RESOLUÇÃO DOS EXERCÍCIOS – FOLHA 1

1.1. (a) $r(\mathbf{A}) = r(\mathbf{A}^{\mathtt{T}}) = 2$, $n(\mathbf{A}) = 2$ e $n(\mathbf{A}^{\mathtt{T}}) = 1$. Bases:

$$\mathcal{R}(\mathbf{A}) : \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \right\}, \qquad \mathcal{N}(\mathbf{A}) : \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \end{bmatrix} \right\}.$$

$$\mathcal{R}(\mathbf{A}^{\mathtt{T}}) : \left\{ \begin{bmatrix} 1 \\ 2 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 1 \\ 3 \end{bmatrix} \right\}, \qquad \mathcal{N}(\mathbf{A}^{\mathtt{T}}) : \left\{ \begin{bmatrix} 1 \\ -5 \\ 3 \end{bmatrix} \right\}.$$

(b) $r(\mathbf{A}) = r(\mathbf{A}^{\mathtt{T}}) = 2$, $n(\mathbf{A}) = 3$ e $n(\mathbf{A}^{\mathtt{T}}) = 1$. Bases:

$$\mathcal{R}(\mathbf{A}) : \left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \right\}, \qquad \mathcal{N}(\mathbf{A}) : \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -4 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

$$\mathcal{R}(\mathbf{A}^{\mathtt{T}}) : \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 3 \\ 4 \end{bmatrix} \right\}, \qquad \mathcal{N}(\mathbf{A}^{\mathtt{T}}) : \left\{ \begin{bmatrix} 4 \\ 1 \\ -2 \end{bmatrix} \right\}.$$

1.2. Sejam $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \Bbbk^{n\times 1}$ tais que $\mathbf{A} = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix}$, de modo que $\begin{bmatrix} \mathbf{A}\ \mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 & \cdots \mathbf{v}_n\ \mathbf{b} \end{bmatrix}$. Como $\mathcal{R}(\mathbf{A}) \subseteq \mathcal{R}\big(\begin{bmatrix} \mathbf{A}\ \mathbf{b} \end{bmatrix}\big)$, obtemos

$$\begin{aligned} \mathbf{b} \in \mathcal{R}(\mathbf{A}) \quad &\iff \quad \mathcal{R}\big(\begin{bmatrix} \mathbf{A}\ \mathbf{b} \end{bmatrix}\big) = \langle \mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{b} \rangle = \langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle = \mathcal{R}(\mathbf{A}) \\ &\iff \quad r\big(\begin{bmatrix} \mathbf{A}\ \mathbf{b} \end{bmatrix}\big) = \dim \mathcal{R}\big(\begin{bmatrix} \mathbf{A}\ \mathbf{b} \end{bmatrix}\big) = \dim \mathcal{R}(\mathbf{A}) = r(\mathbf{A}). \end{aligned}$$

1.3. Pondo $\mathbf{b} = \begin{bmatrix} -2 \\ -5 \\ -6 \\ -7 \\ -7 \end{bmatrix}$, tem-se $r\big([\mathbf{A}\ \mathbf{b}]\big) = r(\mathbf{A}) = 3$

1.4. $\mathbf{A}$ é invertível se e só se $r(\mathbf{A}) = n$. Para qualquer $\mathbf{v} \in \Bbbk^{n\times 1}$, temos

$$\begin{aligned} \mathbf{v} \in \mathcal{N}(\mathbf{A}) \quad &\iff \quad \mathbf{A}\mathbf{v} = \mathbf{0} \quad \iff \quad \mathbf{A}_1\mathbf{v} = \mathbf{0} \quad \text{e} \quad \mathbf{A}_2\mathbf{v} = \mathbf{0} \\ &\iff \quad \mathbf{v} \in \mathcal{N}(\mathbf{A}_1) \quad \text{e} \quad \mathbf{v} \in \mathcal{N}(\mathbf{A}_2). \end{aligned}$$

Como $\mathcal{N}(\mathbf{A}_1) = \mathcal{R}(\mathbf{A}_2^{\mathsf{T}})$, existe $\mathbf{w} \in \Bbbk^{n\times 1}$ tal que $\mathbf{A}_2^{\mathsf{T}}\mathbf{w} = 0$. Assim,

$$\mathbf{v}^{\mathsf{T}}\mathbf{v} = (\mathbf{A}_2^{\mathsf{T}}\mathbf{w})^{\mathsf{T}}\mathbf{v} = \mathbf{w}^{\mathsf{T}}\mathbf{A}_2\mathbf{v} = \mathbf{0},$$

logo $\sum_{1\leqslant i\leqslant n} \mathbf{v}_i^2 = 0$ e, portanto, $\mathbf{v} = \mathbf{0}$. Segue-se que $\mathcal{N}(\mathbf{A}) = \{\mathbf{0}\}$, logo

$$r(\mathbf{A}) = n - n(\mathbf{A}) = n - \dim \mathcal{N}(\mathbf{A}) = n.$$

1.5. Para qualquer $\mathbf{v} \in \Bbbk^{m\times 1}$, temos $\mathbf{v} \in \mathcal{R}\big([\mathbf{A}\ \ \mathbf{B}]\big)$ se e só se $\mathbf{v} = \big([\mathbf{A}\ \ \mathbf{B}]\big)$ para algum $\mathbf{w} \in \Bbbk^{(n+p)\times 1}$. Pondo $\mathbf{w} = \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix}$, em que $\mathbf{w}_1 \in \Bbbk^{n\times 1}$ e $\mathbf{w}_2 \in \Bbbk^{p\times 1}$, vem

$$\mathbf{v} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix} = \mathbf{A}\mathbf{w}_1 + \mathbf{B}\mathbf{w}_2 \in \mathcal{R}(\mathbf{A}) + \mathcal{R}(\mathbf{B}),$$

logo

$$\mathbf{v} \in \mathcal{R}\big([\mathbf{A}\ \mathbf{B}]\big) \quad \iff \quad \mathbf{v} \in \mathcal{R}(\mathbf{A}) + \mathcal{R}(\mathbf{B}).$$

1.6. (a) Para qualquer $\mathbf{v} \in \mathbb{C}^{n\times 1}$, tem-se

$$\begin{aligned} \mathbf{v} \in \mathcal{N}(\mathbf{A}^*) \cap \mathcal{R}(\mathbf{A}) \quad &\iff \quad \mathbf{A}^*\mathbf{v} = \mathbf{0} \quad \text{e} \quad \mathbf{v} = \mathbf{A}\mathbf{w}, \quad \text{para algum } \mathbf{w} \in \mathbb{C}^{n\times 1} \\ &\implies \quad \mathbf{v}^*\mathbf{v} = (\mathbf{A}\mathbf{w})^*\mathbf{v} = \mathbf{w}^*\mathbf{A}^*\mathbf{w} = \mathbf{0} \\ &\implies \quad \sum_{1\leqslant i\leqslant n} |v_i|^2 = 0 \quad \implies \quad \mathbf{v} = \mathbf{0} \end{aligned}$$

e, portanto,

$$\mathcal{N}(\mathbf{A}^*) \cap \mathcal{R}(\mathbf{A}) = \{\mathbf{0}\}.$$

Deste modo,

$$r(\mathbf{A}^*\mathbf{A}) = r(\mathbf{A}) - \dim\big(\mathcal{N}(\mathbf{A}^*) \cap \mathcal{R}(\mathbf{A})\big) = r(\mathbf{A}).$$

Substituindo $\mathbf{A}$ por $\mathbf{A}^*$, obtemos

$$r(\mathbf{A}\mathbf{A}^*) = r((\mathbf{A}^*)^*) = r(\mathbf{A}^*) = r(\mathbf{A}).$$

(b),(c) Para quaisquer $\mathbf{X}, \mathbf{Y} \in \mathbb{C}^{n\times p}$, temos

$$\mathcal{R}(\mathbf{XY}) \subseteq \mathcal{R}(\mathbf{X}) \quad \text{e} \quad \mathcal{N}(\mathbf{Y}) \subseteq \mathcal{N}(\mathbf{XY}).$$

Em particular,

$$\mathcal{R}(\mathbf{A}^*\mathbf{A}) \subseteq \mathcal{R}(\mathbf{A}^*) \qquad \text{e} \qquad \mathcal{N}(\mathbf{A}) \subseteq \mathcal{N}(\mathbf{A}^*\mathbf{A}).$$

Como $r(\mathbf{A}^*\mathbf{A}) = r(\mathbf{A}) = r(\mathbf{A}^*)$, concluímos que $\mathcal{R}(\mathbf{A}^*\mathbf{A}) = \mathcal{R}(\mathbf{A}^*)$. Por outro lado,

$$n(\mathbf{A}) = n - r(\mathbf{A}) = n - r(\mathbf{A}^*\mathbf{A}) = n(\mathbf{A}^*\mathbf{A}),$$

logo $\mathcal{N}(\mathbf{A}) = \mathcal{N}(\mathbf{A}^*\mathbf{A})$.

Para as outras igualdades, basta trocar $\mathbf{A}$ por $\mathbf{A}^*$.

1.7. Temos $\mathcal{N}(\mathbf{B}) \subseteq \mathcal{N}(\mathbf{AB})$, logo $n(\mathbf{B}) \leqslant n(\mathbf{AB})$. Por outro lado, $\mathcal{R}(\mathbf{AB}) \subseteq \mathcal{R}(\mathbf{A})$, logo

$$n(\mathbf{A}) = n - r(\mathbf{A}) \leqslant n - r(\mathbf{AB}) = n(\mathbf{AB})$$

e, portanto, $\max\{n(\mathbf{A}), n(\mathbf{B})\} \leqslant n(\mathbf{AB})$.

Para a outra desigualdade, sabemos que $r(\mathbf{A}) + r(\mathbf{B}) - n \leqslant r(\mathbf{AB})$, logo

$$\begin{aligned} n(\mathbf{AB}) &= n - r(\mathbf{AB}) \leqslant n - \big(r(\mathbf{A}) + r(\mathbf{B}) - n\big) \\ &= \big(n - r(\mathbf{A})\big) + \big(n - r(\mathbf{B})\big) = n(\mathbf{A}) + n(\mathbf{B}). \end{aligned}$$

1.8. (a) Temos

$$r(\mathbf{B}) = n \quad \Longrightarrow \quad \mathcal{R}(\mathbf{B}) = \mathbb{k}^{n\times n} \quad \Longrightarrow \quad \mathcal{N}(\mathbf{A}) \cap \mathcal{R}(\mathbf{B}) = \mathcal{N}(\mathbf{A}).$$

Sendo assim, se $r(\mathbf{B}) = n$, então

$$r(\mathbf{AB}) = r(\mathbf{B}) - \dim\big(\mathcal{N}(\mathbf{A}) \cap \mathcal{R}(\mathbf{B})\big) = n - \dim \mathcal{N}(\mathbf{A}) = n - n(\mathbf{A}) = r(\mathbf{A}),$$

logo $\mathcal{R}(\mathbf{AB}) = \mathcal{R}(\mathbf{A})$ (porque $\mathcal{R}(\mathbf{AB}) \subseteq \mathcal{R}(\mathbf{A})$).

(b) Temos

$$r(\mathbf{A}) = n \quad \Longrightarrow \quad n(\mathbf{A}) = 0 \quad \Longrightarrow \quad \mathcal{N}(\mathbf{A}) = \{\mathbf{0}\} \quad \Longrightarrow \quad \mathcal{N}(\mathbf{A}) \cap \mathcal{R}(\mathbf{B}) = \{\mathbf{0}\}.$$

Sendo assim, se $r(\mathbf{A}) = n$, então

$$r(\mathbf{AB}) = r(\mathbf{B}) - \dim\big(\mathcal{N}(\mathbf{A}) \cap \mathcal{R}(\mathbf{B})\big) = r(\mathbf{B}),$$

logo

$$n(\mathbf{AB}) = n - r(\mathbf{AB}) = n - r(\mathbf{B}) = n(\mathbf{B})$$

e, portanto, $\mathcal{N}(\mathbf{AB}) = \mathcal{N}(\mathbf{B})$ (porque $\mathcal{N}(\mathbf{B}) \subseteq \mathcal{N}(\mathbf{AB})$).

RESOLUÇÃO DOS EXERCÍCIOS – FOLHA 2

2.1. $\|\mathbf{u}\| = \|\mathbf{v}\| = \|\mathbf{w}\| = 1$ e $\langle \mathbf{u}|\mathbf{v}\rangle = \langle \mathbf{u}|\mathbf{w}\rangle = \langle \mathbf{v}|\mathbf{w}\rangle = 0$.

Os coeficientes de Fourier de $\mathbf{a}$ são

$$\langle \mathbf{a}|\mathbf{u}\rangle = -\frac{3}{\sqrt{2}}, \quad \langle \mathbf{a}|\mathbf{v}\rangle = \frac{2}{\sqrt{3}}, \quad \langle \mathbf{a}|\mathbf{w}\rangle = \frac{1}{\sqrt{6}},$$

logo

$$\mathbf{a} = -\frac{3}{\sqrt{2}}\,\mathbf{u} + \frac{2}{\sqrt{3}}\,\mathbf{v} + \frac{1}{\sqrt{6}}\,\mathbf{w}.$$

Os coeficientes de Fourier de $\mathbf{b}$ são

$$\langle \mathbf{b}|\mathbf{u}\rangle = \frac{1}{\sqrt{2}}, \quad \langle \mathbf{b}|\mathbf{v}\rangle = -\frac{1}{\sqrt{3}}, \quad \langle \mathbf{b}|\mathbf{w}\rangle = -\frac{5}{\sqrt{6}},$$

logo

$$\mathbf{b} = \frac{1}{\sqrt{2}}\,\mathbf{u} - \frac{1}{\sqrt{3}}\,\mathbf{v} - \frac{5}{\sqrt{6}}\,\mathbf{w}.$$

2.2. Dado que $\mathbf{y} = \sum_{1\leqslant i\leqslant m} \langle \mathbf{v}_i|\mathbf{y}\rangle \mathbf{v}_i$, vem

$$\langle \mathbf{x}|\mathbf{y}\rangle = \left\langle \mathbf{x} \;\Big|\; \sum_{1\leqslant i\leqslant m} \langle \mathbf{v}_i|\mathbf{y}\rangle \mathbf{v}_i \right\rangle = \sum_{1\leqslant i\leqslant m} \langle \mathbf{v}_i|\mathbf{y}\rangle \langle \mathbf{x}|\mathbf{v}_i\rangle = \sum_{1\leqslant i\leqslant m} \langle \mathbf{x}|\mathbf{v}_i\rangle \overline{\langle \mathbf{y}|\mathbf{v}_i\rangle}$$

2.3. Aplicando o método de Gram-Schmidt às bases do Exercício 1.1(b), obtemos bases ortonormadas seguintes:

$$\mathcal{R}(\mathbf{A}) : \left\{ \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{14}} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\}, \qquad \mathcal{N}(\mathbf{A}) : \left\{ \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \sqrt{\frac{5}{54}} \begin{bmatrix} 2/5 \\ 4/5 \\ -3 \\ 1 \\ 0 \end{bmatrix}, \frac{1}{2\sqrt{682}} \begin{bmatrix} -17 \\ -34 \\ -21 \\ -29 \\ 27 \end{bmatrix} \right\}.$$

$$\mathcal{R}(\mathbf{A}^{\mathrm{T}}) : \left\{ \frac{1}{\sqrt{10}} \begin{bmatrix} 1 \\ 2 \\ 0 \\ -2 \\ 1 \end{bmatrix}, \frac{1}{4} \begin{bmatrix} -1 \\ -2 \\ 1 \\ 1 \\ 3 \end{bmatrix} \right\}, \qquad \mathcal{N}(\mathbf{A}^{\mathrm{T}}) : \frac{1}{\sqrt{21}} \left\{ \begin{bmatrix} 4 \\ 1 \\ -2 \end{bmatrix} \right\}.$$

2.4. $\mathbf{U}^* = \begin{bmatrix} \dfrac{1-i}{\sqrt{3}} & -\dfrac{i}{\sqrt{3}} \\ \dfrac{1-i}{\sqrt{6}} & \dfrac{2i}{\sqrt{6}} \end{bmatrix}$ e $\mathbf{U}\mathbf{U}^* = \mathbf{I}_2$, logo $\mathbf{U}$ é unitária.

2.5. (a) Temos

$$\mathbf{P}\mathbf{P}^{\mathsf{T}} = \begin{bmatrix} \alpha+\beta & -(\alpha-\beta) \\ \alpha-\beta & \alpha+\beta \end{bmatrix} \begin{bmatrix} \alpha+\beta & \alpha-\beta \\ -(\alpha-\beta) & \alpha+\beta \end{bmatrix} = \begin{bmatrix} 2(\alpha^2+\beta^2) & 0 \\ 0 & 2(\alpha^2+\beta^2) \end{bmatrix},$$

logo

$$\mathbf{P}^{\mathsf{T}} = \mathbf{P}^{-1} \quad \Longleftrightarrow \quad \mathbf{P}\mathbf{P}^{\mathsf{T}} = \mathbf{I}_2 \quad \Longleftrightarrow \quad 2(\alpha^2+\beta^2) = 1.$$

(b) Temos

$$\mathbf{U}\mathbf{U}^{\mathsf{T}} = \begin{bmatrix} 0 & \alpha & 0 & \beta i \\ \alpha & 0 & \beta i & 0 \\ 0 & \beta i & 0 & \alpha \\ \beta i & 0 & \alpha & 0 \end{bmatrix} \begin{bmatrix} 0 & \alpha & 0 & -\beta i \\ \alpha & 0 & -\beta i & 0 \\ 0 & -\beta i & 0 & \alpha \\ -\beta i & 0 & \alpha & 0 \end{bmatrix} = \begin{bmatrix} \alpha^2+\beta^2 & 0 & 0 & 0 \\ 0 & \alpha^2+\beta^2 & 0 & 0 \\ 0 & 0 & \alpha^2+\beta^2 & 0 \\ 0 & 0 & 0 & \alpha^2+\beta^2 \end{bmatrix},$$

logo

$$\mathbf{U}^* = \mathbf{U}^{-1} \quad \Longleftrightarrow \quad \mathbf{U}\mathbf{U}^* = \mathbf{I}_4 \quad \Longleftrightarrow \quad \alpha^2+\beta^2 = 1.$$

2.6. (a) $(\mathbf{U}\mathbf{V})(\mathbf{U}\mathbf{V})^* = (\mathbf{U}\mathbf{V})(\mathbf{V}^*\mathbf{U}^*) = \mathbf{U}(\mathbf{V}\mathbf{V}^*)\mathbf{U}^* = \mathbf{U}\mathbf{I}_n\mathbf{U}^* = \mathbf{U}\mathbf{U}^* = \mathbf{I}_n$.

(b) $\begin{bmatrix} \mathbf{U} & \mathbf{0} \\ \mathbf{0} & \mathbf{V} \end{bmatrix} \begin{bmatrix} \mathbf{U} & \mathbf{0} \\ \mathbf{0} & \mathbf{V} \end{bmatrix}^* = \begin{bmatrix} \mathbf{U} & \mathbf{0} \\ \mathbf{0} & \mathbf{V} \end{bmatrix} \begin{bmatrix} \mathbf{U}^* & \mathbf{0} \\ \mathbf{0} & \mathbf{V}^* \end{bmatrix} = \begin{bmatrix} \mathbf{U}\mathbf{U}^* & \mathbf{0} \\ \mathbf{0} & \mathbf{V}\mathbf{V}^* \end{bmatrix} = \begin{bmatrix} \mathbf{I}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_n \end{bmatrix} = \mathbf{I}_{m+n}$.

2.7. $\langle \mathbf{U}\mathbf{u} | \mathbf{U}\mathbf{v} \rangle = (\mathbf{U}\mathbf{u})^*(\mathbf{U}\mathbf{v}) = \mathbf{u}^*\mathbf{U}^*\mathbf{U}\mathbf{v} = \mathbf{u}^*\mathbf{I}_n\mathbf{v} = \mathbf{u}^*\mathbf{v} = \langle \mathbf{u} | \mathbf{v} \rangle$.

RESOLUÇÃO DOS EXERCÍCIOS – FOLHA 3

3.1. (a) Temos

$$(\mathbf{R}_u)^2 = \big(\mathbf{I}_n - \frac{2}{\|\mathbf{u}\|^2}\,\mathbf{u}\mathbf{u}^*\big)^2 = \mathbf{I}_n - \frac{4}{\|\mathbf{u}\|^2}\,\mathbf{u}\mathbf{u}^* + \frac{4}{\|\mathbf{u}\|^2}\,(\mathbf{u}\mathbf{u}^*)(\mathbf{u}\mathbf{u}^*)$$
$$= \mathbf{I}_n - \frac{4}{\|\mathbf{u}\|^2}\,\mathbf{u}\mathbf{u}^* + \frac{4}{\|\mathbf{u}\|^2}\,\mathbf{u}(\mathbf{u}^*\mathbf{u})\mathbf{u}^* = \mathbf{I}_n - \frac{4}{\|\mathbf{u}\|^2}\,\mathbf{u}\mathbf{u}^* + \frac{4}{\|\mathbf{u}\|^2}\,\|\mathbf{u}\|^2\mathbf{u}\mathbf{u}^* = \mathbf{I}_n.$$

Sendo assim, $\mathbf{R}_\mathbf{u}$ é invertível e $(\mathbf{R}_\mathbf{u})^{-1} = \mathbf{R}_\mathbf{u}$. Também,

$$(\mathbf{R}_\mathbf{u})^* = \big(\mathbf{I}_n - \frac{2}{\|\mathbf{u}\|^2}\,\mathbf{u}\mathbf{u}^*\big)^* = (\mathbf{I}_n)^* - \frac{2}{\|\mathbf{u}\|^2}\,(\mathbf{u}\mathbf{u}^*)^* = \mathbf{I}_n - \frac{2}{\|\mathbf{u}\|^2}\,\mathbf{u}\mathbf{u}^* = \mathbf{R}_\mathbf{u}.$$

(b) (i) Temos

$$\mathbf{R}_\mathbf{u}\mathbf{v} = \big(\mathbf{I}_n - \frac{2}{\|\mathbf{u}\|^2}\,\mathbf{u}\mathbf{u}^*\big)\mathbf{v} = \mathbf{v} - \frac{2}{\|\mathbf{u}\|^2}\,\mathbf{u}\mathbf{u}^*\mathbf{v} = \mathbf{v} - \frac{2\langle\mathbf{u}|\mathbf{v}\rangle}{\|\mathbf{u}\|^2}\,\mathbf{u}.$$

Por outro lado,

$$\langle\mathbf{u}|\mathbf{v}\rangle = \langle\mathbf{v}|\mathbf{v}\rangle \pm \overline{\mu}\|\mathbf{v}\|\langle\mathbf{e}_1|\mathbf{v}\rangle = \|\mathbf{v}\|^2 \pm \overline{\mu}\|\mathbf{v}\|\,v_1$$
$$= \|\mathbf{v}\|\big(\|\mathbf{v}\| \pm \overline{\mu}\,v_1\big) = \begin{cases} \|\mathbf{v}\|(1 \pm v_1), & \text{se } v_1 \in \mathbb{R}, \\ \|\mathbf{v}\|\big(\|\mathbf{v}\| \pm |v_1|\big), & \text{se } v_1 \notin \mathbb{R}, \end{cases}$$

Por outro lado,

$$\langle\mathbf{u}|\mathbf{u}\rangle = \langle\mathbf{v}|\mathbf{v}\rangle \pm \mu\|\mathbf{v}\|\langle\mathbf{v}|\mathbf{e}_1\rangle \pm \overline{\mu}\|\mathbf{v}\|\langle\mathbf{e}_1|\mathbf{v}\rangle + |\mu|^2\|\mathbf{v}\|^2$$
$$= (1 + |\mu|^2)\|\mathbf{v}\|^2 \pm \|\mathbf{v}\|(\mu\overline{v_1} + \overline{\mu}v_1)$$
$$= \|\mathbf{v}\|\big((1 + |\mu|^2)\|\mathbf{v}\| \pm (\mu\overline{v_1} + \overline{\mu}v_1)\big)$$
$$= \begin{cases} 2\|\mathbf{v}\|(1 \pm v_1), & \text{se } v_1 \in \mathbb{R}, \\ 2\|\mathbf{v}\|\big(\|\mathbf{v}\| \pm |v_1|\big), & \text{se } v_1 \notin \mathbb{R}. \end{cases}$$

Em conclusão,

$$\|\mathbf{u}\|^2 = \langle\mathbf{u}|\mathbf{u}\rangle = 2\langle\mathbf{u}|\mathbf{v}\rangle$$

e, portanto,

$$\mathbf{R}_\mathbf{u}\mathbf{v} = \mathbf{v} - \mathbf{u} = \mp\mu\|\mathbf{v}\|\mathbf{e}_1.$$

(ii) Pela alínea (a), temos

$$\mathbf{U}^* = \big(\mp\mu\mathbf{R_u}\big)^* = \mp\overline{\mu}(\mathbf{R_u})^* = \mp\overline{\mu}\mathbf{R_u},$$

logo

$$\mathbf{U}\mathbf{U}^* = (\mp\mu\mathbf{R_u})(\mp\overline{\mu}\mathbf{R_u}) = |\mu|^2(\mathbf{R_u})^2 = |\mu|^2\mathbf{I}_n = \mathbf{I}_n$$

e, portanto, $\mathbf{U}$ é unitária.

Pela alínea anterior, sabemos que $\mathbf{R_u}\mathbf{v} = \mp\mu\|\mathbf{v}\|\mathbf{e}_1 = \mp\mu\mathbf{e}_1$, logo

$$\mathbf{v} = \mathbf{I}_n\mathbf{v} = (\mathbf{R_u})^2\mathbf{v} = \mathbf{R_u}(\mp\mu\mathbf{e}_1) = \mp\mu\mathbf{R_u}\mathbf{e}_1 = \mathbf{U}\mathbf{e}_1,$$

o que significa que $\mathbf{v}$ é a primeira coluna de $\mathbf{U}$.

3.2. Consideremos $\mathbf{u} = \mathbf{v} - \mathbf{e}_1 = -\frac{2}{3}\big[2\ -1\ 0\ 1\big]^{\mathsf{T}}$. Pelo exercício anterior, a matriz que se quer é

$$\mathbf{U} = \mathbf{R_u} = \mathbf{I}_4 - \frac{2}{\|\mathbf{u}\|^2}\mathbf{u}\mathbf{u}^* = \frac{1}{3}\begin{bmatrix} -1 & -2 & 0 & -2 \\ 2 & 2 & 0 & 1 \\ 0 & 0 & 3 & 0 \\ -2 & -1 & 0 & 2 \end{bmatrix}.$$

3.3. 1 é um valor próprio da matriz $\mathbf{A} = \begin{bmatrix} 13 & -9 \\ 16 & -11 \end{bmatrix}$ e $\mathbf{v} = \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix}$ é um vector próprio que lhe está associado. Temos $\|\mathbf{v}\| = 1$ e $v_1 = 3/5 \in \mathbb{R}$. Tomando $\mathbf{u} = \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 4/5 \end{bmatrix}$, o Exercício 3.1 garante que a matriz

$$\mathbf{P} = \mathbf{R_u} = \mathbf{I}_2 - \frac{2}{\|\mathbf{u}\|^2}\mathbf{u}\mathbf{u}^* = \begin{bmatrix} 3/5 & 4/5 \\ 4/5 & -3/5 \end{bmatrix}$$

é ortogonal (e tem $\mathbf{v}$ como primeira coluna). Além disso,

$$\mathbf{P}^{\mathsf{T}}\mathbf{A}\mathbf{P} = \begin{bmatrix} 1 & 25 \\ 0 & 1 \end{bmatrix}.$$

3.4. • Aplicamos o algoritmo descrito no teorema da decomposição de Schur.

Ora, 2 é um valor próprio da matriz $\mathbf{A} = \begin{bmatrix} 3 & 2 & 1 \\ 0 & 2 & 0 \\ -2 & -3 & 0 \end{bmatrix}$ e $\mathbf{v} = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ -\frac{\sqrt{2}}{2} \end{bmatrix}$ é um vector próprio que lhe está associado. Temos $\|\mathbf{v}\| = 1$ e $v_1 = \frac{\sqrt{2}}{2} \in \mathbb{R}$. Tomando $\mathbf{u} = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ -\frac{\sqrt{2}}{2} \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}-2}{2} \\ 0 \\ -\frac{\sqrt{2}}{2} \end{bmatrix}$, o

Exercício 3.1 garante que a matriz

$$\mathbf{V} = \mathbf{R_u} = \mathbf{I}_3 - \frac{2}{\|\mathbf{u}\|^2}\mathbf{u}\mathbf{u}^* = \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \\ 0 & 1 & 0 \\ -\frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \end{bmatrix}$$

é unitária (e tem $\mathbf{v}$ como primeira coluna). Além disso,

$$\mathbf{V}^*\mathbf{A}\mathbf{V} = \begin{bmatrix} 2 & \frac{5\sqrt{2}}{2} & -3 \\ 0 & 2 & 0 \\ 0 & \frac{\sqrt{2}}{2} & 1 \end{bmatrix}.$$

Agora, consideramos a matriz $\mathbf{A}_0 = \begin{bmatrix} 2 & 0 \\ \frac{\sqrt{2}}{2} & 1 \end{bmatrix}$. 2 é um valor próprio de $\mathbf{A}_0$ e $\mathbf{v} = \begin{bmatrix} \frac{\sqrt{6}}{3} \\ \frac{\sqrt{3}}{3} \end{bmatrix}$ é um vector próprio associado a 2. Temos $\|\mathbf{v}\| = 1$ e $v_1 = \frac{\sqrt{6}}{3} \in \mathbb{R}$. Tomando $\mathbf{u} = \begin{bmatrix} \frac{\sqrt{6}}{3} \\ \frac{\sqrt{3}}{3} \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{6}-3}{3} \\ \frac{\sqrt{3}}{3} \end{bmatrix}$, obtemos a matriz unitária

$$\mathbf{U}_0 = \mathbf{R_u} = \mathbf{I}_2 - \frac{2}{\|\mathbf{u}\|^2}\mathbf{u}\mathbf{u}^* = \begin{bmatrix} \frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} & -\frac{\sqrt{6}}{3} \end{bmatrix}$$

(que tem $\mathbf{v}$ como primeira coluna) e temos

$$\mathbf{U}_0^*\mathbf{A}_0\mathbf{U}_0 = \begin{bmatrix} 2 & \frac{\sqrt{2}}{2} \\ 0 & 1 \end{bmatrix}.$$

Para terminar, a matriz

$$\mathbf{U} = \mathbf{V}\begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_0 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} \\ 0 & \frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{3} \\ -\frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} \end{bmatrix}$$

é unitária e temos

$$\mathbf{U}^*\mathbf{A}\mathbf{U} = \begin{bmatrix} 2 & \frac{2\sqrt{3}}{3} & \frac{11\sqrt{6}}{6} \\ 0 & 2 & \frac{\sqrt{2}}{2} \\ 0 & 0 & 1 \end{bmatrix}.$$

- -1 é um valor próprio da matriz $\mathbf{A} = \begin{bmatrix} -4 & -3 & -3 \\ 0 & -1 & 0 \\ 6 & 6 & 5 \end{bmatrix}$ e $\mathbf{v} = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ -\frac{\sqrt{2}}{2} \end{bmatrix}$ é um vector próprio que lhe está associado. Temos $\|\mathbf{v}\| = 1$ e $v_1 = \frac{\sqrt{2}}{2} \in \mathbb{R}$. Como acima, tomamos $\mathbf{u} = \begin{bmatrix} \frac{\sqrt{2}-2}{2} \\ 0 \\ -\frac{\sqrt{2}}{2} \end{bmatrix}$ e obtemos

$$\mathbf{V}=\mathbf{R_u}=\begin{bmatrix}\frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2}\\ 0 & 1 & 0\\ -\frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2}\end{bmatrix} \qquad \text{e} \qquad \mathbf{V}^*\mathbf{A}\mathbf{V}=\begin{bmatrix}-1 & 5\sqrt{2} & 9\\ 0 & -1 & 0\\ 0 & -\sqrt{2} & 2\end{bmatrix}.$$

Agora, consideramos a matriz $\mathbf{A}_0=\begin{bmatrix}-1 & 0\\ -\sqrt{2} & 2\end{bmatrix}$. 2 é um valor próprio de $\mathbf{A}_0$ e $\mathbf{v}=\begin{bmatrix}0\\1\end{bmatrix}$ é um vector próprio que lhe está associado. Como $v_1=0$, não podemos aplicar o Exercício 3.1. No entanto, a matriz $\mathbf{U}_0=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$ é unitária e temos

$$\mathbf{U}_0^*\mathbf{A}_0\mathbf{U}_0=\begin{bmatrix}2 & -\sqrt{2}\\ 0 & -1\end{bmatrix}.$$

Para terminar, a matriz

$$\mathbf{U}=\mathbf{V}\begin{bmatrix}1 & \mathbf{0}\\ \mathbf{0} & \mathbf{U}_0\end{bmatrix}=\begin{bmatrix}\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0\\ 0 & 0 & 1\\ -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0\end{bmatrix}$$

é unitária e temos

$$\mathbf{U}^*\mathbf{A}\mathbf{U}=\begin{bmatrix}-1 & 9 & 5\sqrt{2}\\ 0 & 2 & -\sqrt{2}\\ 0 & 0 & -1\end{bmatrix}.$$

- -1 é um valor próprio da matriz $\mathbf{A}=\begin{bmatrix}1 & -2 & 0\\ 0 & 1 & -2\\ -2 & 0 & 1\end{bmatrix}$ e $\mathbf{v}=\begin{bmatrix}\frac{\sqrt{3}}{3}\\ \frac{\sqrt{3}}{3}\\ \frac{\sqrt{3}}{3}\end{bmatrix}$ é um vector próprio que lhe está associado. Temos $\|\mathbf{v}\|=1$ e $v_1=\frac{\sqrt{3}}{3}\in\mathbb{R}$. Tomamos $\mathbf{u}=\begin{bmatrix}\frac{\sqrt{3}}{3}\\ \frac{\sqrt{3}}{3}\\ \frac{\sqrt{3}}{3}\end{bmatrix}-\begin{bmatrix}1\\0\\0\end{bmatrix}=\begin{bmatrix}\frac{\sqrt{3}-3}{3}\\ \frac{\sqrt{3}}{3}\\ \frac{\sqrt{3}}{3}\end{bmatrix}$ e obtemos

$$\mathbf{V}=\mathbf{R_u}=\begin{bmatrix}\frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3}\\ \frac{\sqrt{3}}{3} & \frac{3-\sqrt{3}}{6} & -\frac{3+\sqrt{3}}{6}\\ \frac{\sqrt{3}}{3} & -\frac{3+\sqrt{3}}{6} & \frac{3-\sqrt{3}}{6}\end{bmatrix} \qquad \text{e} \qquad \mathbf{V}^*\mathbf{A}\mathbf{V}=\begin{bmatrix}-1 & 0 & 0\\ 0 & 2 & \sqrt{3}\\ 0 & -\sqrt{3} & 2\end{bmatrix}.$$

Agora, consideramos a matriz $\mathbf{A}_0=\begin{bmatrix}2 & \sqrt{3}\\ -\sqrt{3} & 2\end{bmatrix}$. $2-i\sqrt{3}\in\mathbb{C}$ é um valor próprio de $\mathbf{A}_0$ e $\mathbf{v}=\begin{bmatrix}\frac{\sqrt{2}}{2}\\ -\frac{\sqrt{2}}{2}i\end{bmatrix}\in\mathbb{C}^{2\times 1}$ é um vector próprio que lhe está associado. Temos $\|\mathbf{v}\|=1$ e $v_1=\frac{\sqrt{2}}{2}\in\mathbb{R}$.

Tomando $\mathbf{u} = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2}\,i \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{2-\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2}\,i \end{bmatrix}$, obtemos a matriz unitária

$$\mathbf{U}_0 = \mathbf{R}_{\mathbf{u}} = \mathbf{I}_2 - \frac{2}{\|\mathbf{u}\|^2}\,\mathbf{u}\mathbf{u}^* = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\,i \\ -\frac{\sqrt{2}}{2}\,i & -\frac{\sqrt{2}}{2} \end{bmatrix}$$

e temos

$$\mathbf{U}_0^*\mathbf{A}_0\mathbf{U}_0 = \begin{bmatrix} 2-\sqrt{3}\,i & 0 \\ 0 & 2+\sqrt{3}\,i \end{bmatrix}.$$

Para terminar, a matriz

$$\mathbf{U} = \mathbf{V}\begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_0 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} - \frac{\sqrt{6}}{6}\,i & -\frac{\sqrt{6}}{6} + \frac{\sqrt{6}}{6}\,i \\ \frac{\sqrt{3}}{3} & \frac{3\sqrt{2}-\sqrt{6}}{12} + \frac{3\sqrt{2}+\sqrt{6}}{12}\,i & -\frac{3\sqrt{2}+\sqrt{6}}{12} + \frac{3\sqrt{2}-\sqrt{6}}{12}\,i \\ \frac{\sqrt{3}}{3} & -\frac{3\sqrt{2}+\sqrt{6}}{12} - \frac{3\sqrt{2}-\sqrt{6}}{12}\,i & -\frac{3\sqrt{2}-\sqrt{6}}{12} - \frac{3\sqrt{2}+\sqrt{6}}{12}\,i \end{bmatrix}$$

é unitária e temos

$$\mathbf{U}^*\mathbf{A}\mathbf{U} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2-\sqrt{3}\,i & 0 \\ 0 & 0 & 2+\sqrt{3}\,i \end{bmatrix}.$$

RESOLUÇÃO DOS EXERCÍCIOS – FOLHA 4

4.1. É normal: $\mathbf{AA}^* = \begin{bmatrix} 10 & 6-6i \\ 6+6i & 24 \end{bmatrix} = \mathbf{A}^*\mathbf{A}$.

