Criteri di diagonalizzabilità

Maurizio Cornalba

18/12/2013

Sia K un campo e sia $P(t)$ un polinomio a coefficienti in K. Se $a \in K$, la *molteplicità* di a come radice di $P(t)$, che indicheremo con m_a, è il massimo intero m tale che $P(t)$ sia divisibile per $(t-a)^m$. Dunque $P(t) = (t-a)^{m_a}Q(t)$, dove $Q(t)$ non è divisibile per $(t-a)$, cioè non ha a come radice. È chiaro che a è radice di $P(t)$ se e solo se $m_a > 0$; si dice che a è una radice *semplice* di $P(t)$ se $m_a = 1$ e che è una radice *multipla* se $m_a > 1$.

Sia V uno spazio vettoriale di dimensione finita su K, e sia $f : V \to V$ un endomorfismo. Sia $P(t)$ il polinomio caratteristico di f. Se $\lambda \in K$ è una radice di $P(t)$, cioè se è un autovalore di f, diremo che m_λ è la molteplicità dell'autovalore stesso. Per ogni $\lambda \in K$ indichiamo con

$$V_\lambda = \{v \in V : f(v) = \lambda v\}$$

l'autospazio di f relativo a λ.

Lemma 1. $m_\lambda \geq \dim(V_\lambda)$

Dimostrazione. Scelta una base $(v_1, \dots, v_k)$ di V_λ, completiamola a una base $(v_1, \dots, v_n)$ di V. Dato che $f(v_i) = \lambda v_i$ per $i = 1, \dots, k$, la matrice di f rispetto a questa base è della forma

$$A = \begin{pmatrix} \lambda I_k & C \\ 0 & B \end{pmatrix}$$

dove I_k è la matrice identità $k \times k$, C è una matrice $k \times (n-k)$ e B è una matrice $(n-k) \times (n-k)$. Allora

$$P(t) = \det(tI - A) = \det((t-\lambda)I_k)\det(tI - B) = (t-\lambda)^k \det(tI - B)$$

Quindi $\dim(V_\lambda) = k \leq m_\lambda$. $\square$

Lemma 2. *Siano $\lambda_1, \dots, \lambda_h$ autovalori distinti di f. Se*

$$\sum_{i=1}^{h} v_i = 0 \qquad (1)$$

dove $v_i \in V_{\lambda_i}$, allora $v_i = 0$ per ogni i.

Dimostrazione. Ragioniamo per induzione su h. Se $h = 1$ non c'è niente da dimostrare. Il passo induttivo è il seguente. Applichiamo f ai due lati di (1) e otteniamo

$$0 = \sum_{i=1}^{h} f(v_i) = \sum_{i=1}^{h} \lambda_i v_i$$

Moltiplichiamo poi per λ_h i due lati della (1) ottenendo

$$0 = \sum_{i=1}^{h} \lambda_h v_i$$

Sottraendo dalla prima la seconda di queste uguaglianze ricaviamo che

$$0 = \sum_{i=1}^{h} (\lambda_i - \lambda_h) v_i = \sum_{i=1}^{h-1} (\lambda_i - \lambda_h) v_i$$

A questo punto l'ipotesi induttiva ci assicura che $(\lambda_i - \lambda_h)v_i = 0$ per $i = 1, \ldots, h-1$. Dato che $\lambda_i \neq \lambda_h$ per $i < h$ se ne deduce che $v_1 = \cdots = v_{h-1} = 0$, e la (1) si riduce a $v_h = 0$. □

Ricordiamo che un endomorfismo $f : V \to V$ come sopra si dice *diagonalizzabile* se V ha una base costituita da autovettori di f; perché ciò si verifichi è sufficiente che V sia generato da autovettori.

Proposizione 1. *Sia V uno spazio vettoriale di dimensione finita sul campo K. Un endomorfismo $f : V \to V$ è diagonalizzabile se e solo se*

$$\dim(V) = \sum_{\lambda \in K} \dim(V_\lambda) \tag{2}$$

dove V_λ è l'autospazio di f relativo a λ.

