PERTURBATION THEOREMS FOR GENERALIZED SINGULAR VALUES *1)

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

Abstract

Let A and B be $m\times n$ and $p\times n$ complex matrices respectively. This paper, as a continuation of the author's papers [7] (*Math. Numer. Sinica*, 4(1982), 229—233) and [8] (*SIAM J. Numer. Anal.*, to appear), discusses perturbation bounds for the generalized singular values of the matrix-pair $\{A, B\}$ in the case of rank $\binom{A}{B}<n$.

Let m, p and n be arbitrary natural numbers, A and B be $m\times n$ and $p\times n$ complex matrices respectively. Van Loan[10], Paige and Saunders[4] have suggested forms of the generalized singular value decomposition (GSVD) of the matrix-pair $\{A, B\}$. In two later papers[7,8] the author has analysed the perturbation of the singular values and the singular subspaces of $\{A, B\}$ in the case of rank$\binom{A}{B}=n$. In this paper we investigate the perturbation of the singular values of $\{A, B\}$ in the case of rank $\binom{A}{B}<n$ (Perturbation bounds for generalized singular subspaces of $\{A, B\}$ in this case have been given by the author in "The $\sin\theta$ theorems for generalized singular subspaces").

It is well-known that the singular values of an $m\times n$ matrix A are the non-negative square roots of the n eigenvalues of the positive semi-definite matrix A^HA (A^H is the conjugate transpose of A). In § 1 we generalize the singular value concept and derive the GSVD exactly from this point of view. Formerly, any pair (α, β) with $\alpha, \beta\geqslant 0$ and $(\alpha, \beta)\neq(0, 0)$ was regarded as a singular value of $\{A, B\}$ in the case of rank $\binom{A}{B}<n$ (Ref. [10], [4], [7]), and consequently it is difficult to investigatethe perturbation of singular values in this case; we shall clarify this problem in § 1. In § 2 and § 3 we prove a Weyl type theorem and a Hoffman-Wielandt type theorem respectively. The results show that, in the case where $\binom{A}{B}$ is acutely perturbed, if we use the chordal metric to describe the perturbation of singular values, then the singular values of $\{A, B\}$ are insensitive to perturbations in the elements of A and B.

Notation. Capital case is used for matrices and lower case Greek letters for

* Received October 12, 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.

scalars. The symbol $\mathbb{C}^{m\times n}$ denotes the set of complex $m\times n$ matrices, $\mathbb{C}^m=\mathbb{C}^{m\times 1}$ and $\mathbb{C}=\mathbb{C}^1$. A^T and A^H stand for transpose and conjugate transpose of A, respectively. $I^{(n)}$ is $n\times n$ identity matrix, and $0^{(n)}$ $n\times n$ null matrix. $A>0(\geq 0)$ denotes that A is a positive definite (positive semi-definite) Hermitian matrix. The column space of A is denoted by $R(A)$. $\|\ \|_2$ denotes the usual Euclidean vector norm and the spectral norm, and $\|\ \|_F$ the Frobenius matrix norm. $\sigma_{\max}(A)$ and $\sigma_{\min}(A)$ are the maximal singular value and the minimal singular value of A, respectively; and $\sigma^+_{\min}(A)$ is the minimal non-zero singular value of 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

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

§ 1. Generalized Singular Values and GSVD

We begin with the generalized eigenvalue concept.

Definition 1.1[9]. *Let A, $B\in\mathbb{C}^{m\times n}$, and*

$$\max_{(\lambda,\mu)\in G_{1,2}} \operatorname{rank}(\mu A-\lambda B)=k.$$

A number-pair $(\alpha, \beta)\in G_{1,2}$ is an eigenvalue of the pencil $\mu A-\lambda B$ if rank $(\beta A-\alpha B)<k$.

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

Theorem 1.1. *Let H, $K\in\mathbb{C}^{n\times n}$, and H, $K\geq 0$. If*

$$\max_{\sigma,\tau\geq 0}\operatorname{rank}(\tau H+\sigma K)=k, \tag{1.1}$$

then there exists a non-singular $S\in\mathbb{C}^{n\times n}$ such that

$$H=S\Lambda S^H,\ K=S\Omega S^H, \tag{1.2}$$

where

$$\Lambda=\operatorname{diag}(\Lambda_1, 0),\ \Omega=\operatorname{diag}(\Omega_1, 0), \tag{1.3}$$

