# AULA 17

Sumário. Polinómio mínimo de uma matriz.

▷ Dizemos que um polinómio $p(x) \in \mathbb{C}[x]$, $p(x) = a_0 + a_1x + \cdots + a_nx^n$, ANULA uma matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$ se

$$p(\mathbf{A}) = a_0\mathbf{I}_n + a_1\mathbf{A} + \cdots + a_n\mathbf{A}^n = \mathbf{0};$$

por exemplo, para qualquer matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$, o polinómio característico $p_{\mathbf{A}}(x) \in \mathbb{C}[x]$ é um polinómio anulador de $\mathbf{A}$ (teorema de Cayley-Hamilton).

Teorema 17.1. *Seja $\mathbf{A} \in \mathbb{C}^{n\times n}$. Então:*

(a) *Existe um e um só polinómio mónico*[(†)] $m_{\mathbf{A}}(x) \in \mathbb{C}[x]$ *de grau mínimo que anula* $\mathbf{A}$*.*

(b) $\operatorname{gr} m_{\mathbf{A}}(x) \leqslant n$.

(c) *Se $p(x) \in \mathbb{C}[x]$ é um polinómio que anula $\mathbf{A}$, então existe um (e um só) polinómio $q(x) \in \mathbb{C}[x]$ tal que $p(x) = m_{\mathbf{A}}(x)q(x)$.*

Demonstração. O conjunto $\mathcal{A} = \{p(x) \in \mathbb{C}[x] \colon p(x) \neq 0,\ p(\mathbf{A})\}$ é não-vazio (porque $p_{\mathbf{A}}(x) \in \mathcal{A}$) e, portanto, existe um polinómio $m(x) \in \mathcal{A}$ tal que

$$\operatorname{gr} m(x) = \min \{\operatorname{gr} p(x) \colon p(x) \in \mathcal{A}\}.$$

Se $\operatorname{gr} m(x) = r$ e $m(x) = a_0 + a_1x + \cdots + a_rx^r$, então $m_{\mathbf{A}}(x) = a_r^{-1}m(x)$ é um polinómio mónico com $\operatorname{gr} a_r^{-1}m(x) = \operatorname{gr} m(x)$. Provámos assim que existe um polinómio nas condições de (a). A sua unicidade resulta de (c).

Seja $p(x) \in \mathcal{A}$ arbitrário. Dividindo $p(x)$ por $m_{\mathbf{A}}(x)$, obtemos polinómios $q(x), r(x) \in \mathbb{C}[x]$ tais que

$$p(x) = m_{\mathbf{A}}(x)q(x) + r(x) \qquad \text{e} \qquad \operatorname{gr} r(x) < \operatorname{gr} m_{\mathbf{A}}(x)$$

(permitimos que $r(x) = 0$ convencionando este caso acontece se e só se $\operatorname{gr} r(x) = -\infty$). Temos $r(x) = p(x) - m_{\mathbf{A}}(x)q(x)$ e, portanto,

$$r(\mathbf{A}) = p(\mathbf{A}) - m_{\mathbf{A}}(\mathbf{A})q(\mathbf{A}) = \mathbf{0}$$

(porque $p(x), m_{\mathbf{A}}(x) \in \mathcal{A}$). Como $\operatorname{gr} r(x) < \operatorname{gr} m_{\mathbf{A}}(x)$, tem de ser $r(x) = 0$, logo $p(x) = m_{\mathbf{A}}(x)q(x)$, o que prova (c).

---

[(†)]Um polinómio $p(x) \in \mathbb{C}[x]$, $p(x) = a_0 + a_1x + \cdots + a_nx^n$, diz-se MÓNICO se $a_n = 1$.

