国 防 科 技 大 学 学 报
JOURNAL OF NATIONAL UNIVERSITY OF DEFENSE TECHNOLOGY
第10卷第1期 1988年3月 Vol.10.No.1

关于正定矩阵的一类不等式

游 光 荣
（系统工程与应用数学系）

摘　要　本文考虑了常见的Hölder不等式和算术——几何平均不等式在矩阵中的类似，对R.Bellman提出的两个问题[1]给出了肯定的回答。同时建立了Cauchy—Schwarz不等式及其增补在矩阵中的类似。

关键词　不等式，矩阵，迹，特征值。

1. 引　　言

定义 1　对n阶实对称矩阵A、B，如果$A-B$是非负定矩阵，则记为$A\geqslant B$；如果$A-B$是正定的，记为$A>B$。

于是，$A>0$、$A\geqslant 0$分别表示A是正定阵和非负定阵。

定义 2　n阶矩阵A的主对角之和称为A的迹(trace)，记为$\mathrm{tr}A$。

引理 1　(1) 迹有线性性，即对任意常数λ、μ和n阶矩阵A、B，有

$$\mathrm{tr}(\lambda A+\mu B)=\lambda\mathrm{tr}A+\mu\mathrm{tr}B;$$

(2) $\mathrm{tr}A'=\mathrm{tr}A$，这里$A'$表示矩阵$A$的转置；

(3) 对任意n阶矩阵A、B，有$\mathrm{tr}(AB)=\mathrm{tr}(BA)$。

在第二届国际不等式会议上，R.Bellman再次提出了[1]：

定理 1　设A、B是n阶（实对称）正定矩阵，则

$$2\mathrm{tr}(AB)\leqslant\mathrm{tr}A^2+\mathrm{tr}B^2 \tag{1}$$

$$\mathrm{tr}(AB)\leqslant(\mathrm{tr}A^2)^{\frac{1}{2}}(\mathrm{tr}B^2)^{\frac{1}{2}} \tag{2}$$

其中，(1)式等号成立当且仅当$A=B$；(2)式等号成立当且仅当$A=kB$，这里k是一个常数。

R.Bellman称(2)式为Cauchy—Schwarz不等式在矩阵中的类似。[1]中借助(1)式证明了不等式(2)。实际上，由$(\sqrt{\mathrm{tr}A^2}-\sqrt{\mathrm{tr}B^2})^2\geqslant 0$，有

$$(\mathrm{tr}A^2)^{\frac{1}{2}}(\mathrm{tr}B^2)^{\frac{1}{2}}\leqslant\frac{\mathrm{tr}A^2+\mathrm{tr}B^2}{2}$$

于是，(1)、(2)可联立为：

1987年6月1日收到。作者为硕士生

$$\operatorname{tr}(AB)\leqslant(\operatorname{tr}A^2)^{\frac{1}{2}}\cdot(\operatorname{tr}B^2)^{\frac{1}{2}}\leqslant\frac{\operatorname{tr}A^2+\operatorname{tr}B^2}{2} \tag{3}$$

定理 1　可以推广至 n 阶直至无穷阶 Hermite 矩阵[2]。

R.Bellman 在[1]中最后提出了两个问题：(1) 是否存在算术——几何平均不等式在矩阵中的类似？(2) 定理 1 的结论是否对更高的乘方 (higher power) 成立？

2. 答 Bellman 的两个问题

本节将建立若干矩阵不等式，给R.Bellman 的上述问题一个肯定的回答。为此，首先引入下述定义和引理。

定义 3[3]　对实向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$和 $\boldsymbol{y}=(y_1,y_2,\cdots,y_n)$，如果可能重新排列分量，使得$x_{[1]}\geqslant x_{[2]}\geqslant\cdots\geqslant x_{[n]}$，$y_{[1]}\geqslant y_{[2]}\geqslant\cdots\geqslant y_{[n]}$，并且

$$\sum_{i=1}^{k}x_{[i]}\geqslant\sum_{i=1}^{k}y_{[i]},\ (k=1,2,\cdots,n-1);\quad \sum_{i=1}^{n}x_{[i]}=\sum_{i=1}^{n}y_{[i]}$$

则称向量$\boldsymbol{x}$优势于$\boldsymbol{y}$，记为$\boldsymbol{x}\succ\boldsymbol{y}$或$\boldsymbol{y}\prec\boldsymbol{x}$。

