7. Il polinomio minimale di una matrice

Situazione 7.1. Sia $A \in \mathbb{C}_n^n$.

Lemma 7.2. *U sia un'algebra su $\mathbb{C}$ con elemento neutro 1_U.*

Per $f = a_0x^n + a_1x^{n+1} + \cdots + a_n \in \mathbb{C}[x]$ ed $u \in U$ poniamo

$$f(u) := a_0u^n + a_1u^{n+1} + \cdots + a_n 1_U.$$

Allora per ogni $u \in U$, l'applicazione $\underset{f}{\bigcirc} f(u) : \mathbb{C}[x] \longrightarrow U$ è un omomorfismo di $\mathbb{C}$-algebre.

Abbiamo quindi $f(1) = 1_U$ e, per $f, g \in \mathbb{C}[x]$ e $\alpha \in \mathbb{C}$, le relazioni

$$(f+g)(u) = f(u) + g(u)$$
$$(\alpha f)(u) = \alpha f(u)$$
$$(f \cdot g)(u) = f(u) \cdot g(u)$$

Nel seguito useremo questi risultati per $U = \mathbb{C}_n^n$.

Dimostrazione. Dimostriamo solo l'ultima relazione, in quanto le altre sono evidenti.

Siano $f = a_0x^n + a_1x^{n+1} + \cdots + a_n$, $g = b_0x^m + b_1x^{m+1} + \cdots + b_m$ ed $h := f \cdot g$. Allora $h = \sum_{k=0}^{n+m} c_kx^k$ con $c_k = \sum_{j=0}^{k} a_jb_{n-j}$ per ogni k.

Ma questi sono proprio i coefficienti che si ottengono raccogliendo i coefficienti delle potenze di x con lo stesso esponente nel prodotto espanso $(a_0x^n + a_1x^{n-1} + \cdots + a_n)(b_0x^m + b_1x^{m-1} + \cdots + b_m)$ ed è quindi chiaro che, sostituendo x con u, si ottiene proprio $h(u) = \sum_{k=0}^{n+m} c_ku^k$.

Corollario 7.3. *Siano $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ ed $f = (x - \lambda_1) \cdots (x - \lambda_n)$.*

Allora $f(A) = (A - \lambda_1\delta) \cdots (A - \lambda_n\delta)$.

Definizione 7.4. Se calcoliamo l'espressione $\det(x\delta - A) =: \mathcal{P}_A$ in $\mathbb{C}[x]$, otteniamo un polinomio monico in $\mathbb{C}[x]$ che si chiama il *polinomio caratteristico* di A.

Definizione 7.5. Sia $B \in \mathbb{C}_n^n$. Le matrici A e B si dicono *simili*, se esiste $T \in GL(n, U)$ tale che $B = T^{-1}AT$.

Proposizione 7.6. *Matrici simili hanno lo stesso polinomio caratteristico.*

Dimostrazione. Segue immediatamente dalla definizione.

Teorema 7.7. *Sia $\lambda \in \mathbb{C}$. Le condizioni* (1) *e* (2) *sono equivalenti tra di loro e, quando $n \geq 2$, equivalenti alla* (3)*:*

(1) *λ è un autovalore di A.*
(2) *λ è radice del polinomio caratteristico di A :* $\mathcal{P}_A(\lambda) = 0$.
(3) *Esiste una matrice $B \in \mathbb{C}^{n-1}_{n-1}$ tale che A è simile a una matrice della forma $\begin{pmatrix} \lambda & * \\ 0 & B \end{pmatrix}$. In questo caso $\mathcal{P}_A = (x - \lambda) \cdot \mathcal{P}_B$.*

Dimostrazione. Corsi di geometria oppure Koecher, pag.234.

Proposizione 7.8. *I sia un ideale (possibilmente improprio) di $\mathbb{C}[x]$. Se I $\neq 0$, esiste un unico polinomio monico $p \in \mathbb{C}[x]$ che genera I.*

Il grado di p è il più piccolo grado di un polinomio non nullo contenuto in I.

Dimostrazione. Corso di algebra.

Definizione 7.9. $\mathbb{C}[A] := \{f(A) \mid f \in \mathbb{C}[x]\}$.

È immediato che $\mathbb{C}[A]$ è una sotto-$\mathbb{C}$-algebra di $\mathbb{C}^n_n$.

Lemma 7.10. *Esiste un polinomio $f \in \mathbb{C}[x]$ con $f \neq 0$ ed $f(A) = 0$.*

Dimostrazione. Siccome $\mathbb{C}[A]$ è un sottospazio vettoriale di $\mathbb{C}^n_n$, sicuramente $\dim_{\mathbb{C}} \mathbb{C}^n_n \leq n^2 =: s$. Ciò significa che le $s{+}1$ matrici $\delta, A, A^2, \ldots, A^s$ sono linearmente dipendenti, perciò esistono $a_0, a_1, \ldots, a_s \in \mathbb{C}$ non tutti nulli, tali che $a_0\delta + a_1 A + a_2 A^2 + \ldots + a_s A^s = 0$. Se poniamo $f := a_0 + a_1 x + \ldots + a_s x^s$ abbiamo trovato un polinomio $f \neq 0$ con $f(A) = 0$.