Para (a), seja $m'(x) \in \mathbb{C}[x]$ for um polinómio mónico de grau mínimo que anula $\mathbf{A}$. Então, por (c), $m'(x) = m_{\mathbf{A}}(x)q(x)$ para algum $q(x) \in \mathbb{C}[x]$, logo

$$\operatorname{gr} m'(x) = \operatorname{gr} m_{\mathbf{A}}(x) + \operatorname{gr} q(x) \leqslant \operatorname{gr} m_{\mathbf{A}}(x) \operatorname{gr} m'(x)$$

e, portanto, $\operatorname{gr} q(x) = 0$. Segue-se que $q(x) = a \in \mathbb{C}$. Como $m'(x)$ e $m_{\mathbf{A}}(x)$ são mónicos, tem de ser $a = 1$, logo $m'(x) = m_{\mathbf{A}}(x)$.

A alínea (b) resulta imediatamente de ser ter $p_{\mathbf{A}}(x) \in \mathcal{A}$ e $\operatorname{gr} p_{\mathbf{A}}(x) = n$. □

▷ Para qualquer $\mathbf{A} \in \mathbb{C}^{n\times n}$, dizemos que $m_{\mathbf{A}}(x)$ é o POLINÓMIO MÍNIMO de $\mathbf{A}$.

COROLÁRIO 17.2. *Se $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$ forem matrizes semelhantes, então $m_{\mathbf{A}}(x) = m_{\mathbf{B}}(x)$.*

DEMONSTRAÇÃO. Sejam Se $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$ matrizes semelhantes e seja $\mathbf{P} \in \mathbb{C}^{n\times n}$ uma matriz invertível tal que $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. Pondo $m_{\mathbf{B}}(x) = a_0 + a_1 x + \cdots + a_r x^r$ em que $a_r = 1$, temos

$$\begin{aligned} m_{\mathbf{B}}(\mathbf{A}) &= a_0\mathbf{I}_n + a_1\mathbf{A} + \cdots + a_r\mathbf{A}^r \\ &= a_0\mathbf{I}_n + a_1(\mathbf{P}\mathbf{B}\mathbf{P}^{-1}) + \cdots + a_r(\mathbf{P}\mathbf{B}\mathbf{P}^{-1})^r \\ &= \mathbf{P}\big(a_0\mathbf{I}_n + a_1\mathbf{B} + \cdots + a_r\mathbf{B}^r\big)\mathbf{P} = \mathbf{P}m_{\mathbf{B}}(\mathbf{B})\mathbf{P}^{-1} = \mathbf{0}, \end{aligned}$$

logo $m_{\mathbf{B}}(x) = m_{\mathbf{A}}(x)q(x)$ para algum $q(x) \in \mathbb{C}[x]$ (pela alínea (c) do teorema anterior); em particular, $\operatorname{gr} m_{\mathbf{A}}(x) \leqslant \operatorname{gr} m_{\mathbf{B}}(x)$.

Analogamente, concluímos que $m_{\mathbf{A}}(x) = m_{\mathbf{B}}(x)q'(x)$ para algum $q'(x) \in \mathbb{C}[x]$ e, em particular, $\operatorname{gr} m_{\mathbf{B}}(x) \leqslant \operatorname{gr} m_{\mathbf{A}}(x)$. Sendo assim, tem de ser

$$\operatorname{gr} m_{\mathbf{B}}(x) = \operatorname{gr} m_{\mathbf{A}}(x).$$

Como $m_{\mathbf{B}}(x) = m_{\mathbf{A}}(x)q(x)$ e como $m_{\mathbf{A}}(x)$ e $m_{\mathbf{B}}(x)$ são polinómios mónicos, concluímos $q(x) = 1$, logo $m_{\mathbf{B}}(x) = m_{\mathbf{A}}(x)$. □

▷ O recíproco do teorema anterior nem sempre é verdadeiro: as matrizes

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \in \mathbb{C}^{4\times 4} \qquad \text{e} \qquad \mathbf{B} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \in \mathbb{C}^{4\times 4}$$

não são semelhantes, mas têm o mesmo polinómio mínimo.