引理 2　(Schur)[3]设 n 阶实对称矩阵$A=(a_{ij})$的特征值是$\lambda_1,\lambda_2,\cdots,\lambda_n$，则

$$(\lambda_1,\lambda_2,\cdots,\lambda_n)\succ(a_{11},a_{22},\cdots,a_{nn}).$$

定义 4[3]　设$I\subset R$是一个区间，f 是 I^n 到R的一个函数，且对一切$\boldsymbol{x}$，$\boldsymbol{y}\in I^n$，由$\boldsymbol{x}\prec\boldsymbol{y}$，可以推出$f(\boldsymbol{x})\leqslant f(\boldsymbol{y})$，则称 f 是I^n上的Schur凸函数，简记为 s 凸函数。

引理 3[3]　若 f 是 $I^n\to R$的连续可微函数，则函数 f 在 I^n 上是 s 凸的充要条件是：

(1) f 是变量的对称函数；

(2) 对一切$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)\in I^n$，有

$$(x_i-x_j)\left(\frac{\partial f}{\partial x_i}-\frac{\partial f}{\partial x_j}\right)\geqslant 0$$

引理 4　设 $x_1,x_2,\cdots,x_m$；$\alpha_1,\alpha_2,\cdots,\alpha_m$ 是两组正实数，而且 $\alpha_1+\alpha_2+\cdots+\alpha_m=1$，则

$$x_1^{\alpha_1}x_2^{\alpha_2}\cdots x_m^{\alpha_m}\leqslant\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_mx_m.$$

定理 2　设A、B是 n 阶实对称正定矩阵，而且 $\frac{1}{p}+\frac{1}{q}=1$，$p$、$q>0$，则

$$\operatorname{tr}(AB)\leqslant(\operatorname{tr}A^p)^{\frac{1}{p}}(\operatorname{tr}B^q)^{\frac{1}{q}}\leqslant\frac{1}{p}\operatorname{tr}A^p+\frac{1}{q}\operatorname{tr}B^q \tag{4}$$

证明　根据引理 4 (令$m=2$)，即知(4)式右端成立。下面证(4)式左端成立。

因为A、B是 n 阶实对称正定矩阵，故存在正交阵T、S，使 $A=T'\operatorname{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)T$，$B=S'\operatorname{diag}(\mu_1,\mu_2,\cdots,\mu_n)S$，其中$\lambda_i,\mu_i>0$，分别是$A$、$B$的特征值。从而，对任意实数 p、q 有

$$A^p=T'\operatorname{diag}(\lambda_1^p,\lambda_2^p,\cdots,\lambda_n^p)T,\ B^q=S'\operatorname{diag}(\mu_1^q,\mu_2^q,\cdots,\mu_n^q)S,$$ 于是

$$\operatorname{tr}A^p=\sum_{i=1}^{n}\lambda_i^p,\qquad \operatorname{tr}B^q=\sum_{i=1}^{n}\mu_i^q, \tag{5}$$

另一方面，记$TBT'=(C_{ij})$，由于矩阵B在正交变换下特征值不变，根据引理2，有：

$$(\mu_1,\mu_2,\cdots,\mu_n)>(C_{11},C_{22},\cdots,C_{nn}),$$

根据引理3，易知$f(\boldsymbol{x})=\sum\limits_{i=1}^{n}x_i^q(q>1)$是$s$凸函数，根据定义，有

$$\sum_{i=1}^{n}\mu_i^q\geqslant\sum_{i=1}^{n}C_{ii}^q \tag{6}$$

而 $\operatorname{tr}(AB)=\operatorname{tr}[T'\operatorname{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)TB]=\operatorname{tr}[\operatorname{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)TBT']$

$$=\operatorname{tr}\left[\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}\begin{pmatrix}C_{11} & * & & * \\ * & C_{22} & & \\ & & \ddots & * \\ * & & * & C_{nn}\end{pmatrix}\right]=\sum_{i=1}^{n}\lambda_iC_{ii} \tag{7}$$

由Hölder 不等式：$\sum\limits_{i=1}^{n}\lambda_iC_{ii}\leqslant\left(\sum\limits_{i=1}^{n}\lambda_i^p\right)^{\frac{1}{p}}\left(\sum\limits_{i=1}^{n}C_{ii}^q\right)^{\frac{1}{q}}$，并结合(5)、(6)、(7)式，即得(4)式左端。 **证毕**

不难考察(4)式等号成立的充要条件。不等式(4)是 R.Bellman 得到的(3)式的一个推广，实际上也是对他提出的问题(2)的一个回答。

注意到$\operatorname{tr}(A\otimes B)=\operatorname{tr}A\cdot\operatorname{tr}B$，这里$A\otimes B$是矩阵$A$与$B$的Kronecker 积（或直积、张量积），而且函数$\left(\sum\limits_{i=1}^{n}x_i^p\right)^{\frac{1}{p}}(x_i>0,\ p>0)$是关于$p$单调下降的，于是有：

