1 Introduzione

Il teorema dimostrato da Schur nel 1923 descrive (vedi [6]) una sorprendente relazione che intercorre fra gli autovalori e gli elementi diagonali di una matrice Hermitiana; esso infatti mostra che, data una tale matrice e fissato un indice $k \leq n$ la somma dei k più piccoli autovalori è minore o uguale della somma dei più piccoli k elementi diagonali. È d'altra parte ben noto che per $k = n$ le due somme coincidono. In formula, data una matrice di ordine n con $A = A^H$, posto $\lambda_1 \leq \ldots \leq \lambda_n$, $a_{i_1} \leq \ldots \leq a_{i_n}$ (dove λ_i è l'iesimo autovalore di A e a_{i_j} il j-esimo elemento diagonale), vale

$$\begin{gathered}\lambda_1 \leq a_{i_1} \\ \lambda_1 + \lambda_2 \leq a_{i_1} + a_{i_2} \\ \vdots \\ \lambda_1 + \ldots + \lambda_{n-1} \leq a_{i_1} + \ldots + a_{i_n-1} \\ \lambda_1 + \ldots + \lambda_n = a_{i_1} + \ldots + a_{i_n}.\end{gathered}$$

Nel 1954 Horn dimostrò poi l'implicazione inversa di tale teorema (vedi [2]), ovvero che, dati due vettori reali $\underline{d} = [d_i], \underline{\lambda} = [\lambda_i]$ di dimensione n (e per comodità di notazione li supponiamo qui ordinati in modo crescente $\lambda_1 \leq \ldots \leq \lambda_n$, $d_1 \leq \ldots \leq d_n$) che soddisfino le disuguaglianze sopracitate

$$\begin{gathered}\lambda_1 \leq d_1 \\ \lambda_1 + \lambda_2 \leq d_1 + d_2 \\ \vdots \\ \lambda_1 + \ldots + \lambda_{n-1} \leq d_1 + \ldots + d_{n-1} \\ \lambda_1 + \ldots + \lambda_n = d_1 + \ldots + d_n\end{gathered}$$

esiste una matrice Hermitiana di ordine n che abbia per autovalori i λ_i e come elementi diagonali i d_i.
Nel 1964 Mirsky, cercando una generalizzazione del Teorema di Schur-Horn alle matrici quadrate non necessariamente Hermitiane, trovò il seguente risultato (vedi [5]):

Teorema 1.1 *Condizione necessaria e sufficiente per l'esistenza di una matrice con autovalori $\lambda_1, \ldots, \lambda_n$ ed elementi diagonali $a_{11}, \ldots, a_{nn}$ preassegnati è che*

$$\sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i.$$

Tale teorema non ebbe però importanti sviluppi perché la condizione imposta alla matrice è troppo debole, e lo stesso Mirsky si chiese (sempre in [5]) se esistesse qualche legame tra gli elementi diagonali ed i valori singolari di una matrice. Un'importante risposta a questo quesito fu trovata separatamente da Sing nel 1976 [7] e da Thompson nel 1977 [8]. Essi mostrarono che, data una matrice A di ordine n i suoi elementi diagonali (che non è restrittivo supporre ordinati in modo decrescente) $a_{11} \geq \ldots \geq a_{nn}$ ed i suoi valori singolari (ordinati sempre in modo decrescente) $s_1 \geq \ldots \geq s_n$ soddisfano le seguenti condizioni:

$$|a_{11}| \leq s_1$$

$$|a_{11}| + |a_{22}| \leq s_1 + s_2$$

$$\vdots$$

$$|a_{11}| + \ldots + |a_{nn}| \leq s_1 + \ldots + s_n$$

$$|a_{11}| + \ldots + |a_{n-1,n-1}| - |a_{nn}| \leq s_1 + \ldots + s_{n-1} - s_n.$$

Thompson mostrò inoltre nel medesimo lavoro anche l'implicazione inversa di questo teorema (mentre Sing la dimostrò solo per il caso 2×2). Egli infatti trovò che, dati due vettori $\underline{d} = [d_1 \ldots d_n]$ a coefficienti in $\mathbb{C}$, $\underline{s} = [s_1 \ldots s_n]$ a coefficienti in $\mathbb{R}$ con $|d_1| \geq \ldots \geq |d_n|$, $s_1 \geq \ldots \geq s_n$ che soddisfino alle condizioni sopracitate

$$|d_1| \leq s_1$$

$$|d_1| + |d_2| \leq s_1 + s_2$$

$$\vdots$$

$$|d_1| + \ldots + |d_n| \leq s_1 + \ldots + s_n$$

$$|d_1| + \ldots + |d_{n-1}| - |d_n| \leq s_1 + \ldots + s_{n-1} - s_n,$$

esiste una matrice che ha per elementi diagonali i d_i e valori singolari gli s_i.
In questo lavoro ci occuperemo di dare una dimostrazione del Teorema di Schur-Horn e del Teorema di Sing-Thompson. Per il primo Teorema seguiremo la trattazione datane nel libro di Horn e Jhonson [3]. Per il secondo Teorema utilizzeremo:
a) per la sufficienza nel caso 2×2 e per la necessità nel caso $n \times n$ (tranne l'ultima disuguagliaza) il lavoro di Sing [7];
b) per completare la necessità nel caso $n \times n$ il lavoro di Thompson [8];
c) per la sufficienza nel caso $n \times n$ il lavoro di Chu [1], utilizzato anche per una interpretazione geometrica del caso 2×2 reale.

2 Notazioni e preliminari

Vengono qui presentati alcuni risultati ben noti (per una loro dimostrazione vedere anche [3] e [4]), alcuni meno noti (di cui daremo qui la dimostrazione) e alcune definizioni preliminari necessarie per poter affrontare i teoremi discussi nel seguito:

Definizione 2.1 (Decomposizione in valori singolari) *Data una matrice $A \in M_{m\times n}[\mathbb{C}]$ di rango k, una sua decomposizione in valori singolari è*

$$A = U \cdot S \cdot V^H$$

con U, V matrici unitarie, $U \in M_m[\mathbb{C}]$, $V \in M_n[\mathbb{C}]$ e $S \in M_{m\times n}[\mathbb{R}]$, $S = psdiag([s_1, \ldots, s_k, 0, \ldots, 0])$, con $s_1 \geq s_2 \geq \ldots \geq s_k > 0$. Gli s_i (assieme agli eventuali 0) si dicono valori singolari di A e sono univocamente individuati da A, mentre U e V non lo sono. Nel seguito indicheremo con $\underline{s}(A)$ il vettore formato dai valori singolari della matrice A.

Data una matrice $A \in M_{m\times n}[\mathbb{C}]$ esiste sempre una sua decomposizione in valori singolari; inoltre, nel caso di matrici Hermitiane i valori singolari coincidono con il modulo degli autovalori.
Nota: nel seguito talvolta si userà, per comodità di notazione, $A = U \cdot S \cdot W$ (con U, W matrici unitarie, $U \in M_m[\mathbb{C}]$, $W \in M_n[\mathbb{C}]$ e S definita come prima) come decomposizione in valori singolari di una matrice A. Essa è ovviamente del tutto equivalente alla definizione data prima; basta infatti prendere $W = V^H$, essendo questa ancora una matrice unitaria.

Lemma 2.2 *Sia $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = A \in M_2[\mathbb{C}]$ e siano s_1, s_2 i suoi valori singolari; allora si ha*

$$s_1^2 + s_2^2 = |a|^2 + |b|^2 + |c|^2 + |d|^2$$

e

$$s_1 s_2 = |ad - bc|.$$

Dim.
Basta prendere una decomposizione in valori singolari di A ed osservare che

$$s_1^2 + s_2^2 = Tr(A \cdot A^H) = |a|^2 + |b|^2 + |c|^2 + |d|^2$$

e

$$s_1 s_2 = |det(A)| = |ad - bc|.$$

□

Definizione 2.3 *Dati due vettori* $\underline{a} = [a_i]$, $\underline{b} = [b_i] \in \mathbb{R}^n$, *con* $a_{i_1} \geq \ldots \geq a_{i_n}$ *e* $b_{i_1} \geq \ldots \geq b_{i_n}$, *si scriverà* $\underline{a} \preceq \underline{b}$ *se*

$$\sum_{j=1}^{k} a_{i_j} \geq \sum_{j=1}^{k} b_{i_j} \quad \forall k = 1, \ldots, n \tag{1}$$

e

$$\sum_{j=1}^{n} a_{i_j} = \sum_{j=1}^{n} b_{i_j}. \tag{2}$$

Attesa la condizione (2), la condizione (1) è equivalente alla seguente:

$$\sum_{j=k}^{n} a_{i_j} \leq \sum_{j=k}^{n} b_{i_j} \quad \forall k = 2, \ldots, n, \tag{3}$$

ovvero, se i vettori $\underline{a}, \underline{b}$ *sono ordinati in modo crescente,* $a_{i'_1} \leq \ldots \leq a_{i'_n}$ *e* $b_{i'_1} \leq \ldots \leq b_{i'_n}$, *allora* $\underline{a} \preceq \underline{b}$ *se e solo se*

$$\sum_{j=1}^{k} a_{i'_j} \leq \sum_{j=1}^{k} b_{i'_j} \quad \forall k = 1, \ldots, n \tag{4}$$

e

$$\sum_{j=1}^{n} a_{i'_j} = \sum_{j=1}^{n} b_{i'_j}. \tag{5}$$

In $\mathbb{R}^n$ $\preceq$ *è solo un preordine, non godendo della proprietà antisimmetrica; esso induce però un ordine (parziale) su* $\mathbb{R}^n/\sim$, *dove* $\underline{a} \sim \underline{b} \Leftrightarrow \exists\, P$ *matrice di permutazione tale che* $\underline{a} = P \cdot \underline{b}$.

Nel seguito, oltre a $\preceq$ useremo anche una seconda relazione di preordine $\prec$, dove si dirà che $\underline{a} \prec \underline{b}$ se, ordinati $\underline{a}$ e $\underline{b}$ in modo decrescente, cioè $a_{i_1} \geq \ldots \geq a_{i_n}$ e $b_{i_1} \geq \ldots \geq b_{i_n}$ si ha

$$\sum_{j=1}^{k} a_{i_j} \geq \sum_{j=1}^{k} b_{i_j} \quad \forall\, k = 1, \ldots, n, \tag{6}$$

ovvero, rispetto alla relazione $\preceq$, l'uguaglianza in (2) diventa disuguaglianza.

Definizione 2.4 (Matrici doppiamente substocastiche) *Una matrice* $D = (d_{ij}) \in M_n[\mathbb{R}]$, $D \geq 0$, *si dice doppiamente substocastica se* $\sum_{j=1}^{n} d_{ij} \leq 1\ \forall\, i$ *e* $\sum_{i=1}^{n} d_{ij} \leq 1\ \forall\, j$.

Nota: $D \geq 0$ significa $d_{ij} \geq 0 \, \forall \, i, j$.

Lemma 2.5 *Siano $U = (u_{ij}), V = (v_{ij})$ matrici di ordine n unitarie; posto $W = (w_{ij}) = (|u_{ij} v_{ji}|)$ si ha allora che W è una matrice doppiamente substocastica.*

Dim.
Innanzitutto è banalmente $W \geq 0$. Supponiamo ora per assurdo che esista un indice i tale che $\sum_{j=1}^{n} w_{ij} > 1$; allora

$$1 < \left(\sum_{j=1}^{n} |u_{ij} v_{ji}| \right)^2 = \sum_{j=1}^{n} |u_{ij} v_{ji}|^2 + \sum_{1 \leq j < h \leq n} 2 |u_{ij} v_{ji}| |u_{ih} v_{hi}| = A.$$

Inoltre, poiché U, V sono matrici unitarie, si ha che $|U|^2 = (|u_{ij}|^2)$ e $|V|^2 = (|v_{ij}|^2)$ sono matrici ortostocastiche per definizione, da cui segue $\sum_{j=1}^{n} |u_{ij}|^2 = 1$ e $\sum_{j=1}^{n} |v_{ji}|^2 = 1$. Si ha quindi

$$1 = \sum_{j=1}^{n} |u_{ij}|^2 \cdot \sum_{j=1}^{n} |v_{ji}|^2 = \sum_{j=1}^{n} |u_{ij} v_{ji}|^2 + \sum_{j,h,j \neq h} |u_{ij} v_{hi}|^2 = B$$

da cui segue

$$0 > B - A = \sum_{1 \leq j < h \leq n} (|u_{ij} v_{hi}| - |u_{ih} v_{ji}|)^2,$$

assurdo. Dev'essere dunque $\sum_{j=1}^{n} |w_{ij}| \leq 1 \; \forall \, i = 1, \ldots, n$, e in modo analogo si dimostra $\sum_{i=1}^{n} |w_{ij}| \leq 1 \; \forall j = 1, \ldots, n$.

