Diagonalizzazione

Maurizio Citterio

Marco Boella Alan Cigoli

Politecnico di Milano
beep.metid.polimi.it

2019/2020, I semestre

AVVERTENZA:
quanto segue è soltanto l'elenco di definizioni, proposizioni, osservazioni, esempi ed esercizi che sono stati discussi in aula. Questo non sostituisce in alcun modo la lezione, le esercitazioni, gli appunti, le dispense, il testo.

- Un **autovettore** dell'endomorfismo (lineare) $F : V \to V$ è un vettore non nullo $\mathbf{v}$ di V tale che

$$F(\mathbf{v}) = \lambda \mathbf{v} \ ,$$

dove λ è un numero reale (positivo negativo o nullo), detto **autovalore** dell'autovettore $\mathbf{v}$.

- Si chiama **autospazio** relativo all'autovalore λ il sottospazio vettoriale di V

$$V_\lambda = \ker(F - \lambda \mathrm{Id}_V) = \{\mathbf{v} \in V \mid F\mathbf{v} = \lambda \mathbf{v}\} \ .$$

- La dimensione di V_λ si chiama **molteplicità geometrica** dell'autovalore λ e si denota $\mathrm{mg}\,(\lambda)$:

$$\mathrm{mg}\,(\lambda) = \dim V_\lambda \ .$$

Se $\dim V = n$, dal teorema di nullità più rango, risulta

$$\mathrm{mg}\,(\lambda) = n - \mathrm{rk}(F - \lambda \mathrm{Id}_V) \ .$$

Le definizioni date per l'endomorfismo F si adattano alla matrice reale A quadrata di ordine n considerando l'endomorfismo $L_A : \mathbb{R}^n \to \mathbb{R}^n$. In dettaglio:

- Un **autovettore** della matrice reale A quadrata di ordine n è un vettore non nullo $\mathbf{v}$ di $\mathbb{R}^n$ tale che

$$A\mathbf{v} = \lambda\mathbf{v} \ ,$$

dove λ è un numero reale (positivo negativo o nullo), detto **autovalore** dell'autovettore $\mathbf{v}$.

- Si chiama **autospazio** relativo all'autovalore λ il sottospazio vettoriale di V

$$V_\lambda = \mathrm{Sol}\,\{(A - \lambda \mathrm{I})\mathbf{v} = \mathbf{0}\} = \{\mathbf{v} \in \mathbb{R}^n \mid A\mathbf{v} = \lambda\mathbf{v}\} \ .$$

- La dimensione di V_λ si chiama **molteplicità geometrica** dell'autovalore λ e si denota $\mathrm{mg}\,(\lambda)$; risulta

$$\mathrm{mg}\,(\lambda) = \dim V_\lambda = n - \mathrm{rk}(A - \lambda \mathrm{I}) \ .$$

Basi di autovettori e matrici diagonali

Supponiamo che esista una base $\mathcal{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ di V costituita da autovettori dell'endomorfismo $F : V \to V$, con autovalori (non necessariamente distinti) $\lambda_1, \dots, \lambda_n$.

Allora la matrice $M^{\mathcal{B}}_{\mathcal{B}}(F)$ che rappresenta F rispetto la base $\mathcal{B}$ è la matrice diagonale $\operatorname{diag}(\lambda_1, \dots, \lambda_n)$:

$$M^{\mathcal{B}}_{\mathcal{B}}(F) = \operatorname{diag}(\lambda_1, \dots, \lambda_n) = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

Infatti:

$$\begin{aligned} F(\mathbf{v}_1) &= \lambda_1 \mathbf{v}_1 = \lambda_1 \mathbf{v}_1 + 0 \cdot \mathbf{v}_2 + \cdots + 0 \cdot \mathbf{v}_n \\ F(\mathbf{v}_2) &= \lambda_2 \mathbf{v}_2 = 0 \cdot \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + 0 \cdot \mathbf{v}_n \\ &\vdots \\ F(\mathbf{v}_n) &= \lambda_n \mathbf{v}_n = 0 \cdot \mathbf{v}_1 + 0 \cdot \mathbf{v}_2 + \cdots + \lambda_n \mathbf{v}_n \end{aligned}$$

Basi di autovettori e matrici diagonali

Viceversa, se rispetto ad una base $\mathcal{B} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ di V l'endomorfismo $F : V \to V$ ha una matrice rappresentativa diagonale $M^{\mathcal{B}}_{\mathcal{B}}(F) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, allora $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sono autovettori di autovalori $\lambda_1, \ldots, \lambda_n$.

