Operators
and
Matrices
Volume 2, Number 3 (2008), 435–442

PROOF OF BÖTTCHER AND WENZEL'S CONJECTURE

SEAK-WENG VONG AND XIAO-QING JIN

(*communicated by L. Rodman*)

Abstract. In 2005, Böttcher and Wenzel raised a conjecture that if X and Y are any two real n-by-n matrices, then $\|XY - YX\|_F^2 \leqslant 2\|X\|_F^2\|Y\|_F^2$ where $\|\cdot\|_F$ denotes the Frobenius norm. They proved this for the case of 2-by-2 matrices. Later, László proved the conjecture for the case of 3-by-3 matrices. In this paper, we prove the conjecture for general n-by-n matrices.

1. Introduction

Böttcher and Wenzel proposed the following conjecture in [1]: the upper bound of the Frobenius norm of the commutator of any two n-by-n real matrices X and Y is given by

$$\|XY - YX\|_F^2 \leqslant 2\|X\|_F^2\|Y\|_F^2.$$

They also proved that the inequality is true for the case of $n = 2$. Very recently, László [3] proved the conjecture for the case of $n = 3$. In this paper, we prove the conjecture for general n. In next section, for simplicity, we will give a detailed proof of the conjecture for the case of $n = 3$. An outline of the proof for general n can be found in Section 3. We emphasize that our proof for the case of $n = 3$ is completely different from that provided in [3]. As to our knowledge, the technique of [3] cannot be applied to prove the conjecture in the general case.

2. The proof

Our analysis will show that the conjecture has close connection with the Cauchy-Schwarz inequality [2]. To get a sharper bound, the difference between the two sides of the Cauchy-Schwarz inequality has to be taken into account. The following Lagrange

---

Mathematics subject classification (2000): 15A52, 15A45, 60H25, 65F35.

Key words and phrases: Commutator of two matrices, the Cauchy-Schwarz inequality, the Lagrange identity.

The research is supported by the research grant 050/2005/A from FDCT.

© ELEMENT, Zagreb
Paper OAM-02-26

identity [3] serves this purpose:

$$\begin{aligned}\left(\sum_{i=1}^{n}\alpha_i\beta_i\right)^2 &= \sum_{i=1}^{n}\alpha_i^2\sum_{i=1}^{n}\beta_i^2-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\alpha_i\beta_j-\alpha_j\beta_i)^2\\ &= \sum_{i=1}^{n}\alpha_i^2\sum_{i=1}^{n}\beta_i^2-\sum_{i=1}^{n}\sum_{j>i}\left|\begin{array}{cc}\alpha_i & \alpha_j\\ \beta_i & \beta_j\end{array}\right|^2.\end{aligned}$$

It will be used repeatedly throughout our analysis.

For $n=3$, let $X=USV$ be the singular value decomposition of X where $S\equiv\mathrm{diag}(s_1,s_2,s_3)$ and U, V are (real) orthogonal. Denote

$$B\equiv VYV^T\equiv[b_{ij}]_{i,j=1}^3,\qquad Q\equiv VU,\qquad C\equiv Q^TBQ\equiv[c_{ij}]_{i,j=1}^3.$$

Note that Q is also orthogonal. By the Cauchy-Schwarz inequality,

$$\begin{aligned}\|XY-YX\|_F^2 &= \|USVY-YUSV\|_F^2=\|SB-CS\|_F^2\\ &= (s_1b_{11}-s_1c_{11})^2+(s_1b_{12}-s_2c_{12})^2+(s_1b_{13}-s_3c_{13})^2\\ &\quad+(s_2b_{21}-s_1c_{21})^2+(s_2b_{22}-s_2c_{22})^2+(s_2b_{23}-s_3c_{23})^2\\ &\quad+(s_3b_{31}-s_1c_{31})^2+(s_3b_{32}-s_2c_{32})^2+(s_3b_{33}-s_3c_{33})^2\\ &\leqslant (s_1^2b_{11}^2+s_1^2c_{11}^2-2s_1^2b_{11}c_{11})+(s_1^2+s_2^2)(b_{12}^2+c_{12}^2)\\ &\quad+(s_1^2+s_3^2)(b_{13}^2+c_{13}^2)+(s_2^2+s_1^2)(b_{21}^2+c_{21}^2)\\ &\quad+(s_2^2b_{22}^2+s_2^2c_{22}^2-2s_2^2b_{22}c_{22})+(s_2^2+s_3^2)(b_{23}^2+c_{23}^2)\\ &\quad+(s_3^2+s_1^2)(b_{31}^2+c_{31}^2)+(s_3^2+s_2^2)(b_{32}^2+c_{32}^2)\\ &\quad+(s_3^2b_{33}^2+s_3^2c_{33}^2-2s_3^2b_{33}c_{33})\\ &= \sum_{j=1}^{3}s_j^2\Delta_j,\end{aligned}$$

