Matrici complesse e prodotti Hermitiani

28 ottobre 2009

Richiami sui numeri complessi

In questo paragrafo richiamiamo alcune proprietà del campo dei numeri complessi $\mathbb{C}$.

1. Ogni numero complesso z si può scrivere in modo unico come $z = z_0 + i\, z_1$, ove $z_0, z_1 \in \mathbb{R}$ e $i = \sqrt{-1}$; l'elemento $z_0 = \Re z$ è detto *parte reale* di z, mentre $z_1 = \Im z$ viene chiamato *parte immaginaria*.

2. Il coniugio $\bar{\cdot} : \mathbb{C} \to \mathbb{C}$ è la funzione che associa a $z = z_0 + i\, z_1$ il numero complesso
$$\bar{z} = z_0 - i\, z_1.$$

3. Il coniugio è un automorfismo del campo $\mathbb{C}$: in particolare, per ogni $a, b, c \in \mathbb{C}$:
$$\overline{a + bc} = \bar{a} + \bar{b}\bar{c}.$$

4. L'insieme degli elementi $x \in \mathbb{C}$ per cui $\bar{x} = x$ coincide con $\mathbb{R}$.

5. Per ogni $z = z_0 + i\, z_1 \in \mathbb{C}$, il numero $z\bar{z} = z_0^2 + z_1^2$ è reale e positivo. Si definisce la *norma* di z come
$$|z| = \sqrt{z\bar{z}}.$$

6. La norma di un numero reale coincide col suo valore assoluto.

7. In generale, $|z + w| \leq |z| + |w|$.

8. La *traccia* di $z \in \mathbb{C}$ è il numero reale
$$z + \bar{z} = 2z_0.$$

9. In generale
$$\Re z = \frac{1}{2}(z + \bar{z}); \qquad \Im z = \frac{i}{2}(\bar{z} - z).$$

10. Ogni numero complesso può rappresentarsi in *forma polare* come
$$z = a(\cos\theta + i\sin\theta) = a\, e^{i\theta},$$
ove $a, \theta \in \mathbb{R}$, $a \geq 0$ mentre $-\pi < \theta \leq \pi$. Se $z \neq 0$ tale rappresentazione è unica. Il numero $a = |z|$ il *modulo* di z, mentre θ è detto *argomento*.

11. Il campo complesso è algebricamente chiuso. In particolare, ogni polinomio $p(x) \in \mathbb{C}[x]$ di grado almeno 1 a coefficienti in $\mathbb{C}$ ammette sempre almeno una soluzione $\hat{x} \in \mathbb{C}$.

12. Conseguentemente, un polinomio di grado n in $\mathbb{C}$ ammette sempre esattamente n soluzioni complesse, contate con la debita molteplicità.

13. Applicando il teorema di Ruffini si ha che per ogni $p(x) \in \mathbb{C}[x]$ di grado n esistono $\lambda_1, \dots, \lambda_n$ (non necessariamente tutti distinti) tali che
$$p(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n).$$
In altre parole, si dice che $p(x)$ *si spezza* in fattori lineari (i.e. di primo grado) su $\mathbb{C}$.

Prodotti Hermitiani

Vogliamo definire un prodotto fra vettori di uno spazio vettoriale V sul campo complesso $\mathbb{C}$ che subordini una norma su V. In particolare, si desidera $\forall \mathbf{x} \in V$,
$$0 \leq \langle \mathbf{x}, \mathbf{x} \rangle \in \mathbb{R}.$$

È chiaro che $\langle \cdot, \cdot \rangle$ non potrà essere una forma bilineare ψ, in quanto, anche se per qualche vettore $\mathbf{x} \in V$, si ha $\psi(\mathbf{x}, \mathbf{x}) \in \mathbb{R}$, comunque
$$\psi((1+i)\mathbf{x}, (1+i)\mathbf{x}) = 2i\psi(\mathbf{x}, \mathbf{x}) \notin \mathbb{R}$$

Definizione 1. Sia V uno spazio vettoriale sul campo $\mathbb{C}$. Si dice *forma sesquilineare* su V ogni applicazione $\phi : V \times V \to \mathbb{C}$ tale che, per ogni $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ e $\lambda \in \mathbb{C}$:

1. ϕ è lineare nella prima componente, cioè
$$\phi(\mathbf{x} + \lambda\mathbf{y}, \mathbf{z}) = \phi(\mathbf{x}, \mathbf{z}) + \lambda\phi(\mathbf{y}, \mathbf{z});$$

2. ϕ è antilineare nella seconda componente, cioè

$$\phi(\mathbf{z}, \mathbf{x} + \lambda\mathbf{y}) = \phi(\mathbf{z}, \mathbf{x}) + \overline{\lambda}\phi(\mathbf{z}, \mathbf{y}).$$

Se vale inoltre la condizione

$$\phi(\mathbf{x}, \mathbf{y}) = \overline{\phi(\mathbf{y}, \mathbf{x})},$$

allora ϕ è detta *simmetrica*.

Data una forma sesquilineare simmetrica ϕ abbiamo

$$\phi(\mathbf{x}, \mathbf{x}) = \overline{\phi(\mathbf{x}, \mathbf{x})},$$

da cui segue $\phi(\mathbf{x}, \mathbf{x}) \in \mathbb{R}$.

Definizione 2. Sia $M \in \mathrm{Mat}_n(\mathbb{C})$. Indichiamo con il simbolo $M^\star$ la matrice *aggiunta Hermitiana* di M, ottenuta coniugando la trasposta di M. In simboli

$$M^\star = \overline{M^T} = \overline{M}^T.$$

Definizione 3. Una matrice $H \in \mathrm{Mat}_n(\mathbb{C})$ è detta *Hermitiana* se $H^* = H$.

