ORTHOGONAL PROJECTIONS AND THE PERTURBATION OF THE EIGENVALUES OF SINGULAR PENCILS*

SUN JI-GUANG(孙继广)
(*Computing Center, Academia Sinica*)

Abstract

In this paper we obtain a Hoffman-Wielandt type theorem and a Bauer-Fike type theorem for singular pencils of matrices. These results delineate the relations between the perturbation of the eigenvalues of a singular diagonalizable pencil $A-\lambda B$ and the variation of the orthogonal projection onto the column space $\mathfrak{R}\begin{pmatrix}A^H\\B^H\end{pmatrix}$.

1. Introduction

Let A and B be complex $m\times n$ matrices. A pencil of matrices $A-\lambda B$ is called singular if $m\neq n$ or $m=n$ but $\det(A-\lambda B)\equiv 0$[4]. A prevalent viewpoint is that in this case any complex number λ is an eigenvalue of $A-\lambda B$ (ref. [6]), consequently it is difficult to investigate the perturbation of the eigenvalues of singular pencils. In this paper we adopt a new definition for the eigenvalues of a singular pencil which is due to P. van Dooren[2], and relate the perturbation of the eigenvalues of $A-\lambda B$ and the variation of the orthogonal projection onto the column space $\mathfrak{R}\begin{pmatrix}A^H\\B^H\end{pmatrix}$ to each other, thus obtain a Hoffman-Wielandt type theorem (§ 3) and a Bauer-Fike type theorem (§ 4) for singular pencils which are generalizations of the main results for regular pencils in [8] and [3].

Notation: Capital case is used for matrices and lower case Greek letters for scalars. The symbol $\mathbb{C}^{m\times n}$ denotes the set of complex $m\times n$ matrices. $\bar{A}$ and A^T stand for conjugate and transpose of A, respectively; $A^H=\bar{A}^T$. $I^{(n)}$ is the $n\times n$ identity matrix, and 0 is the null matrix. The matrix $|A|$ has elements $|a_{ij}|$ if $A=(a_{ij})$. $A>0$ $(\geqslant 0)$ denotes that H is a positive definite (semi-positive definite) Hermitian matrix. The colume space of A is denoted by $\mathfrak{R}(A)$ and the null space by $N(A)$. $\mathfrak{R}(A)^\perp$ is the orthogonal complement space of $\mathfrak{R}(A)$. $G_{1,2}$ denotes the complex projective plane. The chordal distance between the points (α,β) and (γ,δ) on $G_{1,2}$ is

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

* Received August 17, 1982.

1) This work was done while the author was visiting the University of Bielefeld (FRG) and assisted by the Alexander von Humboldt Foundation in the Federal Republic of Germany.

2) P. van Dooren has advanced a new definition for the eigenvalues of a singular pencil in his lecture "A numerical method to compute reducing subspaces of a singular pencil" at "The Conference on Matrix Pencils" in March 1982, Piteå, Sweden.

The matrix pencil formed of A and B can be expressed more precisely as $\mu A-\lambda B$, $(\lambda, \mu)\in G_{1,2}$.

2. Preliminaries

In this section we give some definitions and basic results.

2.1. Eigenvalues and eigenvectors

Definition 2.1. *Let $A, B\in \mathbb{C}^{m\times n}$, and $\max\limits_{(\lambda,\mu)\in G_{1,2}} \operatorname{rank}(\mu A-\lambda B)=k$. A number-pair $(\alpha, \beta)\in G_{1,2}$ is an eigenvalue of the pencil $\mu A-\lambda B$ if* $\operatorname{rank}(\beta A-\alpha B)<k$.

The set of all eigenvalues of $\mu A-\lambda B$ is denoted by $\lambda(A, B)$.

The following consequences of Definition 2.1 can easily be verified.

i) If $\mu A-\lambda B$ is a regular pencil (i. e. $m=n$ and $\det(\mu A-\lambda B)\neq 0$, $(\lambda, \mu)\in G_{1,2}$)[4], then Definition 2.1 is coincide with the usual definition[8].

ii) If $P\in\mathbb{C}^{m\times m}$ and $Q\in\mathbb{C}^{n\times n}$ are non-singular, then

$$\lambda(PAQ, PBQ)=\lambda(A, B). \tag{1.1}$$

Kronecker showed that if $\mu A-\lambda B$ is a singular pencil, then there exist non-singular matrices P and Q such that[4,10,2]

$$P(\mu A-\lambda B)Q=\begin{bmatrix} L_{\varepsilon_1}(\lambda,\mu) & & & & & \\ & \ddots & & & & \\ & & L_{\varepsilon_p}(\lambda,\mu) & & & \\ & & & L^T_{\eta_1}(\lambda,\mu) & & \\ & & & & \ddots & \\ & & & & L^T_{\eta_q}(\lambda,\mu) & \\ & & & & & \mu A_0-\lambda B_0\end{bmatrix},\quad (\lambda,\mu)\in G_{1,2}. \tag{1.2}$$

Where $L_\varepsilon(\lambda, \mu)\in\mathbb{C}^{\varepsilon\times(\varepsilon+1)}$ and $L^T_\eta(\lambda, \mu)\in\mathbb{C}^{(\eta+1)\times\eta}$ are elementary Kronecker blocks, e. g. $L_2(\lambda, \mu)=\begin{pmatrix}\mu & -\lambda & 0\\ 0 & \mu & -\lambda\end{pmatrix}$; $\mu A_0-\lambda B_0$ is a regular pencil.

From the Kronecker's canonical form (1.2) it follows that

iii) $$\lambda(A, B)=\lambda(A_0, B_0).$$

We say that a singular pencil $\mu A-\lambda B$ contains an r-order regular part if $\mu A_0-\lambda B_0\in\mathbb{C}^{r\times r}$ in the form (1.2). The symbol $\mathfrak{S}^{m\times n}_r$ is used to denote the set of all $m\times n$ singular pencils, each of which contains an r-order regular part.

