第24卷第4期
2004年11月
数学研究与评论
JOURNAL OF MATHEMATICAL RESEARCH AND EXPOSITION
Vol. 24 No. 4
Nov. 2004

关于矩阵和与矩阵积的特征值的关系*

席博彦[1]， 张晓明[2]

(1. 内蒙古民族大学数学系，内蒙古 通辽 028043; 2. 北京师范大学数学系，北京 100875)

摘　要：本文给出了两个矩阵积的迹与其特征值之间的若干等价关系，并推广到两个矩阵和的情形.

关键词：控制不等式；矩阵；特征值；迹.

分类号：AMS(2000) 15A18,15A42/CLC number: O151.21

文献标识码：A　　　**文章编号**：1000-341X(2004)04-0689-08

1 引　言

两个矩阵积的迹与其特征值和奇异值之间的不等式是有重要应用的一类不等式，又称为 Neumann 不等式，在许多文献中有多种不同的证明方法，但关于其等式成立的等价条件却讨论较少. 本文就这个问题，从等式成立的角度出发，利用控制不等式理论给出了相对应于 Hermite 矩阵、非负定矩阵及一般矩阵的一系列等价条件，其中包括了以往很少涉及到的矩阵和与矩阵积的特征值关系问题.

矩阵 $A=(a_{ij})\in C^{n\times n}$ 为 Hermite 矩阵时，记作 $A\in H^{n\times n}$，若 $A\in H^{n\times n}$ 为非负定矩阵时，记作 $A\geqslant 0$；具有实特征值的矩阵 $A\in C^{n\times n}$ 的特征值规定为 $\lambda_1(A)\geqslant\lambda_2(A)\geqslant\cdots\geqslant\lambda_n(A)$，且记 $\lambda(A)=(\lambda_1(A),\lambda_2(A),\cdots,\lambda_n(A))$，$\wedge(A)=\mathrm{diag}(\lambda_1(A),\lambda_2(A),\cdots,\lambda_n(A))$；对于一般 $A\in C^{n\times n}$，其特征值排列为 $|\lambda_1(A)|\geqslant|\lambda_2(A)|\geqslant\cdots\geqslant|\lambda_n(A)|$，并记 $|\lambda(A)|=(|\lambda_1(A)|,|\lambda_2(A)|,\cdots,|\lambda_n(A)|)$；矩阵 $A\in C^{n\times n}$ 的奇异值规定为 $\sigma_1(A)\geqslant\sigma_2(A)\geqslant\cdots\geqslant\sigma_n(A)$，且记 $\sigma(A)=(\sigma_1(A),\sigma_2(A),\cdots,\sigma_n(A))$，$\triangle(A)=\mathrm{diag}(\sigma_1(A),\sigma_2(A),\cdots,\sigma_n(A))$；又记 $d(A)=(a_{11},a_{22},\cdots,a_{nn})$；$n$ 阶酉矩阵全体的集合记作 $U^{n\times n}$；R^n,R^n_+ R^n_{++} 分别表示实分量，非负分量，正分量的行向量集合.

对于 $x,y\in R^n$，把 x,y 的分量重新按降序排列为 $x_{[1]}\geqslant\cdots\geqslant x_{[n]}$ 和 $y_{[1]}\geqslant\cdots\geqslant y_{[n]}$，并记为 $x\downarrow=(x_{[1]},x_{[2]},\cdots,x_{[n]})$，$y\downarrow=(y_{[1]},y_{[2]},\cdots,y_{[n]})$. 若满足 $\sum_{i=1}^{k}x_{[i]}\leqslant\sum_{i=1}^{k}y_{[i]}$，$k=1,\cdots,$

* 收稿日期：2002-01-15
基金项目：内蒙古自然科学基金(20001301)和内蒙古高校科学研究基金(NJ03169)资助项目.
作者简介：席博彦(1959-)，男，教授.