Dimostrazione. Siano $\lambda_1, \ldots, \lambda_h$ gli autovalori distinti di f. Poniamo $n_i = \dim(V_{\lambda_i})$. La (2) dice che

$$\dim(V) = \sum_{i=1}^{h} n_i$$

Per ogni i scegliamo una base $v_{i,1}, \ldots, v_{i,n_i}$ di V_{λ_i}. Dimostreremo che

$$v_{1,1}, \ldots, v_{1,n_1}, \ldots, v_{i,1}, \ldots, v_{i,n_i}, \ldots, v_{h,1}, \ldots, v_{h,n_h}$$

è una base di V. Dato che abbiamo a che fare con $\sum n_i = \dim(V)$ vettori, basterà mostrare che sono indipendenti. Supponiamo che

$$\sum_{j=1}^{n_1} a_{1,j} v_{1,j} + \cdots + \sum_{j=1}^{n_i} a_{i,j} v_{i,j} + \cdots + \sum_{j=1}^{n_h} a_{h,j} v_{h,j} = 0$$

dove i coefficienti $a_{i,j}$ appartengono a K. Per ogni i, $\sum_j a_{i,j} v_{i,j} \in V_{\lambda_i}$. Quindi il lemma 2 implica che

$$\sum_{j=1}^{n_i} a_{i,j} v_{i,j} = 0$$

per ogni i. Dato che $v_{i,1}, \ldots, v_{i,n_i}$ sono indipendenti si conclude che $a_{i,j} = 0$ per ogni i e ogni j. □

Corollario 1. *Sia V uno spazio vettoriale di dimensione finita su $\mathbb{C}$. Un endomorfismo $f : V \to V$ è diagonalizzabile se e solo se* $\dim(V_\lambda) = m_\lambda$ *per ogni* $\lambda \in \mathbb{C}$.

Dimostrazione. In virtù della proposizione 1 basta osservare che $\dim(V) = \sum m_\lambda$. In effetti, dato che ogni polinomio a coefficienti complessi si fattorizza completamente, il polinomio caratteristico di f è della forma

$$P(t) = \prod_{i=1}^{h} (t - \lambda_i)^{m_{\lambda_i}} = \prod_{\lambda \in \mathbb{C}} (t - \lambda)^{m_\lambda}$$

dove $\lambda_1, \ldots, \lambda_h$ sono gli autovalori di f, e quindi

$$\dim(V) = \deg(P(t)) = \sum_{\lambda \in \mathbb{C}} m_\lambda$$

□

Sia $f : V \to V$ un endomorfismo dello spazio vettoriale V sul campo K, che supponiamo come sempre di dimensione finita. Ricordiamo che il *polinomio minimo* di f (su K) è quello di grado minimo tra i polinomi $Q(t)$ monici a coefficienti in K per cui $Q(f) = 0$. Per il teorema di Cayley-Hamilton il polinomio minimo divide il polinomio caratteristico di f e quindi il suo grado non supera la dimensione di V. Se λ è un autovalore di f, v un autovettore per λ e indichiamo con $Q(t)$ il polinomio minimo di f, allora

$$0 = Q(f)(v) = Q(\lambda) \cdot v$$

e quindi $Q(\lambda) = 0$. In altre parole, ogni autovalore di f è radice del polinomio minimo.

Proposizione 2. *Sia V uno spazio vettoriale di dimensione finita su $\mathbb{C}$ e sia f un suo endomorfismo. Sono condizioni equivalenti:*

(i) *f è diagonalizzabile;*
(ii) *il polinomio minimo di f non ha radici multiple;*
(iii) *esiste un polinomio non nullo senza radici multiple $P(t)$ tale che $P(f) = 0$.*

Dimostrazione. Siano $\lambda_1, \ldots, \lambda_h$ gli autovalori distinti di f. Se m_{λ_i} è la molteplicità dell'autovalore λ_i il polinomio minimo di f è della forma

$$\prod (t - \lambda_i)^{r_i}$$

dove $1 \leq r_i \leq m_{\lambda_i}$, e ha radici semplici se e solo se $r_i = 1$ per ogni i. Poniamo $Q(t) = \prod (t - \lambda_i)$. Se f è diagonalizzabile la sua matrice A rispetto a una opportuna base è diagonale e gli elementi diagonali sono autovalori di f. Quindi se d è uno di questi elementi diagonali $Q(d) = 0$. Ne segue che $Q(A) = 0$, cioè che $Q(f) = 0$, e perciò che $Q(t)$ è il polinomio minimo di f.

Supponiamo viceversa che $Q(t)$ sia il polinomio minimo di f e dimostriamo che f è diagonalizzabile. Useremo il seguente semplice lemma algebrico.