$$\Lambda_1=\operatorname{diag}(I^{(r)}, \Lambda_{10}, 0^{(k-r-s)}),\ \Omega_1=\operatorname{diag}(0^{(r)}, \Omega_{10}, I^{(k-r-s)}), \tag{1.4}$$

$$\left.\begin{array}{c}\Lambda_{10}=\operatorname{diag}(\alpha^2_{r+1}, \cdots, \alpha^2_{r+s}),\ \Omega_{10}=\operatorname{diag}(\beta^2_{r+1}, \cdots, \beta^2_{r+s}),\\ 1>\alpha_{r+1}\geq\cdots\geq\alpha_{r+s}>0,\ 0<\beta_{r+1}\leq\cdots\leq\beta_{r+s}<1,\\ \alpha_i^2+\beta_i^2=1,\ r+1\leq i\leq r+s\end{array}\right\} \tag{1.5}$$

and r, $s\geq 0$, $r+s\leq k\leq n$.

Proof. From (1.1) there exist σ, $\tau\geq 0$ satisfying $\sigma^2+\tau^2=1$ such that $\operatorname{rank}(\tau H+\sigma K)=k$. Let

$$\tilde{H}=\sigma H-\tau K,\ \tilde{K}=\tau H+\sigma K. \tag{1.6}$$

Then there is a non-singular $Q\in\mathbb{C}^{n\times n}$ such that

$$K_0=Q\tilde{K}Q^H=\begin{pmatrix}I^{(k)} & 0\\ 0 & 0\end{pmatrix},\ H_0=Q\tilde{H}Q^H=\begin{pmatrix}H_{11} & H_{12}\\ H_{12}^H & H_{22}\end{pmatrix}. \tag{1.7}$$

Suppose that $\eta_0 I+H_{11}>0$ for a certain $\eta_0>0$. Let

$$L=\begin{pmatrix}I & 0\\ -H_{12}^H(\eta I+H_{11})^{-1} & I\end{pmatrix},\ \eta\geq\eta_0.$$

Then

$$L(\eta K_0+H_0)L^H=\text{diag}\ (\eta I+H_{11},\ H_{22}-H_{12}^H(\eta I+H_{11})^{-1}H_{12}).$$

From (1.1),

$$H_{22}-H_{12}^H(\eta I+H_{11})^{-1}H_{12}\equiv 0,\ \forall\eta\geq\eta_0;$$

and so

$$H_{12}=0,\ H_{22}=0. \tag{1.8}$$

Decomposing

$$H_{11}=U_1T_1U_1^H,\ T_1=\text{diag}(\tau_1,\ \cdots,\ \tau_k),\ \tau_1\geq\cdots\geq\tau_k,$$

where U_1 is unitary; substituting into (1.7) and combining with (1.6) and (1.8), and writing $R=Q^{-1}\begin{pmatrix}U_1 & 0\\ 0 & I\end{pmatrix}$, then we obtain

$$H=R\ \text{diag}\ (\sigma T_1+\tau I,\ 0)R^H,\ K=R\ \text{diag}(-\tau T_1+\sigma I,\ 0)R^H. \tag{1.9}$$

It follows from $H,\ K\geq 0$ and

$$(\sigma T_1+\tau I)^2+(-\tau T_1+\sigma I)^2=I+T_1^2>0$$

that

$$\Delta\equiv(\sigma T_1+\tau I)+(-\tau T_1+\sigma I)>0.$$

Hence, if we set

$$S=R\ \text{diag}(\Delta^{\frac{1}{2}},\ I),\ \Lambda_1=(\sigma T_1+\tau I)\Delta^{-1},\ \Omega=(-\tau T_1+\sigma I)\Delta^{-1}$$

and

$$\Lambda=\text{diag}(\Lambda_1,\ 0^{(n-k)}),\ \Omega=\text{diag}(\Omega_1,\ 0^{(n-k)}),$$

then from (1.9) we obtain (1.2)—(1.5) at once.

Theorem 1.1 shows that $\lambda(H,\ K)=\{(\alpha_i^2,\ \beta_i^2)\}_{i=1}^k$ for the above mentioned H and K, where α_i and β_i satisfy

$$\left.\begin{array}{c}1=\alpha_1=\cdots=\alpha_r>\alpha_{r+1}\geq\cdots\geq\alpha_{r+s}>\alpha_{r+s+1}=\cdots=\alpha_k=0,\\ 0=\beta_1=\cdots=\beta_r<\beta_{r+1}\leq\cdots\leq\beta_{r+s}<\beta_{r+s+1}=\cdots=\beta_k=1,\\ \alpha_i^2+\beta_i^2=1,\ i=1,\ \cdots,\ k.\end{array}\right\} \tag{1.10}$$

This fact suggests the following generalization of the singular value concept.