Teorema 1. *Una forma* $\phi : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ *è sesquilineare simmetrica se, e solamente se, esiste una matrice Hermitiana* H *tale che*

$$\phi(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T H \overline{\mathbf{y}}.$$

Dimostrazione. Innanzi tutto verifichiamo che, data una matrice H con $H^\star = H$, l'applicazione $\theta_H(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T H \overline{\mathbf{y}}$ è sesquilineare simmetrica. Infatti, per ogni $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{C}^n$ e $\lambda \in \mathbb{C}$ si ha

1. $\theta(\mathbf{x} + \lambda\mathbf{z}, \mathbf{y}) = (\mathbf{x} + \lambda\mathbf{z})^T H \overline{\mathbf{y}} = \mathbf{x}^T H \overline{\mathbf{y}} + \lambda \mathbf{z}^T H \overline{\mathbf{y}} = \theta(\mathbf{x}, \mathbf{y}) + \lambda\theta(\mathbf{z}, \mathbf{y})$.
2. $\overline{\theta(\mathbf{x}, \mathbf{y})} = \overline{\mathbf{x}^T H \overline{\mathbf{y}}} = \mathbf{x}^\star \overline{H} \mathbf{y} = \mathbf{y}^T H^\star \overline{\mathbf{x}} = \overline{\theta(\mathbf{y}, \mathbf{x})}$.

Viceversa, sia ϕ una forma sesquilineare simmetrica e fissiamo una base di $\mathbb{C}^n$

$$\mathfrak{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}.$$

Consideriamo la matrice $H = (h_{ij})$ le cui entrate sono

$$h_{ij} = \phi(\mathbf{b}_i, \mathbf{b}_j).$$

Per la simmetria di ϕ, abbiamo che $h_{ij} = \overline{h_{ji}}$, cioè $H^\star = H$. Inoltre, scrivendo i vettori $\mathbf{x}$ e $\mathbf{y}$ in componenti rispetto $\mathfrak{B}$,

$$\phi(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \overline{y_j} \phi(\mathbf{b}_i, \mathbf{b}_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \overline{y_j} h_{ij} = \mathbf{x}^T H \overline{\mathbf{y}}.$$

La tesi segue. □

Definizione 4. Sia $\phi : V \times V \to \mathbb{C}$ una forma sesquilineare simmetrica. L'applicazione $h : V \to \mathbb{R}$ descritta da

$$h(\mathbf{x}) = \phi(\mathbf{x}, \mathbf{x})$$

è detta *forma Hermitiana* associata a ϕ.

Definizione 5. Una forma Hermitiana h su V è detta *definita positiva* se, per ogni $\mathbf{x} \in V$ si ha $h(\mathbf{x}) = \phi(\mathbf{x}, \mathbf{x}) \geq 0$ e $h(\mathbf{x}) = 0$ se, e solamente se $\mathbf{x} = \mathbf{0}$. In tale caso la forma sesquilineare ϕ che induce h sarà detta *prodotto Hermitiano*. Una matrice Hermitiana H è *definita positiva* se, e solamente se, la forma Hermitiana

$$h(\mathbf{x}) = \mathbf{x}^T H \overline{\mathbf{x}}$$

ad essa associata è definita positiva.

Osserviamo che poiché $h(\mathbf{x}) \in \mathbb{R}$ per ogni $\mathbf{x} \in V$ e $H^T = \overline{H}$,

$$h(\mathbf{x}) = \overline{h(\mathbf{x})} = \mathbf{x}^\star \overline{H} \mathbf{x} = \mathbf{x}^\star H^T \mathbf{x}.$$

Teorema 2. *Sia h una forma Hermitiana definita positiva associata ad una forma sesquilineare ϕ. Allora, $||\mathbf{x}|| = \sqrt{h(\mathbf{x})}$ è una norma su V.*

Dimostrazione. Rammentiamo che il modulo di un numero complesso λ è il numero reale $|\lambda| = \sqrt{\lambda\overline{\lambda}}$. Innanzi tutto osserviamo che

$$||\lambda\mathbf{x}||^2 = \lambda\overline{\lambda}||\mathbf{x}||^2 = |\lambda|^2||\mathbf{x}||^2.$$

Dobbiamo verificare ora la disuguaglianza triangolare, cioè (elevando tutto al quadrato)

$$h(\mathbf{x} + \mathbf{y}) \leq h(\mathbf{x}) + h(\mathbf{y}).$$

Dimostriamo dapprima che la disuguaglianza di Schwartz vale anche nel caso di forme sesquilineari. Poiché ϕ è definita positiva abbiamo, per ogni $\alpha \in \mathbb{C}$

$$\phi(\mathbf{x} + \alpha\mathbf{y}, \mathbf{x} + \alpha\mathbf{y}) = \phi(\mathbf{x}, \mathbf{x}) + |\alpha|^2\phi(\mathbf{y}, \mathbf{y}) + \overline{\alpha}\phi(\mathbf{x}, \mathbf{y}) + \alpha\overline{\phi(\mathbf{x}, \mathbf{y})} \geq 0$$

In particolare, ponendo $\alpha = -\frac{\phi(\mathbf{x},\mathbf{y})}{\phi(\mathbf{y},\mathbf{y})}$ otteniamo

$$\phi(\mathbf{x}, \mathbf{x}) + \frac{|\phi(\mathbf{x}, \mathbf{y})|^2}{\phi(\mathbf{y}, \mathbf{y})} - 2\frac{|\phi(\mathbf{x}, \mathbf{y})|^2}{\phi(\mathbf{y}, \mathbf{y})} \geq 0$$

da cui si deduce
$$|\phi(\mathbf{x},\mathbf{y})|^2 \leq \phi(\mathbf{x},\mathbf{x})\phi(\mathbf{y},\mathbf{y}),$$
cioè
$$|\phi(\mathbf{x},\mathbf{y})| \leq \|\mathbf{x}\|\,\|\mathbf{y}\|.$$
D'altro canto,
$$\begin{aligned}
\|\mathbf{x}+\mathbf{y}\|^2 &= h(\mathbf{x}+\mathbf{y}) = \phi(\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y}) = \\
&h(\mathbf{x}) + \phi(\mathbf{x},\mathbf{y}) + \overline{\phi(\mathbf{x},\mathbf{y})} + h(\mathbf{y}) \leq |h(\mathbf{x}) + \phi(\mathbf{x},\mathbf{y}) + \overline{\phi(\mathbf{x},\mathbf{y})} + h(\mathbf{y})| \leq \\
&h(\mathbf{x}) + |\phi(\mathbf{x},\mathbf{y})| + |\overline{\phi(\mathbf{x},\mathbf{y})}| + h(\mathbf{y}) = h(\mathbf{x}) + 2|\phi(\mathbf{x},\mathbf{y})| + h(\mathbf{y}) \leq \\
&\|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\,\|\mathbf{y}\| + \|\mathbf{y}\|^2.
\end{aligned}$$
Quindi,
$$\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$
e la dimostrazione è conclusa. □

