Journal of Mathematics Research
www.ccsenet.org/journal.html
Vol. 1, No. 1
March 2009

On the Proof of Bushell-Trustrum Inequality

Shiqing Wang & Yan Lou
College of Mathematics and Information Science
North China University of Water Conservancy and Electric power
Zhengzhou 450011, China
Tel: 86-371-6579-0758 E-mail: wangshiqing@ncwu.edu.cn

Foundation item: Support by the Natural Science Foundation of Henan (0611052600), the Basic and the Forward Position Technology Research Project of Henan.

Abstract

Bushell and Trustrum (Bushell, 1990, p. 173-178) give the famous Bushell-Trustrum inequality, but their proof exists two main mistakes which make their proof process can not establish. This paper corrects these mistakes and gives the correct proof.

Keywords: Bushell-Trustrum inequality, Positive semi-definite Hermite matrix, Unitary matrix

1. Introduction

Let A and B be two positive semi-definite Hermite matrix with rank n, there eigenvalues are $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$ and $\mu_1 \geq \cdots \geq \mu_n \geq 0$, respectively. Then for any positive integer k, (Marcus, 1956, p. 173-174. Marshall, 1979).

$$\sum_{i=1}^{n} \lambda_i^k \mu_{n-i+1}^k \leq tr(A^k B^k) \leq \sum_{i=1}^{n} \lambda_i^k \mu_i^k$$

And (Lieb and Thirring, 1976, see the third reference of (Bushell P J, 1990)).

$$tr(AB)^k \leq tr(A^k B^k)$$

In 1990, Bushell and Trustrum proved

$$\sum_{i=1}^{n} \lambda_i^k \mu_{n-i+1}^k \leq tr(AB)^k \leq tr(A^k B^k) \leq \sum_{i=1}^{n} \lambda_i^k \mu_i^k$$

Whereas the result proved by Lieb and Thirring, Bushell and Trustum only need to prove

$$\sum_{i=1} \lambda_i^k \mu_{n-i+1}^k \leq tr(AB)^k \leq \sum_{i=1} \lambda_i^k \mu_i^k$$

They construct $B_i = U_i B U_i^* (i = 1, 2)$ in their proof firstly, then $tr(AB_1)^k$, $tr(AB_2)^k$ are the smallest and largest values of $tr(AB)^k$, here U_i is unitary matrix. The mistakes in their proof are mainly in the following two points: (1) Exist unitary matrix X with rank n, such that X^*AX, X^*B_1X, X^*B_2X become diagonal at the same time; (2)

$$tr(AB)^k = \sum_{i=1}^{n} \lambda_{\pi(i)}^k \mu_i^k \tag{1}$$

We will point out that unitary matrix X with rank n, which makes X^*AX, X^*B_1X, X^*B_2X become diagonal at the same time does not definitely exist, and for general positive semi-definite Hermite matrix, (1) does also not definitely exist.

We give the following conclusions:

Exist unitary matrixX with rank n, such that X^*AX and X^*B_1X become diagonal at the same time; Exist unitary matrixY with rank n, such that Y^*AY and Y^*B_2Y become diagonal at the same time. And for B_1, B_2,

$$tr(AB_1)^k = \sum_{i=1}^{n} \lambda_{\pi(i)}^k \mu_i^k \tag{2}$$

$$tr(AB_2)^k = \sum_{i=1}^{n} \lambda_{\pi(i)}^k \mu_i^k \tag{3}$$

Thus complete the certification of Bushell-Trustrum inequality. We need to use the following Lemma:

Lemma (Wang Song-Gui, 2006, p. 143) Assuming that $\alpha_1 \geq \cdots \geq \alpha_n$, $\mu_1 \geq \cdots \geq \mu_n$. If $\pi(1), \cdots, \pi(n)$ is any permutation of $1, \cdots, n$, then

$$\sum_{i=1}^{n} \alpha_i \mu_{n-i+1} \leq \sum_{i=1}^{n} \alpha_{\pi(i)} \mu_i \leq \sum_{i=1}^{n} \alpha_i \mu_i$$

2. Our proof

Suppose $A > 0$, $B > 0$, otherwise for any $c > 0$, There must be $A + cI > 0$, $B + cI > 0$, finally we take limit to the result obtained when $c \to 0$, then we conclude the proof.