Definition 2.2. *Let $\mu A-\lambda B\in\mathfrak{S}^{m\times n}_r$. A non-zero vector $x\in\mathbb{C}^n$ is called an eigenvector of the singular pencil $\mu A-\lambda B$ corresponding to the eigenvalue (α, β) if*

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

2.2. Singular diagonalizable pencils and singular normal pencils

Definition 2.3. *A pencil $\mu A-\lambda B\in\mathfrak{S}^{m\times n}_r$ is called diagonalizable if there exist r linearly independent eigenvectors $x_1, \cdots, x_r$ of $\mu A-\lambda B$ and a complement space $\mathfrak{R}(x_1, \cdots, x_r)^c$ of $\mathfrak{R}(x_1, \cdots, x_r)$ satisfying*

$$\mathfrak{R}(x_1, \cdots, x_r)^c\subseteq\mathscr{N}(A)\cap\mathscr{N}(B).$$

The set of all such pencils is denoted by $\mathscr{D}^{m\times n}_r$.

Definition 2.4. *A pencil $\mu A-\lambda B\in\mathfrak{S}^{m\times n}_r$ is called normal if there exist r ortho-*

normal eigenvectors $u_1, \cdots, u_r$ *of* $\mu A-\lambda B$ *and*

$$\mathfrak{R}(u_1, \cdots, u_r)^{\perp}\subseteq\mathcal{N}(A)\cap\mathcal{N}(B).$$

The set of all such pencils is denoted by $\mathcal{N}_r^{m\times n}$.

Theorem 2.1. *Suppose that* $\mu A-\lambda B\in\mathfrak{S}_r^{m\times n}$. *Then* $\mu A-\lambda B\in\mathscr{D}_r^{m\times n}$ *if and only if there exist non-singular* $S\in\mathbb{C}^{m\times m}$, $Q\in\mathbb{C}^{n\times n}$ *and diagonal matrices* $\Lambda_1=\mathrm{diag}(\alpha_1, \cdots, \alpha_r)$, $\Omega_1=\mathrm{diag}(\beta_1, \cdots, \beta_r)$ *satisfying* $|\Lambda_1|^2+|\Omega_1|^2>0$ *such that*

$$A=S\begin{pmatrix}\Lambda_1 & 0\\ 0 & 0\end{pmatrix}Q^H,\quad B=S\begin{pmatrix}\Omega_1 & 0\\ 0 & 0\end{pmatrix}Q^H, \tag{2.1}$$

i. e.,

$$A=S_1\Lambda_1Q_1^H,\quad B=S_1\Omega_1Q_1^H, \tag{2.2}$$

where $S=(\underset{r}{S_1}, \underset{m-r}{S_2})$, $Q=(\underset{r}{Q_1}, \underset{n-r}{Q_2})$.

Proof. Suppose that (2.1) holds. Writing $X=Q^{-H}=(X_1, X_2)$, $X_1=(x_1, \cdots, x_r)$, from (2.1) we know that

$$\beta_iAx_i=\alpha_iBx_i,\ (Ax_i, Bx_i)\neq(0, 0),\ 1\leqslant i\leqslant r \tag{2.3}$$

and $\mathfrak{R}(X_2)$ is a complement space of $\mathfrak{R}(X_1)$ satisfying $\mathfrak{R}(X_2)\subseteq\mathcal{N}(A)\cap\mathcal{N}(B)$. This shows that $\mu A-\lambda B\in\mathscr{D}_r^{m\times n}$.

Suppose now that $\mu A-\lambda B\in\mathscr{D}_r^{m\times n}$. By Definition 2.3 there exist r linearly independent eigenvectors x_i satisfying (2.3). Let

$$s_i=\begin{cases}Ax_i/\alpha_i \text{ if } \alpha_i\neq0,\\ Bx_i/\beta_i \text{ if } \alpha_i=0,\end{cases}$$

and let $S_1=(s_1, \cdots, s_r)$, $X_1=(x_1, \cdots, x_r)$, $\Lambda_1=\mathrm{diag}(\alpha_1, \cdots, \alpha_r)$ and $\Omega_1=\mathrm{diag}(\beta_1, \cdots, \beta_r)$, then we have

$$AX_1=S_1\Lambda_1,\ BX_1=S_1\Omega_1,\ |\Lambda_1|^2+|\Omega_1|^2>0.$$

By the hypothesis there exists $X_2\in\mathbb{C}^{n\times(n-r)}$ such that $X=(X_1, X_2)$ is non-singular and $AX_2=BX_2=0$. Hence, if we set $Q=X^{-H}=(Q_1, Q_2)$, then A and B have the decompositions (2.2). Therefore

$$\mu A-\lambda B=S_1(\mu\Lambda_1-\lambda\Omega_1)Q_1^H,\ (\lambda, \mu)\in G_{1,2}. \tag{2.4}$$

Observe that $\mu A-\lambda B\in\mathfrak{S}_r^{m\times n}$, $\mu\Lambda_1-\lambda\Omega_1$ is a regular pencil and $\mathrm{rank}(Q_1)=r$, so from (2.4), rank $(S_1)=r$. We take $S_2\in\mathbb{C}^{m\times(m-r)}$ such that $S=(S_1, S_2)$ is non-singular, then from (2.2) we obtain (2.1).

Similarly one can prove

Theorem 2.2. *Suppose that* $\mu A-\lambda B\in\mathfrak{S}_r^{m\times n}$. *Then* $\mu A-\lambda B\in\mathcal{N}_r^{m\times n}$ *if and only if there exist a non-singular matrix* S, *a unitary matrix* U *and diagonal matrices* $\Lambda_1=\mathrm{diag}(\alpha_1, \cdots, \alpha_r)$, $\Omega_1=\mathrm{diag}(\beta_1, \cdots, \beta_r)$ *satisfying* $|\Lambda_1|^2+|\Omega_1|^2>0$ *such that*

$$A=S\begin{pmatrix}\Lambda_1 & 0\\ 0 & 0\end{pmatrix}U^H,\quad B=S\begin{pmatrix}\Omega_1 & 0\\ 0 & 0\end{pmatrix}U^H. \tag{2.5}$$

2.3. Orthogonal projections and metrics

The symbol $Z^{\dagger}$ denotes the pseudo-inverse (or Moore-Penrose generalized inverse) of a matrix Z. It is well-known that

$$P_Z=ZZ^{\dagger}$$

is the orthogonal projection onto $\mathfrak{R}(Z)$, and

$$P_{Z^H}=Z^H Z^{H\dagger}\equiv Z^\dagger Z \tag{2.6}$$

is the orthogonal projection onto $\mathfrak{R}(Z^H)$. Using the MacDuffee theorem (see [1], p. 23) one can directly verify the identity in (2.6).