Sia $\mathbf{x}$ un vettore colonna di $V = \mathbb{C}^n$. Il prodotto Hermitiano standard su V è definito come
$$\langle \mathbf{x},\mathbf{y}\rangle = \mathbf{x}^T\overline{\mathbf{y}}.$$
Pertanto, la forma Hermitiana standard su V risulta
$$h(\mathbf{x}) = \mathbf{x}^T\overline{\mathbf{x}}.$$
Tenuto conto che $h(\mathbf{x}) \in \mathbb{R}$ si vede che
$$h(\mathbf{x}) = \overline{h(\mathbf{x})} = \overline{\mathbf{x}^T\overline{\mathbf{x}}} = \mathbf{x}^\star\mathbf{x}.$$
Pertanto, tale forma associa al vettore $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ il numero reale
$$h(\mathbf{x}) = \sum_i \overline{x_i}x_i = \sum_i |x_i|^2;$$
ed essa risulta definita positiva. La norma indotta da h verrà indicata col simbolo
$$\|\mathbf{x}\|_2 = \sqrt{h(\mathbf{x})}.$$
Osserviamo che, se $\mathbf{x} \in \mathbb{R}^n$, tale norma coincide esattamente con la $\|\cdot\|_2$ precedentemente introdotta. Similmente, ristretto a vettori $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$ un qualsiasi prodotto Hermitiano si comporta come un prodotto scalare.

Diremo che due vettori $\mathbf{x},\mathbf{y} \in \mathbb{C}^n$ sono *ortogonali* se, e solamente se,
$$\langle \mathbf{x},\mathbf{y}\rangle = 0,$$

ove $\langle\cdot,\cdot\rangle$ denota il prodotto Hermitiano standard.

Sia ora $M \in \mathrm{Mat}_{n,m}(\mathbb{C})$ una matrice qualsiasi. Osserviamo che M induce in modo naturale un omomorfismo fra gli spazi vettoriali $W = \mathbb{C}^m$ e $V = \mathbb{C}^n$. Prendiamo $\mathbf{x} \in \mathbb{C}^n$ e $\mathbf{y} \in \mathbb{C}^m$. Inoltre,

$$\langle \mathbf{x}, M\mathbf{y}\rangle = \mathbf{x}^T\overline{M\mathbf{y}} = \mathbf{x}^T\overline{M^T}^T\overline{\mathbf{y}} = \mathbf{x}^T(M^\star)^T\overline{\mathbf{y}} = \langle M^\star\mathbf{x}, \mathbf{y}\rangle\,.$$

Teorema 3 (Teorema di rappresentazione)**.** *Sia V uno spazio vettoriale su $\mathbb{C}$ di dimensione finita, e V^* il suo duale. Per ogni $\varphi \in V^*$ esiste un vettore $\mathbf{y} \in V$ tale che, per ogni $\mathbf{x} \in V$,*

$$\varphi(\mathbf{x}) = \langle \mathbf{x}, \mathbf{y}\rangle\,.$$

Dimostrazione. Per definizione di forma sesquilineare, l'applicazione $\langle\cdot,\mathbf{y}\rangle$ è un elemento del duale V^* per ogni possibile scelta del vettore $\mathbf{y}$. Fissiamo una base ortonormale per V

$$\mathfrak{E} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$$

e siano $\alpha_i = \varphi(\mathbf{e}_i)$ per $i = 1,\ldots n$. Poniamo

$$\mathbf{y} = \sum_{i=1}^{n} \overline{\alpha_i}\mathbf{e}_i.$$

Allora,

$$\langle \mathbf{e}_i, \mathbf{y}\rangle = \sum_{i=1}^{n} \langle \mathbf{e}_i, \overline{\alpha_i}\mathbf{e}_i\rangle = \alpha_i\langle \mathbf{e}_i, \mathbf{e}_i\rangle = \alpha_i.$$

Poiché le due forme lineari φ e $\langle\cdot,\mathbf{y}\rangle$ coincidono su di una base di V, esse sono la stessa. □

Osserviamo che, esattamente come nel caso del prodotto scalare per spazi vettoriali reali, si che, per ogni $\mathbf{y} \in \mathbb{C}^n$, l'insieme

$$\mathbf{y}^\perp = \{\mathbf{x} \in \mathbb{C}^n : \langle \mathbf{x}, \mathbf{y}\rangle = 0\}$$

è un sottospazio vettoriale di $\mathbb{C}^n$ avente dimensione n se $\mathbf{y} = \mathbf{0}$ e $n-1$ se $\mathbf{y} \neq \mathbf{0}$, in quanto nucleo di una forma lineare $\varphi \in V^*$.

Matrici normali e teorema spettrale

Definizione 6. Una matrice $M \in \mathrm{Mat}_n(\mathbb{C})$ si dice

Hermitiana	se	$M^\star = M$
Unitaria	se	$MM^\star = I$
Normale	se	$M^\star M = MM^\star$

Osserviamo che matrici Hermitiane e Unitarie sono sempre normali. Per matrici reali $R \in \mathrm{Mat}_n(\mathbb{R})$ vale la condizione $R^\star = R^T$, pertanto vigono le le seguenti implicazioni:

$$\begin{array}{lcl} R \text{ Ortogonale} & \Leftrightarrow & R \text{ Unitaria} \\ R \text{ Simmetrica} & \Leftrightarrow & R \text{ Hermitiana} \end{array}$$

Osserviamo che se U è una matrice unitaria e $\mathbf{x} \in \mathbb{C}^n$:

$$\|U\mathbf{x}\|_2^2 = \mathbf{x}^T U^T \overline{U\mathbf{x}} = \overline{\overline{\mathbf{x}^T U^T} U\mathbf{x}} = \mathbf{x}^\star U^\star U\mathbf{x} = \mathbf{x}^\star \mathbf{x} = \mathbf{x}^T \overline{\mathbf{x}} = \|\mathbf{x}\|_2^2,$$

cioè le matrici unitarie corrispondono a trasformazioni lineari dello spazio vettoriale che conservano la $\|\cdot\|_2$ dei vettori. Vedremo in seguito, nel Teorema 14, che vale anche il viceversa.

Dimostriamo ora un risultato che riguarda tutte le matrici a coefficienti in $\mathbb{C}$.

