CRITERI DI DIAGONALIZZABILITÀ

Queste note presentano una versione semplificata dei risultati già esposti nel contributo precedente sulla relazione fra la dimensione degli autospazi e la molteplicità, e la loro applicazione alla formulazione di criteri di diagonalizzabilità.

TEOREMA: Sia $A: X \to X$, $0 < \dim X < \infty$, sia λ_0 un autovalore di A, e sia k la dimensione del relativo autospazio, cioè la dimensione di $Ker(A - \lambda_0 I)$. Allora la molteplicità (algebrica) μ di λ_0 come soluzione dell'equazione caratteristica di A verifica $\mu \geq k$.

DIM. Sia $u_1 \dots u_k$ una base di $Ker(A - \lambda_0 I)$ e sia

$$u_1\, u_2 \dots u_k\, v_{k+1} \dots v_n$$

un suo completamento ad una base di X.

Poiché il polinomio caratteristico è invariante per cambio di base, può essere calcolato impiegando la matrice associata ad una qualunque base, ed in particolare $u_1 \dots u_k\, v_{k+1} \dots v_n$.

Poiché $A(u_i) = \lambda_0 u_i = \lambda_0 u_i + \sum_{j \neq i} 0 \cdot u_j$ ne segue che le prime k colonne della matrice associata sono formate da zeri salvo l'elemento i-esimo, che vale λ_0.

La matrice associata sarà dunque

$$\begin{array}{r l}
 & \overbrace{\hphantom{\lambda_0 \quad 0 \quad \cdots \quad 0 \quad 0}}^{k \text{ colonne}} \\
k \text{ righe} \left\{ \begin{array}{c} \\ \\ \\ \\ \\ \end{array} \right. &
\begin{array}{ccccc|c}
\lambda_0 & 0 & \cdots & 0 & 0 & \\
0 & \lambda_0 & & \vdots & \vdots & \\
\vdots & & \ddots & 0 & 0 & A \\
0 & \cdots & 0 & \lambda_0 & 0 & \\
0 & \cdots & 0 & 0 & \lambda_0 & \\
\hline
0 & & 0 & 0 & 0 & \\
\vdots & & \vdots & \vdots & \vdots & B \\
0 & & 0 & 0 & 0 &
\end{array}
\end{array}$$

A e B dipendono dal completamento scelto, e nulla più dirsi sulla loro struttura.

Il corrispondente polinomio caratteristico si può calcolare mediante lo sviluppo di Laplace rispetto alle prime k colonne (una alla volta) ottenendo

$$\left| \begin{array}{cccc|c}
\lambda_0-\lambda & 0 & \cdots & 0 & \\
0 & \lambda_0-\lambda & \cdots & 0 & \\
\vdots & & & \vdots & A \\
0 & \cdots & 0 & \lambda_0-\lambda & \\
\hline
0 & \cdots & \cdots & 0 & \\
\vdots & & O & \vdots & B-\lambda I \\
0 & \cdots & \cdots & 0 &
\end{array} \right| = (\lambda_0-\lambda)^k \det(B-\lambda I)$$

che coincide, per l'invarianza del polinomio caratteristico, con quello originale, da cui la tesi. $\blacksquare$

<u>NOTA</u>: $\det(B-\lambda I)$ può ancora annullarsi per $\lambda=\lambda_0$, oppure no.

Il passo successivo richiede un teorema fondamentale, la cui prova è riferita in un altro contributo:

TEOREMA: Ogni sistema di autovettori relativi ad autovalori a due a due distinti è indipendente.

Una conseguenza (quasi) immediata è l'importante

COROLLARIO: La somma di un numero arbitrario di autospazi relativi ad autovalori distinti è diretta.