推论 2.1 假设同定理2，则

$$\operatorname{tr}(AB)\leqslant(\operatorname{tr}A^p)^{\frac{1}{p}}(\operatorname{tr}B^q)^{\frac{1}{q}}\leqslant\operatorname{tr}(A\otimes B) \tag{8}$$

引理 5[4] 实（复）数域上n维向量空间V的一组（有限或无限多个）两两可交换的线性变换在V中有一公共特征向量。

引理 6 对n阶实对称正定矩阵$A_1,A_2,\cdots,A_m(m\geqslant2)$，存在共同的正交矩阵$Q$，使$QA_iQ'$同时化为对角形的充要条件是任意$A_i$、$A_j$可交换即

$$A_iA_j=A_jA_i,\ i,j=1,2,\cdots,m。$$

必要性是显然的。对阶数n进行数学归纳并由引理5可证得充分性。具体过程从略。

必须指出，引理6中的Q同时能使$QA_iQ'=\operatorname{diag}(\lambda_{i1},\lambda_{i2},\cdots,\lambda_{in})$，其中$\lambda_{i1},\lambda_{i2},\cdots,\lambda_{in}$是$A_i$的特征值，但不能保证对角元按$\lambda_{i1}\geqslant\lambda_{i2}\geqslant\cdots\geqslant\lambda_{in}$排列。例如取$m=2$，$A_1$、$A_2$分别为对角阵$\left(\begin{smallmatrix}1&0\\0&2\end{smallmatrix}\right)$和$\left(\begin{smallmatrix}2&0\\0&1\end{smallmatrix}\right)$，即知不可能存在正交阵$Q$，使$QA_1Q'=QA_2Q'=\left(\begin{smallmatrix}2&0\\0&1\end{smallmatrix}\right)$。文[2]恰恰忽略了这一点。

定理 3 设$A_1,A_2,\cdots,A_m$是一组两两可交换的n阶实对称正定矩阵，而且

$$\frac{1}{r_1}+\frac{1}{r_2}+\cdots+\frac{1}{r_m}=1,\ r_i>0,\ i=1,2,\cdots,m$$

则

$$\operatorname{tr}(A_1A_2\cdots A_m)\leqslant\prod_{i=1}^{m}(\operatorname{tr}A_i^{r_i})^{\frac{1}{r_i}}\leqslant\sum_{i=1}^{m}\frac{1}{r_i}\operatorname{tr}A_i^{r_i} \tag{9}$$

证明：根据引理6，存在正交矩阵Q，使 $QA_iQ'=\operatorname{diag}(\lambda_{i1},\lambda_{i2},\cdots,\lambda_{in})$，从而

$$QA_1A_2\cdots A_mQ'=(QA_1Q')(QA_2Q')\cdots(QA_mQ')$$
$$=\mathrm{diag}(\lambda_{11},\lambda_{12},\cdots,\lambda_{1n})\cdot\mathrm{diag}(\lambda_{21},\lambda_{22},\cdots\lambda_{2n})\cdots\mathrm{diag}(\lambda_{m1},\lambda_{m2},\cdots,\lambda_{mn})$$
$$=\mathrm{diag}\left(\prod_{i=1}^{m}\lambda_{i1},\prod_{i=1}^{m}\lambda_{i2},\cdots,\prod_{i=1}^{m}\lambda_{in}\right)$$

于是

$$\mathrm{tr}(A_1A_2\cdots A_m)=\mathrm{tr}(QA_1A_2\cdots A_mQ')=\sum_{j=1}^{n}\prod_{i=1}^{m}\lambda_{ij}$$

另一方面

$$A_i^{r_i}=Q'\mathrm{diag}(\lambda_{i1}^{r_i},\ \lambda_{i2}^{r_i},\ \cdots,\ \lambda_{in}^{r_i})Q,\quad \mathrm{tr}A_i^{r_i}=\sum_{j=1}^{n}\lambda_{ij}^{r_j},\quad i=1,2,\cdots,m,$$

根据Hölder不等式，有

$$\sum_{j=1}^{n}\lambda_{1j}\lambda_{2j}\cdots\lambda_{mj}\leqslant\left(\sum_{j=1}^{n}\lambda_{1j}^{r_1}\right)^{\frac{1}{r_1}}\left(\sum_{j=1}^{n}\lambda_{2j}^{r_2}\right)^{\frac{1}{r_2}}\cdots\left(\sum_{j=1}^{n}\lambda_{mj}^{r_m}\right)^{\frac{1}{r_m}},$$