COROLÁRIO 17.3. *Para qualquer $\mathbf{A} \in \mathbb{C}^{n\times n}$ e qualquer $\lambda \in \mathbb{C}$, $m_{\mathbf{A}}(\lambda) = 0$ se e só se $\lambda \in \sigma(\mathbf{A})$.*

DEMONSTRAÇÃO. Como $p_{\mathbf{A}}(\mathbf{A}) = \mathbf{0}$ (pelo teorema de Cayley-Hamilton), existe $q(x) \in \mathbb{C}[x]$ tal que $p_{\mathbf{A}}(x) = m_{\mathbf{A}}(x)q(x)$. Como $\lambda \in \sigma(\mathbf{A})$ se e só se $p_{\mathbf{A}}(\lambda) = 0$, temos $\lambda \in \sigma(\mathbf{A})$ sempre que $m_{\mathbf{A}}(\lambda) = 0$.

Reciprocamente, suponhamos que $\lambda \in \sigma(\mathbf{A})$ e seja $\mathbf{0} \neq \mathbf{v} \in \mathbb{C}^{n\times 1}$ tal que $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$. Se $m_{\mathbf{A}}(x) = a_0 + a_1 x + \cdots + a_r x^r$ em que $a_r = 1$, então

$$\begin{aligned} m_{\mathbf{A}}(\lambda)\mathbf{v} &= \big(a_0 + a_1\lambda + \cdots + a_r\lambda^r\big)\mathbf{v} = a_0\mathbf{v} + a_1\lambda\mathbf{v} + \cdots + a_r\lambda^r\mathbf{v} \\ &= a_0\mathbf{v} + a_1\mathbf{A}\mathbf{v} + \cdots + a_r\mathbf{A}^r\mathbf{v} = \big(a_0\mathbf{I}_n + a_1\mathbf{A} + \cdots + a_r\mathbf{A}^r\big)\mathbf{v} \\ &= m_{\mathbf{A}}(\mathbf{A})\mathbf{v} = \mathbf{0}\mathbf{v} = \mathbf{0} \end{aligned}$$

e, portanto, $m_{\mathbf{A}}(\lambda) = 0$ (porque $\mathbf{v} \neq \mathbf{0}$). $\square$

$\triangleright$ Seja $\mathbf{A} \in \mathbb{C}^{n\times n}$ e sejam $\lambda_1, \ldots, \lambda_t \in \mathbb{C}$ tais que $\lambda_i \neq \lambda_j$, para $1 \leqslant i \neq j \leqslant t$, e $\sigma(\mathbf{A}) = \{\lambda_1, \ldots, \lambda_r\}$. Então,

$$p_{\mathbf{A}}(x) = (x-\lambda_1)^{m_1}\cdots(x-\lambda_t)^{m_t}$$

onde $m_i = \text{m.a.}(\lambda_i)$ para $1 \leqslant i \leqslant t$. Pelo corolário anterior, temos

$$m_{\mathbf{A}}(x) = (x-\lambda_1)^{r_1}\cdots(x-\lambda_t)^{r_t}, \qquad 0 < r_i \leqslant m_i,\ 1 \leqslant i \leqslant t.$$

Teorema 17.4. *Seja $\mathbf{A} \in \mathbb{C}^{n\times n}$ e sejam $\lambda_1, \ldots, \lambda_t \in \mathbb{C}$ tais que $\lambda_i \neq \lambda_j$, para $1 \leqslant i \neq j \leqslant t$, e $\sigma(\mathbf{A}) = \{\lambda_1, \ldots, \lambda_r\}$. Então,*

$$m_{\mathbf{A}}(x) = (x-\lambda_1)^{r_1}\cdots(x-\lambda_t)^{r_t}$$

*onde $r_1, \ldots, r_t \in \mathbb{N}$ são tais que, para cada $1 \leqslant i \leqslant t$, $\mathbf{J}_{r_i}(\lambda_i)$ é o maior bloco de Jordan na forma canónica de Jordan de $\mathbf{A}$.*