Since entire unitary matrix with rank n constitutes a closed set and mapping $U \to tr(AUBU^*)^k$ is a continuous function defined on this closed set, so there must be the smallest and largest values in U_1 and U_2, Then

$$tr(AU_2BU_2^*)^k \leq tr(AUBU^*)^k \leq tr(AU_1BU_1^*)^k \tag{4}$$

Especially, in (4), take $U = I$, then we have

$$tr(AU_2BU_2^*)^k \leq tr(AB)^k \leq tr(AU_1BU_1^*)^k \tag{5}$$

If let $B_i = U_iBU_i^*(i = 1, 2)$, we will prove first: Exist unitary matrix X with rank n, such that X^*AX and X^*B_1X are diagonal.

Let

$$R = \begin{pmatrix} R_{12} & 0 \\ 0 & I \end{pmatrix} \tag{6}$$

$$R = \begin{pmatrix} F_{12} & 0 \\ 0 & 0 \end{pmatrix} \tag{7}$$

where

$$R_{12} = (1 + |\varepsilon|^2)^{-\frac{1}{2}} \begin{bmatrix} 1 & -\varepsilon \\ \overline{\varepsilon} & 1 \end{bmatrix} \tag{8}$$

$$F_{12} = \frac{1}{|\varepsilon|} \begin{bmatrix} 0 & -\varepsilon \\ \overline{\varepsilon} & 0 \end{bmatrix} \tag{9}$$

R, F are $n \times n$ rectangular matrix, 0, I are zero matrix and unit matrix on some degree.

Obviously, R is an unitary matrix, and to infinitely small $\varepsilon \neq 0$, R can denoted as

$$R = I + |\varepsilon|F + o(|\varepsilon|^2) \tag{10}$$

Here $o(|\varepsilon|^2)$ is $n \times n$ rectangular matrix, everyone of its element is infinitesimal of higher order of $|\varepsilon|$. For convenient we use $o(|\varepsilon|^2)$ to denote either matrix or number.

In fact,

$$R - I - |\varepsilon|F = \begin{bmatrix} R_{12} - I - |\varepsilon|F_{12} & 0 \\ 0 & 0 \end{bmatrix} \tag{11}$$

$$R_{12} - I - |\varepsilon|F_{12} = \begin{bmatrix} \frac{1}{\sqrt{1+|\varepsilon|^2}} - 1 & -\frac{\varepsilon}{\sqrt{1+|\varepsilon|^2}} - \varepsilon \\ \frac{\overline{\varepsilon}}{\sqrt{1+|\varepsilon|^2}} - \overline{\varepsilon} & \frac{1}{\sqrt{1+|\varepsilon|^2}} - 1 \end{bmatrix} \tag{12}$$

From mathematical analysis, when $x \to 0$,

$$1 - \frac{1}{\sqrt{1+x^2}} = \frac{\sqrt{1+x^2} - 1}{\sqrt{1+x^2}} = \frac{x^2}{\sqrt{1+x^2}(\sqrt{1+x^2}+1)} \sim x^2$$

so elements in (11) and (12) are infinitesimal of higher order of $|\varepsilon|$, thus (10) holds. For any unitary matrix T, define

$$\tilde{B} = (TRT^*)B(TR^*T^*) \tag{13}$$

Since R is unitary matrix, TRT^* is unitary matrix. Because B is positive semi-definite Hermite matrix, $TR^*T^* = (TRT^*)^*$, $\tilde{B}$ is positive semi-definite Hermite matrix too. From (10), we get

$$TRT^* = T(I + |\varepsilon|F + o(|\varepsilon|^2))T^* = I + |\varepsilon|TFT^* + o(|\varepsilon|^2) \tag{14}$$

Notice that $F^* = -F$,

$$TR^*T^* = T(I + |\varepsilon|F^* + o(|\varepsilon|^2))T^* = I - |\varepsilon|TFT^* + o(|\varepsilon|^2) \tag{15}$$

Then

$$\begin{aligned} \tilde{B} &= B + |\varepsilon|(TFT^*B - BTFT^*) + o(|\varepsilon|^2) \\ &= B + |\varepsilon|T(FT^*BT - T^*BTF)T^* + o(|\varepsilon|^2) \\ &= B + |\varepsilon|T(FC - CF)T^* + o(|\varepsilon|^2) \end{aligned} \tag{16}$$

Here

$$C = T^*BT \tag{17}$$

It is easy to prove that for any two unitary matrix with rank n P and Q, have

$$tr(P + |\varepsilon|Q)^k = trP^k + k|\varepsilon|trP^{k-1}Q + o(|\varepsilon|^2) \tag{18}$$

Then from (16), (18)

$$\begin{aligned} tr(A\tilde{B})^k &= tr(AB + |\varepsilon|AT(FC - CF)T^* + o(|\varepsilon|^2))^k \\ &= tr(AB)^k + k|\varepsilon|tr(AB)^{k-1}AT(FC - CF)T^* + o(|\varepsilon|^2) \\ &= tr(AB)^k + k|\varepsilon|tr[D(FC - CF)] + o(|\varepsilon|^2) \end{aligned} \tag{19}$$

Here

$$D = T^*(AB)^{k-1}AT \tag{20}$$

We can prove that $(AB)^{k-1}A \geq 0$

In fact, notice that A and B are both positive semi-definite Hermite matrixes.