Lemma 3. *Sia K un campo e siano $\lambda_1, \ldots, \lambda_h$ elementi distinti di K. Per ogni i poniamo*

$$Q_i(t) = \prod_{j=1,\ j \neq i}^{h} (t - \lambda_j)$$

Allora

$$\sum_{i=1}^{h} \frac{1}{Q_i(\lambda_i)} Q_i(t) = 1$$

Dimostrazione. Notiamo innanzitutto che

$$Q_i(\lambda_i) = \prod_{j \neq i} (\lambda_i - \lambda_j) \neq 0\,; \quad Q_i(\lambda_h) = \prod_{j \neq i} (\lambda_h - \lambda_j) = 0 \text{ se } h \neq i$$

Poniamo poi

$$Z(t) = \sum_{i=1}^{h} \frac{1}{Q_i(\lambda_i)} Q_i(t) - 1$$

e notiamo che si tratta di un polinomio di grado $h-1$ in t o del polinomio nullo. Inoltre

$$Z(\lambda_i) = \frac{1}{Q_i(\lambda_i)} Q_i(\lambda_i) - 1 = 0$$

per ogni i e quindi $Z(t)$ ha almeno h radici. L'unica possibilità è dunque che $Z(t)$ sia nullo. $\square$

Ora possiamo dimostrare che f è diagonalizzabile. Usiamo le notazioni del lemma precedente. Il lemma implica che

$$\sum_{i=1}^{h} \frac{1}{Q_i(\lambda_i)} Q_i(f)$$

è l'applicazione identità e quindi che per ogni $v \in V$

$$v = \sum_{i=1}^{h} \frac{1}{Q_i(\lambda_i)} Q_i(f)(v) = \sum_{i=1}^{h} v_i$$

dove abbiamo posto

$$v_i = \frac{1}{Q_i(\lambda_i)} Q_i(f)(v)$$

D'altra parte, visto che $(t-\lambda_i)Q_i(t) = Q(t)$,

$$f(v_i) - \lambda_i v_i = \frac{1}{Q_i(\lambda_i)} Q(f)(v) = 0$$

cioè $v_i \in V_{\lambda_i}$. Questo mostra che ogni elemento di V è somma di autovettori, e dunque che f è diagonalizzabile.

La condizione (iii) è chiaramente conseguenza della (ii); basta infatti prendere come $P(t)$ il polinomio minimo di f. Supponiamo invece che sia soddisfatta la (iii) e deduciamone la (ii) ragionando per assurdo. Ricordiamo che il polinomio minimo $Q(t)$ divide $P(t)$. Se $Q(t)$ avesse una radice multipla λ, cioè se fosse divisibile per $(t-\lambda)^2$, lo stesso sarebbe quindi vero per $P(t)$, contro l'ipotesi. $\square$

Esempio 1. Sia n un intero positivo e sia c un numero complesso diverso da zero. Sia A una matrice complessa quadrata e supponiamo che $A^n = cI$. Allora segue dalla proposizione 2 che A è diagonalizzabile. Infatti la condizione soddisfatta da A è che $P(A) = 0$, dove $P(t) = t^n - c$. D'altra parte, se b è un numero complesso tale che $b^n = c$, le radici di $P(t)$ sono i numeri complessi della forma $b\zeta$ dove ζ è una radice n-esima dell'unità, e dunque $P(t)$ ha $n = \deg(P)$ radici distinte.

Dato un polinomio a coefficienti complessi $P(t) = \sum a_i t^i$ definiamo la sua *derivata* $P'(t)$ ponendo

$$P'(t) = \sum_i i a_i t^{i-1}$$

Segue dalla definizione che la derivata di un polinomio costante è nulla. Inoltre l'operazione di derivazione, come si può facilmente verificare, gode delle usuali proprietà:

$$(P+Q)' = P' + Q'$$
$$(PQ)' = PQ' + P'Q$$

Lemma 4. *Una radice di $P(t)$ è multipla se e solo se è radice anche di $P'(t)$.*

Dimostrazione. Sia λ una radice di $P(t)$. Possiamo dunque scrivere $P(t) = (t-\lambda)Q(t)$ per qualche polinomio $Q(t)$. Derivando otteniamo che

$$P'(t) = Q(t) + (t-\lambda)Q'(t)$$

Ne segue che λ è radice di $P'(t)$ se e solo se è radice di $Q(t)$. Ma d'altro canto λ è radice di $Q(t)$ se e solo se $(t-\lambda)$ divide $Q(t)$, cioè se e solo se $(t-\lambda)^2$ divide $P(t)$, il che equivale a dire che λ è radice multipla di $P(t)$, $\square$

Esempio 2. Sia A una matrice complessa quadrata tale che $A^3 - 3A + 1 = 0$. Allora A è diagonalizzabile. Per dimostrarlo, in base al criterio 2, basta mostrare che $P(t) = t^3 - 3t + 1$ non ha radici multiple. Per il lemma 4 ciò equivale a dire che $P(t)$ e la sua derivata $P'(t)$ non hanno radici comuni. Ma $P'(t) = 3t^2 - 3$ ha come sole radici 1 e -1, mentre $P(1) = -1$ e $P(-1) = 3$.