ESTIMATION OF THE SEPARATION OF TWO MATRICES*

SUN JI-GUANG (孙继广)

(Computing Center, Academia Sinica, Beijing, China)

Abstract

The estimation of the solution to the matrix equation $AX-XB=C$ is primarily dependent on the quantity $\operatorname{sep}(A, B)$ introduced by Stewart[5]. Varah[6] has given some examples to show that $\operatorname{sep}_F(A, B)$ can be very small even though the eigenvalues of A and B are well separated. In this paper we give some lower bounds of $\operatorname{sep}_F(A, B)$.

§ 1. Introduction

Investigations of perturbation bounds for invariant subspaces are frequently reduced to estimations of upper bounds for the solution X to the matrix equation[1,5]

$$AX-XB=C \quad (\lambda(A)\cap\lambda(B)=\emptyset),$$

where $\lambda(\cdot)$ denotes the set of all eigenvalues of a matrix, $\emptyset$ is the empty set. Stewart[5] has difined the separation between A and B

$$\operatorname{sep}(A, B)=\min_{\|X\|=1}\|AX-XB\|, \tag{1.1}$$

where $\|\ \|$ is any matrix norm; thus we obtain

$$\|X\|\leqslant\|C\|/\operatorname{sep}(A, B).$$

Therefore it is necessary to find lower bounds of $\operatorname{sep}(A, B)$ whenever one is investigating perturbations of invariant subspaces.

For A and B normal, Stewart[5] shows that if $\lambda(A)=\{\lambda_i\}$ and $\lambda(B)=\{\mu_i\}$ then

$$\operatorname{sep}_F(A, B)\equiv\min_{\|X\|_F=1}\|AX-XB\|_F=\min_{i,j}|\lambda_i-\mu_j|, \tag{1.2}$$

where $\|\ \|_F$ is the Frobenius norm. However, for A and B non-normal, and

$$\lambda(A)\cap\lambda(B)=\emptyset,$$

up to now we have only the following estimation[5]

$$0<\operatorname{sep}(A, B)\leqslant\min_{i,j}|\lambda_i-\mu_j|; \tag{1.3}$$

and Varah[6] has given some examples to show that $\operatorname{sep}_F(A, B)$ can be very small even though the eigenvalues of A and B are well separated.

In this paper we try to give some lower bounds of $\operatorname{sep}_F(A, B)$. We use reductions of A and B to Jordan canonical forms in § 2 and to some block diagonal forms in § 3.

Notation. The symbol $\mathbf{C}^{m\times n}$ denotes the set of complex $m\times n$ matrices. $I^{(n)}$ is the $n\times n$ identity matrix, and O is the null matrix. Sometimes we express the block

* Received June 1, 1982.

diagonal matrix $[A_1, \cdots, A_p]$ as $[\cdots, A_i, \cdots]_{(p)}$. $\begin{pmatrix} 0 & [A_1, \cdots, A_s] \\ 0 & 0 \end{pmatrix}_{(p)}$ denotes the matrix $\begin{pmatrix} 0 \cdots 0 & A_1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & A_s \\ & & & 0 \\ & & & \vdots \\ 0 & & & 0 \end{pmatrix}$ in which every row and column contains p submatrices.

Let $\| \; \|_2$ denote the spectral norm and $\kappa(Q) = \|Q\|_2 \|Q^{-1}\|_2$. For $A \in \mathbb{C}^{m\times m}$ with $\lambda(A) = \{\lambda_i\}$ we write $\Delta_F(A) = \{\|A\|_F^2 - \sum_{i=1}^{m} |\lambda_i|^2\}^{\frac{1}{2}}$.

§ 2. Lower bounds of $\mathrm{sep}_F(A, B)$ (I)

Let $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$, $\lambda(A) \cap \lambda(B) = \emptyset$ and $X \in \mathbb{C}^{m\times n}$. Now we consider to estimate lower bounds of the separation

$$\mathrm{sep}_F(A, B) \equiv \min_{\|X\|_F=1} \|AX - XB\|_F. \tag{2.1}$$

First of all we use the Kronecker product to get another representation for $\mathrm{sep}_F(A, B)$. The Kronecker product of any two matrices $C = (c_{ij}) \in \mathbb{C}^{p\times q}$ and $D \in \mathbb{C}^{r\times s}$ is the matrix $C \otimes D = (c_{ij}D) \in \mathbb{C}^{pr\times qs}$. We associate the matrix X in (2.1) with the mn-vector x which is the direct sum of the column vectors of X. Use the same method we associate the matrix $AX - XB$ in (2.1) with the mn-vector Tx, where

$$T = I^{(n)} \otimes A - B^T \otimes I^{(m)} \in \mathbb{C}^{mn\times mn} \tag{2.2}$$

(see [4], 8—9), B^T stand for transpose of B. From $\lambda(A) \cap \lambda(B) = \emptyset$, the matrix T is nonsingular (see [3], 259. Theorem 8.3.1), thus we obtain

$$\mathrm{sep}_F(A, B) = \min_{\|x\|_2=1} \|Tx\|_2 = \min_{x\neq 0} \frac{\|Tx\|_2}{\|x\|_2} = \left(\max_{y\neq 0} \frac{\|T^{-1}y\|_2}{\|y\|_2}\right)^{-1} = \|T^{-1}\|_2^{-1}. \tag{2.3}$$

Suppose that the Jordan canonical decomposition of A and B^T are

$$A = Q_A J_A Q_A^{-1}, \qquad B^T = Q_B J_B Q_B^{-1}, \tag{2.4}$$

where

$$\begin{aligned} &J_A = \Lambda_A + N_A,\ \Lambda_A = [\cdots, \lambda_i I^{(m_i)}, \cdots]_{(p)},\ N_A = [\cdots, N_i(A), \cdots]_{(p)}, \\ &\qquad N_i(A) = [\cdots, N_{i,k}(A), \cdots]_{(k_i)} \in \mathbb{C}^{m_i\times m_i},\ 1 \leqslant i \leqslant p; \\ &J_B = \Lambda_B + N_B,\ \Lambda_B = [\cdots, \mu_j I^{(n_j)}, \cdots]_{(q)},\ N_B = [\cdots, N_j(B), \cdots]_{(q)}, \\ &\qquad N_j(B) = [\cdots, N_{j,l}(B), \cdots]_{(l_j)} \in \mathbb{C}^{n_j\times n_j},\ 1 \leqslant j \leqslant q. \end{aligned} \tag{2.5}$$