4.2. Sejam $\mathbf{u} \in \mathcal{R}(\mathbf{A})$ e $\mathbf{v} \in \mathcal{N}(\mathbf{A})$. Pelo Exercício 1.6, sabemos que $\mathcal{R}(\mathbf{A}) = \mathcal{R}(\mathbf{AA}^*)$, logo $\mathbf{u} \in \mathcal{R}(\mathbf{AA}^*)$ e, portanto, existe $\mathbf{w} \in \mathbb{C}^{n\times 1}$ tal que $\mathbf{u} = \mathbf{AA}^*\mathbf{w}$. Sendo assim,

$$\langle \mathbf{u}|\mathbf{v}\rangle = \mathbf{u}^*\mathbf{v} = (\mathbf{AA}^*\mathbf{w})^*\mathbf{v} = \mathbf{w}^*\mathbf{AA}^*\mathbf{v} = \mathbf{w}^*\mathbf{A}^*\mathbf{Av} = 0$$

(porque $\mathbf{AA}^* = \mathbf{A}^*\mathbf{A}$ e $\mathbf{Av} = \mathbf{0}$).

4.3. Como $\mathbf{A}$ é normal, também $\mathbf{A} - \lambda\mathbf{I}_n$ é normal: porque $\mathbf{AA}^* = \mathbf{A}^*\mathbf{A}$ e

- $(\mathbf{A} - \lambda\mathbf{I}_n)(\mathbf{A} - \lambda\mathbf{I}_n)^* = (\mathbf{A} - \lambda\mathbf{I}_n)(\mathbf{A}^* - \overline{\lambda}\mathbf{I}_n) = \mathbf{AA}^* - \overline{\lambda}\mathbf{A} - \lambda\mathbf{A}^* + |\lambda|^2\mathbf{I}_n$,
- $(\mathbf{A} - \lambda\mathbf{I}_n)^*(\mathbf{A} - \lambda\mathbf{I}_n) = (\mathbf{A}^* - \overline{\lambda}\mathbf{I}_n)(\mathbf{A} - \lambda\mathbf{I}_n) = \mathbf{A}^*\mathbf{A} - \lambda\mathbf{A}^* - \overline{\lambda}\mathbf{A} + |\lambda|^2\mathbf{I}_n$.

Seja $u \in \mathcal{N}(\mathbf{A} - \lambda\mathbf{I}_n)$; sendo assim, $(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{u} = \mathbf{0}$, ou seja, $\mathbf{Au} = \lambda\mathbf{u}$. Provemos que $u \in \mathcal{N}\big((\mathbf{A} - \lambda\mathbf{I}_n)^*\big) = \mathcal{N}(\mathbf{A}^* - \overline{\lambda}\mathbf{I}_n)$ (isto é, que $\mathbf{A}^*\mathbf{u} = \overline{\lambda}\mathbf{u}$). Ora, pelo exercício anterior, sabemos que $\langle \mathbf{u}|\mathbf{u}'\rangle = 0$ para todo $\mathbf{u}' \in \mathcal{R}(\mathbf{A} - \lambda\mathbf{I}_n)$ (porque $\mathbf{A} - \lambda\mathbf{I}_n$ é normal). Por conseguinte,

$$\mathbf{u}^*(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{w} = \langle \mathbf{u}|(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{w}\rangle = 0, \qquad \mathbf{w} \in \mathbb{C}^{n\times 1};$$

em particular, escolhendo $\mathbf{w} = (\mathbf{A} - \lambda\mathbf{I}_n)^*\mathbf{u}$, deduzimos que

$$0 = \mathbf{u}^*(\mathbf{A} - \lambda\mathbf{I}_n)(\mathbf{A} - \lambda\mathbf{I}_n)^*\mathbf{u} = \langle (\mathbf{A} - \lambda\mathbf{I}_n)^*\mathbf{u}|(\mathbf{A} - \lambda\mathbf{I}_n)^*\mathbf{u}\rangle$$

e, portanto, tem de ser $(\mathbf{A} - \lambda\mathbf{I}_n)^*\mathbf{u} = \mathbf{0}$, como se queria.

Finalmente, para qualquer $\mathbf{v} \in \mathcal{N}(\mathbf{A} - \mu\mathbf{I}_n)$, obtemos

$$\begin{aligned}(\lambda - \mu)\langle \mathbf{u}|\mathbf{v}\rangle &= \lambda\langle \mathbf{u}|\mathbf{v}\rangle - \mu\langle \mathbf{u}|\mathbf{v}\rangle = \langle \overline{\lambda}\mathbf{u}|\mathbf{v}\rangle - \langle \mathbf{u}|\mu\mathbf{v}\rangle \\ &= \langle \mathbf{A}^*\mathbf{u}|\mathbf{v}\rangle - \langle \mathbf{u}|\mathbf{Av}\rangle = \mathbf{u}^*\mathbf{Av} - \mathbf{u}^*\mathbf{Av} = 0\end{aligned}$$

e, portanto, $\langle \mathbf{u}|\mathbf{v}\rangle = 0$ (porque $\lambda \neq \mu$).

4.4. Suponhamos que $\mathbf{A}^\mathsf{T} = \mathbf{A}$. Então,

- $\mathbf{AA}^\mathsf{T} = \mathbf{A}^2 = \mathbf{A}^*\mathbf{A}$, isto é, $\mathbf{A}$ é (como $\mathbf{A} \in \mathbb{R}^{n\times n}$, temos $\mathbf{A}^* = \mathbf{A}^\mathsf{T}$).
- $\mathbf{A}$ é hermítica, logo todos os seus valores próprios são reais (Lema 5.1).

Reciprocamente, suponhamos que $\mathbf{A}$ é normal e que $\sigma(\mathbf{A}) \subseteq \mathbb{R}$. Sejam $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ os valores próprios de $\mathbf{A}$ (com repetições). Pelo teorema da decomposição espectral, existe uma matriz unitária $\mathbf{U} \in \mathbb{C}^{n\times n}$ tal que

$$\mathbf{U}^*\mathbf{A}\mathbf{U} = \mathbf{D}, \qquad \mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Temos $\mathbf{D} \in \mathbb{R}^{n\times n}$, logo $\mathbf{D}^* = \mathbf{D}$; além disso, $\mathbf{U}\mathbf{U}^* = \mathbf{U}^*\mathbf{U} = \mathbf{I}_n$ (porque $\mathbf{U}$ é unitária), logo

$$\mathbf{A} = \mathbf{U}\mathbf{U}^*\mathbf{A}\mathbf{U}\mathbf{U}^* = \mathbf{U}\mathbf{D}\mathbf{U}^* = \mathbf{U}\mathbf{D}^*\mathbf{U}^* = (\mathbf{U}\mathbf{D}\mathbf{U}^*)^* = (\mathbf{U}\mathbf{U}^*\mathbf{A}\mathbf{U}\mathbf{U}^*)^* = \mathbf{A}^* = \mathbf{A}^{\mathsf{T}}.$$

4.5. Seja $\lambda \in \mathbb{C}$ um valor próprio de $\mathbf{A}$ e seja $\mathbf{v} \in \mathbb{C}^{n\times 1}$, $\mathbf{v} \neq \mathbf{0}$, tal que $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$. Como $\mathbf{A}^* = -\mathbf{A}$, temos

$$\begin{aligned} \langle \mathbf{v}|\mathbf{v}\rangle(\lambda + \overline{\lambda}) &= \lambda\langle \mathbf{v}|\mathbf{v}\rangle + \overline{\lambda}\langle \mathbf{v}|\mathbf{v}\rangle = \langle \mathbf{v}|\lambda\mathbf{v}\rangle + \langle \lambda\mathbf{v}|\mathbf{v}\rangle = \langle \mathbf{A}\mathbf{v}|\mathbf{v}\rangle + \langle \mathbf{v}|\mathbf{A}\mathbf{v}\rangle \\ &= \mathbf{v}^*\mathbf{A}^*\mathbf{v} + \mathbf{v}^*\mathbf{A}\mathbf{v} = -\mathbf{v}\mathbf{A}\mathbf{v} + \mathbf{v}\mathbf{A}\mathbf{v} = 0 \end{aligned}$$

e, portanto, $\lambda + \overline{\lambda} = 0$ (porque $\mathbf{v} \neq \mathbf{0}$, logo $\langle \mathbf{v}|\mathbf{v}\rangle = 0$). Por conseguinte, λ é imaginário puro (a sua parte real tem de ser zero).

4.6. Supomos que $\mathbf{T}$ é triangular superior.

$\mathbf{T}$ será normal se e só se $\mathbf{T}^*\mathbf{T} = \mathbf{T}\mathbf{T}^*$. Considerando a entrada $(1,1)$ desta matriz, obtemos a equação

$$\overline{t_{1,1}}t_{1,1} = t_{1,1}\overline{t_{1,1}} + t_{1,2}\overline{t_{1,2}} + \cdots + t_{1,n}\overline{t_{1,n}},$$

de onde resulta que

$$|t_{1,2}|^2 + \cdots + |t_{1,n}|^2 = 0$$

e, portanto, $t_{1,2} = \cdots = t_{1,n} = 0$. Deste modo, $\mathbf{T} = \begin{bmatrix} t_{1,1} & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_0 \end{bmatrix}$ onde $\mathbf{T}_0 \in \mathbb{C}^{(n-1)\times(n-1}$ é triangular superior; além disso, tem de ser $\mathbf{T}_0^*\mathbf{T}_0 = \mathbf{T}_0\mathbf{T}_0^*$ e, portanto, podemos usar indução para concluir que $\mathbf{T}_0$ é diagonal. Em conclusão, $\mathbf{T}$ é uma matriz diagonal.

O recíproco é claramente verdadeiro.

4.7. (a) Suponhamos que $\mathbf{A}$ é invertível. Nesta situação, tem de ser $\lambda \neq 0$ (caso contrário, $r(\mathbf{A}) \leqslant n$), logo

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \iff \mathbf{v} = \mathbf{A}^{-1}\mathbf{A}\mathbf{v} = \lambda\mathbf{A}^{-1}\mathbf{v} \iff \mathbf{A}^{-1}\mathbf{v} = \lambda^{-1}\mathbf{v}.$$

(b) Na resolução do Exercício 4.3, provámos que, se $\mathbf{A}$ for normal, então

$$\mathbf{v} \in \mathbb{N}(\mathbf{A} - \lambda\mathbf{I}_n) \iff \mathbf{v} \in \mathcal{N}(\mathbf{A}^* - \overline{\lambda}\mathbf{I}_n),$$

o que significa que $\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \iff \mathbf{A}^*\mathbf{v} = \overline{\lambda}\mathbf{v}$.

(c) $\mathbf{A}^k\mathbf{v} = \mathbf{A}^{k-1}(\mathbf{A}\mathbf{v}) = \lambda\mathbf{A}^{k-1}\mathbf{v}$, pelo que basta argumentar por indução sobre k.

(d) $p(\mathbf{A})\mathbf{v} = \alpha_0\mathbf{v} + \alpha_1\mathbf{A}\mathbf{v} + \cdots + \alpha_m\mathbf{A}^m\mathbf{v} = \alpha_0\mathbf{v} + \alpha_1\lambda\mathbf{v} + \cdots + \alpha_m\lambda^m\mathbf{v} = p(\lambda)\mathbf{v}$.

4.8. $(\mathbf{A}+\mathbf{E})^* = \mathbf{A}^* + \mathbf{E}^* = \mathbf{A}+\mathbf{E}$, logo $\mathbf{B} = \mathbf{A}+\mathbf{E}$ é hermítica.

Seja $\mathbf{U} \in \mathbb{C}^{n\times n}$ uma matriz unitária tal que

$$\mathbf{U}^*\mathbf{A}\mathbf{U} = \mathbf{D}, \qquad \mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix};$$

deste modo, temos

$$\mathbf{U}^*\mathbf{B}\mathbf{U} = \mathbf{U}^*(\mathbf{A}+\mathbf{E})\mathbf{U} = \mathbf{U}^*\mathbf{A}\mathbf{U} + \mathbf{U}^*\mathbf{E}\mathbf{U} = \mathbf{D} + \mathbf{U}^*\mathbf{E}\mathbf{U}.$$

Para simplificar, pomos $\tilde{\mathbf{B}} = \mathbf{U}^*\mathbf{B}\mathbf{U}$ e $\tilde{\mathbf{E}} = \mathbf{U}^*\mathbf{E}\mathbf{U}$; observamos que $\tilde{\mathbf{B}}$ e $\tilde{\mathbf{E}}$ têm os mesmos valores próprios que $\mathbf{B}$ e $\mathbf{E}$, respectivamente.

Seja $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ a base canónica de $\mathbb{C}^{n\times 1}$ e, para cada $1 \leqslant i \leqslant n$, sejam $\mathcal{V}_i = \langle \mathbf{e}_1, \ldots, \mathbf{e}_i \rangle$ e $\mathcal{S}_i = \{\mathbf{v} \in \mathcal{V}_i \colon \|\mathbf{v}\| = 1\}$; em particular, temos $\mathcal{V}_n = \mathbb{C}^{n\times 1}$ e $\mathcal{S}_n = \mathcal{S} = \{\mathbf{v} \in \mathbb{C}^{n\times 1} \colon \|\mathbf{v}\| = 1\}$. Notemos que

$$\mathbf{v}^*\mathbf{D}\mathbf{v} = \sum_{1\leqslant k\leqslant i} \lambda_k |v_k|^2 \geqslant \lambda_i \sum_{1\leqslant k\leqslant i} |v_k|^2 = \lambda_i\|\mathbf{v}\| = \lambda_i, \qquad \mathbf{v} \in \mathcal{S}_i.$$

Aplicando a condição max-min do teorema de Courant-Fischer à matriz $\tilde{\mathbf{B}}$, deduzimos que

$$\mu_i \geqslant \min_{\mathbf{v}\in\mathcal{S}_i} \mathbf{v}^*\tilde{\mathbf{B}}\mathbf{v} \geqslant \min_{\mathbf{v}\in\mathcal{S}_i} \mathbf{v}^*\mathbf{D}\mathbf{v} + \min_{\mathbf{v}\in\mathcal{S}_i} \mathbf{v}^*\tilde{\mathbf{E}}\mathbf{v} \geqslant \lambda_i + \min_{\mathbf{v}\in\mathcal{S}} \mathbf{v}^*\tilde{\mathbf{E}}\mathbf{v} = \lambda_i + \varepsilon_n$$

(usando a Proposição 5.2).

De modo análogo, para cada $1 \leqslant i \leqslant n$, sejam $\mathcal{W}_i = \langle \mathbf{e}_i, \ldots, \mathbf{e}_n \rangle$ e $\mathcal{T}_i = \{\mathbf{v} \in \mathcal{W}_i \colon \|\mathbf{v}\| = 1\}$. Temos que

$$\mathbf{v}^*\mathbf{D}\mathbf{v} = \sum_{i\leqslant k\leqslant n} \lambda_k |v_k|^2 \leqslant \lambda_i \sum_{i\leqslant k\leqslant n} |v_k|^2 = \lambda_i\|\mathbf{v}\| = \lambda_i, \qquad \mathbf{v} \in \mathcal{T}_i.$$

Aplicando a condição min-max do teorema de Courant-Fischer à matriz $\tilde{\mathbf{B}}$, deduzimos que

$$\mu_i \leqslant \max_{\mathbf{v}\in\mathcal{T}_i} \mathbf{v}^*\tilde{\mathbf{B}}\mathbf{v} \leqslant \max_{\mathbf{v}\in\mathcal{T}_i} \mathbf{v}^*\mathbf{D}\mathbf{v} + \max_{\mathbf{v}\in\mathcal{T}_i} \mathbf{v}^*\tilde{\mathbf{E}}\mathbf{v} \leqslant \lambda_i + \max_{\mathbf{v}\in\mathcal{S}} \mathbf{v}^*\tilde{\mathbf{E}}\mathbf{v} = \lambda_i + \varepsilon_1$$

(usando a Proposição 5.2).

4.9. Seja $\mathbf{U} \in \mathbb{C}^{n\times n}$ uma matriz unitária tal que

$$\mathbf{U}^*\mathbf{A}\mathbf{U} = \mathbf{D}, \qquad \mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Como a matriz $\mathbf{V} = \begin{bmatrix} \mathbf{V} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \in \mathbb{C}^{(n+1)\times(n+1)}$ é unitária, os valores próprios de $\mathbf{B}$ são os mesmos (com repetições) que os da matriz

$$\tilde{\mathbf{B}} = \mathbf{V}^*\mathbf{B}\mathbf{V} = \begin{bmatrix} \mathbf{U}^*\mathbf{A}\mathbf{U} & \mathbf{U}^*\mathbf{u} \\ \mathbf{u}^*\mathbf{U} & \alpha \end{bmatrix}.$$

Seja $\{\mathbf{e}_1, \ldots, \mathbf{e}_{n+1}\}$ a base canónica de $\mathbb{C}^{(n+1)\times 1}$ e, para $1 \leqslant i \leqslant n$, sejam $\mathcal{V}_i = \langle \mathbf{e}_1, \ldots, \mathbf{e}_i \rangle$ e $\mathcal{S}_i = \{\mathbf{v} \in \mathcal{V}_i \colon \|\mathbf{v}\| = 1\}$. Temos que

$$\mathbf{v}^*\tilde{\mathbf{B}}\mathbf{v} = \sum_{1\leqslant k\leqslant i} \lambda_k |v_k|^2 \geqslant \lambda_i \sum_{1\leqslant k\leqslant i} |v_k|^2 = \lambda_i \|\mathbf{v}\| = \lambda_i, \qquad \mathbf{v} \in \mathcal{S}_i.$$

Aplicando a condição max-min do teorema de Courant-Fischer à matriz $\tilde{\mathbf{B}}$, deduzimos que

$$\mu_i \geqslant \min_{\mathbf{v}\in\mathcal{S}_i} \mathbf{v}^*\tilde{\mathbf{B}}\mathbf{v} \geqslant \lambda_i.$$

De modo análogo, para cada $1 \leqslant i \leqslant n$, sejam $\mathcal{W}_i = \langle \mathbf{e}_{i-1}, \ldots, \mathbf{e}_n \rangle \subseteq \mathbb{C}^{(n+1)\times 1}$ e $\mathcal{T}_i = \{\mathbf{v} \in \mathcal{W}_i \colon \|\mathbf{v}\| = 1\}$. Temos que

$$\mathbf{v}^*\tilde{\mathbf{B}}\mathbf{v} = \sum_{i-1\leqslant k\leqslant n} \lambda_k |v_k|^2 \leqslant \lambda_{i-1} \sum_{i-1\leqslant k\leqslant n} |v_k|^2 = \lambda_{i-1} \|\mathbf{v}\| = \lambda_{i-1}, \qquad \mathbf{v} \in \mathcal{T}_i.$$

Aplicando a condição min-max do teorema de Courant-Fischer à matriz $\tilde{\mathbf{B}}$, deduzimos que

$$\mu_i \leqslant \max_{\mathbf{v}\in\mathcal{T}_i} \mathbf{v}^*\tilde{\mathbf{B}}\mathbf{v} \leqslant \lambda_{i-1}.$$

RESOLUÇÃO DOS EXERCÍCIOS – FOLHA 5

5.1. (a) Temos

$$\mathbf{I}_n = \mathbf{P}\mathbf{P}^{-1} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_r \end{bmatrix} \begin{bmatrix} \mathbf{Y}_1^{\mathsf{T}} \\ \mathbf{Y}_2^{\mathsf{T}} \\ \cdots \\ \mathbf{Y}_r^{\mathsf{T}} \end{bmatrix} = \sum_{1 \leqslant i \leqslant r} \mathbf{X}_i \mathbf{Y}_i^{\mathsf{T}} = \sum_{1 \leqslant i \leqslant n} \mathbf{G}_i.$$

Analogamente,

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_r \end{bmatrix} \begin{bmatrix} \lambda_1 \mathbf{I}_{m_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \lambda_2 \mathbf{I}_{m_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \lambda_r \mathbf{I}_{m_r} \end{bmatrix} \begin{bmatrix} \mathbf{Y}_1^{\mathsf{T}} \\ \mathbf{Y}_2^{\mathsf{T}} \\ \cdots \\ \mathbf{Y}_r^{\mathsf{T}} \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 \mathbf{X}_1 & \lambda_2 \mathbf{X}_2 & \cdots & \lambda_r \mathbf{X}_r \end{bmatrix} \begin{bmatrix} \mathbf{Y}_1^{\mathsf{T}} \\ \mathbf{Y}_2^{\mathsf{T}} \\ \cdots \\ \mathbf{Y}_r^{\mathsf{T}} \end{bmatrix} = \sum_{1 \leqslant i \leqslant r} \lambda_i \mathbf{X}_i \mathbf{Y}_i^{\mathsf{T}} = \sum_{1 \leqslant i \leqslant r} \lambda_i \mathbf{G}_i. \end{aligned}$$

(b) Temos

$$\mathbf{I}_n = \mathbf{P}^{-1}\mathbf{P} = \begin{bmatrix} \mathbf{Y}_1^{\mathsf{T}} \\ \mathbf{Y}_2^{\mathsf{T}} \\ \cdots \\ \mathbf{Y}_r^{\mathsf{T}} \end{bmatrix} \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_r \end{bmatrix} = \begin{bmatrix} \mathbf{Y}_1^{\mathsf{T}}\mathbf{X}_1 & \mathbf{Y}_1^{\mathsf{T}}\mathbf{X}_2 & \cdots & \mathbf{Y}_1^{\mathsf{T}}\mathbf{X}_r \\ \mathbf{Y}_2^{\mathsf{T}}\mathbf{X}_1 & \mathbf{Y}_2^{\mathsf{T}}\mathbf{X}_2 & \cdots & \mathbf{Y}_2^{\mathsf{T}}\mathbf{X}_r \\ \vdots & \vdots & & \vdots \\ \mathbf{Y}_r^{\mathsf{T}}\mathbf{X}_1 & \mathbf{Y}_r^{\mathsf{T}}\mathbf{X}_2 & \cdots & \mathbf{Y}_r^{\mathsf{T}}\mathbf{X}_r \end{bmatrix},$$

logo

$$\mathbf{Y}_i^{\mathsf{T}}\mathbf{X}_j = \begin{cases} \mathbf{I}_{m_i}, & \text{se } i = j, \\ \mathbf{0}, & \text{se } i \neq j, \end{cases} \qquad 1 \leqslant i, j \leqslant r.$$

Daqui, obtemos

$$\mathbf{G}_i\mathbf{G}_j = \mathbf{X}_i\mathbf{Y}_i^{\mathsf{T}}\mathbf{X}_j\mathbf{Y}_j^{\mathsf{T}} = \begin{cases} \mathbf{X}_i\mathbf{Y}_i^{\mathsf{T}} = \mathbf{G}_i, & \text{se } i = j, \\ \mathbf{0}, & \text{se } i \neq j, \end{cases} \qquad 1 \leqslant i, j \leqslant r.$$

(c) Seja $1 \leqslant i \leqslant r$. Temos

$$\mathcal{R}(\mathbf{G}_i) = \mathcal{R}(\mathbf{X}_i\mathbf{Y}_i^{\mathsf{T}}) \subseteq \mathcal{R}(\mathbf{X}_i) = \mathcal{R}(\mathbf{X}_i\mathbf{Y}_i^{\mathsf{T}}\mathbf{X}_i) = \mathcal{R}(\mathbf{G}_i\mathbf{X}_i) \subseteq \mathcal{R}(\mathbf{G}_i)$$

(porque $\mathbf{Y}_i^{\mathsf{T}}\mathbf{X}_i = \mathbf{I}_{m_i}$), pelo que

$$\mathcal{R}(\mathbf{G}_i) = \mathcal{R}(\mathbf{X}_i) = \mathcal{N}(\mathbf{A} - \lambda_i\mathbf{I}_n)$$

(por definição de $\mathcal{R}(\mathbf{X}_i)$ e de $\mathbf{X}_i$, temos $\mathcal{R}(\mathbf{X}_i) = \mathcal{N}(\mathbf{A} - \lambda_i\mathbf{I}_n)$).

Por outro lado, usando as alíneas anteriores, deduzimos que

$$\begin{aligned}
\mathbf{G}_i(\mathbf{A} - \lambda_i\mathbf{I}_n) &= \mathbf{G}_i\Big(\sum_{1\leqslant k\leqslant r}\lambda_k\mathbf{G}_k - \lambda_i\sum_{1\leqslant k\leqslant r}\mathbf{G}_k\Big)\\
&= \sum_{1\leqslant k\leqslant r}\lambda_k\mathbf{G}_i\mathbf{G}_k - \lambda_i\sum_{1\leqslant k\leqslant r}\mathbf{G}_i\mathbf{G}_k\\
&= \lambda_i\mathbf{G}_i - \lambda_i\mathbf{G}_i = \mathbf{0},
\end{aligned}$$

o que significa que qualquer coluna da matriz $\mathbf{A} - \lambda_i\mathbf{I}_n$ está em $\mathcal{N}(\mathbf{G}_i)$, logo

$$\mathcal{R}(\mathbf{A} - \lambda_i\mathbf{I}_n) \subseteq \mathcal{N}(\mathbf{G}_i).$$

Para a inclusão recíproca, como $\mathcal{R}(\mathbf{G}_i) = \mathcal{N}(\mathbf{A} - \lambda_i\mathbf{I}_n)$, temos $r(\mathbf{G}_i) = n(\mathbf{A} - \lambda_i\mathbf{I}_n)$, logo

$$\dim\mathcal{R}(\mathbf{A} - \lambda_i\mathbf{I}_n) = r(\mathbf{A} - \lambda_i\mathbf{I}_n) = n - n(\mathbf{A} - \lambda_i\mathbf{I}_n) = n - r(\mathbf{G}_i) = n(\mathbf{G}_i) = \dim\mathcal{N}(\mathbf{G}_i)$$

e, portanto, $\mathcal{R}(\mathbf{A} - \lambda_i\mathbf{I}_n) = \mathcal{N}(\mathbf{G}_i)$.

5.2. • $\lambda \in \mathbb{C}$ é valor próprio de $\mathbf{A} = \begin{bmatrix} 1 & -4 & -4\\ 8 & -11 & -8\\ -8 & 8 & 5\end{bmatrix}$ se e só se

$$\det(\lambda\mathbf{I}_3 - \mathbf{A}) = \lambda^3 + 5\lambda^2 + 3\lambda - 9 = (\lambda - 1)(\lambda + 3)^2 = 0.$$

Por conseguinte, $\sigma(\mathbf{A}) = \{-3, 1\}$ com m.a.$(-3) = 2$ e m.a.$(1) = 1$. Por outro lado,

$$\left\{\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}\right\} \qquad \text{e} \qquad \left\{\begin{bmatrix}1\\2\\-2\end{bmatrix}\right\}$$

são bases de $\mathcal{N}(\mathbf{A} + 3\mathbf{I}_3)$ e $\mathcal{N}(\mathbf{A} - \mathbf{I}_3)$, de modo que

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & 1\\ 1 & 0 & 2\\ 0 & 1 & -2\end{bmatrix}$$

é uma matriz invertível tal que

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} -3 & 0 & 0\\ 0 & -3 & 0\\ 0 & 0 & 1\end{bmatrix} = \begin{bmatrix} -3\mathbf{I}_2 & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_1\end{bmatrix}.$$

Pondo $\mathbf{X}_1 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$ e $\mathbf{X}_2 = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$, temos $\mathbf{P} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 \end{bmatrix}$. Por outro lado,

$$\mathbf{P}^{-1} = \begin{bmatrix} -2 & 3 & 2 \\ 2 & -2 & -1 \\ 1 & -1 & -1 \end{bmatrix} = \begin{bmatrix} \mathbf{Y}_1^{\mathtt{T}} \\ \mathbf{Y}_2^{\mathtt{T}} \end{bmatrix}$$

onde $\mathbf{Y}_1^{\mathtt{T}} = \begin{bmatrix} -2 & 3 & 2 \\ 2 & -2 & -1 \end{bmatrix}$ e $\mathbf{Y}_2^{\mathtt{T}} = \begin{bmatrix} 1 & -1 & -1 \end{bmatrix}$. Sendo assim, os projectores espectrais de $\mathbf{A}$ são

- $\mathbf{G}_1 = \mathbf{X}_1\mathbf{Y}_1^{\mathtt{T}} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -2 & 3 & 2 \\ 2 & -2 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ -2 & 3 & 2 \\ 2 & -2 & -1 \end{bmatrix}$,
- $\mathbf{G}_2 = \mathbf{X}_2\mathbf{Y}_2^{\mathtt{T}} = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 \\ 2 & -2 & -2 \\ -2 & 2 & 2 \end{bmatrix}$.

- $\lambda \in \mathbb{C}$ é valor próprio de $\mathbf{A} = \begin{bmatrix} 3 & 2 & 1 \\ 0 & 2 & 0 \\ -2 & -3 & 0 \end{bmatrix}$ se e só se

$$\det(\lambda\mathbf{I}_3 - \mathbf{A}) = (\lambda - 1)(\lambda - 2)^2 = 0.$$

Por conseguinte, $\sigma(\mathbf{A}) = \{1, 2\}$ com m.a.$(1) = 1$ e m.a.$(2) = 2$. $\mathcal{N}(\mathbf{A} - 2\mathbf{I}_2)$ tem dimensão 1 (com base $\left\{\begin{bmatrix} 1 & 0 & -1 \end{bmatrix}^{\mathtt{T}}\right\}$, logo $\mathbf{A}$ não é diagonalizável (porque não existe uma base de $\mathbb{C}^{n\times n}$ inteiramente constituída por vectores próprios de $\mathbf{A}$).

- $\lambda \in \mathbb{C}$ é valor próprio de $\mathbf{A} = \begin{bmatrix} -4 & -3 & -3 \\ 0 & -1 & 0 \\ 6 & 6 & 5 \end{bmatrix}$ se e só se

$$\det(\lambda\mathbf{I}_3 - \mathbf{A}) = (\lambda - 2)(\lambda + 1)^2 = 0.$$

Por conseguinte, $\sigma(\mathbf{A}) = \{-1, 2\}$ com m.a.$(-1) = 2$ e m.a.$(2) = 1$. Por outro lado,

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\} \qquad \text{e} \qquad \left\{ \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} \right\}$$

são bases de $\mathcal{N}(\mathbf{A} + 3\mathbf{I}_3)$ e $\mathcal{N}(\mathbf{A} - \mathbf{I}_3)$, de modo que

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & -1 & -2 \end{bmatrix}$$

é uma matriz invertível tal que

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} -\mathbf{I}_2 & \mathbf{0} \\ \mathbf{0} & 2\mathbf{I}_1 \end{bmatrix}.$$

Pondo $\mathbf{X}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{bmatrix}$ e $\mathbf{X}_2 = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$, temos $\mathbf{P} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 \end{bmatrix}$. Por outro lado,

$$\mathbf{P}^{-1} = \begin{bmatrix} 5/2 & -3/2 & -3/2 \\ 0 & -1 & 0 \\ 3 & 3 & 2 \end{bmatrix} = \begin{bmatrix} \mathbf{Y}_1^{\mathrm{T}} \\ \mathbf{Y}_2^{\mathrm{T}} \end{bmatrix}$$

onde $\mathbf{Y}_1^{\mathrm{T}} = \begin{bmatrix} 5/2 & -3/2 & -3/2 \\ 0 & -1 & 0 \end{bmatrix}$ e $\mathbf{Y}_2^{\mathrm{T}} = \begin{bmatrix} 3 & 3 & 2 \end{bmatrix}$. Sendo assim, os projectores espectrais de $\mathbf{A}$ são

- $\mathbf{G}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 5/2 & -3/2 & -3/2 \\ 0 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 5/2 & -3/2 & -3/2 \\ 0 & -1 & 0 \\ -5/2 & 5/2 & 3/2 \end{bmatrix}$,
- $\mathbf{G}_2 = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} \begin{bmatrix} 3 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 2 \\ 0 & 0 & 0 \\ -6 & -6 & -4 \end{bmatrix}$.

- Vimos no Exercício 3.4 que $\mathbf{A} = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & -2 \\ -2 & 0 & 1 \end{bmatrix}$ é diagonalizável com $\sigma(\mathbf{A}) = \{-1, 2 - \sqrt{3}\,i, 2 + \sqrt{3}\,i\}$. Sejam $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{C}^{3\times 1}$ vectores próprios associados a -1, $2 - \sqrt{3}\,i$ e $2 + \sqrt{3}\,i$, respectivamente; de modo que $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ é uma base de $\mathbb{C}^{3\times 1}$. Considerando a matriz $\mathbf{P} = \begin{bmatrix} \mathbf{u} & \mathbf{v} & \mathbf{w} \end{bmatrix}$, temos

$$\mathbf{P}^* = \begin{bmatrix} \mathbf{u}^* \\ \mathbf{v}^* \\ \mathbf{w}^* \end{bmatrix} \qquad \text{e} \qquad \mathbf{P}^*\mathbf{P} = \begin{bmatrix} \mathbf{u}^*\mathbf{u} & \mathbf{u}^*\mathbf{v} & \mathbf{u}^*\mathbf{w} \\ \mathbf{v}^*\mathbf{u} & \mathbf{v}^*\mathbf{v} & \mathbf{v}^*\mathbf{w} \\ \mathbf{w}^*\mathbf{u} & \mathbf{w}^*\mathbf{v} & \mathbf{w}^*\mathbf{w} \end{bmatrix} = \begin{bmatrix} \|\mathbf{u}\|^2 & 0 & 0 \\ 0 & \|\mathbf{v}\|^2 & 0 \\ 0 & 0 & \|\mathbf{w}\|^2 \end{bmatrix}$$

(usando o Exercício 4.3), logo

$$\mathbf{P}^{-1} = \begin{bmatrix} \frac{1}{\|\mathbf{u}\|^2} & 0 & 0 \\ 0 & \frac{1}{\|\mathbf{v}\|^2} & 0 \\ 0 & 0 & \frac{1}{\|\mathbf{w}\|^2} \end{bmatrix} \mathbf{P}^* = \begin{bmatrix} \frac{1}{\|\mathbf{u}\|^2}\,\mathbf{u}^* \\ \frac{1}{\|\mathbf{u}\|^2}\,\mathbf{v}^* \\ \frac{1}{\|\mathbf{u}\|^2}\,\mathbf{w}^* \end{bmatrix}.$$

Por conseguinte, os projectores espectrais de $\mathbf{A}$ são

$$\mathbf{G}_1 = \frac{1}{\|\mathbf{u}\|^2}\,\mathbf{u}\mathbf{u}^*, \quad \mathbf{G}_2 = \frac{1}{\|\mathbf{v}\|^2}\,\mathbf{v}\mathbf{v}^* \quad \text{e} \quad \mathbf{G}_3 = \frac{1}{\|\mathbf{w}\|^2}\,\mathbf{w}\mathbf{w}^*.$$

Como $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ é um vector próprio associado a -1, obtemos

$$\mathbf{G}_1 = \frac{1}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Agora, sabemos que

$$\mathbf{A} = -\mathbf{G}_1 + \alpha\mathbf{G}_2 + \overline{\alpha}\mathbf{G}_3 \qquad \mathbf{I}_3 = \mathbf{G}_1 + \mathbf{G}_2 + \mathbf{G}_3$$

onde $\alpha = 2 - \sqrt{3}\,i$. Sendo assim, temos

$$\alpha\mathbf{G}_2 + \overline{\alpha}\mathbf{G}_3 = \frac{1}{3}\begin{bmatrix} 4 & -5 & 1 \\ 1 & 4 & -5 \\ -5 & 1 & 4 \end{bmatrix} \qquad \text{e} \qquad \mathbf{G}_2 + \mathbf{G}_3 = \frac{1}{3}\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix},$$

de onde resulta que

$$\mathbf{G}_2 = -\frac{1}{6}\begin{bmatrix} -2 & 3-\sqrt{3}\,i & 3+\sqrt{3}\,i \\ 3+\sqrt{3}\,i & -2 & 3+\sqrt{3}\,i \\ 3-\sqrt{3}\,i & 3+\sqrt{3}\,i & -2 \end{bmatrix} \quad \text{e} \quad \mathbf{G}_3 = \frac{1}{6}\begin{bmatrix} 2 & 1-\sqrt{3}\,i & 1+\sqrt{3}\,i \\ 1+\sqrt{3}\,i & 2 & 1+\sqrt{3}\,i \\ 1-\sqrt{3}\,i & 1+\sqrt{3}\,i & 2 \end{bmatrix}.$$

5.3. Sejam $a_1, \ldots, a_n \in \mathbb{k}$ tais que $p_{\mathbf{A}}(t) = t^n + a_1 t^{n-1} + \cdots + a_1 t + a_0$. Como

$$\mathbf{G}_i\mathbf{G}_j = \begin{cases} \mathbf{G}_i, & \text{se } i = j, \\ \mathbf{0}, & \text{se } i \neq j, \end{cases} \qquad 1 \leqslant i, j \leqslant r,$$

temos

$$\Big(\sum_{1\leqslant i\leqslant n} \lambda_i \mathbf{G}_i\Big)^k = \sum_{1\leqslant i\leqslant n} \lambda_i^k \mathbf{G}_i, \qquad 1 \leqslant k \leqslant n.$$

Por conseguinte, pondo $a_0 = 1$ e $\mathbf{A}^0 = \mathbf{I}_n$, deduzimos que

$$\begin{aligned} p_{\mathbf{A}}(\mathbf{A}) &= \sum_{0\leqslant k\leqslant n} a_k \mathbf{A}^k = \sum_{0\leqslant k\leqslant n} \Big(\sum_{1\leqslant i\leqslant r} \lambda_i \mathbf{G}_i\Big)^k = \sum_{1\leqslant k\leqslant n}\sum_{1\leqslant i\leqslant r} \lambda_i^k \mathbf{G}_i \\ &= \sum_{1\leqslant i\leqslant n} \Big(\sum_{0\leqslant k\leqslant n} a_k \lambda_i^k\Big)\mathbf{G}_i = \sum_{1\leqslant i\leqslant r} p_{\mathbf{A}}(\lambda_i)\mathbf{G}_i = \mathbf{0}. \end{aligned}$$

5.4. Escolhendo uma matriz invertível $\mathbf{P} \in \mathbb{C}^{n\times n}$ tal que

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}, \qquad \mathbf{A} = \begin{bmatrix} 1 & -4 & -4 \\ 8 & -11 & -8 \\ -8 & 8 & 5 \end{bmatrix}, \quad \mathbf{D} = \mathbf{P}\begin{bmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

temos $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$, logo

$$\begin{aligned} p_{\mathbf{A}}(\mathbf{A}) &= \mathbf{A}^3 + 5\mathbf{A}^2 + 3\mathbf{A} - 9\mathbf{I}_3 \\ &= (\mathbf{P}\mathbf{D}\mathbf{P}^{-1})^3 + 5(\mathbf{P}\mathbf{D}\mathbf{P}^{-1})^2 + 3(\mathbf{P}\mathbf{D}\mathbf{P}^{-1}) - 9\mathbf{I}_3 \\ &= \mathbf{P}\big(\mathbf{D}^3 + 5\mathbf{D}^2 + 3\mathbf{D} - 9\mathbf{I}_3\big)\mathbf{P}^{-1} = \mathbf{P}p_{\mathbf{A}}(\mathbf{D})\mathbf{P}^{-1}. \end{aligned}$$

Ora,

$$\mathbf{D}^k = \begin{bmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{bmatrix}^k = \begin{bmatrix} (-3)^k & 0 & 0 \\ 0 & (-3)^k & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

logo

$$p_{\mathbf{A}}(\mathbf{D}) = \begin{bmatrix} p_{\mathbf{A}}(-3) & 0 & 0 \\ 0 & p_{\mathbf{A}}(-3) & 0 \\ 0 & 0 & p_{\mathbf{A}}(1) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

e, portanto,

$$p_{\mathbf{A}}(\mathbf{A}) = \mathbf{P}p_{\mathbf{A}}(\mathbf{D})\mathbf{P}^{-1} = \mathbf{P}\mathbf{0}\mathbf{P}^{-1} = \mathbf{0}.$$

[Para as outras matrizes é análogo.]

