Note per il corso di *Geometria e algebra lineare* 2023-24 LT in Informatica

8 Autovalori e autovettori

8.1

Il problema dell'esistenza di una matrice *diagonale* associata a una funzione lineare, utilissima nelle applicazioni, è legato a due concetti importanti: quelli di autovalore e di autovettore.

Definizione 1. Sia $A \in M_n(\mathbb{K})$ una matrice quadrata. Uno scalare $\lambda \in \mathbb{K}$ si dice *autovalore* di A se esiste un vettore non nullo $x \in \mathbb{K}^n$ (detto *autovettore*) tale che $Ax = \lambda x$.

Se $T : V \to V$ è una funzione lineare (endomorfismo) su uno spazio vettoriale V su $\mathbb{K}$, uno scalare $\lambda \in \mathbb{K}$ si dice *autovalore* di T se esiste un vettore non nullo $v \in V$ (detto *autovettore* di T) tale che $T(v) = \lambda v$.

Gli autovalori e gli autovettori si incontrano frequentemente in problemi applicativi dell'ingegneria, della fisica, ecc. Gli autovettori di A possono rappresentare, ad esempio, gli assi di simmetria di un corpo o di una curva, le direzioni di massimo (o di minimo) sforzo, le funzioni corrispondenti ai livelli di energia di un sistema fisico.

Teorema 1. *La funzione lineare $T : V \to V$ ha una matrice associata diagonale se e solo se V ha una base costituita da autovettori di T.*

Dimostrazione. Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ base di V con $M_{\mathcal{B}}(T) = D$ diagonale. Allora $T(v_i)$ ha coordinate $(0, 0, \dots, d_{ii}, \dots 0)$ rispetto a $\mathcal{B}$ (con d_{ii} in posizione i-esima), e dunque $T(v_i) = d_{ii} v_i$ e v_i è un autovettore per ogni i. Viceversa, se $\mathcal{B} = \{v_1, \dots, v_n\}$ è base di autovettori, si ha $T(v_i) = \lambda_i v_i$ per $i = 1, \dots, n$ e quindi

$$M_{\mathcal{B}}(T) = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

□

Definizione 2. Diremo che T è *diagonalizzabile* se ha una matrice associata diagonale. Una matrice $A \in M_n(\mathbb{K})$ sarà detta *diagonalizzabile* se la funzione lineare T_A è diagonalizzabile.

Due matrici quadrate rappresentano lo stesso endomorfismo se e solo se sono simili (cf.§7.5). Quindi una matrice è diagonalizzabile quando è simile alla matrice diagonale contenente gli autovalori λ_i sulla diagonale principale: $D = P^{-1}AP$, con $P = M^{\mathcal{E}}_{\mathcal{B}}(id)$. L'esistenza di una matrice diagonale associata a T semplifica molto i calcoli. Ad esempio, se si vuole calcolare una potenza D^k (corrisponde a k iterazioni di T) basta elevare alla k-esima potenza gli elementi diagonali (= gli autovalori) di D:

$$D^k = \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{bmatrix}.$$

Se $A = M_{\mathcal{B}}(T)$, la potenza A^k è la matrice associata a $T^k = T \circ T \circ \cdots \circ T$ (k volte) rispetto a $\mathcal{B}$ ed è simile a D^k:

$$A^k = (PDP^{-1})^k = (PDP^{-1})(PDP^{-1}) \cdots (PDP^{-1}) = PD^kP^{-1}.$$

Esempi. (1) La matrice $A = \begin{bmatrix} 0 & 2 \\ -1 & 3 \end{bmatrix}$ ha $v_1 = (1,1)$ e $v_2 = (2,1)$ come autovettori. Infatti $Av_1 = (2,2) = 2v_1$ e $Av_2 = (2,1) = v_2$. Dunque 2 e 1 sono autovalori di A e A è diagonalizzabile, poiché $\{v_1, v_2\}$ è una base di $\mathbb{R}^2$ formata da autovettori di T.