Suppose that p, q and r are natural numbers satisfying $r\leqslant\min\{p, q\}$. Let

$$\mathbb{C}_r^{p\times q}=\{Z\in\mathbb{C}^{p\times q}: \text{rank } (Z)=r\}.$$

Elements of $\mathbb{C}_r^{p\times q}$ are divided into equivalence classes as follows: two elements Z and W are said to belong to the same equivalence class (symbolically $Z\sim W$), if $\mathfrak{R}(Z)=\mathfrak{R}(W)$. We consider every equivalence class of $\mathbb{C}_r^{p\times q}$ as a point and consequently obtain a complex projective space, symbolically G_r^p. Similarly, one can utilize $\mathfrak{R}(Z^H)=\mathfrak{R}(W^H)$ to define $Z\sim W$ and obtain a complex projective space $G_{r,q}$. We usually use one representative of equivalence classes, i. e., a $p\times q$ matrix (or a $q\times p$ matrix) whose rank is r, to represent a point of G_r^p (or $G_{r,q}$).

Theorem 2.3. *Let* $\|\ \|$ *be any unitary-invariant norm on* $\mathbb{C}^{p\times p}$. *Then* $\|P_Z-P_W\|$ *is a unitary-invariant metric on* G_r^p, *i. e.*

(1) $\|P_Z-P_W\|\geqslant 0$, $\|P_Z-P_W\|=0$ *iff* $Z\sim W$;

(2) $\|P_Z-P_W\|=\|P_W-P_Z\|$;

(3) $\|P_Z-P_W\|\leqslant\|P_Z-P_X\|+\|P_X-P_W\|$;

(4) *For any unitary matrix* $Q\in\mathbb{C}^{p\times p}$ *and any non-singular matrices* P, $R\in\mathbb{C}^{q\times q}$, $\|P_{QZP}-P_{QWR}\|=\|P_Z-P_W\|$, *Where* Z, W *and* X *are any points on* G_r^p.

Proof. We only need to prove the later conclusion of (1) and (4).

1. Prove "$\|P_Z-P_W\|=0$ iff $Z\sim W$". We take the full-rank factorizations of Z and W[1, p.22]

$$Z=FG,\ W=ST,\ F,\ S\in\mathbb{C}^{p\times r},\ G,\ T\in\mathbb{C}^{r\times q}. \tag{2.7}$$

Obviously, $\qquad \mathfrak{R}(Z)=\mathfrak{R}(F),\ \mathfrak{R}(W)=\mathfrak{R}(S).$

If $Z\sim W$, then $\mathfrak{R}(F)=\mathfrak{R}(S)$, i. e., there exists a non-singular matrix $K\in\mathbb{C}^{r\times r}$ such that $S=FK$. Substituting this relation into

$$P_Z=F(F^HF)^{-1}F^H,\ P_W=S(S^HS)^{-1}S^H, \tag{2.8}$$

we obtain $P_Z=P_W$.

Conversely, suppose that $P_Z=P_W$. Let $F(F^HF)^{-\frac{1}{2}}=U_1$ and $S(S^HS)^{-\frac{1}{2}}=V_1$. Taking matrices U_2, $V_2\in\mathbb{C}^{p\times(p-r)}$ such that $U=(U_1,\ U_2)$ and $V=(V_1,\ V_2)$ are unitary, then from $U_1U_1^H=V_1V_1^H$ we obtain

$$U\begin{pmatrix} I & 0\\ 0 & 0\end{pmatrix}U^H=V\begin{pmatrix} I & 0\\ 0 & 0\end{pmatrix}V^H,$$

thus $U_1^HV_2=0$, $U_2^HV_1=0$. This shows that $\mathfrak{R}(U_1)=\mathfrak{R}(V_1)$, i. e. $\mathfrak{R}(F)=\mathfrak{R}(S)$, and so $Z\sim W$.

2. Prove (4). Taking the full-rank factorizations (2.7) for Z and W, and using the representations (2.8), after some calculations we have

$$\|P_{QZP}-P_{QWR}\|=\|Q(P_Z-P_W)Q^H\|=\|P_Z-P_W\|.$$

Similarly one can prove

Theorem 2.4. *Let* $\|\ \|$ *be any unitary-invariant norm on* $\mathbb{C}^{q\times q}$. *Then* $\|P_{Z^H}-P_{W^H}\|$ *is a unitary-invariant metric on* $G_{r,q}$. *Here the meaning of "unitary-invariant" is as* *follows:* (4) *For any unitary matrix* $Q\in\mathbb{C}^{q\times q}$ *and any non-singular matrices* $P,\ R\in\mathbb{C}^{p\times p}$, *we have*

$$\|P_{(PZQ)^H}-P_{(RWQ)^H}\|=\|P_{Z^H}-P_{W^H}\|,\ \forall Z,\ W\in G_{r,q}.$$

In the sections § 3 and § 4 we shall adopt the following notations:

$$d_F(Z,\ W)=\frac{1}{\sqrt{2}}\|P_{Z^H}-P_{W^H}\|_F,\ d_2(Z,\ W)=\|P_{Z^H}-P_{W^H}\|_2, \tag{2.9}$$

where $\|\ \|_F$ and $\|\ \|_2$ denote the Frobenius norm and the spectral norm, respectively.

2.4. Acute perturbations

Definition 2.5[7, p.641, 9]. *Let* $Z,\ W\in\mathbb{C}^{p\times q}$. *W is an acute perturbation of Z if*

$$\|P_Z-P_W\|_2<1,\ \|P_{Z^H}-P_{W^H}\|_2<1.$$

Stewart[7] has proved the following theorem.

