6. Matrici normali

Sia V uno spazio vettoriale reale o complesso di dimensione finita, munito di prodotto scalare, o hermitiano, definito positivo $\langle\ ,\ \rangle$. Una applicazione lineare $f : V \to V$ si dice *normale* se commuta con la sua aggiunta f^*, cioè se $ff^* = f^*f$. Ad esempio sono normali le applicazioni lineari autoaggiunte o quelle unitarie. Sia $v_1, \ldots, v_n$ una base ortonormale di V e sia A la matrice di f rispetto a questa base. Dire che f è normale equivale a dire che $A\,{}^t\overline{A} = {}^t\overline{A}A$; una matrice con questa proprietà verrà detta normale.

TEOREMA (6.1). *Sia V uno spazio vettoriale complesso di dimensione finita munito di un prodotto hermitiano definito positivo e sia $f : V \to V$ una applicazione lineare normale. Allora vi è una base ortonormale di V interamente costituita da autovettori di f.*

È chiaro che la traduzione di questo teorema nel linguaggio delle matrici è la seguente: *se A è una matrice normale vi è una matrice unitaria U tale che $U^{-1}AU$ sia diagonale.*

Dimostreremo (6.1) per induzione sulla dimensione di V. Se questa vale 1 non vi è nulla da dimostrare dato che A è già diagonale. Per il passo induttivo ci serve la seguente osservazione, che è utile e interessante anche di per sè.

LEMMA (6.2). *Siano f e g applicazioni lineari di V in sè tali che $fg = gf$. Allora, se $V \neq \{0\}$, vi è un elemento non nullo di V che è autovettore sia per f che per g.*

Sia λ un autovalore di f, e sia V_λ l'autospazio corrispondente. Se $v \in V_\lambda$,

$$f(g(v)) = g(f(v)) = g(\lambda v) = \lambda g(v)\,.$$

In altre parole, $g(v) \in V_\lambda$. Dunque $g(V_\lambda) \subset V_\lambda$ e g induce, per restrizione, una applicazione lineare di V_λ in sè. Questa applicazione ha almeno un autovettore $v \in V_\lambda$. Dunque v è un autovettore sia per f che per g. Questo dimostra (6.2).

Torniamo alla dimostrazione di (6.1). Il lemma che abbiamo appena dimostrato ci dice che vi è un vettore non nullo $v_1 \in V$ tale che $f(v_1) = \lambda v_1$ e $f^*(v_1) = \mu v_1$, dove λ e μ sono opportuni numeri complessi. Notiamo subito che $\mu = \overline{\lambda}$. Infatti

$$\lambda\langle v_1, v_1\rangle = \langle f(v_1), v_1\rangle = \langle v_1, f^*(v_1)\rangle = \overline{\mu}\langle v_1, v_1\rangle\,.$$

In definitiva si ha che

$$\begin{aligned} f(v_1) &= \lambda v_1\,, \\ f^*(v_1) &= \overline{\lambda} v_1\,. \end{aligned}$$

Possiamo anche supporre che $\|v_1\| = 1$. Sia ora $W = \{w \in V : \langle w, v_1\rangle = 0\}$ il complemento ortogonale di v_1. Per ogni $w \in W$ si ha che

$$\langle f(w), v_1\rangle = \langle w, f^*(v_1)\rangle = \lambda\langle w, v_1\rangle = 0\,,$$

e quindi $f(W) \subset W$. Analogamente $f^*(W) \subset W$, e l'aggiunta dell'applicazione lineare da W in sè indotta da f non è altro che l'applicazione lineare indotta da f^*. Dato che la dimensione di W è $\dim(V) - 1$, sappiamo per ipotesi induttiva che W ha una base ortonormale $v_2, \ldots, v_n$ costituita da autovettori di f. La base di V cercata è $v_1, \ldots, v_n$. Questo conclude la dimostrazione del teorema (6.1).

Ci si può chiedere se nel caso reale valga un analogo del teorema appena dimostrato. In questi termini, la risposta è negativa. Infatti, se A è una matrice reale e U è una matrice ortogonale tale che $U^{-1}AU$ sia diagonale, gli autovalori di A sono necessariamente tutti reali. D'altra parte la matrice ortogonale, e quindi normale,

$$\begin{pmatrix} \cos\vartheta & \sin\vartheta \\ -\sin\vartheta & \cos\vartheta \end{pmatrix}$$

ha come autovalori $\exp(\mathrm{i}\vartheta)$ e $\exp(-\mathrm{i}\vartheta)$, che non sono reali a meno che ϑ non sia della forma $2k\pi$ o $(2k+1)\pi$, dove k è un intero. Qualcosa si può tuttavia fare anche nel caso reale. Per spiegarlo useremo il linguaggio delle matrici.