Dim. Siano $A_{\lambda_1} \dots A_{\lambda_h}$ h autospazi relativi ad autovalori distinti, e siano $u_i \in A_{\lambda_i}$ altrettanti loro elementi (e quindi autovettori, o nulli). Supponiamo che $\sum_1^h u_i = 0$ e proviamo che $u_i = 0$ $\forall i = 1..h$. Per assurdo, siano $u_{i_1} \dots u_{i_p}$ tutti e soli i vettori non nulli (e quindi autovettori) fra $u_1 \dots u_h$. Poiché $\sum_1^p u_{i_j} = \sum_{i=1}^h u_i$ (perché le somme differiscono solo per vettori nulli) si ha $\sum^p u_{i_j} = 0$, con u_{i_j} autovettori in autospazi distinti, il che è in contrasto col teorema precedente, perché $\sum_1^p u_{i_j} = 0$ implica che $u_{i_1} \dots u_{i_k}$ siano dipendenti.

Ne segue ancora (quasi) subito un naturale criterio di diagonalizzabilità:

TEOREMA: Condizione necessaria e sufficiente perché $A : X \to X$, $0 < \dim X < \infty$, sia diagonalizzabile è che

$$\sum_{\lambda \in \sigma(A)} \dim A_\lambda = \dim X$$

In sostanza occorre che la somma delle dimensioni degli autospazi di tutti gli autovalori disponibili faccia $\dim X$.

DIM. $\boxed{\text{C.S.} \quad \sum_{\lambda \in \sigma(A)} \dim A_\lambda = \dim X \Rightarrow A \text{ diagonalizzabile}}$

Siano $\lambda_1 \dots \lambda_h$ gli elementi dello spettro $\sigma(A)$, e siano

$u_1^1 \dots u_{k_1}^1$ una base di $A_{\lambda_1} \Rightarrow \dim A_{\lambda_1} = k_1$

$u_1^2 \dots u_{k_2}^2$ una base di $A_{\lambda_2} \Rightarrow \dim A_{\lambda_2} = k_2$

$\vdots$

$u_1^h \dots u_{k_h}^h$ una base di $A_{\lambda_h} \Rightarrow \dim A_{\lambda_h} = k_h$

Dall'ipotesi abbiamo $\sum k_i = \dim X$.

Poiché la somma degli A_{λ_i} è diretta, ne segue che

$$\dim\left(\bigoplus_{j=1}^{h} A_{\lambda_i}\right) = \sum_{j=1}^{h} \dim A_{\lambda_j} = \dim X$$ e dunque,

essendo $\bigoplus_{j=1}^{h} A_{\lambda_h}$ un sottospazio d'X che ha la stessa dimensione d'X, ne segue $X = \bigoplus_{j=1}^{h} A_{\lambda_h}$. La base spettrale richiesta è allora costituita dall'unione delle basi u_j^i di tutti gli autospazi, che sono indipendenti nel loro complesso perché la somma di A_{λ_i} è diretta.

$$\boxed{\text{C.N.} \quad A \text{ diagonalizzabile} \Longrightarrow \sum_{\lambda \in \sigma(A)} \dim A_\lambda = \dim X}$$

Sia $u_1^1 .. u_{k_1}^1\, u_1^2 ... u_{k_2}^2 ... u_1^h .. u_{k_h}^h$ una base spettrale, ordinata in modo che $u_1^i ... u_{k_i}^i$ siano autovettori relativi allo stesso autovalore λ_i. La tesi segue subito dalla proprietà delle somme dirette e delle sue basi se si prova che

$$A_{\lambda_i} = \langle u_1^i, .., u_{k_i}^i \rangle$$

Infatti, sia $u \in A_{\lambda_i}$, sicché $A(u) = \lambda_i u$. Poiché $\{u_j^i\}$ è una base si avrà, per opportuni scalari α_q^p, $u = \sum_{\substack{p=1..h \\ q=1..k_h}} \alpha_q^p u_q^p$ e dunque

$$\begin{array}{ccl} A(u) = \lambda_i u & = & \sum_{\substack{p=1..h \\ q=1..k_p}} \lambda_i \alpha_q^p u_q^p \\ \| & & \\ \sum_{\substack{p=1..h \\ q=1..k_p}} \alpha_q^p A(u_q^p) & = & \sum_{\substack{p=1..h \\ q=1..k_p}} \alpha_q^p \lambda_p u_q^p \end{array}$$

da cui si ha

$$\sum_{\substack{p=1..h\\ q=1..k_p}} \alpha^p_q(\lambda_i-\lambda_p)\,u^p_q = 0$$