Teorema 4 (Schur)**.** *Sia* $M \in \mathrm{Mat}_n(\mathbb{C})$. *Allora, esiste una matrice unitaria* U *tale che* $U^\star M U$ *sia triangolare superiore.*

Dimostrazione. Procediamo per induzione sull'ordine n della matrice.

- Per $n = 1$, il teorema è banale.

- Supponiamo che il teorema valga per tutte le matrici di ordine $(n-1) \times (n-1)$. Poiché $\mathbb{C}$ è algebricamente chiuso, la matrice M ammette almeno un autovalore λ. Sia $\mathbf{x}$ un autovettore associato a λ. È possibile supporre senza perdere in generalità $\|\mathbf{x}\|_2 = 1$. Costruiamo una base ortonormale di $\mathbb{C}^n$ che abbia come primo vettore proprio $\mathbf{x}$ e consideriamo la matrice di cambiamento di base V. Allora, V è unitaria e, posto $M_1 = V^\star M V$ si ha

  $$VM_1\mathbf{e_1} = MV\mathbf{e_1} = M\mathbf{x} = \lambda\mathbf{x} = \lambda V\mathbf{e_1},$$

  dal che, tenuto conto del fatto che V è invertibile, si deduce $M_1\mathbf{e_1} = \lambda\mathbf{e_1}$, cioè $\mathbf{e_1}$ è un autovettore per M_1. Ne segue che

  $$M_1 = \begin{pmatrix} \lambda & \dots \\ 0 & M_2 \end{pmatrix}$$

  con M_2 matrice di ordine $(n-1) \times (n-1)$. Per ipotesi induttiva esiste una matrice unitaria $\widetilde{W}$ tale che $\widetilde{W}^\star M_2 \widetilde{W}$ sia triangolare superiore. Poniamo

  $$W = \begin{pmatrix} 1 & 0 \\ 0 & \widetilde{W} \end{pmatrix}.$$

Chiaramente W è unitaria e trasforma in modo unitario M_1 in una matrice triangolare superiore T. Pertanto, posto $U = VW$ si ha

$$U^\star M U = W^\star(V^\star M V)W = W^\star M_1 W = T.$$

La tesi segue.

□

Per le matrici Hermitiane e ortogonali si può stabilire di più.

Teorema 5. *Gli autovalori di matrici Hermitiane sono tutti reali.*

Dimostrazione. Sia $M \in \mathrm{Mat}_n(\mathbb{C})$ una matrice Hermitiana e supponiamo che λ sia un suo autovalore con autovettore $\mathbf{x}$. Allora, tenuto conto che $M^\star = M$,

$$M\mathbf{x} = \lambda, \qquad \mathbf{x}^\star M = \overline{\lambda}.$$

Pertanto,

$$\lambda \mathbf{x}^\star \mathbf{x} = \mathbf{x}^\star(M\mathbf{x}) = (\mathbf{x}^\star M)\mathbf{x} = \overline{\lambda}\mathbf{x}^\star\mathbf{x},$$

da cui

$$(\lambda - \overline{\lambda})\mathbf{x}^\star\mathbf{x} = 0.$$

Osserviamo che

$$\mathbf{x}^\star\mathbf{x} = \langle \mathbf{x}, \mathbf{x} \rangle = \sum_{i=1}^{n} |x_i|^2 > 0.$$

Pertanto, $\overline{\lambda} = \lambda$ e dunque $\lambda \in \mathbb{R}$. □

Premettiamo al teorema fondamentale di questo paragrafo alcuni lemmi.

Lemma 6. *Sia* $M \in \mathrm{Mat}_n(\mathbb{C})$ *una matrice normale. Allora, per ogni* $\mathbf{x} \in \mathbb{C}^n$ *si ha*

$$||M\mathbf{x}||_2 = ||M^\star\mathbf{x}||_2.$$

Dimostrazione. Tenuto conto che la norma è un numero reale positivo,

$$||M\mathbf{x}||_2^2 = \mathbf{x}^T M^T \overline{M\mathbf{x}} = \overline{\mathbf{x}^T M^T \overline{M\mathbf{x}}} = \mathbf{x}^\star M^\star M \mathbf{x} =$$
$$\mathbf{x}^\star M M^\star \mathbf{x} = \overline{\mathbf{x}^\star M M^\star \mathbf{x}} = (M^\star\mathbf{x})^T \overline{M^\star\mathbf{x}} = ||M^\star\mathbf{x}||_2^2.$$

□

Lemma 7. *Sia* $M \in \mathrm{Mat}_n(\mathbb{C})$ *una matrice normale e supponiamo che* $\mathbf{x}$ *sia un suo autovettore di autovalore* λ. *Allora,* $\mathbf{x}$ *è autovettore di* $M^\star$ *di autovalore* $\overline{\lambda}$.

Dimostrazione. Verifichiamo dapprima che $M - \lambda I$ è una matrice normale. Infatti

$$(M - \lambda I)^\star(M - \lambda I) = (M^\star - \bar{\lambda} I)(M - \lambda I) = M^\star M - \lambda M^\star - \bar{\lambda} M + \lambda\bar{\lambda} I = \\ MM^\star - \lambda M^\star - \bar{\lambda} M + \lambda\bar{\lambda} I = (M - \lambda I)(M - \lambda I)^\star.$$

In particolare, se $\mathbf{x}$ è autovettore di M si ha

$$0 = \|(M - \lambda I)\mathbf{x}\|_2 = \|(M^\star - \bar{\lambda} I)\mathbf{x}\|_2$$

da cui si deduce

$$M^\star \mathbf{x} = \bar{\lambda}\mathbf{x},$$

che è la tesi. □

Teorema 8 (Teorema Spettrale)**.** *Ogni matrice normale* $M \in \mathrm{Mat}_n(\mathbb{C})$ *è diagonalizzabile mediante matrici unitarie.*

Dimostrazione. Ragioniamo per induzione sull'ordine n di M.

