$+\overline{B}_2$ from (3.3) where $\overline{B}_1 \in S_r$, $\overline{B}_2 \in S_{k-r}$. By $\mathrm{tr}(AB)^m = \mathrm{tr}(A^mB^m)$, we have $\mathrm{tr}(D_1\overline{B})^m = \mathrm{tr}(D_1^m\overline{B}^m)$. So,

$$\mathrm{tr}(D'\overline{B}_1)^m = \mathrm{tr}(D'^m\overline{B}_1^m)$$

and

$$\mathrm{tr}(D''\overline{B}_2)^m = \mathrm{tr}(D''^m\overline{B}_2^m).$$

Here

$$D = D' + D'',$$

corresponding with the block of $\overline{B}$. By the hypothesis,

$$D'\overline{B}_1 = \overline{B}_1D',\ D''\overline{B}_2 = \overline{B}_2D''.$$

We have $D_1\overline{B} = \overline{B}D_1$, that is $AB = BA$. □

References

[1] R. Bellman, *Introduction to Matrix Analysis*, Mcgraw-Hill Book Company, New York, 1970.

[2] Man Kam Kwong, *Some results on matrix monotone functions*, Linear Alg. Appl., **118** (1989), 129-153.

[3] R. A. Horn & C. R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, 1985, 464--471.

[4] Xu Changqin, *Bellman's Inequality*, Linear Alg. Appl., 229 (1995), 9-4.

Bellman 不等式(II)

徐常青

(中国科技大学数学系, 合肥 230026)

摘　要

本文考察不等式:

$$\mathrm{tr}(AB)^m \le \mathrm{tr}(A^mB^m),\ m = 1, 2, 3, \cdots,$$

其中A,B **为**K **阶方阵. 证明了当**A **正定**B **对称幂等条件下上述不等式成立. 还考察了**A,B **为非负矩阵时的情形.**

Journal of Mathematical Research & Exposition
Vol. 17, No. 1, 30-34, Feb. 1997

Bellman's Inequality (II)

Xu Changqing
(Dept. of Math., Univ. of Science and Technology of China, Hefei 230026)

Abstract This article concerns a conjecture:

$$\mathrm{tr}(AB)^m \le \mathrm{tr}(A^m B^m)$$

which was put forward by Bellman R. We prove it under the additional condition that $A \in S_k$ and $B \in I_k$. We also investigate the case when A, B are nonnegative.

Keywords definite positive matrices, idempotent matrices, trace.
Classification AMS (1991) 15A15, 15A42/CCL O151.21

1. Introduction

In [4], it is shown that if $A, B \in S_k$, **we have**

$$(\text{for any } m \in N):\ \mathrm{tr}(AB)^m \le \mathrm{tr}(A^m B^m) \tag{1.1}$$

for the case of $k=2$ and $k=3$ for all $m=1,2,3,\cdots$. We will continue to use the notations described in [4]. That is, $S_k(\widetilde{S}_k)$ denotes the set of all real symmetric definite positive matrices of order k: k is any member in N: I_k denotes the set of all $k\times k$ symmetric idempotent matrices; $\mathrm{tr}(X)$ denotes the trace of matrix X, and A^T for the transpose of matrix A. $\underline{K}$ for $1,2,\cdots,k$ for $k\in N$, where N is the set of all the integers. $A>B$ $(A\ge B)$ means $A-B\in S_k$. A_r and $A^{(1/n)}$ denotes the rth leading principal submatrix of A and $A^{(1/n)}$, respectively.

Our main result is

Theorem 1 *Let $A\in S_k$, $B\in I_k$, then (1.1) holds for all $m=1,2,3,\cdots$.*

We will use the following theorem 1' (theorem 1 in [4]) to prove some results

Theorem 1' *Suppose $A, B\in S_k$, set $A=UDU^t$, $D=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_k)$ and $\widetilde{B}=U^tBU=(b_{ij})$. Here σ, $\Phi_{\sigma,m}(\lambda)$ and b_σ are defined by*

$$\Phi_{\sigma,m}(\lambda)=\sum_\sigma(\lambda_{i1}^m+\lambda_{i2}^m+\cdots+\lambda_{im}^m)/m-\lambda_{i1}\lambda_{i2}\cdots\lambda_{im},$$

$$b_\sigma=b_{i_1i_2}b_{i_2i_3}\cdots b_{i_mi_1}.$$

Received May. 5, 1994.

2 Main result

The following result is an improtant proposition of definite positive matrices

Lemma 1 *Let* $A \in S^{k\times k}$, *then* $A_r^{1/n} \geq A_r^{(1/n)} > 0$ *for* $r \in \underline{K}$ *and* $n \in N$.