All the matrices $N_{i,k}(A) \in \mathbb{C}^{m_{i,k}\times m_{i,k}}$ and $N_{j,l}(B) \in \mathbb{C}^{n_{j,l}\times n_{j,l}}$ are nilpotent as $\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}$, $\sum_{k=1}^{k_i} m_{i,k} = m_i$, $\sum_{i=1}^{p} m_i = m$, $\sum_{l=1}^{l_j} n_{j,l} = n_j$ and $\sum_{j=1}^{q} n_j = n$. $\lambda_s \neq \lambda_t$ and $\mu_s \neq \mu_t$ if $s \neq t$. The highest orders of the Jordan blocks of J_A and J_B are r_A and r_B respectively, i.e.

$$r_A = \max_{i,k} \{m_{i,k}\}, \quad r_B = \max_{j,l} \{n_{j,l}\}. \tag{2.6}$$

Thus by (2.1) and (2.4) we have

$$\begin{aligned}\operatorname{sep}_F(A, B) &= \inf_{X\neq 0} \|AX - XB\|_F / \|X\|_F \\ &= \inf_{X\neq 0} \|Q_A(J_A Q_A^{-1} X Q_B^{-T} - Q_A^{-1} X Q_B^{-T} J_B^T) Q_B^T\|_F / \|X\|_F \\ &\geqslant \operatorname{sep}_F(J_A, J_B^T) / [\varkappa(Q_A)\varkappa(Q_B)].\end{aligned} \tag{2.7}$$

Moreover, from (2.2), (2.3) and (2.5),

$$\begin{aligned}\operatorname{sep}_F^{-1}(J_A, J_B^T) &= \|(I^{(n)}\otimes J_A - J_B\otimes I^{(m)})^{-1}\|_2 \\ &\leqslant \|[I + (I\otimes \Lambda_A - \Lambda_B\otimes I)^{-1}(I\otimes N_A - N_B\otimes I)]^{-1}\|_2 \|(I\otimes\Lambda_A - \Lambda_B\otimes I)^{-1}\|_2 \\ &= \frac{\|[I + (\hat{N}_A - \hat{N}_B)]^{-1}\|_2}{\delta(A, B)},\end{aligned} \tag{2.8}$$

where $\delta(A, B) = \min_{i,j} |\lambda_i - \mu_j|$, and

$$\hat{N}_A = \Omega\cdot I\otimes N_A,\ \hat{N}_B = \Omega\cdot N_B\otimes I,\ \Omega = (I\otimes\Lambda_A - \Lambda_B\otimes I)^{-1} = [\cdots, \Omega_j, \cdots]_{(n)}, \tag{2.9}$$

all the $\Omega_j (1\leqslant j\leqslant n)$ are $m\times m$ diagonal matrices.

Lemma 2.1. *If $r > r_A + r_B - 2$, then $(\hat{N}_A - \hat{N}_B)^r = 0$.*

Proof. Develop

$$(\hat{N}_A - \hat{N}_B)^r = \sum_{\alpha+\beta=r} (-1)^\beta \hat{N}_A^{\alpha_1}\hat{N}_B^{\beta_1}\cdots\hat{N}_A^{\alpha_h}\hat{N}_B^{\beta_h},\quad \alpha = \sum_{i=1}^h \alpha_i,\ \beta = \sum_{i=1}^h \beta_i. \tag{2.10}$$

i) According to the definition of the Kronecker product we have

$$\hat{N}_A = [\cdots, N_i, \cdots]_{(n)},\quad N_i = \Omega_i N_A,\ 1\leqslant i\leqslant n.$$

Now we rewrite $N_1, N_2, \cdots, N_n$ as

$$N^{(1)}_{1,1,1}, \cdots, N^{(1)}_{1,1,n_{1,1}}, N^{(1)}_{1,2,1}, \cdots, N^{(1)}_{1,2,n_{1,2}}, \cdots, N^{(1)}_{q,l_q,1}, \cdots, N^{(1)}_{q,l_q,n_{q,l_q}},$$

thus

$$\hat{N}_A = [\cdots, N^{(1)}_{j,l,t}, \cdots]_{(n)}, \tag{2.11}$$

and by (2.5), for every principal square submatrix of $\hat{N}_A$ we have

$$N^{(1)}_{j,l,t} = [N^{(1)}_{j,l,t,1}, \cdots, N^{(1)}_{j,l,t,k_1+\cdots+k_p}] \in \mathbb{C}^{m\times m}, \tag{2.12}$$

the order of every $N^{(1)}_{j,l,t,s}$ is not greater than r_A, and

$$N^{(1)}_{j,l,t,s} = \begin{pmatrix} 0 & * & 0 \\ & \ddots & \ddots & \\ & & \ddots & * \\ 0 & & & 0 \end{pmatrix},\quad 1\leqslant s\leqslant k_1 + \cdots + k_p. \tag{2.13}$$

With same way as above we rewrite $\Omega_1, \Omega_2, \cdots, \Omega_n$ as

$$\Omega^{(1)}_{1,1,1}, \cdots, \Omega^{(1)}_{1,1,n_{1,1}}, \Omega^{(1)}_{1,2,1}, \cdots, \Omega^{(1)}_{1,2,n_{1,2}}, \cdots, \Omega^{(1)}_{q,l_q,1}, \cdots, \Omega^{(1)}_{q,l_q,n_q,l_q},$$

where all the $\Omega^{(1)}_{j,l,t}$ are $m\times m$ diagonal matrices. By the definition of the Kronecker product,

$$\hat{N}_B = \left[\cdots, \begin{pmatrix} 0 & [\Omega^{(1)}_{j,l,1}, \cdots, \Omega^{(1)}_{j,l,n_{j,l}-1}] \\ 0 & 0 \end{pmatrix}_{(n_{j,l})}, \cdots\right]_{(l_1+\cdots+l_q)}. \tag{2.14}$$

ii) From (2.11)—(2.13), $\hat{N}_A^{\alpha_1} = 0$ if $\alpha_1 \geqslant r_A$; and for $\alpha_1 < r_A$, if we write $(N^{(1)}_{j,l,t})^{\alpha_1} = N^{(\alpha_1)}_{j,l,t}$ and $(N^{(1)}_{j,l,t,s})^{\alpha_1} = N^{(\alpha_1)}_{j,l,t,s}$, then

$$\hat{N}_A^{\alpha_1} = [\cdots, N^{(\alpha_1)}_{j,l,t}, \cdots]_{(n)},$$

$$N^{(\alpha_1)}_{j,l,t} = [N^{(\alpha_1)}_{j,l,t,1}, \cdots, N^{(\alpha_1)}_{j,l,t,k_1+\cdots+k_p}] \in \mathbb{C}^{m\times m},\quad 1\leqslant t\leqslant n_{j,l},\ 1\leqslant l\leqslant l_j,\ 1\leqslant j\leqslant q.$$