where

$$\Delta_j\equiv b_{jj}^2+c_{jj}^2-2b_{jj}c_{jj}+\sum_{k\neq j}b_{jk}^2+\sum_{k\neq j}c_{kj}^2+\sum_{k\neq j}c_{jk}^2+\sum_{k\neq j}b_{kj}^2,\tag{1}$$

for $j=1,2,3$. The conjecture follows if we can prove that

$$\Delta_j\leqslant 2\|B\|_F^2,\qquad j=1,2,3.$$

The rest of the proof will concentrate on this claim. Without loss of generality, we only carry out our analysis on Δ_1.

Let $\mathbf{q}\equiv(q_1,q_2,q_3)$ be the first row of Q^T and $\mathbf{r}_i$ be the i-th row of B, i.e.,

$$B\equiv[b_{ij}]_{i,j=1}^3=\begin{bmatrix}\mathbf{r}_1\\ \mathbf{r}_2\\ \mathbf{r}_3\end{bmatrix}.$$

Then we have by the fact that Q is (real) orthogonal,

$$\begin{aligned}\sum_{k=1}^{3} c_{1k}^2 &= \|\mathbf{q}BQ\|_2^2 = \|\mathbf{q}B\|_2^2 \\ &= q_1^2\|\mathbf{r}_1\|_2^2 + q_2^2\|\mathbf{r}_2\|_2^2 + q_3^2\|\mathbf{r}_3\|_2^2 + 2q_1q_2\mathbf{r}_1\cdot\mathbf{r}_2^T + 2q_1q_3\mathbf{r}_1\cdot\mathbf{r}_3^T + 2q_2q_3\mathbf{r}_2\cdot\mathbf{r}_3^T,\end{aligned}$$

and

$$\sum_{k=1}^{3} c_{k1}^2 = \|(Q^TB\mathbf{q}^T)^T\|_2^2 = \|\mathbf{q}B^T\|_2^2 = (\mathbf{q}\cdot\mathbf{r}_1^T)^2 + (\mathbf{q}\cdot\mathbf{r}_2^T)^2 + (\mathbf{q}\cdot\mathbf{r}_3^T)^2. \tag{2}$$

On the other hand, since

$$c_{11} = \mathbf{q}\cdot(\mathbf{r}_1\cdot\mathbf{q}^T,\, \mathbf{r}_2\cdot\mathbf{q}^T,\, \mathbf{r}_3\cdot\mathbf{q}^T)^T,$$

we have by using the Lagrange identity and $\|\mathbf{q}\|_2 = 1$,

$$\begin{aligned}c_{11}^2 &= \|\mathbf{q}\|_2^2\Big[(\mathbf{r}_1\cdot\mathbf{q}^T)^2 + (\mathbf{r}_2\cdot\mathbf{q}^T)^2 + (\mathbf{r}_3\cdot\mathbf{q}^T)^2\Big] \\ &\quad -\begin{vmatrix} q_1 & q_2 \\ \mathbf{r}_1\cdot\mathbf{q}^T & \mathbf{r}_2\cdot\mathbf{q}^T\end{vmatrix}^2 - \begin{vmatrix} q_1 & q_3 \\ \mathbf{r}_1\cdot\mathbf{q}^T & \mathbf{r}_3\cdot\mathbf{q}^T\end{vmatrix}^2 - \begin{vmatrix} q_2 & q_3 \\ \mathbf{r}_2\cdot\mathbf{q}^T & \mathbf{r}_3\cdot\mathbf{q}^T\end{vmatrix}^2 \\ &= (\mathbf{q}\cdot\mathbf{r}_1^T)^2 + (\mathbf{q}\cdot\mathbf{r}_2^T)^2 + (\mathbf{q}\cdot\mathbf{r}_3^T)^2 - I_{12} - I_{13} - I_{23},\end{aligned} \tag{3}$$

where

$$I_{12} \equiv \begin{vmatrix} q_1 & q_2 \\ \mathbf{r}_1\cdot\mathbf{q}^T & \mathbf{r}_2\cdot\mathbf{q}^T\end{vmatrix}^2, \quad I_{13} \equiv \begin{vmatrix} q_1 & q_3 \\ \mathbf{r}_1\cdot\mathbf{q}^T & \mathbf{r}_3\cdot\mathbf{q}^T\end{vmatrix}^2, \quad I_{23} \equiv \begin{vmatrix} q_2 & q_3 \\ \mathbf{r}_2\cdot\mathbf{q}^T & \mathbf{r}_3\cdot\mathbf{q}^T\end{vmatrix}^2.$$