Definition 1.2. *Let $A\in\mathbb{C}^{m\times n}$, $B\in\mathbb{C}^{p\times n}$. A non-negative number-pair $(\alpha,\ \beta)$ is a singular value of the matrix-pair $\{A,\ B\}$ if $(\alpha^2,\ \beta^2)\in\lambda(A^HA,\ B^HB)$.*

The set of all singular values of $\{A,\ B\}$ is denoted by $\sigma\{A,\ B\}$.

Definition 1.3. *Let $A\in\mathbb{C}^{m\times n}$, $B\in\mathbb{C}^{p\times n}$. The matrix-pair $\{A,\ B\}$ is called an $(m,\ p,\ n;\ k)-MP$ if $\text{rank}\begin{pmatrix}A\\ B\end{pmatrix}=k$.*

The set of all $(m,\ p,\ n;\ k)-MP$ is denoted by $\mathbb{P}(m,\ p,\ n;\ k)$.

From Theorem 1.1 we can derive the following result which is due to Van Loan, Paige and Saunders.

Theorem 1.2 (GSVD)[10,4]. *Let $\{A,\ B\}\in\mathbb{P}(m,\ p,\ n;\ k)$. Then there exist unitary $U\in\mathbb{C}^{m\times m}$, $V\in\mathbb{C}^{p\times p}$ and non-singular $Q\in\mathbb{C}^{n\times n}$ such that*

$$U^HAQ=(\Sigma_A,\ 0),\ V^HBQ=(\Sigma_B,\ 0), \tag{1.11}$$

$$\Sigma_A=\begin{pmatrix} \Lambda_A & \\ \hline & 0_A \end{pmatrix}\begin{matrix} r+s \\ m-r-s \end{matrix}, \quad \Sigma_B=\begin{pmatrix} 0_B & \\ \hline & \Lambda_B \end{pmatrix}\begin{matrix} p+r-k \\ k-r, \end{matrix} \tag{1.12}$$
$$\quad r+s \quad k-r-s \qquad\qquad\qquad r \quad k-r$$

where 0_A *and* 0_B *are null matrices, and*

$$\Lambda_A=\operatorname{diag}(\alpha_1, \cdots, \alpha_{r+s}), \ \Lambda_B=\operatorname{diag}(\beta_{r+1}, \cdots, \beta_k) \tag{1.13}$$

satisfy (1.10).

Proof. From $\{A, B\}\in\mathbb{P}(m, p, n; k)$, $\max\limits_{\sigma,\tau>0} \operatorname{rank}(\tau A^HA+\sigma B^HB)=k$. Hence by Theorem 1.1, there exists a non-singular $Q\in\mathbb{C}^{n\times n}$ such that

$$Q^HA^HAQ=\Lambda, \ Q^HB^HBQ=\Omega, \tag{1.14}$$

where Λ and Ω are represented by (1.3)—(1.5).

Writing $Q=(\underset{r+s}{Q_1}, \underset{n-r-s}{Q_2})=(\underset{r}{Q_1'}, \underset{k-r}{Q_2'}, \underset{n-k}{Q_3'})$, then from (1.14), (1.3)—(1.5) and (1.13) we have

$$\Lambda_A^{-1}Q_1^HA^HAQ_1\Lambda_A^{-1}=I, \ Q_2^HA^HAQ_2=0$$

and

$$\Lambda_B^{-1}Q_2'^HB^HBQ_2'\Lambda_B^{-1}=I, \ Q_i'^HB^HBQ_i'=0, \ i=1, 3.$$

Now let $U_1=AQ_1\Lambda_A^{-1}$, $V_2=BQ_2'\Lambda_B^{-1}$, then $U_1^HU_1=I$, $V_2^HV_2=I$. Let U_2 and V_1 be chosen so that $U=(U_1, U_2)$ and $V=(V_1, V_2)$ are unitary. Then

$$U^HAQ=\begin{pmatrix} \Lambda_A & 0 \\ 0 & 0 \end{pmatrix}\begin{matrix} r+s \\ m-r-s \end{matrix}, \quad V^HBQ=\begin{pmatrix} 0 & 0 & 0 \\ 0 & \Lambda_B & 0 \end{pmatrix}\begin{matrix} p+r-k \\ k-r \end{matrix}. \tag{1.15}$$
$$r+s \quad n-r-s \qquad\qquad r \quad k-r \quad n-k$$

The decompositions (1.15) are exactly (1.11)—(1.13).