Infatti le coordinate di $\mathbf{v}_1$ rispetto la base $\mathcal{B}$ sono

$$[\mathbf{v}_1]_{\mathcal{B}} = \mathbf{e}_1 \ ,$$

mentre le coordinate di $F(\mathbf{v}_1)$ sono

$$[F(\mathbf{v}_1)]_{\mathcal{B}} = M^{\mathcal{B}}_{\mathcal{B}}(F)[\mathbf{v}_1]_{\mathcal{B}} = M^{\mathcal{B}}_{\mathcal{B}}(F)\mathbf{e}_1 = \lambda_1\mathbf{e}_1 = \lambda_1[\mathbf{v}_1]_{\mathcal{B}} \ ,$$

dunque

$$F(\mathbf{v}_1) = \lambda_1\mathbf{v}_1 \ .$$

Allo stesso modo si mostra che
$F(\mathbf{v}_2) = \lambda_2\mathbf{v}_2 \ , \ \ldots \ , \ F(\mathbf{v}_n) = \lambda_n\mathbf{v}_n \ .$

Questa osservazione dà senso alle seguenti definizioni:

Definizione (Endomorfismo diagonalizzabile)

Un endomorfismo $F : V \to V$ si dice **diagonalizzabile** quando esiste una base di V costituita da autovettori di F .

La diagonalizzabilità dell'endomorfismo L_A , porta alla seguente

Definizione (Matrice diagonalizzabile)

Una matrice reale A quadrata di ordine n si dice **diagonalizzabile** quando esiste una base di $\mathbb{R}^n$ costituita da autovettori di A .

Matrici simili

Ricordiamo che una matrice A si dice simile alla matrice A' quando esiste una matrice quadrata invertibile P per la quale

$$A' = P^{-1}AP \; .$$

Osservazione

Interpretando la matrice P come una matrice di cambio base, abbiamo che A e A' sono simili se e solo se rappresentano lo stesso endomorfismo rispetto a basi (eventualmente) diverse.

Questa osservazione dimostra il seguente

Teorema

Sia A una matrice reale quadrata di ordine n.
A è diagonalizzabile se e solo se A è simile a una matrice diagonale.

Polinomio caratteristico

Definizione (Polinomio caratteristico)

Il **polinomio caratteristico** di una matrice A quadrata di ordine n è il polinomio di grado n nell'indeterminata λ dato da

$$\det(A - \lambda I)$$

Esempio (Polinomio caratteristico di una 2×2)

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \lambda \mathrm{I} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}, \quad A - \lambda \mathrm{I} = \begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix}$$

Il polinomio caratteristico di A è

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix} \\ &= \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{21}a_{12} \\ &= \lambda^2 - (\operatorname{tr} A)\lambda + \det A \end{aligned}$$

Teorema

Matrici simili hanno lo stesso polinomio caratteristico.

Dimostrazione.

Sia $A' = P^{-1}AP$. Allora

$$\begin{aligned}\det(A' - \lambda I) &= \det(P^{-1}AP - P^{-1}\lambda I P)\\ &= \det[P^{-1}(A - \lambda I)P]\\ &= \det P^{-1} \det(A - \lambda I) \det P\\ &= \det(A - \lambda I)\end{aligned}$$

□

Data la matrice quadrata A, un minore principale di ordine k è una sottomatrice quadrata di A, di ordine k, con la diagonale principale sulla diagonale principale di A.

Per esempio, i minori principali della matrice $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ sono:

- ordine 1: $[a_{11}]$, $[a_{22}]$;
- ordine 2: la stessa A.

i minori principali della matrice $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ sono:

- ordine 1: $[a_{11}]$, $[a_{22}]$, $[a_{33}]$;
- ordine 2: $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $\begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix}$, $\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}$;
- ordine 3: la stessa A.

Si dimostra che il polinomio caratteristico ha la seguente forma:

$$\det(A - \lambda I) = (-1)^n \lambda^n + (-1)^{n-1} \alpha_1 \lambda^{n-1} + \cdots + \alpha_{n-2} \lambda^2 - \alpha_{n-1} \lambda + \alpha_n \ ,$$

dove α_k è la somma dei determinanti dei minori principali di ordine k.