5.5. Temos

$$\begin{aligned} \mathbf{P}^{-1}\mathbf{A}\mathbf{P} &= \mathbf{P}^{-1}\begin{bmatrix} \mathbf{A}\mathbf{v}_1 & \mathbf{A}\mathbf{v}_2 & \cdots & \mathbf{A}\mathbf{v}_n \end{bmatrix} = \mathbf{P}^{-1}\begin{bmatrix} \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & \cdots & \lambda_n\mathbf{v}_n \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{w}_1^{\mathsf{T}} \\ \mathbf{w}_2^{\mathsf{T}} \\ \vdots \\ \mathbf{w}_n^{\mathsf{T}} \end{bmatrix} \begin{bmatrix} \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & \cdots & \lambda_n\mathbf{v}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{w}_1^{\mathsf{T}}\mathbf{v}_1 & \lambda_2\mathbf{w}_1^{\mathsf{T}}\mathbf{v}_2 & \cdots & \lambda_n\mathbf{w}_1^{\mathsf{T}}\mathbf{v}_n \\ \lambda_1\mathbf{w}_2^{\mathsf{T}}\mathbf{v}_1 & \lambda_2\mathbf{w}_2^{\mathsf{T}}\mathbf{v}_2 & \cdots & \lambda_n\mathbf{w}_2^{\mathsf{T}}\mathbf{v}_n \\ \vdots & \vdots & & \vdots \\ \lambda_1\mathbf{w}_n^{\mathsf{T}}\mathbf{v}_1 & \lambda_2\mathbf{w}_n^{\mathsf{T}}\mathbf{v}_2 & \cdots & \lambda_n\mathbf{w}_n^{\mathsf{T}}\mathbf{v}_n \end{bmatrix}. \end{aligned}$$

Dado que

$$\mathbf{I}_n = \mathbf{P}^{-1}\mathbf{P} = \begin{bmatrix} \mathbf{w}_1^{\mathsf{T}} \\ \mathbf{w}_2^{\mathsf{T}} \\ \vdots \\ \mathbf{w}_n^{\mathsf{T}} \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix} = \begin{bmatrix} \mathbf{w}_1^{\mathsf{T}}\mathbf{v}_1 & \mathbf{w}_1^{\mathsf{T}}\mathbf{v}_2 & \cdots & \mathbf{w}_1^{\mathsf{T}}\mathbf{v}_n \\ \mathbf{w}_2^{\mathsf{T}}\mathbf{v}_1 & \mathbf{w}_2^{\mathsf{T}}\mathbf{v}_2 & \cdots & \mathbf{w}_2^{\mathsf{T}}\mathbf{v}_n \\ \vdots & \vdots & & \vdots \\ \mathbf{w}_n^{\mathsf{T}}\mathbf{v}_1 & \mathbf{w}_n^{\mathsf{T}}\mathbf{v}_2 & \cdots & \mathbf{w}_n^{\mathsf{T}}\mathbf{v}_n \end{bmatrix},$$

temos

$$\mathbf{w}_i^{\mathsf{T}}\mathbf{v}_j = \begin{cases} 1, & \text{se } i = j, \\ 0, & \text{se } i \neq j, \end{cases} \qquad 1 \leqslant i, j \leqslant r,$$

e, portanto,

$$\mathbf{P}^{-1}\mathbf{A} = (\mathbf{P}^{-1}\mathbf{A}\mathbf{P})\mathbf{P}^{-1} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{w}_1^{\mathsf{T}} \\ \mathbf{w}_2^{\mathsf{T}} \\ \vdots \\ \mathbf{w}_n^{\mathsf{T}} \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{w}_1^{\mathsf{T}} \\ \lambda_2\mathbf{w}_2^{\mathsf{T}} \\ \vdots \\ \lambda_n\mathbf{w}_n^{\mathsf{T}} \end{bmatrix}.$$

Como

$$\mathbf{P}^{-1}\mathbf{A} = \begin{bmatrix} \mathbf{w}_1^{\mathsf{T}} \\ \mathbf{w}_2^{\mathsf{T}} \\ \vdots \\ \mathbf{w}_n^{\mathsf{T}} \end{bmatrix} \mathbf{A} = \begin{bmatrix} \mathbf{w}_1^{\mathsf{T}}\mathbf{A} \\ \mathbf{w}_2^{\mathsf{T}}\mathbf{A} \\ \vdots \\ \mathbf{w}_n^{\mathsf{T}}\mathbf{A} \end{bmatrix},$$

concluímos que

$$\mathbf{A}^{\mathsf{T}}\mathbf{w}_i = (\mathbf{w}_i^{\mathsf{T}}\mathbf{A})^{\mathsf{T}} = (\lambda_i \mathbf{w}_i^{\mathsf{T}})^{\mathsf{T}} = \lambda_i \mathbf{w}_i, \qquad 1 \leqslant i \leqslant r.$$

5.6. $p_{\mathbf{A}}(\mathbf{A}) = \mathbf{0}$ é uma matriz de tipo $n \times n$ e $\det(\mathbf{0}) = 0$ é um escalar.

5.7. • $\mathbf{A} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix}$ tem valores próprios 1 e $\frac{1}{4}$ que estão associados aos vectores próprios $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ e $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$, respectivamente. Assim, pondo $\mathbf{P} = \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}$, temos

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{bmatrix}$$

e, portanto,

$$\mathbf{P}^{-1}\mathbf{A}^k\mathbf{P} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{bmatrix}^k = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{4^k} \end{bmatrix}, \qquad k \in \mathbb{N},$$

donde

$$\mathbf{A}^k = \mathbf{P} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{4^k} \end{bmatrix} \mathbf{P}^{-1} = \frac{1}{3} \begin{bmatrix} 1 + \frac{2}{4^k} & 2 - \frac{2}{4^k} \\ 1 - \frac{1}{4^k} & 2 + \frac{1}{4^k} \end{bmatrix}.$$

Como as sucessão (de números reais) $\left(\frac{1}{4^k}\right)_{k\in\mathbb{N}}$ é convergente com $\lim_{k\to\infty} \frac{1}{4^k}$, concluímos que a sucessão (de matrizes) $\left(\mathbf{A}^k\right)_{k\in\mathbb{N}}$ também é convergente com

$$\lim_{k\to\infty} \mathbf{A}^k = \frac{1}{3} \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}.$$

• Seguimos um processo diferente (usando os projectores espectrais). $\mathbf{A} = \begin{bmatrix} \frac{7}{5} & \frac{1}{5} \\ -1 & \frac{1}{2} \end{bmatrix}$ tem valores próprios 1 e $\frac{9}{10}$, de modo que, se $\mathbf{G}_1$ e $\mathbf{G}_2$ forem os projectores espectrais de $\mathbf{A}$, então

$$\mathbf{A} = \mathbf{G}_1 + (9/10)\mathbf{G}_2 \qquad \text{e} \qquad \mathbf{I}_2 = \mathbf{G}_1 + \mathbf{G}_2.$$

Temos

$$\mathbf{A} = (\mathbf{G}_1 + \frac{9}{10}\mathbf{G}_2)^k = \mathbf{G}_1 + (9/10)^k\mathbf{G}_2$$

e, portanto, a sucessão $\left(\mathbf{A}^k\right)_{k\in\mathbb{N}}$ é convergente com

$$\lim_{k\to\infty} = \mathbf{G}_1$$

(porque a sucessão $\left((9/10)^k\right)_{k\in\mathbb{N}}$ é convergente com $\lim_{k\to\infty}(9/10)^k = 0$).

Determinemos $\mathbf{G}_1$. Ora, das equações acima, temos

$$10\mathbf{A} - 9\mathbf{I}_2 = 10\big(\mathbf{G}_1 + (9/10)\mathbf{G}_2\big) - 9\big(\mathbf{G}_1 + \mathbf{G}_2\big) = \mathbf{G}_1,$$

logo

$$\lim_{k\to\infty} \mathbf{A}^k = \mathbf{G}_1 = \begin{bmatrix} 5 & 2 \\ -10 & -4 \end{bmatrix}.$$

5.8. Suponhamos que, no início do processo, o Norte tem n_0 habitantes e o Sul tem s_0 habitantes. Passado um ano, o Norte terá $n_1 = (1-1/2)n_0 + (1/4)s_0 = (1/2)n_0 + (1/4)s_0$ habitantes e o Sul terá $n_1 = (1/2)n_0 + (1-1/4)s_0 = (1/2)n_0 + (3/4)s_0$ habitantes. Dado que o processo é uniforme ao longo do tempo, no k-ésimo ano (para $k \in \mathbb{N}$), o Norte terá $n_k = (1/2)n_{k-1} + (1/4)s_{k-1}$ habitantes e o Sul terá $n_k = (1/2)n_{k-1} + (3/4)s_{k-1}$ habitantes, ou seja, temos

$$\begin{bmatrix} n_k \\ s_k \end{bmatrix} = \begin{bmatrix} 1/2 & 1/4 \\ 1/2 & 3/4 \end{bmatrix} \begin{bmatrix} n_{k-1} \\ s_{k-1} \end{bmatrix}.$$

Pondo $\mathbf{v}_k = \begin{bmatrix} n_k \\ s_k \end{bmatrix}$ (para $k \in \mathbb{N}_0$) e $\mathbf{A} = \begin{bmatrix} 1/2 & 1/4 \\ 1/2 & 3/4 \end{bmatrix}$, temos

$$\mathbf{v}_k = \mathbf{A}\mathbf{v}_{k-1} = \mathbf{A}^2\mathbf{v}_{k-2} = \cdots = \mathbf{A}^{k-1}\mathbf{v}_1 = \mathbf{A}^k\mathbf{v}_0, \qquad k \in \mathbb{N}.$$

Pelo que vimos no exercício anterior, a sucessão $\big((\mathbf{A}^{\mathsf{T}})^k\big)_{k\in\mathbb{N}}$ é convergente com $\lim_{k\to\infty}(\mathbf{A}^{\mathsf{T}})^k = \begin{bmatrix} 1/3 & 2/3 \\ 1/3 & 2/3 \end{bmatrix}$, de modo que a sucessão $\big(\mathbf{A}^k\big)_{k\in\mathbb{N}}$ também é convergente com

$$\lim_{k\to\infty} \mathbf{A}^k = \begin{bmatrix} 1/3 & 2/3 \\ 1/3 & 2/3 \end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} 1/3 & 1/3 \\ 2/3 & 2/3 \end{bmatrix}.$$

Deste modo, a sucessão $\big(\mathbf{v}^k\big)_{k\in\mathbb{N}}$ é convergente, o que significa que as populações do Norte e do Sul tendem a estabilizar no limite

$$\mathbf{v}_\infty = \lim_{k\to\infty} \mathbf{v}_k = \lim_{k\to\infty} \mathbf{A}^k\mathbf{v}_0 = \big(\lim_{k\to\infty} \mathbf{A}^k\big)\mathbf{v}_0 = \begin{bmatrix} 1/3 & 1/3 \\ 2/3 & 2/3 \end{bmatrix} \begin{bmatrix} n_0 \\ s_0 \end{bmatrix} = \begin{bmatrix} (1/3)(n_0 + s_0) \\ (2/3)(n_0 + s_0) \end{bmatrix}.$$

5.9. (a) Seja $\mathbf{v} \in \mathbb{C}^{n\times 1}$ com $\|\mathbf{v}\| = 1$ e tal que $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$. Deste modo, pondo $\mathbf{V} = \begin{bmatrix} \mathbf{v} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}$, temos

$$\mathbf{A}\mathbf{V} = \begin{bmatrix} \mathbf{A}\mathbf{v} & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} \lambda\mathbf{v} & 0 & \cdots & 0 \end{bmatrix} = \lambda \begin{bmatrix} \mathbf{v} & 0 & \cdots & 0 \end{bmatrix} = \lambda\mathbf{V}$$

e, portanto,

$$\begin{aligned} \|\mathbf{A}\mathbf{V}\| &= \sum_{1\leqslant i,l\leqslant n} \Big| \sum_{1\leqslant k\leqslant n} a_{i,k}v_{k,l} \Big|^2 \leqslant \sum_{1\leqslant i,l\leqslant n} \sum_{1\leqslant k\leqslant n} \big|a_{i,k}v_{k,l}\big|^2 \\ &= \sum_{1\leqslant i,k,l\leqslant n} \big|a_{i,k}\big|^2 \big|v_{k,l}\big|^2 \leqslant \sum_{1\leqslant i,j,k,l\leqslant n} \big|a_{i,j}\big|^2 \big|v_{k,l}\big|^2 \end{aligned}$$

$$= \Big(\sum_{1\leqslant i,j\leqslant n} \big|a_{i,j}\big|^2\Big)\Big(\sum_{1\leqslant k,l\leqslant n} \big|v_{k,l}\big|^2\Big) = \|\mathbf{A}\|\,\|\mathbf{V}\|.$$

Como $\|\mathbf{AV}\| = \|\lambda\mathbf{V}\| = |\lambda|\,\|\mathbf{V}\|$, concluímos que

$$|\lambda|\,\|\mathbf{V}\| \leqslant \|\mathbf{A}\|\,\|\mathbf{V}\|$$

e, portanto, $|\lambda| \leqslant \|\mathbf{A}\|$ (porque $\mathbf{v} \neq \mathbf{0}$, logo $\mathbf{V} \neq \mathbf{0}$). Como $\lambda \in \sigma(\mathbf{A})$ é arbitrário, segue-se que

$$\rho(\mathbf{A}) = \max_{\lambda\in\sigma(\mathbf{A})} |\lambda| \leqslant \|\mathbf{A}\|.$$

(b) Suponhamos que $\mathbf{A}$ é diagonalizável e sejam $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ os valores próprios de $\mathbf{A}$ (com repetições). Seja $\mathbf{P} \in \mathbb{C}^{n\times n}$ uma matriz invertível tal que

$$\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}, \qquad \mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix};$$

sendo assim,

$$\mathbf{P}^{-1}\mathbf{A}^k\mathbf{P} = (\mathbf{P}^{-1}\mathbf{AP})^k = \mathbf{D}^k = \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{bmatrix}, \qquad k \in \mathbb{N}.$$

Suponhamos que a sucessão $\big(\mathbf{A}^k\big)_{k\in\mathbb{N}}$ é convergente com $\lim_{k\to\infty} \mathbf{A}^k = \mathbf{0}$. Então, a sucessão $\big(\mathbf{P}^{-1}\mathbf{A}^k\mathbf{P}\big)_{k\in\mathbb{N}}$ também é convergente com

$$\lim_{k\to\infty} \mathbf{P}^{-1}\mathbf{A}^k\mathbf{P} = \mathbf{P}^{-1}\big(\lim_{k\to\infty} \mathbf{A}^k\big)\mathbf{P} = \mathbf{0} \ ^{(*)}$$

Como

$$\lim_{k\to\infty} \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{bmatrix} = \begin{bmatrix} \lim_{k\to\infty} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lim_{k\to\infty} \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lim_{k\to\infty} \lambda_n^k \end{bmatrix},$$

concluímos que, para cada $1 \leqslant i \leqslant n$, a sucessão $\big(\lambda_i^k\big)_{k\in\mathbb{N}}$ é convergente com $\lim_{k\to\infty} \lambda_i^k = 0$, o que só pode acontecer quando $|\lambda_i| < 1$. Segue-se que

$$\rho(\mathbf{A}) = \max_{1\leqslant i\leqslant n} |\lambda_i| < 1.$$

Reciprocamente, suponhamos que $\rho(\mathbf{A}) < 1$. Então, para cada $1 \leqslant i \leqslant n$, temos $|\lambda_i| < 1$ e, portanto, a sucessão $\big(\lambda_i^k\big)_{k\in\mathbb{N}}$ é convergente com $\lim_{k\to\infty} \lambda_i^k = 0$. Segue-se que a sucessão

(*) Este facto deve-se à continuidade da função $\mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ definida pela correspondência $\mathbf{X} \mapsto \mathbf{P}^{-1}\mathbf{XP}$ (exercício).

$\left(\mathbf{D}^k\right)_{k\in\mathbb{N}}$ também é convergente com $\lim_{k\to\infty}\mathbf{D}^k = \mathbf{0}$ e portanto, a sucessão $\left(\mathbf{A}^k\right)_{k\in\mathbb{N}}$ também é convergente com

$$\lim_{k\to\infty}\mathbf{A}^k = \lim_{k\to\infty}\mathbf{P}\mathbf{D}^k\mathbf{P}^{-1} = \mathbf{P}\big(\lim_{k\to\infty}\mathbf{D}^k\big)\mathbf{P}^{-1} = \mathbf{0}.$$

(c) Sejam $\lambda_1,\dots,\lambda_r \in \mathbb{C}$ os valores próprios, distintos dois-a-dois, de $\mathbf{A}$ e sejam $\mathbf{G}_1,\dots,\mathbf{G}_r$ os projectores espectrais de $\mathbf{A}$, de modo que $\mathbf{A} = \lambda_1\mathbf{G}_1 + \cdots + \lambda_r\mathbf{G}_r$. Sem perda de genariladade, suponhamos que $|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_r|$, de modo que$\rho(\mathbf{A}) = |\lambda_1|$. Para cada $k \in \mathbb{N}$, seja $\nu_k = \|\mathbf{A}^k\|/|\lambda_1^k|$. Como

$$\mathbf{A}^k = \lambda_1^k\mathbf{G}_1 + \cdots + \lambda_r^k\mathbf{G}_r,$$

temos

$$\begin{aligned}\nu_k = \frac{\|\mathbf{A}\|^k}{|\lambda_1|^k} = \frac{\|\lambda_1^k\mathbf{G}_1 + \cdots + \lambda_r^k\mathbf{G}_r\|}{|\lambda_1^k|} &= \big\|\mathbf{G}_1 + (\lambda_2/\lambda_1)^k\,\mathbf{G}_2 + \cdots + (\lambda_r/\lambda_1)^k\,\mathbf{G}_r\big\| \\ \leqslant \|\mathbf{G}_1\| + \big(|\lambda_2|/|\lambda_1|\big)^k\,\|\mathbf{G}_2\| + \cdots + \big(|\lambda_r|/|\lambda_1|\big)^k\,\|\mathbf{G}_r\| &\leqslant \|\mathbf{G}_1\| + \|\mathbf{G}_2\| + \cdots + \|\mathbf{G}_r\|.\end{aligned}$$

Pondo $\nu = \|\mathbf{G}_1\| + \cdots + \|\mathbf{G}_r\|$, temos

$$1 \leqslant \nu_k \leqslant \nu, \qquad k \in \mathbb{N};$$

notemos que, como $\lambda_1^k \in \sigma(\mathbf{A}^k)$, temos $|\lambda_1|^k = |\lambda_1^k| \leqslant \|\mathbf{A}^k|$ (pela alínea (a)), logo $1 \leqslant \nu_k$. Por conseguinte, temos

$$1 \leqslant \sqrt[k]{\nu_k} \leqslant \sqrt[k]{\nu}, \qquad k \in \mathbb{N}.$$

Como a sucessão $\left(\sqrt[k]{\nu}\right)_{k\in\mathbb{N}}$ é convergente com $\lim_{k\to\infty}\sqrt[k]{\nu} = 1$, concluímos que a sucessão $\left(\sqrt[k]{\nu_k}\right)_{k\in\mathbb{N}}$ também é convergente com

$$\lim_{k\to\infty}\sqrt[k]{\nu_k} = 1.$$

Como $\|\mathbf{A}^k\| = |\lambda_1|\sqrt[k]{\nu_k}$, segue-se que a sucessão $\left(\sqrt[k]{\|\mathbf{A}^k\|}\right)_{k\in\mathbb{N}}$ é convergente com

$$\lim_{k\to\infty}\sqrt[k]{\|\mathbf{A}^k\|} = |\lambda_1| = \rho(\mathbf{A}).$$

RESOLUÇÃO DOS EXERCÍCIOS – FOLHA 6

6.1. Fácil.

6.2. (a) Para qualquer $\mathbf{v} \in \mathbb{C}^{n\times 1}$,

- $|\mathbf{v}|_i = \sqrt{|v_i|^2} \leqslant \sqrt{\sum_{1\leqslant i\leqslant n} |\mathbf{v}_i|^2} = \|\mathbf{v}\|_2$, logo $\max_{1\leqslant i\leqslant n} |\mathbf{v}_i| \leqslant \|\mathbf{v}\|_2$
- $\|\mathbf{v}\|_2 = \sqrt{\sum_{1\leqslant i\leqslant n} |v_i|^2} \leqslant \sqrt{\big(\sum_{1\leqslant i\leqslant n} |v_i|\big)^2} = \sum_{1\leqslant i\leqslant n} |v_i| = \|\mathbf{v}\|_1$.

(a) Seja $\mathbf{v} \in \mathbb{C}^{n\times 1}$ e seja $1 \leqslant k \leqslant n$.

- Consideremos os vectores $\mathbf{u} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ e $\mathbf{w} = \begin{bmatrix} |\mathbf{v}_1| \\ \vdots \\ |v_n| \end{bmatrix}$. Sabemos (da Álgebra Linear) que $\big|\langle \mathbf{u}|\mathbf{w}\rangle\big| \leqslant \|\mathbf{u}\|_2 \|\mathbf{w}\|_2$, ou seja,

$$\|\mathbf{v}\|_1 = \sum_{1\leqslant i\leqslant n} |v_i| = \langle \mathbf{u}|\mathbf{w}\rangle \leqslant \sqrt{n}\sqrt{\sum_{1\leqslant i\leqslant n} |v_i|^2} = \sqrt{n}\,\|\mathbf{v}\|_2.$$

- $\|\mathbf{v}\|_2^2 = \sum_{1\leqslant i\leqslant n} |v_i|^2 \leqslant n \max_{1\leqslant i\leqslant n} |v_i|^2 = n\|\mathbf{v}\|_\infty$, logo $\|\mathbf{v}\|_2 \leqslant \sqrt{n}\|\mathbf{v}\|_\infty$

6.3. Para quaisquer $\mathbf{u}, \mathbf{v} \in \mathbb{C}^{n\times 1}$, temos:

- $\|\mathbf{u}\| = \|(\mathbf{u}-\mathbf{v}) + \mathbf{v}\| \leqslant \|\mathbf{u}-\mathbf{v}\| + \|\mathbf{v}\|$, logo $\|\mathbf{u}\| - \|\mathbf{v}\| \leqslant \|\mathbf{u}-\mathbf{v}\|$;
- $-\big(\|\mathbf{u}\| - \|\mathbf{v}\|\big) = \|\mathbf{v}\| - \|\mathbf{u}\| \leqslant \|\mathbf{v}-\mathbf{u}\| = \|\mathbf{u}-\mathbf{v}\|$.

6.4. Seja $\lambda \in \mathbb{R}$, $0 < \lambda < 1$. A função $f(x) = (1-\lambda) + \lambda x - x^\lambda$ é contínua e tem derivada $f'(x) = \lambda - \lambda x^{\lambda-1} = -\lambda(x^{1-\lambda} - 1)$. Temos $f'(x) < 0$, para $x < 1$, e $f'(x) > 0$, para $x > 1$, logo a função $f(x)$ é decrescente no intervalo $]-\infty, 1[$ e crescente no intervalo $]1, +\infty[$; deste modo, $f(x)$ tem um máximo (absoluto) em $x = 1$, ou seja $f(x) \geqslant f(1) = 0$ para todo $x \in \mathbb{R}$, $x \neq 1$. Tomando $\alpha, \beta \in \mathbb{R}^+$ arbitrários, obtemos

$$0 \leqslant f(\alpha/\beta) = (1-\lambda) + \lambda(\alpha/\beta) - (\alpha/\beta)^\lambda,$$

logo

$$\alpha^\lambda \beta^{1-\lambda} = (\alpha/\beta)^\lambda \beta \leqslant \lambda\alpha + (1-\lambda)\beta.$$

Agora, sejam $\mathbf{u}, \mathbf{v} \in \mathbb{C}^{n\times 1}$ e consideremos os vectores $\hat{\mathbf{u}} = \frac{1}{\|\mathbf{u}\|_p}\mathbf{u}$ e $\hat{\mathbf{v}} = \frac{1}{\|\mathbf{v}\|_q}\mathbf{v}$; notemos que $\|\hat{\mathbf{u}}\|_p = \|\hat{\mathbf{v}}\|_q = 1$. Tomando $\alpha = |\hat{u}_i|^p$, $\beta = |\hat{v}_i|^q$ e $\lambda = 1/p$ (de modo que $1-\lambda = 1/q$), obtemos

$$|\hat{u}_i\hat{v}_i| = \big(|\hat{u}_i|^p\big)^{1/p}\big(|\hat{v}_i|^q\big)^{1/q} \leqslant (1/p)|\hat{u}_i|^p + (1/q)|\hat{v}_i|^q, \quad 1 \leqslant i \leqslant n,$$

de modo que

$$\sum_{1\leqslant i\leqslant n} |\hat{u}_i\hat{v}_i| \leqslant \frac{1}{p}\sum_{1\leqslant i\leqslant n} |\hat{u}_i|^p + \frac{1}{q}\sum_{1\leqslant i\leqslant n} |\hat{v}_i|^q = \frac{1}{p}\|\hat{\mathbf{u}}\|_p + \frac{1}{q}\|\hat{\mathbf{v}}\|_q = \frac{1}{p} + \frac{1}{q} = 1.$$

Sendo assim,

$$\begin{aligned}\sum_{1\leqslant i\leqslant n} |u_iv_i| &= \sum_{1\leqslant i\leqslant n} \big|\big(\|\mathbf{u}\|_p\hat{u}_i\big)\big(\|\mathbf{v}\|_p\hat{v}_i\big)\big| = \|\mathbf{u}\|_p\|\mathbf{v}\|_q \sum_{1\leqslant i\leqslant n} |\hat{u}_i\hat{v}_i| \\ &\leqslant \|\mathbf{u}\|_p\|\mathbf{v}\|_q = \Bigg(\sum_{1\leqslant i\leqslant n} |u_i|^p\Bigg)^{1/p}\Bigg(\sum_{1\leqslant i\leqslant n} |v_i|^q\Bigg)^{1/q}.\end{aligned}$$

Finalmente, deduzimos que

$$\begin{aligned}|\mathbf{u}^*\mathbf{v}| &= \Bigg|\sum_{1\leqslant i\leqslant n} \overline{u_i}v_i\Bigg| \leqslant \sum_{1\leqslant i\leqslant n} |\overline{u_i}|\,|v_i| = \sum_{1\leqslant i\leqslant n} |u_i|\,|v_i| = \sum_{1\leqslant i\leqslant n} \big|\|\mathbf{u}\|_p\hat{u}_i\big|\,\big|\|\mathbf{v}\|_q\hat{v}_i\big| \\ &= \|\mathbf{u}\|_p\|\mathbf{v}\|_q \sum_{1\leqslant i\leqslant n} |\hat{u}_i|\,|\hat{v}_i| = \|\mathbf{u}\|_p\|\mathbf{v}\|_q \sum_{1\leqslant i\leqslant n} |\hat{u}_i\hat{v}_i| \leqslant \|\mathbf{u}\|_p\|\mathbf{v}\|_q.\end{aligned}$$

Como se queria.

6.5. Seja $\mathbf{u} \in \mathbb{R}^{n\times 1}$ tal que $\sum_{1\leqslant i\leqslant n} u_i = 0$; assim, $\mathbf{u}^*\mathbf{e} = 0$ onde $\mathbf{e} = \begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}^{\mathsf{T}}$. Sejam $\mathbf{v} \in \mathbb{C}^{n\times 1}$ e $\alpha \in \mathbb{R}^+$ arbitrários. Então, $\mathbf{u}^*(\mathbf{v} - \alpha\mathbf{e}) = \mathbf{u}^*\mathbf{v} - \alpha\mathbf{u}^*\mathbf{e} = \mathbf{u}^*\mathbf{v}$ e, portanto,

$$\begin{aligned}|\mathbf{u}^*\mathbf{v}| = |\mathbf{u}^*(\mathbf{v} - \alpha\mathbf{e})| &= \Bigg|\sum_{1\leqslant i\leqslant n} \overline{u_i}(v_i - \alpha)\Bigg| \leqslant \sum_{1\leqslant i\leqslant n} |u_i|\,|v_i - \alpha| \\ &\leqslant \Bigg(\sum_{1\leqslant i\leqslant n} |u_i|\Bigg)\|\mathbf{v} - \alpha\mathbf{e}\|_\infty = \|\mathbf{u}\|_1\|\mathbf{v} - \alpha\mathbf{e}\|_\infty\end{aligned}$$

Como α é arbitrário, concluímos que

$$|\mathbf{u}^*\mathbf{v}| \leqslant \|\mathbf{u}\|_1 \inf_{\alpha\in\mathbb{R}^+} \|\mathbf{v} - \alpha\mathbf{e}\|_\infty.$$

Ora, temos

$$\inf_{\alpha\in\mathbb{R}^+} \|\mathbf{v} - \alpha\mathbf{e}\|_\infty = \frac{v_{\max} - v_{\min}}{2}$$

e, de facto,

$$\frac{v_{\max} - v_{\min}}{2} = \|\mathbf{v} - \alpha\mathbf{e}\|_\infty$$

para $\alpha = \frac{v_{\max}+v_{\min}}{2}$ (de modo que $\inf_{\alpha\in\mathbb{R}^+} \|\mathbf{v} - \alpha\mathbf{e}\|_\infty = \min_{\alpha\in\mathbb{R}^+} \|\mathbf{v} - \alpha\mathbf{e}\|_\infty$). Em conclusão,

$$|\mathbf{u}^*\mathbf{v}| \leqslant \|\mathbf{u}\|_1\frac{v_{\max} - v_{\min}}{2} = \frac{v_{\max} - v_{\min}}{2}\sum_{1\leqslant i\leqslant n} |u_i|,$$

como se queria.

6.6. (a) Supomos que $p \neq 1$ e observamos que $p/q = p-1$ (quando $q = p/(p-1)$). Para quaisquer $\alpha, \beta \in \mathbb{C}$, temos

$$|\alpha+\beta|^p = |\alpha+\beta|\,|\alpha+\beta|^{p-1} \leqslant |\alpha|\,|\alpha+\beta|^{p/q} + |\beta|\,|\alpha+\beta|^{p/q}.$$

Sejam $\mathbf{u}, \mathbf{v} \in \mathbb{C}^{n\times 1}$ quaisquer. Temos

$$\sum_{1\leqslant i\leqslant n} |u_i+v_i|^p = \sum_{1\leqslant i\leqslant n} |u_i+v_i|\,|u_i+v_i|^{p-1} \leqslant \sum_{1\leqslant i\leqslant n} |u_i|\,|\,|u_i+v_i|^{p-1} + \sum_{1\leqslant i\leqslant n} |v_i|\,|\,|u_i+v_i|^{p-1}.$$

Usando a desigualdade de Hölder, deduzimos que

$$\begin{aligned}\sum_{1\leqslant i\leqslant n} |u_i|\,|\,|u_i+v_i|^{p/q} &= \sum_{1\leqslant i\leqslant n} |u_i(u_i+v_i)^{p/q}| \leqslant \Bigg(\sum_{1\leqslant i\leqslant n} |u_i|^p\Bigg)^{1/p} \Bigg(\sum_{1\leqslant i\leqslant n} \big(|u_i+v_i|^{p/q})^q\Bigg)^{1/q} \\ &= \Bigg(\sum_{1\leqslant i\leqslant n} |u_i|^p\Bigg)^{1/p} \Bigg(\sum_{1\leqslant i\leqslant n} |u_i+v_i|^p\Bigg)^{(p-1)/p} = \|\mathbf{u}\|_p \|\mathbf{u}+\mathbf{v}\|_p^{p-1}\end{aligned}$$

e, de maneira análoga,

$$\sum_{1\leqslant i\leqslant n} |v_i|\,|\,|u_i+v_i|^{p/q} \leqslant \|\mathbf{v}\|_p \|\mathbf{u}+\mathbf{v}\|_p^{p-1}.$$

Por conseguinte, obtemos

$$\|\mathbf{u}+\mathbf{v}\|_p^p \leqslant \big(\|\mathbf{u}\|_p + \|\mathbf{v}\|_p\big)\|\mathbf{u}+\mathbf{v}\|_p^{p-1}$$

e, portanto,

$$\|\mathbf{u}+\mathbf{v}\|_p \leqslant \|\mathbf{u}\|_p + \|\mathbf{v}\|_p.$$

(b) Tendo em conta a alínea anterior, é fácil concluir que $\| \star \|_p$ é uma norma em $\mathbb{C}^{n\times 1}$. Seja $\mathbf{v} \in \mathbb{C}^{n\times 1}$ e provemos que $\lim_{p\to\infty} \|\mathbf{v}\|_p$ existe e que $\lim_{p\to\infty} \|\mathbf{v}\|_p = \|\mathbf{v}\|_\infty$. Por um lado, para qualquer $p \in \mathbb{R}$, $p \geqslant 1$, temos

$$\|\mathbf{v}\|_\infty = \max_{1\leqslant i\leqslant n} |v_i| = \big(\max_{1\leqslant i\leqslant n} |v_i|^p\big)^{1/p} \leqslant \Bigg(\sum_{1\leqslant i\leqslant n} |v_i|^p\Bigg)^{1/p} = \|\mathbf{v}\|_p$$

e, por outro lado,

$$\|\mathbf{v}\|_p = \Bigg(\sum_{1\leqslant i\leqslant n} |v_i|^p\Bigg)^{1/p} \leqslant \Bigg(\sum_{1\leqslant i\leqslant n} \big(\max_{1\leqslant i\leqslant n} |v_i|\big)^p\Bigg)^{1/p} = n^{1/p}\|\mathbf{v}\|_\infty.$$

Por conseguinte,

$$\|\mathbf{v}\|_\infty \leqslant \|\mathbf{v}\|_p \leqslant n^{1/p}\|\mathbf{v}\|_\infty$$

e, portanto, como $\lim_{p\to\infty} n^{1/p}$ existe e $\lim_{p\to\infty} = 1$, concluímos que $\lim_{p\to\infty} \|\mathbf{v}\|_p$ também existe e tem-se

$$\lim_{p\to\infty} \|\mathbf{v}\|_p = \|\mathbf{v}\|_\infty.$$

6.7. Para $\mathbf{A} = \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}$, temos $\|\mathbf{A}\|_2 = \sqrt{1+4+4+1} = \sqrt{10}$.

Para $\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$, temos $\|\mathbf{A}\|_2 = \sqrt{3}$.