Substituting (3) into (2), we obtain

$$c_{21}^2 + c_{31}^2 = I_{12} + I_{13} + I_{23}.$$

Thus for $j = 1$, (1) can be rewritten as

$$\begin{aligned}\Delta_1 &= \|B\|_F^2 - b_{22}^2 - b_{33}^2 - b_{32}^2 - b_{23}^2 - 2\Big[\mathbf{q}\cdot(\mathbf{r}_1\cdot\mathbf{q}^T,\, \mathbf{r}_2\cdot\mathbf{q}^T,\, \mathbf{r}_3\cdot\mathbf{q}^T)^T\Big]b_{11} \\ &\quad + q_1^2\|\mathbf{r}_1\|_2^2 + q_2^2\|\mathbf{r}_2\|_2^2 + q_3^2\|\mathbf{r}_3\|_2^2 + 2q_1q_2\mathbf{r}_1\cdot\mathbf{r}_2^T + 2q_1q_3\mathbf{r}_1\cdot\mathbf{r}_3^T + 2q_2q_3\mathbf{r}_2\cdot\mathbf{r}_3^T \\ &\quad + I_{12} + I_{13} + I_{23}.\end{aligned} \tag{4}$$

Now notice that the sum of the terms coming from c_{ij}^2 must be less than $\|B\|_F^2$. To find the difference between them, the Lagrange identity comes into play again. More precisely, we have

$$\begin{aligned}
&2q_1q_2\mathbf{r}_1\cdot\mathbf{r}_2^T+I_{12}\\
&\quad=2q_1q_2\mathbf{r}_1\cdot\mathbf{r}_2^T+[q_1(q_1b_{21}-q_2b_{11})+q_2(q_1b_{22}-q_2b_{12})+q_3(q_1b_{23}-q_2b_{13})]^2\\
&\quad\leqslant 2q_1q_2\mathbf{r}_1\cdot\mathbf{r}_2^T+\|\mathbf{q}\|_2^2\Big[(q_1b_{21}-q_2b_{11})^2+(q_1b_{22}-q_2b_{12})^2+(q_1b_{23}-q_2b_{13})^2\Big]\\
&\qquad-\left|\begin{matrix} q_1 & q_2\\ (q_1b_{21}-q_2b_{11}) & (q_1b_{22}-q_2b_{12})\end{matrix}\right|^2-\left|\begin{matrix} q_1 & q_3\\ (q_1b_{21}-q_2b_{11}) & (q_1b_{23}-q_2b_{13})\end{matrix}\right|^2\\
&\quad=q_1^2\|\mathbf{r}_2\|_2^2+q_2^2\|\mathbf{r}_1\|_2^2-J_{12,2}-J_{12,3}
\end{aligned}$$

with

$$\begin{aligned}
J_{12,2}&\equiv\left|\begin{matrix} q_1 & q_2\\ (q_1b_{21}-q_2b_{11}) & (q_1b_{22}-q_2b_{12})\end{matrix}\right|^2,\\
J_{12,3}&\equiv\left|\begin{matrix} q_1 & q_3\\ (q_1b_{21}-q_2b_{11}) & (q_1b_{23}-q_2b_{13})\end{matrix}\right|^2;
\end{aligned}$$