Definizione 7.11. Dal lemma 7.2 è immediato che l'insieme $\{f \in \mathbb{C}[x] \mid f(A) = 0\}$ è un ideale (possibilmente improprio) di $\mathbb{C}[x]$.

Dal lemma 7.10 sappiamo che questo ideale è $\neq 0$, quindi per la proposizione 7.8 esiste un unico polinomio monico $\mathcal{M}_A \in \mathbb{C}[x]$ che genera questo ideale.

Per ogni $f \in \mathbb{C}[x]$ si ha quindi$f(A) = 0$ se e solo se esiste $g \in \mathbb{C}[x]$ con $f = g \cdot \mathcal{M}_A$. Il grado di $\mathcal{M}_A$ è allo stesso tempo il più piccolo grado di un polinomio $\neq 0$ che annulla A.

$\mathcal{M}_A$ si chiama il *polinomio minimale* di A. Per definizione $\mathcal{M}_A \neq 0$.

Corollario 7.12. *m sia il grado di $\mathcal{M}_A$. Allora:*

(1) *Le matrici $\delta, A, A^2, ..., A^{m-1}$ formano una base di $\mathbb{C}[A]$.*
(2) $\dim_{\mathbb{C}} \mathbb{C}[A] = m$.

Dimostrazione. Dal lemma 7.10 sappiamo che $\mathcal{M}_A \neq 0$, quindi $\mathcal{M}_A$ è della forma $\mathcal{M}_A = x^m + a_1 x^{m-1} + \ldots + a_m$.

Abbiamo perciò $A^m = -a_1 A^{m-1} - \ldots - a_m\delta$ e ciò implica che $\delta, A, A^2, \ldots, A^{m-1}$ generano lo spazio vettoriale $\mathbb{C}[A]$.

Se queste matrici fossero linearmente dipendenti, ad esempio $b_0\delta + b_1A + \ldots + b_{m-1}A^{m-1} = 0$ con i coefficienti b_j non tutti nulli, allora con $g := b_0 + b_1x + \ldots + b_{m-1}x^{m-1}$ avremmo trovato un polinomio $g \neq 0$ con $g(A) = 0$ e grado minore di m, in contraddizione alla prop.7.8.

Proposizione 7.13. *A è invertibile se e solo se $\mathcal{M}_A(0) \neq 0$.*

In tal caso $A^{-1} \in \mathbb{C}[A]$.

Dimostrazione. Sia di nuovo $\mathcal{M}_A = x^m + a_1x^{m-1} + \ldots + a_m$. Allora $A^m + a_1A^{m-1} + \ldots + a_{m-1}A = -a_m\delta$ e quindi $AB = -a_m\delta$ con $B := A^{m-1} + a_1A^{m-2} + \ldots + a_{m-1}\delta$. Per il corollario 7.12 $B \neq 0$.

Si noti che $\mathcal{M}_A(0) = a_m$.

(1) Sia A invertibile.
Allora $-a_mA^{-1} = B \neq 0$ e quindi $a_m \neq 0$.

(2) Sia $a_m \neq 0$.
Allora $A^{-1} = -\dfrac{1}{a_m}B \in \mathbb{C}[A]$.

Esempio 7.14. Assumiamo di sapere che $\mathcal{M}_A = x^3 + 2x + 5$. Allora

$$-5\delta = A^3 + 2A, \text{ ovvero } A(A^2 + 2\delta) = -5\delta$$

e quindi

$$A^{-1} = -\frac{1}{5}(A^2 + 2\delta) \in \mathbb{C}[A].$$

Corollario 7.15. *A sia invertibile. Allora esiste $f \in \mathbb{C}[x]$ con $f(A) = \delta$ ed $f(0) = 0$.*

Dimostrazione. Per la proposizione 7.13 esiste $g \in \mathbb{C}[x]$ con $A^{-1} = g(A)$. Perciò $A \cdot g(A) = \delta$ e se poniamo $f = x \cdot g$, allora $f(A) = \delta$ ed $f(0) = 0$.

Esempio 7.16. Nell'esempio 7.14 possiamo porre

$$f = -\frac{1}{5}x(x^2 + 2) = -\frac{1}{5}(x^3 + 2).$$

Infatti allora $f(0) = 0$ ed $f(A) = -\dfrac{1}{5}(A^3 + 2A) = -\dfrac{1}{5}(-5\delta) = \delta$.

Osservazione 7.17. Siano $f \in \mathbb{C}[x]$ e $T \in GL(n, \mathbb{C})$.

Allora $f(T^{-1}AT) = f(A)$.

Dimostrazione. È sufficiente osservare che $(T^{-1}AT)^k = T^{-1}A^kT$ per ogni $k \in \mathbb{N}$.