- Se $n = 1$ non vi è nulla da dimostrare e la matrice $M = (m_{11})$ è già diagonale.
- Supponiamo che ogni matrice normale $(n-1) \times (n-1)$ sia diagonalizzabile mediante matrici unitarie. Per il Teorema 4 esiste sicuramente una matrice unitaria V tale che $M_1 = V^\star M V$ sia triangolare superiore. In particolare, vi è un λ tale che $M_1\mathbf{e_1} = \lambda\mathbf{e_1}$. D'altro canto,

  $$M_1^\star M_1 = (V^\star M^\star V)(V^\star M V) = V^\star(M^\star M)V = \\ V^\star(MM^\star)V = (V^\star M V)(V^\star)M^\star V) = M_1 M_1^\star,$$

  per cui anche M_1 è normale. Per il Lemma 7, il vettore $\mathbf{e_1}$ è anche autovettore di $M_1^\star$ e $M_1^\star\mathbf{e_1} = \bar{\lambda}\mathbf{e_1}$. Ne segue che M_1 è diagonale a blocchi della forma

  $$M_1 = \begin{pmatrix} \lambda & 0 \\ 0 & M_2 \end{pmatrix}.$$

  Tenuto conto del fatto che

  $$\begin{pmatrix} \bar{\lambda}\lambda & \\ 0 & M_2^\star M_2 \end{pmatrix} = M_1^\star M_1 = M_1 M_1^\star = \begin{pmatrix} \lambda\bar{\lambda} & \\ & M_2 M_2^\star \end{pmatrix}$$

  si vede che M_2 è una matrice normale $(n-1) \times (n-1)$.
- La tesi ora segue dall'ipotesi induttiva.

$\square$

In particolare, assegnata una matrice normale M è sempre possibile trovare una base ortonormale di $\mathbb{C}^n$ formata da autovettori della stessa.

Nel caso di matrici reali simmetriche si può dimostrare che è possibile costruire una base di autovettori reali.

Teorema 9. *Sia* $M \in \mathrm{Mat}_n(\mathbb{R})$ *una matrice simmetrica. Allora, esiste una matrice ortogonale* $O \in \mathrm{Mat}_n(\mathbb{R})$ *tale che* $D = OMO^{-1}$ *è diagonale.*

Dimostrazione. Innanzi tutto osserviamo che, per il Teorema 5, tutti gli autovalori di M sono reali. Procediamo per induzione sull'ordine n di M:

- Se $n = 1$, M è già diagonale.
- Supponiamo che ogni matrice reale simmetrica di ordine $(n-1)$ sia diagonalizzabile mediante trasformazioni lineari. Mostriamo che allora anche una matrice di ordine n lo è. Sia λ un autovalore di M. Allora, esiste un vettore $\mathbf{x}$ tale che

  $$M\mathbf{x} = \lambda\mathbf{x}.$$

  Poiché $\overline{M} = M$, abbiamo anche

  $$\overline{M\mathbf{x}} = \lambda\overline{\mathbf{x}}.$$

  Cio significa che $\mathbf{x}$ e $\overline{\mathbf{x}}$ sono entrambi autovettori associati al medesimo autovalore. Distinguiamo due casi:

  1. se $\overline{\mathbf{x}} = -\mathbf{x}$, allora il vettore $\mathbf{y} = i\mathbf{x}$ è un autovettore reale di autovalore λ;
  2. se $\overline{\mathbf{x}} \neq -\mathbf{x}$, allora il vettore $\mathbf{y} = \mathbf{x} + \overline{\mathbf{x}}$ è reale, diverso da $\mathbf{0}$ e

     $$M\mathbf{y} = M(\mathbf{x} + \overline{\mathbf{x}}) = \lambda\mathbf{x} + \lambda\overline{\mathbf{x}} = \lambda\mathbf{y}.$$

     Pertanto, anche in questo caso abbiamo trovato un autovettore reale.

  Consideriamo ora il vettore $\mathbf{y}' = \frac{1}{\|\mathbf{y}\|_2}\mathbf{y}$ e sia O_1 la matrice ortogonale che manda $\mathbf{y}'$ in $\mathbf{e_1}$. Allora,

  $$O_1 M O_1^{-1}\mathbf{e_1} = O_1 M\mathbf{y}' = \lambda O_1\mathbf{y}' = \lambda\mathbf{e_1}$$

Cioè $\mathbf{e_1}$ è autovettore di $O_1MO_1^{-1}$. Esattamente come nella dimostrazione del Teorema 8, vediamo che $\mathbf{e_1}$ deve essere anche autovettore di $O_1M^TO_1^{-1}$. Pertanto,

$$O_M^TO_1^{-1} = \begin{pmatrix} \lambda & 0 \\ 0 & M' \end{pmatrix},$$

ove M' è una matrice reale e simmetrica di ordine $n-1$

- La tesi segue per induzione.

□

Una conseguenza del Teorema 9 è la seguente caratterizzazione.

Teorema 10 (Teorema dell'asse principale)**.** *Sia* $M \in \mathrm{Mat}_n(\mathbb{R})$ *una matrice reale. Allora, le seguenti tre condizioni sono equivalenti:*

(1) $\mathbb{R}^n$ *ammette una base ortonormale di autovettori per* M*;*

(2) M *è diagonalizzabile mediante una matrice ortogonale;*

(3) M *è simmetrica.*

Dimostrazione. Le condizioni (1) e (2) sono equivalenti, per costruzione di matrice diagonalizzante. Per il Teorema 9, la condizione (3) implica le prime due. D'altro canto, se M è diagonalizzabile mediante una matrice ortogonale O, allora

$$D^T = (OMO^{-1})^T = (OM^TO^{-1}),$$

da cui

$$M = O^{-1}(OMO^{-1})O = O^{-1}(OM^TO^{-1})O = M^T,$$

cioè M è simmetrica. □

Esiste un metodo per calcolare gli autovalori di matrici Hermitiane (e quindi anche reali simmetriche) in termini del valore minimo che le relative forme Hermitiane assumono su opportuni insiemi.

Sia H una matrice Hermitiana $n \times n$ e indichiamo, come al solito, con $h(\mathbf{x}) = \mathbf{x}^\star H\mathbf{x}$ la forma Hermitiana ad essa associata.