$$\begin{aligned}
&2q_1q_3\mathbf{r}_1\cdot\mathbf{r}_3^T+I_{13}\\
&\quad=2q_1q_3\mathbf{r}_1\cdot\mathbf{r}_3^T+[q_1(q_1b_{31}-q_3b_{11})+q_2(q_1b_{32}-q_3b_{12})+q_3(q_1b_{33}-q_3b_{13})]^2\\
&\quad\leqslant 2q_1q_3\mathbf{r}_1\cdot\mathbf{r}_3^T+\|\mathbf{q}\|_2^2\Big[(q_1b_{31}-q_3b_{11})^2+(q_1b_{32}-q_3b_{12})^2+(q_1b_{33}-q_3b_{13})^2\Big]\\
&\qquad-\left|\begin{matrix} q_1 & q_2\\ (q_1b_{31}-q_3b_{11}) & (q_1b_{32}-q_3b_{12})\end{matrix}\right|^2-\left|\begin{matrix} q_1 & q_3\\ (q_1b_{31}-q_3b_{11}) & (q_1b_{33}-q_3b_{13})\end{matrix}\right|^2\\
&\quad=q_1^2\|\mathbf{r}_3\|_2^2+q_3^2\|\mathbf{r}_1\|_2^2-J_{13,2}-J_{13,3}
\end{aligned}$$

with

$$\begin{aligned}
J_{13,2}&\equiv\left|\begin{matrix} q_1 & q_2\\ (q_1b_{31}-q_3b_{11}) & (q_1b_{32}-q_3b_{12})\end{matrix}\right|^2,\\
J_{13,3}&\equiv\left|\begin{matrix} q_1 & q_3\\ (q_1b_{31}-q_3b_{11}) & (q_1b_{33}-q_3b_{13})\end{matrix}\right|^2;
\end{aligned}$$

and similarly

$$2q_2q_3\mathbf{r}_2\cdot\mathbf{r}_3^T+I_{23}\leqslant q_3^2\|\mathbf{r}_2\|_2^2+q_2^2\|\mathbf{r}_3\|_2^2.$$

Putting all these inequalities into (4), we get

$$\begin{aligned}
\Delta_1\ \leqslant\ & 2\|B\|_F^2-b_{22}^2-b_{33}^2-b_{32}^2-b_{23}^2-2q_1^2b_{11}^2-2q_2^2b_{11}b_{22}-2q_3^2b_{11}b_{33}\\
&-2q_2q_3(b_{23}+b_{32})b_{11}-2q_1q_2(b_{12}+b_{21})b_{11}-2q_1q_3(b_{13}+b_{31})b_{11} \qquad (5)\\
&-J_{12,2}-J_{12,3}-J_{13,2}-J_{13,3}.
\end{aligned}$$

We must find negative terms to dominate $-2q_1q_2(b_{12}+b_{21})b_{11}-2q_1q_3(b_{13}+b_{31})b_{11}$ and those terms, in fact, can be absorbed in the "J" terms. However, it is not trivial to estimate the proportion to which these terms should be associated to the "J" terms although it turns out to be quite intuitive.

We first notice that $\Delta_1 \leqslant 2\|B\|_F^2$ obviously holds when $q_1^2 = 1$ (i.e., $q_2 = q_3 = 0$). Therefore, in the following, we assume that $q_1^2 \neq 1$ and denote $M^2 \equiv q_2^2 + q_3^2 > 0$. We have

$$\begin{aligned}
&-J_{12,2} - 2q_1q_2\tfrac{q_2^2}{M^2}(b_{12}+b_{21})b_{11}\\
&\quad = -q_2^4b_{11}^2 - (q_1q_2b_{21}+q_1q_2b_{12}-q_1^2b_{22})^2 + 2q_2^2b_{11}(q_1q_2b_{21}+q_1q_2b_{12}-q_1^2b_{22})\\
&\qquad -2q_1q_2\tfrac{q_2^2}{M^2}(b_{12}+b_{21})b_{11} + 2q_2^2\tfrac{q_1^2}{M^2}b_{11}b_{22} - 2q_2^2\tfrac{q_1^2}{M^2}b_{11}b_{22}\\
&\quad = -\Big[q_1q_2b_{21}+q_1q_2b_{12}-q_1^2b_{22} - \tfrac{(M^2-1)}{M^2}q_2^2b_{11}\Big]^2\\
&\qquad +\tfrac{q_1^4}{M^4}q_2^4b_{11}^2 - q_2^4b_{11}^2 - 2q_2^2\tfrac{q_1^2}{M^2}b_{11}b_{22};
\end{aligned}$$

