Journal of Computational Mathematics, Vol.9, No.3, 1991, 247–255.

RAYLEIGH QUOTIENT AND RESIDUAL OF A DEFINITE PAIR*1)

Sun Ji-guang
(*Computing Center, Academia Sinica, Beijing, China*)

Abstract

Let $\{A, B\}$ be a definite matrix pair of order n, and let Z be an l-dimensional subspace of $\mathbb{C}^n$. In this paper we introduce the Rayleigh quotient matrix pair $\{H_1, K_1\}$ and residual matrix pair $\{R_A, R_B\}$ of $\{A, B\}$ with respect to Z, and use the norm of $\{R_A, R_B\}$ to bound the difference between the eigenvalues of $\{H_1, K_1\}$ and that of $\{A, B\}$, and to bound the difference between Z and an l-dimensional eigenspace of $\{A, B\}$. The corresponding classical theorems on the Hermitian matrices can be derived from the results of this paper.

§1. Preliminaries

Notation. $\mathbb{C}^{m\times n}$: the set complex $m \times n$ matrices, and $\mathbb{C}^n = \mathbb{C}^{n\times 1}$; $\mathbb{C}_r^{m\times n}$: the set of matrices with rank r in $\mathbb{C}^{m\times n}$. $\mathbb{R}$: the set of real numbers. $\bar{A}$: the conjugate of A. A^T: the transpose of A. $A^H = \bar{A}^T$. $A^\dagger$: the Moore-Penrose inverse of a matrix A. $\lambda(A)$: the set of the eigenvalues of A. $\mathcal{R}(X_1)$: the column space of a matrix X_1. $\|\cdot\|_2$: the Euclidean norm for vectors and the spectral norm for matrices. $\|\cdot\|_F$: the Frobenius matrix norm.

In this section we give some definitions and basic results on definite pairs.

Let $A, B \in \mathbb{C}^{n\times n}$ be Hermitian. The matrix pair $\{A, B\}$ is a definite pair if [3]–[6]

$$c(A, B) \equiv \min_{\|x\|_2=1} |x^H(A + iB)x| > 0, \tag{1.1}$$

where $i = \sqrt{-1}$. The set of all definite pairs of order n will be denoted by $\mathbb{D}(n)$.

Let $\{A, B\} \in \mathbb{D}(n)$. A non-zero vector $x \in \mathbb{C}^n$ is an eigenvector of $\{A, B\}$ belonging to the eigenvalue (α, β), if

$$(\alpha, \beta) \neq (0, 0), \quad \beta Ax = \alpha Bx.$$

The set of the eigenvalues of $\{A, B\}$ will be denoted by $\lambda(A, B)$.

Let $\{A, B\} \in \mathbb{D}(n)$, and let $\mathcal{X} = \mathcal{R}(X_1)$ be an l-dimensional subspace of $\mathbb{C}^n$, in which $X_1 \in \mathbb{C}_l^{n\times l}$. $\mathcal{X}$ is called an eigenspace of $\{A, B\}$ if[3]

$$\dim\,(A\mathcal{X} + B\mathcal{X}) \leq \dim\,(\mathcal{X}).$$

Let $\mathcal{X} = \mathcal{R}(X_1)$, $X_1 \in \mathbb{C}_l^{n\times l}$. It is easy to prove that the following statements are equivalent (ref. [3]–[5]):

1) $\mathcal{X}$ is an l-dimensional eigenspace of $\{A, B\}$;

2) $\mathcal{X}$ is spanned by l linearly independent eigenvectors of $\{A, B\}$;

* Received May 27, 1988.

1) The Project Supported by National Natural Science Foundation of China.

3) there is an l-dimensional subspace $\mathcal{Y} \subseteq \mathbb{C}^n$ such that

$$A\mathcal{X}, \quad B\mathcal{X} \subseteq \mathcal{Y};$$

4) there is $\{A', B'\} \in \mathbb{D}(l)$ such that

$$AX_1B' = BX_1A';$$

5) there are $Y_1 \in \mathbb{C}^{n\times l}$ and $\{A_1, B_1\} \in \mathbb{D}(l)$ such that $Y_1^H X_1 = I$, and

$$AX_1 = Y_1A_1, \quad BX_1 = Y_1B_1. \tag{1.2}$$

Now we introduce the Rayleigh quotient and the residual of a definite pair.