□

Per le matrici doppiamente substocastiche si ha il seguente lemma:

Lemma 2.6 *Dato un vettore $\underline{y} = [y_i] \in \mathbb{R}^n$ tale che $y_i \geq 0$ per $i = 1, \ldots, n$, allora si ha $\underline{y} \prec S \cdot \underline{y}$ per ogni matrice $S = (s_{ij})$ doppiamente substocastica.*

Dim:
Sia $\underline{x} = [x_i] = S \cdot \underline{y}$; senza perdita di generalità possimo suppore $x_1 \geq \ldots \geq x_n$ e $y_1 \geq \ldots \geq y_n$. Infatti, premoltiplicando $\underline{x}, \underline{y}$ per le opportune matrici di permutazione P, P', ottenendo

$$P \cdot \underline{x} = P \cdot S \cdot P' \cdot (P')^T \cdot \underline{y},$$

dove le coordinate di $P \cdot \underline{x}$ e $(P')^T \cdot \underline{y}$ sono ordinate in modo decrescente e, se S è doppiamente substocastica, lo è anche $P \cdot S \cdot P'$.

Verifichiamo ora che $\forall\, k = 1, \ldots, n$ dev'essere $\sum_{i=1}^{k} x_i \leq \sum_{i=1}^{k} y_i$; infatti fissato k si ha

$$\sum_{i=1}^{k} x_i - \sum_{i=1}^{k} y_i = \sum_{i=1}^{k}\sum_{j=1}^{n} s_{ij} y_j - \sum_{i=1}^{k} y_i = \sum_{j=1}^{n} \underbrace{\left(\sum_{i=1}^{k} s_{ij} \right)}_{=\sigma_j^{(k)}} y_j - \sum_{i=1}^{k} y_i =$$

$$= \sum_{j=1}^{n} \sigma_j^{(k)} y_j - \sum_{i=1}^{k} y_i \leq \sum_{j=1}^{n} \sigma_j^{(k)} y_j - \sum_{i=1}^{k} y_i + \left(k - \sum_{j=1}^{n} \sigma_j^{(k)} \right) y_k =$$

$$= \sum_{j=1}^{k} \sigma_j^{(k)} y_j + \sum_{j=k+1}^{n} \sigma_j^{(k)} y_j - \sum_{j=1}^{k} y_j + \sum_{j=1}^{k} y_k - \sum_{j=1}^{k} \sigma_j^{(k)} y_k - \sum_{j=k+1}^{n} \sigma_j^{(k)} y_k =$$

$$= \sum_{j=1}^{k} \underbrace{\left(1 - \sigma_j^{(k)}\right)}_{\geq 0} \underbrace{(y_k - y_j)}_{\leq 0} + \sum_{j=k+1}^{n} \underbrace{\sigma_j^{(k)}}_{\geq 0} \underbrace{(y_j - y_k)}_{\leq 0} \leq 0,$$

dove nella disuguaglianza centrale si è usato $\left(k - \sum_{j=1}^{n} \sigma_j^{(k)}\right) y_k \geq 0$, in quanto $\sum_{j=1}^{n} \sigma_j^{(k)} \leq k$ poiché è la somma degli elementi delle prime k righe di S, e $y_k \geq 0$ per ipotesi.

$\square$

Definizione 2.7 (Matrici doppiamente stocastiche) *Una matrice $D = (d_{ij}) \in M_n[\mathbb{R}]$, $D \geq 0$ si dice doppiamente stocastica se è stocastica per righe e per colonne, se cioè $\sum_{j=1}^{n} d_{ij} = 1\ \forall\, i$ e $\sum_{i=1}^{n} d_{ij} = 1\ \forall\, j$.*

Definizione 2.8 (Matrici ortostocastiche) *Una matrice $D = (d_{ij})$ si dice ortostocastica se esiste $U = (u_{ij})$ matrice unitaria tale che $d_{ij} = |u_{ij}|^2$.*

Nota: Come si vede facilmete, ogni matrice ortostocastica è anche doppiamente stocastica.
Per le matrici doppiamente stocastiche si ha un lemma analogo al Lemma 2.6, che è il seguente lemma:

Lemma 2.9 *Dato un vettore $\underline{y} = [y_i] \in \mathbb{R}^n$ si ha $\underline{y} \preceq D \cdot \underline{y}$ per ogni matrice $D = (d_{ij})$ doppiamente stocastica.*

Ci serviranno inoltre per la dimostrazione dei teoremi di Schur e Horn i seguenti risultati:

Teorema 2.10 (Teorema dell'intreccio) *Sia $A \in M_n[\mathbb{C}]$ e sia A' una sua sottomatrice principale, $A' \in M_{n-1}[\mathbb{C}]$. Detti λ_i e μ_j rispettivamente gli autovalori di A e A' ordinati in modo non decrescente, si ha*

$$\lambda_1 \leq \mu_1 \leq \lambda_2 \leq \mu_2 \leq \ldots \leq \mu_{n-1} \leq \lambda_n.$$

Lemma 2.11 (Lemma combinatorio dell'intreccio) *Sia $n \geq 1$ e sia $\underline{d} \succeq \underline{\lambda} \in \mathbb{R}^n$, con $d_1 \leq \ldots \leq d_n$ e $\lambda_1 \leq \ldots \leq \lambda_n$. Allora esiste $\underline{\mu} \in \mathbb{R}^{n-1}$ tale che*

$$\lambda_1 \leq \mu_1 \leq \lambda_2 \leq \mu_2 \leq \ldots \leq \mu_{n-1} \leq \lambda_n$$

e

$$[d_1\, d_2\, \ldots\, d_{n-1}]^T \succeq \underline{\mu}.$$

Per una dimostrazione del lemma, che è molto tecnica e non fa uso della teoria delle matrici, si veda Horn e Johnson, Matrix Analysis, 1985, pag 194-6.
Nel seguito, data una matrice $A = (a_{ij}) \in M_n[\mathbb{C}]$, indicheremo con $\underline{d}(A)$ il vettore formato dagli elementi diagonali di A e il vettore spettro di A con $\underline{\lambda}(A)$.

Lemma 2.12 (Formula del determinante a blocchi) *Sia $X \in M_{m+n}[\mathbb{C}]$ matrice a blocchi, $X = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)$, con $A \in M_n[\mathbb{C}]$ matrice invertibile. Si ha allora*

$$det(X) = det(A)\, det(D - C \cdot A^{-1} \cdot B)$$

e $(D - C \cdot A^{-1} \cdot B)$ si dice complemento di Schur.

Dim:
Basta osservare che

$$det(X) = det\left(\left(\begin{array}{c|c} \mathbb{I}_n & \mathbb{O} \\ \hline -C \cdot A^{-1} & \mathbb{I}_m \end{array}\right) \cdot \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)\right) =$$

$$= det\left(\begin{array}{c|c} A & B \\ \hline \mathbb{O} & D - C \cdot A^{-1} \cdot B \end{array}\right).$$

$\square$

Teorema 2.13 (Polinomio interpolatore di Lagrange) *Sia K un campo. Fissati n elementi distinti $a_1, \ldots, a_n \in K, n \geq 2$, e assegnati $\alpha_1, \ldots, \alpha_n \in K$ esiste $f(t) \in K[t]$ con $\deg(f) < n$, tale che:*

$$f(a_i) = \alpha_i \,\forall\, i = 1, \ldots, n.$$

Tale polinomio, detto ***polinomio interpolatore di Lagrange****, è unico e si scrive nel seguente modo:*

$$f(t) = \frac{(t-a_2)(t-a_3)\dots(t-a_n)}{(a_1-a_2)(a_1-a_3)\dots(a_1-a_n)}\,\alpha_1 + \frac{(t-a_1)(t-a_3)\dots(t-a_n)}{(a_2-a_1)(a_2-a_3)\dots(a_2-a_n)}\,\alpha_2 + \dots +$$

$$+\frac{(t-a_1)(t-a_2)\dots(t-a_{n-1})}{(a_n-a_1)(a_n-a_2)\dots(a_n-a_{n-1})}\,\alpha_n.$$

Per dimostrare il Teorema di Sing-Thompson avremo inoltre bisogno dei seguenti risultati e definizioni:

Definizione 2.14 (Prodotto interno di Frobenius) *Date due matrici $A, B \in M_{m\times n}[\mathbb{C}]$, definiremo* ***prodotto interno di Frobenius*** *su $M_{m\times n}[\mathbb{C}]$ il numero complesso*

$$(A|B)_{\mathcal{F}} = Tr(B^H \cdot A).$$

Definizione 2.15 (Norma di Frobenius) *Data una matrice $A = (a_{ij}) \in M_{m\times n}[\mathbb{C}]$, chiamasi* ***norma di Frobenius*** *di A il numero reale:*

$$|||A|||_{\mathcal{F}} = (A|A)_{\mathcal{F}}^{1/2} = \left(\sum_{i,j} |a_{ij}|^2\right)^{1/2}.$$

Lemma 2.16 *Sia $A \in M_{m\times n}[\mathbb{C}]$ e siano U, V matrici unitarie di rango m e n rispettivamente; allora*

$$|||U \cdot A|||_{\mathcal{F}} = |||A|||_{\mathcal{F}} = |||A \cdot V|||_{\mathcal{F}}$$

Quindi, se $A = U \cdot psdiag([s_1 \dots s_k 0 \dots 0]) \cdot V^H$ è una decomposizione in valori singolari di A, si ha

$$|||A|||_{\mathcal{F}} = \left(\sum_i s_i^2\right)^{1/2}.$$

Nota: la norma di Frobenius è una norma matriciale, cioè oltre alle proprietà delle norme vettoriali soddisfa anche la proprietà:

$$|||A \cdot B|||_{\mathcal{F}} \leq |||A|||_{\mathcal{F}}\, |||B|||_{\mathcal{F}}$$

se $A \in M_{m\times n}[\mathbb{C}]$ e $B \in M_{n\times k}[\mathbb{C}]$.
Dato un numero $a \in \mathbb{C}$ indicheremo d'ora in avanti con $\Re(a)$ la sua parte reale e con $\Im(a)$ la sua parte immaginaria.

Lemma 2.17 *Siano* $\underline{w} = [w_i], \underline{s} = [s_i], \underline{s}' = [s_i'] \in \mathbb{R}^n$ *vettori tali che* $w_1 \geq \ldots \geq w_n \geq 0$*,* $s_1 \geq \ldots \geq s_n \geq 0$*,* $s_1' \geq \ldots \geq s_n' \geq 0$ *e* $\underline{w} \succ \underline{s}$*; si ha allora*

$$\sum_{i=1}^{n} s_i' s_i \geq \sum_{i=1}^{n} s_i' w_i$$

Dim.
Per induzione sulla dimensione n dei vettori:
Caso $n = 1$: vero perché $0 \leq w_1 \leq s_1$ e $s_1' \geq 0$.
$n-1 \Rightarrow n$: siano $\underline{v} = [v_i], \underline{t} = [t_i], \underline{s}'' = [s_i'']$ vettori di dimensione $n-1$ con $v_1 = w_1 + w_2$, $v_j = w_{j+1}$ per $j \geq 2$, $t_1 = s_1 + s_2, t_j = s_{j+1}$ per $j \geq 2$ e $s_j'' = s_{j+1}'$ per ogni $j = 1, \ldots, n-1$; si ha che essi soddisfano all'ipotesi induttiva, per cui vale

$$\sum_{i=1}^{n-1} s_i'' v_i \leq \sum_{i=1}^{n-1} s_i'' t_i$$

ovvero

$$s_2' w_1 + \sum_{i=2}^{n} s_i' w_i \leq s_2' s_1 + \sum_{i=2}^{n} s_i' s_i. \tag{7}$$

Poiché vale anche

$$s_1' w_1 - s_2' w_1 \leq s_1' s_1 - s_2' s_1 \tag{8}$$

($s_1' - s_2' \geq 0$ e $w_1 \leq s_1$) si ha, sommando (7) con (8)

$$\sum_{i=1}^{n} s_i' w_i \leq \sum_{i=1}^{n} s_i' s_i,$$

che è la tesi.

□

I seguenti due risultati sono i punti b) e c) dell'esercizio 18 a pp.185-186 di [4].