$$\begin{aligned}
&-J_{13,3} - 2q_1q_3\tfrac{q_3^2}{M^2}(b_{13}+b_{31})b_{11}\\
&\quad = -q_3^4b_{11}^2 - (q_1q_3b_{31}+q_1q_3b_{13}-q_1^2b_{33})^2 + 2q_3^2b_{11}(q_1q_3b_{31}+q_1q_3b_{13}-q_1^2b_{33})\\
&\qquad -2q_1q_3\tfrac{q_3^2}{M^2}(b_{13}+b_{31})b_{11} + 2q_3^2\tfrac{q_1^2}{M^2}b_{11}b_{33} - 2q_3^2\tfrac{q_1^2}{M^2}b_{11}b_{33}\\
&\quad = -\Big[q_1q_3b_{31}+q_1q_3b_{13}-q_1^2b_{33} - \tfrac{(M^2-1)}{M^2}q_3^2b_{11}\Big]^2\\
&\qquad +\tfrac{q_1^4}{M^4}q_3^4b_{11}^2 - q_3^4b_{11}^2 - 2q_3^2\tfrac{q_1^2}{M^2}b_{11}b_{33};\\
&-J_{12,3} - 2q_2q_3\tfrac{q_1}{M^2}(q_3b_{21}+q_2b_{13})b_{11}\\
&\quad = -q_2^2q_3^2b_{11}^2 - (q_1q_3b_{21}+q_1q_2b_{13}-q_1^2b_{23})^2 + 2q_2q_3b_{11}(q_1q_3b_{21}+q_1q_2b_{13}-q_1^2b_{23})\\
&\qquad -2q_2q_3\tfrac{q_1}{M^2}(q_3b_{21}+q_2b_{13})b_{11} + 2q_2q_3\tfrac{q_1^2}{M^2}b_{11}b_{23} - 2q_2q_3\tfrac{q_1^2}{M^2}b_{11}b_{23}\\
&\quad = -\Big[q_1q_3b_{21}+q_1q_2b_{13}-q_1^2b_{23} - \tfrac{(M^2-1)}{M^2}q_2q_3b_{11}\Big]^2\\
&\qquad +\tfrac{q_1^4}{M^4}q_2^2q_3^2b_{11}^2 - q_2^2q_3^2b_{11}^2 - 2q_2q_3\tfrac{q_1^2}{M^2}b_{11}b_{23};
\end{aligned}$$

and

$$\begin{aligned}
&-J_{13,2} - 2q_2q_3\tfrac{q_1}{M^2}(q_2b_{31}+q_3b_{12})b_{11}\\
&\quad = -q_2^2q_3^2b_{11}^2 - (q_1q_2b_{31}+q_1q_3b_{12}-q_1^2b_{32})^2 + 2q_2q_3b_{11}(q_1q_2b_{31}+q_1q_3b_{12}-q_1^2b_{32})\\
&\qquad -2q_2q_3\tfrac{q_1}{M^2}(q_2b_{31}+q_3b_{12})b_{11} + 2q_2q_3\tfrac{q_1^2}{M^2}b_{11}b_{32} - 2q_2q_3\tfrac{q_1^2}{M^2}b_{11}b_{32}\\
&\quad = -\Big[q_1q_2b_{31}+q_1q_3b_{12}-q_1^2b_{32} - \tfrac{(M^2-1)}{M^2}q_2q_3b_{11}\Big]^2\\
&\qquad +\tfrac{q_1^4}{M^4}q_2^2q_3^2b_{11}^2 - q_2^2q_3^2b_{11}^2 - 2q_2q_3\tfrac{q_1^2}{M^2}b_{11}b_{32}.
\end{aligned}$$

It is easy to see that

$$\begin{aligned}
&2q_1q_2\tfrac{q_2^2}{M^2}(b_{12}+b_{21})b_{11} + 2q_1q_3\tfrac{q_3^2}{M^2}(b_{13}+b_{31})b_{11}\\
&\quad +2q_2q_3\tfrac{q_1}{M^2}(q_3b_{21}+q_2b_{13})b_{11} + 2q_2q_3\tfrac{q_1}{M^2}(q_2b_{31}+q_3b_{12})b_{11}\\
&\quad = 2q_1q_2(b_{12}+b_{21})b_{11} + 2q_1q_3(b_{13}+b_{31})b_{11}.
\end{aligned}$$

