6.b Riduzione a matrice triangolare

6.10 Teorema. *Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione finita e $A : V \to V$ un'applicazione lineare. Sono fatti equivalenti:*

(i) *esiste una base nella quale A si rappresenta con una matrice triangolare superiore,*
(ii) *il polinomio caratteristico di A si decompone in n fattori di primo grado; esistono cioè $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{K}$ tali che*

$$p_A(s) = (s - \lambda_1) \cdots (s - \lambda_n),$$

Dimostrazione. (i) $\Rightarrow$ (ii). Osserviamo che ogni matrice $\mathbf{T}$ triangolare superiore ha come autovalori gli elementi della diagonale $\lambda_i := \mathbf{T}^i_i$, $i = 1, \dots, n$ (lo si verifica facilmente, ad esempio utilizzando le formule di Laplace per il calcolo del determinante). In particolare $\mathbf{T}$ ha n autovalori $\lambda_1, \lambda_2, \dots, \lambda_n$ in $\mathbb{K}$ e il suo polinomio caratteristico è il prodotto di fattori di primo grado

$$p_{\mathbf{T}}(s) = (s - \lambda_1) \cdots (s - \lambda_n).$$

Ora (i) implica che $p_A(s) = p_{\mathbf{T}}(s)$ per una matrice $\mathbf{T}$ triangolare superiore. Segue quindi (ii).

(ii) $\Rightarrow$ (i) Procediamo per induzione sulla dimensione di V. Se $\dim V = 1$ la tesi è ovvia. Supponiamo per induzione di avere dimostrato (ii) $\Rightarrow$ (i) per tutti gli operatori lineari su spazi di dimensione $(n - 1)$ e proviamolo per $A : V \to V$ quando $\dim V = n$. Poiché per ipotesi

$$p_A(s) = (s - \lambda_1) \dots (s - \lambda_n),$$

λ_1 è un autovalore per A. Esiste quindi un autovettore non nullo u_1 di A. Sia $(v_2, \dots, v_n)$ una base di un supplementare V_1 di $\operatorname{Span}\{u_1\}$. Allora $(u_1, v_2, \dots, v_n)$ è una base di V e la matrice associata a A in questa base è

$$\mathbf{A} = \begin{pmatrix} a^1_1 & \boxed{a^1_2\ a^1_3\ \dots a^1_n} \\ 0 & \boxed{\mathbf{B}} \end{pmatrix}, \qquad a^1_1 = \lambda_1.$$

Pertanto $p_{\mathbf{A}}(s) = (s - \lambda_1)\, p_{\mathbf{B}}(s)$ e quindi $p_{\mathbf{B}}(s) = (s - \lambda_2) \dots (s - \lambda_n)$ è prodotto di $n - 1$ fattori di primo grado. Sia $B : V_1 \to V_1$ l'applicazione lineare rappresentata nella base $(v_2, \dots, v_n)$ da $\mathbf{B}$. Per l'ipotesi induttiva, esiste una base $(u_2, \dots, u_n)$ di V_1 nella quale B si rappresenta con una matrice $\mathbf{T}$ triangolare superiore. Pertanto A si rappresenta nella base $(u_1, u_2, \dots, u_n)$ con una matrice della forma

$$\mathbf{A}' = \begin{pmatrix} a^1_1 & \boxed{c^1_2\ c^1_3\ \dots c^1_n} \\ 0 & \boxed{\mathbf{T}} \end{pmatrix}, \qquad a^1_1 = \lambda_1,$$

che, come $\mathbf{T}$, è triangolare superiore. □

Se V è euclideo o hermitiano, il Teorema 6.10 si precisa meglio.

6.11 Teorema. *Sia $A : V \to V$ un operatore lineare su uno spazio euclideo o hermitiano di dimensione n. Sono fatti equivalenti:*

(i) *esiste una base* ortonormale *nella quale A si rappresenta con una matrice triangolare superiore,*
(ii) *il polinomio caratteristico di A si decompone in n fattori di primo grado: esistono cioè $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{K}$ tali che*

$$p_A(s) = (s - \lambda_1) \cdots (s - \lambda_n),$$

Dimostrazione. La dimostrazione è la stessa del Teorema 6.10. Basta osservare che nel passo (ii) $\Rightarrow$ (i) della dimostrazione del Teorema 6.10 è possibile scegliere u_1 con $|u_1| = 1$ e come suo supplementare $V_1 = \operatorname{Span}\{u_1\}^\perp$. □