Theorem 2.5. *Let* $Z,\ W\in\mathbb{C}^{p\times q}$. *Then W is an acute perturbation of Z if and only if*

$$\operatorname{rank}(Z)=\operatorname{rank}(W)=\operatorname{rank}(P_ZWP_{Z^H})$$

3. The Hoffman-Wielandt Type Theorem

From the decompositions (2.1) and (2.5) we know that if we set $Z=(A,\ B)\in\mathbb{C}^{m\times 2n}$ for $\mu A-\lambda B\in\mathscr{D}_r^{m\times n}$ (specially, $\mathscr{N}_r^{m\times n}$), then $\operatorname{rank}(Z)=r$. Moreover, we notice that $PZ=(PA,\ PB)$ for a non-singular matrix $P\in\mathbb{C}^{m\times m}$ is corresponding to the pencil $\mu(PA)-\lambda(PB)\in\mathscr{D}_r^{m\times n}$ which has the same eigenvalues and eigenvectors as $\mu A-\lambda B$. Hence we can regard Z as a point on $G_{r,2n}$. In § 3 and § 4 we shall use the variation of Z on $G_{r,2n}$ to bound the perturbation of the eigenvalues of $\mu A-\lambda B$.

Theorem 3.1. *Let* $\mu A-\lambda B,\ \mu C-\lambda D\in\mathscr{N}_r^{m\times n}\,(r\geq 1)$, $\lambda(A,\ B)=\{(\alpha_i,\ \beta_i)\}_{i=1}^r$ *and* $\lambda(C,\ D)=\{(\gamma_i,\ \delta_i)\}_{i=1}^r$. *If we set* $Z=(A,\ B)$, $W=(C,\ D)$ *and*

$$\rho_{i,j}=\rho((\alpha_i,\ \beta_i),\ (\gamma_j,\ \delta_j)),\ 1\leq i,\ j\leq r,$$

then there exists a permutation $k_1,\ \cdots,\ k_r$ *of* $1,\ \cdots,\ r$ *such that*

$$\sqrt{\sum_{i=1}^r\rho_{i,k_i}^2}\leq d_F(Z,\ W), \tag{3.1}$$

where $d_F(Z,\ W)$ *is defined by* (2.9).

Proof.

1. In the decompositions (2.5) of A and B we may assume, without loss of generality, that $|A_1|^2+|\Omega_1|^2=I^{(r)}$. Writing

$$S=(\underset{r}{S_1},\ \underset{m-r}{S_2}),\ U=(\underset{r}{U_1},\ \underset{n-r}{U_2})$$

then from (2.5) we get the full-rank factorization of Z:

$$Z=S_1(A_1U_1^H,\ \Omega_1U_1^H).$$

Utilizing the MacDuffee theorem we obtain

$$Z^\dagger=\begin{pmatrix}U_1 & \bar{A}_1\\ U_1 & \bar{\Omega}_1\end{pmatrix}(S_1^HS_1)^{-1}S_1^H,$$

and so

$$P_{Z^H}=\begin{pmatrix}U_1 & \bar{A}_1\\ U_1 & \bar{\Omega}_1\end{pmatrix}(A_1U_1^H,\ \Omega_1U_1^H). \tag{3.2}$$

By Theorem 2.2, C and D have the decompositions

$$C=T\begin{pmatrix}\Gamma_1 & 0\\ 0 & 0\end{pmatrix}V^H,\quad D=T\begin{pmatrix}\Delta_1 & 0\\ 0 & 0\end{pmatrix}V^H \tag{3.3}$$

where $T\in\mathbb{C}^{m\times m}$ is non-singular, $V\in\mathbb{C}^{n\times n}$ is unitary, $\Gamma_1=\mathrm{diag}(\gamma_1,\cdots,\gamma_r)$, $\Delta_1=\mathrm{diag}(\delta_1,\cdots,\delta_r)$, and we may assume, without loss of generality, that $|\Gamma_1|^2+|\Delta_1|^2=I^{(r)}$. Hence, if we write $V=(V_1, V_2), V_1\in\mathbb{C}^{n\times r}$, then with the same argument as the above we obtain

$$P_{W^H}=\begin{pmatrix}V_1 & \overline{\Gamma}_1\\ V_1 & \overline{\Delta}_1\end{pmatrix}(\Gamma_1V_1^H,\ \Delta_1V_1^H). \tag{3.4}$$

2. According to (2.9),

$$d_F^2(Z, W)=\frac{1}{2}[\mathrm{tr}(P_{Z^H})+\mathrm{tr}(P_{W^H})]-\mathrm{tr}(P_{Z^H}P_{W^H}). \tag{3.5}$$

Utilizing the expressions (3.2) and (3.4) we obtain

$$\mathrm{tr}(P_{Z^H})=\mathrm{tr}(P_{W^H})=r \tag{3.6}$$

and

$$\mathrm{tr}(P_{Z^H}P_{W^H})=f(R), \tag{3.7}$$

where

$$R=U_1^HV_1=(r_{ij})\in\mathbb{C}^{r\times r} \tag{3.8}$$

and

$$f(R)=\mathrm{tr}[(\Lambda_1R\overline{\Gamma}_1+\Omega_1R\overline{\Delta}_1)(\Lambda_1R\overline{\Gamma}_1+\Omega_1R\overline{\Delta}_1)^H]. \tag{3.9}$$

Therefore from (3.5)—(3.7),

$$d_F^2(Z, W)=r-f(R). \tag{3.10}$$

3. From (3.9),

$$f(R)=\sum_{i,j=1}^{r}\theta_{i,j}y_{ij}, \tag{3.11}$$

where

$$\theta_{i,j}=|\alpha_i\overline{\gamma}_j+\beta_i\delta_j|^2,\quad y_{ij}=|r_{ij}|^2,\ 1\leqslant i,\ j\leqslant r. \tag{3.12}$$