$n-1$，且 $\sum_{i=1}^{n} x_{[i]} = \sum_{i=1}^{n} y_{[i]}$，则称 x 被 y 所控制，记作 $x \prec y$. 如果 $x \prec y$，且 x 不是 y 的重排，则称 x 被 y 严格控制，记作 $x \prec\prec y$. 如果 $\sum_{i=1}^{k} x_{[i]} \leqslant \sum_{i=1}^{k} y_{[i]}, k=1,\cdots,n$，则称 x 被 y 下（弱）控制，记作 $x \prec_w y$. 相应地，对于 $x \in R^n$，把 x 的分量重新按增序排列为 $x_{(1)} \leqslant \cdots \leqslant x_{(n)}$，并记作 $x\uparrow = (x_{(1)}, x_{(2)}, \cdots, x_{(n)})$. 显然 $x_{[i]} = x_{(n-i+1)}, i=1,2,\cdots,n$.

对于 $x \in R_+^n$，记 $\ln x = (\ln x_1, \ln x_2, \cdots, \ln x_n)$，由于用于控制不等式，因此其中有零分量出现时，对数形式就按乘积形式理解. 对于 $x, y \in R^n$，记 $x \circ y = (x_1y_1, x_2y_2, \cdots, x_ny_n)$.

关于矩阵乘积的特征值、奇异值与迹之间的关系，有着著名的 Neumann 不等式：

设 $A, B \in H^{n\times n}$ 的特征值分别为 $\lambda(A), \lambda(B)$，则

$$\sum_{i=1}^{n} \lambda_i(A)\lambda_{n-i+1}(B) \leqslant \operatorname{tr}(AB) \leqslant \sum_{i=1}^{n} \lambda_i(A)\lambda_i(B). \tag{1}$$

设 $A, B \in C^{n\times n}$ 的奇异值分别为 $\sigma(A), \sigma(B)$，则

$$-\sum_{i=1}^{n} \sigma_i(A)\sigma_i(B) \leqslant \operatorname{Re}(\operatorname{tr}(AB)) \leqslant \sum_{i=1}^{n} \sigma_i(A)\sigma_i(B). \tag{2}$$

许多文献[1-6]中都有上述定理的证明，但对等式成立的条件却未见有清晰的证明. 例如有文献叙述为(1)右边等式成立 $\Leftrightarrow B = \sum_{i=1}^{n} \lambda_i(B)\varphi_i\varphi_i^*$，这里 $\varphi_1, \cdots, \varphi_n$ 为 A 的对应于 $\lambda_1(A), \cdots, \lambda_n(A)$ 的标准正交化特征向量.

可以看到这个叙述（包括其证明）是不够准确的. 例如对于 $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$，显然 $\varphi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \varphi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ 是 A 的对应于 $\lambda_1(A), \lambda_2(A)$ 的标准正交化特征向量，且 $\operatorname{tr}(AB) = \sum_{i=1}^{n} \lambda_i(A)\lambda_i(B)$，但 $B \neq \sum_{i=1}^{n} \lambda_i(B)\varphi_i\varphi_i^* \left(= \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \right)$.

我们认为可以叙述为：(1)右边等式成立 $\Leftrightarrow$ 存在对应于 $\lambda_1(A), \cdots, \lambda_n(A)$ 的标准正交化特征向量 $\varphi_1, \cdots, \varphi_n$，使得 $B = \sum_{i=1}^{n} \lambda_i(B)\varphi_i\varphi_i^*$，或者存在酉矩阵 $U \in U^{n\times n}$，使得

$$U^*AU = \operatorname{diag}(\lambda_1(A), \cdots, \lambda_n(A)), \quad U^*BU = \operatorname{diag}(\lambda_1(B), \cdots, \lambda_n(B)).$$

前面所述问题的所在是缺少存在量的证明. 为此本文将给出两种证明方法，同时给出了等式成立的一系列等价条件，特别是包含有与矩阵和的特征值关系问题.

2 主要结果

下面所要用到的控制不等式的一些基本性质，都可以从[7]和[8]中找到，为简便只列出[7]的页数.

引理 1 设 $x, y \in R_+^n$，$\prod_{i=1}^{k} x_{[i]} \leqslant \prod_{i=1}^{k} y_{[i]}, k=1,\cdots,n$，且 $\sum_{i=1}^{n} x_{[i]} = \sum_{i=1}^{n} y_{[i]}$，则 $x\downarrow = y\downarrow$.

证明 由[7, P. 51]可得 $\sum_{i=1}^{k} x_{[i]} \leqslant \sum_{i=1}^{k} y_{[i]}, k=1,\cdots,n$.