Definition 1.1. *Let $\{A, B\} \in \mathbb{D}(n), Z_1 \in \mathbb{C}^{n\times l}$, and $Z_1^H Z_1 = I$. Let*

$$H_1 = Z_1^H AZ_1, \quad K_1 = Z_1^H BZ_1. \tag{1.3}$$

Then $\{H_1, K_1\}$ is called the Rayleigh quotient matrix pair (or simply, the Rayleigh quotient) of $\{A, B\}$ with respect to Z_1.

Suppose that $\mathcal{R}(X_1)$ is an l-dimensional eigenspace of $\{A, B\} \in \mathbb{D}(n)$. From 5), there are $Y_1 \in \mathbb{C}^{n\times l}$ and $\{A, B\} \in \mathbb{D}(l)$ such that $Y_1^H X_1 = I$, and the relations (1.2) hold. From (1.2) we get

$$A_1 = X_1^H AX_1, \quad B_1 = X_1^H BX_1$$

and

$$Y_1 = (AX_1A_1 + BX_1B_1)(A_1^2 + B_1^2)^{-1}.$$

This suggests the following definition.

Definition 1.2. *Let $\{A, B\} \in \mathbb{D}(n)$. For any fixed $Z_1 \in \mathbb{C}^{n\times l}$ satisfying $Z_1^H Z_1 = I$, construct H_1 and K_1 by (1.3), and let*

$$W_1 = (AZ_1H_1 + BZ_1K_1)(H_1^2 + K_1^2)^{-1} \tag{1.4}$$

and

$$R_A(Z_1) = AZ_1 - W_1H_1, \quad R_B(Z_1) = BZ_1 - W_1K_1. \tag{1.5}$$

Then $\{R_A(Z_1), R_B(Z_1)\}$ is called the residual matrix pair (or simply, the residual) of $\{A, B\}$ with respect to Z_1.

It is easy to see that if $Z_1 \in \mathbb{C}^{n\times l}$ and $Z_1^H Z_1 = I$, then $\mathcal{R}(Z_1)$ is an eigenspace of $\{A, B\} \in \mathbb{D}(n)$ if and only if

$$R_A(Z_1) = 0, \quad R_B(Z_1) = 0.$$

We have proved in [6] that the precision of the eigenvalues of $\{H_1, K_1\}$ as l approximate eigenvalues of $\{A, B\}$ is higher than that of $\mathcal{R}(Z_1)$ as its approximate eigenspace. In this paper we shall use the norm of $\{R_A(Z_1), R_B(Z_1)\}$ to bound the difference between the eigenvalues of $\{H_1, K_1\}$ and that of $\{A, B\}$ to bound the difference between $\mathcal{R}(Z_1)$ and an l-dimensional eigenspace of $\{A, B\}$. From the results of this paper one can derive the corresponding classical theorems on Hermitian matrices.

At the end of this section we cite two perturbation theorems which will be used in the following.

Theorem 1.3[3]. *Let $\{A, B\}, \{\tilde{A}, \tilde{B}\} \in \mathbb{D}(n)$, and let $\lambda(A, B) = \{(\alpha_i, \beta_i)\}$, $\lambda(\tilde{A}, \tilde{B}) = \{(\tilde{\alpha}_i, \tilde{\beta}_i)\}$. Then there is a permutation π of $\{1, \cdots, n\}$ such that*

$$\rho\big((\alpha_i, \beta_i), (\tilde{\alpha}_{\pi(i)}, \tilde{\beta}_{\pi(i)})\big) \le \frac{\sqrt{\|\tilde{A} - A\|_2^2 + \|\tilde{B} - B\|_2^2}}{c(A, B)}, \quad i = 1, \cdots, n.$$

Here

$$\rho((\alpha,\beta),(\gamma,\delta)) \equiv \frac{|\alpha\delta-\beta\gamma|}{\sqrt{(|\alpha|^2+|\beta|^2)(|\gamma|^2+|\delta|^2)}}$$

is the chordal distance between (α,β) *and* (γ,δ).

Theorem 1.4[4]. *Let* $\{A,B\},\{\tilde{A},\tilde{B}\}\in \mathbb{D}(n)$, *and let*

$$X^HAX=\begin{pmatrix} A_1 & 0\\ 0 & A_2\end{pmatrix},\quad X^HBX=\begin{pmatrix} B_1 & 0\\ 0 & B_2\end{pmatrix},$$

$$\tilde{X}^H\tilde{A}\tilde{X}=\begin{pmatrix} \tilde{A}_1 & 0\\ 0 & \tilde{A}_2\end{pmatrix},\quad \tilde{X}^H\tilde{B}\tilde{X}=\begin{pmatrix} \tilde{B}_1 & 0\\ 0 & \tilde{B}_2\end{pmatrix},$$

where $X=(X_1,X_2),\tilde{X}=(\tilde{X}_1,\tilde{X}_2)\in\mathbb{C}^{n\times n}$, *and* $X_1,\tilde{X}_1\in\mathbb{C}^{n\times l}$. *Let*

$$\delta\equiv\min_{i,j}\left\{\rho((\alpha_i,\beta_i),(\tilde{\alpha}_j,\tilde{\beta}_j)) : \begin{array}{l}(\alpha_i,\beta_i)\in\lambda(A_1,B_1)\\ (\tilde{\alpha}_j,\tilde{\beta}_j)\in\lambda(\tilde{A}_2,\tilde{B}_2)\end{array}\right\}.$$