Para $\mathbf{A} = \begin{bmatrix} 4 & -2 & 4 \\ -2 & 1 & -2 \\ 4 & -2 & 4 \end{bmatrix}$, temos $\|\mathbf{A}\|_2 = \sqrt{4\cdot 16 + 4 \cdot 4 + 1} = \sqrt{81} = 9$.

6.8. Sabemos que $\|\mathbf{A}\|_s = \sqrt{\lambda_{\max}}$ onde $\lambda_{\max}$ é o maior valor próprio da matriz hermítica

$$\mathbf{A}^*\mathbf{A} = \frac{1}{3}\begin{bmatrix} 3 & 0 \\ -1 & \sqrt{8} \end{bmatrix}\begin{bmatrix} 3 & -1 \\ 0 & \sqrt{8} \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}.$$

Os valores próprios desta matriz são 2 e 4, logo $\|\mathbf{A}\|_s = \sqrt{4} = 2$.

$\mathbf{A}$ é invertível porque $\det(\mathbf{A}) \neq 0$. Tem-se $\|\mathbf{A}^{-1}\|_s = \frac{1}{\sqrt{\lambda_{\min}}}$ onde $\lambda_{\min}$ é o menor valor próprio de $\mathbf{A}^*\mathbf{A}$, logo $\|\mathbf{A}^{-1}\|_s = \frac{1}{\sqrt{2}}$.

6.9. • Seja $\mathbf{A} \in \mathbb{C}^{n\times n}$. Por definição, temos

$$\|\mathbf{A}\|_1' = \max_{\mathbf{v}\in\mathcal{S}} \|\mathbf{A}\mathbf{v}\|_1$$

onde $\mathcal{S} = \{\mathbf{v} \in \mathbb{C}^{n\times 1} \colon \|\mathbf{v}\|_1 = 1\}$. Para qualquer $\mathbf{v} \in \mathcal{S}$, temos

$$\begin{aligned}
\|\mathbf{A}\mathbf{v}\|_1 &= \sum_{1\leqslant i\leqslant n}\left|\sum_{1\leqslant j\leqslant n} a_{i,j}v_j\right| \leqslant \sum_{1\leqslant i,j\leqslant n} |a_{i,j}|\,|v_j| = \sum_{1\leqslant j\leqslant n}\Bigg(\sum_{1\leqslant i\leqslant n}|a_{i,j}|\Bigg)|v_j| \\
&\leqslant \sum_{1\leqslant j\leqslant n}\Bigg(\max_{1\leqslant k\leqslant n}\sum_{1\leqslant i\leqslant n}|a_{i,k}|\Bigg)|v_j| = \Bigg(\max_{1\leqslant k\leqslant n}\sum_{1\leqslant i\leqslant n}|a_{i,k}|\Bigg)\sum_{1\leqslant j\leqslant n}|v_j| \\
&= \Bigg(\max_{1\leqslant k\leqslant n}\sum_{1\leqslant i\leqslant n}|a_{i,k}|\Bigg)\|\mathbf{v}\|_1 = \max_{1\leqslant k\leqslant n}\sum_{1\leqslant i\leqslant n}|a_{i,k}|
\end{aligned}$$

e, portanto,

$$\|\mathbf{A}\|_1' = \max_{\mathbf{v}\in\mathcal{S}}\|\mathbf{A}\mathbf{v}\| \leqslant \max_{1\leqslant j\leqslant n}\sum_{1\leqslant i\leqslant n}|a_{i,j}|.$$

Por outro lado, se $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ for a base canónica de $\mathbb{C}^{n\times 1}$ temos $\mathbf{e}_1, \ldots, \mathbf{e}_n \in \mathcal{S}$ e

$$\mathbf{A}\mathbf{e}_j = \begin{bmatrix} a_{1,j} \\ \vdots \\ a_{n,j} \end{bmatrix}, \qquad 1 \leqslant j \leqslant n.$$

Sendo assim,

$$\|\mathbf{A}\|_1' = \max_{\mathbf{v}\in\mathcal{S}}\|\mathbf{A}\mathbf{v}\|_1 \geqslant \max_{1\leqslant j\leqslant n}\|\mathbf{A}\mathbf{e}_j\|_1 = \max_{1\leqslant j\leqslant n}\sum_{1\leqslant i\leqslant n}|a_{i,j}| \geqslant \|\mathbf{A}\|_1'$$

e, portanto,

$$\|\mathbf{A}\|_1' = \max_{1\leqslant j\leqslant n} \sum_{1\leqslant i\leqslant n} |a_{i,j}|.$$

- Seja $\mathbf{A} \in \mathbb{C}^{n\times n}$. Por definição, temos

$$\|\mathbf{A}\|_\infty' = \max_{\mathbf{v}\in\mathcal{S}} \|\mathbf{A}\mathbf{v}\|_\infty$$

onde $\mathcal{S} = \{\mathbf{v} \in \mathbb{C}^{n\times 1} : \|\mathbf{v}\|_\infty = 1\}$. Para qualquer $\mathbf{v} \in \mathcal{S}$, temos

$$\begin{aligned} \|\mathbf{A}\mathbf{v}\|_\infty &= \max_{1\leqslant i\leqslant n} \left| \sum_{1\leqslant j\leqslant n} a_{i,j} v_j \right| \leqslant \max_{1\leqslant i\leqslant n} \sum_{1\leqslant j\leqslant n} |a_{i,j}|\,|v_j| \\ &\leqslant \max_{1\leqslant i\leqslant n} \sum_{1\leqslant j\leqslant n} |a_{i,j}| \|\mathbf{v}|_\infty = \max_{1\leqslant i\leqslant n} \sum_{1\leqslant j\leqslant n} |a_{i,j}| \end{aligned}$$

e, portanto,

$$\|\mathbf{A}\|_\infty' = \max_{\mathbf{v}\in\mathcal{S}} \|\mathbf{A}\mathbf{v}\| \leqslant \max_{1\leqslant i\leqslant n} \sum_{1\leqslant j\leqslant n} |a_{i,j}|.$$

No caso em que $\mathbf{A} = \mathbf{0}$, temos (trivialmente) $\|\mathbf{A}\|_\infty' = 0 = \max_{1\leqslant i\leqslant n} \sum_{1\leqslant j\leqslant n} |a_{i,j}|$. Assim, suponhamos que $\mathbf{A} \neq \mathbf{0}$ e seja $1 \leqslant i \leqslant n$ tal que a i-ésima linha de $\mathbf{A}$ é não-nula (isto é, existe $1 \leqslant j \leqslant n$ tal que $a_{i,j} \neq 0$). Definamos o vector $\mathbf{w} \in \mathbb{C}^{n\times 1}$ por

$$w_k = \begin{cases} \frac{\overline{a_{i,k}}}{|a_{i,k}|}, & \text{se } a_{i,k} \neq 0, \\ 1, & \text{se } a_{i,k} = 0, \end{cases}, \qquad 1 \leqslant k \leqslant n;$$

notemos que $|w_k| = 1$ para qualquer $1 \leqslant k \leqslant n$, logo $\|\mathbf{w}\|_\infty = 1$ (isto é, $\mathbf{w} \in \mathcal{S}$). Sendo assim, temos

$$\|\mathbf{A}\|_\infty' = \max_{\mathbf{v}\in\mathcal{S}} \|\mathbf{A}\mathbf{v}\|_\infty \geqslant \|\mathbf{A}\mathbf{w}\|_\infty \max_{1\leqslant j\leqslant n} \left| \sum_{1\leqslant k\leqslant n} a_{j,k} w_k \right| \geqslant \left| \sum_{1\leqslant k\leqslant n} a_{i,k} w_k \right| = \sum_{1\leqslant k\leqslant n} |a_{i,k}|.$$

Daqui, resulta que

$$\|\mathbf{A}\|_\infty' \geqslant \max_{1\leqslant i\leqslant n} \sum_{1\leqslant j\leqslant n} |a_{i,j}|$$

(repetindo o raciocínio para qualquer linha não-nula de $\mathbf{A}$) e, portanto,

$$\|\mathbf{A}\|_\infty' = \max_{1\leqslant i\leqslant n} \sum_{1\leqslant j\leqslant n} |a_{i,j}|,$$

como se queria.

- $\|\mathbf{A}\|_1' = \max\left\{\sqrt{3}, \frac{\sqrt{3}}{3}\left(1+\sqrt{8}\right)\right\} = \sqrt{3}$ e $\|\mathbf{A}\|_\infty' = \max\left\{\frac{2\sqrt{3}}{3}, \frac{2\sqrt{6}}{3}\right\} = \frac{2\sqrt{6}}{3}$.

6.10. • Para $\mathbf{A} = \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}$, temos $\|\mathbf{A}\|_1' = \max\{2, 4\} = 4$ e $\|\mathbf{A}\|_\infty' = \max\{3, 3\} = 3$.

Quanto à norma espectral, a matriz $\mathbf{A}^*\mathbf{A} = \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}$ tem valores próprios 0 e 3, de modo que $\|\mathbf{A}\|_s = \sqrt{3}$.

- Para $\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$, temos $\|\mathbf{A}\|_1' = 1$ e $\|\mathbf{A}\|_\infty' = 1$. Quanto à norma espectral, a matriz $\mathbf{A}^*\mathbf{A} = \mathbf{I}_3$, logo $\|\mathbf{A}\|_s = 1$.

- Para $\mathbf{A} = \begin{bmatrix} 4 & -2 & 4 \\ -2 & 1 & -2 \\ 4 & -2 & 4 \end{bmatrix}$, temos $\|\mathbf{A}\|_1' = \max\{10, 5\} = 10$ e $\|\mathbf{A}\|_\infty' = \max\{10, 5\} = 10$.

Quanto à norma espectral, a matriz $\mathbf{A}^*\mathbf{A} = \begin{bmatrix} 36 & -18 & 36 \\ -18 & 9 & -18 \\ 36 & -18 & 36 \end{bmatrix}$ tem valores próprios 0 e 81, de modo que $\|\mathbf{A}\|_s = \sqrt{81} = 9$.

6.11. • É claro que, para qualquer $\mathbf{A} \in \mathbb{C}^{n\times n}$, se tem $\|\mathbf{A}\|_\mathbf{P} = \|\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\| \geqslant 0$ e

$$\|\mathbf{A}\|_\mathbf{P} = \|\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\| = 0 \quad \Longleftrightarrow \quad \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{0} \quad \Longleftrightarrow \quad \mathbf{A} = \mathbf{0}.$$

- Para qualquer $\mathbf{A} \in \mathbb{C}^{n\times n}$ e qualquer $\alpha \in \mathbb{C}$, temos

$$\|\alpha\mathbf{A}\|_\mathbf{P} = \|\mathbf{P}^{-1}(\alpha\mathbf{A})\mathbf{P}\| = \|\alpha(\mathbf{P}^{-1}\mathbf{A}\mathbf{P})\| = |\alpha|\,\|\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\| = |\alpha|\,\|\mathbf{A}\|_\mathbf{P}.$$

- Para quaisquer $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$, temos

$*$ $\|\mathbf{A}+\mathbf{B}\|_\mathbf{P} = \|\mathbf{P}^{-1}(\mathbf{A}+\mathbf{B})\mathbf{P}\| = \|\mathbf{P}^{-1}\mathbf{A}\mathbf{P} + \mathbf{P}^{-1}\mathbf{B}\mathbf{P}\| \leqslant \|\mathbf{P}^{-1}\mathbf{A}\mathbf{P} + \mathbf{P}^{-1}\mathbf{B}\mathbf{P}\| = \|\mathbf{A}\|_\mathbf{P} + \|\mathbf{B}\|_\mathbf{P}$;

$*$ $\|\mathbf{A}\mathbf{B}\|_\mathbf{P} = \|\mathbf{P}^{-1}(\mathbf{A}\mathbf{B})\mathbf{P}\| = \|(\mathbf{P}^{-1}\mathbf{A}\mathbf{P})(\mathbf{P}^{-1}\mathbf{B}\mathbf{P})\| \leqslant \|\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\|\,\|\mathbf{P}^{-1}\mathbf{B}\mathbf{P}\| = \|\mathbf{A}\|_\mathbf{P}\|\mathbf{B}\|_\mathbf{P}$

6.12. Seja $\mathbf{A} \in \mathbb{k}^{n\times n}$ qualquer.

(a) Para quaisquer $\mathbf{v}, \mathbf{w} \in \mathcal{S}$, temos

$$|\mathbf{w}^*\mathbf{A}\mathbf{v}| = |\langle \mathbf{w}|\mathbf{A}\mathbf{v}\rangle| \leqslant \|\mathbf{w}\|_2\|\mathbf{A}\mathbf{v}\|_2 = \|\mathbf{A}\mathbf{v}\|_2,$$

logo

$$\sup_{\mathbf{v}\in\mathcal{S}}\sup_{\mathbf{w}\in\mathcal{S}} |\mathbf{w}^*\mathbf{A}\mathbf{v}| \leqslant \sup_{\mathbf{v}\in\mathcal{S}}\sup_{\mathbf{w}\in\mathcal{S}} \|\mathbf{A}\mathbf{v}\|_2 = \max_{\mathbf{v}\in\mathcal{S}} \|\mathbf{A}\mathbf{v}\|_2 = \|\mathbf{A}\|_s.$$

Agora, seja $\mathbf{v}_0 \in \mathcal{S}$ tal que

$$\|\mathbf{A}\mathbf{v}_0\|_2 = \max_{\mathbf{v}\in\mathcal{S}} \|\mathbf{A}\mathbf{v}\|_2 = \|\mathbf{A}\|_s$$

e seja

$$\mathbf{w}_0 = \frac{1}{\|\mathbf{A}\mathbf{v}_0\|_2}\,\mathbf{A}\mathbf{v}_0 \in \mathcal{S}.$$

Então,

$$\mathbf{w}_0^*\mathbf{A}\mathbf{v}_0 = \frac{1}{\|\mathbf{A}\mathbf{v}_0\|_2}\,\mathbf{v}_0^*\mathbf{A}^*\mathbf{A}\mathbf{v}_0 = \frac{1}{\|\mathbf{A}\mathbf{v}_0\|_2}\,\|\mathbf{A}\mathbf{v}_0\|_2^2 = \|\mathbf{A}\mathbf{v}_0\|_2 = \|\mathbf{A}\|_s,$$

de onde resulta que

$$\|\mathbf{A}\|_s = \mathbf{w}_0^*\mathbf{A}\mathbf{v}_0 \leqslant \sup_{\mathbf{v}\in\mathcal{S}}\sup_{\mathbf{w}\in\mathcal{S}} |\mathbf{w}^*\mathbf{A}\mathbf{v}| \leqslant \|\mathbf{A}\|_s$$

e, portanto,

$$\|\mathbf{A}\|_s = \mathbf{w}_0^*\mathbf{A}\mathbf{v}_0 = \max_{\mathbf{v}\in\mathcal{S}}\max_{\mathbf{w}\in\mathcal{S}} |\mathbf{w}^*\mathbf{A}\mathbf{v}|.$$

(b) Pela alínea (a), temos

$$\|\mathbf{A}^*\|_s = \max_{\mathbf{v}\in\mathcal{S}}\max_{\mathbf{w}\in\mathcal{S}} |\mathbf{w}^*\mathbf{A}^*\mathbf{v}| = \max_{\mathbf{w}\in\mathcal{S}}\max_{\mathbf{v}\in\mathcal{S}} |\mathbf{v}^*\mathbf{A}\mathbf{w}| = \|\mathbf{A}\|_s.$$

(c) Temos

$$\begin{aligned}\|\mathbf{A}^*\mathbf{A}\|_s &= \max_{\mathbf{v}\in\mathcal{S}}\max_{\mathbf{w}\in\mathcal{S}} |\mathbf{w}^*\mathbf{A}^*\mathbf{A}\mathbf{v}| = \max_{\mathbf{v}\in\mathcal{S}}\max_{\mathbf{w}\in\mathcal{S}} |(\mathbf{A}\mathbf{w})^*(\mathbf{A}\mathbf{v})| \\ &= \max_{\mathbf{v}\in\mathcal{S}}\max_{\mathbf{w}\in\mathcal{S}} |\langle\mathbf{A}\mathbf{w}|\mathbf{A}\mathbf{v}\rangle| \leqslant \max_{\mathbf{v}\in\mathcal{S}}\max_{\mathbf{w}\in\mathcal{S}} \|\mathbf{A}\mathbf{w}\|_2\|\mathbf{A}\mathbf{v}|\|_2 \\ &= \Big(\max_{\mathbf{v}\in\mathcal{S}} \|\mathbf{A}\mathbf{v}\|_2\Big)\Big(\max_{\mathbf{v}\in\mathcal{S}} \|\mathbf{A}\mathbf{v}\|_2\Big) = \|\mathbf{A}\|_s^2.\end{aligned}$$

Para provarmos a igualdade, basta considerar $\mathbf{v}_0 \in \mathcal{S}$ tal que

$$\|\mathbf{A}\mathbf{v}_0\|_2 = \max_{\mathbf{v}\in\mathcal{S}} \|\mathbf{A}\mathbf{v}\|_2 = \|\mathbf{A}\|_s$$

e observar que

$$|\mathbf{v}_0^*\mathbf{A}^*\mathbf{A}\mathbf{v}_0| = \mathbf{v}_0^*\mathbf{A}^*\mathbf{A}\mathbf{v}_0 = \|\mathbf{A}\mathbf{v}_0\|_2^2 = \|\mathbf{A}\|_s^2.$$

(d) Sejam $\mathbf{U}, \mathbf{V} \in \mathbb{k}^{n\times n}$ matrizes unitárias. Como $\mathbf{U}^*\mathbf{U} = \mathbf{I}_n$, deduzimos que

$$\|\mathbf{U}\mathbf{A}\|_s = \max_{\mathbf{v}\in\mathcal{S}} \|\mathbf{U}\mathbf{A}\mathbf{v}\|_2 = \max_{\mathbf{v}\in\mathcal{S}} \sqrt{\mathbf{v}^*\mathbf{A}^*\mathbf{U}^*\mathbf{U}\mathbf{A}\mathbf{v}} = \max_{\mathbf{v}\in\mathcal{S}} \sqrt{\mathbf{v}^*\mathbf{A}^*\mathbf{A}\mathbf{v}} = \max_{\mathbf{v}\in\mathcal{S}} \|\mathbf{A}\mathbf{v}\|_2 = \|\mathbf{A}\|_s$$

e, portanto,

$$\|\mathbf{U}\mathbf{A}\mathbf{V}\|_s = \|\mathbf{U}(\mathbf{A}\mathbf{V})\|_s = \|\mathbf{A}\mathbf{V}\|_s = \|\mathbf{A}\|_s.$$

6.13. Seja $\mathbf{A} \in \mathbb{C}^{n\times n}$ qualquer. Para qualquer $z \in \mathbb{C}$ temos

$$z \notin \sigma(\mathbf{A}) \quad\Longleftrightarrow\quad \det(z\mathbf{I}_n - \mathbf{A}) \neq 0 \quad\Longleftrightarrow\quad z\mathbf{I}_n - \mathbf{A} \text{ é uma matriz invertível.}$$

Deste modo, a função $\mathbf{R}\colon \mathbb{C}\backslash\sigma(\mathbf{A}) \to \mathbb{C}^{n\times n}$ está bem-definida. Seja $z \in \mathbb{C}\backslash\sigma(\mathbf{A})$ tal que $|z| > \|\mathbf{A}\|$. Então,

$$\|\mathbf{R}(z)\| = \big\|(z\mathbf{I}_n - \mathbf{A})\big\| = \frac{1}{\min_{\mathbf{v}\in\mathcal{S}} \|(z\mathbf{I}_n - \mathbf{A})\mathbf{v}\|}$$

onde $\mathcal{S} = \{\mathbf{v} \in \mathbb{C}^{n\times 1} \colon \|\mathbf{v}\| = 1\}$. Seja $\mathbf{v}_0 \in \mathcal{S}$ tal que $\|(z\mathbf{I}_n - \mathbf{A})\mathbf{v}_0\| = \min_{\mathbf{v}\in\mathcal{S}} \|(z\mathbf{I}_n - \mathbf{A})\mathbf{v}\|$. Temos

$$\|(z\mathbf{I}_n - \mathbf{A})\mathbf{v}_0\| = \|z\mathbf{v}_0 - \mathbf{A}\mathbf{v}_0\| \geqslant \big|\|z\mathbf{v}_0\| - \|\mathbf{A}\mathbf{v}_0\|\big| = \big||z|\,\|\mathbf{v}_0\| - \|\mathbf{A}\mathbf{v}_0\|\big| = \big||z| - \|\mathbf{A}\mathbf{v}_0\|\big|.$$

Como $|z| > \|\mathbf{A}\|$ e $\|\mathbf{A}\mathbf{v}_0\| \leqslant \max_{\mathbf{v}\in\mathcal{S}} \|\mathbf{A}\mathbf{v}\| = \|\mathbf{A}\|$, concluímos que

$$\min_{\mathbf{v}\in\mathcal{S}} \|(z\mathbf{I}_n - \mathbf{A})\mathbf{v}\| = \|(z\mathbf{I}_n - \mathbf{A})\mathbf{v}_0\| \geqslant |z| - \|\mathbf{A}\|$$

e, portanto,

$$\|\mathbf{R}(z)\| = \frac{1}{\min_{\mathbf{v}\in\mathcal{S}} \|(z\mathbf{I}_n - \mathbf{A})\mathbf{v}\|} \leqslant \frac{1}{|z| - \|\mathbf{A}\|}.$$

6.14. Por definição, temos

$$\|\mathbf{A}\|_2 = \sqrt{\langle \mathbf{A}|\mathbf{A}\rangle} = \sqrt{\operatorname{tr}(\mathbf{A}^*\mathbf{A})} = \sqrt{\lambda_1 + \lambda_2 + \cdots + \lambda_n}$$

onde $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n \geqslant 0$ são todos os valores próprios (com repetições) de $\mathbf{A}^*\mathbf{A}$. Para concluir, basta observar que os valores singulares de $\mathbf{A}$ são as raízes quadradas dos valores próprios positivos de $\mathbf{A}^*\mathbf{A}$.

6.15. • Como $\mathbf{A}$ é normal, o teorema da decomposição espectral garante que existe uma matriz unitária $\mathbf{U} \in \mathbb{C}^{n\times n}$ tal que

$$\mathbf{U}^*\mathbf{A}\mathbf{U} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

onde $\lambda_1, \ldots, \lambda_n$ são os valores próprios de $\mathbf{A}$. Sendo assim, temos

$$\mathbf{U}^*\mathbf{A}^*\mathbf{A}\mathbf{U} = (\mathbf{U}^*\mathbf{A}\mathbf{U})^*(\mathbf{U}^*\mathbf{A}\mathbf{U}) = \begin{bmatrix} |\lambda_1|^2 & 0 & \cdots & 0 \\ 0 & |\lambda_2|^2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & |\lambda_n|^2 \end{bmatrix}$$

e, portanto, $|\lambda_1|^2, \ldots, |\lambda_n|^2$ são os valores próprios de $\mathbf{A}^*\mathbf{A}$. Segue-se que

$$\|\mathbf{A}\|_s = \sqrt{\max_{1\leqslant i\leqslant n} |\lambda_i|^2} = \max_{1\leqslant i\leqslant n} |\lambda_i| = \rho(\mathbf{A}).$$

• Como $\mathbf{A}$ e $\mathbf{B}$ são normais, obtemos $\rho(\mathbf{AB}) = \|\mathbf{AB}\|_s \leqslant \|\mathbf{A}\|_s\|\mathbf{B}\|_s = \rho(\mathbf{A})\rho(\mathbf{B})$.

6.16. (a) Por definição, $\sigma_1 = \sqrt{\lambda_1}$ onde λ_1 é o maior valor próprio de $\mathbf{A}^*\mathbf{A}$, de modo que $\|\mathbf{A}\|_s = \sigma_1$. Como $|\lambda| \leqslant \|\mathbf{A}\|_s$, concluímos que $|\lambda| \leqslant \sigma_1$ para qualquer $\lambda \in \sigma(\mathbf{A})$.

(b) Suponhamos que $\mathbf{A}$ é invertível. Por definição, os valores singulares de $\mathbf{A}^{-1}$ são as raízes quadradas dos valores próprios da matriz $(\mathbf{A}^{-1})^*(\mathbf{A}^{-1}) = (\mathbf{AA}^*)^{-1}$ e, além disso, $\sigma\big((\mathbf{AA}^*)^{-1}\big) = \{\lambda^{-1} \colon \lambda \in \sigma(\mathbf{AA}^*)\}$. Pela Proposição 6.3, as matrizes $\mathbf{AA}^*$ e $\mathbf{A}^*\mathbf{A}$ têm os mesmos valores próprios (positivos) e, portanto, os valores singulares de $\mathbf{A}^{-1}$ são $\sigma_1^{-1} < \sigma_2^{-1} < \cdots < \sigma_r^{-1}$. Sendo assim, como $\sigma(\mathbf{A}^{-1}) = \{\lambda^{-1} \colon \lambda \in \sigma(\mathbf{A})\}$, a alína (a) garante que $|\lambda|^{-1} \leqslant \sigma_r^{-1}$ e, portanto, $\sigma_r \leqslant |\lambda|$ para qualquer $\lambda \in \sigma(\mathbf{A})$.

6.17. (a) $\mathbf{A}^k = \dfrac{1}{2^k}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}^k = \dfrac{1}{2^k}\begin{bmatrix} 1 & 2k \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1/2^k & k/2^{k-1} \\ 0 & 1/2^k \end{bmatrix}$ e $\rho(\mathbf{A}^k) = \dfrac{1}{2^k}$.

(b) $\lim_{k\to\infty} \mathbf{A}^k = \begin{bmatrix} \lim_{k\to\infty}(1/2^k) & \lim_{k\to\infty}(k/2^{k-1}) \\ 0 & \lim_{k\to\infty}(1/2^k) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$; de facto, como $\rho(\mathbf{A}) < 1$, o Teorema 9.4 garante que $\lim_{k\to\infty} \mathbf{A}^k = 0$.

(b) Como qualquer norma $\| \star \| \colon \mathbb{C}^{n\times n} \to \mathbb{R}$ é uma função contínua (pelo Lema 7.2[(*)]), tem de ser $\lim_{k\to\infty} \|\mathbf{A}^k\|_s = \| \lim_{k\to\infty} \mathbf{A}^k\|_s = 0$, $\lim_{k\to\infty} \|\mathbf{A}^k\|_1' = 0$ e $\lim_{k\to\infty} \|\mathbf{A}^k\|_\infty' = 0$.

6.18. Temos $\mathbf{v}_k = \mathbf{A}^k\mathbf{v}_0$ para qualquer $k \in \mathbb{N}$. A matriz $\mathbf{A}$ tem valores próprios $(2 \pm \sqrt{2})/4$, logo $\rho(\mathbf{A}) = (2+\sqrt{2})/4 < 1$ e, portanto, $\lim_{k\to\infty} \mathbf{A}^k = 0$. Como

$$0 \leqslant \|\mathbf{A}^k\mathbf{v}_0\|_2 \leqslant \|\mathbf{A}\|_2\|\mathbf{v}_0\|_2,$$

concluímos que $\lim_{k\to\infty} \|\mathbf{A}^k\mathbf{v}_0\|_2 = 0$ e, portanto, $\lim_{k\to\infty} \mathbf{A}^k\mathbf{v}_0 = \mathbf{0}$, isto é, $\lim_{k\to\infty} \mathbf{v}_k = 0$.

[(*)] Que é verdadeiro para qualquer espaço vectorial.

RESOLUÇÃO DOS EXERCÍCIOS – FOLHA 7

7.1. Usando o teorema de Geršgorin, os valores próprios de $\mathbf{A}$ estão na união

$$\begin{aligned}\mathcal{G} &= \{z \in \mathbb{C}\colon |z-5| \leqslant 2\} \cup \{z \in \mathbb{C}\colon |z-6| \leqslant 1\} \cup \{z \in \mathbb{C}\colon |z+5| \leqslant 1\} \\ &= \{z \in \mathbb{C}\colon |z-5| \leqslant 2\} \cup \{z \in \mathbb{C}\colon |z+5| \leqslant 1\}\,.\end{aligned}$$

Usando a matriz transposta $\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 5 & 0 & 1 \\ 0 & 6 & 1 \\ 1 & 1 & -5 \end{bmatrix}$, garantimos que os valores próprios de $\mathbf{A}$ estão na união

$$\mathcal{G}' = \{z \in \mathbb{C}\colon |z-5| \leqslant 1\} \cup \{z \in \mathbb{C}\colon |z-6| \leqslant 1\} \cup \{z \in \mathbb{C}\colon |z+5| \leqslant 2\}\,,$$

de modo que

$$\sigma(\mathbf{A}) \subseteq \mathcal{G} \cap \mathcal{G}' = \{z \in \mathbb{C}\colon |z-5| \leqslant 1\} \cup \{z \in \mathbb{C}\colon |z-6| \leqslant 1\} \cup \{z \in \mathbb{C}\colon |z+5| \leqslant 1\}\,.$$

7.2. Pelo teorema de Geršgorin, temos

$$\sigma(\mathbf{A}) \subseteq \{z \in \mathbb{C}\colon |z-n| \leqslant n-1\}\,,$$

de modo que $0 \notin \sigma(\mathbf{A})$. Como $0 \notin \sigma(\mathbf{A})$, temos $\det(\mathbf{A}) \neq 0$, logo $\mathbf{A}$ é invertível.

7.3. Temos

$$\begin{aligned}\sigma(\mathbf{A}) &\subseteq \{z \in \mathbb{C}\colon |z-1| \leqslant 2\} \cup \{z \in \mathbb{C}\colon |z-12| \leqslant 4\} \cup \{z \in \mathbb{C}\colon |z+1| \leqslant 1\} \cup \{z \in \mathbb{C}\colon |z| \leqslant 5\} \\ &= \{z \in \mathbb{C}\colon |z| \leqslant 5\} \cup \{z \in \mathbb{C}\colon |z-12| \leqslant 4\}\,.\end{aligned}$$

Pelo Teorema 12.3, temos

$$\#\big(\sigma(\mathbf{A}) \cap \{z \in \mathbb{C}\colon |z-12| \leqslant 4\}\big) = 1$$

e, portanto, existe $\lambda \in \sigma(\mathbf{A})$ tal que $|\lambda - 12| \leqslant 4$; analogamente, temos

$$\#\big(\sigma(\mathbf{A}) \cap \{z \in \mathbb{C}\colon |z| \leqslant 5\}\big) = 3$$

e, portanto, existem $\mu_1, \mu_2, \mu_3 \in \sigma(\mathbf{A})$ tais que $|\mu_i| \leqslant 5$, $1 \leqslant i \leqslant 3$ (μ_1, μ_2, μ_3 não têm de ser distintos dois-a-dois). Como $\mathbf{A}$ é uma matriz com coeficientes reais, temos $p_{\mathbf{A}}(z) \in \mathbb{R}[z]$ e, portanto,

$$p_{\mathbf{A}}(\lambda) = 0 \qquad \Longleftrightarrow \qquad p_{\mathbf{A}}(\overline{\lambda}) = 0,$$

ou seja, $\overline{\lambda} \in \sigma(\mathbf{A})$. Como $|\overline{\lambda} - 12| = |\lambda - 12| \leqslant 4$, tem de ser $\overline{\lambda} = \lambda$ e, portanto, $\lambda \in \mathbb{R}$. De modo análogo, $\overline{\mu_1}, \overline{\mu_2}, \overline{\mu_3} \in \sigma(\mathbf{A})$, logo

$$\{\overline{\mu_1}, \overline{\mu_2}, \overline{\mu_3}\} = \{\mu_1, \mu_2, \mu_3\},$$

o que obriga a que $\mu_i \in \mathbb{R}$ para pelo menos um $1 \leqslant i \leqslant 3$. Como $\lambda \notin \{\mu_1, \mu_2, \mu_3\}$, concluímos que $\mathbf{A}$ tem pelo menos dois valores próprios reais.

7.4. (a) Suponhamos que $\mathbf{A}$ não é invertível. Então, $0 \in \sigma(\mathbf{A})$ e, portanto, pelo teorema de Geršgorin, existe $1 \leqslant i \leqslant n$ tal que

$$|a_{i,i}| = |0 - a_{i,i}| \leqslant \sum_{1 \leqslant j \neq i \leqslant n} |a_{i,j}| = r_i(\mathbf{A}),$$

o que não acontece.

(b) Para qualquer $\lambda \in \sigma(\mathbf{A})$, existe $1 \leqslant i \leqslant n$ tal que

$$|\lambda - a_{i,i}| \leqslant r_i(\mathbf{A}) < |a_{i,i}| = a_{i,i}.$$

Pondo $\lambda = \mu + i\nu$ com $\mu, \nu \in \mathbb{R}$, temos $\lambda - a_{i,i} = (\mu - a_{i,i}) + i\nu$, logo

$$|\mu - a_{i,i}| \leqslant \sqrt{(\mu - a_{i,i})^2 + \nu^2} = |\lambda - a_{i,i}| < a_{i,i}$$

e, portanto, $-a_{i,i} \leqslant \mu - a_{i,i} \leqslant a_{i,i}$, pelo que $\mu \geqslant 0$.

(c) Suponhamos que $\mathbf{A}$ é hermítica e que $a_{i,i} \in \mathbb{R}^+$ para qualquer $1 \leqslant i \leqslant n$. Então, $\sigma(\mathbf{A}) \subseteq \mathbb{R}$ e, portanto, $\lambda \in \mathbb{R}^+$ (pela alínea anterior). Como $\mathbf{A}$ é uma matriz normal, existe uma matriz unitária $\mathbf{U} \in \mathbb{C}^{n \times n}$ tal que $\mathbf{U}^*\mathbf{A}\mathbf{U}$ é diagonal e, portanto, existe uma base ortonormada $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ de $\mathbb{C}^{n \times 1}$ constituída por vectores próprios de $\mathbf{A}$. Para cada $1 \leqslant i \leqslant n$, seja $\lambda_i \in \mathbb{C}$ tal que $\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i$. Temos

$$\mathbf{v}_i^*\mathbf{A}\mathbf{v}_i = \langle \mathbf{v}_i | \mathbf{A}\mathbf{v}_i \rangle = \langle \mathbf{v}_i | \lambda_i \mathbf{v}_i \rangle = \lambda_i \langle \mathbf{v}_i | \mathbf{v}_i \rangle = \lambda_i \|\mathbf{v}_i\|^2 > 0.$$

Como $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ é uma base ortonormada de $\mathbb{C}^{n \times 1}$, concluímos que

$$\mathbf{v}^*\mathbf{A}\mathbf{v} = \langle \mathbf{v} | \mathbf{A}\mathbf{v} \rangle > 0, \qquad \mathbf{v} \in \mathbb{C}^{n \times 1}.$$

7.5. (a) Óbvia.

(b) É claro que $\mathbf{A}$ não tem diagonal estritamente dominante (porque $1 - \varepsilon < 1$). No entanto, $\det(\mathbf{A}) = 1 - (1 - \varepsilon) = \varepsilon \neq 0$, logo $\mathbf{A}$ é invertível.

7.6. Por hipótese, existe $1 \leqslant k \leqslant n$ tal que $|a_{k,k}| \geqslant r_k(\mathbf{A})$ e

$$|a_{i,i}| > r_i(\mathbf{A}), \qquad 1 \leqslant i \leqslant n,\ i \neq k.$$

Se $|a_{k,k}| > r_k(\mathbf{A})$, então $\mathbf{A}$ terá diagonal estritamente dominante e, portanto, $\mathbf{A}$ será invertível. Por outro lado, suponhamos que $|a_{k,k}| = r_k(\mathbf{A})$. Seja $\varepsilon \in \mathbb{R}^+$ e sejam $p_k = 1 + \varepsilon$ e

$$p_i = 1, \qquad 1 \leqslant i \leqslant n,\ i \neq k.$$

Temos
$$\frac{1}{p_k}\sum_{1\leqslant j\neq k\leqslant n} p_j|a_{k,j}| = \frac{r_k(\mathbf{A})}{1+\varepsilon}$$
e
$$\frac{1}{p_i}\sum_{1\leqslant j\neq i\leqslant n} p_j|a_{i,j}| = r_i(\mathbf{A})+\varepsilon a_{k,i}, \qquad 1\leqslant i\leqslant n,\ i\neq k.$$
Como $r_i(\mathbf{A})<|a_{i,i}|$, podemos escolher $\varepsilon\in\mathbb{R}^+$ tal que
$$r_i(\mathbf{A})+\varepsilon|a_{k,i}|<|a_{i,i}|, \qquad 1\leqslant i\leqslant n,\ i\neq k.$$
Com esta escolha, temos $0\notin\mathcal{G}(\mathbf{D}^{-1}\mathbf{A}\mathbf{D})$ para $\mathbf{D}=\begin{bmatrix} p_1 & \cdots & 0\\ \vdots & & \vdots\\ 0 & \cdots & p_n\end{bmatrix}$. Sendo assim, $0\notin\sigma(\mathbf{D}^{-1}\mathbf{A}\mathbf{D})=\sigma(\mathbf{A})$ e, portanto, $\mathbf{A}$ é invertível.