首先若 $y_{[m]}=0$,则明显有 $x_{[m]}=0$;又若 $x_{[m]}=0$,则由

$$\sum_{i=1}^{m-1}x_{[i]}=\sum_{i=1}^{n}x_{[i]}=\sum_{i=1}^{n}y_{[i]}\leqslant\sum_{i=1}^{m-1}y_{[i]},$$

得 $y_{[m]}=0$,故 $x_{[m]}=0$ 当且仅当 $y_{[m]}=0,1\leqslant m\leqslant n$.

不失一般性,可假设 $x,y\in R^n_{++}$,则问题转化为 $\ln x\prec_w \ln y, x\prec y\Rightarrow x\downarrow=y\downarrow$.

若 $x\downarrow\neq y\downarrow$,则由定义知 $x\prec\prec y$. 设 $g(t)=\ln t(t>0)$,显然 $g(t)$ 为$(0,+\infty)$上的严格凹函数,那么由[7,P.48]知,

$$x\prec\prec y\Rightarrow\sum_{i=1}^{n}\ln(x_i)>\sum_{i=1}^{n}\ln(y_i),$$

而这与 $\ln x\prec_w\ln y$ 相矛盾,因此 $x\downarrow=y\downarrow$. □

引理 2 设 $A=(a_{ij})\in C^{n\times n}$,则 $|a_{ij}|\leqslant\sigma_1(A)$. 进而,若 $|a_{ij}|=\sigma_1(A)$,则当 $s\neq i$ 时,有 $a_{sj}=0$,当 $t\neq j$ 时,有 $a_{it}=0$.

证明 因 A^*A 的主对角线元素为

$$d(A^*A)=(\sum_{i=1}^{n}|a_{i1}|^2,\sum_{i=1}^{n}|a_{i2}|^2,\cdots,\sum_{i=1}^{n}|a_{in}|^2).$$

由[7,P100]可知,当 $H\in H^{n\times n}$ 时,$d(H)\prec\lambda(H)$,那么,

$$\sigma_1(A)=(\lambda_1(A^*A))^{\frac{1}{2}}\geqslant\max_j\Big(\sum_{i=1}^{n}|a_{ij}^2|\Big)^{\frac{1}{2}}\geqslant|a_{ij}|.$$

并且,当 $|a_{ij}|=\sigma_1(A)$ 时,有 $|a_{ij}|^2=\sigma_1^2(A)=\lambda_1(A^*A)\geqslant\sum_{i=1}^{n}|a_{ij}|^2$,因此可得 $s\neq i$ 时,有 $a_{sj}=0$. 又由 $\lambda_1(A^*A)=\lambda_1(AA^*)$ 可证得当 $t\neq j$ 时,有 $a_{it}=0$. □

引理 3 设 $A\in C^{n\times n}$,则下列条件等价:

(i) $A\geqslant 0$.

(ii) $\lambda(A)=\sigma(A)$.

(iii) $\mathrm{tr}(A)=\sum_{i=1}^{n}\sigma_i(A)$.

证明 (i)与(ii)的等价性是明显的,(ii)⇒(iii)也是显然的. 下面只证(iii)⇒(ii). 由[7,P104]有 $\ln|\lambda(A)|\prec\ln\sigma(A),|\lambda(A)|\prec_w\sigma(A)$,于是

$$\sum_{i=1}^{n}\sigma_i(A)=\mathrm{tr}(A)=\sum_{i=1}^{n}\lambda_i(A)\leqslant\Big|\sum_{i=1}^{n}\lambda_i(A)\Big|\leqslant\sum_{i=1}^{n}|\lambda_i(A)|\leqslant\sum_{i=1}^{n}\sigma_i(A),$$

从而有 $\sum_{i=1}^{n}|\lambda_i(A)|=\sum_{i=1}^{n}\sigma_i(A)$,再由引理1即知 $|\lambda(A)|=\sigma(A)$. 又由 $\sum_{i=1}^{n}\lambda_i(A)=\sum_{i=1}^{n}|\lambda_i(A)|$ 可知 $\lambda(A)=|\lambda(A)|$,这表明 $\lambda(A)=\sigma(A)$. □

引理 4 设 $A,B\in C^{n\times n},A\geqslant 0,B\geqslant 0,U,V\in U^{n\times n}$,且 $A=U^*BV$,则 $A=U^*BU=V^*BV$.