Teorema 11 (Quozienti di Rayleigh)**.** *Siano* $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ *gli* n *autovalori di* H*, posti in ordine crescente e contati con le debite molteplicità. Allora,*

$$\lambda_k = \min_{\dim F = k} \max_{\mathbf{x} \in F \setminus \{\mathbf{0}\}} \frac{h(\mathbf{x})}{||\mathbf{x}||_2^2} = \min_{\dim F = k} \max_{\substack{\mathbf{x} \in F \\ ||\mathbf{x}||_2 = 1}} h(\mathbf{x}).$$

Dimostrazione. Per ogni sottospazio $F \leq \mathbb{C}^n$ poniamo

$$R(F) = \max_{\mathbf{x} \in F \setminus \{\mathbf{0}\}} \frac{h(\mathbf{x})}{||\mathbf{x}||_2^2} = \max\{h(\mathbf{x}) : \mathbf{x} \in F, ||\mathbf{x}||_2 = 1\}$$

e sia

$$\mathfrak{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$$

una base di autovettori, con $H\mathbf{b}_i = \lambda_i$. Chiaramente, per ogni vettore $\mathbf{x} \in \mathbb{C}^n$ possiamo scrivere $\mathbf{x} = \sum_i x_i \mathbf{b}_i$ e si ottiene

$$h(\mathbf{x}) = \sum_i \lambda_i |x_i|^2.$$

Consideriamo ora un generico sottospazio F di $\mathbb{C}^n$ con $\dim F = k$ e poniamo $\mathfrak{B}_k = \{\mathbf{b}_k, \mathbf{b}_{k+1}, \ldots, \mathbf{b}_n\}$. Poiché la dimensione del sottospazio vettoriale B_k generato da $\mathfrak{B}_k$ è proprio $n-k+1$, si ha che esiste almeno un vettore $\mathbf{x} \in F \cap B_k$ con $\mathbf{x} \neq \mathbf{0}$. In particolare per $i < k$ si ha $x_i = 0$.

$$h(\mathbf{x}) = \sum_{i=0}^{n} \lambda_i |x_i|^2 = \sum_{i=k}^{n} \lambda_i |x_i|^2 \geq \lambda_k \sum_{i=k}^{n} |x_i|^2 = \lambda_k ||\mathbf{x}||_2^2.$$

Pertanto $R(F) \geq \lambda_k$ e quindi

$$\min_{\dim F = k} R(F) \geq \lambda_k.$$

D'altro canto, se consideriamo lo spazio vettoriale F' generato da $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k\}$ vediamo che:

1. esso ha dimensione k;

2. per ogni $\mathbf{x} \in F'$ si ha

   $$h(\mathbf{x}) = \sum_{i=1}^{k} |x_i|^2 \lambda_i \leq \lambda_k \sum_{i=1}^{k} |x_i|^2 = \lambda_k ||\mathbf{x}||_2^2;$$

   pertanto $R(F') \leq \lambda_k$. D'altro canto $\mathbf{b}_k \in F'$ implica proprio $R(F') = \lambda_k$.

Ne segue che

$$\min_{\dim F = k} R(F) = \lambda_k$$

che è la tesi. □

Teorema 12. *Una matrice Hermitiana H è definita positiva se, e solamente se, tutti i suoi autovalori sono strettamente maggiori di 0.*

Dimostrazione. Gli autovalori di H coincidono con quelli di $H^T = \overline{H}$. Supponiamo essi siano tutti strettamente maggiori di 0. Poiché $\overline{H}$ è Hermitiana, esiste una matrice U tale che $U^\star DU = \overline{H}$, con D matrice diagonale con $d_{ii} = \lambda_i > 0$ Sia

$$\mathfrak{U} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$$

la base di $\mathbb{C}^n$ data dalle colonne di U. Per ogni $\mathbf{x} \in \mathbb{C}^n$ si può scrivere

$$\mathbf{x} = x_1\mathbf{u}_1 + \ldots + x_n\mathbf{u}_n$$

e quindi

$$h(\mathbf{x}) = \mathbf{x}^\star\overline{H}\mathbf{x} = \sum_{i,j=1}^{n} \overline{x_j}\mathbf{u}_j^\star\overline{H}x_i\mathbf{u}_i = \sum_{i,j=1}^{n} |x_j|^2\lambda_i\mathbf{u}_j^\star\mathbf{u}_i.$$

Tenuto conto che la base $\mathfrak{U}$ è ortonormale, si deduce che

$$\sum_{i,j=1}^{n} |x_j|^2\lambda_i\mathbf{u}_j^\star\mathbf{u}_i = \sum_{i=1}^{n} |x_i|^2\lambda_i \geq 0,$$

e $\mathbf{x}^\star\overline{H}\mathbf{x} = 0$ se, e solamente se, $\mathbf{x} = \mathbf{0}$. Pertanto, H è definita positiva.

Viceversa supponiamo che H sia definita positiva e sia $\mathfrak{U} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ una base ortonormale di autovettori, con la condizione $H\mathbf{u}_i = \lambda_i\mathbf{u}_i$. Allora,

$$h(\mathbf{u}_i) = \mathbf{u}_i^T H\overline{\mathbf{u}_i} = \overline{\lambda_i}\mathbf{u}_i^T\overline{\mathbf{u}_i} = \lambda_i > 0.$$

Pertanto tutti gli autovalori di H sono tutti positivi. □

Corollario 13. *Una matrice Hermitiana* $H \in GL_n(\mathbb{C})$ *è definita positiva se, e solamente se, per ogni* $\mathbf{x} \in \mathbb{C}^n$ *con* $\mathbf{x} \neq \mathbf{0}$ *si ha*

$$\mathbf{x}^\star H\mathbf{x} > 0.$$

Dimostrazione. Osserviamo che H è definita positiva se, e solamente se, $H^T = \overline{H}$ lo è, ovvero

$$\mathbf{x}^T H^T\overline{\mathbf{x}} = \mathbf{x}^\star H\mathbf{x} > 0,$$

per ogni $\mathbf{x} \neq \mathbf{0}$. □

Come conseguenza diretta del Teorema 12 osserviamo che se due matrici Hermitiane H_1 e H_2 sono simili, allora esse hanno i medesimi autovalori e sono simili alla stessa matrice diagonale D. Pertanto, una matrice Hermitiana simile ad una matrice Hermitiana definita positiva è a sua volta definita positiva.

Isometrie

Sia $V = \mathbb{C}^n$ uno spazio vettoriale su $\mathbb{C}$.

Definizione 7. Una trasformazione lineare $M : V \to V$ è detta *isometria* se, per ogni $\mathbf{x} \in V$:

$$||M\mathbf{v}||_2 = ||\mathbf{v}||_2.$$

Caratterizziamo ora tutte le isometrie di V.