(2) Una rotazione nel piano di un angolo θ in senso antiorario attorno all'origine ha autovettori se e solo se $\theta = k\pi$, con $k \in \mathbb{Z}$. Infatti solo in questo caso i vettori ruotati non cambiano direzione (al più il verso) e quindi sono autovettori. La rotazione è diagonalizzabile se e solo se $\theta = k\pi$.

(3) Sia T la derivazione sullo spazio dei polinomi $\mathbb{R}_n[x]$. Esiste un polinomio $p \neq 0$ tale che $T(p) = p' = \lambda p$ se e solo se $\lambda = 0$ e $p = a_0$ è costante. Quindi 0 è l'unico autovalore e gli elementi non nulli del nucleo $N(T)$ sono gli autovettori di T. In particolare, T non è diagonalizzabile se $n > 0$.

(4) Nello spazio $C^\infty(\mathbb{R}, \mathbb{R})$ delle funzioni reali di una variabile reale che possiedono tutte le derivate, la *derivazione* ha anche autovalori diversi da zero: ad esempio, $(e^{\lambda x})' = \lambda e^{\lambda x}$ per ogni $\lambda \in \mathbb{R}$ fissato.

Definizione 3. Sia λ un autovalore di T. Il sottospazio

$$E(\lambda) = N(T - \lambda\, id_V) = \{v \in V \mid T(v) = \lambda v\}$$

contenente il vettore nullo e gli autovettori di T con autovalore λ, è detto *autospazio* di T relativo all'autovalore λ. La sua dimensione (sempre maggiore o uguale a 1) è detta *molteplicità geometrica* di λ, denotata con $m_{geo}(\lambda)$.

Similmente, data una matrice quadrata A, con λ autovalore di A, il sottospazio

$$E(\lambda) = N(A - \lambda I_n) = \{x \in \mathbb{K}^n \mid Ax = \lambda x\}$$

contenente il vettore nullo e gli autovettori con autovalore λ, è detto *autospazio* di A relativo all'autovalore λ.

8.2 Polinomio caratteristico

Sia T un endomorfismo di uno spazio n-dimensionale V e sia $A = M_{\mathcal{B}}(T)$. Lo scalare λ è autovalore di T se e solo se λ è autovalore di A: se $x = T_{\mathcal{B}}(v)$,

$$T(v) = \lambda v \Leftrightarrow Ax = \lambda x.$$

Dunque λ è autovalore $\Leftrightarrow$ il sistema lineare omogeneo $(A - \lambda I_n)x = 0$ ha una soluzione non nulla $\Leftrightarrow \det(A - \lambda I_n) = 0$.

Se $A' = PAP^{-1}$ è un'altra matrice associata a T, la matrice $A' - \lambda I_n$ è simile alla matrice $A - \lambda I_n$:

$$P(A - \lambda I_n)P^{-1} = PAP^{-1} - \lambda I_n = A' - \lambda I_n$$

e dunque $\det(A' - \lambda I_n) = \det(A - \lambda I_n)$.

Definizione 4. Il polinomio di grado n nella variabile λ

$$p_A(\lambda) = \det(A - \lambda I_n)$$

si dice *polinomio caratteristico* della matrice A e anche *polinomio caratteristico* $p_T(\lambda)$ *dell'endomorfismo* T.

Teorema 2. *Uno scalare* λ_0 *è un autovalore di* A *(di* T *) se e solo se* λ_0 *è* radice *del polinomio caratteristico. L'autospazio* $E(\lambda_0)$ *ha dimensione*

$$m_{geo}(\lambda) = \dim E(\lambda_0) = n - rg(A - \lambda_0 I_n).$$

Dimostrazione. La prima parte è immediata dalla definizione di p_A e di p_T. Inoltre $\dim E(\lambda_0) = \dim N(T - \lambda_0 id_V) = \dim N(A - \lambda_0 I_n)$, poiché $A - \lambda_0 I_n$ è la matrice associata a $T - \lambda_0 id_V$ rispetto a $\mathcal{B}$. □

La molteplicità di λ_0 come radice di p_A o di p_T è detta *molteplicità algebrica* di λ_0, denotata con $m_{alg}(\lambda_0)$. Si può vedere che è sempre maggiore o uguale alla molteplicità geometrica: $m_{alg}(\lambda_0) \geq m_{geo}(\lambda_0)$.