Where the order of every $N_{j,l,t,s}^{(\alpha_1)}$ is not greater than r_A, and

$$N_{j,l,t,s}^{(\alpha_1)}=\begin{pmatrix} 0\cdots 0 & * & 0 \\ & \ddots\ddots & \ddots\ * \\ & & \ddots\ 0 \\ 0 & & \vdots \\ & & 0 \end{pmatrix}\Bigg\}\alpha_1, \quad 1\leqslant s\leqslant k_1+\cdots+k_p.$$

By (2.6), $n_{j,l}\leqslant r_B$ ($1\leqslant l\leqslant l_j$, $1\leqslant j\leqslant q$). Hence, it follows from (2.14) that $\hat{N}_B^{\beta_1}=0$ if $\beta_1\geqslant r_B$; and for $\beta_1<r_B$ we have

$$\hat{N}_B^{\beta_1}=\left[\cdots,\begin{pmatrix}0 & [\Omega_{j,l,1}^{(\beta_1)},\ \cdots,\ \Omega_{j,l,n_{j,l}-\beta_1}^{(\beta_1)}] \\ 0 & 0\end{pmatrix}_{(n_{j,l})},\ \cdots\right]_{(l_1+\cdots+l_q)},$$

where all the $\Omega_{j,l,t}^{(\beta_1)}$ ($1\leqslant t\leqslant n_{j,l}-\beta_1$) are $m\times m$ diagonal matrices.

iii) Therefore, we have $\hat{N}_A^{\alpha_1}\hat{N}_B^{\beta_1}=0$ if $\alpha_1\geqslant r_A$ or $\beta_1\geqslant r_B$; and for $\alpha_1<r_A$ and $\beta_1<r_B$, if we write $N_{j,l,t}^{(\alpha_1)}\Omega_{j,l,t}^{(\beta_1)}=N_{j,l,t}^{(\alpha_1,\beta_1)}$, then

$$\hat{N}_A^{\alpha_1}\hat{N}_B^{\beta_1}=\left[\cdots,\begin{pmatrix}0 & [N_{j,l,1}^{(\alpha_1,\beta_1)},\ \cdots,\ N_{j,l,n_{j,l}-\beta_1}^{(\alpha_1,\beta_1)}] \\ 0 & 0\end{pmatrix}_{(n_{j,l})},\ \cdots\right]_{(l_1+\cdots+l_q)}, \tag{2.15}$$

$$N_{j,l,t}^{(\alpha_1,\beta_1)}=[N_{j,l,t,1}^{(\alpha_1,\beta_1)},\ \cdots,\ N_{j,l,t,k_1+\cdots+k_p}^{(\alpha_1,\beta_1)}]\in\mathbb{C}^{m\times m},\quad 1\leqslant t\leqslant n_{j,l},\ 1\leqslant l\leqslant l_j,\ 1\leqslant j\leqslant q, \tag{2.16}$$

the order of every $N_{j,l,t,s}^{(\alpha_1,\beta_1)}$ is not greater than r_A, and

$$N_{j,l,t,s}^{(\alpha_1,\beta_1)}=\begin{pmatrix} 0\cdots 0 & * & 0 \\ & \ddots\ddots & \ddots\ * \\ & & \ddots\ 0 \\ 0 & & \vdots \\ & & 0 \end{pmatrix}\Bigg\}\alpha_1, \quad 1\leqslant t\leqslant k_1+\cdots+k_p. \tag{2.17}$$

iv) Consequently, from (2.15)—(2.17) we see that

$$\hat{N}_A^{\alpha_1}\hat{N}_B^{\beta_1}\cdots\hat{N}_A^{\alpha_h}\hat{N}_B^{\beta_h}=0\quad\text{if}\quad \alpha=\alpha_1+\cdots+\alpha_h\geqslant r_A\quad\text{or}\quad \beta=\beta_1+\cdots+\beta_h\geqslant r_B$$

in (2.10). But $r>r_A+r_B-2$ implies $\alpha\geqslant r_A$ or $\beta\geqslant r_B$, therefore the lemma is valid. ∎

Applying Lemma 2.1 to the inequality (2.8), we get

$$\begin{aligned}\mathrm{sep}_F^{-1}(J_A,\ J_B^T)&\leqslant\frac{\left\|I+\sum_{r=1}^{r_A+r_B-2}(-1)^r[(I\otimes A_A-A_B\otimes I)^{-1}(I\otimes N_A-N_B\otimes I)]^r\right\|_2}{\delta(A,\ B)}\\ &\leqslant\left[1+\sum_{r=1}^{r_A+r_B-2}\left(\frac{\nu(r_A,\ r_B)}{\delta(A,\ B)}\right)^r\right]\Big/\delta(A,\ B),\end{aligned}\tag{2.18}$$

where

$$\nu(r_A,\ r_B)=\|I\otimes N_A\|_2+\|N_B\otimes I\|_2=\begin{cases}0 & \text{if } r_A=r_B=1,\\ 1 & \text{if } r_A=1 \text{ or } r_B=1 \text{ but } r_Ar_B>1,\\ 2 & \text{if } r_A,\ r_B\geqslant 2.\end{cases}\tag{2.19}$$

Combining the inequalities (1.3), (2.7) and (2.18), we obtain

Theorem 2.1. *Let A and B be any two matrices with $\lambda(A)=\{\lambda_i\}$ and $\lambda(B)=\{\mu_j\}$, and*

$$\delta(A,\ B)=\min_{i,j}|\lambda_i-\mu_j|.$$

Suppose that Q_A and Q_B are the transformation matrices which transpose A and B to the

Jordan canonical forms, and the highest orders of the Jordan blocks of A and B are r_A and r_B, respectively. Then

$$\frac{\nu(r_A, r_B)-\delta(A, B)}{(\nu(r_A, r_B))^{r_A+r_B-1}-(\delta(A, B))^{r_A+r_B-1}}\cdot\frac{(\delta(A, B))^{r_A+r_B-1}}{\varkappa(Q_A)\varkappa(Q_B)}\leqslant \mathrm{sep}_F(A, B)\leqslant\delta(A, B). \tag{2.20}$$

Where $\nu(r_A, r_B)$ is denoted by (2.19).