Lemma 2.18 *Siano* $A, B \in M_{m \times n}[\mathbb{C}]$ *matrici con valori singolari rispettivamente* s_i, s_i'*; posto* $q = \min\{m, n\}$*, si ha*

$$|(A|B)_{\mathcal{F}}| \leq \sum_{i=1}^{q} s_i s_i' \leq |||A|||_{\mathcal{F}}\, |||B|||_{\mathcal{F}}$$

Dim.
Dimostreremo innanzitutto la prima disuguaglianza nel caso quadrato, ovvero con $n = m = q$; mostreremo poi che si generalizza facilmente. Sia dunque $m = n = q$ e siano $A = U \cdot S \cdot V$, $B = U' \cdot S' \cdot V'$ decomposizioni in valori singolari di A e B con $s_1 \geq \ldots \geq s_q$, $s'_1 \geq \ldots \geq s'_q$; si ha

$$|Tr(B^H \cdot A)| = |Tr(U'^H \cdot U \cdot S \cdot V \cdot V'^H \cdot S')| \leq Tr(|W'| \cdot S \cdot |W''| \cdot S')$$

dove $W' = (w'_{ij}) = U'^H \cdot U$, $W'' = (w''_{ij}) = V \cdot V'^H$. Si ha

$$Tr(|W'| \cdot S \cdot |W''| \cdot S') = \sum_{i=1}^{q} s'_i \sum_{j=1}^{q} |w'_{ij} w''_{ji}| s_j = \sum_{i=1}^{q} s'_i (W \cdot \underline{s})_{ii}$$

dove si è posto $W = (w_{ij}) = (|w'_{ij} w''_{ji}|)$. Poiché $s'_i (W \cdot \underline{s})_{ii} \geq 0$ per ogni i e poiché gli s'_i sono in ordine decrescente si ha (si intende $(W \cdot \underline{s})_{i_1, i_1} \geq \ldots \geq (W \cdot \underline{s})_{i_q, i_q}$)

$$\sum_{i=1}^{q} s'_i (W \cdot \underline{s})_{ii} \leq \sum_{i=1}^{q} s'_i (W \cdot \underline{s})_{i_j, i_j};$$

inoltre W è matrice doppiamente substocastica per il Lemma 2.5, per cui per il Lemma 2.6 si ha $W \cdot \underline{s} \succ \underline{s}$ (si è posto $\underline{s} = [s_i]$). Il Lemma 2.17 assicura allora

$$\sum_{i=1}^{q} s'_i (W \cdot \underline{s})_{i_j, i_j} \leq \sum_{i=1}^{q} s'_i s_i,$$

e risalendo le disuguaglianze si ottiene

$$|Tr(B^H \cdot A)| \leq \sum_{i=1}^{q} s'_i s_i.$$

Supponiamo ora che sia invece $m \neq n$; se si considerano le matrici quadrate di ordine $m+n$ $A_0 = \left(\begin{array}{c|c} A & \mathbb{O} \\ \hline \mathbb{O} & \mathbb{O} \end{array} \right)$ e $B_0 = \left(\begin{array}{c|c} B & \mathbb{O} \\ \hline \mathbb{O} & \mathbb{O} \end{array} \right)$, si ha che A_0 e B_0 hanno i valori singolari non nulli uguali a quelli non nulli rispettivamente di A e B, e $Tr(B_0^H \cdot A_0) = Tr(B^H \cdot A)$. Questo mostra che la prima disuguaglianza vale per A e B non necessariamente quadrate.
Verifichiamo ora, nel caso generale, che vale

$$\sum_{i=1}^{q} s_i s'_i \leq |||A|||_{\mathcal{F}} \, |||B|||_{\mathcal{F}};$$

poiché entrambi i membri della disequazione sono ≥ 0, si può quadrare ottenendo una disequazione equivalente. Si ha poi

$$\left(\sum_{i=1}^{q} s_i s_i'\right)^2 = \sum_{i=1}^{q} (s_i s_i')^2 + \sum_{i \neq j} s_i s_i' s_j s_j',$$

$$|||A|||_{\mathcal{F}}^2 \, |||B|||_{\mathcal{F}}^2 = \sum_{i=1}^{q} (s_i s_i')^2 + \sum_{i \neq j} (s_i s_j')^2$$

da cui

$$|||A|||_{\mathcal{F}} \, |||B|||_{\mathcal{F}} - \sum_{i=1}^{q} s_i s_i' = \sum_{i \neq j} (s_i - s_j)^2 \geq 0,$$

che è quanto voluto.

$\square$

Proposizione 2.19 *Siano $A, B \in M_{m\times n}[\mathbb{C}]$ matrici con valori singolari rispettivamente s_i, s_i'; posto $q = \min\{m, n\}$, si ha allora:*

$$|||A - B|||_{\mathcal{F}}^2 \geq \left(\sum_{i=1}^{q} (s_i - s_i')^2\right)^{1/2}.$$

Dim.
Per il Lemma 2.18 si ha

$$|||A - B|||_{\mathcal{F}}^2 = |||A|||_{\mathcal{F}}^2 + |||B|||_{\mathcal{F}}^2 - 2\Re(Tr(B^H \cdot A)) \geq$$

$$\geq |||A|||_{\mathcal{F}}^2 + |||B|||_{\mathcal{F}}^2 - 2|Tr(B^H \cdot A)| \geq \sum_{i=1}^{q} s_i^2 + \sum_{i=1}^{q} s_i^2 - 2\sum_{i=1}^{q} s_i s_i' =$$

$$= \sum_{i=1}^{q} (s_i - s_i')^2.$$

$\square$

3 Il teorema di Schur

Vediamo ora il Teorema di Schur del 1923:

Teorema 3.1 (Teorema di Schur) *Sia* $A = A^H \in M_n[\mathbb{C}]$ *con autovalori* $\lambda_1 \leq \ldots \leq \lambda_n$. *Allora* $\underline{d}(A) \succeq \underline{\lambda}(A)$.

Dim.
Notiamo per prima cosa che, senza perdita di generalità, possiamo supporre $a_{11} \leq \ldots \leq a_{nn}$; se così non fosse basterebbe infatti considerare una cogredienza che permuti le righe e le colonne di A in modo tale che si abbia $\widehat{a}_{11} \leq \ldots \leq \widehat{a}_{nn}$. Poiché le cogredienze sono particolari similitudini esse non cambiano gli autovalori.
Mostriamo ora il teorema per induzione:
Caso n=1: banalmente $\lambda_1 = a_{11}$.
$n-1 \Rightarrow n$:

$$A = \left(\begin{array}{c|c} A' & \underline{x} \\ \hline \underline{x}^H & a_{nn} \end{array}\right),$$

dove $A' \in M_{n-1}[\mathbb{C}]$. Siano $\mu_1 \leq \ldots \leq \mu_{n-1}$ gli autovalori di A'. Per ipotesi induttiva si ha:

$$\begin{aligned} a_{11} &\geq \mu_1 \\ a_{11} + a_{22} &\geq \mu_1 + \mu_2 \\ &\vdots \\ a_{11} + \ldots + a_{n-2,n-2} &\geq \mu_1 + \ldots + \mu_{n-2} \\ a_{11} + \ldots + a_{n-1,n-1} &= \mu_1 + \ldots + \mu_{n-1} \end{aligned}$$

Inoltre, per il teorema dell'intreccio è

$$\lambda_1 \leq \mu_1 \leq \lambda_2 \leq \mu_2 \leq \ldots \leq \mu_{n-1} \leq \lambda_n,$$

da cui si ricava

$$\begin{aligned} a_{11} &\geq \lambda_1 \\ a_{11} + a_{22} &\geq \lambda_1 + \lambda_2 \\ &\vdots \\ a_{11} + \ldots + a_{n-1n-1} &\geq \lambda_1 + \ldots + \lambda_{n-1}; \end{aligned}$$

si ha inoltre

$$a_{11} + \ldots + a_{nn} = Tr(A) = \lambda_1 + \ldots + \lambda_n,$$

e quindi la tesi.

$\square$

Vediamo ora anche una dimostrazione alternativa del Teorema di Schur; tale dimostrazione non utilizza il Teorema dell'intreccio.

Dim.

A è Hermitiana, quindi per il Teorema spettrale esiste $U = (u_{ij})$ matrice unitaria tale che $A = U \cdot \Lambda \cdot U^H$, con $\Lambda = diag([\lambda_1 \dots \lambda_n]) \in M_n[\mathbb{R}]$; da tale uguaglianza segue

$$a_{ii} = \sum_{j=1}^{n} u_{ij} \lambda_j \overline{u}_{ij} = \sum_{j=1}^{n} |u_{ij}|^2 \lambda_j,$$

ovvero $\underline{d}(A) = D \cdot \underline{\lambda}(A)$ (dove si è posto $D = (d_{ij}) = (|u_{ij}|^2)$). Per costruzione D è una matrice ortostocastica, perciò si ha $\underline{\lambda}(A) \preceq \underline{d}(A)$ per il Lemma 2.9.

$\square$

4 Il teorema di Horn

Come visto, il teorema di Schur dice che, data una matrice complessa Hermitiana A, il suo vettore degli autovalori e quello diagonale soddisfano alla disuguaglianza $\underline{d}(A) \succeq \underline{\lambda}(A)$; è vero anche il viceversa? A questa domanda rispose Horn nel 1954 dimostrando che, dati due vettori reali $\underline{d}, \underline{\lambda}$ tali che $\underline{d} \succeq \underline{\lambda}$, esiste sempre una matrice reale simmetrica A tale che $\underline{d}(A) = \underline{d}$ e $\underline{\lambda}(A) = \underline{\lambda}$. Vedremo qui tale risultato seguendo la dimostrazione data da Horn e Johnson [3].

Teorema 4.1 (Teorema di Horn) *Dati* $\underline{d}\,, \underline{\lambda} \in \mathbb{R}^n, \underline{d} \succeq \underline{\lambda}$, *esiste* $A \in M_n[\mathbb{R}], A = A^T$ *tale che* $\underline{d}(A) = \underline{d}$ *e* $\underline{\lambda}(A) = \underline{\lambda}$.

Prima di dimostrare il Teorema di Horn abbiamo bisogno di vedere un altro risultato, da cui il teorema seguirà facilmente. Si tratta del

Teorema 4.2 (Teorema inverso dell'intreccio) *Sia* $A_1 \in M_{n-1}[\mathbb{R}], A_1 = A_1^T$, *con spettro* $\underline{\lambda}(A_1) = \underline{\mu} = [\mu_i]$. *Se* $\underline{\lambda} = [\lambda_i] \in \mathbb{R}^n$ *soddisfa alla condizione*

$$\lambda_1 \leq \mu_1 \leq \lambda_2 \leq \mu_2 \leq \ldots \leq \mu_{n-1} \leq \lambda_n, \tag{9}$$

allora esistono $a \in \mathbb{R}$, $\underline{b} \in \mathbb{R}^{n-1}$ *tali che la matrice*

$$A = \left(\begin{array}{c|c} a & \underline{b}^T \\ \hline \underline{b} & A_1 \end{array}\right)$$

ha spettro $\underline{\lambda}(A) = \underline{\lambda}\,$.

Dim.
Questa dimostrazione si compone di due passi, il primo in cui si dimostra il teorema per le matrici diagonali, cioè nel caso sia $A_1 = diag(\mu_1, \ldots, \mu_{n-1})$, il secondo in cui viene dimostrato per una matrice A_1 qualsiasi riconducendosi al caso diagonale.