Esempi. (1) La matrice reale $A = \begin{bmatrix} 0 & -3 & -1 \\ 6 & 11 & 3 \\ 10 & 15 & 7 \end{bmatrix}$ ha polinomio caratteristico

$$p_A(\lambda) = \det(A - \lambda I_3) = \det \begin{bmatrix} -\lambda & -3 & -1 \\ 6 & 11-\lambda & 3 \\ 10 & 15 & 7-\lambda \end{bmatrix} = \det \begin{bmatrix} -\lambda & 0 & -1 \\ 6 & 2-\lambda & 3 \\ 10 & 3\lambda - 6 & 7-\lambda \end{bmatrix}$$

$$= (2-\lambda) \det \begin{bmatrix} -\lambda & 0 & -1 \\ 6 & 1 & 3 \\ 10 & -3 & 7-\lambda \end{bmatrix} = (2-\lambda)(-\lambda(16-\lambda) + 28) = (2-\lambda)(\lambda^2 - 16\lambda + 28)$$

$$= -(\lambda - 2)^2(\lambda - 14).$$

Dunque A ha autovalori 2 e 14, con autovettori rispettivamente

$$E(2) = N(A - 2I_3) = N \begin{bmatrix} -2 & -3 & -1 \\ 6 & 9 & 3 \\ 10 & 15 & 5 \end{bmatrix} = N \begin{bmatrix} -2 & -3 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \langle (1,0,-2), (0,1,-3) \rangle$$

e $$E(14) = N(A - 14I_3) = N \begin{bmatrix} -14 & -3 & -1 \\ 6 & -3 & 3 \\ 10 & 15 & -7 \end{bmatrix} = N \begin{bmatrix} 1 & 0 & 1/5 \\ 0 & 1 & -3/5 \\ 0 & 0 & 0 \end{bmatrix} = \langle (1,-3,-5) \rangle .$$

Le molteplicità algebriche e geometriche coincidono per i due autovalori. La matrice è diagonalizzabile, poiché l'insieme di autovettori

$$\mathcal{B} = \{(1,0,-2), (0,1,-3), (1,-3,-5)\}$$

forma una base di $\mathbb{R}^3$, rispetto alla quale l'endomorfismo definito da A ha matrice associata diagonale

$$D = M_{\mathcal{B}}(T_A) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 14 \end{bmatrix}.$$

Mettendo in colonna la base di autovettori, si ottiene la matrice di transizione $P = M^{\mathcal{E}}_{\mathcal{B}}(id) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -3 \\ -2 & -3 & -5 \end{bmatrix}$, da cui si ricava la matrice $P^{-1} = M^{\mathcal{B}}_{\mathcal{E}}(id) = \frac{1}{12} \begin{bmatrix} 14 & 3 & 1 \\ -6 & 3 & -3 \\ -2 & -3 & -1 \end{bmatrix}$ tale che $P^{-1}AP = D$.

(2) Sia $T : \mathbb{R}^3 \to \mathbb{R}^3$ definita da

$$T(x_1, x_2, x_3) = \left(\frac{8}{5}x_1 + \frac{3}{5}x_3, x_1 + 2x_2 + x_3, \frac{2}{5}x_1 + \frac{7}{5}x_3\right).$$

Essendo $M_{\mathcal{E}}(T) = B = \begin{bmatrix} 8/5 & 0 & 3/5 \\ 1 & 2 & 1 \\ 2/5 & 0 & 7/5 \end{bmatrix}$, T ha polinomio caratteristico

$$p_T(\lambda) = p_B(\lambda) = \det(B - \lambda I_3) = \det \begin{bmatrix} 8/5-\lambda & 0 & 3/5 \\ 1 & 2-\lambda & 1 \\ 2/5 & 0 & 7/5-\lambda \end{bmatrix}$$

$$= (2-\lambda)(\lambda^2 - 3\lambda + 2) = -(\lambda-2)^2(\lambda-1) \text{ e autospazi}$$

$$E(1) = N(B - I_3) = N \begin{bmatrix} 3/5 & 0 & 3/5 \\ 1 & 1 & 1 \\ 2/5 & 0 & 2/5 \end{bmatrix} = N \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \langle(-1, 0, 1)\rangle$$

$$E(2) = N(B - 2I_3) = N \begin{bmatrix} -2/5 & 0 & 3/5 \\ 1 & 0 & 1 \\ -3/5 & 0 & 2/5 \end{bmatrix} = N \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \langle(0, 1, 0)\rangle$$