7.7. (a) Temos
$$\operatorname{int}\big(\mathcal{G}(\mathbf{A})\big)=\bigcup_{1\leqslant i\leqslant n}\operatorname{int}\big(\mathcal{G}_i(\mathbf{A})\big)=\bigcup_{1\leqslant i\leqslant n}\{z\in\mathbb{C}\colon |z-a_{i,i}|<r_i(\mathbf{A})\}$$
e, portanto,
$$\lambda\notin\operatorname{int}\big(\mathcal{G}(\mathbf{A})\big) \qquad\Longleftrightarrow\qquad |\lambda-a_{i,i}|\geqslant r_i(\mathbf{A}),\quad 1\leqslant i\leqslant n.$$
(b) Se $\lambda\in\operatorname{fr}\big(\mathcal{G}(\mathbf{A})\big)$, então $\lambda\notin\operatorname{int}\big(\mathcal{G}(\mathbf{A})\big)$ e, portanto,
$$\lambda\in\operatorname{int}\big(\mathcal{G}(\mathbf{A})\big) \qquad\Longleftrightarrow\qquad |\lambda-a_{i,i}|\geqslant r_i(\mathbf{A}),\quad 1\leqslant i\leqslant n.$$

(c) $\mathbf{A}$ terá diagonal dominante se e só se $|a_{i,i}|\geqslant r_i(\mathbf{A})$ para qualquer $1\leqslant i\leqslant n$, ou seja, se e só se $0\notin\operatorname{int}\big(\mathcal{G}(\mathbf{A})\big)$ (pela alínea (a)).

7.8. (a) Seja $1\leqslant k\leqslant n$ for tal que $|v_k|=\|\mathbf{v}\|_\infty$. Temos
$$\lambda v_k=(\lambda\mathbf{v})_k=(\mathbf{A}\mathbf{v})_k=\sum_{1\leqslant j\leqslant n}a_{k,j}v_j=a_{k,k}v_k-\sum_{1\leqslant j\neq k\leqslant n}a_{k,j}v_j$$
e, portanto,
$$\begin{aligned}|\lambda-a_{k,k}|\,\|\mathbf{v}\|_\infty=|\lambda-a_{k,k}|\,|v_k|=|\lambda v_k-a_{k,k}v_k|&=\Bigg|\sum_{1\leqslant j\neq k\leqslant n}a_{k,j}v_j\Bigg|\\ &\leqslant\sum_{1\leqslant j\neq k\leqslant n}|a_{k,j}|\,|v_j|\leqslant\sum_{1\leqslant j\neq k\leqslant n}|a_{k,j}|\,\|\mathbf{v}|_\infty=r_k(\mathbf{A})\|\mathbf{v}\|_\infty.\end{aligned}$$
Sendo assim, $|\lambda-a_{k,k}|\leqslant r_k(\mathbf{A})$ e, portanto, $|\lambda-a_{k,k}|=r_k(\mathbf{A})$ (pelo exercício anterior, porque $\lambda\notin\operatorname{int}\big(\mathcal{G}(\mathbf{A})\big)$).

(b) Pela alínea anterior, obtemos
$$|\lambda-a_{k,k}|\,\|\mathbf{v}\|_\infty=\sum_{1\leqslant j\neq k\leqslant n}|a_{k,j}|\,|v_j|=\sum_{1\leqslant j\neq k\leqslant n}|a_{k,j}|\,\|\mathbf{v}|_\infty=r_k(\mathbf{A})\|\mathbf{v}\|_\infty$$

e, portanto,
$$\sum_{1\leqslant j\neq k\leqslant n} |a_{k,j}|\big(\|\mathbf{v}\|_\infty - |v_j|\big) = 0.$$
Como todas as parcelas são não-negativas, concluímos que
$$|a_{k,j}|\,\big(\|\mathbf{v}\|_\infty - |v_j|\big) = 0, \qquad 1 \leqslant j \leqslant n,\ j \neq k.$$
Daqui, resulta que $\|\mathbf{v}\|_\infty - |v_j| = 0$ sempre que $1 \leqslant j \leqslant n$ for tal que $a_{k,j} \neq 0$.

7.9. Dizemos que uma matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$ tem a PROPRIEDADE SC se, para quaisquer $1 \leqslant i, j \leqslant n$, existirem $1 \leqslant k_1, \dots, k_t \leqslant n$ tais que $k_1 = i$, $k_t = j$ e $a_{k_s,k_{s+1}} \neq 0$ para qualquer $1 \leqslant s < t$. Sejam $\mathbf{A} \in \mathbb{C}^{n\times n}$, $\lambda \in \sigma(\mathbf{A})$ e $\mathbf{0} \neq \mathbf{v} \in \mathbb{C}^{n\times 1}$ tal que $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$. Sejam $\mathcal{G}_1(\mathbf{A}), \dots, \mathcal{G}_n(\mathbf{A})$ os discos de Grešgorin de $\mathbf{A}$, seja $\mathcal{G}(\mathbf{A}) = \mathcal{G}_1(\mathbf{A}) \cup \cdots \cup \mathcal{G}_n(\mathbf{A})$ e suponhamos que $\lambda \notin \operatorname{int}\big(\mathcal{G}(\mathbf{A})\big)$. Prove que, se $\mathbf{A}$ tiver a propriedade SC, então:

(a) $|\lambda - a_{i,i}| = r_i(\mathbf{A})$ para qualquer $1 \leqslant i \leqslant n$.

(b) $\|\mathbf{v}\|_\infty = |v_i|$ para qualquer $1 \leqslant i \leqslant n$.

(c) $\mathbf{A}$ é invertível sempre que $\mathbf{A}$ tiver diagonal dominante e existir $1 \leqslant i \leqslant n$ tal que $|a_{i,i}| > r_i(\mathbf{A})$.

(a) Seja $1 \leqslant k \leqslant n$ tal que $|v_k| = \|\mathbf{v}\|_\infty$. Então, pelo exercício anterior, temos $|\lambda - a_{k,k}| = r_k(\mathbf{A})$. Por outro lado, seja $1 \leqslant i \neq k \leqslant n$. Como $\mathbf{A} \in \mathbb{C}^{n\times n}$ tem a propriedade SC, existem $1 \leqslant k_1, \dots, k_t \leqslant n$ tais que $k_1 = k$, $k_t = i$ e $a_{k_s,k_{s+1}} \neq 0$ para qualquer $1 \leqslant s < t$. De novo, pelo exercício anterior, temos $\|\mathbf{v}\|_\infty = |v_{k_2}|$ (porque $a_{k,k_2} \neq 0$) e, portanto, $|\lambda - a_{k_2,k_2}| = r_{k_2}(\mathbf{A})$. Repetindo o argumento, concluímos que $\|\mathbf{v}\|_\infty = |v_{k_3}|$ (porque $a_{k_2,k_3} \neq 0$) e, portanto, $|\lambda - a_{k_3,k_3}| = r_{k_3}(\mathbf{A})$ e assim sucessivamente; na última etapa, concluímos que $\|\mathbf{v}\|_\infty = |v_i|$ e que $|\lambda - a_{i,i}| = r_i(\mathbf{A})$.

(b) Foi provada em (a).

(c) Suponhamos que $\mathbf{A}$ tem diagonal dominante e que $|a_{i,i}| > r_i(\mathbf{A})$ para algum $1 \leqslant i \leqslant n$. Se $\mathbf{A}$ não fosse invertível, então $0 \in \sigma(\mathbf{A})$ e, portanto, $|a_{i,i}| = r_i(\mathbf{A})$ (pela alínea (a)), o que não acontece.

RESOLUÇÃO DOS EXERCÍCIOS – FOLHA 8

8.1. (a) Seja $\mathbf{P} \in \mathbb{C}^{n\times n}$ uma matriz invertível tal que

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} \lambda_1\mathbf{I}_{m_1} & 0 & \cdots & 0 \\ 0 & \lambda_2\mathbf{I}_{m_2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_r\mathbf{I}_{m_r} \end{bmatrix}, \qquad m_i = \text{m.a.}(\lambda_i),\ 1 \leqslant i \leqslant r.$$

Sejam $\mathbf{X}_i, \mathbf{Y}_i \in \mathbb{k}^{n\times m_i}$, para $1 \leqslant i \leqslant r$, tais que

$$\mathbf{P} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_r \end{bmatrix} \qquad \text{e} \qquad \mathbf{P}^{-1} = \begin{bmatrix} \mathbf{Y}_1 & \mathbf{Y}_2 & \cdots & \mathbf{Y}_r \end{bmatrix}^{\mathsf{T}},$$

de modo que $\mathbf{G}_i = \mathbf{X}_i\mathbf{Y}_i^{\mathsf{T}}$ para qualquer $1 \leqslant i \leqslant r$ (ver Exercício 5.1). Então,

$$\begin{aligned} f(\mathbf{A}) &= \mathbf{P} \begin{bmatrix} f(\lambda_1)\mathbf{I}_{m_1} & 0 & \cdots & 0 \\ 0 & f(\lambda_2)\mathbf{I}_{m_2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & f(\lambda_r)\mathbf{I}_{m_r} \end{bmatrix} \mathbf{P}^{-1} \\ &= \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_r \end{bmatrix} \begin{bmatrix} f(\lambda_1)\mathbf{I}_{m_1} & 0 & \cdots & 0 \\ 0 & f(\lambda_2)\mathbf{I}_{m_2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & f(\lambda_r)\mathbf{I}_{m_r} \end{bmatrix} \begin{bmatrix} \mathbf{Y}_1^{\mathsf{T}} \\ \mathbf{Y}_2^{\mathsf{T}} \\ \cdots \\ \mathbf{Y}_r^{\mathsf{T}} \end{bmatrix} \\ &= f(\lambda_1)\mathbf{X}_1\mathbf{Y}_1^{\mathsf{T}} + f(\lambda_2)\mathbf{X}_2\mathbf{Y}_2^{\mathsf{T}} + \cdots + f(\lambda_r)\mathbf{X}_r\mathbf{Y}_r^{\mathsf{T}} \\ &= f(\lambda_1)\mathbf{G}_1 + f(\lambda_2)\mathbf{G}_2 + \cdots + f(\lambda_r)\mathbf{G}_r. \end{aligned}$$

(b) Consideremos o polinómio interpolador de Lagrange

$$p(x) = \sum_{1\leqslant i\leqslant r} \frac{f(\lambda_i)}{\prod_{1\leqslant j\neq i\leqslant r}(\lambda_i - \lambda_j)} \prod_{1\leqslant j\neq i\leqslant r} (x - \lambda_j);$$

este polinómio tem coeficientes complexos (isto é, $p(x) \in \mathbb{C}[x]$), tem grau $\leqslant r-1$ e verifica $p(\lambda_i) = f(\lambda_i)$ para qualquer $1 \leqslant i \leqslant r$. Sendo assim, pela alínea anterior, temos

$$f(\mathbf{A}) = f(\lambda_1)\mathbf{G}_1 + \cdots + f(\lambda_r)\mathbf{G}_r = p(\lambda_1)\mathbf{G}_1 + \cdots + p(\lambda_r)\mathbf{G}_r = p(\mathbf{A}).$$

(c) Sabemos que $|\lambda_i| \leqslant \rho(\mathbf{A}) < R$ para qualquer $1 \leqslant i \leqslant r$, de modo que a série $\sum_{k\geqslant 0} a_k\lambda_i^k$ é convergente com

$$f(\lambda_i) = \sum_{k\geqslant 0} a_k\lambda_i^k, \qquad 1 \leqslant i \leqslant r.$$

Sendo assim, obtemos

$$\begin{aligned} f(\mathbf{A}) &= f(\lambda_1)\mathbf{G}_1 + \cdots + f(\lambda_r)\mathbf{G}_r \\ &= \Big(\sum_{k\geqslant 0} a_k\lambda_1^k\Big)\mathbf{G}_1 + \cdots + \Big(\sum_{k\geqslant 0} a_k\lambda_r^k\Big)\mathbf{G}_r \\ &= \sum_{k\geqslant 0} a_k\big(\lambda_1^k\mathbf{G}_1 + \cdots + \lambda_r^k\mathbf{G}_r\big) = \sum_{k\geqslant 0} a_k\mathbf{A}^k \end{aligned}$$

(uma vez que $\mathbf{A}^k = \lambda_1^k\mathbf{G}_1 + \cdots + \lambda_r^k\mathbf{G}_r$ para qualquer $k \in \mathbb{N}_0$).

(d) Para cada $1 \leqslant i \leqslant r$, consideramos o polinómio

$$g_i(t) = \frac{1}{\prod_{1\leqslant j\neq i\leqslant r}(\lambda_i - \lambda_j)} \prod_{1\leqslant j\neq i\leqslant r}(x - \lambda_j) \in \mathbb{C}[x].$$

Temos

$$g_i(\lambda_j) = \begin{cases} 1, & \text{se } j = i, \\ 0, & \text{se } j \neq i, \end{cases} \qquad 1 \leqslant j \leqslant r,$$

de maneira que

$$g_i(\mathbf{A}) = \frac{1}{\prod_{1\leqslant j\neq i\leqslant r}(\lambda_i - \lambda_j)} \prod_{1\leqslant j\neq i\leqslant r}(\mathbf{A} - \lambda_j\mathbf{I}_n) = g_i(\lambda_1)\mathbf{G}_1 + \cdots + g_i(\lambda_r)\mathbf{G}_r = \mathbf{G}_i$$

para qualquer $1 \leqslant i \leqslant r$ (na primeira igualdade, usamos a alínea anterior e, na segunda igualdade, usamos a alínea (a)).

(e) Basta observar que os projectores espectrais de $\mathbf{A}$ são univocamente determinados (pela alínea anterior), de modo que a matriz $f(\mathbf{A})$ é univocamente determinada pelas imagens $f(\lambda_1), \ldots, f(\lambda_r)$ e pelos projectores espectrais $\mathbf{G}_1, \ldots, \mathbf{G}_r$.

8.2. Temos $\rho(\mathbf{I}_n - \mathbf{A}) = 0 < 1$, logo

$$\mathbf{A}^{-1} = \sum_{k\geqslant 0}(\mathbf{I}_n - \mathbf{A})^k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 2 & -1 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & -6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & -7 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}.$$

8.3. A matriz $\mathbf{A} = \begin{bmatrix} -\pi/2 & \pi/2 \\ \pi/2 & -\pi/2 \end{bmatrix}$ tem valores próprios 0 e $-\pi$, logo $\mathbf{A}$ é diagonalizável e, portanto, podemos usar o Exercício 8.1:

$$\cos(\mathbf{A}) = \cos(0)\mathbf{G}_1 + \cos(-\pi)\mathbf{G}_2 = \mathbf{G}_1 - \mathbf{G}_2$$

onde $\mathbf{G}_1$ e $\mathbf{G}_2$ são os projectores espectrais de $\mathbf{A}$. Pela alínea (d) do Exercício 8.1, temos

$$\mathbf{G}_1 = \frac{1}{\pi}(\mathbf{A}+\pi\mathbf{I}_2) = \frac{1}{\pi}\begin{bmatrix} \pi/2 & \pi/2 \\ \pi/2 & \pi/2 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

$$\mathbf{G}_2 = -\frac{1}{\pi}\mathbf{A} = -\frac{1}{\pi}\begin{bmatrix} -\pi/2 & \pi/2 \\ \pi/2 & -\pi/2 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix},$$

logo

$$\cos(\mathbf{A}) = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

8.4. Os valores próprios de $\mathbf{A}$ são 0 e $-(\alpha+\beta)$, de modo que $\mathbf{A}$ é diagonalizável (porque $\alpha+\beta \neq 0$) e, portanto, considerando a função $z \mapsto e^{tz}$, $z \in \mathbb{C}$, obtemos

$$e^{t\mathbf{A}} = e^0\mathbf{G}_1 + e^{-t(\alpha+\beta)}\mathbf{G}_2 = \mathbf{G}_1 + e^{-t(\alpha+\beta)}\mathbf{G}_2$$

onde $\mathbf{G}_1$ e $\mathbf{G}_2$ são os projectores espectrais de $\mathbf{A}$. Pela alínea (d) do Exercício 8.1, temos

$$\mathbf{G}_1 = \frac{1}{\alpha+\beta}(\mathbf{A}+(\alpha+\beta)\mathbf{I}_2) = \frac{1}{\alpha+\beta}\begin{bmatrix} \beta & \beta \\ \alpha & \alpha \end{bmatrix},$$

$$\mathbf{G}_2 = -\frac{1}{\alpha+\beta}\mathbf{A} = -\frac{1}{\alpha+\beta}\begin{bmatrix} -\alpha & \beta \\ \alpha & -\beta \end{bmatrix},$$

logo

$$\begin{aligned} e^{t\mathbf{A}} &= \frac{1}{\alpha+\beta}\begin{bmatrix} \beta & \beta \\ \alpha & \alpha \end{bmatrix} - \frac{e^{-t(\alpha+\beta)}}{\alpha+\beta}\begin{bmatrix} -\alpha & \beta \\ \alpha & -\beta \end{bmatrix} \\ &= \frac{1}{\alpha+\beta}\left(\begin{bmatrix} \beta- & \beta \\ \alpha & \alpha \end{bmatrix} - e^{-t(\alpha+\beta)}\begin{bmatrix} -\alpha & \beta \\ \alpha & -\beta \end{bmatrix}\right). \end{aligned}$$

8.5. Seja $\mathbf{A} \in \mathbb{C}^{n\times n}$ uma matriz diagonalizável com decomposição espectral $\mathbf{A} = \lambda_1\mathbf{G}_1 + \cdots + \lambda_r\mathbf{G}_r$ onde $\sigma(\mathbf{A}) = \{\lambda_1,\dots,\lambda_r\}$ e onde $\mathbf{G}_1,\dots,\mathbf{G}_r$ são os projectores espectrais de $\mathbf{A}$. Então,

$$\begin{aligned} \cos^2(\mathbf{A}) + \operatorname{sen}^2(\mathbf{A}) &= \sum_{1\leqslant i\leqslant r}\cos^2(\lambda_i)\mathbf{G}_i + \sum_{1\leqslant i\leqslant r}\operatorname{sen}^2(\lambda_i)\mathbf{G}_i \\ &= \sum_{1\leqslant i\leqslant r}\big(\cos^2(\lambda_i)+\operatorname{sen}^2(\lambda_i)\big)\mathbf{G}_i \\ &= \sum_{1\leqslant i\leqslant r}\mathbf{G}_i = \mathbf{I}_n. \end{aligned}$$

Por outro lado, suponhamos que a matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$ não é diagonalizável. Nesta situação, temos

$$\cos(\mathbf{A}) = \sum_{k\geqslant 0}\frac{(-1)^k}{(2k)!}\mathbf{A}^{2k} \qquad \text{e} \qquad \operatorname{sen}(\mathbf{A}) = \sum_{k\geqslant 0}\frac{(-1)^k}{(2k+1)!}\mathbf{A}^{2k+1}.$$

Para concluir que $\cos^2(\mathbf{A}) + \operatorname{sen}^2(\mathbf{A}) = \mathbf{I}_n$, basta notar que $\cos^2(z)+\operatorname{sen}^2(z) = 1$ para qualquer $z \in \mathbb{C}$.

8.6. (a) Ponhamos $\mathbf{A}_k = \big(a_{i,j}^{(k)}\big)$ para $k \in \mathbb{N}_0$, de modo que

$$\mathbf{F}(z) = \begin{bmatrix} \sum_{k\geqslant 0} a_{1,1}^{(k)} z^k & \cdots & \sum_{k\geqslant 0} a_{1,2}^{(k)} z^k \\ \vdots & & \vdots \\ \sum_{k\geqslant 0} a_{n,1}^{(k)} z^k & \cdots & \sum_{k\geqslant 0} a_{n,n}^{(k)} z^k \end{bmatrix}.$$

Sendo assim, para quaisquer $1 \leqslant i, j \leqslant n$, a (i,j)-ésima entrada da matriz $\mathbf{F}(z)$ é dada por uma função $\mathbf{F}_{i,j} \colon \mathcal{B} \to \mathbb{C}$ que admite o desenvolvimento em série de potências

$$\mathbf{F}_{i,j}(z) = \sum_{k\geqslant 0} a_{i,j}^{(k)} z^k, \qquad z \in \mathcal{B}.$$

Da Análise Complexa, sabemos que uma função nestas condições é diferenciável com

$$\mathbf{F}'_{i,j}(z) = \sum_{k\geqslant 1} k a_{i,j}^{(k)} z^{k-1}, \qquad 1 \leqslant i, j \leqslant n,$$

e, portanto,

$$\mathbf{F}'(z) = \begin{bmatrix} \mathbf{F}'_{1,1}(z) & \cdots & \mathbf{F}'_{1,n}(z) \\ \vdots & & \vdots \\ \mathbf{F}'_{n,1}(z) & \cdots & \mathbf{F}'_{n,n}(z) \end{bmatrix} = \begin{bmatrix} \sum_{k\geqslant 1} k a_{1,1}^{(k)} z^{k-1} & \cdots & \sum_{k\geqslant 1} k a_{1,2}^{(k)} z^{k-1} \\ \vdots & & \vdots \\ \sum_{k\geqslant 1} k a_{n,1}^{(k-1)} z^{k-1} & \cdots & \sum_{k\geqslant 1} k a_{n,n}^{(k)} z^{k-1} \end{bmatrix} = \sum_{k\geqslant 1} k \mathbf{A}_k z^k.$$

Repetindo o processo, concluímos que $\mathbf{F}(z)$ é indefinidamente diferenciável com

$$\mathbf{F}^{(t)}(z) = \sum_{k\geqslant t} k(k-1)\cdots(k-t+1)\mathbf{A}_k z^{k-t} = \sum_{k\geqslant t} \frac{k!}{(k-t)!}\,\mathbf{A}_k z^{k-t}, \qquad t \in \mathbb{N}.$$

Além disso, para qualquer $t \in \mathbb{N}$, temos

$$\mathbf{F}(t)(0) = t!\,\mathbf{A}_t \quad\iff\quad \mathbf{A}_t = \frac{1}{t!}\,\mathbf{F}^{(t)}(0).$$

(b) Suponhamos que $\mathbf{A} \in \mathbb{C}^{n\times n}$ é tal que $\mathbf{F}'(z) = \mathbf{A}\mathbf{F}(z)$. Então, $\mathbf{F}^{(k+1)}(z) = \mathbf{A}\mathbf{F}^{(k)}(z)$, logo $\mathbf{F}(k+1)(0) = \mathbf{A}\mathbf{F}^{(k)}(0)$ e, portanto,

$$\mathbf{A}_{k+1} = \frac{1}{(k+1)!}\,\mathbf{F}^{(k+1)}(0) = \frac{1}{(k+1)!}\,\mathbf{A}\mathbf{F}^{(k)}(0) = \frac{1}{k+1}\,\mathbf{A}\mathbf{A}_k, \qquad k \in \mathbb{N}_0.$$

Daqui, resulta que

$$\mathbf{A}_k = \frac{1}{k}\,\mathbf{A}\mathbf{A}_{k-1} = \frac{1}{k(k-1)}\,\mathbf{A}^2\mathbf{A}_{k-2} = \cdots = \frac{1}{k!}\,\mathbf{A}^k\mathbf{A}_0, \qquad k \in \mathbb{N},$$

logo

$$\mathbf{F}(t) = \sum_{k\geqslant 0} \mathbf{A}_k z^k = \sum_{k\geqslant 0} \frac{1}{k!}\,\mathbf{A}_0\mathbf{A}^k z^k = \mathbf{A}_0 \sum_{k\geqslant 0} \frac{1}{k!}\,(z\mathbf{A})^k = \mathbf{A}_0 e^{z\mathbf{A}}.$$

Como se queria.

8.7. Usando séries de potências, é fácil verificar que $\big(e^{z\mathbf{A}}\big)' = \mathbf{A}e^{z\mathbf{A}}$. Por conseguinte, usando as regras de derivação usuais, obtemos

$$\mathbf{F}'(z) = \big(e^{z(\mathbf{A}+\mathbf{B})}\big)' - \big(e^{z\mathbf{A}}\big)'e^{z\mathbf{B}} - e^{z\mathbf{A}}\big(e^{z\mathbf{B}}\big)' = (\mathbf{A}+\mathbf{B})e^{z(\mathbf{A}+\mathbf{B})} - \mathbf{A}e^{z\mathbf{A}}e^{z\mathbf{B}} - e^{z\mathbf{A}}\big(\mathbf{B}e^{z\mathbf{B}}\big).$$

Como $\mathbf{AB} = \mathbf{BA}$, temos

$$e^{\mathbf{A}}\mathbf{B} = \bigg(\sum_{k\geqslant 0}\frac{1}{k!}\,\mathbf{A}^k\bigg)\mathbf{B} = \sum_{k\geqslant 0}\frac{1}{k!}\,(\mathbf{A}^k\mathbf{B}) = \sum_{k\geqslant 0}\frac{1}{k!}\,(\mathbf{B}\mathbf{A}^k) = \mathbf{B}\bigg(\sum_{k\geqslant 0}\frac{1}{k!}\,\mathbf{A}^k\bigg) = \mathbf{B}e^{z\mathbf{A}}$$

e, portanto,

$$\begin{aligned}\mathbf{F}'(z) &= (\mathbf{A}+\mathbf{B})e^{z(\mathbf{A}+\mathbf{B})} - \mathbf{A}e^{z\mathbf{A}}e^{z\mathbf{B}} - \mathbf{B}e^{z\mathbf{A}}e^{z\mathbf{B}}\\ &= (\mathbf{A}+\mathbf{B})\big(e^{z(\mathbf{A}+\mathbf{B})} - e^{z\mathbf{A}}e^{z\mathbf{B}}\big) = (\mathbf{A}+\mathbf{B})\mathbf{F}(z)\end{aligned}$$

para qualquer $z \in \mathbb{C}$.

Pelo exercício anterior, existe uma matriz $\mathbf{C} \in \mathbb{C}^{n\times n}$ tal que

$$\mathbf{F}(z) = \mathbf{C}e^{z(\mathbf{A}+\mathbf{B})}, \qquad z \in \mathbb{C}.$$

Como $e^{\mathbf{0}} = \mathbf{I}_n$, concluímos que

$$\mathbf{C} = \mathbf{C}e^{\mathbf{0}} = \mathbf{C}e^{0\cdot\mathbf{A}} = \mathbf{F}(0) = e^{\mathbf{0}} - e^{\mathbf{0}}e^{\mathbf{0}} = \mathbf{I}_n - \mathbf{I}_n = \mathbf{0}$$

e, portanto, $\mathbf{F}(z) = \mathbf{0}$. Em particular,

$$e^{\mathbf{A}+\mathbf{B}} = e^{\mathbf{A}+\mathbf{B}} - e^{\mathbf{A}}e^{\mathbf{B}} + e^{\mathbf{A}}e^{\mathbf{B}} = \mathbf{F}(1) + e^{\mathbf{A}}e^{\mathbf{A}} = e^{\mathbf{A}}e^{\mathbf{B}},$$

como queríamos.

8.8. Por um lado, como $\mathbf{A}(-\mathbf{A}) = (-\mathbf{A})\mathbf{A}$, o exercício anterior garante que

$$e^{\mathbf{A}}e^{-\mathbf{A}} = e^{\mathbf{A}-\mathbf{A}} = e^{\mathbf{0}} = \mathbf{I}_n,$$

logo $e^{\mathbf{A}}$ é invertível com inversa $(e^{\mathbf{A}})^{-1} = e^{-\mathbf{A}}$.

Por outro lado, temos

$$(e^{\mathbf{A}})^* = \bigg(\sum_{k\geqslant 0}\frac{1}{k!}\mathbf{A}^k\bigg)^* = \sum_{k\geqslant 0}\frac{1}{k!}(\mathbf{A}^k)^* = \sum_{k\geqslant 0}\frac{1}{k!}(\mathbf{A}^*)^k = e^{\mathbf{A}^*}$$

e, portanto, $(e^{\mathbf{A}})^* = e^{\mathbf{A}^*} = e^{-\mathbf{A}} = (e^{\mathbf{A}})^{-1}$, o que garante que $e^{\mathbf{A}}$ é unitária.

8.9. (a) Temos

$$e^{\mathbf{P}^{-1}\mathbf{AP}} = \sum_{k\geqslant 0}\frac{1}{k!}(\mathbf{P}^{-1}\mathbf{AP})^k = \sum_{k\geqslant 0}\frac{1}{k!}(\mathbf{P}^{-1}\mathbf{A}^k\mathbf{P}) = \mathbf{P}^{-1}\bigg(\sum_{k\geqslant 0}\frac{1}{k!}\mathbf{A}^k\bigg)\mathbf{P} = \mathbf{P}^{-1}e^{\mathbf{A}}\mathbf{P};$$

notemos que, como a correspondência $\mathbf{X} \mapsto \mathbf{P}^{-1}\mathbf{XP}$ define uma função contínua de $\mathbb{C}^{n\times n}$ em $\mathbb{C}^{n\times n}$, tem de ser

$$\lim_{k\to\infty}\sum_{0\leqslant i\leqslant k}\frac{1}{i!}(\mathbf{P}^{-1}\mathbf{A}^i\mathbf{P}) = \lim_{k\to\infty}\mathbf{P}^{-1}\bigg(\sum_{0\leqslant i\leqslant k}\frac{1}{i!}\,\mathbf{A}^i\bigg)\mathbf{P} = \mathbf{P}^{-1}\bigg(\lim_{k\to\infty}\sum_{0\leqslant i\leqslant k}\frac{1}{i!}\,\mathbf{A}^i\bigg)\mathbf{P}.$$

(b) Pelo teorema da decomposição de Schur, existe uma matriz unitária $\mathbf{U} \in \mathbb{C}^{n\times n}$ tal que $\mathbf{T} = \mathbf{U}^*\mathbf{A}\mathbf{U}$ é uma matriz triangular superior. Como $\mathbf{U}$ é invertível e $\mathbf{U}^{-1} = \mathbf{U}^*$, a alínea anterior garante que

$$\mathbf{U}^* e^{\mathbf{A}} \mathbf{U} = e^{\mathbf{U}^*\mathbf{A}\mathbf{U}} = e^{\mathbf{T}} = \sum_{k\geqslant 0} \frac{1}{k!}\,\mathbf{T}^k.$$

Como $\mathbf{T}$ é triangular superior, qualquer potência de $\mathbf{T}$ também é triangular superior e, portanto, $e^{\mathbf{T}}$ também é triangular superior.

Para terminar, como $\mathbf{T} = \mathbf{U}^*\mathbf{A}\mathbf{U} = \mathbf{U}^{-1}\mathbf{A}\mathbf{U}$, temos $\sigma(\mathbf{T}) = \sigma(\mathbf{A})$ e, portanto, as entradas da diagonal de $\mathbf{T}$ são exactamente os valores próprios $\lambda_1, \dots, \lambda_n$ de $\mathbf{A}$. Analogamente, como $e^{\mathbf{T}} = \mathbf{U}^* e^{\mathbf{A}}\mathbf{U} = \mathbf{U}^{-1} e^{\mathbf{A}}\mathbf{U}$, as entradas da diagonal de $e^{\mathbf{T}}$ são os valores próprios de $e^{\mathbf{A}}$. Pondo $\mathbf{T} = \begin{bmatrix} \lambda_1 & \star & \cdots & \star \\ 0 & \lambda_2 & \cdots & \star \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$, temos $\mathbf{T}^k = \begin{bmatrix} \lambda_1^k & \star & \cdots & \star \\ 0 & \lambda_2^k & \cdots & \star \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{bmatrix}$ para qualquer $k \in \mathbb{N}_0$, logo

$$e^{\mathbf{T}} = \sum_{k\geqslant 0} \frac{1}{k!}\,\mathbf{T}^k = \begin{bmatrix} \sum_{k\geqslant 0} \frac{1}{k!}\,\lambda_1^k & \star & \cdots & \star \\ 0 & \sum_{k\geqslant 0} \frac{1}{k!}\,\lambda_2^k & \cdots & \star \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sum_{k\geqslant 0} \frac{1}{k!}\,\lambda_n^k \end{bmatrix} = \begin{bmatrix} e^{\lambda_1} & \star & \cdots & \star \\ 0 & e^{\lambda_2} & \cdots & \star \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n} \end{bmatrix}.$$

Em conclusão, os valores próprios de $e^{\mathbf{T}}$ e de $e^{\mathbf{A}}$ são $e^{\lambda_1}, \dots, e^{\lambda_n}$; além disso, tem de ser $\text{m.a.}(e^{\lambda_i}) = \text{m.a.}(\lambda_i)$ para todo $1 \leqslant i \leqslant n$.

(c) Por um lado, temos

$$\det(e^{\mathbf{A}}) = \det(\mathbf{U}^{-1} e^{\mathbf{A}}\mathbf{U}) = \det(e^{\mathbf{T}}) = e^{\lambda_1} e^{\lambda_2} \cdots e^{\lambda_n} = e^{\lambda_1+\lambda_2+\cdots+\lambda_n}.$$

Por outro lado,

$$\operatorname{tr}(\mathbf{A}) = \operatorname{tr}(\mathbf{U}\mathbf{T}\mathbf{U}^{-1}) = \operatorname{tr}(\mathbf{T}\mathbf{U}^{-1}\mathbf{U}) = \operatorname{tr}(\mathbf{T}) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$$

e, portanto, $\det(e^{\mathbf{A}}) = e^{\operatorname{tr}(\mathbf{A})}$.

RESOLUÇÃO DOS EXERCÍCIOS – FOLHA 9

9.1. Seja $\lambda \in \sigma(\mathbf{A})$ e seja $\mathbf{0} \neq \mathbf{v} \in \mathbb{C}^{n\times 1}$ tal que $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$. Então,

$$\mathbf{0} = \mathbf{A}^k\mathbf{v} = \mathbf{A}^{k-1}(\mathbf{A}\mathbf{v}) = \lambda\mathbf{A}^{k-1}\mathbf{v} = \lambda\mathbf{A}^{k-2}(\mathbf{A}\mathbf{v}) = \lambda^2\mathbf{A}^{k-2}\mathbf{v} = \cdots = \lambda^{k-1}\mathbf{A}\mathbf{v} = \lambda^k\mathbf{v},$$

logo $\lambda^k = 0$ (porque $\mathbf{v} \neq \mathbf{0}$) e, portanto, $\lambda = 0$.

Pelo teorema da decomposição de Schur, existe uma matriz unitária $\mathbf{U} \in \mathbb{C}^{n\times n}$ tal que $\mathbf{T} = \mathbf{U}^*\mathbf{A}\mathbf{A}$ é triangular superior. Como $\sigma(\mathbf{T}) = \sigma(\mathbf{A}) = \{0\}$, tem de ser $t_{1,1} = t_{2,2} = \cdots = t_{n,n} = 0$, logo $\mathbf{T}$ é estritamente triangular superior. Sendo assim, existe uma matriz invertível $\mathbf{Q} \in \mathbb{C}^{n\times n}$ tal que

$$\mathbf{Q}^{-1}\mathbf{T}\mathbf{Q} = \begin{bmatrix} \mathbf{J}_{n_1}(0) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{n_2}(0) & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{J}_{n_r}(0) \end{bmatrix}$$

onde $n_1, \ldots, n_r \in \mathbb{N}$ são tais que $n_1 \geqslant n_2 \geqslant \cdots \geqslant n_r$ e $n_1 + n_2 + \cdots + n_r = n$, pelo que basta considerar $\mathbf{P} = \mathbf{U}\mathbf{Q}$.

9.2. (a) Temos

- $\mathcal{V}_k = \mathcal{R}(\mathbf{A}^k) \cap \mathcal{N}(\mathbf{A}) = \mathcal{R}(\mathbf{0}) \cap \mathcal{N}(\mathbf{A}) = \{\mathbf{0}\} \cap \mathcal{N}(\mathbf{A}) = \{\mathbf{0}\}$.
- $\mathcal{V}_0 = \mathcal{R}(\mathbf{A}^0) \cap \mathcal{N}(\mathbf{A}) = \mathcal{R}(\mathbf{I}_n) \cap \mathcal{N}(\mathbf{A}) = \mathbb{C}^{n\times n} \cap \mathcal{N}(\mathbf{A}) = \mathcal{N}(\mathbf{A})$.

Por outro lado, se $\mathbf{v} \in \mathcal{R}(\mathbf{A}^{k-1})$, então $\mathbf{v} = \mathbf{A}^{k-1}\mathbf{w}$ para algum $\mathbf{w} \in \mathbb{C}^{n\times 1}$, logo

$$\mathbf{A}\mathbf{v} = \mathbf{A}(\mathbf{A}^{k-1}\mathbf{w}) = \mathbf{A}^k\mathbf{w} = \mathbf{0}\mathbf{w} = \mathbf{0}$$

e, portanto, $\mathbf{v} \in \mathcal{N}(\mathbf{A})$. Sendo assim, temos $\mathcal{R}(\mathbf{A}^{k-1}) \subseteq \mathcal{N}(\mathbf{A})$, o que é equivalente a

$$\mathcal{V}_{k-1} = \mathcal{R}(\mathbf{A}^{k-1}) \cap \mathcal{N}(\mathbf{A}) = \mathcal{R}(\mathbf{A}^{k-1}).$$

(b) Seja $1 \leqslant i \leqslant k$ e seja $\mathbf{v} \in \mathcal{V}_i = \mathcal{R}(\mathbf{A}^i) \cap \mathcal{N}(\mathbf{A})$. Como $\mathbf{v} \in \mathcal{R}(\mathbf{A}^i)$, existe $\mathbf{w} \in \mathbb{C}^{n\times 1}$ tal que $\mathbf{v} = \mathbf{A}^i\mathbf{w}$. Sendo assim, $\mathbf{v} = \mathbf{A}^{i-1}(\mathbf{A}\mathbf{w}) \in \mathcal{R}(\mathbf{A}^{i-1})$ e, portanto, $\mathbf{v} \in \mathcal{R}(\mathbf{A}^{i-1}) \cap \mathcal{N}(\mathbf{A}) = \mathcal{V}_{i-1}$.