Teorema 14. *Le isometrie di $V = \mathbb{C}^n$ sono tutti e soli gli endomorfismi indotti da matrici unitarie; similmente le isometrie di $W = \mathbb{R}^n$ sono tutti e soli gli endomorfismi indotti da matrici ortogonali.*

Dimostrazione. Sia $\mathbf{x} \in V$ e U una matrice unitaria. Allora,

$$||U\mathbf{x}||_2^2 = \langle U\mathbf{x}, U\mathbf{x}\rangle = \mathbf{x}^\star U^\star U\mathbf{x} = \mathbf{x}^\star \mathbf{x} = \langle \mathbf{x}, \mathbf{x}\rangle = ||\mathbf{x}||_2^2$$

Pertanto, ogni matrice unitaria induce una isometria.

Viceversa, supponiamo che M induca una isometria su V; allora

$$||M\mathbf{x}||_2^2 = \langle M\mathbf{x}, M\mathbf{x}\rangle = \langle M^\star M\mathbf{x}, \mathbf{x}\rangle = \langle \mathbf{x}, \mathbf{x}\rangle = ||\mathbf{x}||_2^2.$$

In particolare, usando la linearità del prodotto Hermitiano nella prima componente, abbiamo

$$\langle (M^\star M - I)\mathbf{x}, \mathbf{x}\rangle = 0.$$

Ora, la matrice $S = M^\star M - I$ è normale in quanto Hermitiana (poiché $S^\star = S$); pertanto, per il Teorema 9, essa è diagonalizzabile. Sia ora λ un qualsiasi autovalore di $S^\star$ e $\mathbf{y} \neq \mathbf{0}$ un suo autovettore. Si ottiene

$$0 = \langle S\mathbf{y}, \mathbf{y}\rangle = \langle \lambda\mathbf{y}, \mathbf{y}\rangle = \lambda \langle \mathbf{y}, \mathbf{y}\rangle = \lambda||\mathbf{y}||_2^2.$$

Poiché $||\mathbf{y}|| \neq 0$, si deduce $\lambda = 0$ per tutti gli autovalori di S, da cui $S = 0$ e $M^\star M = I$, ovvero M è unitaria.

Il caso reale si dimostra esattamente al medesimo modo, tenendo conto del fatto che una matrice reale unitaria è ortogonale. $\square$

Una conseguenza del Teorema 14 è la seguente caratterizzazione degli autovalori di matrici unitarie.

Teorema 15. *Gli autovalori di una matrice unitaria U o di una matrice reale ortogonale sono numeri complessi di modulo 1.*

Dimostrazione. Sia λ un autovalore di U con autovettore $\mathbf{x}$. Allora,

$$||U\mathbf{x}||_2 = |\lambda|^2||\mathbf{x}||_2^2$$

Per il Teorema 14, U è una isometria; pertanto $|\lambda| = 1$. $\square$

Decomposizione Polare di una matrice

Come richiamato nel primo paragrafo, ogni numero complesso z può essere identificato dalla coppia $(a, \theta) \in \mathbb{R}^+ \times]-\pi, \pi]$ mediante la corrispondenza

$$z = |a| (\cos\theta + i\sin(\theta)) .$$

Un risultato analogo vale anche per le matrici invertibili. Premettiamo due lemmi.

Lemma 16. *Sia* $H \in GL_n(\mathbb{C})$ *una matrice Hermitiana definita positiva. Allora, anche* H^{-1} *è Hermitiana definita positiva.*

Dimostrazione. Poiché H è Hermitiana, esiste una matrice unitaria U tale che

$$H = U^\star D U,$$

con D matrice diagonale definita positiva. Chiaramente, anche D^{-1} è diagonale e definita positiva. D'altro canto, posto $M = U^\star D^{-1} U$ abbiamo

$$HN = (U^\star D U)(U^\star D^{-1} U) = U^\star D D^{-1} U = U^\star U = I.$$

Pertanto $N = H^{-1}$ e la tesi segue. □

Lemma 17. *Sia* H *una matrice Hermitiana definita positiva. Allora, esiste una matrice Hermitiana definita positiva* $N = \sqrt{H}$ *tale che* $N^2 = H$. *Tale matrice è detta* radice quadrata *di* H.

Dimostrazione. Per il Teorema 8, esiste una matrice unitaria U tale che

$$H = U^\star D U,$$

ove D è una matrice diagonale, con tutte le entrate $d_1, d_2, \ldots, d_n$ sulla diagonale principale reali e strettamente positive. Sia

$$\sqrt{D} = \begin{pmatrix} \sqrt{d_1} & & & \\ & \sqrt{d_2} & & \\ & & \ddots & \\ & & & \sqrt{d_n} \end{pmatrix}.$$

Chiaramente, $\sqrt{D}^2 = D$. Poniamo $N = U^\star \sqrt{D} U$. Allora,

$$N^2 = (U^\star \sqrt{D} U)((U^\star \sqrt{D} U) = U^\star (\sqrt{D})^2 U = U^\star D U = H.$$

Pertanto, $N = \sqrt{H}$ è la matrice cercata. □

Teorema 18. *Sia* $M \in \mathrm{GL}_n(\mathbb{C})$*. Allora esiste un'unica coppia* (H, Q) *ove*

1. H *è una matrice Hermitiana positiva definita;*
2. Q *è una matrice Unitaria;*

tale che $M = HQ$*. Se* $M \in \mathrm{GL}_n(\mathbb{R})$*, allora* H *risulta essere una matrice simmetrica definita positiva e* Q *una matrice ortogonale.*

Dimostrazione. Innanzi tutto osserviamo che la matrice $MM^\star$ è Hermitiana. Inoltre, se $\mathbf{x} \neq \mathbf{0}$,

$$\mathbf{x}^* MM^\star \mathbf{x} = \langle M^\star \mathbf{x}, M^\star \mathbf{x} \rangle > 0,$$

per cui $MM^\star$ è anche definita positiva.

Usando il Lemma 17, poniamo

$$H = \sqrt{MM^\star}, \qquad Q = H^{-1}M$$

Chiaramente, $M = HQ$. Per costruzione, H è Hermitiana, positiva definita e

$$H^2 = HH^\star = U^\star \sqrt{D} U U^\star \sqrt{D} U = U^\star \sqrt{D}^2 U = U^\star D U = MM^\star$$

D'altro canto,

$$Q^\star Q = M^\star (H^{-1})^\star H^{-1} M = M^\star H^{-2} M = M^\star (MM^\star)^{-1} M = I,$$

per cui U è unitaria e si è ottenuta una decomposizione polare per M.