再由引理 4，有

$$\left(\sum_{j=1}^{n}\lambda_{1j}^{r_1}\right)^{\frac{1}{r_1}}\left(\sum_{j=1}^{n}\lambda_{2j}^{r_2}\right)^{\frac{1}{r_2}}\cdots\left(\sum_{j=1}^{n}\lambda_{mj}^{r_m}\right)^{\frac{1}{r_m}}\leqslant\sum_{i=1}^{m}\frac{1}{r_i}\mathrm{tr}A_i^{r_i}.$$

综合上述两式，即可推得不等式(9)。　　证毕

在(9)式中令$r_i=m$，得到：

推论 3.1　设$A_1,A_2,\cdots,A_m$是一组两两可交换的n阶实对称正定矩阵，则

$$\mathrm{tr}(A_1A_2\cdots A_m)\leqslant(\mathrm{tr}A_1^m)^{\frac{1}{m}}(\mathrm{tr}A_2^m)^{\frac{1}{m}}\cdots(\mathrm{tr}A_m^m)^{\frac{1}{m}}\leqslant\frac{1}{m}(\mathrm{tr}A_1^m+\mathrm{tr}A_2^m+\cdots+\mathrm{tr}A_m^m)$$

不难看出，推论3.1可视为对 R.Bellman 的两个问题的一个统一的回答。

减弱不等式(9)成立的条件是一个值得考虑的问题。

3. tr(AB)的上、下界估计

对常见的Cauchy-Schwarz 不等式，G.Polya 和 G.Szegö 曾得到一个增补即所谓的反问Schwarz 不等式：设 a_i，b_i 满足 $0<m_1\leqslant a_i\leqslant M_1$，$0<m_2\leqslant b_i\leqslant M_2$，$i=1,2,\cdots,n$，则

$$\frac{4m_1m_2M_1M_2}{(m_1m_2+M_1M_2)^2}\leqslant\frac{\left(\sum_{i=1}^{n}a_ib_i\right)^2}{\sum_{i=1}^{n}a_i^2\cdot\sum_{i=1}^{n}b_i^2}\leqslant1 \tag{10}$$

下面考虑此类不等式在矩阵中的类似。

定理 4　设A、B是n阶实对称正定矩阵，且$AB=BA$，λ_i、μ_i分别是A、B的特征值，而且$0<m_1\leqslant\lambda_i\leqslant M_1$，$0<m_2\leqslant\mu_i\leqslant M_2$，$i=1,2,\cdots,n$，则

$$\frac{m_1M_1\mathrm{tr}B^2+m_2M_2\mathrm{tr}A^2}{m_1m_2+M_1M_2}\leqslant\mathrm{tr}(AB)\leqslant(\mathrm{tr}A^2)^{\frac{1}{2}}(\mathrm{tr}B^2)^{\frac{1}{2}} \tag{11}$$

证明　上式右端显然成立。

根据引理 6，存在正交矩阵Q，使

$$QAQ'=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n),\quad QBQ'=\mathrm{diag}(\mu_1,\mu_2,\cdots,\mu_n)$$

注意到

$$M_2QAQ'-m_1QBQ'=\mathrm{diag}(\lambda_1M_2-m_1\mu_1,\cdots,\lambda_iM_2-m_1\mu_i,\cdots,\lambda_nM_2-m_1\mu_n)$$

其第 i 个主对角元为 $\lambda_iM_2-m_1\mu_i\geqslant0$，从而 $M_2QAQ'-m_1QBQ'\geqslant0$。

同理 $$M_1QBQ'-m_2QAQ'\geqslant0$$

于是 $$(M_2QAQ'-m_1QBQ')(M_1QBQ'-m_2QAQ')=\mathrm{diag}(v_1,v_2,\cdots,v_n)$$

其中 $$v_i=(\lambda_iM_2-\mu_im_1)(\mu_iM_1-\lambda_im_2)\geqslant0,\quad i=1,2,\cdots,n,$$

故 $$(M_2QAQ'-m_1QBQ')(M_1QBQ'-m_2QAQ')\geqslant0$$

整理，即得 $$M_1M_2QABQ'+m_1m_2QBAQ'\geqslant m_1M_1QB^2Q'+m_2M_2QA^2Q'$$

两端取迹，并由引理 1，即可推出

$$(m_1m_2+M_1M_2)\mathrm{tr}(AB)\geqslant m_1M_1\mathrm{tr}B^2+m_2M_2\mathrm{tr}A^2$$