Si osservi che:
α_1 è la traccia di A: $\quad \alpha_1 = \mathrm{tr} A = a_{11} + a_{22} + \cdots + a_{nn}$;
α_n è il determinate di A: $\quad \alpha_n = \det A$.

Si osservi inoltre che i numeri α_k sono invarianti per similitudine perché matrici simili hanno lo stesso polinomio caratteristico.

Teorema (Equazione caratteristica)

Il numero λ *è autovalore per la matrice quadrata* A *se e solo se è una soluzione dell'equazione caratteristica*

$$\det(A - \lambda \mathrm{I}) = 0 \; .$$

Dimostrazione.

λ è autovalore di A se e solo se esiste $\mathbf{v} \neq \mathbf{0}$ in $\mathbb{R}^n$ per il quale $A\mathbf{v} = \lambda\mathbf{v}$, cioè $A\mathbf{v} = (\lambda \mathrm{I})\mathbf{v}$, ossia $(A - \lambda I)\mathbf{v} = \mathbf{0}$.
Per il Teorema di Cramer, questo sistema lineare *omeogeneo* $(A - \lambda I)\mathbf{v} = \mathbf{0}$ ha una soluzione non nulla se e solo se $\det(A - \lambda \mathrm{I}) = 0$. □

Molteplicità algebrica, molteplicità geometrica

- Se nella scomposizione del polinomio $P(\lambda)$ il fattore $(\lambda - \lambda_1)$ compare esattamente k volte, allora si dice che λ_1 è una radice di ordine k, o che la **molteplicità algebrica** di λ_1 è k:

$$\mathrm{ma}(\lambda_1) = k$$

Esempio

$P(\lambda) = (\lambda - 3)^2(\lambda - 5)$.
Radici: $\lambda_1 = \lambda_2 = 3$; $\lambda_3 = 5$.
$\mathrm{ma}(\lambda_1) = 2$, $\mathrm{ma}(\lambda_3) = 1$.

- Se $\lambda_1 \in \mathbb{R}$ è radice di $\det(A - \lambda \mathrm{I})$, la **molteplicità geometrica** di λ è, per definizione,

$$\mathrm{mg}(\lambda_1) = \dim V_{\lambda_1}$$

Teorema

Per ogni autovalore λ si ha

$$1 \leq \mathrm{mg}(\lambda) \leq \mathrm{ma}(\lambda) .$$

Idea dell'argomentazione: caso $n = 3$.

Supponiamo che λ_1 sia autovalore di un endomorfismo F, con molteplicità geometrica 2:

$$\mathrm{mg}(\lambda_1) = \dim V_{\lambda_1} = 2$$

Allora esistono due autovettori, relativi a λ_1, linearmente indipendenti. Completiamoli a una base. Rispetto a tale base, la matrice rappresentativa di F sarà del tipo:

$$B = \begin{bmatrix} \lambda_1 & 0 & * \\ 0 & \lambda_1 & * \\ 0 & 0 & * \end{bmatrix}$$

Allora il fattore $(\lambda - \lambda_1)$ compare, nella fattorizzazione di $\det(B - \lambda \mathrm{I})$, almeno due volte. Dunque $\mathrm{mg}(\lambda) \leq \mathrm{ma}(\lambda)$. □

Teorema
Sia $F : V \to V$ *un endomorfismo lineare. Siano* $\mathbf{v}_1, \dots, \mathbf{v}_m$ *autovettori di* F *, con rispettivi autovalori* $\lambda_1, \dots, \lambda_m$ *.*
Se gli autovalori $\lambda_1, \dots, \lambda_m$ *sono a due a due distinti, allora gli autovettori* $\mathbf{v}_1, \dots, \mathbf{v}_m$ *sono linearmente indipendenti.*

Dimostrazione
Per induzione su m . Se $m = 1$, un autovettore v_1 è un vettore non nullo, e quindi è linearmente indipendente.
Assumiamo (ipotesi induttiva) che l'enunciato sia vero per $m - 1$ autovettori.
Dimostriamo la tesi per m autovettori.