Supponiamo ora che $M = H'Q'$ sia un'altra decomposizione polare per M. Allora

$$N = H^{-1}H' = Q(Q')^{-1}$$

è ovviamente unitaria; pertanto tutti i suoi autovalori hanno modulo 1. Mostriamo ora che N è anche Hermitiana (rammentiamo che, in generale, il prodotto di due matrici Hermitiane non è Hermitiano), di modo da poter asserire che tutti i suoi autovalori sono *reali*. Per il Lemma 16, H^{-1} è Hermitiana e definita positiva. Pertanto, a norma del Lemma 17 esiste una matrice S Hermitiana definita positiva tale che $S^2 = H^{-1}$. Allora,

$$S^{-1}NS = S^{-1}H^{-1}H'S = S^{-1}S^2H'S = SH'S,$$

per cui N è simile ad $P = SH'S$. D'altro canto, usando l'Hermitianità di H' e S si ha

$$P^\star = S^\star H'^\star S^\star = SH'S = P,$$

per cui P è Hermitiana e, conseguentemente, lo è pure N.

Mostriamo ora che N è definita positiva. Come sopra, ragioniamo su P. Osserviamo che S è invertibile, per cui $\mathbf{x} \neq \mathbf{0}$ implica $S\mathbf{x} \neq \mathbf{0}$, e, chiaramente, per ogni $\mathbf{y} \in \mathbb{C}^n$ con $\mathbf{y} \neq 0$ si ha $\mathbf{y}^T H' \mathbf{y} > 0$. Da ciò si deduce che

$$\mathbf{x}^T P \overline{\mathbf{x}} = \mathbf{x}^T (SH'S) \overline{\mathbf{x}} = \mathbf{x}^T \overline{S}^T H' \overline{\overline{S}\mathbf{x}} = (\overline{S}\mathbf{x})^T H' (\overline{S}\mathbf{x}) = \mathbf{y}^T H' \overline{\mathbf{y}} \geq 0$$

Ne segue che P è definita positiva e dunque anche N lo è. Pertanto, $N = I$. □

Incidentalmente, si noti che dalla dimostrazione dell'unicità della decomposizione polare segue anche che la radice quadrata di una matrice Hermitiana definita positiva è unica.

Definizione 8. Sia $M \in GL_n(\mathbb{C})$ e indichiamo con (H, Q) la sua decomposizione polare. Gli autovalori di H (ovvero le radici quadrate degli autovalori di $MM^\star$) sono detti *valori singolari* di M.

Decomposizione di Choleski

Definizione 9. Sia $M \in \mathrm{Mat}_n(\mathbb{C})$. Si dice *minore principale* p*–esimo* di M il minore ottenuto da M prendendo le prime p–righe e le prime p–colonne. Si indicherà tale minore con il simbolo $M^{(p)}$.

Lemma 19. *Se una matrice Hermitiana* $M \in \mathrm{Mat}_n(\mathbb{C})$ *è definita positiva allora anche tutti i suoi minori principali* $M^{(p)}$ *lo sono. In particolare, per ogni* $p = 1, \ldots, n$,

$$\det M^{(p)} \neq 0.$$

Dimostrazione. Se esistesse un vettore $\widetilde{\mathbf{x}} \in \mathbb{C}^p$ con $\widetilde{\mathbf{x}} \neq \mathbf{0}$ tale che

$$\widetilde{\mathbf{x}}^\star M^{(p)} \widetilde{\mathbf{x}} \leq 0,$$

allora sarebbe possibile scrivere il vettore $\mathbf{x} = (\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_p, 0, \ldots 0) \in \mathbb{C}^n \setminus \{\mathbf{0}\}$ per cui

$$\mathbf{x}^\star M \mathbf{x} = \widetilde{\mathbf{x}}^\star M^{(p)} \widetilde{\mathbf{x}} \leq 0.$$

Pertanto, affinché M sia definita positiva è necessario che tutti i suoi minori principali lo siano.

Infine, osserviamo che se una matrice N è singolare, allora esiste un vettore $\mathbf{x} \in \ker N$ con $\mathbf{x} \neq \mathbf{0}$. Pertanto,

$$\mathbf{x}^\star N \mathbf{x} = 0,$$

e N non può essere definita positiva. Da questo deduciamo che tutti i minori $M^{(p)}$ devono essere non singolari. □

Teorema 20. *Sia* $M \in GL_n(\mathbb{C})$ *una matrice Hermitiana definita positiva. Allora, esiste un'unica matrice triangolare inferiore* $L \in Mat_n(\mathbb{C})$ *con entrate sulla diagonale principale reali positive tale che* $M = LL^\star$. *Se* M *è una matrice reale simmetrica, allora anche la matrice* L *ottenuta è reale e* $M = LL^T$.

Dimostrazione. Poiché M è definita positiva, tutti i suoi minori principali sono non singolari. Pertanto esiste una fattorizzazione LU di M del tipo $M = L_0U_0$ con U_0 invertibile. Sia D la diagonale di U_0 e poniamo $U_0 = DU_1$, ove U_1 è una matrice triangolare superiore con tutti 1 sulla diagonale principale. Osserviamo che

$$M = M^\star = (LDU_1)^\star = U_1^\star D^\star L^\star.$$

Dall'unicità della fattorizzazione LU segue

$$U_1 = L_0^\star.$$

Pertanto, $M = L_0(DL_0^\star)$. Posto $P = (L_0^{-1})^\star$ si ha

$$\mathbf{x}^\star D\mathbf{y} = \mathbf{x}^\star P^\star M P\mathbf{y} = (P\mathbf{x})^\star M(P\mathbf{y}).$$

Quindi la matrice D è Hermitiana e, per il Corollario 13, visto che M è definita positiva, è anche essa definita positiva; in particolare D ammette radice quadrata. Poiché D è diagonale, tutte le sue entrate non nulle sono reali positive e, conseguentemente, lo sono anche tutte quelle di $\sqrt{D}$. Posto ora $L = L_0\sqrt{D}$, si ottiene

$$LL^\star = L_0\sqrt{D}(\sqrt{D})^\star L_0^\star = L_0DL_0^\star = M,$$

come desiderato.

Eè chiaro che è possibile ripetere la stessa costruzione per una matrice S reale simmetrica; in questo caso tutte le matrici (incluse $\sqrt{D}$ e L) sono reali, per cui si ottiene

$$S = LL^\star = LL^T.$$

□