If $\delta>0$, *then*

$$\|\sin\Theta(X_1,\tilde{X}_1)\|_F\le\frac{\|(A,B)\|_2}{c(A,B)c(\tilde{A},\tilde{B})}\cdot\frac{\|((\tilde{A}-A)X_1,(\tilde{B}-B)X_1\|_F}{\delta}.$$

Here

$$\Theta(X_1,\tilde{X}_1)\equiv\arccos(Z_1^H\tilde{Z}_1\tilde{Z}_1^HZ_1)^{\frac{1}{2}}\ge 0 \text{ (positive semi-definite)},$$

in which

$$Z_1=X_1(X_1^HX_1)^{-\frac{1}{2}},\quad \tilde{Z}_1=\tilde{X}_1(\tilde{X}_1^H\tilde{X}_1)^{-\frac{1}{2}}.$$

§2. An Extremum Property

This section will show that the matrix W_1 defined by (1.4) has an extremum property.

Theorem 2.1. *Let* $\{A,B\}\in\mathbb{D}(n)$, *and let* $Z_1\in\mathbb{C}^{n\times l}$. *Construct* H_1,K_1 *and* W_1 *by (1.3) and (1.4), respectively. Then* $F=W_1$ *is the unique solution of the problem*

$$\begin{cases}\left\|\begin{pmatrix} AZ_1-FH_1\\ BZ_1-FK_1\end{pmatrix}\right\|_F=\min,\\ \|F\|_F=\min,\quad F\in\mathbb{C}^{n\times l}.\end{cases}\tag{2.1}$$

Proof. First rewrite (2.1) as

$$\begin{cases}\|H_1^TF^T-P^T\|_F^2+\|K_1^TF^T-Q^T\|_F^2=\min,\\ \|F^T\|_F=\min,\end{cases}\tag{2.2}$$

where $P=AZ_1,Q=BZ_1$.

We associate the matrices F^T,P^T and Q^T with the nl-dimensional vectors f,p and q which are the direct sums of the column vectors of F^T,P^T and Q^T, respectively. Then the problem (2.2) takes the form

$$\begin{cases}\left\|\begin{pmatrix} I^{(n)}\otimes H_1^T\\ I^{(n)}\otimes K_1^T\end{pmatrix}f-\begin{pmatrix} p\\ q\end{pmatrix}\right\|_2=\min,\\ \|f\|_2=\min.\end{cases}\tag{2.3}$$

This is a linear least squares problem, which has a unique solution

$$f = \begin{pmatrix} I \otimes H_1^T \\ I \otimes K_1^T \end{pmatrix}^{\dagger} \begin{pmatrix} p \\ q \end{pmatrix} = \left[\begin{pmatrix} I \otimes H_1^T \\ I \otimes K_1^T \end{pmatrix}^H \begin{pmatrix} I \otimes H_1^T \\ I \otimes K_1^T \end{pmatrix} \right]^{-1} \begin{pmatrix} I \otimes H_1^T \\ I \otimes K_1^T \end{pmatrix}^H \begin{pmatrix} p \\ q \end{pmatrix}$$
$$= (I \otimes \bar{H}_1 H_1^T + I \otimes \bar{K}_1 K_1^T)^{-1} [(I \otimes \bar{H}_1)p + (I \otimes \bar{K}_1)q]$$
$$= [I \otimes (\bar{H}_1 H_1^T + \bar{K}_1 K_1^T)]^{-1} [(I \otimes \bar{H}_1)p + (I \otimes \bar{K}_1)q]$$
$$= [I \otimes (\bar{H}_1 H_1^T + \bar{K}_1 K_1^T)^{-1} \bar{H}_1]p + [I \otimes (\bar{H}_1 H_1^T + \bar{K}_1 K_1^T)^{-1} \bar{K}_1]q,$$

where $\otimes$ denotes the Kronecker product. Consequently,

$$F^T = (\bar{H}_1 H_1^T + \bar{K}_1 K_1^T)^{-1} (\bar{H}_1 P^T + \bar{K}_1 Q^T),$$

i.e., $F = W_1$. □

§3. The Eigenvalues of the Rayleigh Quotient

Let $\{A, B\} \in \mathbb{D}(n)$, $Z_1 \in \mathbb{C}^{n \times l}$, and $Z_1^H Z_1 = I$. If $\mathcal{R}(Z_1)$ is an eigenspace of $\{A, B\}$, then $\lambda(H_1, K_1) \subseteq \lambda(A, B)$ (see [3]). Now we assume that $\mathcal{R}(Z_1)$ is not any eigenspace of $\{A, B\}$.