证明 由于 $A\geqslant 0,B\geqslant 0$,所以 $A^2=AA^*=U^*BV\cdot V^*B^*U=U^*B^2U=(U^*BU)^2$,显然 $U^*BU\geqslant 0$. 因非负定 Hermite 矩阵的平方根是唯一的,所以 $A=U^*BU$.

同理可证,$A=V^*BV$. □

引理 5 设 $A,B\in C^{n\times n}$,则 $\sigma_1(AB)=\sigma_1(A)\sigma_1(B)$ 的充要条件是存在酉矩阵 $V,U,W\in$

$U^{n\times n}$,使得

$$V^*AU=\operatorname{diag}(\sigma_1(A),A_1),\ U^*BW=\operatorname{diag}(\sigma_1(B),B_1),\ A_1\geqslant 0,B_1\geqslant 0.$$

证明 充分性显然.下证必要性.

若 $\sigma_1(AB)=0$,则 A,B 中至少有一个是零矩阵,那么结论成立.若 $\sigma_1(AB)\neq 0$,由奇异值分解定理,存在酉矩阵 $V,P\in U^{n\times n}$,使得

$$V^*ABP=\triangle(AB).$$

设 $V=(V_1,V_2)$,其中 $V_1\in C^{n\times 1},V_2\in C^{n\times(n-1)},P=(P_1,P_2)$,其中 $P_1\in C^{n\times 1},P_2\in C^{n\times(n-1)}$,那么

$$V^*ABP=\begin{pmatrix}V_1^*\\V_2^*\end{pmatrix}AB(P_1,P_2)=\begin{pmatrix}V_1^*ABP_1 & V_1^*ABP_2\\ V_2^*ABP_1 & V_2^*ABP_2\end{pmatrix}=\begin{pmatrix}\sigma_1(AB) & 0\\ 0 & V_2^*ABP_2\end{pmatrix},$$

记 $d=\|BP_1\|_E$, 则 $d>0$,设 $Q_1=\dfrac{BP_1}{d}$,从而 Q_1 为单位列向量,把 Q_1 扩充为酉矩阵 $Q=(Q_1,Q_2)\in U^{n\times n}$,那么

$$V^*AQ=\begin{pmatrix}V_1^*AQ_1 & V_1^*AQ_2\\ V_2^*AQ_1 & V_2^*AQ_2\end{pmatrix}=\begin{pmatrix}\dfrac{V_1^*ABP_1}{d} & V_1^*AQ_2\\ V_2^*AQ_1 & V_2^*AQ_2\end{pmatrix},$$

设 $a=\dfrac{V_1^*ABP_1}{d}=\dfrac{\sigma_1(AB)}{d}>0$,即 $ad=\sigma_1(AB)$. 且有

$$Q^*BP=\begin{pmatrix}Q_1^*BP_1, & Q_1^*BP_2\\ Q_2^*BP_1, & Q_2^*BP_2\end{pmatrix}=\begin{pmatrix}d & *\\ 0 & Q_2^*BP_2\end{pmatrix},$$

根据引理 2 可知 $a\leqslant\sigma_1(A),d\leqslant\sigma_1(B)$,而 $\sigma_1(A)\sigma_1(B)\geqslant ad=\sigma_1(AB)=\sigma_1(A)\sigma_1(B)$.因此 $a=\sigma_1(A),d=\sigma_1(B)$,并且由引理 2 可知,它们所在的行,列的其它元素为零.此时已将 A,B 化为

$$V^*AQ=\operatorname{diag}(\sigma_1(A),A_2),\quad Q^*BP=\operatorname{diag}(\sigma_1(B),B_2),$$

先对 A_2 做极分解,使得 $A_2=A_1R_1$,其中 $A_1\geqslant 0,R_1\in U^{(n-1)\times(n-1)}$.设 $R=\operatorname{diag}(1,R_1^*)$ 时,显然 $R\in U^{n\times n}$,且

$$V^*AQR=\operatorname{diag}(\sigma_1(A),A_1),\ R^*Q^*BP=\operatorname{diag}(\sigma_1(B),B_3),$$

再对 B_3 做极分解,使得 $B_3=B_1H_1$,其中 $B_1\geqslant 0,H_1\in U^{(n-1)\times(n-1)}$.设 $H=\operatorname{diag}(1,H_1^*)$ 时,显然 $H\in U^{n\times n}$,且

$$R^*Q^*BPH=\operatorname{diag}(\sigma_1(B),B_1),$$

其中 $A_1\geqslant 0,B_1\geqslant 0$,且 $U=QR,W=PH$ 即为所求. □

推论 1 设 $A,B\in C^{n\times n},A\geqslant 0,B\geqslant 0$,则下列条件等价:

(i) $\sigma_1(AB)=\sigma_1(A)\sigma_1(B)$.