Passo 1
Sia A_1 matrice diagonale, cioè $A_1 = D = diag([\mu_1, \ldots, \mu_{n-1}])$. La matrice A cercata sarà del tipo

$$A = \left(\begin{array}{c|ccc} a & & \underline{b}^T & \\ \hline & \underline{\mu}_1 & & \\ \underline{b} & & \ddots & \\ & & & \underline{\mu}_{n-1} \end{array}\right)$$

L'elemento a è individuato, infatti

$$a = Tr(A) - Tr(D) = \sum_{i=1}^{n} \lambda_i - \sum_{j=1}^{n-1} \mu_j.$$

Inoltre il polinomio caratteristico di A risulta (utilizzando la formula del determinante a blocchi, Lemma 2.12)

$$p_A(t) = \prod_{i=1}^{n} (t - \lambda_i) = det \left(\begin{array}{c|c} t-a & -\underline{b}^T \\ \hline -\underline{b} & t\,\mathbb{I}_{n-1} - D \end{array} \right) =$$

$$= \prod_{j=1}^{n-1} (t-\mu_j)[t - a - \underline{b}^T (t\,\mathbb{I}_{n-1} - D)^{-1} \underline{b}] = \prod_{j=1}^{n-1} (t-\mu_j) \left[t - a - \sum_{i=1}^{n-1} \frac{b_j^2}{t-\mu_j} \right] =$$

$$= g(t) \left[t - a - \sum_{i=1}^{n-1} \frac{b_j^2}{t-\mu_j} \right], \tag{10}$$

dove si è posto $g(t) = \prod_{j=1}^{n-1} (t-\mu_j)$. $g(t)$ è polinomio di grado $n-1$, $p_A(t)$ è poinomio di grado n, si ha quindi:

$$p_A(t) = g(t) \cdot (t-c) + r(t), \, \deg(r) < n-1$$

e si vede facilmente che $c = a$ (infatti $p_A(t) = g(t)\,(t-a) - g(t) \left[\sum_{i=1}^{n-1} \frac{b_j^2}{t-\mu_j} \right]$ e $g(t) \left[\sum_{i=1}^{n-1} \frac{b_j^2}{t-\mu_j} \right]$ ha grado minore di $n-1$).
Inoltre $\forall\, j = 1, \ldots, n-1$ si ha che $p_A(\mu_j) = r(\mu_j)$. Supponiamo per ora che i μ_j siano tutti distinti. In questo caso $r(t)$ è univocamente individuato e calcolabile tramite i polinomi interpolatori di Lagrange (Teorema 2.13), ovvero

$$r(t) = \sum_{j=1}^{n-1} \frac{\prod_{i \neq j} (t-\mu_i)}{\prod_{i \neq j} (\mu_j - \mu_i)} p_A(\mu_j) = \sum_{j=1}^{n-1} p_A(\mu_j) \cdot \frac{g(t)}{g'(\mu_j)\,(t-\mu_j)}$$

Si ha quindi

$$\frac{p_A(t)}{g(t)} = t - a + \frac{r(t)}{g(t)} = t - a - \sum_{j=1}^{n-1} - \frac{p_A(\mu_j)}{g'(\mu_j)} \cdot \frac{1}{t-\mu_j}$$

e confrontando con (10) si vede che basta porre

$$b_j^2 = -\frac{p_A(\mu_j)}{g'(\mu_j)} \quad \forall\, j = 1, \ldots, n-1$$

Bisogna quindi verificare che $\frac{p_A(\mu_j)}{g'(\mu_j)} \leq 0 \, \forall \, j = 1, \ldots, n-1$. Si ha

$$p_A(\mu_j) = \prod_{i=1}^{n} (\mu_j - \lambda_i) = \prod_{i=1}^{j} (\mu_j - \lambda_i) \cdot \prod_{i=j+1}^{n} (\mu_j - \lambda_i)$$

dove i primi j fattori sono tutti maggiori o uguali a 0, gli ultimi $n-j$ negativi o nulli; inoltre

$$g'(\mu_j) = \prod_{i=1, i\neq j}^{n} (\mu_j - \mu_i) = \prod_{i=1}^{j-1} (\mu_j - \mu_i) \cdot \prod_{i=j+1}^{n} (\mu_j - \mu_i),$$

che, attesa l'ipotesi dei μ_j tutti distinti, ha i primi $j-1$ fattori strettamente maggiori di 0, gli ultimi $n-j+1$ strettamente negativi. Quindi $p_A(\mu_j)$ e $g'(\mu_j)$ hanno segno opposto (o è $p_A(\mu_j) = 0$), e di conseguenza $\frac{p_A(\mu_j)}{g'(\mu_j)} \leq 0 \, \forall \, j = 1, \ldots, n-1$ come richiesto.
Infine, nel caso i μ_j non siano tutti distinti, per esempio $\mu_1 = \mu_2 = \ldots = \mu_k < \mu_{k+1}$ si ha $\lambda_2 = \lambda_3 = \ldots = \lambda_k = \mu_1$, e quindi $(t-\lambda_1) \cdot (t-\mu_1)^{k-1} | p_A(t)$, $(t-\mu_1)^k || g(t)$ (dove il simbolo $||$ significa 'divide esattamente'). Dividendo quindi $p_A(t), g(t)$ e $r(t)$ per $(t-\mu_1)^{k-1}$ e ripetendo tale procedimento per tutti i μ_j non distinti ci si riconduce al caso precedente.

Passo 2
Per il Teorema spettrale si ha una decomposizione di A_1 in $A_1 = Q \cdot D \cdot Q^T$, con Q matrice reale ortogonale e $D = diag([\mu_1 \ldots \mu_{n-1}])$, $\mu_i \in \mathbb{R}$. Per il *Passo 1* esistono $a \in \mathbb{R}, \underline{b}_0 \in \mathbb{R}^{n-1}$ tali che:

$$A_0 = \left(\begin{array}{c|c} a & \underline{b}_0^T \\ \hline \underline{b}_0 & D \end{array} \right)$$

con $\underline{\lambda}(A_0) = \underline{\lambda}$. Basta allora prendere

$$A = \left(\begin{array}{c|c} 1 & \underline{0}^T \\ \hline \underline{0} & Q \end{array} \right) \cdot A_0 \cdot \left(\begin{array}{c|c} 1 & \underline{0}^T \\ \hline \underline{0} & Q^T \end{array} \right) =$$

$$= \left(\begin{array}{c|c} 1 & \underline{0}^T \\ \hline \underline{0} & Q \end{array} \right) \cdot \left(\begin{array}{c|c} a & \underline{b}_0^T \\ \hline \underline{b}_0 & D \end{array} \right) \cdot \left(\begin{array}{c|c} 1 & \underline{0}^T \\ \hline \underline{0} & Q^T \end{array} \right) = \left(\begin{array}{c|c} a & \underline{b}_0^T Q^T \\ \hline Q\,\underline{b}_0 & QDQ^T \end{array} \right) =$$

$$= \left(\begin{array}{c|c} a & \underline{b}^T \\ \hline \underline{b} & A_1 \end{array} \right),$$

dove si è posto $\underline{b} = Q\,\underline{b}_0$. La matrice A così definita è reale simmetrica e si ha $\underline{\lambda}(A) = \underline{\lambda}$ in quanto A è simile ad A_0 per costruzione.

$\square$

Possiamo ora vedere la dimostrazione del Teorema di Horn:
Dim (Teorema di Horn):
A meno di moltiplicare per matrici di permutazioni, possiamo supporre di avere gli elementi dei vettori $\underline{d} = [d_i], \underline{\lambda} = [\lambda_i]$ ordinati in modo non decrescente, cioè $d_i \leq d_j\,, \lambda_i \leq \lambda_j$ se $i < j$. Quindi, per induzione su n:
Caso $n = 1$*:* l'asserto è banalmente vero.
$n - 1 \Rightarrow n$*:* per il lemma combinatorio dell'intreccio, esiste $\underline{\mu} \in \mathbb{R}^{n-1}$ vettore soddisfacente a (9) e quindi, per ipotesi induttiva, esiste $A_1 \in M_{n-1}[\mathbb{R}]^{n-1}$ simmetrica tale che $\underline{\lambda}(A_1) = \mu$ e $\underline{d}(A_1) = [d_1 d_2 \ldots d_{n-1}]^T$. Per il Teorema inverso dell'intreccio esiste allora

$$A = \left(\begin{array}{c|c} a & \underline{b}^T \\ \hline \underline{b} & A_1 \end{array}\right)$$

matrice reale simmetrica di ordine n con $\underline{\lambda}(A) = \underline{\lambda}$; inoltre $\underline{d}(A_1) = [d_1 d_2 \ldots d_{n-1}]^T$ e

$$Tr(A) = Tr(A_1) + a = \sum_{i=1}^{n-1} d_i + a = \sum_{i=1}^{n-1} \lambda_i = \sum_{i=1}^{n} d_i,$$

da cui $a = d_n$ e quindi la tesi.

$\square$

5 Il Teorema di Sing-Thompson: necessità

Dopo aver visto ai paragrafi 3 e 4 la dimostrazione del teorema di Schur-Horn, vediamo ora il teorema di Sing-Thompson; esso assomiglia al precedente, ma considera i valori singolari in luogo degli autovalori.
Nel seguito si userà come definizione di $\preceq$ quella data nella Definizione 2.3, in quanto i valori singolari, al contrario degli autovalori, vengono tradizionalmente scritti in ordine non crescente.

Teorema 5.1 (Il Teorema di Sing-Thompson) *Siano* $\underline{d} = [d_i] \in \mathbb{C}^n$, $\underline{s} = [s_i] \in \mathbb{R}^n$ *vettori ordinati in modo tale che si abbia* $s_1 \geq s_2 \geq \ldots \geq s_n$ *e* $|d_1| \geq |d_2| \geq \ldots \geq |d_n|$; *esiste* $A \in M_n[\mathbb{C}]$ *tale che* $\underline{s}(A) = \underline{s}$ *e* $\underline{d}(A) = \underline{d}$ *se e solo se*

$$\sum_{i=1}^{k} |d_i| \leq \sum_{i=1}^{k} s_i \quad \forall\, k = 1, 2, \ldots, n \tag{11}$$

$$\sum_{i=1}^{n-1} |d_i| - |d_n| \leq \sum_{i=1}^{n-1} s_i - s_n. \tag{12}$$

Osserviamo che se $|\underline{d}| \succeq \underline{s}$ allora (11) e (12) sono sicuramente verificate, ma non è detto il viceversa; ad esempio, se si prende $\underline{d} = [4,\, 3,\, 0.5]^T$ e $\underline{s} = [5,\, 3,\, 1]^T$ si ha che $|\underline{d}|$ e $\underline{s}$ non sono confrontabili con la relazione $\preceq$, mentre soddisfano (11) e (12).
Vediamo ora per prima cosa la dimostrazione della parte necessaria del teorema di Sing-Thompson; innanzitutto dimostriamo le disuguaglianze in (11).

Proposizione 5.2 *Sia* $A \in M_n[\mathbb{C}]$ *matrice tale che* $\underline{d}(A) = \underline{d} = [d_i]$ *e* $\underline{s}(A) = \underline{s} = [s_i]$; *allora si ha* $|\underline{d}| \succ \underline{s}$.

Dim.
Innanzitutto notiamo che, a meno di cogredienze (le quali lasciano inalterati i valori singolari), possiamo supporre $|d_1| \geq \ldots \geq |d_n|$. Sia $A = U\, diag([s_1 \ldots s_n])\, V$ con U, V una decomposizione in valori singolari di A, con $s_1 \geq \ldots \geq s_n$; sviluppando i conti per gli elementi diagonali di A si trova $d_i = \sum_{j=1}^{n} u_{ij} s_j v_{ji}$, ovvero

$$\underline{d} = W \cdot \underline{s},$$

ove si è posto $W = (u_{ij} v_{ji})$. Si ha inoltre che

$$|d_i| = |(W \cdot \underline{s})_i| = \left| \sum_{j=1}^{n} w_{ij} s_j \right| \leq \sum_{j=1}^{n} |w_{ij}| s_j = (|W| \cdot \underline{s})_i$$

(si intende $|W| = (|w_{ij}|)$); non è detto che gli elementi diagonali di $|W| \cdot \underline{s}$ siano ordinati in modo decrescente, ma posto $(|W| \cdot \underline{s})_{i_1} \geq \ldots \geq (|W| \cdot \underline{s})_{i_n}$, si ha:

$$|d_1| \leq (|W| \cdot \underline{s})_1 \leq (|W| \cdot \underline{s})_{i_1}$$

$$|d_1| + |d_2| \leq (|W| \cdot \underline{s})_1 + (|W| \cdot \underline{s})_2 \leq (|W| \cdot \underline{s})_{i_1} + (|W| \cdot \underline{s})_{i_2}$$

$$\vdots$$

$$|d_1| + \ldots + |d_{n-1}| \leq (|W| \cdot \underline{s})_1 + \ldots + (|W| \cdot \underline{s})_{n-1} \leq (|W| \cdot \underline{s})_{i_1} + \ldots + (|W| \cdot \underline{s})_{i_{n-1}}$$

$$|d_1| + \ldots + |d_n| \leq (|W| \cdot \underline{s})_1 + \ldots + (|W| \cdot \underline{s})_n = (|W| \cdot \underline{s})_{i_1} + \ldots + (|W| \cdot \underline{s})_{i_n}.$$

Da questo risulta $|\underline{d}| \succ |W| \cdot \underline{s} \succ \underline{s}$ per il Lemma 2.5 e il Lemma 2.6, quindi $|\underline{d}| \succ \underline{s}$ per la transitività di $\succ$.

$\square$

Per mostrare la necessità della disuguaglianza (12) abbiamo prima bisogno di vedere alcuni risultati preliminari.
D'ora in poi indicheremo con $\mathcal{M}(\mathrm{S})$ l'insieme delle matrici $A \in M_n[\mathbb{C}]$ tali che $A = U \cdot S \cdot V$, con $U, V \in M_n[\mathbb{C}]$ matrici unitarie, $S = diag([s_1 \ldots s_n])$ con $s_1 \ldots s_n$ sono numeri reali fissati ≥ 0.