Specially, for normal matrices A and B, the inequality (2.20) becomes the equality (1.2); and it follows from (2.20) that for diagonalizable matrices A and B we have[5]

$$\frac{\delta(A, B)}{\varkappa(Q_A)\varkappa(Q_B)}\leqslant \mathrm{sep}_F(A, B)\leqslant\delta(A, B).$$

Varah[6] has given the following example to show that $\mathrm{sep}_F(A, B)$ can be very small even though $\lambda(A)$ and $\lambda(B)$ are well separated. Varah has computed the $\mathrm{sep}_F(A, B)$ for some A and B (these values of $\mathrm{sep}_F(A, B)$ are denoted by sep in Table 1), applying Theorem 2.1 we give satisfactory lower bounds for these $\mathrm{sep}_F(A, B)$ and the lower bounds are denoted by s in Table 1.

Example 2. 1.

$$A=\begin{pmatrix} 1 & -1 & & \\ & \ddots & \ddots & \\ & & \ddots & -1 \\ & & & 1\end{pmatrix}_{m\times m}, \qquad B=\begin{pmatrix} 1-\alpha & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 1-\alpha\end{pmatrix}_{n\times n}, \quad \alpha>0.$$

For $Q_A=[i, -i, \cdots, (-1)^{m-1}i]$ and $Q_B=\begin{pmatrix} & & 1\\ & 1 & \\ 1^{\cdot^{\cdot^{\cdot}}} & & \end{pmatrix}$, $i=\sqrt{-1}$, we have

$$Q_A^{-1}AQ_A=\begin{pmatrix} 1 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 1\end{pmatrix}, \qquad Q_B^{-1}B^TQ_B=\begin{pmatrix} 1-\alpha & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 1-\alpha\end{pmatrix}.$$

Obviously, $\varkappa(Q_A)=\varkappa(Q_B)=1$, $\delta(A, B)=\alpha$, $r_A=m$ and $r_B=n$. Therefore, by Theorem 2.1 we have

$$\mathrm{sep}_F(A, B)\geqslant\frac{(2-\alpha)\alpha^{m+n-1}}{2^{m+n-1}-\alpha^{m+n-1}}\equiv s.$$

Some computed results are given in Table 1.

Table 1

m	n	α	sep	s
4	4	$\frac{1}{2}$	$3.4_{10}-4$	$9.2_{10}-5$
6	3	$\frac{1}{4}$	$7.0_{10}-7$	$1.0_{10}-7$
6	4	$\frac{1}{8}$	$1.3_{10}-10$	$2.7_{10}-11$
6	6	$\frac{1}{16}$	$2.2_{10}-16$	$5.4_{10}-17$

§ 3. Lower bounds of $\operatorname{sep}_F(A, B)$ (II)

From [2] we gain inspiration and suggest a method to estimate lower bounds of $\operatorname{sep}_F(A, B)$. The main points of this method is that the estimation of lower bounds of $\operatorname{sep}_F(A, B)$ is reduced to estimation of upper bounds for the condition numbers $\varkappa(Q_A)$ and $\varkappa(Q_B)$ of the transformation matrices Q_A and Q_B which transform A and B to some block diagonal forms respectively. But it differs from [2] in the following two points: 1° We reduce estimations of upper bounds for $\varkappa(Q_A)$ and $\varkappa(Q_B)$ to estimations of lower bounds for a series of sep_F corresponding to certain matrices of lower order, and the lower bounds of these sep_F can be obtained by the same argument in Theorem 2.1; 2° The way to estimate upper bounds for $\|Q_A^{-1}\|_2$ is similar to that for $\|Q_A\|_2$, and the development of $(I-X)^{-1}$ is unnecessary (see Lemma 3.1). The results obtained in this section show that for any matrix A with closed eigenvalues (by "closed eigenvalues" we mean that the smallest distance between the eigenvalues of A is smaller than $\Delta_F(A)$) the upper bound for $\varkappa(Q_A)$ obtained by our method is sharper than that in [2]. For this reason the way to estimate upper bound for $\varkappa(Q_A)$ clarified in this section may also be useful for perturbation analysis of eigenvalues (ref. [2]).

Lemma 3.1. *Let $A\in\mathbb{C}^{m\times m}$ be an upper triangular matrix:*

$$A=\begin{pmatrix} A_1 & \multicolumn{3}{c}{\boxed{\quad P_1 \quad}} \\ & A_2 & \multicolumn{2}{c}{\boxed{\quad P_2 \quad}} \\ & & \ddots & \vdots \\ & & A_{p-1} & P_{p-1} \\ & & & A_p \end{pmatrix}=\begin{pmatrix} A_1 & P_1' & \multirow{3}{*}{$\boxed{P_{p-2}'}$} & \multirow{4}{*}{$\boxed{P_{p-1}'}$} \\ & A_2 & & \\ & & \ddots & \\ & & A_{p-1} & \\ & & & A_p \end{pmatrix}. \tag{3.1}$$

Here $A_i=\lambda_i I^{(m_i)}+H_i$ and H_i is an upper triangular matrix with zeros on its diagonal, $1\leqslant i\leqslant p$; and $\lambda_i\neq\lambda_j$ for $i\neq j$. Then there is a nonsingular matrix Q_A such that

$$Q_A^{-1}AQ_A=[A_1, \cdots, A_p]\equiv A_{(1)}. \tag{3.2}$$

Moreover, if we set

$$A_{(i)}=[A_i, \cdots, A_p],\ 2\leqslant i\leqslant p;\quad A^{(i)}=[A_1, \cdots, A_i],\ 1\leqslant i\leqslant p-1, \tag{3.3}$$

and

$$\xi_{p-1}=\frac{\|P_{p-1}\|_F}{\operatorname{sep}_F(A_{p-1}, A_p)},\qquad \eta_1=\frac{\|P_1'\|_F}{\operatorname{sep}_F(A_1, A_2)}, \tag{3.4}$$

$$\xi_{p-k}=\frac{\left(1+\sqrt{\sum_{i=1}^{k-1}\xi_{p-i}^2}\right)\|P_{p-k}\|_F}{\operatorname{sep}_F(A_{p-k}, A_{(p-k+1)})},\quad k=2, 3, \cdots, p-1, \tag{3.5}$$

$$\eta_k=\frac{\left(1+\sqrt{\sum_{i=1}^{k-1}\eta_i^2}\right)\|P_k'\|_F}{\operatorname{sep}_F(A^{(k)}, A_{k+1})},\quad k=2, 3, \cdots, p-1; \tag{3.6}$$

then

$$\varkappa(Q_A)\leqslant\left(1+\sqrt{\sum_{k=1}^{p-1}\xi_{p-k}^2}\right)\left(1+\sqrt{\sum_{k=1}^{p-1}\eta_k^2}\right). \tag{3.7}$$