(ii) 存在酉矩阵 $U\in U^{n\times n}$,使得 $U^*AU=\operatorname{diag}(\sigma_1(A),A_1),U^*BU=\operatorname{diag}(\sigma_1(B),B_1)$.

(iii) $\sigma_1(A+B)=\sigma_1(A)+\sigma_1(B)$.

证明 (i)$\Leftrightarrow$(ii) 由引理 4 及引理 5 给出.

(ii)$\Rightarrow$(iii) 显然成立.

(iii)$\Rightarrow$(ii) 由于 $A+B\geqslant 0$,那么存在酉矩阵 $U\in U^{n\times n}$,使得

$$U^*(A+B)U=\triangle(A+B)=\wedge(A+B)=U^*AU+U^*BU,$$

设$(U^*AU)_{(1,1)}=a,(U^*BU)_{(1,1)}=b$,则由引理2知,有

$$\sigma_1(A+B)=a+b\leqslant\sigma_1(A)+\sigma_1(B)=\sigma_1(A+B),$$

即得$a=\sigma_1(A)$,且$b=\sigma_1(B)$.同时,有$U^*AU=\mathrm{diag}(\sigma_1(A),A_1)$,$U^*BU=\mathrm{diag}(\sigma_1(B),B_1)$.

□

定理1 设$A,B\in C^{n\times n}$,$A\geqslant 0$,$B\geqslant 0$,则下列条件等价:

(i) $\mathrm{tr}(AB)=\sum_{i=1}^{n}\lambda_i(A)\lambda_i(B)$.

(ii) $\sigma(AB)=\lambda(A)\circ\lambda(B)$.

(iii) 存在酉矩阵$U\in U^{n\times n}$,使得$U^*AU=\wedge(A)$,$U^*BU=\wedge(B)$.

(iv) $\lambda(A+B)=\lambda(A)+\lambda(B)$.

(v) $\lambda(AB)=\lambda(A)\circ\lambda(B)$.

证明 (i)$\Rightarrow$(ii) 从[7,P120]知,$\ln\sigma(AB)\prec\ln(\sigma(A)\circ\sigma(B))$,$\sigma(AB)\prec_w\sigma(A)\circ\sigma(B)$,于是

$$\sum_{i=1}^{n}\sigma_i(A)\sigma_i(B)=\sum_{i=1}^{n}\lambda_i(A)\lambda_i(B)=\mathrm{tr}(AB)=\sum_{i=1}^{n}\lambda_i(AB)$$
$$\leqslant\sum_{i=1}^{n}|\lambda_i(AB)|\leqslant\sum_{i=1}^{n}\sigma_i(AB)\leqslant\sum_{i=1}^{n}\sigma_i(A)\sigma_i(B),$$

因此,

$$\sum_{i=1}^{n}\sigma_i(AB)=\sum_{i=1}^{n}\sigma_i(A)\sigma_i(B),$$

再用引理1得,$\sigma(AB)=\sigma(A)\circ\sigma(B)=\lambda(A)\circ\lambda(B)$.

(ii),(iii),(iv)的等价性由推论1用简单的归纳法即得.(iii)$\Rightarrow$(v),(v)$\Rightarrow$(i)是明显的.

□

定理2 设$A,B\in H^{n\times n}$,则下列条件等价:

(i) $\mathrm{tr}(AB)=\sum_{i=1}^{n}\lambda_i(A)\lambda_i(B)$.

(ii) 存在酉矩阵$U\in U^{n\times n}$,使得$U^*AU=\wedge(A)$,$U^*BU=\wedge(B)$.

(iii) $\lambda(A+B)=\lambda(A)+\lambda(B)$.

(iv) $\lambda(AB)=(\lambda(A)\circ\lambda(B))\downarrow$.