Substituting all these into (5), we finally obtain

$$\begin{aligned}\Delta_1 \;\leqslant\;& 2\|B\|_F^2 - b_{22}^2 - b_{33}^2 - b_{32}^2 - b_{23}^2 - 2q_1^2 b_{11}^2 - 2q_2^2 b_{11}b_{22} - 2q_3^2 b_{11}b_{33}\\ &-2q_2q_3b_{23}b_{11} - 2q_2q_3b_{32}b_{11} + \tfrac{q_1^4}{M^4}q_2^4b_{11}^2 - q_2^4b_{11}^2 - 2q_2^2\tfrac{q_1^2}{M^2}b_{11}b_{22} + \tfrac{q_1^4}{M^4}q_3^4b_{11}^2\\ &-q_3^4b_{11}^2 - 2q_3^2\tfrac{q_1^2}{M^2}b_{11}b_{33} + \tfrac{q_1^4}{M^4}q_2^2q_3^2b_{11}^2 - q_2^2q_3^2b_{11}^2 - 2q_2q_3\tfrac{q_1^2}{M^2}b_{11}b_{23} + \tfrac{q_1^4}{M^4}q_2^2q_3^2b_{11}^2\\ &-q_2^2q_3^2b_{11}^2 - 2q_2q_3\tfrac{q_1^2}{M^2}b_{11}b_{32}\\ =\;& 2\|B\|_F^2 - b_{22}^2 - b_{33}^2 - b_{32}^2 - b_{23}^2 + \left[(\tfrac{q_1^4}{M^4}-1)(q_2^2+q_3^2)^2 - 2q_1^2\right]b_{11}^2\\ &-2q_2^2(1+\tfrac{q_1^2}{M^2})b_{11}b_{22} - 2q_3^2(1+\tfrac{q_1^2}{M^2})b_{11}b_{33} - 2q_2q_3(1+\tfrac{q_1^2}{M^2})(b_{23}+b_{32})b_{11}\\ =\;& 2\|B\|_F^2 - b_{22}^2 - b_{33}^2 - b_{32}^2 - b_{23}^2 + [(1-M^2)^2 - M^4 - 2(1-M^2)]b_{11}^2\\ &-\tfrac{2}{M^2}[q_2^2b_{22} + q_3^2b_{33} + q_2q_3(b_{23}+b_{32})]b_{11}\\ \leqslant\;& 2\|B\|_F^2 - b_{11}^2 + \tfrac{1}{M^4}(q_2^4+q_3^4+2q_2^2q_3^2)b_{11}^2 = 2\|B\|_F^2.\end{aligned}$$

3. Outline of the proof for general n

We are going to give an outline of the proof of the conjecture for general n. Denote

$$B \equiv VYV^T \equiv [b_{ij}]_{i,j=1}^n, \qquad Q \equiv VU, \qquad C \equiv Q^TBQ \equiv [c_{ij}]_{i,j=1}^n.$$

For $n \geqslant 3$, we have similarly,

$$\|XY - YX\|_F^2 \leqslant \sum_{j=1}^n s_j^2 \Delta_j,$$

where

$$\Delta_j = b_{jj}^2 + c_{jj}^2 - 2b_{jj}c_{jj} + \sum_{k\neq j} b_{jk}^2 + \sum_{k\neq j} c_{kj}^2 + \sum_{k\neq j} c_{jk}^2 + \sum_{k\neq j} b_{kj}^2,$$

for $j = 1, 2, ..., n$. Once again we consider Δ_1 only. Let the first row of Q^T be $\mathbf{q} \equiv (q_1, ..., q_n)$ and $\mathbf{r}_1, ..., \mathbf{r}_n$ be the rows of B. Then by a similar argument as for $n = 3$, we obtain the following analogue of (4):

$$\begin{aligned}\Delta_1 =\;& \|B\|_F^2 - \sum_{k=2}^n\sum_{\ell=2}^n b_{k\ell}^2 - 2\left[\mathbf{q}\cdot(\mathbf{r}_1\cdot\mathbf{q}^T, ..., \mathbf{r}_n\cdot\mathbf{q}^T)^T\right]b_{11}\\ &+\sum_{i=1}^n q_i^2\|\mathbf{r}_i\|_2^2 + 2\sum_{i=1}^n\sum_{k>i} q_iq_k\mathbf{r}_i\cdot\mathbf{r}_k^T + \sum_{i=1}^n\sum_{k>i} I_{ik},\end{aligned} \tag{6}$$

where

$$I_{ik} \equiv \begin{vmatrix} q_i & q_k \\ \mathbf{r}_i\cdot\mathbf{q}^T & \mathbf{r}_k\cdot\mathbf{q}^T \end{vmatrix}^2.$$