When $k = 2$, $ABA = AB^{\frac{1}{2}}B^{\frac{1}{2}}A = (B^{\frac{1}{2}}A)^*B^{\frac{1}{2}}A \geq 0$.

When $k = 3$, $ABABA = ABA^{\frac{1}{2}}A^{\frac{1}{2}}BA = (A^{\frac{1}{2}}BA)^*A^{\frac{1}{2}}BA \geq 0$. It can be proved by induction.

In(20), because $(AB)^{k-1}A$ is positive semi-definite, T is any unitary matrix, so we can choose unitary matrix T, such that D becomes diagonal,

$$D = diag(d_1, \cdots, d_n), \quad d_1 \geq \cdots \geq d_n \geq 0 \tag{21}$$

Let $C = \begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \end{pmatrix}$, $D = \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix}$, where $C_1 = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$, $D_1 = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}$; Notice that C_1 is Hermite matrix, $F_{12}^* = -F_{12}$, then

$$\begin{aligned}
|\varepsilon| trD(FC - CF) &= |\varepsilon| tr \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix} \left[\begin{pmatrix} F_{12} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \end{pmatrix} - \begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \end{pmatrix} \begin{pmatrix} F_{12} & 0 \\ 0 & 0 \end{pmatrix} \right] \\
&= |\varepsilon| tr \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix} \begin{bmatrix} F_{12}C_1 - C_1F_{12} & F_{12}C_2 \\ -C_2F_{12} & 0 \end{bmatrix} \\
&= |\varepsilon| tr \begin{bmatrix} D_1(F_{12}C_1 - C_1F_{12}) & D_1F_{12}C_2 \\ -D_2C_2F_{12} & 0 \end{bmatrix} \\
&= |\varepsilon| trD_1(F_{12}C_1 - C_1F_{12}) \\
&= |\varepsilon| trD_1(F_{12}C_1 + (F_{12}C_1)^*) \\
&= (d_2 - d_1)(\overline{\varepsilon} c_{12} + \varepsilon \overline{c}_{12})
\end{aligned} \tag{22}$$

The last equation is right because C_1 is Hermite matrix, and

$$|\varepsilon| F_{12}C_1 = \begin{pmatrix} 0 & -\varepsilon \\ \overline{\varepsilon} & 0 \end{pmatrix} \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} = \begin{pmatrix} -\varepsilon c_{21} & -\varepsilon c_{22} \\ \overline{\varepsilon} c_{11} & \overline{\varepsilon} c_{12} \end{pmatrix}$$

By (22) and (19), then

$$tr(A\tilde{B})^k - tr(AB)^k = k(d_2 - d_1)(\overline{\varepsilon} c_{12} + \varepsilon \overline{c}_{12}) + o(|\varepsilon|^2) \tag{23}$$

This formula is correct on arbitrary semi-positive Hermite matrix B and infinitely small $\varepsilon \neq 0$.

Especially, set $B = B_1$, $\varepsilon = \eta c_{12}$, $\eta > 0$, If $d_2 \neq d_1$, then by definition of B_1 and (2), we obtain $\overline{\varepsilon} c_{12} + \varepsilon \overline{c}_{12} = \eta |c_{12}|^2 = 0$, then $c_{12} = \overline{c}_{21} = 0$.

Similarly, we take R, F such that their $i,\ j(i < j)$ row and column have form of (8), (9), and similar to the proof above, then it can be obtained.

$$tr(A\tilde{B})^k - tr(AB)^k = k(d_j - d_i)(\overline{\varepsilon} c_{ij} + \varepsilon \overline{c}_{ij}) + o(|\varepsilon|^2) \tag{24}$$

Set $\varepsilon = \eta c_{ij}$, $\eta > 0$, Use the same method $c_{ij} = \overline{c}_{ji} = 0$ can be obtained.