Il caso degli operatori $A : V \to V$ su uno spazio vettoriale V *complesso* è più semplice perché per il *teorema fondamentale dell'algebra*, vedi Esercizio 15.8 e Teorema 16.2, l'equazione caratteristica di una matrice $n \times n$ si fattorizza in fattori di primo grado. Pertanto

6.12 Corollario. *Sia $A : V \to V$ un'applicazione lineare su uno spazio vettoriale V complesso di dimensione finita. Allora esiste una base di V nella quale A si rappresenta con una matrice triangolare superiore.*

Inoltre,

6.13 Proposizione (Decomposizione di Schur). *Sia $\mathbf{A} \in M_{n,n}(\mathbb{C})$. Allora esiste $\mathbf{S} \in M_{n,n}(\mathbb{C})$ tale che $\overline{\mathbf{S}}^T\mathbf{S} = \mathrm{Id}$ e $\mathbf{T} := \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ è triangolare superiore.*

Dimostrazione. Per il Teorema 6.11, esiste una base *ortonormale* $(u_1, u_2, \dots, u_n)$ di $\mathbb{C}^n$ nella quale A è rappresentata da una matrice triangolare superiore $\mathbf{T}$. Posto $\mathbf{S} := [u_1|\cdots|u_n]$ si ha allora

$$\mathbf{T} := \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$$

Resta da provare che $\mathbf{S}$ è unitaria. Indichiamo con $(e_1, e_2, \dots, e_n)$ la base canonica di $\mathbb{C}^n$. Essendo la base canonica ortonormale. per ogni $i, j = 1, \dots, n$ si ha $(\overline{\mathbf{S}}^T\mathbf{S})^i_j = (\overline{\mathbf{S}}^T\mathbf{S}e_j) \bullet e_i = \overline{e_i}^T\overline{\mathbf{S}}^T\mathbf{S}e_j = (\mathbf{S}e_j) \bullet (\mathbf{S}e_i) = u_j \bullet u_i = \delta_{ij}$. □

6.c Teorema di Cayley–Hamilton

6.14 Teorema (Cayley-Hamilton). *Sia $\mathbf{A} \in M_{n,n}\mathbb{C})$ e sia $p_{\mathbf{A}}(\lambda)$ il suo polinomio caratteristico. Allora $p_{\mathbf{A}}(\mathbf{A}) = 0$.*

6.15 Lemma. *Siano $\mathbf{A}, \mathbf{B} \in M_{n,n}(\mathbb{C})$ matrici triangolari superiori e sia $1 \le j \le n$. Se $\mathbf{A}$ ha le prime $j-1$ colonne nulle e $\mathbf{B}^j_j = 0$, allora $\mathbf{AB}$ ha le prime j colonne nulle.*

Dimostrazione. Essendo $\mathbf{A}$ e $\mathbf{B}$ triangolari superiori, si ha $\mathbf{A}^h_k = \mathbf{B}^h_k = 0$ per ogni $h > k$. Pertanto

$$(\mathbf{AB})^i_j = \sum_{k=1}^n \mathbf{A}^i_k\mathbf{B}^k_j = \sum_{i \le k \le j} \mathbf{A}^i_k\mathbf{B}^k_j = \sum_{i \le k < j} \mathbf{A}^i_k\mathbf{B}^k_l + \mathbf{A}^i_j\mathbf{B}^j_j = 0 + 0 = 0.$$

□

Dimostrazione del Teorema 6.14. (i) Consideriamo prima il caso di una matrice triangolare superiore $\mathbf{T}$. Siano $\lambda_1 = \mathbf{T}^1_1, \dots, \lambda_n = \mathbf{T}^n_n$ gli autovalori di $\mathbf{T}$ e sia $p_{\mathbf{T}}(\lambda) = \prod_{j=1}^n (\lambda - \lambda_j)$ il polinomio caratteristico di $\mathbf{T}$. Posto $\mathbf{B}_j := \mathbf{T} - \lambda_j \,\mathrm{Id}$, si ha $(\mathbf{B}_j)^j_j = 0$ e

$$p_{\mathbf{T}}(\mathbf{T}) = \prod_{j=1}^n (\mathbf{T} - \lambda_j \,\mathrm{Id}) = \mathbf{B}_1\mathbf{B}_2 \cdots \mathbf{B}_n.$$