Proof For any real function $f(\lambda)$, given a real symmetric matrix A, we define

$$f(A) = U^T f(D) U = U^T \operatorname{diag}(f(\lambda_1), \cdots, f(\lambda_k)) U,$$

where $A = U^T D U$, U is a real orthogonal matrix, $D = \operatorname{diag}(\lambda_1, \cdots, \lambda_k)$, $(\lambda_i \in R, \ i \in \underline{K})$. The case $k = 1$ or $n = 1$ is trivial. Now we consider $k > 1$ and $n > 1$. For any given $r \in \underline{K}$, if $r = k$, the result is immediate; If $1 \leq r < k$, we block A and $A^{1/n}$ as follows ($A^{1/n}$ is defined as above when we set $f(\lambda) = \lambda^{1/n}$):

$$A = \begin{bmatrix} A_r & A_{12}^T \\ A_{12} & A_r^{(c)} \end{bmatrix}, \ A^{1/n} = \begin{bmatrix} A^{(1/n)} & A_{12}^{(1/n)} \\ [A_{12}^{(r_n)}]^T & \bigstar \end{bmatrix},$$

$A_r, A_r^{(1/n)} \in R^{r\times r}$, others are corresponding block matrices. We know that $f(\lambda)$ is a monotone increasing function on S_k in terms of [2], that is,

$$(\text{for any } A, B \in S_k): \ A > B \rightarrow f(A) > f(B),$$

note that $A_r \in S_r$. By [2], we have:

$$f(A_r) \geq A_r^{(1/n)} > 0, \text{ says, } A^{1/n} \geq A_r^{(1/n)} > 0 \text{ for } r \in \underline{K}, n \in N.$$

Lemma 2 *Suppose* $A, B \in S^{k\times k}$ *and* $A \geq B$, *then*

$$\operatorname{tr}(A^n) \geq \operatorname{tr}(B^n) \quad (n = 1, 2, 3, \cdots).$$

Proof By the positivity of A and B, we can set

$$\sigma(A) = \{\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k\} \ (\lambda_k > 0),$$

$$\sigma(B) = \{\mu_1 \geq \mu_2 \geq \cdots \geq \mu_k\} \ (\mu_k > 0).$$

Corollary 7.7.4 in [3] tells us: $\lambda_i \geq \mu_i > 0$, $i \in \underline{K}$. Therefore

$$\lambda_i^n \geq \mu_i^n > 0, \ i \in \underline{K} \text{ for any } n \in N.$$

So, $\operatorname{tr}(A^n) = \sum_{i=1}^{K} \lambda_i \geq \sum_{i=1}^{K} \mu_i^n = \operatorname{tr}(B^n)$. □

Lemma 3 *Let* $A \in S_K$ $(K \in N)$, *then* $\operatorname{tr}(A_r^m) \leq \operatorname{tr}(A_r^{(m)})$ *for all* m *and all* $r \in \underline{K}$. *Here* A_r *and* $A_r^{(m)}$ *are the rth principal main submatrices of* A *and* A^m, *respectively.*

Proof Note that for the case $r=k$ or $m=1$, it is trivial. Now let $m>1$ and $1\le r<k$, and set $F=A^m$ corresponding to (1.1), we have

$$F=\begin{bmatrix} F_r & F_{12} \\ F_{21} & F_r^c \end{bmatrix}$$

and $F_r=A_r^{(m)}$, also $A=F^{1/m}$ (defined as in Lemma 1), then we may write $A_r=F_r^{(1/m)}$, $F_r^{(1/m)}$ is the rth principal main submatrix of $F^{1/m}$. If the result is not true, that is, $\mathrm{tr}(A_r^{(m)})<\mathrm{tr}(A_r^m)$, then

$$\mathrm{tr}(F_r)=\mathrm{tr}(A_r^{(m)})<\mathrm{tr}(A_r^m)=\mathrm{tr}[(F_r^{(1/m)})^m]. \tag{2.2}$$

By Lemma 1, $0<F_r^{(1/m)}\le F_r^{1/m}$, while we have

$$0<\mathrm{tr}\{[F_r^{(1/m)}]^m\}\le\mathrm{tr}\{(F_r^{1/m})^m\}=\mathrm{tr}(F_r). \tag{2.3}$$

From Lemma 2, a contridiction.

Now we complete this part by theorem 1:

Theorem 1 *Let* $A\in S_k$, $B\in I_k$, *then* (1.1) *holds for all* $m=1,2,3,\cdots$.

Theorem 1 is immediate from Lemma 3.

3. About N_k

Now we deal with the inequality (1.1) for nonnegative matrices. We have

Lemma 4 *Given* $A,B\in R^{k\times k}$, $A\gg B$ *is defined as* $A-B\in N_k$ (N_k *stands for the set of all non-negative matrices of order* K). *Then* $A,B,A-B\in N_k$ *for any* $n\in N$ *implies* $A^n-B^n\in N_k$.

Proof For $n=1$, it is trivial. Now suppose it is valid for $n-1$, then from the formula

$$A^n-B^n=A(A^{n-1}-B^{n-1})+(A-B)B^{n-1}$$

and the hypothesis we know that $A^n-B^n\in N_k$.

Next we use Lemma 4 to get Theorem 2.