By (3.8), the matrix R satisfies $RR^H\leqslant I^{(r)}$ and $R^HR\leqslant I^{(r)}$; combining these relations with (3.12) we know that the matrix (y_{ij}) satisfies

$$y_{ij}\geqslant 0,\ \sum_{j=1}^{r}y_{ij}\leqslant 1,\ \sum_{i=1}^{r}y_{ij}\leqslant 1,\ 1\leqslant i,\ j\leqslant r. \tag{3.13}$$

Let
$$\mathfrak{X}_r=\{X=(x_{ij})\in\mathbb{C}^{r\times r}: x_{ij}\geqslant 0,\ \sum_{j=1}^{r}x_{ij}=\sum_{i=1}^{r}x_{ij}=1,\ 1\leqslant i,\ j\leqslant r\},$$

i. e., $\mathfrak{X}_r$ is the set of all $r\times r$ bistochastic matrices. It is easy to see that for any r^2 non-negative numbers $\{y_{ij}\}$ satisfying the conditions expressed in (3.13), there exists a matrix $X_0=(x_{ij}^{(0)})\in\mathfrak{X}_r$ such that $y_{ij}\leqslant x_{ij}^{(0)}$ for $1\leqslant i,\ j\leqslant r$. Substituting these inequalities into (3.11) we obtain

$$f(R)\leqslant\sum_{i,j=1}^{r}\theta_{ij}x_{ij}^{(0)}=g(X_0),$$

where $g(X)$ is a linear function of X on $\mathfrak{X}_r$. As $\mathfrak{X}_r$ is a convex polyhedron the vertices of which are the permutation matrices (ref. [5]) there exists a permutation matrix $P=(p_{ij})$ (where $p_{ij}=\delta_{jk_i}$, $1\leqslant i,\ j\leqslant r$. δ_{ij} is the Kronecker's symbol) such that

$$f(R)\leqslant g(P)=\sum_{i=1}^{r}\theta_{i,k_i}.$$

Substituting this inequality into (3.10) it follows that

$$d_F^2(Z, W) \geq \sum_{i=1}^{r} (1-\theta_{i,k_i}) = \sum_{i=1}^{r} |\alpha_i \delta_{k_i} - \beta_i \gamma_{k_i}|^2 = \sum_{i=1} \rho_{i,k_i}^2.$$

This is the inequality (3.1).

4. The Bauer–Fike Type Theorem

Let $\mu A - \lambda B \in \mathscr{D}_r^{m\times n} (r \geq 1)$, $\mu D - \lambda D \in \mathfrak{S}_s^{m\times n} (s \geq 1)$, $\lambda(A, B) = \{(\alpha_i, \beta_i)\}_{i=1}^r$ and $\lambda(C, D) = \{(\gamma_i, \delta_i)\}_{i=1}^s$. In this section we search for a upper bound under some appropriate conditions for the generalized spectral variation of $\mu C - \lambda D$ with respect to $\mu A - \lambda B$

$$s_Z(W) = \max_{1<j<s} \min_{1<i<r} \rho((\alpha_i, \beta_i), (\gamma_j, \delta_j)), \tag{4.1}$$

where $Z = (A, B)$, $W = (C, D)$.

First of all we give other expressions of the decompositions (2.1). Let

$$S = UT, \ Q = VR \tag{4.2}$$

be the unitary-triangular factorizations of S and Q, where U and V are unitary matrices, $T = \begin{pmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{pmatrix}$ and $R = \begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix}$ are non-singular upper triangular matrices, and $T_{11}, R_{11} \in \mathbb{C}^{r\times r}$. Substituting the decompositions (4.2) into (2.1) and setting

$$K_1 = T_{11} A_1 R_{11}^H, \ L_1 = T_{11} \Omega_1 R_{11}^H, \tag{4.3}$$

then we obtain

$$A = U\begin{pmatrix} K_1 & 0 \\ 0 & 0 \end{pmatrix} V^H, \quad B = U\begin{pmatrix} L_1 & 0 \\ 0 & 0 \end{pmatrix} V^H. \tag{4.4}$$

Lemma 4.1. *Suppose that* $\mu A - \lambda B \in \mathscr{D}_r^{m\times n} (r \geq 1)$ *with the decompositions* (4.4), $\mu C - \lambda D \in \mathfrak{S}_s^{m\times n} (s \geq 1)$. *Let*

$$\tilde{C} = U^H C V = \begin{pmatrix} \tilde{C}_{11} & \tilde{C}_{12} \\ \tilde{C}_{21} & \tilde{C}_{22} \end{pmatrix}, \quad \tilde{D} = U^H D V = \begin{pmatrix} \tilde{D}_{11} & \tilde{D}_{12} \\ \tilde{D}_{21} & \tilde{D}_{22} \end{pmatrix}, \tag{4.5}$$

where $\tilde{C}_{11}, \tilde{D}_{11} \in \mathbb{C}^{r\times r}$. *If* $W = (C, D)$ *is an acute perturbation of* $Z = (A, B)$, *and* $W' = \begin{pmatrix} C \\ D \end{pmatrix}$ *is an acute perturbation of* $Z' = \begin{pmatrix} A \\ B \end{pmatrix}$, *then*

$$\lambda(C, D) = \lambda(\tilde{C}_{11}, \tilde{D}_{11}).$$

Proof.

1. Let

$$\begin{cases} \tilde{Z} = U^H Z \begin{pmatrix} V & 0 \\ 0 & V \end{pmatrix} = \begin{pmatrix} K_1 & 0 & L_1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\ \tilde{W} = U^H W \begin{pmatrix} V & 0 \\ 0 & V \end{pmatrix} = \begin{pmatrix} \tilde{C}_{11} & \tilde{C}_{12} & \tilde{D}_{11} & \tilde{D}_{12} \\ \tilde{C}_{21} & \tilde{C}_{22} & \tilde{D}_{21} & \tilde{D}_{22} \end{pmatrix}. \end{cases} \tag{4.6}$$