(c) Seja $1 \leqslant i \leqslant k-1$ qualquer. Pelo Teorema 1.4, temos

$$r(\mathbf{A}^{i+1}) = r(\mathbf{A}\mathbf{A}^i) = r(\mathbf{A}^i) - \dim\big(\mathcal{R}(\mathbf{A}^i) \cap \mathcal{N}(\mathbf{A})\big) = r(\mathbf{A}^i) - \dim\mathcal{V}_i$$

e, portanto, $\dim\mathcal{V}_i = r(\mathbf{A}^i) - r(\mathbf{A}^{i+1})$.

9.3. (a) Seja $1 \leqslant i \leqslant k$. Por definição, $\mathcal{S}_{k-1} \cup \cdots \cup \mathcal{S}_{i+1}$ é uma base de $\mathcal{V}_{i+1}$ e $\mathcal{S}_{k-1} \cup \cdots \cup \mathcal{S}_{i+1} \cup \mathcal{S}_i$ é uma base de $\mathcal{V}_i$, logo

$$\begin{aligned}\#(\mathcal{S}_i) = \dim \mathcal{V}_i - \dim \mathcal{V}_{i+1} &= r(\mathbf{A}^i) - r(\mathbf{A}^{i+1}) - r(\mathbf{A}^{i+1}) + r(\mathbf{A}^{i+2}) \\ &= r(\mathbf{A}^i) - 2r(\mathbf{A}^{i+1}) + r(\mathbf{A}^{i+2}).\end{aligned}$$

(b) Seja $1 \leqslant i \leqslant k$ e seja $\mathbf{s} \in \mathcal{S}_i$. Então, $\mathbf{s} \in \mathcal{V}_i = \mathcal{R}(\mathbf{A}^i) \cap \mathcal{N}(\mathbf{A}) \subseteq \mathcal{R}(\mathbf{A}^i)$ e, portanto, existe $\mathbf{v} \in \mathbb{C}^{n\times 1}$ tal que $\mathbf{A}^i\mathbf{v} = \mathbf{s}$.

(c) Sejam $1 \leqslant i \leqslant k$, $\mathbf{s} \in \mathcal{S}_i$ e $\mathbf{v} \in \mathbb{C}^{n\times 1}$ tal que $\mathbf{A}^i\mathbf{v} = \mathbf{s}$. Sejam $\alpha_1, \ldots, \alpha_{i+1} \in \mathbb{C}$ tais que $\alpha_1\mathbf{v} + \alpha_2\mathbf{A}\mathbf{v} + \cdots + \alpha_{i+1}\mathbf{A}^i\mathbf{v} = \mathbf{0}$. Como $\mathbf{s} \in \mathcal{S}_i \subseteq \mathcal{V}_i \subseteq \mathcal{N}(\mathbf{A})$, temos $\mathbf{A}^{i+1}\mathbf{v} = \mathbf{A}(\mathbf{A}^i\mathbf{v}) = \mathbf{A}\mathbf{s} = \mathbf{0}$ e, portanto, $\mathbf{A}^k\mathbf{v} = \mathbf{0}$ para qualquer $k \geqslant i+1$. Sendo assim, uma vez que $\mathbf{s} \neq \mathbf{0}$ (porque pertence a uma base), obtemos

$$\begin{cases} \mathbf{0} = \mathbf{A}^i\big(\alpha_1\mathbf{v} + \alpha_2\mathbf{A}\mathbf{v} + \cdots + \alpha_{i+1}\mathbf{A}^i\mathbf{v}\big) = \alpha_1\mathbf{s} \quad \Longrightarrow \quad \alpha_1 = 0, \\ \mathbf{0} = \mathbf{A}^{i-1}\big(\alpha_1\mathbf{v} + \alpha_2\mathbf{A}\mathbf{v} + \cdots + \alpha_{i+1}\mathbf{A}^i\mathbf{v}\big) = \alpha_2\mathbf{s} \quad \Longrightarrow \quad \alpha_2 = 0, \\ \cdots \\ \mathbf{0} = \mathbf{A}\big(\alpha_1\mathbf{v} + \alpha_2\mathbf{A}\mathbf{v} + \cdots + \alpha_{i+1}\mathbf{A}^i\mathbf{v}\big) = \alpha_i\mathbf{s} \quad \Longrightarrow \quad \alpha_i = 0, \\ \mathbf{0} = \alpha_1\mathbf{v} + \alpha_2\mathbf{A}\mathbf{v} + \cdots + \alpha_{i+1}\mathbf{A}^i = \alpha_{i+1}\mathbf{s} \quad \Longrightarrow \quad \alpha_{i+1} = 0, \end{cases}$$

provando que $\mathcal{J}_\mathbf{s} = \{\mathbf{A}^i\mathbf{v}, \ldots, \mathbf{A}\mathbf{v}, \mathbf{v}\}$ é um subconjunto linearmente independente.

Por outro lado, temos

$$\mathbf{A}\begin{bmatrix}\mathbf{A}^i\mathbf{v} & \cdots & \mathbf{A}\mathbf{v} & \mathbf{v}\end{bmatrix} = \begin{bmatrix}\mathbf{0} & \mathbf{A}^i\mathbf{v} & \cdots & \mathbf{A}\mathbf{v}\end{bmatrix} = \begin{bmatrix}\mathbf{A}^i\mathbf{v} & \cdots & \mathbf{A}\mathbf{v} & \mathbf{v}\end{bmatrix}\mathbf{J}_{i+1}(0).$$

(d) Análoga à alínea anterior: dada uma combinação linear nula dos vectores de $\mathcal{J}$, multiplicamos sucessivamente por potências adequadas de $\mathbf{A}$ de forma a reduzir a uma combinação linear nula de vectores de $\mathcal{S}$ para concluir que os escalares respectivos têm de ser nulos (porque $\mathcal{S}$ é linearmente independente). Para justificarmos que é uma base, observamos que qualquer cadeia construída a partir de um vector $\mathbf{s} \in \mathcal{S}_i$ tem $i+1$ vectores, de modo que

$$\begin{aligned}\sum_{\mathbf{s}\in\mathcal{S}} \#\mathcal{J}_\mathbf{s} = \sum_{0\leqslant i\leqslant k-1}\sum_{\mathbf{s}\in\mathcal{S}_i} \#\mathcal{J}_\mathbf{s} &= \sum_{0\leqslant i\leqslant k-1} (\#\mathcal{S}_i)(i+1) = \sum_{0\leqslant i\leqslant k-1} (i+1)\big(\dim \mathcal{V}_i - \dim \mathcal{V}_{i+1}\big) \\ &= \big(\dim \mathcal{V}_0 - \dim \mathcal{V}_1\big) + 2\big(\dim \mathcal{V}_1 - \dim \mathcal{V}_2\big) + \cdots + k\big(\dim \mathcal{V}_{k-1} - \dim \mathcal{V}_k\big) \\ &= \dim \mathcal{V}_0 + \dim \mathcal{V}_1 + \cdots + \dim \mathcal{V}_{k-1} \\ &= \big(r(\mathbf{A}^0) - r(\mathbf{A}^1)\big) + \big(r(\mathbf{A}^1) - r(\mathbf{A}^2)\big) + \cdots + \big(r(\mathbf{A}^{k-1}) - r(\mathbf{A}^k)\big) \\ &= r(\mathbf{A}^0) - r(\mathbf{A}^k) = r(\mathbf{I}_n) - r(\mathbf{0}) = n = \dim \mathbb{C}^{n\times 1}.\end{aligned}$$

(e) A matriz $\mathbf{P}$ é invertível porque as suas colunas formam uma base de $\mathbb{C}^{n\times 1}$ (pela alínea anterior). Por outro lado, usando a alínea (c), obtemos

$$\begin{aligned}\mathbf{AP} &= \begin{bmatrix}\mathbf{AJ}_1 & \mathbf{AJ}_2 & \cdots & \mathbf{AJ}_r\end{bmatrix} = \begin{bmatrix}\mathbf{J}_1\mathbf{J}_{m_1}(0) & \cdots & \mathbf{J}_r\mathbf{J}_{m_r}(0)\end{bmatrix}\\ &= \mathbf{P}\begin{bmatrix}\mathbf{J}_{m_1}(0) & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \mathbf{J}_{m_2}(0) & \cdots & \mathbf{0}\\ \vdots & \vdots & & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{J}_{m_r}(0)\end{bmatrix}\end{aligned}$$

onde $m_i = \#\mathcal{J}_i$ para qualquer $1 \leqslant i \leqslant r$.

(f) Para justificar que $k = \max\{m_1, \ldots, m_r\}$, basta observar que $J_{m_i}(0)^{m_i} = \mathbf{0}$ para qualquer $1 \leqslant i \leqslant k$.

Por outro lado, para qualquer $1 \leqslant i \leqslant k-1$, a cardinalidade $\#\{1 \leqslant s \leqslant r : m_s = i\}$ é igual ao número de cadeias construídas a partir de um vector de $\mathcal{S}_i$ e, portanto, existem $\#\mathcal{S}_i = r(\mathbf{A}^i) - 2r(\mathbf{A}^{i+1}) + r(\mathbf{A}^{i+2})$ cadeias deste tipo.

9.4. Fixemos $1 \leqslant i \leqslant k$ e seja $\mathbf{B} \in \mathbb{C}^{n\times r_i}$, onde $r_i = r(\mathbf{A}^i)$, uma matriz cujas colunas formam uma base de $\mathcal{R}(\mathbf{A}^i)$. Como as colunas de $\mathbf{B}$ formam uma base de $\mathcal{R}(\mathbf{A}^i)$, é claro que $\mathcal{R}(\mathbf{B}) = \mathcal{R}(\mathbf{A}^i)$ (porque as colunas de $\mathbf{B}$ são linearmente independentes, logo $\dim\mathcal{R}(\mathbf{B}) = r(\mathbf{B}) = r_i = \dim\mathcal{R}(\mathbf{A}^i)$). Por outro lado, é claro que $\mathbf{Bv}_1, \ldots, \mathbf{Bv}_s \in \mathcal{R}(\mathbf{B}) = \mathcal{R}(\mathbf{A}^i)$; além disso, como $\mathbf{v}_1, \ldots, \mathbf{v}_s \in \mathcal{N}(\mathbf{AB})$, temos

$$\mathbf{A}(\mathbf{Bv}_i) = (\mathbf{AB})\mathbf{v}_i = \mathbf{0}, \qquad 1 \leqslant i \leqslant s.$$

Sendo assim, $\mathbf{Bv}_1, \ldots, \mathbf{Bv}_s \in \mathcal{R}(\mathbf{A}^i) \cap \mathcal{N}(\mathbf{A})$. Agora, pondo $\mathbf{V} = \begin{bmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_s\end{bmatrix}$, temos $r(\mathbf{V}) = s$ (porque as colunas de $\mathbf{V}$ são linearmente independentes) e, portanto,

$$r(\mathbf{BV}) = r(\mathbf{V}) - \dim\big(\mathcal{N}(\mathbf{B}) \cap \mathcal{R}(\mathbf{V})\big) = s$$

(porque $n(\mathbf{B}) = r_i - r(\mathbf{B}) = 0$, logo $\mathcal{N}(\mathbf{B}) = \{\mathbf{0}\}$). Como $\mathbf{Bv}_1, \ldots, \mathbf{Bv}_s$ são as colunas de $\mathbf{BV}$, concluímos que $\mathbf{Bv}_1, \ldots, \mathbf{Bv}_s$ são vectores linearmente independentes. Por outro lado, como $\{\mathbf{v}_1, \ldots, \mathbf{v}_s\}$ é uma base de $\mathcal{N}(\mathbf{AB})$, temos

$$s = n(\mathbf{AB}) = r_i - r(\mathbf{AB}) = r(\mathbf{A}^i) - r(\mathbf{A}^{i+1}) = \dim\mathcal{V}_i$$

(pelo Exercício 9.2) e, portanto, $\{\mathbf{Bv}_1, \ldots, \mathbf{Bv}_s\}$ é uma base de $\mathcal{V}_i$.

9.5. Temos

$$\mathbf{A}^2 = \begin{bmatrix}6 & 3 & 3 & 1 & 1 & 2\\ -6 & -3 & -3 & -1 & -1 & -2\\ 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\\ -6 & -3 & -3 & -1 & -1 & -2\\ -6 & -3 & -3 & -1 & -1 & -2\end{bmatrix} \qquad \text{e} \qquad \mathbf{A}^3 = \mathbf{0},$$

de modo que $\mathbf{A}$ é nilpotente com índice 3. Temos também $r(\mathbf{A}) = 3$, $r(\mathbf{A}^2) = 1$ e $r(\mathbf{A}^3) = 0$. Pelo Exercício 9.3,

- o número de blocos de Jordan de tamanho 3×3 é $r(\mathbf{A}^2) - 2r(\mathbf{A}^3) + r(\mathbf{A}^4) = 1$,
- o número de blocos de Jordan de tamanho 2×2 é $r(\mathbf{A}) - 2r(\mathbf{A}^2) + r(\mathbf{A}^3) = 3 - 2 = 1$,
- o número de blocos de Jordan de tamanho 1×1 é $r(\mathbf{I}_6) - 2r(\mathbf{A}) + r(\mathbf{A}^2) = 6 - 6 + 1 = 1$ (note que $\mathbf{A}^0 = \mathbf{I}_6$),

pelo que existe a matriz invertível $\mathbf{P} \in \mathbb{C}^{6\times 6}$ tal que

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} \mathbf{J}_3(0) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2(0) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{J}_1(0) \end{bmatrix} = \left[\begin{array}{ccc|cc|c} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Para determinarmos a matriz $\mathbf{P}$, teremos de construir as cadeias de Jordan como indicado nos exercícios anteriores. Em primeiro lugar, consideramos os subespaços vectoriais

$$\mathcal{V}_2 = \mathcal{R}(\mathbf{A}^2) \cap \mathcal{N}(\mathbf{A}) = \mathcal{R}(\mathbf{A}^2), \quad \mathcal{V}_1 = \mathcal{R}(\mathbf{A}) \cap \mathcal{N}(\mathbf{A}) \quad \text{e} \quad \mathcal{V}_0 = \mathcal{N}(\mathbf{A}),$$

que formam a cadeia $\{\mathbf{0}\} \subsetneq \mathcal{V}_2 \subsetneq \mathcal{V}_1 \subsetneq \mathcal{V}_0$. Como $r(\mathbf{A}^2)$, o subespaço $\mathcal{V}_2$ tem dimensão 1 e podemos escolher a base $\mathcal{S}_2 = \{\mathbf{s}_1\}$ onde

$$\mathbf{s}_1 = \begin{bmatrix} 6 \\ -6 \\ 0 \\ 0 \\ -6 \\ -6 \end{bmatrix}.$$

De seguida, estendemos $\mathcal{S}_1$ a uma base de $\mathcal{V}_1$, isto é, encontramos $\mathcal{S}_1$ tal que $\mathcal{S}_2 \cap \mathcal{S}_1 = \varnothing$ e $\mathcal{S}_2 \cup \mathcal{S}_1$ é base de $\mathcal{V}_1$. Para isto, vamos usar o exercício anterior e considerar a matriz

$$\mathbf{B} = \begin{bmatrix} 1 & 1 & -2 \\ 3 & 1 & 5 \\ -2 & -1 & 0 \\ 2 & 1 & 0 \\ -5 & -3 & -1 \\ -3 & -2 & -1 \end{bmatrix} \in \mathbb{C}^{6\times 3}$$

cujas colunas formam uma base de $\mathcal{R}(\mathbf{A})$. Calculamos

$$\mathbf{AB} = \begin{bmatrix} 6 & 3 & 3 \\ -6 & -3 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -6 & -3 & -3 \\ -6 & -3 & -3 \end{bmatrix}$$

(trata-se da matriz que tem as 3 primeiras colunas de $\mathbf{A}^2$) e determinamos uma base para $\mathcal{N}(\mathbf{AB})$: resolvendo o sistema $(\mathbf{AB})\mathbf{X} = \mathbf{0}$, obtemos a base $\{\mathbf{v}_1, \mathbf{v}_2\}$ onde

$$\mathbf{v}_1 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} \qquad \text{e} \qquad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}.$$

O exercício anterior diz-nos que $\{\mathbf{Bv}_1, \mathbf{Bv}_2\}$ é uma base de $\mathcal{V}_1$: ora temos

$$\mathbf{Bv}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \\ -1 \\ -1 \end{bmatrix} \qquad \text{e} \qquad \mathbf{Bv}_2 = \begin{bmatrix} -5 \\ 7 \\ 2 \\ -2 \\ 3 \\ 1 \end{bmatrix}.$$

Substituindo $\mathbf{s}_1$ nesta base, obtemos a base $\{\mathbf{s}_1, \mathbf{s}_2\}$ de $\mathcal{V}_1$ onde $\mathbf{s}_2 = \mathbf{Bv}_2$ (deste modo, temos $\mathcal{S}_1 = \{\mathbf{s}_2\}$). Finalmente, estendemos esta base a uma base $\mathcal{S} = \{\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3\}$ de $\mathcal{V}_0 = \mathcal{N}(\mathbf{A})$ (note que $n(\mathbf{A}) = 6 - r(\mathbf{A}) = 3$). Resolvendo o sistema $\mathbf{AX} = \mathbf{0}$, obtemos a base $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ onde

$$\mathbf{w}_1 = \begin{bmatrix} 2 \\ -4 \\ -1 \\ 3 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} -4 \\ 5 \\ 2 \\ 0 \\ 3 \\ 0 \end{bmatrix} \qquad \text{e} \qquad \mathbf{w}_3 = \begin{bmatrix} 1 \\ -2 \\ -2 \\ 0 \\ 0 \\ 3 \end{bmatrix}.$$

Se considerarmos a matriz $\begin{bmatrix} \mathbf{s}_1 \ \mathbf{s}_2 \ \mathbf{w}_1 \ \mathbf{w}_2 \ \mathbf{w}_3 \end{bmatrix}$ e a reduzirmos à forma de escada, concluímos que $\{\mathbf{s}_1, \mathbf{s}_2, \mathbf{w}_1\}$ é base de $\mathcal{N}(\mathbf{A})$, logo podemos escolher $\mathbf{s}_3 = \mathbf{w}_1$ (de modo que $\mathcal{S}_0 = \{\mathbf{s}_3\}$).

Para obtermos a matriz $\mathbf{P}$, construímos as cadeias de Jordan com base nos vectores $\mathbf{s}_1$, $\mathbf{s}_2$ e $\mathbf{s}_3$. Para $\mathbf{s}_1 \in \mathcal{S}_2$, determinamos um vector $\mathbf{u}_1 \in \mathbb{C}^{6\times 1}$ tal que $\mathbf{A}^2\mathbf{u}_1 = \mathbf{s}_1$; nesta situação, podemos escolher

$$\mathbf{u}_1 = \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

de modo que a cadeia que queremos corresponde à matriz

$$\mathbf{J}_1 = \begin{bmatrix} \mathbf{A}^2\mathbf{e}_1 & \mathbf{A}\mathbf{e}_1 & \mathbf{e}_1 \end{bmatrix} = \begin{bmatrix} 6 & 1 & 1 \\ -6 & 3 & 0 \\ 0 & -2 & 0 \\ 0 & 2 & 0 \\ -6 & -5 & 0 \\ -6 & -3 & 0 \end{bmatrix}.$$

Para $\mathbf{s}_2 \in \mathcal{S}_1$, determinamos um vector $\mathbf{u}_2 \in \mathbb{C}^{6\times 1}$ tal que $\mathbf{A}\mathbf{u}_2 = \mathbf{s}_2$; nesta situação, podemos escolher

$$\mathbf{u}_2 = \begin{bmatrix} -1 \\ 0 \\ 2 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

de modo que a cadeia que queremos corresponde à matriz

$$\mathbf{J}_2 = \begin{bmatrix} \mathbf{A}\mathbf{u}_2 & \mathbf{u}_2 \end{bmatrix} = \begin{bmatrix} -5 & -1 \\ 7 & 0 \\ 2 & 2 \\ -2 & 0 \\ 3 & 0 \\ 1 & 0 \end{bmatrix}.$$

Por fim, para $\mathbf{s}_3 \in \mathcal{S}_2$, determinamos um vector $\mathbf{u}_3 \in \mathbb{C}^{6\times 1}$ tal que $\mathbf{A}^0\mathbf{u}_3 = \mathbf{s}_3$ e, nesta situação, escolhemos $\mathbf{u}_3 = \mathbf{s}_3$, de modo que a cadeia que queremos corresponde à matriz

$$\mathbf{J}_3 = \begin{bmatrix} \mathbf{s}_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -4 \\ -1 \\ 3 \\ 0 \\ 0 \end{bmatrix}.$$

Em conclusão,

$$\mathbf{P} = \begin{bmatrix} \mathbf{J}_1 & \mathbf{J}_2 & \mathbf{J}_3 \end{bmatrix} = \begin{bmatrix} 6 & 1 & 1 & -5 & 1 & 2 \\ -6 & 3 & 0 & 7 & 0 & -4 \\ 0 & -2 & 0 & 2 & 2 & -1 \\ 0 & 2 & 0 & -2 & 0 & 3 \\ -6 & -5 & 0 & 3 & 0 & 0 \\ -6 & -3 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

9.6. Temos $\mathbf{A}^2 = \mathbf{0}$, logo $\mathbf{A}$ é nilpotente com índice 2. Por conseguinte, o maior bloco de Jordan é $\mathbf{J}_2(0)$. Por outro lado, temos $r_1 = r(\mathbf{A}) = 2$ e $r_i = r(\mathbf{A}^i) = 0$ para qualquer $i \geqslant 2$, de modo que o número de blocos de tipo 2×2 é $r_1 - 2r_2 + r_3 = 2$. Sendo assim, a forma canónica de Jordan de $\mathbf{A}$ é

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_2(0) & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2(0) \end{bmatrix}$$

e, portanto, existe uma matriz invertível $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J}$.

Neste caso, como $\mathbf{A}^2 = \mathbf{0}$, temos $\mathcal{R}(\mathbf{A}) = \mathcal{N}(\mathbf{A})$ (isto é, $\mathcal{V}_1 = \mathcal{V}_0$). Como as duas primeiras colunas de $\mathbf{A}$ são linearmente independentes, podemos tomar $\mathcal{S} = \{\mathbf{s}_1, \mathbf{s}_2\}$ onde

$$\mathbf{s}_1 = \begin{bmatrix} 3 \\ -2 \\ 1 \\ -5 \end{bmatrix} \qquad \text{e} \qquad \mathbf{s}_2 = \begin{bmatrix} 3 \\ -1 \\ -1 \\ -4 \end{bmatrix}.$$

Se $\{\mathbf{e}_1, \ldots, \mathbf{e}_4\}$ for a base canónica de $\mathbb{C}^{4\times 1}$, então $\mathbf{A}\mathbf{e}_1 = \mathbf{s}_1$ e $\mathbf{A}\mathbf{e}_2 = \mathbf{s}_2$. Deste modo, pondo $\mathbf{J}_1 = \begin{bmatrix} \mathbf{A}\mathbf{e}_1 & \mathbf{e}_1 \end{bmatrix}$ e $\mathbf{J}_2 = \begin{bmatrix} \mathbf{A}\mathbf{e}_2 & \mathbf{e}_2 \end{bmatrix}$, obtemos a matriz desejada

$$\mathbf{P} = \begin{bmatrix} \mathbf{J}_1 & \mathbf{J}_2 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 3 & 0 \\ -2 & 0 & -1 & 1 \\ 1 & 0 & -1 & 0 \\ -5 & 0 & -4 & 0 \end{bmatrix}.$$

9.7. Temos

$$r_1 = r(\mathbf{A}) = 4, \quad r_2 = r(\mathbf{A}^2) = 1 \quad \text{e} \quad r_i = r(\mathbf{A}^i) = 0,\ i \geqslant 3,$$

de modo que $\mathbf{A}$ é nilpotente com índice 3 (a matriz nula é a única matriz com característica 0). Sendo assim,

$$\#(\text{blocos de tipo } 3 \times 3) = r_2 - 2r_3 + r_4 = 1 - 0 + 0 = 1,$$
$$\#(\text{blocos de tipo } 2 \times 2) = r_1 - 2r_2 + r_3 = 4 - 2 + 0 = 2,$$
$$\#(\text{blocos de tipo } 1 \times 1) = r_0 - 2r_1 + r_2 = 8 - 8 + 1 = 1,$$

de modo que a forma canónica de Jordan é (a menos da permutação dos blocos)

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_3(0) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2(0) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{J}_2(0) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{J}_1(0) \end{bmatrix} = \left[\begin{array}{ccc|cc|cc|c} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

9.8. Seja $\mathbf{P} \in \mathbb{C}^{n\times n}$ uma matriz invertível tal que $\mathbf{J} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ está na forma canónica de Jordan. Sem perda de generalidade podemos supor que $\mathbf{J} = \begin{bmatrix} \mathbf{J}(\lambda) & \mathbf{0} \\ \mathbf{0} & \mathbf{J}' \end{bmatrix}$ onde

$$\mathbf{J}(\lambda) = \begin{bmatrix} \mathbf{J}_{n_1}(\lambda) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{n_2}(\lambda) & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{J}_{n_r}(\lambda) \end{bmatrix},$$

sendo $\mathbf{J}_{n_1}(\lambda), \ldots, \mathbf{J}_{n_r}(\lambda)$ todos os blocos de Jordan associados a λ, e onde $\mathbf{J}'$ está na forma canónica de Jordan e é tal que $\lambda \notin \sigma(\mathbf{J}')$ (isto é, λ não é valor próprio de $\mathbf{J}'$). Deste modo, temos

$$\mathbf{P}^{-1}(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{P} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} - \lambda\mathbf{I}_n = \begin{bmatrix} \mathbf{J}(0) & \mathbf{0} \\ \mathbf{0} & \mathbf{J}' - \lambda\mathbf{I}_{n-n'} \end{bmatrix},$$

onde

$$\mathbf{J}(0) = \begin{bmatrix} \mathbf{J}_{n_1}(0) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{n_2}(0) & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{J}_{n_r}(0) \end{bmatrix}$$

e $n' = n - (n_1 + \cdots + n_r)$. Mais geralmente, para qualquer $m \in \mathbb{N}$, temos

$$\mathbf{P}^{-1}(\mathbf{A} - \lambda\mathbf{I}_n)^m\mathbf{P} = (\mathbf{P}^{-1}\mathbf{A}\mathbf{P} - \lambda\mathbf{I}_n)^m = \begin{bmatrix} \mathbf{J}(0)^m & \mathbf{0} \\ \mathbf{0} & (\mathbf{J}' - \lambda\mathbf{I}_{n-n'})^m \end{bmatrix},$$

onde

$$\mathbf{J}(0)^m = \begin{bmatrix} \mathbf{J}_{n_1}(0)^m & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{n_2}(0)^m & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{J}_{n_r}(0)^m \end{bmatrix}$$

e, portanto, $k = \max\{n_1, \ldots, n_r\}$. A alínea (a) segue-se porque $(\mathbf{J}' - \lambda\mathbf{I}_{n-n'})^m$ é uma matriz invertível para qualquer $m \in \mathbb{N}$ (caso contrário, 0 seria valor próprio de $(\mathbf{J}' - \lambda\mathbf{I}_{n-n'})^m$). Por outro lado, a alínea (b) é consequência do Exercício 9.3 (porque, para qualquer $1 \leqslant i \leqslant k$, $\nu_i(\lambda)$ é igual ao número de blocos de Jordan de $\mathbf{A} - \lambda\mathbf{I}_n$ que estão associados a 0 e têm tamanho $i \times i$.)

9.9. Considerando a quarta coluna e a quinta linha de $\mathbf{A}$ vemos que $2, -1 \in \sigma(\mathbf{A})$. Temos

$$\begin{aligned} r_1(2) &= r(\mathbf{A} - 2\mathbf{I}_6) = 4, \\ r_2(2) &= r\big((\mathbf{A} - 2\mathbf{I}_6)^2\big) = 3, \\ r_3(2) &= r\big((\mathbf{A} - 2\mathbf{I}_6)^3\big) = 2, \\ r_4(2) &= r\big((\mathbf{A} - 2\mathbf{I}_6)^4\big) = 2, \end{aligned}$$

de modo que o maior bloco de Jordan é de tipo 3×3; além disso,

$$\begin{aligned} \#(\text{blocos de tipo } 3 \times 3) &= r_2(2) - 2r_3(2) + r_4(2) = 3 - 4 + 2 = 1, \\ \#(\text{blocos de tipo } 2 \times 2) &= r_1(2) - 2r_2(2) + r_3(2) = 4 - 6 + 2 = 0, \\ \#(\text{blocos de tipo } 1 \times 1) &= r_0(2) - 2r_1(2) + r_2(2) = 6 - 8 + 3 = 1, \end{aligned}$$

pelo que a forma canónica de Jordan de $\mathbf{A}$ tem os blocos $\mathbf{J}_3(2)$ e $\mathbf{J}_1(2)$.

Por outro lado, temos

$$r_1(-1) = r(\mathbf{A} + \mathbf{I}_6) = 4 \qquad \text{e} \qquad r_2(-1) = r\big((\mathbf{A} + \mathbf{I}_6)^2\big) = 4,$$

de modo que o maior bloco de Jordan é de tipo 1×1 e, portanto, têm de existir 2 blocos de Jordan $\mathbf{J}_1(-1)$; de facto,

$$\#(\text{blocos de tipo } 1 \times 1) = r_0(-1) - 2r_1(-1) + r_2(-1) = 6 - 8 + 4 = 2.$$

Para concluir, como $\mathbf{A}$ é de tipo 6×6, concluímos que a forma canónica de Jordan de $\mathbf{A}$ tem de ser

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_3(2) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_1(2) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{J}_1(-1) & 0 \\ \mathbf{0} & \mathbf{0} & 0 & \mathbf{J}_1(-1) \end{bmatrix} = \left[\begin{array}{ccc|c|c|c} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 2 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & -1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & -1 \end{array}\right] \in \mathbb{C}^{6\times 6}.$$

Daqui, resulta que $\sigma(\mathbf{A}) = \{-1, 2\}$ com m.a.$(2) = 4$, m.g.$(2) = 2$ e m.a.$(-1) =$ m.g.$(-1) = 2$; além disso, o polinómio característico de $\mathbf{A}$ é $p_{\mathbf{A}}(x) = (x+1)^2(x-2)^4$.

9.10. Temos $\sigma(\mathbf{A}) = \sigma(\mathbf{J}) = \{2, 3, 4\}$ com m.a.$(2) =$ m.g.$(2) = 2$, m.a.$(3) = 2$, m.g.$(3) = 1$, m.a.$(4) = 5$ e m.g.$(4) = 2$; mais, $p_{\mathbf{A}}(x) = (x-2)^2(x-3)^2(x-4)^5$ de $\mathbf{A}$.

9.11. Seja $\sigma(\mathbf{A}) = \{\lambda_1, \ldots, \lambda_r\}$. Para cada $1 \leqslant i \leqslant r$, construímos a base $\mathcal{S}_i = \mathcal{S}(\lambda_i) = \{\mathbf{s}_1^{(i)}, \ldots, \mathbf{s}_{n_i}^{(i)}\}$ de $\mathcal{N}(\mathbf{A} - \lambda_i \mathbf{I}_n)$ (de acordo com o processo descrito). Para cada $1 \leqslant i \leqslant r$ e cada $1 \leqslant j \leqslant n_i$, escolhemos o vector $\mathbf{v}_j^{(i)} \in \mathbb{C}^{n\times 1}$ tal que $(\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{v}_j^{(i)} = \mathbf{s}_j^{(i)}$ e a matriz $\mathbf{J}_j^{(i)} = \mathbf{J}_{\mathbf{s}_j}(\lambda_i)$ (como indicado). A matriz pretendida é

$$\mathbf{P} = \begin{bmatrix} \mathbf{J}_1^{(1)} & \cdots & \mathbf{J}_{n_1}^{(1)} & \mathbf{J}_1^{(2)} & \cdots & \mathbf{J}_{n_2}^{(2)} & \cdots & \mathbf{J}_1^{(r)} & \cdots & \mathbf{J}_{n_r}^{(r)} \end{bmatrix}.$$

[A justificação das afirmações é como no Exercício 9.3.]

9.12. Para $\mathbf{A} = \begin{bmatrix} 3 & 0 & 1 \\ -4 & 1 & -2 \\ -4 & 0 & -1 \end{bmatrix}$, temos $\sigma(\mathbf{A}) = \{1\}$. Como $r_1(1) = r(\mathbf{A} - \mathbf{I}_3) = 1$ e $r_2(1) = r\big((\mathbf{A} - \mathbf{I}_3)^2\big) = 0$, a forma canónica de Jordan de $\mathbf{A}$ é

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_2(1) & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_1(1) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Neste caso, consideramos $\mathcal{V}_1 = \mathcal{R}(\mathbf{A} - \mathbf{I}_3) \cap \mathcal{N}(\mathbf{A} - \mathbf{I}_3)$; uma base para este subespaço vectorial é $\mathcal{S}_1 = \{\mathbf{s}_1\}$ onde $\mathbf{s}_1 = \begin{bmatrix} 1 \\ -2 \\ -2 \end{bmatrix}$. Por outro lado, determinamos uma base $\mathcal{S}_0 = \{\mathbf{s}_1, \mathbf{s}_0\}$ para $\mathcal{V}_0 = \mathcal{N}(\mathbf{A} - \mathbf{I}_3)$ que contenha o vector $\mathbf{s}_1$. Ora, uma base para $\mathcal{N}(\mathbf{A} - \mathbf{I}_3)$ é $\{\mathbf{u}_1, \mathbf{u}_2\}$ onde $\mathbf{u}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ e $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$. Se transformarmos a matriz $\begin{bmatrix} \mathbf{s}_1 & \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix}$, obtemos

$$\begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \\ -2 & 0 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix},$$

de modo que podemos escolher $\mathbf{s}_0 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. Resolvendo o sistema $(\mathbf{A} - \mathbf{I}_3)\mathbf{X} = \mathbf{s}_1$, obtemos a solução $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, de modo que a matriz que queremos é

$$\mathbf{P} = \begin{bmatrix} (\mathbf{A} - \mathbf{I}_3)\mathbf{v}_1 & \mathbf{v}_1 & \mathbf{s}_0 \end{bmatrix} = \begin{bmatrix} \mathbf{s}_1 & \mathbf{v}_1 & \mathbf{s}_0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix}.$$

De facto, tem-se

$$\mathbf{AP} = \begin{bmatrix} 3 & 0 & 1 \\ -4 & 1 & -2 \\ -4 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ -2 & -2 & 1 \\ -2 & -1 & 0 \end{bmatrix},$$

$$\mathbf{PJ} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ -2 & -2 & 1 \\ -2 & -1 & 0 \end{bmatrix}.$$

RESOLUÇÃO DOS EXERCÍCIOS – FOLHA 10

10.1. (a) Por indução, prova-se que, para cada $m \in \mathbb{N}_0$, a matriz $\mathbf{U}_m$ é combinação linear das matrizes $\mathbf{I}_n, \mathbf{A}, \ldots, \mathbf{A}^m$ e a matriz $\mathbf{A}^m$ é combinação linear das matrizes $\mathbf{U}_0, \mathbf{U}_1, \ldots, \mathbf{U}_m$; por outras palavras,

$$\langle \mathbf{U}_0, \mathbf{U}_1, \ldots, \mathbf{U}_m \rangle = \langle \mathbf{I}_n, \mathbf{A}, \ldots, \mathbf{A}^m \rangle, \qquad m \in \mathbb{N}_0.$$

(b) Como $\mathbf{A}^k = \sum_{1 \leqslant j \leqslant k-1} \langle \mathbf{U}_j | \mathbf{A}^k \rangle \mathbf{U}_j$, temos

$$\mathbf{A}^k \in \langle \mathbf{U}_0, \mathbf{U}_1, \ldots, \mathbf{U}_{k-1} \rangle = \left\langle \mathbf{I}_n, \mathbf{A}, \ldots, \mathbf{A}^{k-1} \right\rangle$$

e, portanto, existem $\alpha_0, \alpha_1, \ldots, \alpha_{k-1}$ tais que

$$\mathbf{A}^k = \alpha_0 \mathbf{I}_n + \alpha_1 \mathbf{A} + \cdots + \alpha_{k-1} \mathbf{A}^{k-1}.$$

(c) Pondo $m(z) = z^k - \alpha_{k-1} z^{k-1} - \cdots - \alpha_1 z - \alpha_0$, temos

$$m(\mathbf{A}) = \mathbf{A}^k - \alpha_{k-1} \mathbf{A}^{k-1} - \cdots - \alpha_1 \mathbf{A} - \alpha_0 \mathbf{I}_n = \mathbf{0}$$

e, portanto, $m_{\mathbf{A}}(z)$ é um divisor de $m(z)$; em particular, temos $\mathrm{gr}(m_{\mathbf{A}}(z)) \leqslant \mathrm{gr}(m(z)) = k$. Se fosse $\mathrm{gr}(m_{\mathbf{A}}(z)) < k$, teríamos

$$\mathbf{A}^{k'} = \alpha_0' \mathbf{I}_n + \alpha_1' \mathbf{A} + \cdots + \alpha_{k'-1} \mathbf{A}^{k'-1}$$

onde $\alpha_0', \alpha_1', \ldots, \alpha_{k'-1}' \in \mathbb{C}$ e $m_{\mathbf{A}}(z) = z^{k'} - \alpha_{k'-1} z^{k'-1} - \cdots - \alpha_1' z - \alpha_0'$. Sendo assim,

$$\mathbf{A}^{k'} \in \left\langle \mathbf{I}_n, \mathbf{A}, \ldots, \mathbf{A}^{k'-1} \right\rangle = \langle \mathbf{U}_0, \mathbf{U}_1, \ldots, \mathbf{U}_{k'-1} \rangle$$

e, portanto, existem $\beta_0, \beta_1, \ldots, \beta_{k'-1} \in \mathbb{C}$ tais que

$$\mathbf{A}^{k'} = \beta_0 \mathbf{U}_0 + \beta_1 \mathbf{U}_1 + \cdots + \beta_{k'-1} \mathbf{U}_{k'-1}.$$

Como $\{\mathbf{U}_0, \mathbf{U}_1, \ldots, \mathbf{U}_{k'-1}\}$ é um sistema ortonormado de vectores em $\mathbb{C}^{n \times n}$, temos

$$\beta_j = \langle \beta_0 \mathbf{U}_0 + \beta_1 \mathbf{U}_1 + \cdots + \beta_{k'-1} \mathbf{U}_{k'-1} | \mathbf{U}_j \rangle = \langle \mathbf{A}^{k'-1} | \mathbf{U}_j \rangle, \qquad 0 \leqslant j \leqslant k' - 1,$$

pelo que

$$\mathbf{A}^{k'} = \sum_{1 \leqslant j \leqslant k'-1} \langle \mathbf{U}_j | \mathbf{A}^k \rangle \mathbf{U}_j,$$

o que contradiz a minimalidade de k. Segue-se que $\mathrm{gr}(m_{\mathbf{A}}(z)) = k = \mathrm{gr}(m(z))$ e, portanto, $m_{\mathbf{A}}(z) = m(z)$ (porque $m_{\mathbf{A}}(z)$ e $m(z)$ são mónicos).