Proposizione 5.3 $\mathcal{M}(\mathrm{S})$ *è un compatto di* $M_n[\mathbb{C}]$.

Dim.
Si ha che $\mathcal{M}(\mathrm{S})$ è limitato, in quanto se $M \in \mathcal{M}(\mathrm{S})$ allora $|||M|||_{\mathcal{F}} = \left(\sum_{i=1}^n s_i^2\right)^{1/2}$. Vediamo ora che $\mathcal{M}(\mathrm{S})$ è chiuso mostrando che il suo complementare (in $M_n[\mathbb{C}]$) $\mathcal{S}$ è intorno di ogni suo punto. Infatti, sia $B \in \mathcal{S}$ e $B = U' \, diag([s'_1 \ldots s'_n]) \, V'$ una sua decomposizione in valori singolari; poiché $B \notin \mathcal{M}(\mathrm{S})$ esiste k tale che $s'_k \neq s_k$. Per la Proposizione 2.19 si ha poi

$$|||M - B|||_{\mathcal{F}} \geq \left(\sum_{i=1}^{n} (s_i - s'_i)^2 \right)^{1/2} \geq |s_k - s'_k| > 0$$

per ogni matrice $M \in \mathcal{M}(\mathrm{S})$; prendendo quindi la palla di centro B e raggio $\epsilon = \frac{|s_k - s'_k|}{2}$ si ottiene un intorno di B contenuto in $\mathcal{S}$.

$\square$

Le seguenti due proposizioni confrontano i valori singolari e gli elementi diagonali di una matrice 2×2 e servono per la base dell'induzione del caso generale.

Proposizione 5.4 *Sia* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *matrice* 2×2 *con* $a, d \in \mathbb{R}$, $a, d \geq 0$, *valori singolari* $s_1 \geq s_2$. *Si ha allora* $s_1 + s_2 \geq a + d$, *dove vale l'uguaglianza se e solo se* A *è Hermitiana semidefinita positiva.*

Dim. La disuguaglianza cercata $s_1+s_2 \geq a+d$ è equivalente (poiché entrambi i membri sono maggiori o uguali a 0 per ipotesi) a $s_1^2+s_2^2+2s_1s_2 \geq a^2+d^2+2ad$, e quindi per il Lemma 2.2 a

$$|b|^2 + |c|^2 + 2|ad - bc| \geq 2ad.$$

Si ha poi dalla disuguaglianza triangolare

$$|b|^2+|c|^2+2|ad-bc| \geq |b|^2+|c|^2+2(ad-|bc|) = (|b|-|c|)^2+2ad \geq 2ad. \quad (13)$$

Questo prova che vale $s_1 + s_2 \geq a + d$, e l'uguaglianza vale se e solo se le disuguaglianze in (13) sono tutte uguaglianze, ovvero se e solo se $|b| = |c|$ e $|ad - bc| = ad - |bc|$; è un facile esercizio sui numeri complessi verificare che questo succede se e solo se bc è reale non negativo, cioè, atteso $|b| = |c|$, se e solo se $b = \overline{c}$. Pertanto ciò equivale a dire che A è Hermitiana e $ad - bc \geq 0$ e quindi ($a, d \geq 0$ per ipotesi) A semidefinita positiva.

□

Proposizione 5.5 *Sia* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *matrice* 2×2 *con* $a, d \in \mathbb{R}$, $a \geq |d|$, $a \geq 0 \geq d$ *e valori singolari* $s_1 \geq s_2$. *Allora si ha* $s_1 - s_2 \geq a - |d|$, *dove vale l'uguaglianza se e solo se* A *è Hermitiana nel caso in cui* $a > 0 > d$ *e se e solo se* $|b| = |c|$ *nel caso in cui* $ad = 0$.

Dim. La disuguaglianza cercata $s_1 - s_2 \geq a - |d|$ è equivalente (poiché entrambi i membri sono maggiori o uguali a 0 per ipotesi) a $s_1^2+s_2^2-2s_1s_2 \geq a^2 + d^2 - 2a|d|$, e quindi per il Lemma 2.2 a

$$|b|^2 + |c|^2 - 2|ad - bc| \geq -2a|d|.$$

Si ha poi dalla disuguaglianza triangolare

$$|b|^2+|c|^2-2|ad-bc| \geq |b|^2+|c|^2-2(a|d|+|bc|) = (|b|-|c|)^2-2a|d| \geq -2a|d|. \quad (14)$$

Questo prova che vale $s_1 - s_2 \geq a - |d|$, e l'uguaglianza vale se e solo se le disuguaglianze in (14) sono tutte uguaglianze, ovvero se e solo se $|b| = |c|$ e $|ad-bc| = ad+|bc|$; come nella Proposizione 5.4 è un facile esercizio verificare che questo accade se e solo se A è Hermitiana nel caso sia $a > 0 > d$.

$\square$

Indicheremo da qui in poi con f la funzione $f : \mathcal{M}(\mathrm{S}) \to \mathbb{R}$ definita da $f(A) = \sum_{i=1}^{n-1} |a_{ii}| - |a_{nn}|$, dove gli a_{ii} sono gli elementi diagonali di A. Osserviamo che f ha massimo in virtù della Proposizione 5.3.

Lemma 5.6 *Sia* $M = (m_{ij}) \in \mathcal{M}(\mathrm{S})$ *matrice massimizzante* f *con elementi diagonali* $m_{ii} \in \mathbb{R}$, $m_{11} \geq \ldots \geq m_{n-1,n-1} \geq 0 \geq m_{nn}$, $|m_{nn}| \leq m_{n-1,n-1}$. *Per ogni* $1 \leq p < q \leq n$ *sia* $M' = \begin{pmatrix} m_{pp} & m_{pq} \\ m_{qp} & m_{qq} \end{pmatrix}$ *sottomatrice* 2×2 *di* M; *si ha allora:*
1) $m_{pq} = \overline{m}_{qp}$ *se* $q \leq n-1$;
2) $|m_{pn}| = |m_{np}|$;
3) se $m_{nn} < 0 < m_{pp}$ *vale in particolare* $m_{pn} = \overline{m}_{np}$.

Dim. Consideriamo una decomposizione in valori singolari di M', $M' = U'\, diag([s'_1 s'_2])\, V'$; si ha quindi $diag([s'_1 s'_2]) = U'^H \cdot M' \cdot V'^H$ e, detti u'_{ij}, v'_{ij} gli elementi rispettivamente di U' e V' con indici in $\{p,q\}$, consideriamo la matrice $X = (x_{ij})$ con $x_{ij} = \overline{u}'_{ji}$ se $i,j \in \{p,q\}$, $x_{ij} = \delta_{ij}$ altrimenti. Analogamente prendiamo $Y = (y_{ij})$ con $y_{ij} = \overline{v}'_{ji}$ se $i,j \in \{p,q\}$, $y_{ij} = \delta_{ij}$ altrimenti; le matrici X, Y sono unitarie perché lo sono U', V', e considerando $N = (n_{ij}) = X \cdot M \cdot Y$, si ha che N è una matrice con diagonale $n_{ii} = m_{ii}$ (se $i \neq p,q$), $n_{pp} = s'_1, n_{qq} = s'_2$. Poiché M è una matrice che massimizza f deve'essere $\sum_{i=1}^{n-1} n_{ii} - |n_{nn}| \leq \sum_{i=1}^{n-1} m_{ii} - |m_{nn}|$. Consideriamo ora i tre casi, $q \leq n-1$, $q = n$, $q = n$ e $m_{nn} < 0 < m_{pp}$:
1) sia $q \leq n-1$; allora

$$\sum_{i=1}^{n-1} n_{ii} - |n_{nn}| = \sum_{i=1, i\neq p,q}^{n-1} m_{ii} + s'_1 + s'_2 - |m_{nn}|,$$

e quindi dev'essere $m_{pp} + m_{qq} \geq s'_1 + s'_2$. Dalla Proposizione 5.4 dev'essere anche $m_{pp} + m_{qq} \leq s'_1 + s'_2$, e quindi si ha $m_{pp} + m_{qq} = s'_1 + s'_2$; sempre la Proposizione 5.4 assicura allora che risulta $m_{pq} = \overline{m}_{qp}$.
2) Sia ora invece $q = n$; allora

$$\sum_{i=1}^{n-1} n_{ii} - n_{nn} = \sum_{i=1, i\neq p}^{n-1} m_{ii} + s'_1 - s'_2,$$

e quindi dev'essere $m_{pp} - |m_{nn}| \geq s'_1 - s'_2$. Dalla Proposizione 5.5 dev'essere anche $m_{pp} - |m_{nn}| \leq s'_1 - s'_2$, e quindi si ha $m_{pp} - |m_{nn}| = s'_1 - s'_2$; sempre la Proposizione 5.5 assicura allora che è $|m_{pn}| = |m_{np}|$.
3) Se in particolare $m_{nn} < 0 < m_{pp}$ risulta $m_{pn} = \overline{m}_{np}$ per la Proposizione 5.5.

$\square$

Lemma 5.7 *Sia* $M = (m_{ij}) \in \mathcal{M}(\mathrm{S})$ *matrice massimizzante* f *con elementi diagonali* $m_{ii} \in \mathbb{R}$, $m_{11} \geq \ldots \geq m_{n-1,n-1} \geq m_{nn} = 0$, *e tale che* $|m_{pn}| = |m_{np}| \neq 0$ *per qualche* p, $1 \leq p \leq n-1$. *Sia* $M' = \begin{pmatrix} m_{pp} & m_{pn} \\ m_{np} & 0 \end{pmatrix}$ *sottomatrice* 2×2 *di* M; *allora esiste una matrice* $N \in \mathcal{M}(\mathrm{S})$ *che massimizza* f *con diagonale* $n_{ii} = m_{ii}$ *(per* $i \neq p, n$*) e* $n_{pp} > 0 > n_{nn}$.

Dim. Consideriamo una decomposizione in valori singolari di M', $M' = U' \, diag([s'_1 s'_2]) \, V'$, $s'_1 \geq s'_2$, e poiché M' è non singolare (infatti il suo determinante è diverso da 0) si ha $s'_2 > 0$. Si ha poi $diag([s'_1 s'_2]) = U'^H \cdot M' \cdot V'^H$ e, detti u'_{ij}, v'_{ij} gli elementi rispettivamente di U' e V' con indici in $\{p, n\}$, consideriamo la matrice $X = (x_{ij})$ con $x_{ij} = \overline{u}'_{ji}$ se $i, j \in \{p, n\}$, $x_{ij} = \delta_{ij}$ altrimenti. Analogamente prendiamo $Y = (y_{ij})$ con $y_{ij} = \overline{v}'_{ji}$ se $i, j \in \{p, n\}$, $y_{ij} = \delta_{ij}$ altrimenti; le matrici X, Y sono unitarie perché lo sono U', V', e considerando $N = (n_{ij}) = E \cdot X \cdot M \cdot Y$ (dove $E = (\underline{e}_1 \ldots \underline{e}_{n-1} - \underline{e}_n)$), si ha che $N \in \mathcal{M}(\mathrm{S})$ e che i suoi elementi diagonali sono $n_{ii} = m_{ii}$ (se $i \neq p, n$), $n_{pp} = s'_1$, $n_{nn} = -s'_2$. Per la Proposizione 5.5 dev'essere $s'_1 - s'_2 \geq m_{pp}$ da cui

$$f(N) = \sum_{i=1}^{n-1} n_{ii} - |n_{nn}| = \sum_{i=1, i \neq p}^{n-1} m_{ii} + s'_1 - s'_2 \geq \sum_{i=1, i \neq p}^{n-1} m_{ii} + m_{pp} = f(M),$$

e quindi anche N massimizza f.