Decompose $Z_1 = U \begin{pmatrix} I^{(l)} \\ 0 \end{pmatrix}$, where $U \in \mathbb{C}^{n \times n}$ is unitary. Construct W_1 by (1.4). From $W_1^H Z_1 = I$ we get $W_1 = U \begin{pmatrix} I^{(l)} \\ W_{21} \end{pmatrix}$. Let

$$Z = (Z_1, Z_2) = U \begin{pmatrix} I^{(l)} & -W_{21}^H \\ 0 & I \end{pmatrix}, \quad W = (W_1, W_2) = U \begin{pmatrix} I^{(l)} & 0 \\ W_{21} & I \end{pmatrix}. \tag{3.1}$$

We have $W^H Z = I$.

Write

$$Z^H A Z = \begin{pmatrix} H_1 & C^H \\ C & H_2 \end{pmatrix} = \begin{pmatrix} H_1 & 0 \\ 0 & H_2 \end{pmatrix} + \begin{pmatrix} 0 & C^H \\ C & 0 \end{pmatrix},$$
$$Z^H B Z = \begin{pmatrix} K_1 & D^H \\ D & K_2 \end{pmatrix} = \begin{pmatrix} K_1 & 0 \\ 0 & K_2 \end{pmatrix} + \begin{pmatrix} 0 & D^H \\ D & 0 \end{pmatrix}.$$

Clearly, $\{Z^H A Z, Z^H B Z\}, \left\{ \begin{pmatrix} H_1 & 0 \\ 0 & H_2 \end{pmatrix}, \begin{pmatrix} K_1 & 0 \\ 0 & K_2 \end{pmatrix} \right\} \in \mathbb{D}(n)$, and $\lambda(Z^H A Z, Z^H B Z) = \lambda(A, B)$. Let

$$\lambda(A, B) = \{(\alpha_i, \beta_i)\}, \quad \lambda(H_1, K_1) = \{(\gamma_i, \delta_i)\}.$$

Then by Theorem 1.3, there are $1', \cdots, l' \in \{1, \cdots, n\}$ such that

$$\rho((\alpha_{i'}, \beta_{i'}), (\gamma_i, \delta_i)) \le \frac{\sqrt{\left\| \begin{pmatrix} 0 & C^H \\ C & 0 \end{pmatrix} \right\|_2^2 + \left\| \begin{pmatrix} 0 & D^H \\ D & 0 \end{pmatrix} \right\|_2^2}}{c(Z^H A Z, Z^H B Z)}$$
$$= \frac{\sqrt{\|C\|_2^2 + \|D\|_2^2}}{c(Z^H A Z, Z^H B Z)}, \quad i = 1, \cdots, l. \tag{3.2}$$

Observe that for the residual $\{R_A(Z_1), R_B(Z_1)\}$ defined by (1.5), we have

$$Z^H R_A(Z_1) = \begin{pmatrix} H_1 \\ C \end{pmatrix} - \begin{pmatrix} I \\ 0 \end{pmatrix} H_1 = \begin{pmatrix} 0 \\ C \end{pmatrix}, \quad Z^H R_B(Z_1) = \begin{pmatrix} K_1 \\ D \end{pmatrix} - \begin{pmatrix} I \\ 0 \end{pmatrix} K_1 = \begin{pmatrix} 0 \\ D \end{pmatrix}.$$

Hence inequalities (3.2) can be rewritten as

$$\rho((\alpha_{i'}, \beta_{i'}), (\gamma_i, \delta_i)) \le \frac{\sqrt{\|Z^H R_A(Z_1)\|_2^2 + \|Z^H R_B(Z_1)\|_2^2}}{c(Z^H AZ, Z^H BZ)}, \quad i = 1, \cdots, l. \tag{3.3}$$

Let $\sigma = \|W_{21}\|_2$, and

$$\omega = \|W_1\|_2 = \|(AZ_1 H_1 + BZ_1 K_1)(H_1^2 + K_1^2)^{-1}\|_2. \tag{3.4}$$