T (e la matrice B) non è diagonalizzabile, poiché si possono scegliere al più due autovettori linearmente indipendenti. In questo caso un autovalore (quello *doppio* $\lambda = 2$) ha molteplicità geometrica minore di quella algebrica.

Attenzione: negli esempi precedenti sono state applicate le operazioni elementari sulle righe delle matrici $A - \lambda_i I_n$ per calcolare gli autovettori di A. Inoltre ogni matrice triangolare ha sulla diagonale esattamente i propri autovalori. Ma *non* è possibile usare lo stesso procedimento anche per calcolare gli autovalori di A, poiché la matrice EA, con E matrice elementare, ottenuta facendo sulle righe di A un'operazione elementare, ha in generale autovalori differenti da quelli di A.

8.3 Criterio di diagonalizzabilità

Proposizione 1. *Sia $T : V \to V$ lineare. Se $v_1, \dots, v_k$ sono autovettori di T corrispondenti ad autovalori distinti $\lambda_1, \dots, \lambda_k$, allora $v_1, \dots, v_k$ sono linearmente indipendenti.*

Dimostrazione. Vediamo il caso $k = 2$. Sia

$$a_1 v_1 + a_2 v_2 = 0.$$

Applicando T si ha anche $a_1 T(v_1) + a_2 T(v_2) = 0$, cioè

$$a_1 \lambda_1 v_1 + a_2 \lambda_2 v_2 = 0.$$

Sottraendo la prima uguaglianza moltiplicata per λ_1 dall'ultima, si ottiene

$$a_2(\lambda_2 - \lambda_1)v_2,$$

da cui $a_2 = 0$ essendo $\lambda_1 \neq \lambda_2$ e $v_2 \neq 0$ poiché autovettore. Tornando alla prima equaglianza, si ha $a_1 v_1 = 0$, da cui anche $a_1 = 0$ poiché $v_1 \neq 0$. □

In particolare, si ha le seguente condizione *sufficiente* di diagonalizzabilità: se T ha esattamente n autovalori distinti, scegliendo n autovettori relativi agli autovalori distinti si ottiene una base, e quindi T è diagonalizzabile.

Diretta conseguenza della proposizione precedente è la seguente:

Proposizione 2. *L'insieme $\mathcal{B}$ di vettori ottenuto unendo basi degli autospazi di T è linearmente indipendente.*

Dimostrazione. Se $\mathcal{B}$ fosse linearmente dipendente, almeno uno dei vettori sarebbe combinazione lineare degli altri. Essendo gli elementi di ogni base indipendenti, nella combinazione dovrebbero comparire anche elementi di altri autospazi, e si otterrebbe allora un autovettore come combinazione lineare di autovettori appartenenti ad altri autospazi. Ma questo non può accadere per la Proposizione precedente. □

Teorema 3. *Siano $\lambda_1, \dots, \lambda_h$ gli autovalori distinti di T ($h \leq n = \dim V$). T è diagonalizzabile se e solo se la somma delle molteplicità geometriche è uguale alla dimensione di V:*

$$\sum_{i=1}^{h} m_{geo}(\lambda_i) = \sum_{i=1}^{h} \dim E(\lambda_i) = \dim V.$$

Dimostrazione. Se T ha una matrice associata diagonale D, il polinomio caratteristico è

$$p_T(\lambda) = \det(D - \lambda I_n) = \prod_{i=1}^{h} (\lambda_i - \lambda)^{m_i}$$

dove m_i è la molteplicità algebrica di λ_i. In particolare, questo mostra che il polinomio caratteristico di un endomorfismo diagonalizzabile ha tutte le sue radici (gli autovalori) nel campo $\mathbb{K}$ degli scalari.