Proof. We choose

$$Q_A = I - X, \quad X = \begin{pmatrix} O^{(m_1)} & \boxed{\quad X_1 \quad} & & \\ & O^{(m_2)} & \boxed{\quad X_2 \quad} & \\ & & \ddots & \vdots \\ & & O^{(m_{p-1})} & X_{p-1} \\ & & & O^{(m_p)} \end{pmatrix}, \tag{3.8}$$

From (3.2),

$$AQ_A = Q_A A_{(1)}. \tag{3.9}$$

Substituting (3.8) and the first expression of A in (3.1), and comparing the two sides of the equation, we get

$$A_{p-1}X_{p-1} - X_{p-1}A_p = P_{p-1}, \tag{3.10}$$

and for $k=2, 3, \cdots, p-1$, we have

$$A_{p-k}X_{p-k} - X_{p-k}A_{(p-k+1)} = \tilde{P}_{p-k}, \tag{3.11}$$

here

$$\tilde{P}_{p-k} = P_{p-k}\begin{pmatrix} I & \boxed{\quad -X_{p-k+1} \quad} & \\ & I \quad \boxed{\quad -X_{p-k+2}} \\ & \ddots & \vdots \\ & I & -X_{p-1} \end{pmatrix}. \tag{3.12}$$

From $\lambda(A_{p-1}) \cap \lambda(A_p) = \emptyset$, the equation (3.10) has a unique solution X_{p-1} (ref. [3], Theorem 8.5.1), and

$$\|X_{p-1}\|_F \leqslant \frac{\|P_{p-1}\|_F}{\mathrm{sep}_F(A_{p-1}, A_p)} = \xi_{p-1}. \tag{3.13}$$

For $k=2, 3, \cdots, p-1$, it follows from $\lambda(A_{p-k}) \cap \lambda(A_{(p-k+1)}) = \emptyset$ that the equation (3.11) has a unique solution X_{p-k}, and

$$\|X_{p-k}\|_F \leqslant \frac{\|\tilde{P}_{p-k}\|_F}{\mathrm{sep}_F(A_{p-k}, A_{(p-k+1)})}, \tag{3.14}$$

here $\tilde{P}_{p-k}$ is expressed by (3.12), and so

$$\|\tilde{P}_{p-k}\|_F \leqslant \left(1 + \sqrt{\sum_{i=1}^{k-1} \|X_{p-i}\|_F^2}\right) \|P_{p-k}\|_F. \tag{3.15}$$

Hence, if $\xi_{p-1}, \xi_{p-2}, \cdots, \xi_1$ are computed successively according to the formulas (3.4) and (3.5), then from (3.13)—(3.15) we have $\|X_{p-k}\|_F \leqslant \xi_{p-k}$ $(1 \leqslant k \leqslant p-1)$; thus by (3.8) we get

$$\|Q_A\|_2 \leqslant 1 + \|X\|_2 \leqslant 1 + \sqrt{\sum_{k=1}^{p-1} \|X_{p-k}\|_F^2} \leqslant 1 + \sqrt{\sum_{k=1}^{p-1} \xi_{p-k}^2}. \tag{3.16}$$

Now we estimate the $\|Q_A^{-1}\|_2$ from the above. It is easy to see that

$$Q_A^{-1} = (I-X)^{-1} = I + Y, \quad Y = \begin{pmatrix} O^{(m_1)} & Y_1 \cdots & \boxed{} & \boxed{} \\ & O^{(m_2)} & Y_{p-2} & \\ & & \ddots & Y_{p-1} \\ & & O^{(m_{p-1})} & \\ & & & O^{(m_p)} \end{pmatrix}, \tag{3.17}$$

and Q_A^{-1} satisfies

$$Q_A^{-1}A = A_{(1)}Q_A^{-1}. \tag{3.18}$$

Substituting (3.17) and the second expression of A in (3.1) into (3.18), and comparing the two sides of the equation, we get

$$A_1Y_1-Y_1A_2=P_1'; \quad A^{(k)}Y_k-Y_kA_{k+1}=\widetilde{P}_k', \quad k=2,3,\cdots,p-1,$$

here

$$\widetilde{P}_k'=\begin{pmatrix} I & Y_1 & \cdots & \boxed{Y_{k-2}} & \boxed{Y_{k-1}} \\ & I & & & \\ & & \ddots & & \\ & & & I & \\ & & & & I \end{pmatrix} P_k'.$$

It follows from a similar argument as above (see (3.13)—(3.15)) that if $\eta_1, \eta_2, \cdots, \eta_{p-1}$ are computed successively according to the formulas (3.4) and (3.6), then we have $\|Y_k\|_F\leqslant\eta_k$ $(1\leqslant k\leqslant p-1)$; thus by (3.17) we get

$$\|Q_A^{-1}\|_2\leqslant 1+\|Y\|_2\leqslant 1+\sqrt{\sum_{k=1}^{p-1}\|Y_k\|_F^2}\leqslant 1+\sqrt{\sum_{k=1}^{p-1}\eta_k^2}. \tag{3.19}$$

Combining (3.16) and (3.19), we obtain the estimation (3.7). ∎

Lemma 3.2. *Let $A_{(1)}=[A_1, \cdots, A_p]\in\mathbb{C}^{m\times m}$ and $B_{(1)}=[B_1, \cdots, B_q]\in\mathbb{C}^{n\times n}$. Here $A_i=\lambda_iI^{(m_i)}+H_i$ and H_i is an upper triangular matrix with zeros on its diagonal, $1\leqslant i\leqslant p$, and $\lambda_k\neq\lambda_l$ for $k\neq l$; $B_j=\mu_jI^{(n_j)}+K_j$ and K_j is an upper triangular matrix with zeros on its diagonal, $1\leqslant j\leqslant q$, and $\mu_k\neq\mu_l$ for $k\neq l$. Then*

$$\operatorname{sep}_F(A_{(1)}, B_{(1)})\geqslant\frac{[\Delta(A_{(1)}, B_{(1)})-1]\delta(A_{(1)}, B_{(1)})}{[\Delta(A_{(1)}, B_{(1)})]^{s_A+s_B-1}-1}\equiv s(A_{(1)}, B_{(1)}), \tag{3.20}$$

where

$$\delta(A_{(1)}, B_{(1)})=\min_{i,j}|\lambda_i-\mu_j|, \quad \Delta(A_{(1)}, B_{(1)})=\frac{\Delta_F(A_{(1)})+\Delta_F(B_{(1)})}{\delta(A_{(1)}, B_{(1)})},$$

$$s_A=\max_i\{m_i\}, \quad s_B=\max_j\{n_j\}.$$

Proof. Let $S=\begin{pmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{pmatrix}\in\mathbb{C}^{n\times n}$ and $SB_{(1)}S=\widetilde{B}^T$, then by § 2 we have

$$\operatorname{sep}_F(A_{(1)}, B_{(1)})=\operatorname{sep}_F(A_{(1)}, \widetilde{B}^T),$$

where $\widetilde{B}=[\widetilde{B}_q, \cdots, \widetilde{B}_1]$, $\widetilde{B}_j=\mu_jI^{(n_j)}+\widetilde{K}_j$, $\widetilde{K}_j$ is an upper triangular matrix with zeros on its diagonal, and $\Delta_F(\widetilde{B}_j)=\Delta_F(B_j)$, $1\leqslant j\leqslant q$. Therefore

$$\operatorname{sep}_F^{-1}(A_{(1)}, B_{(1)})=\|(I^{(n)}\otimes A_{(1)}-\widetilde{B}\otimes I^{(m)})^{-1}\|_2. \tag{3.21}$$