Supponiamo che

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_{m-1}\mathbf{v}_{m-1} + c_m\mathbf{v}_m = \mathbf{0} \tag{1}$$

dobbiamo dimostrare che $c_1 = \cdots = c_m = 0$.
Applicando F , ricordando che $F(\mathbf{v}_j) = \lambda_j\mathbf{v}_j$, otteniamo:

$$c_1\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2 + \cdots + c_{m-1}\lambda_{m-1}\mathbf{v}_{m-1} + c_m\lambda_m\mathbf{v}_m = \mathbf{0}$$

Moltiplicando la (1) per λ_m otteniamo

$$c_1\lambda_m\mathbf{v}_1 + c_2\lambda_m\mathbf{v}_2 + \cdots + c_{m-1}\lambda_m\mathbf{v}_{m-1} + c_m\lambda_m\mathbf{v}_m = \mathbf{0}$$

Sottraendo le ultime due espressioni abbiamo

$$c_1(\lambda_1 - \lambda_m)\mathbf{v}_1 + \cdots + c_{m-1}(\lambda_{m-1} - \lambda_m)\mathbf{v}_{m-1} = \mathbf{0}\ .$$

L'ipotesi induttiva implica

$$c_1(\lambda_1 - \lambda_m) = \cdots = c_{m-1}(\lambda_{m-1} - \lambda_m) = 0$$

Poiché $\lambda_j - \lambda_m \neq 0$ per $j \neq m$, concludiamo che $c_1 = \cdots = c_{m-1} = 0$.
Tornando alla (1), vediamo che $c_m\mathbf{v}_m = \mathbf{0}$ e quindi anche $c_m = 0$. $\square$

Autovettori indipendenti

Teorema

Sia $F : V \to V$ *un endomorfismo lineare.*
Siano $\lambda_1, \ldots, \lambda_s$ *gli autovalori distinti di* F*.*
Il numero N *di autovettori linearmente indipendenti di* F *è la somma delle molteplicità geometriche degli autovalori:*

$$N = \mathrm{mg}(\lambda_1) + \cdots + \mathrm{mg}(\lambda_s) \ .$$

In breve si enuncia il teorema dicendo che “ad autovalori distinti corrispondono autovettori linearmente indipendenti”.

Dimostrazione (facoltativa)

Si veda il libro di Enrico Schlesinger, Algebra lineare e geometria, Lemma 4.8 del Capitolo 7.

Definizione

Gli autovalori con molteplicità geometrica uguale alla molteplicità algebrica sono detti **regolari**.

Teorema (Matrici diagonalizzabili su $\mathbb{R}$)

Una matrice quadrata A è diagonalizzabile su $\mathbb{R}$
se e solo se
gli autovalori di A sono tutti reali e regolari.

Osservazione

Nell'enunciato del teorema non si chiede che gli autovalori siano distinti.

Criterio di diagonalizzabilità

Dimostrazione.

Siano $\lambda_1, \ldots, \lambda_k$ gli autovalori (tutti reali);
$a_1, \ldots, a_k$ le loro molteplicità algebriche,
$g_1, \ldots, g_k$ le loro molteplicità geometriche.

Gli autovettori si trovano negli autospazi; quindi, se vogliamo una base di $\mathbb{R}^n$ che sia formata da autovettori di A, dobbiamo prendere
g_1 vettori indipendenti (il massimo numero possibile) in V_{λ_1},
g_2 vettori indipendenti in V_{λ_2},
$\ldots$,
g_k vettori indipendenti in V_{λ_k}.

Perché questi $g_1 + \cdots + g_k$ autovettori (sicuramente indipendenti) siano una base di $\mathbb{R}^n$, occorre che siano n.

Poiché $g_i \leq a_i$, abbiamo $g_1 + \cdots + g_k \leq a_1 + \cdots + a_k = n$.

La condizione $g_1 + \cdots + g_k = n$ equivale quindi a $g_i = a_i$, per ogni $i = 1, \ldots, n$. □

Corollario
Se una matrice quadrata reale A di ordine n ha n autovalori reali distinti, allora è diagonalizzabile.

Dimostrazione.
Gli autovalori sono tutti reali per ipotesi e semplici (cioè con molteplicità algebrica uguale a 1), quindi regolari perché

$$1 \leq \mathrm{mg}(\lambda) \leq \mathrm{ma}(\lambda) .$$

$\square$

Attenzione:
Una matrice può essere diagonalizzabile, anche se gli autovalori reali non sono tutti distinti tra loro. Ad esempio, la matrice $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ è diagonalizzabile (è già diagonale), anche se ha due autovalori coincidenti.