$\square$

Dal Lemma 5.6 e dal Lemma 5.7 segue il seguente corollario:

Corollario 5.8 *Se esiste una matrice* $A \in \mathcal{M}(\mathrm{S})$ *che massimizza* f *con elementi diagonali reali di cui i primi* $n-1$ *non negativi e l'ultimo* ≤ 0, *allora esiste una matrice* $B \in \mathcal{M}(\mathrm{S})$ *che massimizza ancora* f, B *Hermitiana e che verifica in particolare uno di questi due casi:*
1) i suoi primi $n-1$ *elementi diagonali sono* > 0 *e l'ultimo* < 0;
2) i suoi primi $n-1$ *elementi diagonali sono* ≥ 0, *l'ultima riga e l'ultima colonna nulle.*

Dim.
Sia $A = (a_{ij})$ matrice come nell'ipotesi; dev'essere $a_{ii} \geq |a_{nn}|$ altrimenti una cogredienza di A incrementerebbe il valore di f (si ricorda che le matrici di permutazione sono unitarie). Possiamo quindi supporre senza perdita di generalità $a_{11} \geq \ldots \geq a_{n-1,n-1} \geq 0 \geq a_{nn}$. Per ogni p tale che $1 \leq p < n$

applichiamo ora il Lemma 5.6 alla sottomatrice $A' = \begin{pmatrix} a_{pp} & a_{pn} \\ a_{np} & a_{nn} \end{pmatrix}$; questo mostra che è $|a_{pn}| = |a_{np}|$. Mostriamo ora che se $a_{nn} = 0$ ed esiste un indice $p < n$ per cui $a_{pn} \neq 0$, allora esiste una matrice $B \in \mathcal{M}(\mathrm{S})$ che massimizza f con $b_{nn} < 0$ (e quindi ragionando come sopra dev'essere anche $b_{11} \geq \ldots \geq b_{n-1,n-1} > 0$); in questo caso infatti la sottomatrice A' diventa $\begin{pmatrix} a_{pp} & a_{pn} \\ a_{np} & 0 \end{pmatrix}$, e grazie al Lemma 5.7 riusciamo ad ottenere una matrice B che massimizza f tale che $b_{11} \geq \ldots \geq b_{n-1,n-1} > 0 > b_{nn}$. Possiamo quindi assumere senza perdita di generalità che sia o $b_{11} \geq \ldots \geq b_{n-1,n-1} > 0 > b_{nn}$ (ovvero una matrice come in 1)) o $b_{nn} = 0 = b_{1n} = \ldots = b_{n-1,n} = b_{n1} = \ldots = b_{n,n-1}$ (ovvero una matrice come in 2)). In entrambi i casi vale $b_{np} = \overline{b}_{pn}$, nel primo caso per il Lemma 5.6, banalmente nel secondo. Per ogni p, q tali che $1 \leq p < q < n$ applichiamo ora il Lemma 5.6 alla matrice B considerando la sua sottomatrice $B'' = \begin{pmatrix} b_{pp} & b_{pq} \\ b_{qp} & b_{qq} \end{pmatrix}$; otteniamo così che dev'essere $b_{pq} = \overline{b}_{qp}\ \forall\, 1 \leq p < q \leq n-1$, ovvero B Hermitiana.

$\square$

Proposizione 5.9 *Si ha che* $max_{A \in \mathcal{M}(\mathrm{S})} f(A) = \sum_{i=1}^{n-1} s_i - s_n$.

Dim.
Notiamo innanzitutto che se si prende $S = diag([s_1 \ldots s_n])$, scegliendo cioè $U = V = \mathbb{I}_n$, si ha $f(S) = \sum_{i=1}^{n-1} s_i - s_n$; basta quindi provare che vale $f(C) \leq \sum_{i=1}^{n-1} s_i - s_n$ per ogni $C \in \mathcal{M}(\mathrm{S})$.
Sia ora $A = (a_{ij})$ una matrice che massimizzi f. Sia C la matrice ottenuta premoltiplicando A per la matrice unitaria $U = (u_{ij})$ così definita: $u_{ii} = \frac{\overline{a}_{ii}}{|a_{ii}|}$ se $a_{ii} \neq 0$, $u_{ij} = \delta_{ij}$ altrimenti. Si ha quindi che C ha valori diagonali reali e $c_{ii} \geq 0\, \forall\, i = 1, \ldots, n-1$; moltiplicando eventualmente l'ultima riga per -1 si ha anche $c_{nn} \leq 0$. Inoltre C massimizza ancora f perché $|c_{ii}| = |a_{ii}|$ per ogni $i = 1, \ldots, n$. Applicando ora il Corollario 5.8 alla matrice C otteniamo una matrice $B \in \mathcal{M}(\mathrm{S})$ Hermitiana che massimizza f, tale che $b_{ii} \geq 0$ per $i = 1, \ldots, n-1$ e $b_{nn} \leq 0$. Poiché gli s_i sono per costruzione i valori singolari delle matrici $M \in \mathcal{M}(\mathrm{S})$, quindi in particolare di B, essi sono i moduli degli autovalori λ_i della matrice B, ovvero $s_i = \pm\lambda_i$; inoltre non tutti gli autovalori di B sono > 0, in quanto B non è definita positiva ($b_{nn} \leq 0$). Sia $|\lambda_1| = s_1 \geq \ldots \geq |\lambda_n| = s_n$; si hanno allora due casi:
1) $s_n = 0 \Leftrightarrow \lambda_n = 0$, da cui

$$f(B) = \sum_{1=1}^{n-1} b_{ii} - |b_{nn}| = \sum_{1=1}^{n} b_{ii} = Tr(B) = \sum_{i=1}^{n-1} \lambda_i \leq \sum_{i=1}^{n-1} |\lambda_i| = \sum_{i=1}^{n-1} s_i,$$

come voluto;
2) $|\lambda_n| = s_n \neq 0$, e in questo caso esiste un indice k per cui $\lambda_k < 0$, da cui

$$f(B) = \sum_{1=1}^{n-1} b_{ii} - |b_{nn}| = \sum_{1=1}^{n} b_{ii} = Tr(B) = \sum_{i=1, i\neq k}^{n} \lambda_i + \lambda_k =$$

$$= \sum_{i=1, i\neq k}^{n} \lambda_i - |\lambda_k| \leq \sum_{i=1}^{n-1} s_i - s_n,$$

poiché $|\lambda_k| = s_k \geq s_n$, e questo termina la dimostrazione.

□

Siamo ora pronti per vedere la dimostrazione della necessità del Teorema di Sing-Thompson:
Dim.
Sia $A = (a_{ij}) \in M_n[\mathbb{C}]$ matrice tale che $\underline{s}(A) = \underline{s}$ e $\underline{d}(A) = \underline{a}$; le condizioni (11) sono allora date dalla Proposizione 5.2, mentre la condizione (12) è data dalla Proposizione 5.9; infatti se A ammette come valori singolari $s_1, \ldots, s_n$ allora esistono $U, V \in M_n[\mathbb{C}]$ unitarie tali che $A = U\, diag([s_1 \ldots s_n])\, V$ con $s_1 \geq \ldots \geq s_n \geq 0$. Si ha quindi

$$\sum_{i=1}^{n-1} |a_{ii}| - |a_{nn}| \leq max_{C \in \mathcal{M}(\mathrm{S})} f(C) = \sum_{i=1}^{n-1} s_i - s_n.$$

□

6 Il Teorema di Sing-Thompson: sufficienza

Si è vista al paragrafo 5 la dimostrazione della necessità delle condizioni (11), (12) perché esista una matrice $A \in M_n[\mathbb{C}]$ con valori singolari $\underline{s}$ e diagonale $\underline{d}$. Vedremo ora la sufficienza; verrà qui presentata una dimostrazione costruttiva di tale teorema che unisce quella data da M.T. Chu nel 1999 (vedi [1]) a quella data da Sing nel 1976 (vedi [7]). L'idea di Chu si basa sulla dimostrazione originale per induzione data da Thompson nel 1977 e si ispira a [9]. Per convenienza denoteremo la matrice diagonale che ha per elementi diagonali gli stessi di una matrice M con $diag(M)$.

Dimostreremo ora la sufficienza del teorema nel caso 2×2; useremo poi tale caso per costruire una matrice $n \times n$.

Il caso 2×2
Le condizioni (11), (12) per l'esistenza di una matrtice A con valori singolari e diagonale preassegnati diventano, nel caso 2×2 :

$$\begin{aligned} s_1 &\geq s_2 && (15)\\ |d_1| &\geq |d_2| && (16)\\ |d_1| &\leq s_1 && (17)\\ |d_1| + |d_2| &\leq s_1 + s_2 && (18)\\ |d_1| - |d_2| &\leq s_1 - s_2 && (19) \end{aligned}$$

Si noti che la disuguaglianza (17) è conseguenza di (18) e (19).
Per trovare una matrice A con valori singolari $s_1 \geq s_2 \geq 0$ ed elementi diagonali d_1, d_2 basta trovare due matrici unitarie $U\,, V \in M_2[\mathbb{C}]$ che soddisfino l'uguaglianza

$$diag(U \cdot S \cdot V^H) = diag([d_1 d_2]), \tag{20}$$

dove $S = diag([s_1 s_2])$.
Notiamo inoltre che possiamo limitarci al caso in cui sia $d_1 \geq d_2 \geq 0$ numeri reali, in quanto se così non fosse basterebbe trovare A che abbia come elementi diagonali $|d_1|, |d_2|$ e premoltiplicare A per la matrice unitaria $W = \begin{pmatrix} \frac{d_1}{|d_1|} & 0 \\ 0 & \frac{d_2}{|d_2|} \end{pmatrix}$ se $d_2 \neq 0$ (o sostituendo 1 a $\frac{d_i}{|d_i|}$ se $d_i = 0$, $i = 1, 2$).
È noto che le matrici del tipo $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ con $\theta \in \mathbb{R}$ sono ortogonali; quindi per dimostrare il teorema nel caso 2×2 è sufficiente trovare due matrici

$$U = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \; V^T = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}$$

che soddisfino (20), ovvero basta trovare θ, ϕ tali che

$$\begin{cases} \cos\theta\cos\phi\, s_1 - \sin\theta\sin\phi\, s_2 &= d_1 \\ -\sin\theta\sin\phi\, s_1 + \cos\theta\cos\phi\, s_2 &= d_2. \end{cases}$$

Questo sistema è equivalente al sistema

$$\begin{cases} (\cos\theta\cos\phi - \sin\theta\sin\phi)s_1 + (\cos\theta\cos\phi - \sin\theta\sin\phi)s_2 &= d_1 + d_2 \\ (\cos\theta\cos\phi + \sin\theta\sin\phi)s_1 - (\cos\theta\cos\phi + \sin\theta\sin\phi)s_2 &= d_1 - d_2 \end{cases},$$

ed usando le formule di addizione e sottrazione del coseno si ottiene

$$\begin{cases} \cos(\theta+\phi)(s_1+s_2) &= d_1 + d_2 \\ \cos(\theta-\phi)(s_1-s_2) &= d_1 - d_2. \end{cases} \tag{21}$$

Se si ha $s_1 + s_2 = 0$ allora dev'essere $s_1 = s_2 = 0$ e quindi per (18) anche $d_1 = d_2 = 0$; basta quindi prendere $A = \mathbb{O}$. Se invece è $s_1 + s_2 \neq 0$ e $s_1 = s_2$, allora per (16) e (19) dev'essere anche $d_1 = d_2$, per cui la seconda equazione del sistema è sempre soddisfatta e basta quindi trovare θ, ϕ che soddisfino la prima, ovvero θ, ϕ tali che

$$\cos(\theta+\phi) = \frac{d_1 + d_2}{s_1 + s_2}.$$

Da (18) si ha $\frac{d_1+d_2}{s_1+s_2} \leq 1$ e quindi l'equazione ha soluzione; di più, si può scegliere θ ad arbitrio e restano di conseguenza determinati due valori (se $\frac{d_1+d_2}{s_1+s_2} < 1$) o un valore (se $d_1 + d_2 = s_1 + s_2$) di ϕ (ovviamente a meno di multipli di 2π). Se infine è $s_1 + s_2, s_1 - s_2 > 0$, il sistema (21) è equivalente a

$$\begin{cases} \cos(\theta+\phi) &= \frac{d_1+d_2}{s_1+s_2} \\ \cos(\theta-\phi) &= \frac{d_1-d_2}{s_1-s_2} \end{cases},$$

che ha soluzione poiché $\frac{d_1+d_2}{s_1+s_2} \leq 1$ per (18) e $\frac{d_1-d_2}{s_1-s_2} \leq 1$ per (19).

In Figura 1 è mostrata graficamente la regione di piano in cui possono variare d_1, d_2 (nel caso reale), fissati $s_1 = s_2 = s$ (s_1, d_1 sono rappresentati in ascissa, s_2, d_2 in ordinata). Analogamente in Figura 2 e Figura 3, in cui però si considerano rispettivamente il caso $s_1 > s_2 = 0$ e $s_1 > s_2 > 0$.