这实际上就是(11)式左端。　　**证毕**

推论 4.1 假设同定理 4，则

$$\frac{4m_1m_2M_1M_2}{(m_1m_2+M_1M_2)^2}\leqslant\frac{[\mathrm{tr}(AB)]^2}{\mathrm{tr}A^2\cdot\mathrm{tr}B^2}\leqslant1 \tag{12}$$

证明是显然的，由$(\sqrt{m_1M_1\mathrm{tr}B^2}-\sqrt{m_2M_2\mathrm{tr}A^2})^2\geqslant0$，并根据(11)式即得(12)。

在不等式(11)、(12)中，m_1,m_2可分别取为$\min\limits_{1\leqslant i\leqslant n}\lambda_i$，$\min\limits_{1\leqslant i\leqslant n}\mu_i$，$M_1$、$M_2$可分别取为$\max\limits_{1\leqslant i\leqslant n}\lambda_i$，$\max\limits_{1\leqslant i\leqslant n}\mu_i$。这样，利用正定矩阵$A$、$B$的最大和最小特征值可以给出表达式$\dfrac{\mathrm{tr}(AB)}{(\mathrm{tr}A^2)^{\frac12}(\mathrm{tr}B^2)^{\frac12}}$的上、下界。而且，(11)、(12)式均可视为数值不等式(10)的推广和加强。因为令$A=\mathrm{diag}(a_1,a_2,\cdots,a_n)$，$B=\mathrm{diag}(b_1,b_2,\cdots,b_n)$，由(12)式立得不等式(10)。

我们称每个元素都非负的矩阵为非负矩阵。现根据非负矩阵（未必是正定的）的最大元和最小元给出$\mathrm{tr}(AB)$的下界，证明方法类似定理4，只需对非负矩阵(M_2A-m_1B)和(M_1B-m_2A)的乘积取迹即可。

定理 5 设 $A=(a_{ij})$、$B=(b_{ij})$为 n 阶非负矩阵，记 m_1，M_1分别为A的最小元和最大元，m_2、M_2分别为B的最小元和最大元，$M_1M_2\neq0$，则

$$\frac{m_1M_1\mathrm{tr}B^2+m_2M_2\mathrm{tr}A^2}{m_1m_2+M_1M_2}\leqslant\mathrm{tr}(AB)\leqslant(\mathrm{tr}A^2)^{\frac12}(\mathrm{tr}B^2)^{\frac12} \tag{13}$$

推论 5.1 假设同定理 5，则

$$\frac{4m_1m_2M_1M_2}{(m_1m_2+M_1M_2)^2}\leqslant\frac{[\mathrm{tr}(AB)]^2}{\mathrm{tr}A^2\cdot\mathrm{tr}B^2}\leqslant1 \tag{14}$$

同理，(13)、(14)式亦可视为数值不等式(10)的推广和加强。定理 4 和定理 5 的条件各有强弱，结论也不能互相包含。

4. 结　束　语

除定理 5，本文的全部结果均可推广至Hermite矩阵的情形，只需将“(实)正交方阵”改为“酉方阵”，“实对称方阵”改为“Hermite阵”，方阵的“转置”改为“共轭转置”即可。其中，推论3.1实际上加强了[2]中定理 2 。鉴于篇幅，本文不再列出有关结论。

致　　谢

刘德铭教授审阅了全文，谨此致谢。

参考文献

[1] Bellman, R.: Some Inequalities for Positive Definite Matrices, General Inequalities 2 (Proceeding of the Second International Confrence on General Inequalities, 1978), Edited by Beckenbach, E.F., Birkhäuser— Verlag—Basel — Boston —Stuttgart, 1980, pp.89—90

[2] 冯慈璜：关于矩阵迹的不等式，杭州大学学报（自然科学版），Vol.13,4(1986):418—420

[3] 方开泰：优势理论及其应用，应用数学与计算数学，5(1984):75—86

[4] 张禾瑞、郝炳新：高等代数，下册，第2版，人民教育出版社，pp.266

A Class of Inequalities for Positive Definite Matrices

You Guangrong

Abstract

In this paper, we first consider a matrix analogue of Hölder's inequality. Furthermore, we obtain a matrix analogue of the arithmetic—geometric—mean inequality. Then we give an affirmative answer to two open questions of R. Bellman.[1] After that, we obtain a matrix analogue of Cauchy-Schwarz inequality and its complementary (or reverse) inequalities. We derive a similar result for nonnegative matrices.

Key words Inequality, Matrix, Trace, Eigenvalue