Writing $A_{(1)}=\Lambda_A+M_A$ and $\widetilde{B}=\Lambda_B+M_B$, here $\Lambda_A=[\lambda_1I^{(m_1)}, \cdots, \lambda_pI^{(m_p)}]$, $\Lambda_B=[\mu_qI^{(n_q)}, \cdots, \mu_1I^{(n_1)}]$, $M_A=[H_1, \cdots, H_p]$ and $M_B=[\widetilde{K}_1, \cdots, \widetilde{K}_q]$. Substituting (3.21), we get

$$\begin{aligned}\operatorname{sep}_F^{-1}(A_{(1)}, B_{(1)})&\leqslant\|[I+(I\otimes\Lambda_A-\Lambda_B\otimes I)^{-1}(I\otimes M_A-M_B\otimes I)]^{-1}\|_2\\&\quad\times\|(I\otimes\Lambda_A-\Lambda_B\otimes I)^{-1}\|_2\\&=\|[I+(\hat{M}_A-\hat{M}_B)]^{-1}\|_2/\delta(A_{(1)}, B_{(1)}),\end{aligned} \tag{3.22}$$

where $\hat{M}_A=(I\otimes\Lambda_A-\Lambda_B\otimes I)^{-1}I\otimes M_A$ and $\hat{M}_B=(I\otimes\Lambda_A-\Lambda_B\otimes I)^{-1}M_B\otimes I$.

Applying the same argument as in Lemma 2.1 we can prove that $(\hat{M}_A-\hat{M}_B)^s=0$ for $s>s_A+s_B-2$. Hence

$$[I+(\hat{M}_A-\hat{M}_B)]^{-1}=I+\sum_{s=1}^{s_A+s_B-2}(-1)^s(\hat{M}_A-\hat{M}_B)^s.$$

Substituting (3.22), we get

$$\operatorname{sep}_F^{-1}(A_{(1)}, B_{(1)})\leqslant\frac{1}{\delta(A_{(1)}, B_{(1)})}\left[1+\sum_{s=1}^{s_A+s_B-2}\left(\frac{\|I\otimes M_A-M_B\otimes I\|_2}{\delta(A_{(1)}, B_{(1)})}\right)^s\right]. \tag{3.23}$$

Observe that

$$\|I\otimes M_A-M_B\otimes I\|_2\leqslant\|I\otimes M_A\|_2+\|M_B\otimes I\|_2=\|M_A\|_2+\|M_B\|_2\leqslant\Delta_F(A_{(1)})+\Delta_F(B_{(1)}),$$

substituting (3.23), we obtain the estimation (3.20). ∎

The following lemma is a corollary of Lemma 3.2.

Lemma 3.3. *Suppose that A, $A_{(i)}$ and $A^{(i)}$ are denoted by (3.1) and (3.3). Then for $k=1, 2, \cdots, p-1$, we have*

$$\mathrm{sep}_F(A_{p-k},\ A_{(p-k+1)})\geqslant\frac{[\Delta_{(p-k)}(A)-1]\delta_{(p-k)}(A)}{[\Delta_{(p-k)}(A)]^{m_{p-k}+m_{(p-k+1)}-1}-1}\equiv s_{(p-k)}(A), \tag{3.24}$$

$$\mathrm{sep}_F(A^{(k)},\ A_{k+1})\geqslant\frac{[\Delta^{(k+1)}(A)-1]\delta^{(k+1)}(A)}{[\Delta^{(k+1)}(A)]^{m^{(k)}+m_{k+1}-1}-1}\equiv s^{(k+1)}(A); \tag{3.25}$$

where

$$\delta_{(p-k)}(A)=\min_{p-k+1\leqslant i\leqslant p}|\lambda_{p-k}-\lambda_i|,\quad \Delta_{(p-k)}(A)=\frac{\Delta_F(A_{p-k})+\Delta_F(A_{(p-k+1)})}{\delta_{(p-k)}(A)}, \tag{3.26}$$

$$\delta^{(k+1)}(A)=\min_{1\leqslant i\leqslant k}|\lambda_i-\lambda_{k+1}|,\quad \Delta^{(k+1)}(A)=\frac{\Delta_F(A^{(k)})+\Delta_F(A_{k+1})}{\delta^{(k+1)}(A)}, \tag{3.27}$$

$$m_{(p-k+1)}=\max_{p-k+1\leqslant i\leqslant p}\{m_i\},\quad m^{(k)}=\max_{1\leqslant i\leqslant k}\{m_i\}. \tag{3.28}$$

Combining Lemma 3.1 with Lemma 3.3, we get

Lemma 3.4. *Suppose that $A\in\mathbb{C}^{m\times m}$ is explained as in Lemma 3.1. Then there is a nonsingular matrix Q_A such that the equality (3.2) is valid, and*

$$\varkappa(Q_A)\leqslant\varkappa_A, \tag{3.29}$$

where the $\varkappa_A$ can be computed by the following steps (1)—(3):

(1) *According to (3.24) and (3.25) compute $s_{(p-k)}(A)$ and $s^{(k+1)}(A)$, $k=1, 2, \cdots, p-1$;*

(2) *Let*

$$x_{p-1}=\|P_{p-1}\|_F/s_{(p-1)}(A),\quad y_1=\|P'_1\|_F/s^{(2)}(A), \tag{3.30}$$

and then compute successively

$$x_{p-k}=\left(1+\sqrt{\sum_{i=1}^{k-1}x_{p-i}^2}\right)\|P_{p-k}\|_F/s_{(p-k)}(A) \tag{3.31}$$

and

$$y_k=\left(1+\sqrt{\sum_{i=1}^{k-1}y_i^2}\right)\|P'_k\|_F/s^{(k+1)}(A), \tag{3.32}$$