Dall'indipendenza di u^p_q (elementi di una base) segue

$$\alpha^p_q(\lambda_i-\lambda_p)=0 \qquad p=1..h \quad q=1..k_p$$

e dunque, $\lambda_p=\lambda_i$ oppure $\alpha^p_q=0$. Dunque, le uniche coordinate non nulle debbono di necessità essere relative a $u^i_1 \dots u^i_{k_i}$, $p=i$, da cui u è una loro combinazione.

$\blacksquare$

Il criterio presentato ha un tallone d'Achille: il calcolo degli autovalori! Se si conoscono, per qualche via traversa, tutti gli autovalori, una semplice applicazione dell'algoritmo di Gauss al sistema lineare associato all'equazione $A(u)=\lambda u$, una volta fissato come valore di λ uno degli autovalori noti, consente di determinare l'autospazio relativo oppure, limitandosi alla sola riduzione a scala senza la riduzione completa, di calcolarne la sua dimensione (numero delle colonne non-pivot).
Purtroppo, la determinazione degli autovalori è un problema di ben altro spessore: occorre risolvere un'equazione algebrica!

Il criterio presentato può comunque essere raffinato in vario modo utilizzando la relazione precedente fra molteplicità algebrica di un autovalore e dimensione del suo autospazio. Osserviamo, infatti, che il grado del polinomio caratteristico coincide con la dimensione di X. Ricordiamo anche che, se $p_1(\lambda)$ e $p_2(\lambda)$ sono due polinomi e se $\deg(p_1) \geq \deg(p_2)$, esistono $q(\lambda)$, il quoziente, e $r(\lambda)$, il resto tali che

$$p_1(\lambda) = q(\lambda)\, p_2(\lambda) + r(\lambda)$$

con il grado del resto strettamente minore di quello del divisore

$$\deg(r) < \deg(p_2)$$

La prova è conseguenza diretta dell'algoritmo di Euclide di divisione con resto, imparato a scuola per i numeri: invece di avere il resto strettamente minore del divisore, si avrà il GRADO del resto strettamente minore del GRADO del divisore. Ne segue subito il teorema di Ruffini:

TEOREMA (Ruffini): **Un polinomio** $p(\lambda)$ **risulta divisibile per** $(\lambda - \lambda_0)$ **se e solo se** $p(\lambda_0) = 0$

DIM. $p(\lambda) = q(\lambda)(\lambda-\lambda_0) + r(\lambda)$ $\quad (*)$

ove $\deg r < \deg(\lambda-\lambda_0)$, e quindi $\deg r = 0$, e così $r(\lambda)$ è costante. Allora, da $(*)$ per $\lambda = \lambda_0$ segue

$$p(\lambda_0) = r(\lambda_0)$$

da cui $r(\lambda_0)=0$ e $p(\lambda) = q(\lambda)(\lambda-\lambda_0)$ se e solo se $p(\lambda_0)=0$.

$\boxed{//}$

Questo risultato, assieme al formidabile teorema fondamentale dell'Algebra di Gauss (forse il primo teorema della Matematica moderna), che asserisce che ogni polinomio non costante a coefficienti complessi ha radici complesse, consente di formulare un teorema di fattorizzazione:

TEOREMA (di fattorizzazione dei polinomi in $\mathbb{C}$):
Sia $p(\lambda)$ un polinomio di grado $n > 0$, a coefficienti in $\mathbb{C}$. Allora esistono $\lambda_1, \lambda_2, \dots, \lambda_n$, eventualmente coincidenti, e $\alpha \in \mathbb{C}$ tali che

$$p(\lambda) = \alpha(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$$

La costante α coincide col coefficiente del termine di grado n.