Theorem 2 *Let* $A,B,AB-BA\in N_k$, *then*

$$(\text{for any } m\in N):\ \mathrm{tr}(AB)^m\le\mathrm{tr}(A^mB^m). \tag{3.1}$$

Proof We use induction on m to verify the inequality $(AB)^m\le A^mB^m$. It is trivial for $m=1$. Now suppose it holds for $m-1$, then for m, we have:

$$A^mB^m=A(A^{m-1}B^{m-1})B\gg A(AB)^{m-1}B. \tag{3.2}$$

From the hypothesis, $AB\gg BA\gg 0$, and by Lemma 4, we get $(AB)^{m-1}\gg(BA)^{m-1}$, so (3.2) is followed by

$$A^mB^m\gg A(AB)^{m-1}B\gg A(BA)^{m-1}B=(AB)^m,$$

that is,

$$A^m B^m \gg (AB)^m.$$

By the definition "$\gg$", (3.1) is immediate. □

Next we present some results related to [4].

Theorem 3 *Let* $A = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_k)$, $\overline{B} = (b_{ij}) \in N_k$, *then* (1.1) *holds. Moreover, the equality in* (1.1) *holds iff*

$$(\text{for any pair}(i, j):\ 1 \le i, j \le k):\ \overline{b_{ij}} \neq 0 \rightarrow \lambda_i = \lambda_j. \tag{3.3}$$

Proof Since $\overline{B}$ is nonnegative, we have $\overline{b_\sigma} \geq 0$ for any σ (see [4]). By arithematic-geometric means we get: (for any σ): $\Phi_{\sigma,m}(\lambda) \geq 0$. So,

$$\sum_\sigma \Phi_{\sigma,m}(\lambda)\overline{b_\sigma} \geq 0.$$

(2.1) is immediate from Theorem 1 in [4].

Now we consider the case $\sum_\sigma \Phi_{\sigma,m}(\lambda)\overline{b_\sigma} = 0$. In this case $\Phi_{\sigma,m}(\lambda)\overline{b_\sigma} = 0$ for all $\sigma = (i_1, i_2, \cdots, i_m)$. Suppose there exists a pair (i, j) in $\underline{K} \times \underline{K}$ such that $\overline{b_{ij}}\,0$, then we set $\sigma_1 = (i, j, j, \cdots, j)$ whose first coordinate is i and the others are all j. We get $\Phi_{\sigma,m}(\lambda)\overline{b_\sigma} = 0$. But $\overline{b_\sigma} = \overline{b_{ij}}\,\overline{b_{jj}}^{m-1} \neq 0$. So, $\Phi_{\sigma,m}(\lambda) = 0$, whereas

$$\Phi_{\sigma,m}(\lambda) = 1/m\,[\lambda_i^m + (m-1)\lambda_j^m] - \lambda_i\lambda_j^{m-1},\ \lambda_i, \lambda_j > 0,$$

from arithematic-geometric mean inequality we get $\lambda_i = \lambda_j$.

Conversely, if we have:

$$(\text{for any pair}(i, j): 1 \le i, j \le k):\ \overline{b_{ij}} \neq 0 \rightarrow \lambda_i = \lambda_j.$$

Then $\overline{b_\sigma} \neq 0 \rightarrow \Phi_{\sigma,m}(\lambda) = 0$ for any possible σ. So, we have $\Phi_{\sigma,m}(\lambda)\overline{b_\sigma} = 0$ for any σ. That is $\sum_\sigma \Phi_{\sigma,m}(\lambda)\overline{b_\sigma} = 0$. From the proof of the above theorem, we get the equality of (1.1).

It is easy to see that (2.1) holds when $A, B \in S_k$ and $\overline{B} \in N_k$.

Theorem 4 *Suppose that* $A, B \in S_k$, $\overline{B}$ *described as the above is nonnegative. Then the equality of* (2.1) *holds iff* $AB = BA$.

Proof The sufficiency is obvious. Now we come to prove its necessity: As $\overline{B}$ is non-negative, we know the equality of (1.1) holds iff: (for any pair (i, j): $1 \le i, j \le k$): $\overline{b_{ij}} \neq 0 \rightarrow \lambda_i = \lambda_j$. (by Theorem 3)

Next we prove $AB = BA$:

For $k = 1$, it is trivial. Suppose it is true for all the matrices of order less than k. Then for k, if for any pair (i, j): $i \neq j$, $1 \le i, j \le k$, $\overline{b_{ij}} = 0$, that is, B is diagnal, then $AB = BA$ is immediate. If there exists a pair (i, j): $i \neq j$, $1 \le i, j \le k$ such that $\overline{b_{ij}} \neq 0$, we have $\lambda_i = \lambda_j$. Without the loss of generality, we set

$$\lambda_1 = \lambda_2 = \cdots = \lambda_r,\ 1 < r \le k.$$

If $r = k$, then $AB = BA$. Now for $r < k$ and $\lambda_1 = \lambda_2 = \cdots = \lambda_r \neq \lambda_{r+1}$ $(i = 1, 2, \cdots, k - r)$, we get $\overline{B} = \overline{B}_1$