Proposizione (6.3). *Sia A una matrice reale normale. Esiste una matrice ortogonale U tale che $U^{-1}AU$ sia la matrice diagonale a blocchi*

$$\begin{pmatrix} B_1 & 0 & \ldots & & \\ 0 & B_2 & 0 & \ldots & \\ \ldots & 0 & B_3 & 0 & \ldots \\ & & \ldots & & \\ & & \ldots & 0 & B_h \end{pmatrix}$$

dove ogni B_i è o un numero reale o una matrice reale 2×2 della forma

$$\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}.$$

Indichiamo con n la dimensione di A. Sappiamo che, se consideriamo A e ${}^tA$ come matrici complesse, vi è un autovettore comune ad entrambe, che indichiamo con X. Sappiamo dunque che

$$\begin{aligned} AX &= \lambda X\,, \\ {}^tAX &= \overline{\lambda}X\,. \end{aligned} \tag{6.4}$$

Se λ è reale questo ci dice che X è soluzione di un sistema omogeneo di equazioni lineari reali. Prendendo le parti reali o immaginarie dei due membri delle (6.4) si trova che anche i vettori reali $\mathrm{Re}(X) = (X + \overline{X})/2$ e $\mathrm{Im}(X) = (X - \overline{X})/2\mathrm{i}$ sono soluzioni dello stesso sistema; dato che $X \neq 0$, almeno uno dei due non è nullo. Si può dunque supporre che X sia reale, e anche che abbia norma 1. Sia ora W il complemento ortogonale di X in $\mathbb{R}^n$; come nel caso complesso si vede che $AW \subset W$ e ${}^tAW \subset W$. Dunque, se $X_2, \ldots, X_n$ è una base ortonormale di W, la matrice U_1 le cui colonne sono $X, X_2, \ldots, X_n$ è ortogonale e

$$U_1^{-1}AU_1 = \begin{pmatrix} \lambda & 0 \\ 0 & B \end{pmatrix},$$

dove B è una matrice normale $(n-1)\times(n-1)$. Induttivamente si può supporre che vi sia una matrice ortogonale U_2 tale che $U_2^{-1}BU_2$ sia come nella tesi della proposizione. Dunque

$$U = U_1 \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix}$$

è ortogonale e

$$\begin{aligned} U^{-1}AU &= \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix}^{-1} U_1^{-1}AU_1 \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix}^{-1} \begin{pmatrix} \lambda & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & U_2^{-1} \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & U_2^{-1}BU_2 \end{pmatrix} \end{aligned}$$

è della forma richiesta.

Occupiamoci ora del caso in cui λ non è reale. Coniugando le (6.4) si ottengono le relazioni

$$\begin{aligned} A\overline{X} &= \overline{\lambda}\,\overline{X}\,, \\ {}^tA\overline{X} &= \lambda\overline{X}\,. \end{aligned}$$

Notiamo che X e $\overline{X}$ sono ortogonali fra loro. Infatti

$$\lambda\langle X,\overline{X}\rangle = \langle AX,\overline{X}\rangle = \langle X,\ {}^tA\overline{X}\rangle = \langle X,\lambda\overline{X}\rangle = \overline{\lambda}\langle X,\overline{X}\rangle\,;$$

dato che $\lambda\neq\overline{\lambda}$ se ne deduce che $\langle X,\overline{X}\rangle = 0$. Poniamo ora

$$\begin{aligned} v &= \mathrm{Re}(X)\,, \\ w &= \mathrm{Im}(X)\,. \end{aligned}$$

Si ha che

$$\begin{aligned} X &= v + \mathrm{i}w\,, \\ \overline{X} &= v - \mathrm{i}w\,. \end{aligned}$$

Inoltre

$$\langle v,w\rangle = \frac{\mathrm{i}}{4}\langle X,X\rangle - \frac{\mathrm{i}}{4}\langle\overline{X},\overline{X}\rangle = 0\,,$$

dato che $\langle X,X\rangle = \langle\overline{X},\overline{X}\rangle$, mentre

$$\|v\|^2 = \frac{1}{4}(\|X\|^2 + \|\overline{X}\|^2) = \|w\|^2\,.$$

Rinormalizzando opportunamente X possiamo dunque supporre che $\|v\| = \|w\| = 1$. Osserviamo ora che, scrivendo a per indicare la parte reale di λ e b per indicare la parte immaginaria,

$$\begin{aligned} Av &= \frac{1}{2}(AX + A\overline{X}) = \frac{1}{2}(\lambda X + \overline{\lambda}\,\overline{X}) = av - bw\,, \\ Aw &= \frac{1}{2\mathrm{i}}(AX - A\overline{X}) = \frac{1}{2\mathrm{i}}(\lambda X - \overline{\lambda}\,\overline{X}) = bv + aw\,, \\ {}^tAv &= \frac{1}{2}(\,{}^tAX + \,{}^tA\overline{X}) = \frac{1}{2}(\overline{\lambda}X + \lambda\,\overline{X}) = av + bw\,, \\ {}^tAw &= \frac{1}{2\mathrm{i}}(\,{}^tAX - \,{}^tA\overline{X}) = \frac{1}{2\mathrm{i}}(\overline{\lambda}X - \lambda\,\overline{X}) = -bv + aw\,. \end{aligned}$$