证明 (i)$\Rightarrow$(ii) 取$a\geqslant\max\{-\lambda_n(A),-\lambda_n(B),0\}$,又设$\widetilde{A}=A+aI_n$,$\widetilde{B}=B+aI_n$,那么$\widetilde{A}\geqslant 0$,$\widetilde{B}\geqslant 0$,且$\lambda_i(\widetilde{A})=\lambda_i(A)+a$,$\lambda_i(\widetilde{B})=\lambda_i(\widetilde{B})+a$,$i=1,\cdots,n$,从而,有

$$\mathrm{tr}(\widetilde{A}\widetilde{B})=\mathrm{tr}(A+aI_n)(B+aI_n)=\mathrm{tr}(AB)+a\mathrm{tr}(A)+a\mathrm{tr}(B)+na^2$$
$$=\sum_{i=1}^{n}(\lambda_i(A)+a)(\lambda_i(B)+a)=\sum_{i=1}^{n}\lambda_i(\widetilde{A})\lambda_i(\widetilde{B}),$$

根据定理1可知,存在酉矩阵$U\in U^{n\times n}$,使得

$$U^*\widetilde{A}U=\wedge(A)+aI_n,\quad U^*\widetilde{B}U=\wedge(B)+aI_n,$$

而

$$U^*\widetilde{A}U = U^*AU + aI_n,\ U^*\widetilde{B}U = U^*BU + aI_n,$$

因此$U^*AU = \wedge(A)$,$U^*BU = \wedge(B)$,即证得(ii).

用类似的方法,借助$\widetilde{A}$,$\widetilde{B}$,使用定理1,可证明(iii)⇒(ii),(ii)⇒(iv),(iv)⇒(i),(ii)⇒(iii)都显然成立. □

推论 2　设$A,B \in H^{n\times n}$,则下列条件等价:

(i)　$\mathrm{tr}(AB) = \sum_{i=1}^{n}\lambda_i(A)\lambda_{n-i+1}(B)$.

(ii)　存在酉矩阵$U \in U^{n\times n}$,使得

$$U^*AU = \wedge(A), U^*BU = \mathrm{diag}(\lambda_n(B), \lambda_{n-1}(B), \cdots, \lambda_1(B)).$$

(iii)　$\lambda(A-B) = \lambda(A) - \lambda(B)\uparrow$.

(iv)　$\lambda(AB) = (\lambda(A) \circ \lambda(B)\uparrow)\downarrow$.

证明　根据特征值性质可知$\lambda(-B) = -\lambda(B)\uparrow$,即$-\lambda_i(-B) = \lambda_{n-i+1}(B)(i=1,2,\cdots,n)$,再由定理2即可得证. □

下面的是明显的推论.

推论 3　设$A,B \in H^{n\times n}$,若$\mathrm{tr}(AB) = \sum_{i=1}^{n}\lambda_i(A)\lambda_i(B)$,或$\mathrm{tr}(AB) = \sum_{i=1}^{n}\lambda_i(A)\lambda_{n-i+1}(B)$,则$AB = BA \in H^{n\times n}$.

定理 3　设$A,B \in C^{n\times n}$,则下列条件等价:

(i)　$\mathrm{Re}(\mathrm{tr}(AB)) = \sum_{i=1}^{n}\sigma_i(A)\sigma_i(B)$.

(ii)　$AB \geqslant 0$,且$\sigma(AB) = \sigma(A) \circ \sigma(B)$.

(iii)　$\lambda(AB) = \sigma(A) \circ \sigma(B)$.

(iv)　存在酉矩阵$V,U \in U^{n\times n}$,使得$V^*AU = \triangle(A)$,$U^*BV = \triangle(B)$.

证明 (i)⇒(ii) 由

$$\sum_{i=1}^{n}\sigma_i(A)\sigma_i(B) = \mathrm{Re}(\mathrm{tr}(AB)) \leqslant \sum_{i=1}^{n}|\lambda_i(AB)| \leqslant \sum_{i=1}^{n}\sigma_i(AB) \leqslant \sum_{i=1}^{n}\sigma_i(A)\sigma_i(B),$$

得

$$\mathrm{tr}(AB) = \sum_{i=1}^{n}\sigma_i(AB) = \sum_{i=1}^{n}\sigma_i(A)\sigma_i(B),$$

由引理3及定理1的证明即可证得(ii).