Then $\omega = \sqrt{1 + \sigma^2}$, and

$$\|W\|_2 = \left\| \begin{pmatrix} I & 0 \\ W_{21} & I \end{pmatrix} \right\|_2 = \frac{\sqrt{\sigma^2 + 4} + \sigma}{2} = \frac{\sqrt{\omega^2 + 3} + \sqrt{\omega^2 - 1}}{2}, \tag{3.5}$$

$$\|Z\|_2 = \left\| \begin{pmatrix} I & -W_{21}^H \\ 0 & I \end{pmatrix} \right\|_2 = \|W\|_2. \tag{3.6}$$

Moreover, observe that

$$\begin{aligned} c(Z^H AZ, Z^H BZ) &= \min_{\|x\|_2 = 1} |x^H Z^H (A + iB) Z x| \\ &= \min_{\|x\|_2 = 1} \|Zx\|_2^2 \left| \frac{(Zx)^H}{\|Zx\|_2} (A + iB) \frac{Zx}{\|Zx\|_2} \right| \ge \|Z^{-1}\|_2^{-2} c(A, B) = \|W\|_2^{-2} c(A, B). \end{aligned} \tag{3.7}$$

Substituting (3.5)–(3.7) into (3.3), we get

$$\rho((\alpha_{i'}, \beta_{i'}), (\gamma_i, \delta_i)) \le \frac{\left(\frac{\sqrt{\omega^2 + 3} - \sqrt{\omega^2 - 1}}{2}\right)^3 \sqrt{\|R_A(Z_1)\|_2^2 + \|R_B(Z_1)\|_2^2}}{c(A, B)},$$
$$i = 1, \cdots, l. \tag{3.8}$$

Hence, we have proved the following result.

Theorem 3.1. *Let $\{A, B\} \in I\!D(n)$, $\lambda(A, B) = \{(\alpha_i, \beta_i)\}$, and let $Z_1 \in \mathbb{C}^{n\times l}$, $Z_1^H Z_1 = I$. Let $\{H_1, K_1\}$ be the Rayleigh qoutient of $\{A, B\}$ with respect to Z_1, and $\lambda(H_1, K_1) = \{(\gamma_i, \delta_i)\}$. Then there are $1', \cdots, l' \in \{1, \cdots, n\}$ such that inequalities (3.8) hold, where ω is defined by* (3.4).

§4. An Approximate Eigenspace

We first prove a lemma.

Lemma 4.1. *Assume $\{H_1, K_1\} \in I\!D(l)$. Let*

$$H = \begin{pmatrix} H_1 & 0 \\ 0 & 0^{(n-l)} \end{pmatrix}, \quad K = \begin{pmatrix} K_1 & 0 \\ 0 & c(H_1, K_1) I^{(n-l)} \end{pmatrix}.$$

Then

$$\|(H, K)\|_2 = \|(H_1, K_1)\|_2, \quad c(H, K) = c(H_1, K_1).$$

Proof. Let

$$H_{\varphi,1} = H_1 \cos\varphi - K_1 \sin\varphi, \quad K_{\varphi,1} = H_1 \sin\varphi + K_1 \cos\varphi.$$

By Stewart [3, Theorem 2.2], there is $\varphi \in [0, 2\pi)$ such that $K_{\varphi,1}$ is positive definite and

$$c(H_1, K_1) = \lambda\min(K_{\varphi,1}),$$

where $\lambda_{\min}(K_{\varphi,1})$ denotes the smallest eigenvalue of $K_{\varphi,1}$. Let

$$H_\varphi = \begin{pmatrix} H_{\varphi,1} & 0 \\ 0 & 0^{(n-l)} \end{pmatrix}, \quad K_\varphi = \begin{pmatrix} K_{\varphi,1} & 0 \\ 0 & c(H_1, K_1)I^{(n-l)} \end{pmatrix}.$$