Demonstração. Seja $\mathbf{P} \in \mathbb{C}^{n\times n}$ tal que

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J} = \begin{bmatrix} \mathbf{J}(\lambda_1) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}(\lambda_2) & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{J}(\lambda_t) \end{bmatrix}$$

onde, para cada $1 \leqslant i \leqslant t$,

$$\mathbf{J}(\lambda_i) = \begin{bmatrix} \mathbf{J}_{r_{i,1}}(\lambda_i) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{r_{i,2}}(\lambda_i) & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{J}_{r_{i,t_i}}(\lambda_i) \end{bmatrix}, \qquad r_{i,1} \geqslant r_{i,2} \geqslant \cdots \geqslant r_{i,t_i} > 0.$$

Pelo Corolário 21.2, temos $m_{\mathbf{A}}(x) = m_{\mathbf{J}}(x)$. Para qualquer $1 \leqslant i \leqslant t$, temos $r_i = r_{i,1}$, $r_{i,1} + \cdots + r_{i,t_i} = m_i = \text{m.a.}(\lambda_i)$ e

$$\big(\mathbf{J}(\lambda_i) - \lambda_i\mathbf{I}_{m_i}\big)^{r_i} = \begin{bmatrix} \mathbf{J}_{r_{i,1}}(0)^{r_i} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{r_{i,2}}(0)^{r_i} & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{J}_{r_{i,t_i}}(0)^{r_i} \end{bmatrix} = \mathbf{0}.$$

Daqui, resulta que

$$(\mathbf{J}-\lambda_i\mathbf{I}_n)^{r_i} = \begin{bmatrix} \big(\mathbf{J}(\lambda_1)-\lambda_i\mathbf{I}_{m_1}\big)^{r_i} & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & \vdots & & \vdots \\ \mathbf{0} & \cdots & \big(\mathbf{J}(\lambda_i)-\lambda_i\mathbf{I}_{m_i}\big)^{r_i} & \cdots & \mathbf{0} \\ \vdots & & \vdots & & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \cdots & \big(\mathbf{J}(\lambda_t)-\lambda_i\mathbf{I}_{m_t}\big)^{r_i} \end{bmatrix}$$

$$= \begin{bmatrix} \big(\mathbf{J}(\lambda_1)-\lambda_i\mathbf{I}_{m_1}\big)^{r_i} & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & \vdots & & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & \vdots & & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \cdots & \big(\mathbf{J}(\lambda_t)-\lambda_i\mathbf{I}_{m_t}\big)^{r_i} \end{bmatrix}$$

e, portanto,

$$(\mathbf{J}-\lambda_1\mathbf{I}_n)^{r_1}(\mathbf{J}-\lambda_2\mathbf{I}_n)^{r_2}\cdots(\mathbf{J}-\lambda_t\mathbf{I}_n)^{r_t}$$

$$= \begin{bmatrix} \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \star & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \star \end{bmatrix} \begin{bmatrix} \star & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \star \end{bmatrix} \cdots \begin{bmatrix} \star & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \star & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} = \mathbf{0}.$$

Sendo assim, pondo

$$m(x) = (x-\lambda_1)^{r_1}\cdots(x-\lambda_t)^{r_t} \in \mathbb{C}[x],$$

temos $m(\mathbf{J}) = \mathbf{0}$, logo existe $q(x) \in \mathbb{C}[x]$ tal que

$$m(x) = m_{\mathbf{J}}(x)q(x).$$

Pelo Corolário 21.3, temos

$$m_{\mathbf{J}}(x) = (x-\lambda_1)^{s_1}\cdots(x-\lambda_t)^{s_t}, \qquad s_1,\ldots,s_t \in \mathbb{N},$$

e, portanto, tem de ser

$$s_i \leqslant r_i, \qquad 1 \leqslant i \leqslant t.$$

Como $m_{\mathbf{J}}(\mathbf{J}) = \mathbf{0}$, obtemos

$$(\mathbf{J}-\lambda_1\mathbf{I}_n)^{s_1}\cdots(\mathbf{J}-\lambda_t\mathbf{I}_n)^{s_t} = \mathbf{0}.$$