Ora $\mathbf{B}_1$ ha la prima colonna nulla. Applicando induttivamente il Lemma 6.15, si ottiene che $\mathbf{B}_1\mathbf{B}_2$ ha le prime due colonne nulle, $\mathbf{B}_1\mathbf{B}_2\mathbf{B}_3$ ha le prime tre colonne nulle, ecc., e infine che $p_{\mathbf{T}}(\mathbf{T}) = \mathbf{B}_1\mathbf{B}_2 \cdots \mathbf{B}_n$ ha tutte le colonne nulle, cioè $p_{\mathbf{T}}(\mathbf{T}) = 0$.

(ii) Sia $\mathbf{A} \in M_{n,n}(\mathbb{C})$. Segue dalla Proposizione 6.13 che $\mathbf{A}$ è simile ad una matrice triangolare superiore $\mathbf{T}$: esiste cioè $\mathbf{S} \in M_{n,n}(\mathbb{C})$ tale che $\mathbf{A} = \mathbf{S}^{-1}\mathbf{T}\mathbf{S}$. Pertanto $p_{\mathbf{A}}(\lambda) = p_{\mathbf{T}}(\lambda)$ e quindi, tenendo conto dell'Esercizio 5.8 e di (i)

$$p_{\mathbf{A}}(\mathbf{A}) = \mathbf{S}^{-1} p_{\mathbf{A}}(\mathbf{T})\mathbf{S} = \mathbf{S}^{-1} p_{\mathbf{T}}(\mathbf{T})\mathbf{S} = \mathbf{S}^{-1} 0 \,\mathbf{S} = 0.$$

$\square$

6.16 Proposizione. *Le matrici in $M_{n,n}(\mathbb{C})$ con autovalori distinti formano un sottoinsieme denso di $M_{n,n}(\mathbb{C})$.*

Dimostrazione. Per ogni $\mathbf{A} = (\mathbf{A}^i_j) \in M_{n,n}(\mathbb{C})$, indichiamo con $||\mathbf{A}||_2$ la norma di $\mathbf{A}$ definita da

$$||\mathbf{A}||_2^2 := \mathrm{tr}\,(\overline{\mathbf{A}}^T \mathbf{A}) = \sum_{i,j=1}^n |\mathbf{A}^i_j|^2.$$

Sia $\mathbf{A} \in M_{n,n}(\mathbb{C})$ e sia $\epsilon > 0$. Dalla Proposizione 6.13 $\mathbf{A} = \mathbf{S}^{-1}\mathbf{T}\mathbf{S}$ con $\mathbf{T}$ triangolare superiore. Siano $\lambda_1, \lambda_2, \dots, \lambda_n$ gli elementi della diagonale di $\mathbf{T}$ e siano $\mu_1, \mu_2, \dots, \mu_n$ n numeri distinti fra loro tali che $|\mu_j - \lambda_j| < \epsilon$. La matrice $\mathbf{T}_\epsilon = (\mathbf{T}_\epsilon)^i_j$ definita da

$$(\mathbf{T}_\epsilon)^i_j = \begin{cases} \mathbf{T}^i_j & \text{se } i \neq j, \\ \mu_j & \text{se } i = j \end{cases}$$

è triangolare superiore, ha autovalori $\mu_1, \mu_2, \dots, \mu_n$ tutti distinti e $||\mathbf{T} - \mathbf{T}_\epsilon||_2 \le \sqrt{n}\,\epsilon$. Pertanto la matrice $\mathbf{A}_\epsilon := \mathbf{S}^{-1}\mathbf{T}_\epsilon\mathbf{S}$ ha autovalori $\mu_1, \mu_2, \dots, \mu_n$ tutti distinti e

$$||\mathbf{A} - \mathbf{A}_\epsilon||_2 = ||\mathbf{S}^{-1}(\mathbf{T} - \mathbf{T}_\epsilon)\mathbf{S}||_2 \le ||\mathbf{S}^{-1}||_2\, ||\mathbf{T} - \mathbf{T}_\epsilon||_2\, ||\mathbf{S}||_2 \le C\epsilon.$$

$\square$