(d) É só aplicar a definição das matrizes $\mathbf{U}_0, \mathbf{U}_1, \ldots, \mathbf{U}_{k-1}$.

(e) Imediata a partir das alíneas anteriores.

10.2. Os valores próprios de $\mathbf{A}$ são 1 e 2 com m.a.(1) = m.g.(1) = 2 e m.a.(2) = m.g.(2) = 1, de modo que $\mathbf{A}$ é diagonalizável. Por conseguinte,

$$m_{\mathbf{A}}(z) = (z-1)(z-2) = z^2 - 3z + 2$$

(pelo Corolário 17.5).

10.3. Usamos o Exercício 10.1. Temos:

$$\begin{aligned}
\nu_0 &= \|\mathbf{I}_4\| = 2, \quad \mathbf{U}_0 = \tfrac{1}{2}\,\mathbf{I}_4,\\
r_{0,1} &= \langle \mathbf{U}_0 | \mathbf{A} \rangle = 2,\\
\nu_1 &= \|\mathbf{A} - r_{0,1}\mathbf{U}_0\| = \sqrt{1209}, \quad \mathbf{U}_1 = \frac{1}{\nu_1}\,(\mathbf{A} - r_{0,1}\mathbf{U}_0) = \frac{1}{\sqrt{1209}}(\mathbf{A} - \mathbf{I}_4)\\
r_{0,2} &= \langle \mathbf{U}_0 | \mathbf{A}^2 \rangle = 2, \quad r_{1,2} = \langle \mathbf{U}_1 | \mathbf{A}^2 \rangle = 2\sqrt{1209},\\
\nu_2 &= \|\mathbf{A}^2 - r_{0,2}\mathbf{U}_0 - r_{1,2}\mathbf{U}_1\| = 0.
\end{aligned}$$

Assim, $\mathbf{A}^2 - r_{0,2}\mathbf{U}_0 - r_{1,2}\mathbf{U}_1 = \mathbf{0}$ e, portanto,

$$\mathbf{R} = \begin{bmatrix} 2 & 2 \\ 0 & \sqrt{1209} \end{bmatrix}, \qquad \mathbf{c} = \begin{bmatrix} 2 \\ 2\sqrt{1209} \end{bmatrix} \qquad \text{e} \qquad \mathbf{R}^{-1}\mathbf{c} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

Por conseguinte,

$$m_{\mathbf{A}}(z) = z^2 - 2z + 1 = (z-1)^2;$$

deste modo, $\sigma(\mathbf{A}) = \{1\}$, m.g.(1) = 3 e

$$\mathbf{J} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \in \mathbb{C}^{4\times 4}$$

é uma forma canónica de Jordan de $\mathbf{A}$.

10.4. Como $p_{\mathbf{A}}(z) = (z-\lambda)^2(z-\mu)^4$, temos $\sigma(\mathbf{A}) = \{\lambda, \mu\}$ com m.a.$(\lambda) = 2$ e m.a.$(\mu) = 4$; além disso, como $m_{\mathbf{A}}(z) = (z-\lambda)(z-\mu)^2$ o maior bloco de Jordan de $\mathbf{A}$ associado a λ é $\mathbf{J}_1(\lambda)$ e o o maior bloco de Jordan associado a μ é $\mathbf{J}_2(\mu)$. Por conseguinte, uma forma canónica de Jordan de $\mathbf{A}$ é

$$\mathbf{J} = \begin{bmatrix} \lambda & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 & 0 & 0 \\ 0 & 0 & \mu & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 1 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{bmatrix} \qquad \text{ou} \qquad \mathbf{J} = \begin{bmatrix} \lambda & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 & 0 & 0 \\ 0 & 0 & \mu & 1 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 1 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{bmatrix}.$$

10.5. Sabemos que matrizes semelhantes têm o mesmo polinómio característico (da Álgebra Linear) e o mesmo polinómio mínimo (pelo Corolário 17.2). Pelo exercício anterior, o recíproco não é necessariamente verdadeiro.

10.6. Sejam $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$ matrizes não-derrogatórias e suponhamos que $m_{\mathbf{A}}(z) = m_{\mathbf{B}}(z)$. Pelo Teorema 18.2, temos $m_{\mathbf{A}}(z) = p_{\mathbf{A}}(z)$ e $m_{\mathbf{B}}(z) = p_{\mathbf{B}}(z)$; além disso, $\mathbf{A}$ é semelhante à matriz companheira $\mathbf{C}$ de $p_{\mathbf{A}}(z)$ e, analogamente, $\mathbf{B}$ é semelhante à matriz companheira $\mathbf{C}'$ de $p_{\mathbf{B}}(z)$. Como $p_{\mathbf{A}}(z) = m_{\mathbf{A}}(z) = m_{\mathbf{B}}(z) = p_{\mathbf{B}}(z)$, concluímos que $\mathbf{C} = \mathbf{C}'$ e, portanto, $\mathbf{A}$ e $\mathbf{B}$ são semelhantes (porque são semelhantes a uma mesma matriz).

10.7. (a) Temos $m_{\mathbf{A}}(\mathbf{A}) = \mathbf{0}$, logo $m_{\mathbf{A}}(\mathbf{A})\mathbf{v} = \mathbf{0}$ e, portanto, $m_{\mathbf{A}}(z)$ é um polinómio anulador de $\mathbf{v}$. Sendo assim, o conjunto $\{\mathrm{gr}(\nu(z)) \colon 0 \neq \nu(z) \in \mathbb{C}[z],\ \nu(\mathbf{A})\mathbf{v} = \mathbf{0}\}$ tem elemento mínimo, isto é, existe um polinómio não-nulo $\nu(z) \in \mathbb{C}[z]$ com grau mínimo tal que $\nu(\mathbf{A})\mathbf{v} = \mathbf{0}$; além disso, é claro que podemos escolher $\nu(z)$ com coeficiente director igual a 1 (de modo que $\nu(z)$ é mónico).

Suponhamos que $\nu'(z) \in \mathbb{C}[z]$ é mónico com $\mathrm{gr}(\nu'(z)) = \mathrm{gr}(z)$ e tal que $\nu'(\mathbf{A})\mathbf{v} = \mathbf{0}$. Pelo algoritmo da divisão, existem $q(z), r(z) \in \mathbb{C}[z]$ tais que

$$\nu'(z) = q(z)\nu(z) + r(z), \qquad \mathrm{gr}(r(z)) < \mathrm{gr}(\nu(z)).$$

Então,

$$\mathbf{0} = \nu'(\mathbf{A})\mathbf{v} = q(\mathbf{A})\nu(\mathbf{A})\mathbf{v} + r(\mathbf{A})\mathbf{v} = r(\mathbf{A})\mathbf{v}$$

e, portanto, $r(z) = 0$ (por escolha de $\nu(z)$). Sendo assim, $\nu'(z) = q(z)\nu(z)$, logo $\nu'(z) = \nu(z)$ (porque $\nu'(z)$ e $\nu(z)$ são mónicos com o mesmo grau).

(b) Sejam $\beta_0, \beta_1, \ldots, \beta_r \in \mathbb{C}$ tais que

$$m_{\mathbf{v},\mathbf{A}}(z) = \beta_0 + \beta_1 z + \cdots + \beta_r z^r.$$

Então,

$$\mathbf{0} = m_{\mathbf{v},\mathbf{A}}(\mathbf{A})\mathbf{v} = \beta_0\mathbf{v} + \beta_1\mathbf{A}\mathbf{v} + \cdots + \beta_r\mathbf{A}^r\mathbf{v}$$

e, portanto, $\{\mathbf{v}, \mathbf{A}\mathbf{v}, \ldots, \mathbf{A}^r\mathbf{v}\}$ são linearmente dependentes. Por escolha de k, concluímos que $k \leqslant r$.

Por outro lado, o polinómio $\nu(z) = \alpha_0 + \alpha_1 z + \cdots + \alpha_{k-1}z^{k-1} + z^k$ verifica $\nu(\mathbf{A})\mathbf{v} = \mathbf{0}$, logo $r = \mathrm{gr}(m_{\mathbf{v},\mathbf{A}}(z)) \leqslant \mathrm{gr}(\nu(z)) = k$. Como $\nu(z)$ e $m_{\mathbf{v},\mathbf{A}}(z)$ são mónicos com o mesmo grau, tem de ser $\nu(z) = m_{\mathbf{v},\mathbf{A}}(z)$.

(c) Seja $1 \leqslant i \leqslant n$. Pelo argumento que usámos na alínea (a), verificamos que, se $\nu(z) \in \mathbb{C}[z]$ for tal que $\nu(\mathbf{A})\mathbf{v}_i = \mathbf{0}$, então $\nu_i(z) = m_{\mathbf{v}_i,\mathbf{A}}(z)$ será um divisor de $\nu(z)$. Em particular, $\nu_i(z)$ é um divisor de $m_{\mathbf{A}}(z)$ (porque $m_{\mathbf{A}}(\mathbf{A}) = \mathbf{0}$, logo $m_{\mathbf{A}}(\mathbf{A})\mathbf{v}_i = \mathbf{0}$). Como $1 \leqslant i \leqslant n$ é qualquer, concluímos que $m_{\mathbf{A}}(z)$ é um divisor de $\nu(z) = \mathrm{mmc}(\nu_1(z), \ldots, \nu_n(z))$.

Por outro lado, seja $\mathbf{v} \in \mathbb{C}^{n\times 1}$ e sejam $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ tais que

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n.$$

Além disso, para cada $1 \leqslant i \leqslant n$, seja $q_i(z) \in \mathbb{C}[z]$ tal que $\nu(z) = q_i(z)\nu_i(z)$, de modo que

$$\nu(\mathbf{A})\mathbf{v}_i = q_i(\mathbf{A})\nu_i(\mathbf{A})\mathbf{v}_i = q_i(\mathbf{A})\mathbf{0} = \mathbf{0}.$$

Então,

$$\nu(\mathbf{A})\mathbf{v} = \alpha_1\nu(\mathbf{A})\mathbf{v}_1 + \cdots + \alpha_n\nu(\mathbf{A})\mathbf{v}_n = \mathbf{0}$$

e, portanto, $\nu(\mathbf{A}) = \mathbf{0}$ (porque $\mathbf{v} \in \mathbb{C}^{n\times 1}$ é qualquer). Segue-se que $m_{\mathbf{A}}(z)$ é um divisor de $\nu(z)$, logo tem de ser $m_{\mathbf{A}}(z) = \nu(z)$ (porque $m_{\mathbf{A}}(z)$ e $\nu(z)$ são mónicos).

10.8. Temos

$$\mathbf{A}\mathbf{v} = \begin{bmatrix} 5 & 1 & 2 \\ -4 & 0 & -2 \\ -4 & -1 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \\ 2 \end{bmatrix} = 2\mathbf{v}$$

e, portanto, $m_{\mathbf{v},\mathbf{A}}(z) = z - 2$.

10.9. Para $\mathbf{A} = \begin{bmatrix} 5 & 1 & 2 \\ -4 & 0 & -2 \\ -4 & -1 & -1 \end{bmatrix}$, consideramos a base $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ de $\mathbb{C}^{3\times 1}$ onde

$$\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 4 \\ 0 \end{bmatrix} \qquad \text{e} \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$$

(trata-se de uma base de vectores próprios). Temos:

$$\begin{aligned} (\mathbf{A} - 2\mathbf{I}_3)\mathbf{v}_1 = \mathbf{0} \qquad &\Longrightarrow \qquad m_{\mathbf{v}_1,\mathbf{A}}(z) = z - 2, \\ (\mathbf{A} - \mathbf{I}_3)\mathbf{v}_2 = \mathbf{0} \qquad &\Longrightarrow \qquad m_{\mathbf{v}_2,\mathbf{A}}(z) = z - 1, \\ (\mathbf{A} - \mathbf{I}_3)\mathbf{v}_3 = \mathbf{0} \qquad &\Longrightarrow \qquad m_{\mathbf{v}_3,\mathbf{A}}(z) = z - 1, \end{aligned}$$

de modo que

$$m_{\mathbf{A}}(z) = \operatorname{mmc}(z-2, z-1) = (z-1)(z-2).$$

Para $\mathbf{A} = \begin{bmatrix} -7 & -4 & 8 & -8 \\ -4 & -1 & 4 & -4 \\ -16 & -8 & 17 & -16 \\ -6 & -3 & 6 & -5 \end{bmatrix}$, notamos que $1 \in \sigma(\mathbf{A})$ e que

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \\ 0 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 2 \\ 1 \\ 0 \end{bmatrix} \qquad \text{e} \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ -2 \\ 0 \\ 1 \end{bmatrix}$$

são vectores próprios de $\mathbf{A}$ associados a 1.

$$\begin{aligned}
(\mathbf{A}-\mathbf{I}_4)\mathbf{v}_1 = \mathbf{0} \quad &\Longrightarrow \quad m_{\mathbf{v}_1,\mathbf{A}}(z) = z-1,\\
(\mathbf{A}-\mathbf{I}_4)\mathbf{v}_2 = \mathbf{0} \quad &\Longrightarrow \quad m_{\mathbf{v}_2,\mathbf{A}}(z) = z-1,\\
(\mathbf{A}-\mathbf{I}_4)\mathbf{v}_3 = \mathbf{0} \quad &\Longrightarrow \quad m_{\mathbf{v}_3,\mathbf{A}}(z) = z-1.
\end{aligned}$$

Consideramos o vector $\mathbf{v}_4 = \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$. Temos

$$\mathbf{A}\mathbf{e}_1 = \begin{bmatrix} -7 \\ -4 \\ -16 \\ -6 \end{bmatrix} \qquad \text{e} \qquad \mathbf{A}^2\mathbf{e}_2 = \begin{bmatrix} -15 \\ -8 \\ -32 \\ -12 \end{bmatrix} = 2\begin{bmatrix} -7 \\ -4 \\ -16 \\ -6 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = 2\mathbf{A}\mathbf{e}_1 - \mathbf{e}_1,$$

de modo que $(\mathbf{A}^2 - 2\mathbf{A} + \mathbf{I}_4)\mathbf{e}_1 = \mathbf{0}$ e, portanto,

$$m_{\mathbf{e}_1,\mathbf{A}}(z) = z^2 - 2z + 1 = (z-1)^2.$$

Como $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{e}_1\}$ é uma base de $\mathbb{C}^{4\times 1}$, concluímos que

$$m_{\mathbf{A}}(z) = \text{mmc}(z-1, (z-1)^2) = (z-1)^2.$$

RESOLUÇÃO DOS EXERCÍCIOS – FOLHA 11

11.1. Os valores próprios de $\mathbf{A}$ são 2 e 4 com $\operatorname{ind}(2) = 1$ e $\operatorname{ind}(4) = 2$ (o que significa que 1 é o menor natural $k \in \mathbb{N}$ tal que $r\big((\mathbf{A} - 2\mathbf{I}_3)^k\big) = r\big((\mathbf{A} - 2\mathbf{I}_3)^{k+1}\big)$ e 2 é o menor natural $k \in \mathbb{N}$ tal que $r\big((\mathbf{A} - 4\mathbf{I}_3)^k\big) = r\big((\mathbf{A} - 4\mathbf{I}_n)^{k+1}\big)$). Por conseguinte, $f(\mathbf{A})$ existe sempre que $f(2)$, $f(4)$ e $f'(4)$ estejam definidos e, nesta situação,

$$f(\mathbf{A}) = f(2)\mathbf{G}_1 + f(4)\mathbf{G}_2 + f'(4)(\mathbf{A} - 4\mathbf{I}_3)\mathbf{G}_2$$

onde $\mathbf{G}_1$ e $\mathbf{G}_2$ são os projectores espectrais de $\mathbf{A}$ associados a 2 e a 4, respectivamente.

Para calcular, $\mathbf{G}_1$ e $\mathbf{G}_2$, considerando a função constante $f(z) = 1$, obtemos

$$\mathbf{I}_3 = f(\mathbf{A}) = \mathbf{G}_1 + \mathbf{G}_2.$$

Por outro lado, para $f(z) = (z-4)^2$, obtemos

$$(\mathbf{A} - 4\mathbf{I}_3)^2 = f(\mathbf{A}) = 4\mathbf{G}_1.$$

Sendo assim,

$$\mathbf{G}_1 = \frac{1}{4}(\mathbf{A} - 4\mathbf{I}_3)^2 = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & -2 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{e} \qquad \mathbf{G}_2 = \mathbf{I}_3 - \mathbf{G}_1 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Em particular, obtemos

$$\sqrt{\mathbf{A}} = \sqrt{2}\mathbf{G}_1 + 2\mathbf{G}_2 + \frac{1}{4}(\mathbf{A} - 4\mathbf{I}_3)\mathbf{G}_2 = \frac{1}{2}\begin{bmatrix} 5 & 1 & 7-2\sqrt{2} \\ -1 & 3 & 5-4\sqrt{2} \\ 0 & 0 & 2\sqrt{2} \end{bmatrix}$$

e

$$e^{\mathbf{A}} = e^2\mathbf{G}_1 + e^4\mathbf{G}_2 + e^4(\mathbf{A} - 4\mathbf{I}_3)\mathbf{G}_2 = e^2\begin{bmatrix} 3e^2 & 2e^2 & 7e^2-1 \\ -2e^2 & -e^2 & -4e^2-2 \\ 0 & 0 & 1 \end{bmatrix}.$$

11.2. Considerando a função polinomial $p(x,y) = x^2 + y^2 - 1$, obtemos a função complexa de variável complexa

$$h(z) = p(\cos(z), \operatorname{sen}(z)) = \cos^2(z) + \operatorname{sen}^2(z) - 1, \qquad z \in \mathbb{C}.$$

Como $h(z) = 0$ para qualquer $z \in \mathbb{C}$, temos

$$h(\mathbf{A}) = \mathbf{0}, \qquad \mathbf{A} \in \mathbb{C}^{n\times n},$$

ou seja,

$$\cos^2(\mathbf{A}) + \operatorname{sen}^2(\mathbf{A}) = 1, \qquad \mathbf{A} \in \mathbb{C}^{n\times n}.$$

11.3. (a) Para $f(z) = e^{tz}$, tem-se

$$\begin{aligned} e^{t\mathbf{A}} = f(\mathbf{A}) &= \sum_{1\leqslant i\leqslant r}\ \sum_{0\leqslant k\leqslant k_i-1} \frac{f^{(k)}(\lambda_i)}{k!}(\mathbf{A}-\lambda_i\mathbf{I}_n)^k\mathbf{G}_i \\ &= \sum_{1\leqslant i\leqslant r}\ \sum_{0\leqslant k\leqslant k_i-1} \frac{t^k e^{t\lambda_i}}{k!}(\mathbf{A}-\lambda_i\mathbf{I}_n)^k\mathbf{G}_i. \end{aligned}$$

(b) Usando a alínea anterior, verificamos que

$$\frac{d}{dt}e^{t\mathbf{A}} = \mathbf{A}e^{t\mathbf{A}}$$

o que prova que, para qualquer $\mathbf{c} \in \mathbb{C}^{n\times 1}$, $\mathbf{u}(t) = e^{t\mathbf{A}}\mathbf{c}$ é uma solução da equação diferencial

$$\mathbf{u}'(t) = \mathbf{A}\mathbf{u}(t), \qquad \mathbf{u}(0) = \mathbf{c}.$$

Por outro lado, seja $\mathbf{v}(t)$ uma solução da mesma equação diferencial. Usando a alínea (a), verificamos que $\mathbf{A}e^{-t\mathbf{A}} = e^{-t\mathbf{A}}\mathbf{A}$ e obtemos

$$\frac{d}{dt}\big(e^{-t\mathbf{A}}\mathbf{v}(t)\big) = -\mathbf{A}e^{-t\mathbf{A}}\mathbf{v}(t) + e^{-t\mathbf{A}}\mathbf{v}'(t) = \mathbf{0}.$$

Por conseguinte, a função $t \mapsto e^{-t\mathbf{A}}\mathbf{v}(t)$ é constante e, portanto,

$$e^{-t\mathbf{A}}\mathbf{v}(t) = e^{-0\mathbf{A}}\mathbf{v}(0) = e^{\mathbf{0}}\mathbf{v}(0) = \mathbf{v}(0) = \mathbf{c}$$

(porque $e^{\mathbf{0}} = \mathbf{I}_n$ (usando a alínea (a)). Como $e^{t\mathbf{A}}e^{-t\mathbf{A}} = \mathbf{I}_n$ (de novo usando a alínea (a)), concluímos que

$$\mathbf{v}(t) = e^{t\mathbf{A}}\mathbf{c} = \mathbf{u}(t).$$

(c) É só uma forma de reescrever

$$\begin{aligned} \mathbf{u}(t) = e^{t\mathbf{A}}\mathbf{c} &= \sum_{1\leqslant i\leqslant r}\ \sum_{0\leqslant k\leqslant k_i-1} \frac{t^k e^{t\lambda_i}}{k!}(\mathbf{A}-\lambda_i\mathbf{I}_n)^k\mathbf{G}_i\mathbf{c} \\ &= \sum_{1\leqslant i\leqslant r}\ \sum_{0\leqslant k\leqslant k_i-1} \frac{t^k e^{t\lambda_i}}{k!}\mathbf{v}_k(\lambda_i) \end{aligned}$$

onde

$$\mathbf{v}_k(\lambda_i) = (\mathbf{A}-\lambda_i\mathbf{I}_n)^k\mathbf{G}_i\mathbf{c}, \qquad 0 \leqslant k \leqslant k_i - 1,\ 1 \leqslant i \leqslant r.$$

11.4. Para $1 \leqslant i \leqslant r$, temos

$$f_i(\mathbf{A}) = \sum_{1\leqslant j\leqslant r}\ \sum_{0\leqslant k\leqslant k_j-1} \frac{f_i^{(k)}(\lambda_j)}{k!}(\mathbf{A}-\lambda_j\mathbf{I}_n)^k\mathbf{G}_j = f_i(\lambda_i)\mathbf{G}_i = \mathbf{G}_i;$$

além disso, sabemos que existe um polinómio $p_i(z) \in \mathbb{C}[z]$ tal que $f_i(\mathbf{A}) = p_i(\mathbf{A})$, logo $\mathbf{G}_i = p_i(\mathbf{A})$.

11.5. Para $f(z) = z^m$, temos

$$\begin{aligned}
f(\mathbf{A}) &= \sum_{1\leqslant i\leqslant r}\sum_{0\leqslant k\leqslant k_i-1} \frac{f^{(k)}(\lambda_i)}{k!}(\mathbf{A}-\lambda_i\mathbf{I}_n)^k\mathbf{G}_i \\
&= \sum_{1\leqslant i\leqslant r}\sum_{0\leqslant k\leqslant k_i-1} \frac{m(m-1)\cdots(m-k+1)\lambda_i^{m-k}}{k!}(\mathbf{A}-\lambda_i\mathbf{I}_n)^k\mathbf{G}_i \\
&= \sum_{1\leqslant i\leqslant r}\sum_{0\leqslant k\leqslant k_i-1} \frac{m!}{(m-k)!k!}\lambda_i^{m-k}(\mathbf{A}-\lambda_i\mathbf{I}_n)^k\mathbf{G}_i \\
&= \sum_{1\leqslant i\leqslant r}\sum_{0\leqslant k\leqslant k_i-1} \binom{m}{k}\lambda_i^{m-k}(\mathbf{A}-\lambda_i\mathbf{I}_n)^k\mathbf{G}_i
\end{aligned}$$

11.6. Os valores próprios de $\mathbf{A}$ são 1 e 4 com $\operatorname{ind}(1) = 1$ e $\operatorname{ind}(4) = 2$, de modo que

$$f(\mathbf{A}) = f(1)\mathbf{G}_1 + f(4)\mathbf{G}_2 + f'(4)(\mathbf{A} - 4\mathbf{I}_3)\mathbf{G}_2$$

onde $\mathbf{G}_1$ e $\mathbf{G}_2$ são os projectores espectrais de $\mathbf{A}$ associados a 1 e 4, respectivamente. Temos

$$\mathbf{G}_1 = \begin{bmatrix} 2 & 2 & 2 \\ -1 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \qquad \text{e} \qquad \mathbf{G}_2 = \mathbf{I}_3 - \mathbf{G}_1 = \begin{bmatrix} -1 & -2 & -2 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

pelo que

$$4\sqrt{\mathbf{A}} - 1 = f(\mathbf{A}) = \begin{bmatrix} -2 & -10 & -11 \\ 6 & 15 & 10 \\ -1 & -2 & 4 \end{bmatrix}.$$

11.7. Sejam $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ tais que $\sigma(\mathbf{A}) = \{\lambda_1, \ldots, \lambda_r\}$ e, para cada $1 \leqslant i \leqslant r$, seja $\Bbbk_i = \operatorname{ind}(\lambda_i)$. Então,

$$f(\mathbf{A}) = \sum_{1\leqslant i\leqslant r}\sum_{0\leqslant k\leqslant k_i-1} \frac{f^{(k)}(\lambda_i)}{k!}(\mathbf{A}-\lambda_i\mathbf{I}_n)^k\mathbf{G}_i$$

onde $\mathbf{G}_1, \ldots, \mathbf{G}_r \in \mathbb{C}^{n\times n}$ são os projectores espectrais de $\mathbf{A}$ associados a $\lambda_1, \ldots, \lambda_r$, respectivamente.

Seja $1 \leqslant s \leqslant r$ e seja $\mathbf{0} \neq \mathbf{v} \in \mathbb{C}^{n\times 1}$ tal que $\mathbf{A}\mathbf{v} = \lambda_s\mathbf{v}$. Usando a Proposição 19.2-(a), verificamos que

$$\mathbf{v} \in \mathcal{N}(\mathbf{A} - \lambda_s\mathbf{I}_n) \subseteq \mathcal{N}\big((\mathbf{A} - \lambda_s\mathbf{I}_n)^{k_s}\big) = \mathcal{R}(\mathbf{G}_s)$$

e, portanto, existe $\mathbf{w} \in \mathbb{C}^{n\times 1}$ tal que $\mathbf{v} = \mathbf{G}_s\mathbf{w}$. Como $\mathbf{G}_s^2 = \mathbf{G}_s$ (pela Proposição 19.2-(b)), concluímos que

$$\mathbf{v} = \mathbf{G}_s\mathbf{w} = \mathbf{G}_s^2\mathbf{w} = \mathbf{G}_s\mathbf{v}.$$

Por outro lado, escolhendo uma matriz invertível $\mathbf{P} \in \mathbb{C}^{n\times n}$tal que $\mathbf{J} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ está na forma canónica de Jordan e pondo

$$\mathbf{P} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_r \end{bmatrix} \qquad \text{e} \qquad \mathbf{P}^{-1} = \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \\ \vdots \\ \mathbf{Y}_r \end{bmatrix}$$

onde, para cada $1 \leqslant i \leqslant r$, temos $\mathbf{X}_i \in \mathbb{C}^{n\times m_i}$, $\mathbf{Y}_i \in \mathbb{C}^{m_i\times n}$ e $m_i = \text{m.a.}(\lambda_i)$, sabemos que

$$\mathbf{G}_i = \mathbf{X}_i\mathbf{Y}_i \in \mathbb{C}^{n\times n}, \qquad 1 \leqslant i \leqslant r.$$

Como $\mathbf{P}^{-1}\mathbf{P} = \mathbf{I}_n$, para quaisquer $1 \leqslant i, j \leqslant r$, temos

$$\mathbf{Y}_i\mathbf{X}_j = \begin{cases} \mathbf{I}_{m_i}, & \text{se } i = j, \\ \mathbf{0}, & \text{se } i \neq j, \end{cases}$$

e, portanto,

$$\mathbf{G}_i\mathbf{G}_j = \mathbf{X}_i\mathbf{Y}_i\mathbf{X}_j\mathbf{Y}_j = \begin{cases} \mathbf{G}_i, & \text{se } i = j, \\ \mathbf{0}, & \text{se } i \neq j. \end{cases}$$

Em particular, concluímos que

$$\mathbf{G}_i\mathbf{v} = \mathbf{G}_i\mathbf{G}_s\mathbf{v} = \mathbf{0}, \qquad 1 \leqslant i \leqslant r,\ i \neq s.$$

Por conseguinte, obtemos

$$\begin{aligned} f(\mathbf{A})\mathbf{v} &= \sum_{1\leqslant i\leqslant r}\ \sum_{0\leqslant k\leqslant k_i-1} \frac{f^{(k)}(\lambda_i)}{k!}\,(\mathbf{A} - \lambda_i\mathbf{I}_n)^k\mathbf{G}_i\mathbf{v} \\ &= \sum_{0\leqslant k\leqslant k_s-1} \frac{f^{(k)}(\lambda_s)}{k!}\,(\mathbf{A} - \lambda_s\mathbf{I}_n)^k\mathbf{v} \\ &= f(\lambda_s)\mathbf{v} + \sum_{1\leqslant k\leqslant k_s-1} \frac{f^{(k)}(\lambda_s)}{k!}\,(\mathbf{A} - \lambda_s\mathbf{I}_n)^k\mathbf{v} = f(\lambda_s)\mathbf{v}. \end{aligned}$$

Como se queria.

11.8. As matrizes $\mathbf{A}$ e $\mathbf{A}^{\mathsf{T}}$ são semelhantes (porque têm a mesma forma canónica de Jordan) e, portanto, têm os mesmos valores próprios; além disso, o índice de um valor próprio de $\mathbf{A}$ é igual ao índice do mesmo valor próprio de $\mathbf{A}^{\mathsf{T}}$ (pela definição de índice). Sendo assim, se $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ forem tais que $\sigma(\mathbf{A}) = \{\lambda_1, \ldots, \lambda_r\}$ e se $\mathbb{k}_i = \text{ind}(\lambda_i)$, $1 \leqslant i \leqslant r$, então

$$f(\mathbf{A}) = \sum_{1\leqslant i\leqslant r}\ \sum_{0\leqslant k\leqslant k_i-1} \frac{f^{(k)}(\lambda_i)}{k!}\,(\mathbf{A} - \lambda_i\mathbf{I}_n)^k\mathbf{G}_i$$

onde $\mathbf{G}_1, \dots, \mathbf{G}_r \in \mathbb{C}^{n\times n}$ são os projectores espectrais de $\mathbf{A}$ associados a $\lambda_1, \dots, \lambda_r$, respectivamente, de modo que

$$f(\mathbf{A})^{\mathsf{T}} = \sum_{1\leqslant i\leqslant r} \sum_{0\leqslant k\leqslant k_i-1} \frac{f^{(k)}(\lambda_i)}{k!} \big((\mathbf{A}-\lambda_i\mathbf{I}_n)^k\mathbf{G}_i\big)^{\mathsf{T}}.$$

Analogamente, temos

$$f(\mathbf{A}^{\mathsf{T}}) = \sum_{1\leqslant i\leqslant r} \sum_{0\leqslant k\leqslant k_i-1} \frac{f^{(k)}(\lambda_i)}{k!} (\mathbf{A}^{\mathsf{T}}-\lambda_i\mathbf{I}_n)^k\mathbf{G}'_i$$

onde $\mathbf{G}'_1, \dots, \mathbf{G}'_r \in \mathbb{C}^{n\times n}$ são os projectores espectrais de $\mathbf{A}^{\mathsf{T}}$ associados a $\lambda_1, \dots, \lambda_r$, respectivamente. Usando a definição, não é difícil verificar que

$$\mathbf{G}'_i = \mathbf{G}_i^{\mathsf{T}}, \qquad 1 \leqslant i \leqslant r,$$

de modo que

$$\begin{aligned} f(\mathbf{A}^{\mathsf{T}}) &= \sum_{1\leqslant i\leqslant r} \sum_{0\leqslant k\leqslant k_i-1} \frac{f^{(k)}(\lambda_i)}{k!} (\mathbf{A}^{\mathsf{T}}-\lambda_i\mathbf{I}_n)^k\mathbf{G}_i^{\mathsf{T}} \\ &= \sum_{1\leqslant i\leqslant r} \sum_{0\leqslant k\leqslant k_i-1} \frac{f^{(k)}(\lambda_i)}{k!} \big(\mathbf{G}_i(\mathbf{A}-\lambda_i\mathbf{I}_n)^k\big)^{\mathsf{T}} \\ &= \Bigg(\sum_{1\leqslant i\leqslant r} \sum_{0\leqslant k\leqslant k_i-1} \frac{f^{(k)}(\lambda_i)}{k!} \mathbf{G}_i(\mathbf{A}-\lambda_i\mathbf{I}_n)^k\Bigg)^{\mathsf{T}}. \end{aligned}$$

Agora, escolhendo uma matriz invertível $\mathbf{P} \in \mathbb{C}^{n\times n}$tal que $\mathbf{J} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ está na forma canónica de Jordan e pondo

$$\mathbf{P} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_r \end{bmatrix} \qquad \text{e} \qquad \mathbf{P}^{-1} = \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \\ \vdots \\ \mathbf{Y}_r \end{bmatrix}$$

onde, para cada $1 \leqslant i \leqslant r$, temos $\mathbf{X}_i \in \mathbb{C}^{n\times m_i}$, $\mathbf{Y}_i \in \mathbb{C}^{m_i\times n}$ e $m_i = \text{m.a.}(\lambda_i)$, obtemos

$$\mathbf{P}^{-1}\mathbf{G}_i = \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{X}_i \\ \vdots \\ \mathbf{0} \end{bmatrix} \qquad \text{e} \qquad \mathbf{G}_i\mathbf{P} = \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{Y}_i & \cdots & \mathbf{0} \end{bmatrix}, \qquad 1 \leqslant i \leqslant r,$$

e daqui resulta facilmente que

$$(\mathbf{A}-\lambda_i\mathbf{I}_n)\mathbf{G}_i = \mathbf{G}_i(\mathbf{A}-\lambda_i\mathbf{I}_n), \qquad 1 \leqslant i \leqslant r$$

(considerando a forma canónica de Jordan $\mathbf{J}$). Segue-se que

$$f(\mathbf{A}^{\mathtt{T}}) = \left(\sum_{1\leqslant i\leqslant r}\sum_{0\leqslant k\leqslant k_i-1}\frac{f^{(k)}(\lambda_i)}{k!}\,\mathbf{G}_i(\mathbf{A}-\lambda_i\mathbf{I}_n)^k\right)^{\mathtt{T}}$$
$$= \sum_{1\leqslant i\leqslant r}\sum_{0\leqslant k\leqslant k_i-1}\frac{f^{(k)}(\lambda_i)}{k!}\,(\mathbf{A}-\lambda_i\mathbf{I}_n)^k\mathbf{G}_i\Bigg)^{\mathtt{T}} = f(\mathbf{A})^{\mathtt{T}}.$$

Como se queria.

11.9. (a) Consideramos o polinómio $p(x,y) = xy - 1$, de modo que

$$p(e^z, e^{-z}) = e^z e^{-z} - 1 = e^0 - 1 = 0, \qquad z \in \mathbb{C}.$$

Deste modo a função $h(z) = p(e^z, e^{-z})$ é constantemente igual a 0 e, portanto,

$$h(\mathbf{A}) = \mathbf{0}, \qquad \mathbf{A} \in \mathbb{C}^{n\times n},$$

ou seja,

$$e^{\mathbf{A}}e^{-\mathbf{A}} = 1, \qquad \mathbf{A} \in \mathbb{C}^{n\times n}.$$

(b) Análoga a (a), usando o polinómio $p(x,y) = x - y$ e a função $h(z) = p(e^{\alpha z}, (e^z)^\alpha)$.

(c) Análoga a (a), usando o polinómio $p(x,y) = x - y$ e a função $h(z) = p(e^{iz}, \cos(z) + i\,\mathrm{sen}(z))$.