Moreover, we have for $k > 1$,

$$\begin{aligned}
&2q_1q_k\mathbf{r}_1\cdot\mathbf{r}_k^T + I_{1k}\\
&\quad = 2q_1q_k\mathbf{r}_1\cdot\mathbf{r}_k^T + \Big[\sum_{j=1}^n q_j(q_1b_{kj} - q_kb_{1j})\Big]^2\\
&\quad \leqslant 2q_1q_k\mathbf{r}_1\cdot\mathbf{r}_k^T + \|\mathbf{q}\|_2^2\Big[\sum_{j=1}^n (q_1b_{kj} - q_kb_{1j})^2\Big] - \sum_{\ell=2}^n \begin{vmatrix} q_1 & q_\ell \\ q_1b_{k1} - q_kb_{11} & q_1b_{k\ell} - q_kb_{1\ell}\end{vmatrix}^2\\
&\quad = q_1^2\|\mathbf{r}_k\|_2^2 + q_k^2\|\mathbf{r}_1\|_2^2 - \sum_{\ell=2}^n J_{1k,\ell}
\end{aligned}$$

with

$$J_{1k,\ell} \equiv \begin{vmatrix} q_1 & q_\ell \\ q_1b_{k1} - q_kb_{11} & q_1b_{k\ell} - q_kb_{1\ell}\end{vmatrix}^2.$$

For $i \neq 1$, $k > i$, we have

$$2q_iq_k\mathbf{r}_i\cdot\mathbf{r}_k^T + I_{ik} \leqslant q_i^2\|\mathbf{r}_k\|_2^2 + q_k^2\|\mathbf{r}_i\|_2^2.$$

Note that

$$\sum_{i=1}^n q_i^2\|\mathbf{r}_i\|_2^2 + \sum_{i=1}^n\sum_{k>i}(q_i^2\|\mathbf{r}_k\|_2^2 + q_k^2\|\mathbf{r}_i\|_2^2) = \sum_{i=1}^n q_i^2\|\mathbf{r}_i\|_2^2 + \sum_{i=1}^n\sum_{k\neq i} q_i^2\|\mathbf{r}_k\|_2^2 = \|B\|_F^2.$$

Putting all of these into (6), we obtain the following inequality which is similar to (5),

$$\begin{aligned}
\Delta_1 &\leqslant 2\|B\|_F^2 - \sum_{k=2}^n\sum_{\ell=2}^n b_{k\ell}^2 - 2\Big[\mathbf{q}\cdot(\mathbf{r}_1\cdot\mathbf{q}^T,\ldots,\mathbf{r}_n\cdot\mathbf{q}^T)^T\Big]b_{11} - \sum_{k=2}^n\sum_{\ell=2}^n J_{1k,\ell}\\
&= 2\|B\|_F^2 - \sum_{k=2}^n\sum_{\ell=2}^n b_{k\ell}^2 - 2q_1^2b_{11}^2 - 2\Big(\sum_{k=2}^n\sum_{\ell=2}^n q_kq_\ell b_{k\ell}\Big)b_{11}\\
&\qquad -2\Big[\sum_{k=2}^n q_1q_k(b_{1k} + b_{k1})\Big]b_{11} - \sum_{k=2}^n\sum_{\ell=2}^n J_{1k,\ell}. \qquad (7)
\end{aligned}$$

To dominate $-2\Big[\sum_{k=2}^n q_1q_k(b_{1k}+b_{k1})\Big]b_{11}$ in (7), once again we note that $\Delta_1 \leqslant 2\|B\|_F^2$ obviously holds when $q_1^2 = 1$ (i.e., $q_2 = \cdots = q_n = 0$). Therefore, we assume that $q_1^2 \neq 1$ and denote $M^2 \equiv q_2^2 + \cdots + q_n^2 > 0$. Thus, we have for $k, \ell \geqslant 2$,