La matrice $D - \lambda_i I_n$ ha rango $n - m_i$ e quindi $\dim E(\lambda_i) = n - (n - m_i) = m_i$. Dunque $\sum_{i=1}^{h} \dim E(\lambda_i) = \text{grado}(p_T) = n$.

Viceversa, se vale la condizione del teorema, l'insieme $\mathcal{B}$ ottenuto unendo basi degli autospazi di T contiene n elementi, che sono linearmente indipendenti per la Proposizione precedente, e quindi forma una base di autovettori di V. □

Osservazione. Nel caso in cui il polinomio caratteristico di T abbia tutte le sue radici nel campo $\mathbb{K}$, la condizione del teorema è equivalente all'uguaglianza delle molteplicità algebriche e geometriche per ogni autovalore di T (e basta controllare l'uguaglianza per gli autovalori multipli di T):

$$m_{alg}(\lambda_i) = m_{geo}(\lambda_i) \quad \text{per ogni } i = 1, \dots, h.$$

Esempio. (1) La matrice di rotazione

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

ha polinomio caratteristico $p_{R_\theta}(\lambda) = (\cos\theta - \lambda)^2 + \sin^2\theta = \lambda^2 - 2(\cos\theta)\lambda + 1$, il cui discriminante è $\Delta = 4(\cos^2\theta - 1)$, che è non negativo solo per $\theta = k\pi$, k intero. Dunque R_θ ha autovalori reali ($\lambda = 1$ e $\lambda = -1$ doppi) solo quando $R_\theta = \pm I_2$. Negli altri casi *non* è diagonalizzabile su $\mathbb{R}$.

Sul campo complesso la matrice ha due autovalori distinti se $R_\theta \neq \pm I_2$ e quindi è diagonalizzabile per *ogni* valore di θ. Gli autovalori sono $\lambda_1 = e^{i\theta} = \cos\theta + i\sin\theta$ e

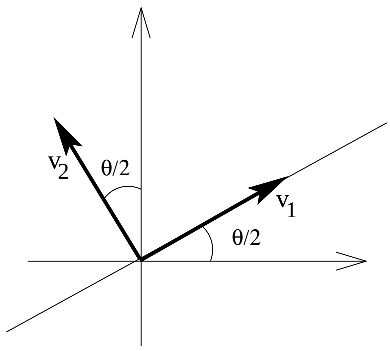
$\lambda_2 = e^{-i\theta} = \cos\theta - i\sin\theta$, con autospazi $E(e^{i\theta}) = \langle(1,-i)\rangle$, $E(e^{-i\theta}) = \langle(1,i)\rangle$. Dunque R_θ è simile a una matrice diagonale:

$$\begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix}^{-1} R_\theta \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} R_\theta \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix} = \begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix}.$$

(2) La matrice

$$M_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

ha polinomio caratteristico $p_{M_\theta}(\lambda) = -(\cos\theta - \lambda)(\cos\theta + \lambda) - \sin^2\theta = \lambda^2 - 1 = (\lambda - 1)(\lambda + 1)$. Dunque M_θ è sempre diagonalizzabile sui reali ed è simile alla matrice diagonale $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Quale trasformazione geometrica del piano rappresenta? (cf. gli esempi in §7.4)



Gli autovettori sono i multipli di

$$v_1 = (-\sin\theta, \cos\theta - 1) \quad \text{e} \quad v_2 = (-\sin\theta, \cos\theta + 1),$$

che sono una coppia di vettori ortogonali di $\mathbb{R}^2$. Si osservi che la matrice M_θ è *simmetrica*, diversamente da R_θ. Vedremo più avanti che ogni matrice reale simmetrica ha proprietà analoghe a quelle di M_θ.

Osservazione. Se 0 è autovalore di una funzione lineare T, allora l'autospazio $E(0)$ coincide con il nucleo di T. In particolare, 0 è autovalore se e solo se T non è iniettiva.