DIM. Infatti, per il teorema di Gauss, il polinomio non costante p si annulla per qualche $\lambda_1 \in \mathbb{C}$ e, per il teorema di Ruffini, è divisibile per $\lambda - \lambda_1$: sia $p_1(\lambda)$ il quoziente. Se $\deg p_1 = 0$, esso sarà costante e, detta α tale costante, si avrà $p(\lambda) = \alpha(\lambda - \lambda_1)$. Altrimenti, si possono applicare di nuovo i teoremi di Gauss e di Ruffini a $p_1(\lambda)$, che ora sarà di grado $n-1$ e, dopo n passi, seguirà la tesi.

$\boxed{//}$

NOTA: per una maggiore "pulizia" formale si può usare l'induzione sul grado del polinomio, abbreviando ulteriormente la prova.

NOTA: La fattorizzazione è "unica", sempre in conseguenza del teorema di Ruffini. Se $\alpha \prod_1^n (\lambda - \lambda_i) = \beta \prod_1^n (\lambda - \mu_i)$ il primo membro è divisibile per $(\lambda - \lambda_i)$, e così deve essere per il secondo membro. Ne segue che $\mu_1 \ldots \mu_n$ possono differire da $\lambda_1 \ldots \lambda_n$ solo per il loro ordine. Dividendo poi ambo i membri per $\prod_1^n (\lambda - \lambda_i)$ segue $\alpha = \beta$.

DEFINIZIONE: Una radice λ_0 del polinomio p si dirà di molteplicità μ se la fattorizzazione di p contiene esattamente μ fattori uguali a $(\lambda - \lambda_0)$.

Concludendo:

TEOREMA: Se p è un polinomio complesso non costante esistono $\lambda_1, \dots, \lambda_k \in \mathbb{C}$ (tutte le radici), $\mu_1 \dots \mu_k \in \mathbb{N}$ (le loro molteplicità), ed una costante $\alpha \in \mathbb{C}$ (il coefficiente del termine di grado massimo) tali che

$$p(\lambda) = \alpha \prod_{i=1}^{k} (\lambda - \lambda_i)^{\mu_i} =$$

$$= \alpha (\lambda - \lambda_1)^{\mu_1} (\lambda - \lambda_2)^{\mu_2} \dots (\lambda - \lambda_k)^{\mu_k}$$

ove le molteplicità $\mu_1 \dots \mu_n$ verificano

$$\sum_{i=1}^{k} \mu_i = \deg p$$

▨

Purtroppo, un simile teorema è falso su $\mathbb{R}$: $p(\lambda) = \lambda^2 + 1$ non si può scrivere come prodotto di 1 e di due fattori di primo grado, perché non ha zeri reali.

Si vede bene come l'apparente astrattezza del teorema fondamentale dell'Algebra abbia conseguenze assai "concrete".

I concetti ora introdotti consentono qualche ulteriore raffinamento del criterio di diagonalizzabilità.

<u>COROLLARIO</u> (<u>criterio di diagonalizzabilità su $\mathbb{C}$</u>):

<u>Condizione necessaria e sufficiente perché $A : X \to X$,</u>
X <u>complesso con</u> $0 < \dim X < \infty$, <u>sia diagonalizzabile è che, per ogni autovalore dello spettro, la sua molteplicità algebrica nell'equazione caratteristica coincida con la dimensione del suo autospazio.</u>

<u>DIM</u>. Poiché stiamo considerando autovalori complessi ($\lambda_1 \dots \lambda_k$ di molteplicità rispettive $\mu_1 \dots \mu_k$) risulta $\sum_1^k \mu_i = \dim X$, in quanto $\dim X$ è il grado del polinomio caratteristico. Ne segue subito che, se $\mu_i = \dim A_{\lambda_i}$ allora $\sum_{\lambda_i \in \sigma(A)} \dim A_{\lambda_i} = \dim X$. ▨

