo $a \in K$, logo $\ell(T)|_W = \ell(T_2)$.

Veja o Exercício 11.

Como U = Núcleo($g(T)$) e W = Núcleo($h(T)$), então obtemos que $0 = g(T)|_U = g(T_1)$ e $0 = h(T)|_W = h(T_2)$. Assim, $g(B) = 0$ e $h(C) = 0$, portanto $m_1(x)$ divide $g(x)$ e $m_2(x)$ divide $h(x)$. Logo, $m_1(x)m_2(x)$ divide $g(x)h(x) = f(x)$.

Se $q(x) \in K[x]$ é irredutível e divide $m_2(x)$, então $q(x)$ divide $h(x)$ e $q(x)$ não divide $g(x)$, pois $\text{mdc}_{K[x]}(g(x), h(x)) = 1$.

Portanto, $f(x) = am_1(x)m_2(x)$, para algum $a \in K$. Como $f(x), m_1(x)$ e $m_2(x)$ são mônicos, então $a = 1$. Assim, $g(x)h(x) = m_1(x)m_2(x)$.

Como $g(x)$ divide $m_1(x)m_2(x)$ e $\text{mdc}_{K[x]}(g(x), m_2(x)) = 1$, então $g(x)$ divide $m_1(x)$. Assim, $g(x) = am_1(x)$, para algum $a \in K$. Sendo ambos os polinômios mônicos, concluímos que $a = 1$ e $g(x) = m_1(x)$.

Portanto, $m_1(x)h(x) = m_1(x)m_2(x)$. Cancelando $m_1(x)$, obtemos $h(x) = m_2(x)$. ■

Agora estamos prontos para a caracterização de operadores diagonalizáveis pelo polinômio mínimo.

Proposição 7 (Operador diagonalizável e o polinômio mínimo)
Sejam V um K-espaço vetorial com $\dim_K(V) = n \geq 1$ e $T : V \longrightarrow V$ um operador K-linear. O operador T é diagonalizável se, e somente se, o polinômio mínimo de T é $m(x) = (x - \lambda_1) \cdot \ldots \cdot (x - \lambda_s)$, onde $\lambda_1, \ldots, \lambda_s \in K$ são os seus autovalores distintos.

Demonstração: Suponhamos que T seja diagonalizável. Pelo Teorema 2, o polinômio característico de T é da forma $p(x) = (x - \lambda_1)^{n_1} \cdot \ldots \cdot (x - \lambda_s)^{n_s}$, onde $\lambda_1, \ldots, \lambda_s$ são os autovalores distintos de T, $n = \dim_K V = n_1 + \cdots + n_s$, $n_j = \dim_K V_{\lambda_j}$, para $j = 1, \ldots, s$, e $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s}$.

Pela Proposição 6, o polinômio mínimo e o polinômio característico têm os mesmos fatores irredutíveis em $K[x]$. Logo, $f(x) = (x - \lambda_1) \cdot \ldots \cdot (x - \lambda_s)$ divide $m(x)$.

Afirmamos que $f(T) = 0$.

De fato, dado $v \in V$, existe $v_k \in V_{\lambda_k}$ unicamente determinado, para todo $k = 1, \ldots, s$, tal que $v = \sum_{k=1}^{s} v_k$. Temos que

Em (1) usamos a definição do operador $f(T)$; em (2), que operadores polinomiais em T comutam; em (3), a definição de composição de funções; em (4), que $v_k \in V_{\lambda_k}$; e em (5), que a composição de operadores lineares é linear.

$$\begin{aligned} f(T)(v_k) &\overset{(1)}{=} \left(\prod_{j=1}^{s} (T - \lambda_j I_V) \right)(v_k) \\ &\overset{(2)}{=} \left(\left(\prod_{1 \leq j \neq k \leq s} (T - \lambda_j I_V) \right) (T - \lambda_k I_V) \right)(v_k) \\ &\overset{(3)}{=} \left(\prod_{1 \leq j \neq k \leq s} (T - \lambda_j I_V) \right) \Big((T - \lambda_k I_V)(v_k) \Big) \end{aligned}$$

$$\overset{(4)}{=} \left(\prod_{1\leq j\neq k\leq s} (T-\lambda_j I_V) \right)(0_V) \overset{(5)}{=} 0_V.$$

Logo, $f(T)(v) = f(T)\left(\sum_{k=1}^{s} v_k\right) = \sum_{k=1}^{s} f(T)(v_k) = \sum_{k=1}^{s} 0_V = 0_V$, para todo $v \in V$, que é equivalente a $f(T) = 0$.

Pela Proposição 5, $m(x)$ divide $f(x)$. Como $f(x)$ e $m(x)$ são mônicos, obtemos $m(x) = f(x)$.

Reciprocamente, suponhamos que o polinômio mínimo do operador T seja $m(x) = (x-\lambda_1)\cdot\ldots\cdot(x-\lambda_s)$, onde $\lambda_1, \ldots, \lambda_s$ são os autovalores distintos de T. Pela Proposição 6, os fatores irredutíveis do polinômio característico de T são os mesmos de $m(x)$, portanto $p(x) = (x-\lambda_1)^{n_1}\cdot\ldots\cdot(x-\lambda_s)^{n_s}$. Sabemos que $\dim_K V = \text{grau}(p(x)) = n_1 + \cdots + n_s$. A demonstração é por indução sobre s, o número de autovalores distintos de T.

Se $s = 1$ e $m(x) = x - \lambda_1$, então $p(x) = (x-\lambda_1)^n$, onde $n = \dim_K V$. Como $m(T) = T - \lambda_1 I = 0$, então $T = \lambda_1 I$ e todo $v \in V$, $v \neq 0_V$ é autovetor de T, isto é, $V = V_{\lambda_1}$ e T é diagonalizável.