Thus $\tilde{Z} = \begin{pmatrix} I^{(r)} \\ 0 \end{pmatrix} (K_1, 0, L_1, 0)$, $\tilde{Z}^\dagger = (K_1, 0, L_1, 0)^H M_1 (I^{(r)}, 0)$,

where $M_1 = (K_1 K_1^H + L_1 L_1^H)^{-1}$. and so

$$P_{\tilde{Z}} = (I^{(r)}, 0)^H (I^{(r)}, 0), \ P_{\tilde{Z}^H} = (K_1, 0, L_1, 0)^H M_1 (K_1, 0, L_1, 0). \tag{4.7}$$

Observe that W is an acute perturbation of Z, U and V are unitary matrices, then by the unitary-invariableness of $\|P_Z-P_W\|_2$ and $\|P_{Z^H}-P_{W^H}\|_2$ (see Theorem 2.3 and Theorem 2.4) as well as Definition 2.5 we know that $\widetilde{W}$ is an acute perturbation of $\widetilde{Z}$. Hence from Theorem 2.5 and (4.6),

$$\operatorname{rank}(P_{\widetilde{Z}}\widetilde{W}P_{\widetilde{Z}^H})=\operatorname{rank}(\widetilde{W})=\operatorname{rank}(\widetilde{Z})=r.$$

This together with (4.7) gives

$$\operatorname{rank}((\widetilde{C}_{11},\ \widetilde{D}_{11})(K_1,\ L_1)^H M_1(K_1,\ L_1))=r,$$

and so we must have

$$\operatorname{rank}(\widetilde{C}_{11},\ \widetilde{D}_{11})=r. \tag{4.8}$$

However, $\operatorname{rank}(\widetilde{W})=r$, therefore there exists $F_1\in\mathbb{C}^{(m-r)\times r}$ such that

$$(\widetilde{C}_{21},\ \widetilde{C}_{22},\ \widetilde{D}_{21},\ \widetilde{D}_{22})=F_1(\widetilde{C}_{11},\ \widetilde{C}_{12},\ \widetilde{D}_{11},\ \widetilde{D}_{12}). \tag{4.9}$$

2. Let

$$\widetilde{Z}'=\begin{pmatrix}U&0\\0&U\end{pmatrix}^H Z'V=\begin{pmatrix}K_1&0\\0&0\\L_1&0\\0&0\end{pmatrix},\ \widetilde{W}'=\begin{pmatrix}U&0\\0&U\end{pmatrix}^H W'V=\begin{pmatrix}\widetilde{C}_{11}&\widetilde{C}_{12}\\\widetilde{C}_{21}&\widetilde{C}_{22}\\\widetilde{D}_{11}&\widetilde{D}_{12}\\\widetilde{D}_{21}&\widetilde{D}_{22}\end{pmatrix}. \tag{4.10}$$

According to the hypothesis with the same argument as the above we can deduce that $\widetilde{W}'$ is an acute perturbation of $\widetilde{Z}'$. Hence from Theorem 2.5 and (4.10),

$$\operatorname{rank}(P_{\widetilde{Z}'}\widetilde{W}'P_{\widetilde{Z}'^H})=\operatorname{rank}(\widetilde{W}')=\operatorname{rank}(\widetilde{Z}')=r. \tag{4.11}$$

Observe that

$$\widetilde{Z}'=\begin{pmatrix}K_1\\0\\L_1\\0\end{pmatrix}(I^{(r)},\ 0),\quad \widetilde{Z}'^{\dagger}=\begin{pmatrix}I^{(r)}\\0\end{pmatrix}N_1(K_1^H,\ 0,\ L_1^H,\ 0),$$

where $N_1=(K_1^HK_1+L_1^HL_1)^{-1}$; and so

$$P_{\widetilde{Z}'}=\begin{pmatrix}K_1\\0\\L_1\\0\end{pmatrix}N_1(K_1^H,\ 0,\ L_1^H,\ 0),\quad P_{\widetilde{Z}'^H}=\begin{pmatrix}I^{(r)}\\0\end{pmatrix}(I^{(r)},\ 0).$$

These together with (4.11) give

$$\operatorname{rank}\left[\begin{pmatrix}K_1\\L_1\end{pmatrix}N_1(K_1^H,\ L_1^H)\begin{pmatrix}\widetilde{C}_{11}\\\widetilde{D}_{11}\end{pmatrix}\right]=r,$$

thus we must have

$$\operatorname{rank}\begin{pmatrix}\widetilde{C}_{11}\\\widetilde{D}_{11}\end{pmatrix}=r. \tag{4.12}$$

However, $\operatorname{rank}(\widetilde{W}')=r$, therefore there exists $G_1\in\mathbb{C}^{r\times(n-r)}$ such that

$$\begin{pmatrix}\widetilde{C}_{12}\\\widetilde{C}_{22}\\\widetilde{D}_{12}\\\widetilde{D}_{22}\end{pmatrix}=\begin{pmatrix}\widetilde{C}_{11}\\\widetilde{C}_{21}\\\widetilde{D}_{11}\\\widetilde{D}_{21}\end{pmatrix}G_1. \tag{4.13}$$

3. As a result from (4.5), (4.9) and (4.13) we obtain

$$\begin{pmatrix} I & 0 \\ -F & I \end{pmatrix} U^H C V \begin{pmatrix} I & -G \\ 0 & I \end{pmatrix} = \begin{pmatrix} \tilde{C}_{11} & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} I & 0 \\ -F & 0 \end{pmatrix} U^H D V \begin{pmatrix} I & -G \\ 0 & I \end{pmatrix} = \begin{pmatrix} \tilde{D}_{11} & 0 \\ 0 & 0 \end{pmatrix}. \tag{4.14}$$

Combining (4.14) with (1.1) we reach the conclusion of Lemma 4.1.