$k=2, 3, \cdots, p-1$;

(3) *Compute*

$$\varkappa_A=\left(1+\sqrt{\sum_{k=1}^{p-1}x_{p-k}^2}\right)\left(1+\sqrt{\sum_{k=1}^{p-1}y_k^2}\right). \tag{3.33}$$

From Lemmas 3.1—3.4 we obtain

Theorem 3.1. *Suppose that $A\in\mathbb{C}^{m\times m}$ and $B\in\mathbb{C}^{n\times n}$ are upper triangular matrices, A is denoted by (3.1), and*

$$B=\begin{pmatrix} B_1 & \boxed{\quad Q_1 \quad} & & \\ & B_2 & \boxed{\quad Q_2 \quad} & \\ & & \ddots & \vdots \\ & & B_{q-1} & Q_{q-1} \\ & & & B_q \end{pmatrix}=\begin{pmatrix} B_1 & Q'_1 & \cdots & \boxed{Q'_{q-2}} & \boxed{Q'_{q-1}} \\ & B_2 & & & \\ & & \ddots & & \\ & & & B_{q-1} & \\ & & & & B_q \end{pmatrix},$$

in which the $B_1, B_2, \cdots, B_q$ *are explained as in Lemma* 3.2. *Then*

$$\mathrm{sep}_F(A, B) \geqslant \frac{s(A_{(1)}, B_{(1)})}{\varkappa_A \cdot \varkappa_B}. \tag{3.34}$$

Here $s(A_{(1)}, B_{(1)})$ *is given by* (3.20), $\varkappa_A$ *and* $\varkappa_B$ *are computed according to the steps stated in Lemma* 3.4.

Proof. By Lemma 3.1, there exist nonsingular matrices Q_A and Q_B such that $Q_A^{-1}AQ_A = A_{(1)}$ and $Q_B^{-1}BQ_B = B_{(1)} = [B_1, \cdots, B_q]$, therefore (see (2.7))

$$\mathrm{sep}_F(A, B) \geqslant \frac{\mathrm{sep}_F(A_{(1)}, B_{(1)})}{\varkappa(Q_A)\varkappa(Q_B)}. \tag{3.35}$$

Substituting the lower bound $s(A_{(1)}, B_{(1)})$ of $\mathrm{sep}_F(A_{(1)}, B_{(1)})$ (see Lemma 3.2) and the upper bounds $\varkappa_A$ and $\varkappa_B$ of $\varkappa(Q_A)$ and $\varkappa(Q_B)$ (see Lemma 3.4) into the right-hand of (3.35), we obtain the inequality (3.34) at once. ∎

If A and B are not upper triangular matrices, by Schur's theorem we can use unitary matrices to transform A and B to upper triangular forms respectively; but the $\{A_i\}$, $\{P_i\}$, $\{P_i'\}$, $\{B_j\}$, $\{Q_j\}$ and $\{Q_j'\}$ are unknown beforehand. In such a case we can use the following inequalities to compute the lower bound of $\mathrm{sep}_F(A_{(1)}, B_{(1)})$ and the upper bounds of $\varkappa(Q_A)$ and $\varkappa(Q_B)$:

$$\Delta_F(A_i),\ \Delta_F(A_{(i)}),\ \Delta_F(A^{(i)}),\ \sqrt{\sum_{i=1}^{p-1}\|P_i\|_F^2},\ \sqrt{\sum_{i=1}^{p-1}\|P_i'\|_F^2} \leqslant \Delta_F(A),$$

$$\Delta_F(A_{p-k}) \leqslant \Delta_F(A_{(p-k+1)}),\quad \Delta_F(A^{(k)}) + \Delta_F(A_{k+1}) \leqslant \sqrt{2}\,\Delta_F(A);$$

for the matrix B there are similar inequalities. Thus Lemma 3.4 can be rewrite as follows.

Lemma 3.5. *Suppose that* $A \in \mathbb{C}^{m\times m}$ *with different eigenvalues* $\lambda_1, \lambda_2, \cdots, \lambda_p$ *of multiplicities* $m_1, m_2, \cdots, m_p$ *respectively. Let* $\delta_{(p-k)}(A)$, $\delta^{(k+1)}(A)$, $m_{(p-k+1)}$ *and* $m^{(k)}$ *be denoted by* (3.26)—(3.28). *Then there is a nonsingular matrix* Q_A *such that the equality* (3.2) *is valid, and*

$$\varkappa(Q_A) \leqslant \varkappa_A', \tag{3.36}$$

where $\varkappa_A'$ *can be computed by the following steps* (1)—(3):

(1) *For* $k = 1, 2, \cdots, p-1$, *let*

$$\Delta_{(p-k)}'(A) = \sqrt{2}\,\Delta_F(A)/\delta_{(p-k)}(A),\quad \Delta^{(k+1)'}(A) = \sqrt{2}\,\Delta_F(A)/\delta^{(k+1)}(A), \tag{3.37}$$

compute

$$s_{(p-k)}'(A) = \frac{[\Delta_{p-k}'(A) - 1]\,\delta_{(p-k)}(A)}{[\Delta_{p-k}'(A)]^{m_{p-k}+m_{(p-k+1)}-1} - 1},\quad s^{(k+1)'}(A) = \frac{[\Delta^{(k+1)'}(A) - 1]\,\delta^{(k+1)}(A)}{[\Delta^{(k+1)'}(A)]^{m^{(k)}+m_{k+1}-1} - 1}; \tag{3.38}$$

(2) *Let*

$$x_{p-1}' = \Delta_F(A)/s_{(p-1)}'(A),\quad y_1' = \Delta_F(A)/s^{(2)'}(A),$$

and then compute successively

$$x_{p-k}' = \frac{\left(1+\sqrt{\sum_{i=1}^{k-1} x_{p-i}'^2}\right)\Delta_F(A)}{s_{(p-k)}'(A)},\quad y_k' = \frac{\left(1+\sqrt{\sum_{i=1}^{k-1} y_i'^2}\right)\Delta_F(A)}{s^{(k+1)'}(A)},$$

$k = 2, 3, \cdots, p-2$;