Since

$$\|(H, K)\|_2 = \|(H_\varphi, K_\varphi)\|_2 = \max\{\|(H_{\varphi,1}, K_{\varphi,1})\|_2, c(H_1, K_1)\}$$

and

$$\|(H_{\varphi,1}, K_{\varphi,1})\|_2 \geq \|K_{\varphi,1}\|_2 \geq c(H_1, K_1),$$

we have

$$\|(H, K)\|_2 = \|(H_{\varphi,1}, K_{\varphi,1})\|_2 = \|(H_1, K_1)\|_2.$$

Moreover, since

$$c(H, K) = c(H_\varphi, K_\varphi) \geq \min_{\|x\|_2=1} |x^H K_\varphi x| = \lambda_{\min}(K_\varphi) = c(H_1, K_1),$$

and for $x = (0, x_2^T)^T \in \mathbb{C}^n$, where $x_2 \in \mathbb{C}^{n-l}$ with $\|x_2\|_2 = 1$,

$$\min_{\|x\|_2=1} |x^H(H_\varphi + iK_\varphi)x| = c(H_1, K_1),$$

we have $c(H, K) = c(H_1, K_1)$. □

Let $\{A, B\} \in \mathbb{D}(n)$, $Z_1 \in \mathbb{C}^{n\times l}$ with $Z_1^H Z_1 = I$. Let H_1, K_1 be defined by (1.3). Theorem 3.1 shows that there are $(\alpha_{1'}, \beta_{1'}), \cdots, (\alpha_{l'}, \beta_{l'}) \in \lambda(A, B)$, which are near the eigenvalues $(\gamma_1, \delta_1), \cdots, (\gamma_l, \delta_l) \in \lambda(H_1, K_1)$. Let $\mathcal{R}(X_1)$ be the eigenspace of $\{A, B\}$ corresponding to the eigenvalues $(\alpha_{1'}, \beta_{1'}), \cdots, (\alpha_{l'}, \beta_{l'})$, where $X_1 \in \mathbb{C}_l^{n\times l}$. This section will use the norm of $\{R_A(Z_1), R_B(Z_1)\}$ to bound $\|\sin\Theta(X_1, Z_1)\|_F$.

By the hypotheses, there is $X = (X_1, X_2) \in \mathbb{C}_n^{n\times n}$ such that

$$X^H AX = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \quad X^H BX = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}, \tag{4.1}$$

in which $\{A_1, B_1\} \in \mathbb{D}(l)$, and $\lambda(A_1, B_1) = \{(\alpha_{i'}, \beta_{i'})\}_{i=1}^l$. Let

$$I = \{1, \cdots, n\} \setminus \{1', \cdots, l'\}, \quad \delta \equiv \min_{\substack{i\in I \\ 1\leq j\leq l}} \rho((\alpha_i, \beta_i), (\gamma_j, \delta_j)), \tag{4.2}$$

and assume $\delta > 0$.

Construct W_1 by (1.4), and then define Z, W by (3.1). Consider the definite pair $\{\tilde{A}, \tilde{B}\}$ of order n defined by

$$\tilde{A} = W \begin{pmatrix} H_1 & 0 \\ 0 & 0 \end{pmatrix} W^H, \quad \tilde{B} = W \begin{pmatrix} K_1 & 0 \\ 0 & c(H_1, K_1)I \end{pmatrix} W^H. \tag{4.3}$$

Let

$$E = A - \tilde{A}, \quad F = B - \tilde{B}.$$

It is easy to see that

$$Z^H \tilde{A} Z = \begin{pmatrix} H_1 & 0 \\ 0 & 0 \end{pmatrix}, \quad Z^H B Z = \begin{pmatrix} K_1 & 0 \\ 0 & c(H_1, K_1) I \end{pmatrix}, \tag{4.4}$$

and

$$\min \left\{ \rho((\alpha_i, \beta_i), (\gamma_j, \delta_j)) : \begin{array}{l} (\alpha_i, \beta_i) \in \lambda(A_2, B_2) \\ (\gamma_j, \delta_j) \in \lambda(H_1, K_1) \end{array} \right\} = \delta > 0.$$

By Theorem 1.4, we have

$$\|\sin \Theta(Z_1, X_1)\|_F \le \frac{\|(\tilde{A}, \tilde{B})\|_2 \|(EZ_1, FZ_1)\|_F}{c(\tilde{A}, \tilde{B}) c(A, B) \delta}. \tag{4.5}$$

Observe that

$$\begin{aligned}
EZ_1 &= (A - \tilde{A}) Z_1 = AZ_1 - W_1 H_1 = R_A(Z_1), \\
FZ_1 &= (B - \tilde{B}) Z_1 = BZ_1 - W_1 K_1 = R_B(Z_1); \\
\|(\tilde{A}, \tilde{B})\|_2 &\le \|W\|_2^2 \left\| \left(\begin{pmatrix} H_1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} K_1 & 0 \\ 0 & c(H_1, K_1) I \end{pmatrix} \right) \right\|_2 \\
&\le \frac{\omega^2 + 1 + \sqrt{\omega^4 + 2\omega^2 - 3}}{2} \|(H_1, K_1)\|_2 \qquad \text{(by (3.5) and Lemma 4.1);} \\
c(\tilde{A}, \tilde{B}) &\ge \|W^{-1}\|_2^{-2} c \left(\begin{pmatrix} H_1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} K_1 & 0 \\ 0 & c(H_1, K_1) I \end{pmatrix} \right) \\
&= \|Z\|_2^{-2} c(H_1, K_1) \qquad \text{(by Lemma 4.1)} \\
&= \left(\frac{\omega^2 + 1 + \sqrt{\omega^4 + 2\omega^2 - 3}}{2} \right)^{-1} c(H_1, K_1) \qquad \text{(by (3.5) and (3.6)).}
\end{aligned}$$