Lemma 4.2. *Suppose that $\mu A - \lambda B \in \mathfrak{S}_r^{m\times n}(r \geqslant 1)$. Let $Z = (A, B)$ and*

$$Z_1 = [(ZZ^H)^\dagger]^{\frac{1}{2}} Z = (A_1, B_1), \tag{4.15}$$

where we take the semi-positive definite Square root for $[(ZZ^H)^\dagger]^{\frac{1}{2}}$. Then

(1) *$\mu A_1 - \lambda B_1$ and $\mu A - \lambda B$ have the same eigenvalues and eigenvectors;*

(2) *$Z_1 Z_1^H = X\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix} X^H$, X is a unitary matrix;*

(3) *$P_{Z_1^H} = P_{Z^H}$.*

Proof. Let $Z = X\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} Y^H$ be the singular value decomposition of Z, where $\Sigma \in \mathbb{C}^{r\times r}$ is a non-singular diagonal matrix, $X = (X_1, X_2)$ and $Y = (Y_1, Y_2)$ are unitary matrices, $X_1 \in \mathbb{C}^{m\times r}$ and $Y_1 \in \mathbb{C}^{2n\times r}$. Therefrom we have

$$[(ZZ^H)^\dagger]^{\frac{1}{2}} = X\begin{pmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} X^H$$

and

$$Z_1 = X\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix} Y^H = PZ, \tag{4.16}$$

where $P = X\begin{pmatrix} \Sigma^{-1} & 0 \\ 0 & I \end{pmatrix} X^H \in \mathbb{C}^{m\times m}$ is non-singular.

From (4.16) we obtain the conclusions (1) and (2) at once.

Moreover, from the singular value decomposition of Z, $Z = X_1 \Sigma Y_1^H$, thus

$$Z^\dagger = Y_1 \Sigma^{-1} X_1^H, \quad P_{Z^H} = Y_1 Y_1^H;$$

on the other hand, from (4.16), $Z_1 = X_1 Y_1^H$, and so

$$Z_1^\dagger = Y_1 X_1^H, \quad P_{Z_1^H} = Y_1 Y_1^H.$$

Therefore the conclusion (3) is also true.

Theorem 4.1. *Suppose that $\mu A - \lambda B \in \mathscr{D}_r^{m\times n}(r \geqslant 1)$ with the decompositions (2.2), and $\mu C - \lambda D \in \mathfrak{S}_s^{m\times n}(s \geqslant 1)$. If $W = (C, D)$ is an acute perturbation of $Z = (A, B)$, and $\begin{pmatrix} C \\ D \end{pmatrix}$ is an acute perturbation of $\begin{pmatrix} A \\ B \end{pmatrix}$, then*

$$s_Z(W) \leqslant \sqrt{\|Q_1^H Q_1\|_2 \|(Q_1^H Q_1)^{-1}\|_2}\, d_2(Z, W), \tag{4.17}$$

where $s_Z(W)$ and $d_2(Z, W)$ are defined as in (4.1) and (2.9), respectively.

Proof. Without loss of generality we may assume that the diagonal matrices Λ_1 and Ω_1 in (2.2) satisfy $|\Lambda_1|^2 + |\Omega_1|^2 = I^{(r)}$. Besides, by Theorem 4.2 we may assume that the matrix Z satisfies

$$ZZ^H = X\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix} X^H, \ X \text{ is unitary.} \tag{4.18}$$

Let (γ, δ) be an eigenvalue of $\mu C - \lambda D$. It is safe to suppose that $|\gamma|^2 + |\delta|^2 = 1$

For a suitable normalized eigenvector x of $\mu C-\lambda D$ corresponding to (γ, δ) (the choice of the vector x will be explained in the following), we have

$$\delta Ax-\gamma Bx=\delta(A-ZW^{\dagger}C)x-\gamma(B-ZW^{\dagger}D)x=(A-ZW^{\dagger}C,\ B-ZW^{\dagger}D)\begin{pmatrix}\delta x\\ -\gamma x\end{pmatrix}$$
$$=(Z-ZW^{\dagger}W)\begin{pmatrix}\delta x\\ -\gamma x\end{pmatrix}=Z(Z^{\dagger}Z-W^{\dagger}W)\begin{pmatrix}\delta x\\ -\gamma x\end{pmatrix}. \tag{4.19}$$

From the transformation (4.5),

$$C=U\begin{pmatrix}C_{11} & C_{12}\\ C_{21} & C_{22}\end{pmatrix}V^{H},\quad D=U\begin{pmatrix}\tilde{D}_{11} & \tilde{D}_{12}\\ \tilde{D}_{21} & \tilde{D}_{22}\end{pmatrix}V^{H}; \tag{4.20}$$

and by Lemma 4.1, $\lambda(C,\ D)=\lambda(\tilde{C}_{11},\ \tilde{D}_{11})$. Hence we can choose a normalized eigenvector u of $\mu\tilde{C}_{11}-\lambda\tilde{D}_{11}$ corresponding to $(\gamma,\ \delta)\in\lambda(\tilde{C}_{11},\ \tilde{D}_{11})$:

$$\delta\tilde{C}_{11}u=\gamma\tilde{D}_{11}u,\quad u\in\mathbb{C}^{r}.$$

Let $x=V\begin{pmatrix}u\\ 0\end{pmatrix}$, then x is a normalized eigenvector of $\mu C-\lambda D$ corresponding to $(\gamma,\ \delta)\in\lambda(C,\ D)$. Substituting this x and the decompositions (4.4) into (4.19), we get

$$U\begin{pmatrix}\delta K_1-\gamma L_1 & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}u\\ 0\end{pmatrix}=Z(P_{Z^H}-P_{W^H})\begin{pmatrix}\delta V\\ -\gamma V\end{pmatrix}\begin{pmatrix}u\\ 0\end{pmatrix}.$$

Thus

$$\|(\delta K_1-\gamma L_1)u\|\leqslant\|Z\|_2\|P_{Z^H}-P_{W^H}\|_2\|u\|\leqslant d_2(Z,\ W), \tag{4.21}$$

where $\|\ \|$ denotes the usual Euclidean vector norm.