(ii)⇒(iv) 由$AB \geqslant 0$,存在酉矩阵$P \in U^{n\times n}$,使得$P^*ABP = \triangle(AB)$,由$\sigma(AB) = \sigma(A) \circ \sigma(B)$及引理5的证明过程,存在酉矩阵$Q \in U^{n\times n}$,使得$P^*AQ = \begin{pmatrix}\sigma_1(A) & 0 \\ 0 & A_2\end{pmatrix}$,$Q^*BP = \begin{pmatrix}\sigma_1(B) & 0 \\ 0 & B_2\end{pmatrix}$,继续即可证得(iv).

(iv)⇒(iii),(iii)⇒(i) 都是明显的. □

推论 4　设$A,B \in C^{n\times n}$,则下列条件等价:

(i)　$\mathrm{tr}(AB) = -\sum_{i=1}^{n}\sigma_i(A)\sigma_i(B)$.

(ii) $-AB \geqslant 0$,且 $\sigma(AB)=\sigma(A)\circ\sigma(B)$.

(iii) $\lambda(AB)\uparrow=-\sigma(A)\circ\sigma(B)$.

(iv) 存在酉矩阵 $V,U\in U^{n\times n}$,使得 $V^*AU=\triangle(A)$, $U^*BV=-\triangle(B)$.

证明 令 $C=-B$,根据定理 3 即可得证推论 4. □

3 关于(1)中等式成立等价条件的另一种证明方法

对于式(1)中等式成立的等价条件,已在定理 2 及推论 2 中证明,下面给出另一种证明方法.

设 $A,B\in H^{n\times n}$,则 $\mathrm{tr}(AB)=\sum_{i=1}^{n}\lambda_i(A)\lambda_i(B)$ 的充要条件是存在酉矩阵 $U\in U^{n\times n}$,使得 $U^*AU=\wedge(A)$, $U^*BU=\wedge(B)$.

证明 充分性是明显的,下证必要性.若 A 的特征值全相等,则结论显然成立.否则,记 A 的不同特征值为 $\mu_1>\mu_2>\cdots>\mu_k$,其中 μ_i 的重数为 $m_i(i=1,\cdots,k)$, $m_1+m_2+\cdots+m_k=n$,因 $A\in H^{n\times n}$,从而存在酉矩阵 $U\in U^{n\times n}$,使得

$$U^*AU=\wedge(A)=\mathrm{diag}(\mu_1I_{m_1},\mu_2I_{m_2},\cdots,\mu_kI_{m_k}),$$

这时设 $U^*BU=S=(s_{ij})=\begin{pmatrix}S_1 & T_1\\ T_1^* & B_1\end{pmatrix}$,其中 $S_1\in H^{m_1\times m_1}$, $B_1\in H^{(n-m_1)\times(n-m_1)}$.那么 $\lambda(B)=\lambda(S)$,且 $U^*ABU=\wedge(A)S$.由 $d(S)\prec\lambda(B)$,有 $\sum_{i=1}^{j}s_{ii}\leqslant\sum_{i=1}^{j}\lambda_i(B)$,因此 $\sum_{i=1}^{j}(\lambda_i(B)-s_{ii})\geqslant 0(j=1,2,\cdots,n)$.令 $t_i=\lambda_i(B)-s_{ii}(i=1,2,\cdots,n)$,则 $\sum_{i=1}^{j}t_i\geqslant 0(j=1,2,\cdots,n)$.由于 $\sum_{i=1}^{n}\lambda_i(A)\lambda_i(B)=\mathrm{tr}(AB)=\sum_{i=1}^{n}\lambda_i(A)s_{ii}$,所以

$$\sum_{i=1}^{n}\lambda_i(A)\lambda_i(B)-\sum_{i=1}^{n}\lambda_i(A)s_{ii}=\sum_{i=1}^{n}\lambda_i(A)t_i$$

$$=\sum_{j=1}^{n-1}(\lambda_j(A)-\lambda_{j+1}(A))\Big(\sum_{i=1}^{j}t_i\Big)+\lambda_n(A)\sum_{i=1}^{n}t_i=0,$$