Ora,

$$(\mathbf{J}-\lambda_i\mathbf{I}_n)^{s_i} = \begin{bmatrix} \big(\mathbf{J}(\lambda_1)-\lambda_i\mathbf{I}_{m_1}\big)^{s_i} & \cdots & \mathbf{0} \\ \vdots & & \vdots \\ \mathbf{0} & \cdots & \big(\mathbf{J}(\lambda_t)-\lambda_i\mathbf{I}_{m_t}\big)^{s_i} \end{bmatrix},$$

logo

$$
\begin{aligned}
\mathbf{0} &= (\mathbf{J}-\lambda_1\mathbf{I}_n)^{s_1}\cdots(\mathbf{J}-\lambda_t\mathbf{I}_n)^{s_t}\\
&= \begin{bmatrix} \prod_{1\leqslant i\leqslant t}\big(\mathbf{J}(\lambda_1)-\lambda_i\mathbf{I}_{m_1}\big)^{s_i} & \cdots & \mathbf{0}\\ \vdots & & \vdots\\ \mathbf{0} & \cdots & \prod_{1\leqslant i\leqslant t}\big(\mathbf{J}(\lambda_t)-\lambda_i\mathbf{I}_{m_1}\big)^{s_i}\end{bmatrix}
\end{aligned}
$$

e, portanto,

$$\prod_{1\leqslant i\leqslant t}\big(\mathbf{J}(\lambda_k)-\lambda_i\mathbf{I}_{m_k}\big)^{s_i}=\mathbf{0},\qquad 1\leqslant k\leqslant t.$$

Para cada $1\leqslant k\leqslant t$, a matriz $\prod_{1\leqslant i\neq k\leqslant t}\big(\mathbf{J}(\lambda_k)-\lambda_i\mathbf{I}_{m_k}\big)^{s_i}$ é invertível, de modo que

$$\big(\mathbf{J}(\lambda_k)-\lambda_k\mathbf{I}_{m_k}\big)^{s_k}=\mathbf{0},\qquad 1\leqslant k\leqslant t.$$

Segue-se que

$$\mathbf{J}_{r_k}(0)^{s_k}=\big(\mathbf{J}_{r_k}(\lambda_k)-\lambda_k\mathbf{I}_{r_k}\big)^{s_k}=\mathbf{0}$$

e, portanto, tem de ser $r_k\leqslant s_k$ para qualquer $1\leqslant k\leqslant t$. Em conclusão,

$$m_{\mathbf{A}}(x)=m_{\mathbf{J}}(x)=(x-\lambda_1)^{r_1}\cdots(x-\lambda_t)^{r_t},$$

como se queria. □

Corolário 17.5. *Seja $\mathbf{A}\in\mathbb{C}^{n\times n}$ e sejam $\lambda_1,\ldots,\lambda_t\in\mathbb{C}$ tais que $\lambda_i\neq\lambda_j$, para $1\leqslant i\neq j\leqslant t$, e $\sigma(\mathbf{A})=\{\lambda_1,\ldots,\lambda_r\}$. Além disso, seja*

$$q(x)=(x-\lambda_1)\cdots(x-\lambda_t)\in\mathbb{C}[x].$$

*Então, $\mathbf{A}$ é diagonalizável se e só se $q(\mathbf{A})=\mathbf{0}$; em particular, $\mathbf{A}$ é diagonalizável se e só se $m_{\mathbf{A}}(x)=(x-\lambda_1)\cdots(x-\lambda_t)$.*

Demonstração. Segue-se imediatamente do teorema anterior uma vez que $\mathbf{A}$ será diagonalizável se e só se todos os blocos de Jordan na sua forma canónica de Jordan forem de tipo $1\times 1$. □