Ne segue in particolare che, se u è un vettore ortogonale sia a v che a w, allora anche Au e ${}^tAu$ sono ortogonali a v e a w; infatti $\langle Au, v\rangle = \langle u, {}^tAv\rangle = 0$, e così via. Se dunque indichiamo con W il sottospazio di $\mathbb{R}^n$ consistente di tutti i vettori ortogonali sia a v che w, si ha che $AW \subset W$ e ${}^tAW \subset W$. Dunque, se $v_3, \ldots, v_n$ è una base ortonormale di W, la matrice U_1 le cui colonne sono $v, w, v_3, \ldots, v_n$ è ortogonale e

$$U_1^{-1}AU_1 = \begin{pmatrix} B_1 & 0 \\ 0 & B \end{pmatrix},$$

dove

$$B_1 = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

e B è una matrice normale $(n-2)\times(n-2)$. Induttivamente si può supporre che vi sia una matrice ortogonale U_2 tale che $U_2^{-1}BU_2$ sia come nella tesi della proposizione. Dunque, indicando con I_2 la matrice identità 2×2,

$$U = U_1\begin{pmatrix} I_2 & 0 \\ 0 & U_2 \end{pmatrix}$$

è ortogonale e

$$\begin{aligned} U^{-1}AU &= \begin{pmatrix} I_2 & 0 \\ 0 & U_2 \end{pmatrix}^{-1} U_1^{-1}AU_1 \begin{pmatrix} I_2 & 0 \\ 0 & U_2 \end{pmatrix} = \begin{pmatrix} I_2 & 0 \\ 0 & U_2 \end{pmatrix}^{-1} \begin{pmatrix} B_1 & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} I_2 & 0 \\ 0 & U_2 \end{pmatrix} \\ &= \begin{pmatrix} I_2 & 0 \\ 0 & U_2^{-1} \end{pmatrix} \begin{pmatrix} B_1 & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} I_2 & 0 \\ 0 & U_2 \end{pmatrix} = \begin{pmatrix} B_1 & 0 \\ 0 & U_2^{-1}BU_2 \end{pmatrix} \end{aligned}$$

è della forma richiesta.

Concludiamo questa sezione con un risultato che asserisce la possibilità, sotto opportune condizioni, di diagonalizzare simultaneamente più matrici.

Proposizione (6.5). *Sia V uno spazio vettoriale complesso di dimensione finita munito di un prodotto hermitiano definito positivo. Siano $f_1, \ldots, f_h$ applicazioni lineari normali di V in sè tali che $f_if_j = f_jf_i$ per ogni i e ogni j. Allora vi è una base ortonormale di V ogni cui elemento è autovettore per ognuna delle f_i.*

Nel linguaggio delle matrici, la proposizione dice che, date matrici normali $A_1, \ldots, A_h$ che commutano, tali cioè che $A_iA_j = A_jA_i$ per ogni i e ogni j, vi è una matrice unitaria U tale che $U^{-1}A_iU$ sia diagonale per ogni i. La dimostrazione, per induzione su h, è simile a quella del lemma (6.2). Il caso $h = 1$ è il teorema (6.1). Se $h > 1$, siano $\lambda_1, \ldots, \lambda_k$ gli autovalori di f_1, elencati senza ripetizioni. Segue da (6.1) che i relativi autospazi V_{λ_i} sono ortogonali fra loro. Inoltre segue dalla dimostrazione del lemma (6.2) che $f_j(V_{\lambda_i}) \subset V_{\lambda_i}$ per ogni j e ogni i. Quindi, per ogni i, $f_2, \ldots, f_h$ inducono applicazioni lineari normali di V_{λ_i} in sè. Induttivamente, si può supporre di sapere che ogni V_{λ_i} ha una base ortonormale $v_{i,1}, \ldots, v_{i,n_i}$ ogni cui elemento è simultaneamente autovettore di $f_2, \ldots, f_h$. Una base ortonormale di V con le caratteristiche cercate è allora

$$v_{1,1}, \ldots, v_{1,n_1}, v_{2,1}, \ldots, v_{2,n_2}, \ldots, v_{k,1}, \ldots, v_{k,n_k}\,.$$

Esercizio (6.6). Mostrare che, date matrici complesse quadrate $A_1, \ldots, A_h$, diagonalizzabili ma non necessariamente normali, che commutano, vi è una matrice complessa invertibile B tale che $B^{-1}A_iB$ sia diagonale per ogni i.