因为 $\mu_2=\lambda_{m_1+1}(A)<\lambda_{m_1}(A)=\mu_1$ 即 $\lambda_{m_1}(A)-\lambda_{m_1+1}(A)>0$,所以 $t_1+t_2+\cdots+t_{m_1}=0$,即 $\mathrm{tr}(S_1)=\sum_{i=1}^{m_1}\lambda_i(B)$.由 Hermite 矩阵特征值交错定理知,有

$$\sum_{i=1}^{m_1}\lambda_i(B)=\mathrm{tr}(S_1)=\sum_{i=1}^{m_1}\lambda_i(S_1)\leqslant\sum_{i=1}^{m_1}\lambda_i(B),$$

而 $\lambda_i(S_1)\leqslant\lambda_i(B)(i=1,\cdots,m_1)$,因此 $\lambda_i(S_1)=\lambda_i(B)(i=1,\cdots,m_1)$.

下面证明 $T_1=0$.

$$\sum_{i=1}^{m_1}\lambda_i^2(B)=\sum_{i=1}^{m_1}\lambda_i(B^2)=\sum_{i=1}^{m_1}\lambda_i(S_1^2)=\mathrm{tr}(S_1S_1^*)$$

$$\leqslant\mathrm{tr}(S_1S_1^*+T_1T_1^*)\leqslant\sum_{i=1}^{m_1}\lambda_i(SS^*)=\sum_{i=1}^{m_1}\lambda_i^2(B),$$

所以上式均取等号，即 $\mathrm{tr}(T_1T_1^*)=0$，因此 $T_1=0$. 并且由 $\lambda_i(S_1)=\lambda_i(B)(i=1,\cdots,m_1)$ 可知，存在 m_1 阶酉矩阵 V_1，使得 $V_1^*S_1V_1=\mathrm{diag}(\lambda_1(B),\lambda_2(B),\cdots,\lambda_{m_1}(B))$，令 $U_1=\mathrm{diag}(V_1,I_{n-m_1})$，显然 $UU_1\in U^{n\times n}$，则

$$U_1^*U^*BUU_1=\mathrm{diag}(\lambda_1(B),\cdots,\lambda_{m_1}(B),B_1),$$

$$U_1^*U^*BUU_1=\mathrm{diag}(\mu_1I_{m_1},\mu_2I_{m_2},\cdots,\mu_kI_{m_k})=\Lambda(A).$$

下面再对 $S_2=B_1[1,\cdots,m_2|1,\cdots,m_2]$，求出相应的酉矩阵 V_2，以此类推，即可得证. □

参考文献：

[1] MARCUS M. *An eigenvalue inequality for the product of normal matrices* [J]. Amer. Math. Monthly, 1956, **63**: 173－174.

[2] MIRSKY L. *On the trace of matrix products* [J]. Math. Nachr., 1959, **20**: 171－174.

[3] MIRSKY L. *A trace inequality of John von Neumann* [J]. Monatsh. Math., 1975, **79**: 303－306.

[4] RICHTER H. *Zur Abschätzung von matrizen normen* [J]. Math. Nachr., 1958, **18**: 178－187.

[5] THEOBALD C M. *An inequality for the trace of the product of two symmetric matrices* [J]. Proc. Cambridge Philos. Soc., 1975, **77**: 265－267.

[6] 王松桂，贾忠贞. 矩阵论中的不等式 [M]. 合肥：安徽教育出版社，1994.
WANG Song-gui, JIA Zhong-zhen. *Inequalities in Matrices* [M]. Hefei: Anhui Education Press, 1994. (in Chinese)

[7] 王伯英. 控制不等式基础 [M]. 北京：北京师范大学出版社，1990.
WANG Bo-ying. *Introduction to Majorization and Matrix Inequalities* [M]. Beijing: Beijing Normal University Press, 1990. (in Chinese)

[8] MARSHALL A W, OIKIN I. *Inequalities: Theory of Majorization and its Applications* [M]. New York: Academic Press, 1979.

Relation between Eigenvalues of Matrix Sum and Matrix Product

XI Bo-yan[1], ZHANG Xiao-ming[2]

(1. Dept. of Math., Inner Mongolia University for Nationalities, Tongliao 028043, China;
2. Dept. of Math., Beijing Normal University, Beijing 100875, China)

Abstract: We prove some equivalence between the trace and the eigenvalue of the product of two matrices, then extend them to the case of the sum of two matrices.

Key words: majorization inequality; matrix; eigenvalue; trace.