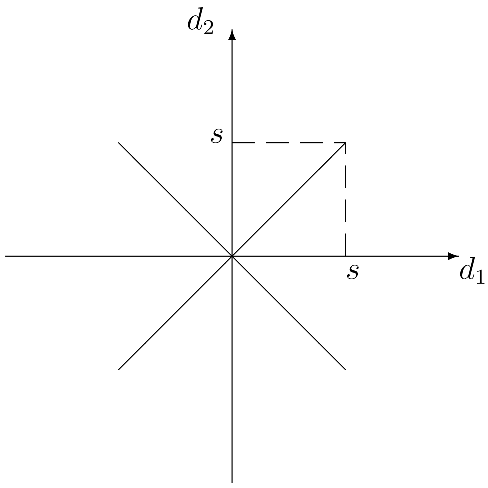

Figura 1: caso ‘banale’ in cui $s_1 = s_2 = s > 0$:

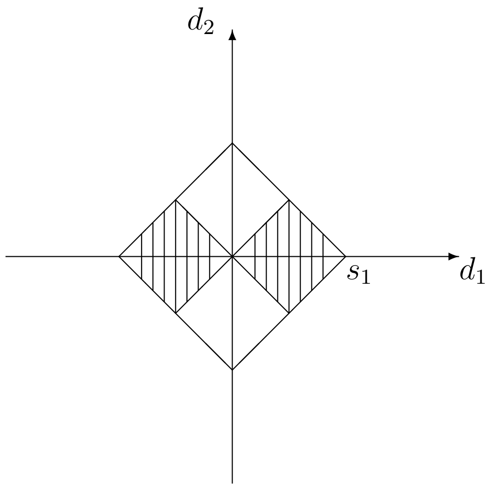

Figura 2: caso ‘banale’ in cui $s_1 > s_2 = 0$

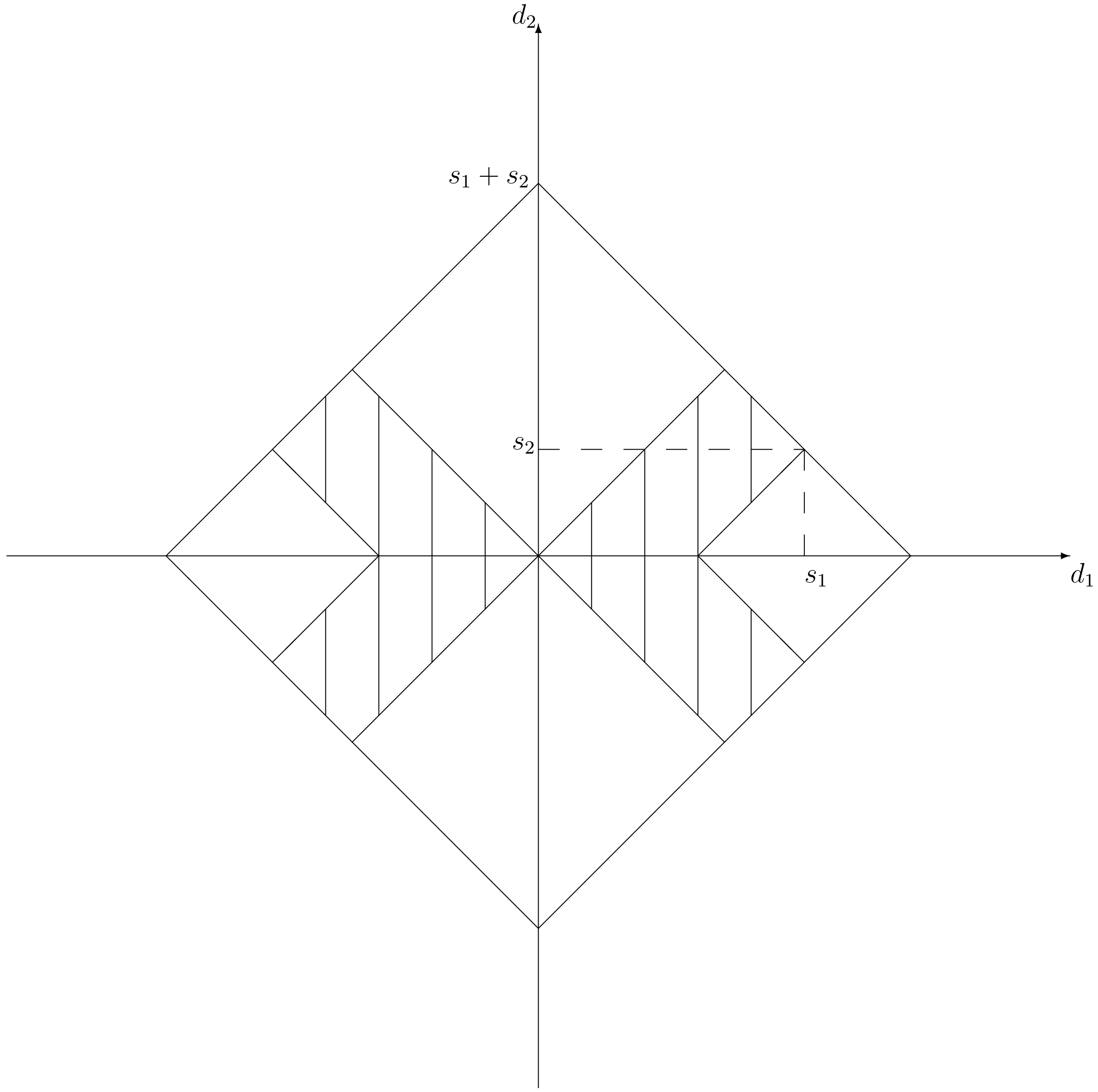

Figura 3: caso ‘generale’ in cui $s_1 > s_2 > 0$

Il caso $n \times n$

La dimostrazione della sufficienza nel caso $n \times n$ è una dimostrazione per induzione sull'ordine n della matrice a partire dal caso 2×2.

Siano $\underline{d}$ e $\underline{s}$ vettori che soddisfino le condizioni (11), (12) del teorema di Sing-Thompson. Supponiamo quindi $n \geq 3$ ed il teorema vero per $n-1$; sia

$$k = \max\{i : |d_1| \leq s_j,\ j = 1, \ldots, i\}; \tag{22}$$

si ha $k \geq 1$ per (11) e si noti che potrebbe essere anche $k = n$; risulta quindi

$$s_1 \geq s_2 \geq \ldots \geq s_k \geq |d_1| > s_{k-1} \geq \ldots \geq s_n$$

oppure

$$s_1 \geq s_2 \ldots \geq s_n \geq |d_1|.$$

Consideriamo ora i due casi, a seconda che sia $k \leq n-2$ oppure $k \in \{n-1, n\}$.

Caso 1: supponiamo $k \leq n-2$. Sia

$$t = s_k + s_{k+1} - |d_1| \tag{23}$$

Si ha $t \geq 0$ in quanto $s_k - |d_1| \geq 0$ per come è definito k. Procediamo ora per passi:

Passo 1. *Proviamo che* $||d_1| - t| \leq s_k - s_{k+1}$.

Infatti, supponiamo sia $|d_1| - t \geq 0$. Si ha

$$\underbrace{|d_1| - s_k}_{\leq 0} - s_{k+1} + |d_1| \leq -s_{k+1} + |d_1| \leq s_k - s_{k+1}.$$

Il caso $|d_1| - t \leq 0$ è analogo:

$$s_k + \underbrace{s_{k+1} - |d_1|}_{<0} - |d_1| < s_k - |d_1| \leq s_k - s_{k+1}.$$

Passo 2. *Le coppie* $\{d_1, t\}$ *e* $\{s_k, s_{k+1}\}$ *soddisfano le condizioni (15)-(19) del caso* 2×2*, per cui esistono* $\widetilde{U}$, $\widetilde{V} \in M_2[\mathbb{C}]$ *matrici unitarie tali che:*

$$diag\left(\widetilde{U}\, diag([s_k\, s_{k+1}])\, \widetilde{V}^H\right) = diag([d_1\, t]).$$

Si noti che se $|d_1| - t \leq 0$, t fa le veci di d_1 e d_1 quelle di t.

Passo 3. *Mostriamo che i due insiemi ordinati, costituiti ciascuno da* $n-1$ *numeri,*

$$s_1 \geq s_2 \geq \ldots \geq s_{k-1} > t \geq s_{k+2} \geq \ldots \geq s_n,$$

$$|d_2| \geq |d_3| \geq \ldots \geq |d_n|,$$

soddifano le condizioni di Sing-Thompson.
Innanzitutto verifichiamo che è proprio $s_{k-1} > t \geq s_{k+2}$; infatti

$$s_{k-1} \geq s_k > s_k + \underbrace{s_{k+1} - |d_1|}_{<0} = t = s_{k+1} + \underbrace{s_k - |d_1|}_{\geq 0} \geq s_{k+1} \geq s_{k+2}.$$

Dopodiché verifichiamo che i due insiemi di numeri soddisfano a (11), verificando prima le sommatorie con indice che varia da 1 fino a k; vogliamo cioè mostrare che

$$\sum_{i=2}^{l} |d_i| \leq \sum_{i=1}^{l-1} s_i \; \forall\, l = 1, \ldots, k;$$

questo è vero in quanto

$$\sum_{i=2}^{l} |d_i| \leq \sum_{i=1}^{l-1} |d_i| \leq \sum_{i=1}^{l-1} s_i \; \forall\, l = 1, \ldots, k.$$

Mostriamo ora che soddisfano a (11) con indice che varia fra $k+1$ e n, verifichiamo cioè

$$\sum_{i=2}^{l} |d_i| \leq \sum_{i=1}^{k-1} s_i + t + \sum_{i=k+2}^{l} s_i \; \forall\, l = k+1, \ldots, n;$$

questo vale poiché

$$\sum_{i=2}^{l} |d_i| \leq \sum_{i=1}^{k-1} s_i + \underbrace{s_k + s_{k+1} - |d_1|}_{=t} + \sum_{i=k+2}^{l} s_i \Leftrightarrow$$

$$\Leftrightarrow \sum_{i=1}^{l} |d_i| \leq \sum_{i=1}^{l} s_i \;\; \forall\, l = k+1, \ldots, n.$$

Infine mostriamo che è soddisfatta anche (12), ovvero che

$$\sum_{i=2}^{n-1} |d_i| - |d_n| \leq \sum_{i=1}^{k-1} s_i + t + \sum_{i=k+2}^{n-1} s_i - s_n;$$

questo segue da

$$\sum_{i=2}^{n-1} |d_i| - |d_n| \leq \sum_{i=1}^{k-1} s_i + \underbrace{s_k + s_{k+1} - |d_1|}_{=t} + \sum_{i=k+2}^{n-1} s_i - s_n \Leftrightarrow$$

$$\Leftrightarrow \sum_{i=1}^{n-1} |d_i| - |d_n| \leq \sum_{i=1}^{n-1} s_i - s_n$$

vero per ipotesi e perché $k < n-1$.

Passo 4. *Utilizziamo l'ipotesi induttiva per ottenere due matrici unitarie di ordine $n-1$ che serviranno per la costruzione delle matrici unitarie U, V di ordine n.*

Dal **Passo 3** segue per ipotesi induttiva che esistono $\widehat{U}\,, \widehat{V} \in M_{n-1}[\mathbb{C}]$ matrici unitarie tali che:

$$diag\left(\widehat{U}\, diag([s_1 \ldots s_{k-1}\, t\, s_{k+2} \ldots s_n])\, \widehat{V}^{\,H}\right) \;=\; diag([d_2 \ldots d_n]).$$

Consideriamo ora le seguenti decomposizioni a blocchi di $\widehat{U}, \widehat{V}$

$$\widehat{U} \;=\; \left(\begin{array}{c|c} \widehat{U}_{11} & \widehat{U}_{12} \\ \hline \widehat{U}_{21} & \widehat{U}_{22} \end{array}\right), \widehat{V} \;=\; \left(\begin{array}{c|c} \widehat{V}_{11} & \widehat{V}_{12} \\ \hline \widehat{V}_{21} & \widehat{V}_{22} \end{array}\right),$$

dove $\widehat{U}_{11}, \widehat{V}_{11} \in M_{k-1}\,[\mathbb{C}]\,.$

Posto $S_1 \;=\; \begin{pmatrix} s_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & s_{k-1} \end{pmatrix}, S_3 \;=\; \begin{pmatrix} t & 0 & 0 & 0 \\ 0 & s_{k+2} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & s_n \end{pmatrix}$ si ha

$$\left(\begin{array}{c|c} \widehat{U}_{11} & \widehat{U}_{12} \\ \hline \widehat{U}_{21} & \widehat{U}_{22} \end{array}\right) \cdot \begin{pmatrix} S_1 & 0 \\ 0 & S_3 \end{pmatrix} \cdot \left(\begin{array}{c|c} \widehat{V}_{11}^{\;H} & \widehat{V}_{21}^{\;H} \\ \hline \widehat{V}_{12}^{\;H} & \widehat{V}_{22}^{\;H} \end{array}\right) = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right),$$

dove

$$\begin{array}{rcl} A & = & \widehat{U}_{11} \cdot S_1 \cdot \widehat{V}_{11}^{\;H} + \widehat{U}_{12} \cdot S_3 \cdot \widehat{V}_{12}^{\;H}, \\ B & = & \widehat{U}_{11} \cdot S_1 \cdot \widehat{V}_{21}^{\;H} + \widehat{U}_{12} \cdot S_3 \cdot \widehat{V}_{22}^{\;H}, \\ C & = & \widehat{U}_{21} \cdot S_1 \cdot \widehat{V}_{11}^{\;H} + \widehat{U}_{22} \cdot S_3 \cdot \widehat{V}_{12}^{\;H}, \\ D & = & \widehat{U}_{21} \cdot S_1 \cdot \widehat{V}_{21}^{\;H} + \widehat{U}_{22} \cdot S_3 \cdot \widehat{V}_{22}^{\;H}. \end{array}$$

Per ipotesi induttiva le matrici A e D devono avere come elementi diagonali rispettivamente $d_2, \ldots, d_k$ e $d_{k+1}, \ldots, d_n$.