Ogni matrice simmetrica è diagonalizzabile

Anticipiamo un risultato molto importante:

Teorema (Teorema spettrale)

Se A è una matrice reale simmetrica di ordine n, allora

1. *Tutti gli autovalori di A sono reali.*
2. *Autovettori relativi ad autovalori distinti sono ortogonali.*
3. *Esiste una base ortonormale di $\mathbb{R}^n$ che è formata da autovettori di A.*

Esempio

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

Polinomio caratteristico: $\det(A - \lambda I) = \lambda^2 - 6\lambda + 8 = (\lambda - 2)(\lambda - 4)$;

Autovalori: $\lambda_1 = 2$, $\lambda_2 = 4$.

La matrice di ordine 2 ha due autovalori reali e distinti, quindi è diagonalizzabile per il corollario al criterio di diagonalizzabilità.

Autospazio $V_{\lambda_1} = \operatorname{Ker}(A - \lambda_1 \mathrm{I}) = \{X \in \mathbb{R}^2 \mid (A - 2\mathrm{I})X = 0\}$, ossia:

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad \text{ossia} \quad \begin{cases} x + y & = & 0 \\ x + y & = & 0 \end{cases}$$

Quindi, V_{λ_1} è la retta di equazione $x + y = 0$.

Dunque, gli autovettori in V_{λ_1} sono tutti i vettori $\begin{bmatrix} t \\ -t \end{bmatrix} = t \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, con $t \neq 0$.

Esempio
Continuazione

Analogamente, $V_{\lambda_2} = \{X \in \mathbb{R}^2 \mid (A - 4\mathrm{I})X = 0\}$ è lo spazio delle soluzioni del sistema lineare

$$\begin{cases} x - y &= 0 \\ x - y &= 0 \end{cases}, \qquad \text{ossia} \quad x - y = 0$$

Gli autovettori relativi a λ_2 sono tutti i vettori $\begin{bmatrix} s \\ s \end{bmatrix} = s \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\mathcal{B} = \big((1, -1), (1, 1)\big)$ è una base di autovettori.

La matrice A è diagonalizzabile ed è simile a $\operatorname{diag}(2, 4) = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$

Una matrice di passaggio P è data allineando in colonna i vettori della base di autovettori.

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

Quindi $P^{-1}AP = D$, cioè $\begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$.

Esempio

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 4 & 1 \\ 0 & 1 & 4 \end{bmatrix}$$

Polinomio caratteristico: $(3-\lambda)(\lambda-3)(\lambda-5)$.
$\lambda_1 = 3\,,\quad \mathrm{ma}(\lambda_1) = 2$ (Radice doppia).
$\lambda_2 = 5\,,\quad \mathrm{ma}(\lambda_2) = 1$ (Radice semplice).
$\mathrm{mg}(\lambda_1) = \dim V_{\lambda_1} = 3 - \mathrm{rk}(A - 3\mathrm{I}) = 3 - 1 = 2$
$\mathrm{mg}(\lambda_2) = 1$
Quindi entrambi gli autovalori sono regolari.
Possiamo trovare due autovettori linearmente indipendenti relativi alla radice doppia $\lambda_1 = 3$.
Troviamo un autovettore relativo a $\lambda_2 = 5$.
Quindi troviamo una base di autovettori di A , rispetto alla quale L_A è rappresentata dalla matrice

$$\mathrm{diag}(3,3,5) = A' = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

Esempio
“Shear matrix”, matrice di taglio

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Polinomio caratteristico: $(1-\lambda)^2 = (1-\lambda)(1-\lambda)$.
Autovalori: $\lambda_1 = \lambda_2 = 1 \ (= \lambda) \qquad\qquad \mathrm{ma}(\lambda) = 2$
Molteplicità geometrica di λ : $\mathrm{mg}(\lambda) = \dim V_\lambda = 2 - \mathrm{rk}(A - \mathrm{I}) = 2 - 1 = 1$
L’autovalore non è regolare.
Non possiamo trovare due autovettori linearmente indipendenti per formare una base.
Quindi A non è diagonalizzabile.

Esempio
Matrice di rotazione di $\pi/2$

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Polinomio caratteristico: $\lambda^2 + 1$.
Radici del polinomio caratteristico: $i, -i$. (Radici complesse, non reali).
La matrice A non è diagonalizzabile (su $\mathbb{R}$) perché il suo polinomio caratteristico ha radici complesse non reali.