Viceversa, se A è diagonalizzabile, $\bigoplus_1^k A_{\lambda_i} = X$ e dunque $\sum_1^k \dim A_{\lambda_i} = \dim X$. Posto $\alpha_i = \dim A_{\lambda_i}$, si ha

$$\sum_1^k \alpha_i = \dim X = \sum_1^k \mu_i$$

da cui $\sum_1^k (\mu_i - \alpha_i) = 0$ e, tenendo conto che $\mu_i \geq \alpha_i$, segue $\mu_i - \alpha_i = 0$ e cioè $\boxed{\mu_i = \dim A_{\lambda_i}}$.

▨

Nel caso della diagonalizzabilità su $\mathbb{R}$, non tutto è perduto: basta aggiungere come ulteriore ipotesi che $p(\lambda)$ abbia deg p radici reali: ciò che non si può avere di diritto dal teorema di Gauss, si introduce come ipotesi. Tale richiesta è meno esagerata di quanto non possa sembrare: le matrici reali simmetriche, ad esempio, oltre ad avere diritto alla fattorizzazione in $\mathbb{C}$ del polinomio caratteristico, hanno anche spettro reale, e dunque verificano immediatamente la condizione precedente senza bisogno di alcun controllo! Dunque:

<u>TEOREMA</u> (<u>diagonalizzabilità in $\mathbb{R}$</u>): <u>Sia</u> $A: X \to X$, X <u>reale con</u> $0 < \dim X < \infty$, $\lambda_1 \dots \lambda_k$ <u>siano i suoi autovalori reali di molteplicità</u> $\mu_1 \dots \mu_k$, <u>con</u> $\sum_1^k \mu_i = \dim X$, <u>allora</u> A <u>è diagonalizzabile se e solo se</u>

$$\mu_i = \dim A_{\lambda_i} \equiv \dim \operatorname{Ker}(A - \lambda_i I) \quad i = 1 \dots k$$

$\boxed{///}$

Il criterio sulla dimensione degli autospazi offre ancora una scorciatoia. Infatti, se λ_0 è un autovalore semplice, e cioè $\mu = 1$, ne segue $\dim A_{\lambda_0} \leq 1$. Poiché ogni autospazio contiene almeno un autovettore ($\neq 0$), si ha che la sua dimensione è almeno 1. Ne segue che la condizione è automaticamente verificata per tutti gli autovalori semplici. Ecco un criterio pratico di diagonalizzabilità:

CRITERIO DI DIAGONALIZZABILITA': Sia $A: X \to X$, $0 < \dim X < \infty$, X reale o complesso. Siano $\lambda_1 \dots \lambda_h$ tutti gli autovalori di $\sigma(A)$ e siano $\mu_1 \dots \mu_h$ le loro molteplicità. Allora A è diagonalizzabile se e solo se:

1) $\sum_{i=1}^{h} \mu_i = \dim X$

2) per ogni autovalore multiplo λ_i (tale che, cioè $\mu_i > 1$)

$$\dim A_{\lambda_i} = \mu_i$$

$\blacksquare$

NOTA: La condizione 1) è automaticamente verificata in $\mathbb{C}$, mentre deve essere verificata in $\mathbb{R}$.

Un'altro risultato, conseguenza immediata del fatto che la condizione $\dim A_\lambda = \mu_\lambda$ è automaticamente verificata per gli autovalori semplici, è il seguente:

TEOREMA (degli autovalori semplici): Sia $A: X \to X$ con $0 < n = \dim X < \infty$. Allora, se A possiede n autovalori distinti, è diagonalizzabile.

DIM. Se A possiede n autovalori distinti, poiché $n = \dim X$ e il grado del polinomio caratteristico è pari a $\dim X$, ne segue che gli autovalori hanno tutti molteplicità 1 e dunque le due condizioni per la diagonalizzabilità sono verificate.