Passo 5. *Costruiamo le matrici unitarie U, V di ordine n a partire dalle matrici di ordine $n-1$ $\widehat{U}, \widehat{V}$ trovate al **Passo 4** e dalle matrici $\widetilde{U}$ e $\widetilde{V}$ trovate al **Passo 2**.*

Poniamo innanzitutto

$$U \;=\; \left(\begin{array}{c|c|c} \widehat{U}_{11} & \underline{0} & \widehat{U}_{12} \\ \hline \underline{0}^T & 1 & \underline{0}^T \\ \hline \widehat{U}_{21} & \underline{0} & \widehat{U}_{22} \end{array}\right) \cdot \begin{pmatrix} \mathbb{I}_{k-1} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \widetilde{U} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{I}_{n-k-1} \end{pmatrix},$$

$$V = \left(\begin{array}{c|c|c} \widehat{V}_{11} & \underline{0} & \widehat{V}_{12} \\ \hline \underline{0}^T & 1 & \underline{0}^T \\ \hline \widehat{V}_{21} & \underline{0} & \widehat{V}_{22} \end{array}\right) \cdot \begin{pmatrix} \mathbb{I}_{k-1} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \widetilde{V} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{I}_{n-k-1} \end{pmatrix},$$

$$S_2 = \begin{pmatrix} s_k & 0 \\ 0 & s_{k+1} \end{pmatrix}, \; S_3' = \begin{pmatrix} s_{k+2} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & s_n \end{pmatrix}, S = \left(\begin{array}{c|c|c} S_1 & 0 & 0 \\ \hline 0 & S_2 & 0 \\ \hline 0 & 0 & S_3' \end{array}\right).$$

Notiamo che U, V sono prodotti di due matrici decomposte a blocchi con blocchi di ordine diverso. Facciamo ora il prodotto del secondo fattore di U per la matrice S per il primo fattore di V^H; risulta

$$\begin{pmatrix} \mathbb{I}_{k-1} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \widetilde{U} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{I}_{n-k-1} \end{pmatrix} \cdot \begin{pmatrix} S_1 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & S_2 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & S_3' \end{pmatrix} \cdot \begin{pmatrix} \mathbb{I}_{k-1} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \widetilde{V}^H & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{I}_{n-k-1} \end{pmatrix} =$$

$$= \begin{pmatrix} S_1 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \widetilde{U} \cdot S_2 \cdot \widetilde{V}^H & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & S_3' \end{pmatrix} = \begin{pmatrix} S_1 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & D_2 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & S_3' \end{pmatrix},$$

dove $D_2 = \begin{pmatrix} d_1 & a \\ b & t \end{pmatrix}$ per opportuni $a, b \in \mathbb{C}$. Dall'uguaglianza precedente si deduce

$$U \cdot S \cdot V^H = \left(\begin{array}{c|c|c} \widehat{U}_{11} & \underline{0} & \widehat{U}_{12} \\ \hline \underline{0}^T & 1 & \underline{0}^T \\ \hline \widehat{U}_{21} & \underline{0} & \widehat{U}_{22} \end{array}\right) \cdot \left(\begin{array}{c|c|c} S_1 & 0 & 0 \\ \hline 0 & d_1 & \underline{a}^H \\ \hline 0 & \underline{b} & S_3 \end{array}\right) \cdot \left(\begin{array}{c|c|c} \widehat{V}_{11}^{\,H} & \underline{0} & \widehat{V}_{21}^{\,H} \\ \hline \underline{0}^T & 1 & \underline{0}^T \\ \hline \widehat{V}_{12}^{\,H} & \underline{0} & \widehat{V}_{22}^{\,H} \end{array}\right) =$$

$$= \left(\begin{array}{c|c|c} A & \widehat{U}_{12} \cdot \underline{b} & B \\ \hline (\widehat{V}_{12} \cdot \underline{a})^H & d_1 & (\widehat{V}_{22} \cdot \underline{a})^H \\ \hline C & \widehat{U}_{22} \cdot \underline{b} & D \end{array}\right),$$

dove si sono posti $\underline{a}, \underline{b} \in \mathbb{C}^{n-k-1}$, $\underline{a} = [\overline{a}\, 0 \dots 0]^T, \underline{b} = [b\, 0 \dots 0]^T$.
Poiché A, B, C, D sono le matrici ottenute induttivamente al **Passo 4**, si ha

$$diag(U\, diag([s_1 \dots s_n])\, V^H) = diag([d_2 \dots d_k\, d_1\, d_{k+1} \dots d_n]),$$

cioè la tesi.

Caso 2: Sia $n - 1 \leq k \leq n$.
Come si accorse Thompson, si ha allora che

$$\max \left\{ -s_{n-1} + s_n + |d_n|, 0, \sum_{i=1}^{n-1} |d_i| - \sum_{i=1}^{n-2} s_i \right\} \leq$$

$$\min \left\{ s_{n-1}, s_{n-1} + s_n - |d_n|, s_{n-1} - s_n + |d_n|, \sum_{i=1}^{n-2} s_i - \sum_{i=1}^{n-2} |d_i| + |d_{n-1}| \right\}.$$

Per mostrare questo basta mostrare che ciascuno degli elementi del primo insieme è minore o uguale di ogni elemento del secondo insieme; tali verifiche sono facili, e ne vengono qui dimostrate solo alcune come esemplificazione. Notiamo innanzitutto che si ha

$$|d_i| \leq |d_1| \leq |s_k| \leq |s_j| \,\forall\, i = 1, \ldots, n \,\forall\, j = 1, \ldots, n-1 \tag{24}$$

per come è stato definito k (e poiché stiamo supponendo $k \geq n-1$). Verifichiamo ora alcune delle disuguaglianze:

$$i) - s_{n-1} + s_n + |d_n| \leq s_{n-1} \Leftrightarrow \underbrace{(s_n - s_{n-1})}_{\leq 0} + \underbrace{(|d_n| - s_{n-1})}_{\leq 0\, per\ (24)} \leq 0.$$

$$ii) - s_{n-1} + s_n + |d_n| \leq \sum_{i=1}^{n-2} s_i - \sum_{i=1}^{n-2} |d_i| + |d_{n-1}| \Leftrightarrow \sum_{i=1}^{n-2} |d_i| - |d_{n-1}| + |d_n| \leq \sum_{i=1}^{n-1} s_i - s_n,$$

ma

$$\sum_{i=1}^{n-2} |d_i| - |d_{n-1}| + |d_n| \leq \sum_{i=1}^{n-2} |d_i| + |d_{n-1}| - |d_n| = \sum_{i=1}^{n-1} |d_i| - |d_n| \leq \sum_{i=1}^{n-1} s_i - s_n$$

per ipotesi.

$$iii) \sum_{i=1}^{n-1} |d_i| - \sum_{i=1}^{n-2} s_i \leq \sum_{i=1}^{n-2} s_i - \sum_{i=1}^{n-2} |d_i| + |d_{n-1}| \Leftrightarrow 2 \left(\sum_{i=1}^{n-2} |d_i| - \sum_{i=1}^{n-2} s_i \right) \leq 0 \Leftrightarrow$$

$$\Leftrightarrow \sum_{i=1}^{n-2} |d_i| \leq \sum_{i=1}^{n-2} s_i,$$

vero per ipotesi.

Le altre seguono da facili passaggi algebrici e da (24).

Sia ora $t \in \mathbb{R}$ compreso fra il massimo del primo insieme e il minimo del secondo; allora è facile vedere che le coppie $\{t, d_n\}, \{s_{n-1}\,, s_n\}$ soddisfano alle ipotesi del teorema di Sing-Thompson caso 2×2, qualunque sia il maggiore fra t e d_n (verifichiamo per esempio la (19): se $t < |d_n|$ allora $|d_n| - t \leq s_{n-1} - s_n \Leftrightarrow t \geq -s_{n-1} + s_n + |d_n|$, vero per come è stato scelto t; se invece $t \geq |d_n|$ si ha $t - |d_n| \leq s_{n-1} - s_n \Leftrightarrow t \leq s_{n-1} - s_n + |d_n|$, ancora vero per la scelta di t). Esistono quindi due matrici unitarie $\widetilde{U}, \widetilde{V} \in M_2[\mathbb{C}]$ tali che

$$diag(\widetilde{U}\, diag([s_{n-1}\, s_n])\, \widetilde{V}^{\,H}) = diag([t\, d_n]);$$

si ha inoltre che i due insiemi ordinati, costituiti ciascuno da $n-1$ numeri,

$$s_1 \geq s_2 \geq \ldots \geq s_{n-2} \geq t,$$

$$|d_1| \geq |d_2| \geq \ldots \geq |d_{n-1}|,$$

soddisfano le condizioni di Sing-Thompson.
Infatti $\sum_{i=1}^{l} |d_i| \leq \sum_{i=1}^{l} s_i \; \forall \, l = 1, \ldots, n-2$ poiché $|d_i| \leq s_i$ per (24); $\sum_{i=1}^{n-1} |d_i| \leq \sum_{i=1}^{n-2} s_i + t \Leftrightarrow t \geq \sum_{i=1}^{n-1} |d_i| - \sum_{i=1}^{n-2} s_i$, vero per la scelta di t.
Infine $\sum_{i=1}^{n-2} |d_i| - |d_{n-1}| \leq \sum_{i=1}^{n-2} s_i - t \Leftrightarrow t \leq \sum_{i=1}^{n-2} s_i - \sum_{i=1}^{n-2} |d_i| + |d_{n-1}|$, vero per la scelta di t.
Per ipotesi induttiva segue quindi che esistono due matrici unitarie $\widehat{U}, \widehat{V} \in M_{n-1}[\mathbb{C}]$ tali che

$$diag(\widehat{U}\, diag([s_1 \ldots s_{n-2}\, t])\, \widehat{V}^{\,H}) = diag([d_1 \ldots d_{n-1}]).$$

Definendo quindi

$$U = \left(\begin{array}{c|c} \widehat{U} & 0 \\ \hline 0 & 1 \end{array}\right) \cdot \left(\begin{array}{c|c} \mathbb{I}_{n-2} & 0 \\ \hline 0 & \widetilde{U} \end{array}\right), \; V = \left(\begin{array}{c|c} \widehat{V} & 0 \\ \hline 0 & 1 \end{array}\right) \cdot \left(\begin{array}{c|c} \mathbb{I}_{n-2} & 0 \\ \hline 0 & \widetilde{V} \end{array}\right),$$

si verifica in modo analogo al *Caso 1* (ed anzi più facilmente!) che

$$diag(U\, diag([s_1 \ldots s_n])\, V^H) = diag([d_1 d_n]),$$

e questo termina la dimostrazione.

$\square$

Riferimenti bibliografici

[1] M. T. Chu, On constructing matrices with prescribed singular values and diagonal elements, *Linear Algebra and its Applications 288*, 1999, pp. 11-22.

[2] R. A. Horn, Doubly stochastic matrices and the diagonal of a rotation matrix, *Amer. J. Math. 76*, 1954, pp. 620-630.

[3] R. A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.

[4] R. A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, New York, 1991.

[5] L. Mirsky, Inequalities and existence theorems in the theory of matrices, *J. Math. Anal. Appl. 9*, 1964, pp. 99-118.

[6] I. Schur, *Ü*ber eine klasse von mittelbildungen mit anwendungen auf der determinantentheorie, *Sitzungsberichte der Berliner Mathematischen Gesellschaft 22*, 1923, pp. 9-20.

[7] F. Y. Sing, Some results on matrices with prescribed diagonal elements and singular values, *Canad. Math. Bull. 19*, 1976, pp. 89-92.

[8] R. C. Thompson, Singular values, diagonal elements and convexity, *SIAM J. Appl. Math. 32*, 1977, pp. 39-63.

[9] H. Zha and Z. Zhang, A note on constructing a symmetric matrix with specified diagonal entries and eigenvalues, *BIT 35*, 1995, pp. 448-452.