Suponhamos o resultado válido para operadores lineares com autovalores distintos $\lambda_1, \ldots, \lambda_r$ e polinômio mínimo $\prod_{j=1}^{r}(x-\lambda_j)$, onde $1 \leq r < s$.

Seja $T : V \longrightarrow V$ um operador linear com autovalores distintos $\lambda_1, \ldots, \lambda_s$ e polinômio mínimo $m(x) = (x-\lambda_1)\cdot\ldots\cdot(x-\lambda_s)$, onde $s > 1$.

Definimos $g(x) = \prod_{j=1}^{s-1}(x-\lambda_j)$ e $h(x) = x-\lambda_s$. Então, $m(x) = g(x)h(x)$ e $\text{mdc}_{K[x]}(g(x), h(x)) = 1$. Pelo Lema anterior, item (a), $V = U \oplus W$, onde $U = \text{Núcleo}(g(T))$ e $W = \text{Núcleo}(h(T)) = \text{Núcleo}(T - \lambda_s I_V)$ são subespaços de V invariantes por T.

Pelo Lema anterior item (b), $g(x)$ e $h(x)$ são os polinômios mínimos de T_1 e T_2, respectivamente.

Por hipótese de indução, T_1 e T_2 são diagonalizáveis. Tomando uma base δ_1 de U formada por autovetores de T_1, associados aos autovalores distintos $\lambda_1, \ldots, \lambda_{s-1}$ de T_1, e uma base δ_2 de W formada por autovetores de T_2, associados ao autovalor λ_s de T_2, temos que $\delta = \delta_1 \cup \delta_2$ é uma base de V formada por autovetores de T. ■

Para todo $u \in U$, temos $T_1(u) = T(u)$ e, para todo $w \in W$, temos $T_2(w) = T(w)$.

Exercícios

1. Seja $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por $T(x, y) = (2x + y, 2y)$.
   (a) Determine os polinômios característico e mínimo de T.
   (b) T é diagonalizável?
   (c) Determine T^{-1} como um polinômio em T com coeficientes reais.
2. Seja $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definida por $T(x, y, z) = (x + y, y, z)$.
   (a) Determine os polinômios característico e mínimo de T.
   (b) T é diagonalizável?
   (c) Determine T^{-1} como um operador polinomial em T.
3. Seja $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por $T(x, y) = (2x, x + y)$. Determine o operador $f(T)$, onde $f(x) = x^2 - x + 3$.

Operadores polinomiais em T comutam.

4. Sejam V um K-espaço vetorial de dimensão $n \geq 1$ e T um operador linear. Sejam $f(x) = a_\ell x^\ell + \cdots + a_1 x + a_0$ e $g(x) = b_m x^m + \cdots + b_1 x + b_0$ em $K[x]$. Mostre que $f(T)g(T) = g(T)f(T)$.
5. Dada $A = T]_\alpha^\alpha$, onde α é a base canônica do $\mathbb{R}^5$, determine os polinômios característico e mínimo de $T : \mathbb{R}^5 \longrightarrow \mathbb{R}^5$.

(a) $\begin{pmatrix} 5 & 2 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 \\ 0 & 0 & 7 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}$ (c) $\begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$

6. Seja $D : P_5(\mathbb{R}) \longrightarrow P_5(\mathbb{R})$ o operador derivação. Determine os polinômios característico e mínimo de D.
7. Seja $D^2 : P_5(\mathbb{R}) \longrightarrow P_5(\mathbb{R})$ o operador derivação de segunda ordem. Determine os polinômios característico e mínimo de D^2.
8. Seja $D^3 : P_5(\mathbb{R}) \longrightarrow P_5(\mathbb{R})$ o operador derivação de terceira ordem. Determine os polinômios característico e mínimo de D^3.
9. Sejam V um K-espaço vetorial de dimensão $n \geq 1$ e $T : V \longrightarrow V$ um operador linear tal que $T^2 = T$.
   (a) Mostre que T é diagonalizável.

(b) Mostre que se $T \neq 0$, então existe uma base β de V tal que $T]_\beta^\beta = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$, onde I_r é a matriz identidade de ordem $r = \dim_K \text{Imagem}(T)$.

10. Sejam $B \in M_{r\times r}(K)$, $C \in M_{s\times s}(K)$, $n = r+s$ e $A \in M_{n\times n}(K)$ a matriz em blocos definida por $A = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}$.

(a) Mostre, por indução sobre $m \geq 1$, que $A^m = \begin{pmatrix} B^m & 0 \\ 0 & C^m \end{pmatrix}$.

(b) Seja $\ell(x) = a_n x^n + \cdots + a_1 x + a_0 \in K[x]$. Mostre que $\ell(A) = \begin{pmatrix} \ell(B) & 0 \\ 0 & \ell(C) \end{pmatrix}$.

11. Seja $T : V \longrightarrow V$ um operador linear K-linear, onde V é um K-espaço vetorial de dimensão $n \geq 1$. Seja U um subespaço de V invariante por T e $T_1 = T|_U$ o operador linear definido por

$$\begin{array}{rcl} T_1 : U & \longrightarrow & U \\ u & \longmapsto & T(u) \end{array}$$

(a) Mostre que $T^j|_U = T_1^j$, para todo $j \geq 1$.

(b) Mostre que $(aT^j)|_U = aT_1^j$, para todo $j \geq 1$ e para todo $a \in K$.

(c) Seja $\ell(x) = \sum_{j=0}^{m} a_j x^j \in K[x]$. Mostre que $\ell(T)|_U = \ell(T_1)$.