Corollario 7.18. *Sia, anche, $B \in \mathbb{C}^n_n$. Se le matrici A e B sono simili, allora $\mathcal{M}_A = \mathcal{M}_B$.*

Lemma 7.19. *Siano $\lambda \in \mathbb{C}$ e $v \in \mathbb{C}^n$ tali che $Av = \lambda v$. Per ogni $f \in \mathbb{C}[x]$ allora $f(A)v = f(\lambda)v$.*

Dimostrazione. Infatti l'ipotesi implica $A^k v = \lambda^k v$ per ogni $k \in \mathbb{N}$. Per linearità si ottiene l'enunciato.

Teorema 7.20. *Per $\lambda \in \mathbb{C}$ sono equivalenti:*

(1) λ è autovalore di A.
(2) λ è radice del polinomio caratteristico di A : $\mathcal{P}_A(\lambda) = 0$.
(3) λ è radice del polinomio minimale di A : $\mathcal{M}_A(\lambda) = 0$.

Dimostrazione.

(1) $\iff$ (2): Teorema 7.7.

(1) $\Rightarrow$ (3) : λ sia un autovalore di A. Allora esiste $v \in \mathbb{C}^n \setminus 0$ tale che $Av = \lambda v$. Dal lemma 7.9 segue che $\mathcal{M}_A(\lambda)v = \mathcal{M}_A(A)v = 0$ e quindi $\mathcal{M}_A(\lambda) = 0$ perché $v \neq 0$.

(3) $\Rightarrow$ (1) : Sia $\mathcal{M}_A(\lambda) = 0$. Allora esiste $f \in \mathbb{C}[x]$ tale che $\mathcal{M}_A = (x - \lambda)f$. Siccome $f \neq 0$ e $\mathrm{grado}\, f < \mathrm{grado}\, \mathcal{M}_A$, la minimalità di $\mathcal{M}_A$ implica $f(A) \neq 0$. Perciò esiste un vettore $w \in \mathbb{C}^n$ tale che $v := f(A)w \neq 0$.

Usando il lemma 7.2 abbiamo adesso

$$Av - \lambda v = Af(A)w - \lambda f(A)w = (A - \lambda\delta)f(A)w = \mathcal{M}_A(A)w = 0.$$

Siccome $v \neq 0$, ciò mostra che λ è un autovalore di A.

Corollario 7.21. *Sia $f \in \mathbb{C}[x]$ con $f(A) = 0$. Allora $f(\lambda) = 0$ per ogni autovalore λ di A .*

Teorema 7.22 (teorema di Cayley-Hamilton). $\mathcal{P}_A(A) = 0$.

Dimostrazione. Corsi di geometria oppure Mondini, pagg. 68-70.

Corollario 7.23. *Il polinomio minimale $\mathcal{M}_A$ divide il polinomio caratteristico $\mathcal{P}_A$.*

Nota 7.24. Dal teorema 7.20 e dal corollario 7.23 vediamo che, se il polinomio caratteristico di A possiede la forma

$$\mathcal{P}_A = (x - \lambda_1)^{n_1} \cdots (x - \lambda_s)^{n_s}$$

con i λ_k distinti e gli esponenti $n_k \geq 1$, allora il polinomio minimale $\mathcal{M}_A$ è della forma

$$\mathcal{M}_A = (x - \lambda_1)^{m_1} \cdots (x - \lambda_s)^{m_s}$$

con $1 \leq m_k \leq n_k$ per ogni k.

Se n non è troppo grande e se gli autovalori λ_k sono noti con le loro molteplicità n_k (come accade nel caso di una matrice triangolare), possiamo così trovare il polinomio minimale tra i polinomi della forma $(x - \lambda_1)^{r_1} \cdots (x - \lambda_s)^{r_s}$ con $1 \leq r_k \leq n_k$ per ogni k, partendo con $r_1 = \ldots = r_s = 1$ e aumentando gli esponenti, fino a quando troviamo un polinomio che annulla A.

Assumiamo ad esempio che $\mathcal{P}_A = (x-\lambda)^3(x-\mu)^2(x-\nu)$ con λ, μ, ν distinti.

Allora proviamo, in questo ordine, i sei polinomi

$$\begin{aligned}
f_1 &= (x-\lambda)(x-\mu)(x-\nu)\\
f_2 &= (x-\lambda)^2(x-\mu)(x-\nu)\\
f_3 &= (x-\lambda)(x-\mu)^2(x-\nu)\\
f_4 &= (x-\lambda)^3(x-\mu)(x-\nu)\\
f_5 &= (x-\lambda)^2(x-\mu)^2(x-\nu)\\
f_6 &= (x-\lambda)^3(x-\mu)^2(x-\nu)
\end{aligned}$$

fino a quando $f_j(A) = 0$. Possiamo eseguire l'algoritmo formando in successione

$$\begin{aligned}
B_1 &= (A-\lambda\delta)(A-\mu\delta)(A-\nu\delta)\\
B_2 &= B_1(A-\lambda\delta)\\
B_3 &= B_1(A-\mu\delta)\\
B_4 &= B_2(A-\lambda\delta)\\
B_5 &= B_2(A-\mu\delta)\\
B_6 &= B_5(A-\lambda\delta)
\end{aligned}$$

fino a quando $B_j = 0$.