(3) *Let*

$$x = \max\left\{\frac{1}{s_{(p-1)}'(A)},\ \max_{2\leqslant k\leqslant p-1}\left\{\frac{1+\sqrt{\sum_{i=1}^{k-1} x_{p-i}'^2}}{s_{(p-k)}'(A)}\right\}\right\} \tag{3.39}$$

and

$$y=\max\left\{\frac{1}{s^{(1)'}(A)},\quad \max_{2<k<p-1}\left\{\frac{1+\sqrt{\sum_{i=1}^{k-1} y_i'^2}}{s^{(k+1)'}(A)}\right\}\right\}, \tag{3.40}$$

compute

$$\varkappa_A'=(1+x\cdot\Delta_F(A))(1+y\cdot\Delta_F(A)). \tag{3.41}$$

Consequently, we obtain

Theorem 3.2. *Let $A\in\mathbb{C}^{m\times m}$ with different eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_p$ of multiplicities $m_1, m_2, \cdots, m_p$, and $B\in\mathbb{C}^{n\times n}$ with different eigenvalues $\mu_1, \mu_2, \cdots, \mu_q$ of multiplicities $n_1, n_2, \cdots, n_q$, respectively. Let*

$$\delta(A, B)=\min_{i,j}|\lambda_i-\mu_j|,\quad s_A=\max_i\{m_i\},\quad s_B=\max_j\{n_j\}.$$

Then

$$\mathrm{sep}_F(A, B)\geqslant\frac{s(A, B)}{\varkappa_A'\cdot\varkappa_B'}, \tag{3.42}$$

where

$$s(A, B)=\frac{[\Delta(A, B)-1]\delta(A, B)}{[\Delta(A, B)]^{s_A+s_B-1}-1},\quad \Delta(A, B)=\frac{\Delta_F(A)+\Delta_F(B)}{\delta(A, B)};$$

$\varkappa_A'$ and $\varkappa_B'$ are computed according to the steps stated in Lemma 3.5.

Finally, we give an example to explain that how shall we apply Lemma 3.4 and Lemma 3.5 as well as a formula in [2] to estimate the condition number of the transformation matrix Q_A which transform a matrix A to a block diagonal form, and compare the corresponding results.

Example 3.1.

$$A=\begin{pmatrix}A_1 & P_1\\ 0 & A_2\end{pmatrix},\quad A_1=\begin{pmatrix}1 & -1 & & \\ & \ddots & \ddots & \\ & & \ddots & -1\\ & & & 1\end{pmatrix}_{m_1\times m_1},\quad A_2=\begin{pmatrix}1-\alpha & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1\\ & & & 1-\alpha\end{pmatrix}_{m_2\times m_2},\quad \alpha>0,$$

$$P_1=(\delta_{ij}),\quad \delta_{ij}=\begin{cases}1 & \text{if } i=j,\\ 0 & \text{if } i\neq j.\end{cases}$$

We choose the following transformation matrix Q_A:

$$Q_A=\begin{pmatrix}I^{(m_1)} & -X_1\\ 0 & I^{(m_2)}\end{pmatrix},\quad Q_A^{-1}AQ_A=\begin{pmatrix}A_1 & 0\\ 0 & A_2\end{pmatrix}.$$

Apply Lemma 3.4. We have

$$\delta_{(1)}(A)=\delta^{(2)}(A)=\alpha,\quad \Delta_{(1)}(A)=\Delta^{(2)}(A)=[\Delta_F(A_1)+\Delta_F(A_2)]/\alpha$$

(see (3.26) and (3.27)), and

$$s_{(1)}(A)=s^{(2)}(A)=\frac{[\Delta_{(1)}(A)-1]\alpha}{[\Delta_{(1)}(A)]^{m_1+m_2-1}-1}$$

(see (3.24) and (3.25)). Substituting (3.29) and (3.33), we get

$$\varkappa(Q_A)\leqslant\left(1+\frac{\|P_1\|_F}{s_{(1)}(A)}\right)^2\equiv s_1. \tag{3.43}$$

Apply Lemma 3.5. We have

$$s_{(1)}'(A)=s^{(2)'}(A)=\frac{[\Delta_{(1)}'(A)-1]\alpha}{[\Delta_{(1)}'(A)]^{m_1+m_2-1}-1},\quad \Delta_{(1)}'(A)=\frac{\sqrt{2}\,\Delta_F(A)}{\alpha}$$

(see (3.38)), and

$$x = y = 1/s'_{(1)}(A)$$

(see (3.39) and (3.40)). Substituting (3.36) and (3.41), we get

$$\varkappa(Q_A) \leqslant \left(1 + \frac{\Delta_F(A)}{s'_{(1)}(A)}\right)^2 \equiv s_2. \tag{3.44}$$

Apply Theorem 1 of [2]. Let

$$g_1 = \frac{[\alpha + \Delta_F(A)]^{m_2-1}}{\alpha^{m_2}}, \qquad g = \frac{[g_1\Delta_F(A)]^{m_1} - 1}{g_1\Delta_F(A) - 1} \cdot g_1,$$

we have

$$\varkappa(Q_A) \leqslant (1 + g\Delta_F(A))^2 \equiv J. \tag{3.45}$$

Let $\chi = \dfrac{\Delta_F(A)}{\alpha}$, we have

$$s_2 = (1 + \chi[1 + \sqrt{2}\chi + \cdots + (\sqrt{2}\chi)^{m_1+m_2-2}])^2$$

and

$$J = (1 + \chi(1+\chi)^{m_2-1}[1 + \chi(1+\chi)^{m_2-1} + \cdots + (\chi(1+\chi)^{m_2-1})^{m_1-1}])^2.$$

Some computed results are given in Table 2.

Table 2

m_1	m_2	α	s_1	s_2	J
4	4	$\frac{1}{2}$	$2.7_{10}11$	$1.8_{10}13$	$1.5_{10}27$
6	3	$\frac{1}{4}$	$1.1_{10}18$	$6.2_{10}19$	$2.9_{10}40$
6	4	$\frac{1}{8}$	$2.9_{10}26$	$2.5_{10}28$	$6.4_{10}69$
6	6	$\frac{1}{16}$	$2.0_{10}40$	$5.7_{10}42$	$2.8_{10}129$

The author wishes to thank Professor L. Elsner for his support.

References

[1] C. Davis, W. Kahan, The rotation of eigenvectors by a perturbation. III, *SIAM J. Numer. Anal.*, 7 (1970), 1—46.
[2] E. Jiang (蒋尔雄), On spectral variation of a nonnormal matrix, *Linear Algebra and Appl.*, 42 (1982), 223—241.
[3] P. Lancaster, Theory of Matrices, Academic Press, New York, 1969.
[4] M. Marcus, H. Minc, A Survey of Matrix Theory and Matrix Inequalities, Allyn and Bacon, Boston, 1964.
[5] G. W. Stewart, Error and perturbation bounds for subspaces associated with certain eigenvalue problems, *SIAM Rev.*, 15 (1973), 727—764.
[6] M. Varah, On the separation of two matrices, *SIAM J. Numer. Anal.*, 16 (1979), 216—222.