Substituting (4.3) into the left side of (4.21), we obtain

$$\|(\delta K_1-\gamma L_1)u\|=\|T_{11}(\delta\Lambda_1-\gamma\Omega_1)R_{11}^{H}u\|$$
$$\geqslant\|T_{11}^{-1}\|_2^{-1}\|R_{11}^{-1}\|_2^{-1}\min_{1\leqslant i\leqslant r}\rho((\alpha_i,\ \beta_i),\ (\gamma,\ \delta)).$$

Therefore, for any $(\gamma,\ \delta)\in\lambda(C,\ D)$ we have

$$\min_{1\leqslant i\leqslant r}\rho((\alpha_i,\ \beta_i),\ (\gamma,\ \delta))\leqslant\|T_{11}^{-1}\|_2\|R_{11}^{-1}\|_2\,d_2(Z,\ W). \tag{4.22}$$

Observe that the matrix ZZ^H has the decomposition (4.18); but from (4.4)

$$ZZ^H=U\begin{pmatrix}K_1K_1^H+L_1L_1^H & 0\\ 0 & 0\end{pmatrix}U^H,\ U\text{ is unitary.}$$

Hence we must have

$$K_1K_1^H+L_1L_1^H=I^{(r)}. \tag{4.23}$$

Substituting (4.3) into (4.23) we obtain

$$(T_{11}^{H}T_{11})^{-1}=\Lambda_1R_{11}^{H}R_{11}\bar{\Lambda}_1+\Omega_1R_{11}^{H}R_{11}\bar{\Omega}_1,$$

thus

$$\|T_{11}^{-1}\|_2\leqslant\|R_{11}\|_2. \tag{4.24}$$

Moreover, from (4.2) and (2.2), $Q_1=V_1R_{11}$. So we get

$$\|R_{11}\|_2=\|Q_1^HQ_1\|_2^{1/2},\quad\|R_{11}^{-1}\|_2=\|(Q_1^HQ_1)^{-1}\|^{\frac{1}{2}}.$$

Substituting these equalities and (4.24) into (4.22) and remembering that $(\gamma,\ \delta)$ is an arbitrary eigenvalue of $\mu C-\lambda D$, then we obtain (4.17).

In case of a singular normal pencil $\mu A-\lambda B$ the matrix Q in (2.1) can be chosen as unitary (see Theorem 2.2), hence we get at once

Theorem 4.2. *Let* $\mu A-\lambda B\in\mathfrak{N}_r^{m\times n}(r\geqslant 1)$, $\mu C-\lambda D\in\mathfrak{S}_s^{m\times n}(s\geqslant 1)$. *If* $W=(C,D)$ *is an acute perturbation of* $Z=(A,B)$, *and* $\begin{pmatrix}C\\D\end{pmatrix}$ *is an acute perturbation of* $\begin{pmatrix}A\\B\end{pmatrix}$, *then*

$$s_Z(W)\leqslant d_2(Z,W).$$

5. Final Remarks

5.1. Theorem 4.1 shows that in the case where (A, B) and $\begin{pmatrix}A\\B\end{pmatrix}$ are acutely perturbed we can use the variation of the orthogonal projection onto $\mathscr{R}\begin{pmatrix}A^H\\B^H\end{pmatrix}$ to bound the perturbation of the eigenvalues of a singular diagonalizable pencil $\mu A-\lambda B$. It is worth-while to point out that under the hypothesis of acute perturbation Stewart[7] has obtained an estimation for the variation of the orthogonal projection. By Theorem 4.1 in [7], if $W=(C,D)$ is an acute perturbation of $Z=(A,B)$, then

$$d_2(Z,W)\leqslant\frac{\bar{\varkappa}p(E)}{[1+(\bar{\varkappa}p(E))^2]^{\frac{1}{2}}}<1, \tag{5.1}$$

where

$$\bar{\varkappa}=\|Z\|_2\|(P_{Z^H}W^HP_Z)^\dagger\|_2$$

and

$$E=(W-Z)^H,\ p(E)=\|(I-P_{Z^H})EP_Z\|_2/\|Z\|_2. \tag{5.2}$$

Therefore from (4.17) and (5.1)—(5.2) we know that, in the case where (A, B) and $\begin{pmatrix}A\\B\end{pmatrix}$ are acutely perturbed, if we use the chordal metric to describe the perturbation of eigenvalues, then the eigenvalues of a singular diagonalizable pencil $\mu A-\lambda B$ are insensitive to perturbations in the elements of A and B.

5.2. If $\mu A-\lambda B$ is a regular pencil, then we can use Definition 2.3 and Definition 2.4 to define the regular diagonalizable pencil and the regular normal pencil, respectively (the corresponding matrix-pairs are called the diagonalizable pair and the normal pair in [8] and [3]); and in these cases the inequalities (3.1) and (4.17) are exactly the conclusions of the Hoffman-Wielandt theorem and the Bauer-Fike theorem for regular pencils, respectively (see [8] and [3]. There are different expressions for $d_F(Z,W)$ and $d_2(Z,W)$, ref. [3], Theorem 1.3. In [8] we have written the $d_2(Z,W)$ as $d_s(Z,W)$).

References

[1] A. Ben-Israel and T. N. E. Greville, Generalized Inverses: Theory and Applications, John Wiley, New York,1974.

[2] P. van Dooren,The computation of Kronecker's canonical form of a singular pencil, Linear Algebra and Appl., 27, 1979, 103—140.

[3] L. Elsner and J. G. Sun, Perturbation theorems for the generalized eigenvalues problem, Linear Algebra and Appl. (to appear)

[4] F. R. Gantmaher, Theory of Matrices, Vol. II, Chelsea, New York, 1959.

[5] M. Marcus and H. Minc, A Survey of Matrix Theory and Matrix Inequalities, Allyn & Bacon, Boston, 1964.
[6] G. W. Stewart, On the sensitivity of the eigenvalue problem $Ax=\lambda Bx$, *SIAM J. Numer. Anal.*, 9, 1972, 669—686.
[7] G. W. Stewart, On the perturbation of pseudo-inverses, projections and linear least squares problems, *SIAM Rev.*, 19, 1977, 634—662.
[8] J. G. Sun, The perturbation bounds of generalized eigenvalues of a class of matrix-pairs, *Math. Numer. Sinica*, 4, 1982, 23—29.
[9] P. —Å. Wedin, Perturbation theory for pseudo-inverses, *BIT*, 13, 1973, 217—232.
[10] J. H. Wilkinson, Kronecker's canonical form and the QZ algorithm, *Linear Algebra and Appl.*, 28, 1979, 285—303.