Substituting these relations into (4.5), we get

$$\|\sin \Theta(Z_1, X_1)\|_F \le \frac{\left(\frac{\omega^2 + 1 + \sqrt{\omega^4 + 2\omega^2 - 3}}{2} \right)^2 \|(H_1, K_1)\|_2 \|(R_A(Z_1), R_B(Z_1))\|_F}{c(H_1, K_1) c(A, B) \delta}. \tag{4.6}$$

Hence, we have proved the following result.

Theorem 4.2. *Let $\{A, B\}, \lambda(A, B), Z_1, \{H_1, K_1\}, \lambda(H_1, K_1)$ and $(\alpha_{1'}, \beta_{1'}), \cdots, (\alpha_{l'}, \beta_{l'})$ be as in Theorem 3.1. Let $\mathcal{R}(X_1)$ be the eigenspace of $\{A, B\}$ corresponding to the eigenvalues $(\alpha_{1'}, \beta_{1'}), \cdots, (\alpha_{l'}, \beta_{l'})$, where $X_1 \in \mathbb{C}^{n \times l}$. Define δ by (4.2), and assume $\delta > 0$. Then inequality (4.6) holds, where ω is defined by (3.4).*

§5. Final Remarks

Let $A \in \mathbb{C}^{n \times n}$ be Hermitian, $X_1 \in \mathbb{C}^{n \times l}$ and $X_1^H X_1 = I$. Then $\mathcal{X} = \mathcal{R}(X_1)$ is called an eigenspace of A if

$$A\mathcal{X} \subseteq \mathcal{X}, \tag{5.1}$$

or equivalently, if there is a Hermitian matrix $A_1 \in \mathbb{C}^{l \times l}$ such that

$$AX_1 = X_1 A_1. \tag{5.2}$$

Clearly, $A_1 = X_1^H A X_1$.

Let $Z_1 \in \mathbb{C}^{n\times l}$, $Z_1^H Z_1 = I$. The matrix $H_1 = Z_1^H A Z_1$ is said to be the Rayleigh quotient of A with respect to Z_1, and

$$R(Z) = AZ_1 - Z_1 H_1 \tag{5.3}$$

is said to be the residual of A with respect to Z_1.

Now we use Theorem 3.1 and Theorem 4.2 to derive the corresponding results on Hermitian matrices (ref. [1], [2], [5]).

5.1. On $\lambda(H_1)$

Let $\lambda(A) = \{\alpha_i\}_{i=1}^n, \lambda(H_1) = \{\gamma_i\}_{i=1}^l$. Take $t > 0$ and $B = I$. Then, obviously $\{A, tI\} \in \mathbb{D}(n), K_1 = tI^{(l)}$, $\lambda(A, tI) = \{(\alpha_i, t)\}$ and $\lambda(H_1, tI) = \{(\gamma_i, t)\}$. Moreover, by (1.4) and (1.5) we have

$$W_1 = (AZ_1H_1 + t^2 Z_1)(H_1^2 + t^2 I)^{-1} = (Z_1 + \frac{1}{t^2}AZ_1H_1)(I + \frac{1}{t^2}H_1^2)^{-1} \equiv W_1(t)$$

and

$$R_A(Z_1) = AZ_1 - W_1(t)H_1, \quad R_{tI}(Z_1) = t(Z_1 - W_1(t)).$$

By Theorem 3.1, there are $1', \cdots, l' \in \{1, \cdots, n\}$ such that

$$\rho((\alpha_{i'}, t), (\gamma_i, t)) \le \frac{\left(\frac{\sqrt{\omega^2(t) + 3} - \sqrt{\omega^2(t) - 1}}{2}\right)^3 \sqrt{\|R_A(Z_1)\|_2^2 + \|R_B(Z_1)\|_2^2}}{c(A, tI)}, \quad i = 1, \cdots, l, \tag{5.4}$$