11.10. Temos $p_{\mathbf{A}}(x) = x^3$ e existe $1 = r(\mathbf{A}) - 2r(\mathbf{A}^2) + r(\mathbf{A}^3)$ blocos de Jordan de tipo 2×2 associados a $0 \in \sigma(\mathbf{A})$, de modo que $\mathbf{A}$ é semelhante a $\mathbf{J} = \begin{bmatrix} 0 & 1 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$. Sendo assim, temos $\mathbf{A}^2 = \mathbf{0}$ e, portanto, $\mathrm{ind}(0) = 2$. Por conseguinte, há que determinar $\alpha, \beta \in \mathbb{C}$, tal que o polinómio $p(z) = \alpha + \beta z$ satisfaça $p(0) = f(0) = 1$ e $p'(0) = f'(0) = 1$. Segue-se que $\alpha = \beta = 1$ e, portanto, $p(z) = 1 + z$ satisfaz

$$e^{\mathbf{A}} = p(\mathbf{A}) = \mathbf{I} + \mathbf{A}.$$

11.11. (a) Para $\mathbf{A} = \begin{bmatrix} -1/2 & 3/2 & -3/2\\ 1 & 0 & -1/2\\ 1 & -1 & 1/2\end{bmatrix}$, temos $p_{\mathbf{A}}(x) = (x-1)(x-1/2)^2$, de modo que $\rho(\mathbf{A}) = 1$ e 1 é o único valor próprio $\lambda \in \sigma(\mathbf{A})$ que satisfaz $|\lambda| = 1$; além disso, $\mathrm{m.a.}(1) = 1$, logo $\mathrm{m.g.}(1) = \mathrm{m.a.}(1)$. Pelo Teorema 20.2, concluímos que $\mathbf{A}$ é convergente e, pelo Teorema 20.3, também é somável à Cesàro.

Para $\mathbf{A} = \begin{bmatrix} 0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{bmatrix}$, temos $p_{\mathbf{A}}(x) = x^3 - 1 = (x-1)(x-(1+\sqrt{5})/2)(x-(1-\sqrt{5})/2)$, de modo que $\rho(\mathbf{A}) = 1$. Como $\left|1 \pm \sqrt{5})/2\right| = 1$, o Teorema 20.2 garante que $\mathbf{A}$ não é convergente.

No entanto, como m.a.$(\lambda) = 1$ para qualquer $\lambda \in \sigma(\mathbf{A})$, temos m.a.$(\lambda) =$ m.g.(λ) para qualquer $\lambda \in \sigma(\mathbf{A})$ e, portanto, $\mathbf{A}$ é somável à Cesàro (pelo Teorema 20.3).

Para $\mathbf{A} = \begin{bmatrix} -1 & -2 & -3/2 \\ 1 & 2 & 1 \\ 1 & 1 & 3/2 \end{bmatrix}$, temos $p_{\mathbf{A}}(x) = (x-1)^2(x-1/2)$, de modo que $\rho(\mathbf{A}) = 1$. Como $r(\mathbf{A} - \mathbf{I}_3) = 2$ e $r\big((\mathbf{A} - \mathbf{I})^2\big) = r\big((\mathbf{A} - \mathbf{I})^3\big) = 1$, temos ind$(1) = 2$, de modo que $\mathbf{A}$ não é convergente, nem é somável à Cesàro.

(b) Para $\mathbf{A} = \begin{bmatrix} -1/2 & 3/2 & -3/2 \\ 1 & 0 & -1/2 \\ 1 & -1 & 1/2 \end{bmatrix}$, temos $\lim_{k\to\infty} \mathbf{A}^k = \mathbf{G}$ onde $\mathbf{G}$ é o projector espectral de $\mathbf{A}$ associado a 1 (pelo Teorema 20.2). Assim,

$$\lim_{k\to\infty} \mathbf{A}^k = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Pelo Teorema 20.3, $\mathbf{G}$ é também a soma de Cesàro de $\mathbf{A}$, isto é,

$$\lim_{k\to\infty} \frac{\mathbf{I} + \mathbf{A} + \cdots + \mathbf{A}^{k-1}}{k} = \mathbf{G}.$$

Para $\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$, temos $\lim_{k\to\infty} \frac{\mathbf{I}+\mathbf{A}+\cdots+\mathbf{A}^{k-1}}{k} = \mathbf{G}$ onde $\mathbf{G}$ é o projector espectral de $\mathbf{A}$ associado a 1 (pelo Teorema 20.3). Assim,

$$\lim_{k\to\infty} \frac{\mathbf{I} + \mathbf{A} + \cdots + \mathbf{A}^{k-1}}{k} = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

11.12. Suponhamos que $\mathbf{A} \in \mathbb{C}^{n\times n}$ é convergente. Então, verifica-se alguma das condições do Teorema 20.2 e, portanto, também se verifica alguma das condições do Teorema 20.3. Por conseguinte, $\mathbf{A}$ é somável à Cesàro.

RESOLUÇÃO DOS EXERCÍCIOS – FOLHA 12

12.1. Como $\mathbf{A} \in \mathbb{R}^{3\times 3}$ é uma matriz positiva, o teorema de Perron garante que as propriedades seguintes são verdadeiras:

(a) O raio espectral $\rho(\mathbf{A}) \in \mathbb{R}_0^+$ é estritamente positivo; além disso, $\rho(\mathbf{A})$ é um valor próprio de $\mathbf{A}$ com m.a.$\big(\rho(\mathbf{A})\big) = 1$ e $\rho(\mathbf{A})$ é o único valor próprio $\lambda \in \sigma(\mathbf{A})$ tal que $|\lambda| = \rho(\mathbf{A})$.

(b) Existe um e um só vector $\mathbf{p} \in \mathbb{R}^{3\times 1}$ tal que

$$\mathbf{p} > \mathbf{0}, \qquad \mathbf{A}\mathbf{p} = \rho(\mathbf{A})\mathbf{p} \qquad \text{e} \qquad \|\mathbf{p}\|_1 = 1$$

e, além disso, qualquer vector próprio $\mathbf{v} \in \mathbb{R}^{3\times 1}$ de $\mathbf{A}$ com $\mathbf{v} \geqslant \mathbf{0}$ é um múltiplo escalar positivo de $\mathbf{p}$.

(c) Se $\mathcal{N} = \{\mathbf{v} \in \mathbb{R}^{3\times 1} : \mathbf{v} \geqslant \mathbf{0},\ \mathbf{v} \neq \mathbf{0}\}$, então

$$\rho(\mathbf{A}) = \max_{\mathbf{v}\in\mathcal{N}} \min_{\substack{1\leqslant i\leqslant 3\\ v_i\neq 0}} \frac{(\mathbf{A}\mathbf{v})_i}{v_i}.$$

Ora, $\sigma(\mathbf{A}) = \{6, 12\}$ com m.a.$(6) = 2$ e m.a.$(12) = 1$. É claro que $\rho(\mathbf{A}) = \max\{6, 12\}$ logo a condição (a) é verdedeira. Temos

$$\mathcal{N}(\mathbf{A} - 12\mathbf{I}_3) = \left\{ \begin{bmatrix} \alpha \\ \alpha \\ \alpha \end{bmatrix} : \alpha \in \mathbb{C} \right\} \subseteq \mathbb{C}^{3\times 1},$$

de modo que

$$\mathbf{p} = \frac{1}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \in \mathbb{R}^{3\times 1}$$

é o único vector $\mathbf{p} \in \mathbb{R}^{3\times 1}$ tal que

$$\mathbf{p} > \mathbf{0}, \qquad \mathbf{A}\mathbf{p} = 12\mathbf{p} = \rho(\mathbf{A})\mathbf{p} = 12 \qquad \text{e} \qquad \|\mathbf{p}\|_1 = p_1 + p_2 + p_3 = 1.$$

Assim, (b) é verdadeira: notemos que, se $\mathbf{v} \in \mathbb{R}^{3\times 1}$ for um valor próprio de $\mathbf{A}$ com $\mathbf{v} \geqslant \mathbf{0}$, então $\mathbf{v} = \alpha\mathbf{p}$ com $\alpha \in \mathbb{R}^+$; com efeito,

$$\mathcal{N}(\mathbf{A} - 6\mathbf{I}_3) = \left\{ \begin{bmatrix} 2\beta - 2\gamma \\ \beta \\ \gamma \end{bmatrix} : \beta, \gamma \in \mathbb{C} \right\} \subseteq \mathbb{C}^{3\times 1},$$

de modo que, se $\mathbf{v} \in \mathcal{N}(\mathbf{A} - 6\mathbf{I}_3)$ for não-negativo, então β, γ e $-2(\beta + \gamma)$ terão de ser reais não-negativos, o que só pode acontecer quando $\mathbf{v} = \mathbf{0}$. Quanto a (c), temos

$$\mathbf{Av} = \begin{bmatrix} 7v_1 + 2v_2 + 3v_3 \\ v_1 + 8v_2 + 3v_3 \\ v_1 + 2v_2 + 9v_3 \end{bmatrix}$$

e o valor $\max_{\mathbf{v}\in\mathcal{N}} \min_{\substack{1 \leqslant i \leqslant 3 \\ v_i \neq 0}} \frac{(\mathbf{Av})_i}{v_i}$ é atingido quando $\mathbf{v} = \mathbf{p}$ (sendo igual a $12 = \rho(\mathbf{A})$.

O vector de Perron à direita é $\mathbf{p}$ (por definição). Quanto ao vector de Perron à esquerda, consideramos $\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 7 & 1 & 1 \\ 2 & 8 & 2 \\ 3 & 3 & 9 \end{bmatrix}$ (que é também uma matriz positiva). Temos $\sigma(\mathbf{A}^{\mathrm{T}}) = \sigma(\mathbf{A}) = \{6, 12\}$, logo $\rho(\mathbf{A}^{\mathrm{T}}) = \rho(\mathbf{A}) = 12$. Temos

$$\mathcal{N}(\mathbf{A}^{\mathrm{T}} - 12\mathbf{I}_3) = \left\{ \begin{bmatrix} \alpha \\ 2\alpha \\ 3\alpha \end{bmatrix} : \alpha \in \mathbb{C} \right\} \subseteq \mathbb{C}^{3\times 1},$$

de modo que o vector de Perron de $\mathbf{A}^{\mathrm{T}}$ é

$$\mathbf{q} = \frac{1}{6} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

e, portanto, este é o vector de Perron à esquerda de $\mathbf{A}$.

12.2. (a) Suponhamos que $\rho(\mathbf{A}) = \max_{\lambda\in\sigma(\mathbf{A})} |\lambda| = 0$. Então, $|\lambda| = 0$ para qualquer $\lambda \in \sigma(\mathbf{A})$ e, portanto, $\sigma(\mathbf{A}) = \{0\}$. Sendo assim, $\mathbf{A}$ tem de ser uma matriz nilpotente (basta considerar uma forma canónica de Jordan de $\mathbf{A}$), logo existe $m \in \mathbb{N}$ tal que $\mathbf{A}^m = \mathbf{0}$. Por conseguinte, para quaisquer $1 \leqslant i, j \leqslant n$, temos

$$0 = (\mathbf{A}^m)_{i,j} = \sum_{1 \leqslant i_1, \ldots, i_{m-1} \leqslant n} a_{i,i_1} a_{i_1,i_2} \cdots a_{i_{m-1},j},$$

o que não pode acontecer (porque $\mathbf{A} > \mathbf{0}$).

(b) Óbvia porque $\rho(\mathbf{A}) \in \mathbb{R}^+$.

(c) Se $\mathbf{A} > 0$ e $\mathbf{u} \geqslant \mathbf{0}$, $\mathbf{u} \neq \mathbf{0}$, então

$$(\mathbf{Au})_i = \sum_{1 \leqslant j \leqslant n} a_{i,j} u_j > 0, \qquad 1 \leqslant i \leqslant n,$$

uma vez que existe $1 \leqslant j \leqslant n$ tal que $u_j > 0$.

(d) Se $\mathbf{u} \geqslant \mathbf{v} \geqslant \mathbf{0}$, então $\mathbf{u} - \mathbf{v} \geqslant \mathbf{0}$, logo $u_j \geqslant v_j$ para qualquer $1 \leqslant j \leqslant n$. Se além disso $\mathbf{A} \geqslant \mathbf{0}$, então

$$(\mathbf{Au})_i = \sum_{1 \leqslant j \leqslant n} a_{i,j} u_j \geqslant \sum_{1 \leqslant j \leqslant n} a_{i,j} v_j = (\mathbf{Av})_i, \qquad 1 \leqslant i \leqslant n.$$

(e) Suponhamos que $\mathbf{A} \geqslant \mathbf{0}$, $\mathbf{u} > \mathbf{0}$ e $\mathbf{A}\mathbf{u} = \mathbf{0}$. Então,

$$0 = (\mathbf{A}\mathbf{u})_i = \sum_{1 \leqslant j \leqslant n} a_{i,j} u_j, \qquad 1 \leqslant i \leqslant n.$$

Como todas as parcelas desta soma são não-negativas, tem de ser

$$a_{i,j} u_j = 0, \qquad 1 \leqslant i, j \leqslant n,$$

logo

$$a_{i,j} = 0, \qquad 1 \leqslant i, j \leqslant n,$$

(porque $\mathbf{u} > \mathbf{0}$).

(f) Suponhamos que $\mathbf{A} \geqslant \mathbf{0}$, $\mathbf{A} \neq \mathbf{0}$ e $\mathbf{u} > \mathbf{v} > \mathbf{0}$. Então, $\mathbf{u} - \mathbf{v} > \mathbf{0}$ e, portanto, $\mathbf{A}(\mathbf{u} - \mathbf{v}) > \mathbf{0}$ (pela alínea (c)). Como $\mathbf{A}(\mathbf{u} - \mathbf{v}) = \mathbf{A}\mathbf{u} - \mathbf{A}\mathbf{v}$, concluímos que $\mathbf{A}\mathbf{u} > \mathbf{A}\mathbf{v}$.

12.3. Sejam $\mathbf{p}_1, \mathbf{p}_2 \in \mathbb{R}^{n \times 1}$ tais que:

- $\mathbf{A}\mathbf{p}_1 = \rho(\mathbf{A})\mathbf{p}_1, \qquad \mathbf{p}_1 > \mathbf{0} \qquad$ e $\qquad \|\mathbf{p}_1\|_1 = 1,$
- $\mathbf{A}\mathbf{p}_2 = \rho(\mathbf{A})\mathbf{p}_2, \qquad \mathbf{p}_2 > \mathbf{0} \qquad$ e $\qquad \|\mathbf{p}_2\|_1 = 1.$

Como m.g.$(\rho(\mathbf{A})) = 1$, existe $\alpha \in \mathbb{R}$ tal que $\mathbf{p}_2 = \alpha \mathbf{p}_1$. Como $\mathbf{p}_1, \mathbf{p}_2 > \mathbf{0}$, tem de ser $\alpha > 0$. Finalmente, como

$$1 = \|\mathbf{p}_2\|_1 = \|\alpha \mathbf{p}_1\|_1 = |\alpha| \, \|\mathbf{p}_1\|_1 = \alpha,$$

concluímos que $\mathbf{p}_2 = \mathbf{p}_1$.

Se $\mathbf{q}_1, \mathbf{q}_2 \in \mathbb{R}^{n \times 1}$ forem vectores de Perron à esquerda de $\mathbf{A}$, então $\mathbf{q}_1$ e $\mathbf{q}_2$ são vectores de Perron de $\mathbf{A}^\mathsf{T}$, logo $\mathbf{q}_1 = \mathbf{q}_2$ (pelo que acabámos de provar).

12.4. Temos $\sigma(\mathbf{A}) = \{0, 1\}$, logo $\rho(\mathbf{A}) = 1$ é a raíz de Perron de $\mathbf{A}$. O vector de Perron é $\mathbf{p} = \frac{1}{\alpha+\beta} \begin{bmatrix} \beta \\ \alpha \end{bmatrix}$.

12.5. (a) e (b) Temos $\sigma(r^{-1}\mathbf{A}) = \{r^{-1}\lambda \colon \lambda \in \sigma(\mathbf{A})\}$, logo $\rho(r^{-1}\mathbf{A}) = r^{-1}\rho(\mathbf{A}) = 1$. Além disso, pelo teorema de Perron (aplicado a $r^{-1}\mathbf{A} > 0$, $1 = \rho(r^{-1}\mathbf{A})$ é o único valor próprio $\lambda \in \sigma(\mathbf{A})$ tal que $|\lambda| = 1$ e tem-se m.g.$(1) = 1$ (logo m.a.$(1) =$ m.g.(1)). Sendo assim, o Teorema 20.2 garante que a sucessão $\big((r^{-1}\mathbf{A})^k\big)_{k \in \mathbb{N}}$ é convergente com $\lim_{k\to\infty} (r^{-1}\mathbf{A})^k = \mathbf{G}$ onde $\mathbf{G} \in \mathbb{C}^{n \times 1}$ é o projector espectral de $r^{-1}\mathbf{A}$ associado a 1. Notemos que $\mathbf{G} \in \mathbb{R}^{n \times n}$ porque $(r^{-1}\mathbf{A})^k \in \mathbb{R}^{n \times n}$ para qualquer $k \in \mathbb{N}$.

(c) Temos $\mathcal{R}(\mathbf{G}) = \mathcal{N}(r^{-1}\mathbf{A} - \mathbf{I}_n)$, logo $r(\mathbf{G}) =$ m.g.$(1) = 1$.

12.6. (a) Para cada $1 \leqslant j \leqslant n$, a aresta E_j liga um vértice P_i, $1 \leqslant i \leqslant m$, a um vértice P_k, $1 \leqslant k \leqslant m$, de modo que a j-ésima coluna de $\mathbf{A}$ tem apenas duas entradas não-nulas: 1 na i-ésima linha e -1 na k-ésima linha. Sendo assim, a soma das entradas de cada qualquer linha de $\mathbf{A}^\mathsf{T}$ (isto é, de qualquer coluna de $\mathbf{A}$) é igual a 0 e, portanto, $\mathbf{A}^\mathsf{T}\mathbf{e} = \mathbf{0}$.

(b) Como $\mathbf{e} \in \mathcal{N}(\mathbf{A}^{\mathrm{T}})$, temos $n(\mathbf{A}^{\mathrm{T}}) \geqslant 1$, logo $r(\mathbf{A}^{\mathrm{T}}) = m - n(\mathbf{A}^{\mathrm{T}}) \leqslant m - 1$. Como $r(\mathbf{A}) = r(\mathbf{A}^{\mathrm{T}})$, concluímos que $r(\mathbf{A}) \leqslant m - 1$.

(c) Suponhamos que G é conexo, com vista a provar que $r(\mathbf{A}) = m - 1$ (ou, equivalentemente, que $n(\mathbf{A}^{\mathrm{T}}) = 1$). Para isto, basta provar que $\mathcal{N}(\mathbf{A}^{\mathrm{T}}) = \langle \mathbf{e} \rangle$. Seja $\mathbf{v} \in \mathcal{N}(\mathbf{A}^{\mathrm{T}})$ e, para quaisquer $1 \leqslant i, k \leqslant m$, consideremos as componentes v_i e v_k de $\mathbf{v}$ e os vértices P_i e P_k de G. Como G é conexo, existem $1 \leqslant i_1, \ldots, i_r \leqslant m$ com $i_1 = i$, $i_r = k$ e tais que, para qualquer $1 \leqslant s \leqslant r - 1$, existe uma aresta E_{j_s}, $1 \leqslant j_s \leqslant n$, que liga P_{i_s} a $P_{i_{s+1}}$. Por conseguinte, para cada $1 \leqslant s \leqslant r-1$, a j_s-ésima coluna de $\mathbf{A}$ tem as entradas a_{i_s,j_s} e a_{i_{s+1},j_s} não nulas ($a_{i_s,j_s} = -1$ e $a_{i_{s+1},j_s} = 1$); nesta coluna, todas as outras entradas têm de ser nulas. Como $\mathbf{A}^{\mathrm{T}}\mathbf{v} = \mathbf{0}$, temos $\mathbf{v}^{\mathrm{T}}\mathbf{A} = (\mathbf{A}^{\mathrm{T}}\mathbf{v})^{\mathrm{T}} = \mathbf{0}$ e, portanto, se $\mathbf{c}_{j_s}$ denotar a j_s-ésima coluna de $\mathbf{A}$, temos $\mathbf{v}^{\mathrm{T}}\mathbf{c}_{j_s} = \mathbf{0}$, o que significa que $v_{i_s} = v_{i_{s+1}}$, Como $1 \leqslant s \leqslant r - 1$ é qualquer, concluímos que

$$v_i = v_{i_1} = v_{i_2} = \cdots = v_{i_r} = v_k.$$

Como $1 \leqslant i, k \leqslant m$ são arbitrários, tem de ser $\mathbf{v} = v_1\mathbf{e}$ e, portanto, $\mathcal{N}(\mathbf{A}^{\mathrm{T}}) = \langle \mathbf{e} \rangle$ como se queria.

Reciprocamente, suponhamos que $r(\mathbf{A}) = m - 1$, com vista a provar que G é conexo. Com vista a absurdo, suponhamos que G não é conexo, de modo que G pode ser decomposto como união de dois subgrafos G_1 e G_2. Sem perda de generalidade, podemos admitir que $P_1, \ldots, P_r$, $1 \leqslant r < m$, são os vértices de G_1 e que $P_{r+1}, \ldots, P_m$ são os vértices de G_2; além disso, podemos ordenar as arestas de G de forma a que

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 \end{bmatrix}$$

onde $\mathbf{A}_1$ é a matriz de incidência de G_1 e $\mathbf{A}_2$ é a matriz de incidência de G_2. Ora, temos

$$m - 1 = r(\mathbf{A}) = r(\mathbf{A}_1) + r(\mathbf{A}_2) \leqslant (r-1) + (m-r-1) = m - 2,$$

o que não pode acontecer. Sendo assim, G é conexo.

12.7. Trata-se do Lema 24.1.

12.8. Trata-se da Proposição 24.2.

12.9. Usamos o critério de Frobenius.

Para simplificar os cálculos, começamos por definir, para qualquer matriz não-negativa $\mathbf{A} \in \mathbb{R}^{n\times n}$, a matriz $\beta(\mathbf{A}) = [b_{i,j}]$ onde

$$b_{i,j} = \begin{cases} 1, & \text{se } a_{i,j} > 0, \\ 0, & \text{se } a_{i,j} = 0, \end{cases} \qquad 1 \leqslant i, j \leqslant n;$$

notemos que, para quaisquer $1 \leqslant i, j \leqslant n$ e qualquer $m \in \mathbb{N}$, se tem $(\mathbf{A}^m)_{i,j} > 0$ se e só se a sequência $(\mathbf{A}^m)_{i,j} > 0$, de modo que $\mathbf{A}$ será primitiva se e só se existir $m \in \mathbb{N}$ tal que $\mathbf{B}_m > \mathbf{0}$, onde

$$B_1 = \beta(\mathbf{A}), \quad \mathbf{B}_2 = \beta(\mathbf{B}_1\mathbf{B}_1), \quad \mathbf{B}_3 = \beta(\mathbf{B}_1\mathbf{B}_2), \quad \mathbf{B}_4 = \beta(\mathbf{B}_1\mathbf{B}_3), \qquad \dots$$

No nosso caso particular, temos

$$\mathbf{B}_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \ \mathbf{B}_2 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \ \mathbf{B}_3 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \ \mathbf{B}_4 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \ \text{e} \ \mathbf{B}_5 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

e, portanto, $\mathbf{A}^5 > \mathbf{0}$, provando que $\mathbf{A}$ é primitiva.

12.10. (a) $\mathbf{L}$ é não-negativa porque $b_1, \dots, b_n, s_1, \dots, s_n > 0$.

O grafo $\mathcal{G}(\mathbf{L})$ de $\mathbf{L}$ tem n vértices $P_1, \dots, P_n$ e, para quaisquer $1 \leqslant i, j \leqslant n$, existe o caminho

- $P_1 \to P_j$ se $i = 1$ (porque $\ell_{1,j} = b_j \neq 0$),
- $P_i \to P_{i-1} \to \cdots \to P_j$ se $j < i$ (porque $\ell_{i,i-1} = s_{i-1}$, $\ell_{i-1,i-2} = s_{i-2}$, ..., $\ell_{j+1,j} = s_j$ são não-nulos),
- $P_i \to P_{i-1} \to \cdots \to P_1 \to P_j$ se $i < j$ (porque $\ell_{i,i-1} = s_{i-1}$, $\ell_{i-1,i-2} = s_{i-2}$, ..., $\ell_{1,j} = b_j$ são não-nulos).

Sendo assim, $\mathcal{G}(\mathbf{L})$ é fortemente conexo e, portanto, $\mathbf{L}$ é irredutível.

$\mathbf{L}$ é primitiva porque $\mathbf{L}^{n+2} > \mathbf{0}$.

(b) $\mathbf{f}(t) = \mathbf{L}\mathbf{f}(t-1) = \mathbf{L}^2\mathbf{f}(t-2) = \cdots = \mathbf{L}^t\mathbf{f}(0)$.

(c) Como $\mathbf{L}$ é não-negativa e irredutível, o teorema de Perron-Frobenius garante que $r \in \mathbb{R}_0^+$ é um valor próprio de $\mathbf{L}$ com m.a.$(r) = 1$ (logo, m.g.$(1) = 1$); além disso, existe um e um só vector $\mathbf{p} \in \mathbb{R}^{n\times 1}$ tal que

$$\mathbf{p} > \mathbf{0}, \qquad \mathbf{L}\mathbf{p} = \rho(\mathbf{L})\mathbf{p} \qquad \text{e} \qquad \|\mathbf{p}\|_1 = 1.$$

Por outro lado, como $\mathbf{L}$ é primitiva, r é o único valor próprio $\lambda \in \sigma(\mathbf{A})$ tal que $|\lambda| = \rho(\mathbf{L}) = r$.

Como $\rho(r^{-1}\mathbf{L}) = 1$, podemos usar o Teorema 20.2 para concluir que a sucessão $\big((r^{-1}\mathbf{L})^t\big)_{t\in\mathbb{N}}$ é convergente com

$$\lim_{t\to\infty} (r^{-1}\mathbf{L})^t = \mathbf{G}$$

onde $\mathbf{G} \in \mathbb{R}^{n\times n}$ é o projector espectral de $r^{-1}\mathbf{L}$ associado a $1 \in \sigma(r^{-1}\mathbf{L})$. Sendo assim, temos

$$\lim_{t\to\infty} \frac{1}{r^t}\,\mathbf{f}(t) = \lim_{t\to\infty} \frac{1}{r^t}\,\mathbf{L}^t\mathbf{f}(0) = \mathbf{G}\mathbf{f}(0).$$

Ora, não é difícil verificar que

$$\mathbf{G} = \frac{1}{\mathbf{q}^{\mathrm{T}}\mathbf{p}}\,\mathbf{p}\mathbf{q}^{\mathrm{T}}$$

(basta usar a definição de projector espectral), de modo que

$$\lim_{t\to\infty} \frac{1}{r^t}\,\mathbf{f}(t) = \frac{1}{\mathbf{q}^{\mathrm{T}}\mathbf{p}}\,\mathbf{p}\mathbf{q}^{\mathrm{T}}\mathbf{f}(0) = \frac{\mathbf{q}^{\mathrm{T}}\mathbf{f}(0)}{\mathbf{q}^{\mathrm{T}}\mathbf{p}}\,\mathbf{p}.$$

Além disso, como $\|\star\|_1$ é uma função contínua, a sucessão $\left(\left\|\frac{1}{r^t}\mathbf{f}(t)\right\|_1\right)_{t\in\mathbb{N}}$ é convergente com

$$\lim_{t\to\infty}\left\|\frac{1}{r^t}\,\mathbf{f}(t)\right\|_1 = \left\|\lim_{t\to\infty}\frac{1}{r^t}\,\mathbf{f}(t)\right\|_1 = \left\|\frac{\mathbf{q}^{\mathrm{T}}\mathbf{f}(0)}{\mathbf{q}^{\mathrm{T}}\mathbf{p}}\,\mathbf{p}\right\|_1 = \frac{\mathbf{q}^{\mathrm{T}}\mathbf{f}(0)}{\mathbf{q}^{\mathrm{T}}\mathbf{p}}\,\|\mathbf{p}\|_1 = \frac{\mathbf{q}^{\mathrm{T}}\mathbf{f}(0)}{\mathbf{q}^{\mathrm{T}}\mathbf{p}}.$$

(d) Para qualquer $t \in \mathbb{N}$, temos

$$\frac{1}{\|\mathbf{f}(t)\|_1}\,\mathbf{f}(t) = \frac{r^t}{\|\mathbf{f}(t)\|_1}\left(\frac{1}{r^t}\,\mathbf{f}(t)\right) = \frac{1}{\left\|\frac{1}{r^t}\,\mathbf{f}(t)\right\|_1}\left(\frac{1}{r^t}\,\mathbf{f}(t)\right)$$

e, portanto,

$$\lim_{t\to\infty}\frac{1}{\|\mathbf{f}(t)\|_1}\,\mathbf{f}(t) = \frac{1}{\lim_{t\to\infty}\left\|\frac{1}{r^t}\,\mathbf{f}(t)\right\|_1}\left(\lim_{t\to\infty}\frac{1}{r^t}\,\mathbf{f}(t)\right) = \frac{\mathbf{q}^{\mathrm{T}}\mathbf{p}}{\mathbf{q}^{\mathrm{T}}\mathbf{f}(0)}\left(\frac{\mathbf{q}^{\mathrm{T}}\mathbf{f}(0)}{\mathbf{q}^{\mathrm{T}}\mathbf{p}}\,\mathbf{p}\right) = \mathbf{p}.$$

12.11. Tal como no Exercício 12.9, consideramos a matriz

$$\mathbf{B} = \beta(\mathbf{A}) = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

Para qualquer $n \in \mathbb{N}$, $n \geqslant 2$, temos

$$\beta(\mathbf{A}^{2n}) = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{bmatrix},$$

de modo que não pode existir $m \in \mathbb{N}$ tal que $\mathbf{A}^m > \mathbf{0}$. Pelo critério de Frobenius, $\mathbf{A}$ não é primitiva.

Alternativamente, temos $\sigma(\mathbf{A}) = \{0, \pm 2, \pm 8\}$, logo $\rho(\mathbf{A}) = 8$. $\mathbf{A}$ não é primitiva porque 8 não é o único valor próprio $\lambda \in \sigma(\mathbf{A})$ com $|\lambda| = 8$.

12.12. (a) Consideramos a norma matricial $\|\star\|'_\infty$ induzida pela norma $\|\star\|_\infty$ em $\mathbb{C}^{n\times 1}$. Temos

$$\|\mathbf{S}\|'_\infty = \max_{1\leqslant i\leqslant n}\sum_{1\leqslant j\leqslant n}|s_{i,j}| = \max_{1\leqslant i\leqslant n}\sum_{1\leqslant j\leqslant n}s_{i,j} \leqslant 1$$

e, portanto,

$$\rho(\mathbf{S}) \leqslant \|\mathbf{S}\|'_\infty \leqslant 1.$$

(b) Suponhamos que $\mathbf{S}$ é irredutível e seja $\mathbf{e} = \begin{bmatrix}1\ 1\ \cdots\ 1\end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^{n\times 1}$. Temos

$$(\mathbf{Se})_i = \sum_{1\leqslant j\leqslant n} s_{i,j} \leqslant 1$$

e, portanto, $\mathbf{Se} \leqslant \mathbf{e}$. Suponhamos que $\rho(\mathbf{S}) = 1$. Como $\mathbf{S}$ é irredutível, também $\mathbf{S}^{\mathsf{T}}$ é irredutível e, portanto, o teorema de Perron-Frobenius aplica-se a $\mathbf{S}^{\mathsf{T}}$. Em particular, $\mathbf{S}^{\mathsf{T}}$ tem um vector de Perron $\mathbf{q} \in \mathbb{R}^{n\times 1}$; por definição, temos $\mathbf{q} > \mathbf{0}$, $\mathbf{S}^{\mathsf{T}}\mathbf{q} = \rho(\mathbf{S}^{\mathsf{T}})\mathbf{q} = \rho(\mathbf{S})\mathbf{q} = \mathbf{q}$ e $\|\mathbf{q}\|_1 = 1$. Agora, temos $(\mathbf{S} - \mathbf{I}_n)\mathbf{e} = \mathbf{Se} - \mathbf{e} \geqslant \mathbf{0}$ e, portanto, $\mathbf{q}^{\mathsf{T}}(\mathbf{S} - \mathbf{I}_n)\mathbf{e} > 0$. No entanto,

$$\mathbf{q}^{\mathsf{T}}(\mathbf{S} - \mathbf{I}_n) = \mathbf{q}^{\mathsf{T}}\mathbf{S} - \mathbf{q}^{\mathsf{T}} = (\mathbf{S}^{\mathsf{T}}\mathbf{q})^{\mathsf{T}} - \mathbf{q}^{\mathsf{T}} = \mathbf{q}^{\mathsf{T}} - \mathbf{q}^{\mathsf{T}} = \mathbf{0},$$

uma contradição. Sendo assim, tem de ser $\rho(\mathbf{S}) < 1$. Como se queria.

12.13. Pelo teorema de Perron-Frobenius, $\mathbf{A}$ tem um vector de Perron, isto é, existe um vector $\mathbf{p} \in \mathbb{R}^{n\times 1}$ tal que

$$\mathbf{p} > \mathbf{0}, \qquad \mathbf{Ap} = \rho(\mathbf{A})\mathbf{p} \qquad \text{e} \qquad \|\mathbf{p}\|_1 = 1.$$

Consideremos a matriz

$$\mathbf{D} = \begin{bmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & p_n \end{bmatrix}$$

onde $p_1, \ldots, p_n \in \mathbb{R}^+$ são as componentes de $\mathbf{p}$. Seja $\mathbf{e} = \begin{bmatrix}1\ 1\ \cdots\ 1\end{bmatrix} \in \mathbb{R}^{n\times 1}$. Temos $\mathbf{De} = \mathbf{p}$, logo

$$\mathbf{D}^{-1}\mathbf{ADe} = \mathbf{D}^{-1}\mathbf{Ap} = \rho(\mathbf{A})\mathbf{D}^{-1}\mathbf{p} = \rho(\mathbf{A})\mathbf{e}.$$

Sendo assim, temos

$$\sum_{1\leqslant j\leqslant n} (\mathbf{D}^{-1}\mathbf{AD})_{i,j} = \rho(\mathbf{A}), \qquad 1 \leqslant i \leqslant n,$$

o que significa que $\mathbf{P} = \rho(\mathbf{A})^{-1}\mathbf{D}^{-1}\mathbf{AD}$ é uma matriz estocástica que verifica $\mathbf{D}^{-1}\mathbf{AD} = \rho(\mathbf{A})\mathbf{P}$; notemos que $\rho(\mathbf{A}) > 0$ (pelo teorema de Perron-Frobenius).

12.14. Como $\mathbf{P} \geqslant \mathbf{0}$ é irredutível, $\rho(\mathbf{P}) = 1$ é tal que m.a.(1) = m.g.$(1) = 1$ e, portanto,

$$r(\mathbf{P} - \mathbf{I}_n) = n - n(\mathbf{P} - \mathbf{I}_n) = n - \text{m.g.}(1) = n - 1.$$

12.15. Consideramos a matriz $\mathbf{A} = \mathbf{P} - \mathbf{I}_n$. Pelo exercício anterior, temos $r(\mathbf{A}) = n - 1$ e, portanto, $\mathbf{A}$ não é invertível. Considerando a matriz adjunta $\operatorname{adj}(\mathbf{A})$ de $\mathbf{A}$, a regra de Laplace para o determinante garante que

$$\mathbf{A}\operatorname{adj}(\mathbf{A}) = \det(\mathbf{A})\mathbf{I}_n = \mathbf{0};$$

além disso, como $r(\mathbf{A}) = n - 1$, tem-se $r(\operatorname{adj}(\mathbf{A})) = 1$. Como $\mathbf{A}\operatorname{adj}(\mathbf{A}) = \mathbf{0}$, o teorema de Perron-Frobenius garante que qualquer coluna de $\operatorname{adj}(\mathbf{A})$ é um múltiplo escalar do vector de

Perron de $\mathbf{P}$ (logo, um múltiplo escalar do vector $\mathbf{e} = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}$). Por conseguinte, temos $\operatorname{adj}(\mathbf{A}) = \mathbf{e}\mathbf{v}^{\mathsf{T}}$ para algum vector $\mathbf{v} \in \mathbb{R}^{n\times 1}$. Como $\operatorname{adj}(\mathbf{A})_{i,i} = P_i$, concluímos que $\mathbf{v} = \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_n \end{bmatrix}$.

De maneira análoga, a igualdade $(\operatorname{adj}(\mathbf{A}))\mathbf{A} = \mathbf{0}$ implica que qualquer linha de $\operatorname{adj}(\mathbf{A})$ é um múltiplo escalar de $\mathbf{q}^{\mathsf{T}}$ e, portanto, $\mathbf{v}^{\mathsf{T}} = \alpha\mathbf{q}^{\mathsf{T}}$ para algum $\alpha \in \mathbb{R}$. Se $\alpha = 0$, então $\mathbf{v} = \mathbf{0}$ e, portanto $\operatorname{adj}(\mathbf{A}) = \mathbf{0}$, o que não acontece (porque $r(\operatorname{adj}(\mathbf{A})) = 1$). Segue-se que $\mathbf{v}^{\mathsf{T}}\mathbf{e} = \alpha \neq 0$ e, portanto, $\mathbf{q} = \frac{1}{\mathbf{v}^{\mathsf{T}}\mathbf{e}}\mathbf{v}$, como se queria.