$$\begin{aligned}
&-J_{1k,\ell} - 2q_kq_\ell\tfrac{q_1}{M^2}(q_\ell b_{k1} + q_kb_{1\ell})b_{11}\\
&\quad = -q_k^2q_\ell^2b_{11}^2 - (q_1q_\ell b_{k1} + q_1q_kb_{1\ell} - q_1^2b_{k\ell})^2 + 2q_kq_\ell b_{11}(q_1q_\ell b_{k1} + q_1q_kb_{1\ell} - q_1^2b_{k\ell})\\
&\qquad -2q_kq_\ell\tfrac{q_1}{M^2}(q_\ell b_{k1} + q_kb_{1\ell})b_{11} + 2q_kq_\ell\tfrac{q_1^2}{M^2}b_{11}b_{k\ell} - 2q_kq_\ell\tfrac{q_1^2}{M^2}b_{11}b_{k\ell}\\
&\quad = -\Big[q_1q_\ell b_{k1} + q_1q_kb_{1\ell} - q_1^2b_{k\ell} - \tfrac{(M^2-1)}{M^2}q_kq_\ell b_{11}\Big]^2 + \tfrac{q_1^4}{M^4}q_k^2q_\ell^2b_{11}^2 - q_k^2q_\ell^2b_{11}^2\\
&\qquad -2q_kq_\ell\tfrac{q_1^2}{M^2}b_{11}b_{k\ell},
\end{aligned}$$

and

$$2\sum_{k=2}^{n}\sum_{\ell=2}^{n} q_k q_\ell \frac{q_1}{M^2}(q_\ell b_{k1} + q_k b_{1\ell})b_{11} = 2\Big[\sum_{k=2}^{n} q_1 q_k (b_{1k} + b_{k1})\Big] b_{11}.$$

We remark that the above choice is inspired by the case when $\ell = k$. In such case, a q_k^2/M^2 proportion of $2q_1q_k(b_{1k}+b_{k1})b_{11}$ is grouped into $J_{1k,k}$. Although it is intuitive, it seems not straightforward to introduce a term in the denominator. We finally arrive at

$$\begin{aligned}
\Delta_1 &\leqslant 2\|B\|_F^2 - \sum_{k=2}^{n}\sum_{\ell=2}^{n} b_{k\ell}^2 - 2q_1^2 b_{11}^2 - 2\Big(\sum_{k=2}^{n}\sum_{\ell=2}^{n} q_k q_\ell b_{k\ell}\Big) b_{11} \\
&\quad + \sum_{k=2}^{n}\sum_{\ell=2}^{n}\Big(\tfrac{q_1^4}{M^4} q_k^2 q_\ell^2 b_{11}^2 - q_k^2 q_\ell^2 b_{11}^2 - 2q_k q_\ell \tfrac{q_1^2}{M^2} b_{11} b_{k\ell}\Big) \\
&= 2\|B\|_F^2 - \sum_{k=2}^{n}\sum_{\ell=2}^{n} b_{k\ell}^2 + \Big[\Big(\tfrac{q_1^4}{M^4} - 1\Big)\Big(\sum_{k=2}^{n} q_k^2\Big)\Big(\sum_{\ell=2}^{n} q_\ell^2\Big) - 2q_1^2\Big] b_{11}^2 \\
&\quad -2\sum_{k=2}^{n}\sum_{\ell=2}^{n} q_k q_\ell \Big(1 + \tfrac{q_1^2}{M^2}\Big) b_{11} b_{k\ell} \\
&= 2\|B\|_F^2 - \sum_{k=2}^{n}\sum_{\ell=2}^{n} b_{k\ell}^2 + \Big[(1-M^2)^2 - M^4 - 2(1-M^2)\Big] b_{11}^2 \\
&\quad - \tfrac{2}{M^2}\sum_{k=2}^{n}\sum_{\ell=2}^{n} q_k q_\ell b_{11} b_{k\ell} \\
&\leqslant 2\|B\||_F^2 - b_{11}^2 + \tfrac{1}{M^4}\Big(\sum_{k=2}^{n}\sum_{\ell=2}^{n} q_k^2 q_\ell^2\Big) b_{11}^2 = 2\|B\|_F^2.
\end{aligned}$$

REFERENCES

[1] A. BÖTTCHER AND D. WENZEL, *How big can the commutator of two matrices be and how big is it typically?* Linear Algebra Appl., Vol. 403 (2005), pp. 216–228.

[2] R. HORN AND C. JOHNSON, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[3] L. LÁSZLÓ, *Proof of Böttcher and Wenzel's conjecture on commutator norms for* 3*-by-*3 *matrices*, Linear Algebra Appl., Vol. 422 (2007), pp. 659–663.

(Received December 11, 2007)

Seak-Weng Vong
Department of Mathematics
University of Macau
Macao, China
e-mail: swvong@umac.mo

Xiao-Qing Jin
Department of Mathematics
University of Macau
Macao, China
e-mail: xqjin@umac.mo