where

$$\omega(t) = \|W_1(t)\|_2 = \|(Z_1 + \frac{1}{t^2}AZ_1H_1)(I + \frac{1}{t^2}H_1^2)^{-1}\|_2,$$

$$c(A, tI) = t \min_{\|x\|_2=1} \left| x^H (I - \frac{i}{t}A)x \right|, \quad i = \sqrt{-1},$$

$$\sqrt{\|R_A(Z_1)\|_2^2 + \|R_{tI}(Z_1)\|_2^2} = \sqrt{\|AZ_1 - W_1(t)H_1\|_2^2 + t^2\|Z_1 - W_1(t)\|_2^2}$$

and

$$\rho((\alpha_{i'}, t), (\gamma_i, t)) = \frac{|\alpha_{i'} - \gamma_i|}{t\sqrt{(1 + (\frac{\alpha_{i'}}{t})^2)(1 + (\frac{\gamma_i}{t})^2)}}.$$

Observe that

$$\lim_{t\to\infty} W_1(t) = Z_1, \quad \lim_{t\to\infty} \omega(t) = 1$$

and

$$Z_1 - W_1(t) = \frac{1}{t^2}(Z_1H_1 - AZ_1)H_1(I + \frac{1}{t^2}H_1^2)^{-1}.$$

Hence, considering $t \to +\infty$, from (5.4) we get

$$|\alpha_{i'} - \gamma_i| \le \|R(Z_1)\|_2, \quad i = 1, \cdots, l. \tag{5.5}$$

5.2. On $\mathcal{R}(Z_1)$

Let $\mathcal{R}(X_1)$ be the eigenspace of A corresponding to $\alpha_{1'}, \cdots, \alpha_{l'}$, where $X_1 \in \mathbb{C}^{n\times l}$. Without loss of generality we may assume that $X_1^H X_1 = I$. Take $t > 0$, and consider the definite pair $\{A, tI\}$ and its Rayleigh quotient $\{H_1, tI\}$, where $H_1 = Z_1^H AZ_1$.

Obviously, there is a unitary matrix $X = (X_1, X_2) \in \mathbb{C}^{n\times n}$ such that

$$X^H AX = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \quad X^H(tI)X = \begin{pmatrix} tI^{(l)} & 0 \\ 0 & tI^{(n-l)} \end{pmatrix},$$

and $\lambda(A_1, tI) = \{(\alpha_{i'}, t)\}_{i=1}^{l}$. Let $I = \{1, \cdots, n\} \setminus \{1', \cdots, l'\}$ and

$$\delta \equiv \min_{\substack{i\in I \\ 1\le j\le l}} |\alpha_i - \delta_j| > 0.$$

Then

$$\delta_t \equiv \min_{\substack{i\in I \\ 1\le j\le l}} \rho((\alpha_i, t), (\gamma_j, t)) = \min_{\substack{i\in I \\ 1\le j\le l}} \frac{|\alpha_i - \gamma_j|}{t\sqrt{(1+(\frac{\alpha_i}{t})^2)(1+(\frac{\gamma_j}{t})^2)}} > 0.$$

By Theorem 4.2, we have

$$\|\sin\Theta(Z_1, X_1)\|_F$$

$$\le \frac{\left(\frac{\omega^2(t)+1+\sqrt{\omega^4(t)+2\omega^2(t)-3}}{2}\right)^2 \|(H_1, tI)\|_2 \|(R_A(Z_1), R_{tI}(Z_1))\|_F}{c(H_1, tI)c(A, tI)\delta_t}. \tag{5.6}$$

Considering $t \to +\infty$, we get

$$\|\sin\Theta(Z_1, X_1)\|_F \le \frac{\|R(Z_1)\|_F}{\delta}. \tag{5.7}$$

References

[1] W. Kahan, Inclusion theorems for clusters of eigenvalues of Hermitian matrices, Tech. Report No. CS42, Computer Science Dept., University of Toronto, 1967.

[2] B.N. Parlett, The Symmetric Eigenvalue Problem, Prentice-Hall, Eaglewood Cliffs, N.J.

[3] G.W. Stewart, Perturbation bounds for the definite generalized eigenvalue problem, *Linear Algebra and Appl.*, **23** (1979), 69–86.

[4] Sun Ji-guang, The perturbation bounds for eigenspaces of a definite matrix-pair, *Numer. Math.*, **41** (1983), 321–343.

[5] Sun Ji-guang, Matrix Perturbation Analysis, Science Press, Beijing, 1987 (in Chinese).

[6] Sun Ji-guang, Relative eigenvalues of a definite matrix-pair, *Linear Algebra and Appl.*, **139** (1990), 253–267.