Suppose $c_1 > c_2 > \cdots > c_l$ are l different value of $d_1, \cdots, d_n$, here $D = diag(c_1 I_{n1}, \cdots c_l I_{nl})$. make $C = T^* B_1 T$ become block matrix

$$C = diag(C_1, \cdots, C_l) \tag{25}$$

Here C_i is positive semi-definite Hermite matrix with rank n_i. Let $V_i (i = 1, \cdots, l)$ is unitary matrix, such that $E_i = V_i^* C_i V_i$, $i = 1, \cdots, l$, becomes diagonal matrix.

Let

$$V = diag(V_1, \cdots, V_l) \tag{26}$$

$$E = diag(E_1, \cdots, E_l) \tag{27}$$

Set $X = TV$, then X is an unitary matrix, and

$$X^* B_1 X = V^* T^* B_1 T V = V^* C V = E \tag{28}$$

This is a diagonal matrix, its diagonal elements are eigenvalues of B_1, furthermore

$$X^* (AB_1)^{k-1} AX = V^* T^* (AB_1)^{k-1} ATV = V^* DV = D \tag{29}$$

The last equation is correct because V and D are block matrix with same degree. By (28) and (29) we know

$$(E^{\frac{1}{2}} (X^* AX) E^{\frac{1}{2}})^k = E^{\frac{1}{2}} X^* (AB_1)^{k-1} AXE^{\frac{1}{2}} = E^{\frac{1}{2}} DE^{\frac{1}{2}} \tag{30}$$

then

$$X^*AX = E^{-\frac{1}{2}}(E^{\frac{1}{2}}DE^{\frac{1}{2}})^{\frac{1}{k}}E^{-\frac{1}{2}} \tag{31}$$

It is a diagonal matrix. It be proved that exist $n \times n$ unitary matrix such that X^*AX, X^*B_1X are all diagonal matrix.

Similarly, in(24), let $B = B_2$, $\varepsilon = \eta c_{ij}$, $\eta < 0$, then $c_{ij} = \overline{c}_{ji} = 0$. Notice that because $C = T^*BT$ and B_i are different, so we write as $G = T^*B_2T$.

Suppose that $g_1 > g_2 > \cdots > g_m$ are m different values of $d_1, \cdots, d_n$, here $D = diag(g_1I_{n1}, \cdots g_mI_{n_m})$, make $G = T^*B_2T$ become block matrix

$$G = diag(G_1, G_2, \cdots G_m) \tag{32}$$

Let $W_i(i = 1, 2, \cdots m)$ be an unitary matrix, such that $W_i^*G_iW_i(i = 1, 2, \cdots m)$ is diagonal matrix. Write

$$W = diag(W_1, W_2, \cdots W_m) \tag{33}$$

Set $Y = TW$, then Y is an unitary matrix, similar to the proof on (28)-(31), it be obtained that exist unitary matrixY such that Y^*AY, Y^*B_2Y are all diagonal matrix.

According to (28) and (31), X^*AX, X^*B_1X, Y^*AY, Y^*B_2Y are all diagonal matrixes. Notice that $X^*B_1X = X^*U_1B_1U_1{}^*X$, $Y^*B_2Y = Y^*U_2B_2U_2{}^*Y$, U_1, U_2, X, Y are all unitary matrix, so diagonal elements of X^*B_1X, Y^*B_2Y are eigenvalues of B. Thus

$$tr(AB_1)^k = tr(X^*(AB_1)^kX) = tr(X^*AXX^*B_1X)^k = \sum_{i=1}^{n} \lambda_{\pi(i)}^k \mu_i^k \tag{34}$$

$$tr(AB_2)^k = tr(Y^*(AB_2)^kY) = tr(Y^*AYY^*B_2Y)^k = \sum_{i=1}^{n} \lambda_{\pi'(i)}^k \mu_i^k \tag{35}$$

Here $\pi(i)$, $\pi'(i)$ is any permutation of $1, 2, \cdots, n$, respectively. From Lemma,

$$tr(AB_1)^k \le \sum_{i=1}^{n} \lambda_i^k \mu_i^k \tag{36}$$

$$tr(AB_2)^k \ge \sum_{i=1}^{n} \lambda_i^k \mu_{n-i+1}^k \tag{37}$$

And using(5), then the proof is completed.

References

Bushell P J, Trustrum G B. (1990). Trace inequalities for positive definite matrix power products. *Linear Algebra and its Applications*, 132, 173-178.

Marcus M, An eigenvalue inequality for the product of normal matrices. *Amer Math Monthly*, (1956), 63:173-174.

Marshall A W and Olkin I. *Inequalities: Majorization and its applications*. Academic, New York, (1979).

Wang Song-Gui, Wu Mi-Xia and Jia Zhong-Zhen. (2006). *Matrix Inequalities*, Beijing: Science Press.