Diagonalizzazione di una matrice quadrata

Sia A una matrice quadrata di ordine n e siano $\lambda_1, \lambda_2, \ldots, \lambda_t$ i t autovalori reali di A. Ricordiamo che $\mu_g(\lambda_i) \leq \mu_a(\lambda_i)$ $\forall$ i e dunque $\sum_{i=1}^{t} \mu_g(\lambda_i) \leq \sum_{i=1}^{t} \mu_a(\lambda_i)$. Inoltre $\sum_{i=1}^{t} \mu_a(\lambda_i) \leq n$, quindi:

$$\sum_{i=1}^{t} \mu_g(\lambda_i) \leq \sum_{i=1}^{t} \mu_a(\lambda_i) \leq n.$$

Teorema. Una matrice A $n \times n$ è diagonalizzabile $\Leftrightarrow$ esiste una base di $\mathbb{R}^n$ formata da autovettori di A.

Dimostrazione. A è diagonalizzabile $\Leftrightarrow$ esiste una matrice invertibile P ed una matrice diagonale D tale che $A = PDP^{-1}$. Moltiplicando a destra ambo i membri per P, si ha $AP = PD$ e P deve essere invertibile. Siano $v_1, v_2, \ldots, v_n$ le colonne di P. Dunque scriviamo P come $(v_1 v_2 \cdots v_n)$, quindi abbiamo suddiviso la matrice in n blocchi ognuno dei quali corrisponde ad una delle n colonne. Quindi $AP = A(v_1 v_2 \cdots v_n) = (Av_1 Av_2 \cdots Av_n)$ (ricordiamo che quando si moltiplicano due matrici suddivise in blocchi, possiamo trattare il prodotto tra blocchi come il prodotto tra numeri; in questo caso A è considerata come un unico blocco). Suddividiamo la matrice D in n blocchi, ognuno corrispondente ad una delle n righe, quindi $D = \begin{pmatrix} w_1 \\ w_2 \\ \cdots \\ w_n \end{pmatrix}$ e w_i è un vettore di $\mathbb{R}^n$ che ha d_i nella i-esima componente e 0 altrove, cioè $w_i = d_i e_i$ ed e_i è un vettore della base canonica. Dunque, $PD = (v_1 v_2 \cdots v_n) \begin{pmatrix} w_1 \\ w_2 \\ \cdots \\ w_n \end{pmatrix} = v_1 w_1 + v_2 w_2 + \ldots + v_n w_n$.

Ricordiamo che v_i è un vettore colonna, quindi $v_i = \begin{pmatrix} v_{i1} \\ v_{i2} \\ \cdots \\ v_{in} \end{pmatrix}$ e $w_i = d_i e_i$, dunque $v_i w_i =$

$$\begin{array}{c} \begin{pmatrix} 0 & 0 & \cdots & d_i v_{i1} & \cdots & 0 \\ 0 & 0 & \cdots & d_i v_{i2} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & d_i v_{in} & \cdots & 0 \end{pmatrix}. \\ \downarrow \\ \text{i-esima colonna} \end{array}$$

Quindi $v_1w_1 + v_2w_2 + \ldots + v_nw_n = (d_1v_1d_2v_2 \cdots d_nv_n)$, dove la suddivisione in blocchi è ancora una volta per colonne. Si ha allora $AP = PD \Leftrightarrow (Av_1Av_2 \cdots Av_n) = (d_1v_1d_2v_2 \cdots d_nv_n)$, dunque $Av_i = d_iv_i \ \forall \ i$. Poichè P è invertibile, nessuno dei vettori v_i è nullo, quindi sono n autovettori di A e, sempre perchè P è invertibile, quindi $\det P \neq 0$, essi sono indipendenti. Quindi A è diagonalizzabile se e solo se esistono n autovettori di A che formano un insieme indipendente, cioè esiste una base di $\mathbb{R}^n$ formata da autovettori di A.