UN OPERATORE NON DIAGONALIZZABILE SU $\mathbb{C}$

NOTA: L'operatore $A\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}0&0\\1&0\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}0\\x\end{pmatrix}$, da $\mathbb{R}^2$ in sé, ha polinomio caratteristico $p(\lambda)=\lambda^2$, e dunque ha solo l'autovalore 0 di molteplicità $\mu=2$, ma la dimensione del relativo autospazio è **uno**, **STRETTAMENTE MINORE DELLA MOLTEPLICITA'**. Di conseguenza, essendo $\mathbb{R}^2$ di dimensione 2, A non è diagonalizzabile.

In definitiva, un operatore può non essere diagonalizzabile tanto perché non ha autovalori in numero sufficiente (polinomio caratteristico con zeri reali in numero strettamente minore del suo grado) quanto perché, pur disponendo di autovalori in numero sufficiente (vedi l'esempio precedente, ove gli autovalori sono due ... o, 0 è doppio!) non dispone di autovettori in "numero" sufficiente: gli infiniti autovettori dell'esempio precedente non bastano a generare lo spazio $\mathbb{R}^2$! Il criterio si rivela dunque molto utile per stabilire la diagonalizzabilità...

... **SE SI CONOSCONO GLI AUTOVALORI!**

Una nota finale sull'ipotesi ricorrente $0 < \dim X < \infty$.
Le dimostrazioni precedenti utilizzano sistematicamente il concetto di base: persino la definizione di diagonalizzabilità lo fa! Se $\dim X = 0 \iff X = \{0\}$ non esistono basi, e non c'è nulla di interessante da studiare! Se invece $\dim X = \infty$, basi non ce ne sono ed è tutto da rifare, perché i problemi ci sono e sono interessantissimi. Le tecniche sono diverse, e sono oggetto di una disciplina matematica che ha perfino un nome diverso: l'Analisi Funzionale, nata e sviluppatasi prima della guerra, e ancora attivo campo di ricerca.
Il nome stesso suggerisce che ci debba entrare in qualche modo la continuità ....

Un'ultima nota di colore: in qualche libro, la dimensione di un autospazio viene definita MOLTEPLICITA' GEOMETRICA, e la molteplicità dell'autovalore ad esso relativo MOLTEPLICITA' ALGEBRICA dell'autovalore...
... molto pittoresco!

<u>Un esempio finale</u>: studiare la diagonalizzabilità su $\mathbb{R}$ di

$$A\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} y \\ y \\ z \end{pmatrix}$$

Adoperando la base canonica per rappresentare A, si può scrivere subito polinomio ed equazione caratteristici:

$$0 = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & 1-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{vmatrix} = -\lambda(1-\lambda)^2$$

Dunque, lo spettro contiene $\lambda = 0$ di molteplicità 1 e $\lambda = 1$ di molteplicità 2, e la somma delle molteplicità è $3 = \dim \mathbb{R}^3$. L'operatore A sarà dunque diagonalizzabile se e solo se la dimensione dell'autospazio relativo a $\lambda = 1$ è 2. Sostituendo $\lambda = 1$ nella matrice $A - \lambda I$, che appare nel determinante precedente, si ottiene per l'autospazio di 1

$$\underbrace{\begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{(A-1I)}\underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_{(u)} = \underbrace{\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}}_{=\ 0} \quad \text{e cioè} \quad \boxed{-x+y=0},$$

che ha soluzioni $\begin{pmatrix} \alpha \\ \alpha \\ \beta \end{pmatrix} = \alpha\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \beta\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, $\alpha, \beta \in \mathbb{R}$. L'autospazio è dunque $\left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle$, ed ha dimensione 2, uguale alla molteplicità algebrica di $\lambda = 1$ nell'equazione caratteristica. L'operatore A è dunque diagonalizzabile su $\mathbb{R}$.