Abbiamo dimostrato che A è diagonalizzabile se esistono n autovettori di A che formano un insieme indipendente. In ogni autospazio V_{λ_i} possiamo prendere insieme indipendenti che contengono al più $\dim V_{\lambda_i} = \mu_g(\lambda_i)$ vettori. Inoltre sappiamo che se uniamo insiemi indipendenti di vettori contenuti in autospazi distinti, otteniamo un insieme indipendente di $\mathbb{R}^n$, quindi, in generale, noi abbiamo che un insieme indipendente di autovettori di A contiene al più $\displaystyle\sum_{i=1}^{t} \mu_g(\lambda_i)$ vettori. Dunque A è diagonalizzabile $\Leftrightarrow \displaystyle\sum_{i=1}^{t} \mu_g(\lambda_i) = n$. Dalla condizione: $\displaystyle\sum_{i=1}^{t} \mu_g(\lambda_i) \leq \sum_{i=1}^{t} \mu_a(\lambda_i) \leq n$, allora si ottiene che se A è diagonalizzabile allora che anche $\displaystyle\sum_{i=1}^{t} \mu_a(\lambda_i) = n$ e $\mu_g(\lambda_i) = \mu_a(\lambda_i) \ \forall \ i$. Quindi, quando vogliamo controllare che A sia diagonalizzabile, dobbiamo procedere in due step:

1. Risolvere l'equazione in λ: $\det(A - \lambda I) = 0$. È un'equazione di grado n, quindi la somma delle molteplicità delle soluzioni è al più n. Se la somma delle molteplicità è proprio n, allora si procede al passo 2. Se si ha $\displaystyle\sum_{i=1}^{t} \mu_a(\lambda_i) < n$, da $\displaystyle\sum_{i=1}^{t} \mu_g(\lambda_i) \leq \sum_{i=1}^{t} \mu_a(\lambda_i)$ si ricava che anche $\displaystyle\sum_{i=1}^{t} \mu_g(\lambda_i) < n$ quindi la matrice non è diagonalizzabile.
2. Per ogni autovalore λ_i, trovare $\mu_g(\lambda_i) = \dim V_{\lambda_i} = n - \rho(A - \lambda_i I)$. Se $\exists \ i$ tale che $\mu_g(\lambda_i) < \mu_a(\lambda_i)$, allora $\displaystyle\sum_{i=1}^{t} \mu_g(\lambda_i) < \sum_{i=1}^{t} \mu_a(\lambda_i) = n$, quindi $\displaystyle\sum_{i=1}^{t} \mu_g(\lambda_i) < n$, quindi A non è diagonalizzabile. Se $\mu_g(\lambda_i) = \mu_a(\lambda_i) \ \forall i$, allora $\displaystyle\sum_{i=1}^{t} \mu_g(\lambda_i) = \sum_{i=1}^{t} \mu_a(\lambda_i) = n$ e dunque A è diagonalizzabile.

Quindi si può anche dire che A è diagonalizzabile se e solo se la molteplicità geometrica di ogni autovalore è uguale alla molteplicità algebrica e la somma delle molteplicità è uguale a n.

Osserviamo, inoltre, che il Teorema ci dice anche come trovare le matrici P e D. Nella matrice D in diagonale vanno gli autovalori ripetuti secondo la loro molteplicità, quindi l'autovalore λ_i andrà ripetuto $\mu_a(\lambda_i)$ volte. Nella matrice P vanno messi per colonna i vettori di $B_1, B_2, \ldots, B_t$, dove B_i è una base dell'autospazio V_{λ_i}. Attenzione! I vettori in P vanno messi in maniera *coerente* rispetto all'ordine scelto per D: nella j–esima colonna va un autovettore associato ad un certo λ se e solo se $d_j = \lambda$.

Teorema. Una matrice simmetrica è diagonalizzabile.

Si omette la dimostrazione.