Theorem 1.2 shows that $\sigma\{A, B\}=\{(\alpha_i, \beta_i)\}_{i=1}^k$ for the matrices A and B mentioned in Theorem 1.2.

The following result is a corollary of Definition 1.1 and Definition 1.2.

Theorem 1.3. *Suppose that* $\{A, B\}\in\mathbb{P}(m, p, n; k)$. *Let* $Z=\begin{pmatrix} A \\ B \end{pmatrix}$, *and* $Z=Z_1F_1^H$ *be a full-rank factorization of* Z (Ref. [1], p. 22), *where* $Z_1=\begin{pmatrix} A_1 \\ B_1 \end{pmatrix}\begin{matrix} m \\ p \end{matrix}$. *Then*

$$\sigma\{A, B\}=\sigma\{A_1, B_1\}. \tag{1.16}$$

Proof. Let F_2 be chosen so that $F=(F_1, F_2)$ is non-singular. Then by Definition 1.1,

$$\begin{aligned}\lambda(A^HA, B^HB)&=\lambda\left(F\begin{pmatrix} A_1^HA_1 & 0 \\ 0 & 0 \end{pmatrix}F^H, F\begin{pmatrix} B_1^HB_1 & 0 \\ 0 & 0 \end{pmatrix}F^H\right)\\ &=\lambda(A_1^HA_1, B_1^HB_1);\end{aligned}$$

and by Definition 1.2 we get (1.16).

§ 2. The Hoffman-Wielandt Type Theorem

Regarding every k-dimensional column subspace in $\mathbb{C}^{m+p}$ as a point we obtain a complex projective space G_k^{m+p} consisting of all such points (Ref. [2], [9]). By [2], we may introduce a projective metric

$$d_L(Z_1, W_1)=\left\{1-\frac{|\det(Z_1^H W_1)|^2}{\det(Z_1^H Z_1)\det(W_1^H W_1)}\right\}^{\frac{1}{2}} \tag{2.1}$$

on G_k^{m+p}, where $Z_1, W_1 \in G_k^{m+p}$ (i. e. $Z_1, W_1 \in \mathbb{C}^{(m+p)\times k}$, and rank $(Z_1)=$ rank $(W_1)=k$). In this section we use $d_L(\cdot, \cdot)$ to bound perturbations of generalized singular values.

Theorem 2.1. *Let* $\{A, B\}, \{C, D\} \in \mathbb{P}(m, p, n; k)$, $0<k<n$, $\sigma\{A, B\}=\{(\alpha_i, \beta_i)\}_{i=1}^k$, $\sigma\{C, D\}=\{(\gamma_i, \delta_i)\}_{i=1}^k$, *and let* (α_i, β_i) *and* (γ_i, δ_i) *be ordered as in Theorem* 1.2 (*see* (1.10)). *Suppose that*

$$Z=Z_1F_1^H, \ W=W_1G_1^H \tag{2.2}$$

are any full-rank factorizations of $Z=\begin{pmatrix}A\\B\end{pmatrix}$ *and* $W=\begin{pmatrix}C\\D\end{pmatrix}$, *then*

$$\prod_{i=1}^k(1-\rho_{i,i}^2)\geqslant 1-d_L^2(Z_1, W_1), \tag{2.3}$$

where

$$\rho_{i,i}=\rho((\alpha_i, \beta_i), (\gamma_i, \delta_i)), \ 1\leqslant i\leqslant k. \tag{2.4}$$

Proof. Utilizing Theorem 2.1 of [8] and the above Theorem 1.3, we obtain (2.3) at once.

Utilizing the arithmetic-mean-geometric-mean inequality, from (2.3) we can obtain the following corollary.

Corollary 2.1. Assume the hypotheses of Theorem 2.1. Then

$$\sum_{i=1}^k \rho_{i,i}^2\leqslant k(1-\sqrt[k]{1-d_L^2(Z_1, W_1)}). \tag{2.5}$$

Now we investigate the relationship between the right-hand side of (2.3) and the elements of A, B, C and D. The symbol $Z^\dagger$ denotes the pseudo-inverse (or Moore-Penrose generalized inverse) of a matrix Z. It is well-known that $P_Z=ZZ^\dagger$ and $P_{Z^H}=Z^\dagger Z$ are the orthogonal projections onto $R(Z)$ and $R(Z^H)$, respectively. For the matrices Z, W, Z_1 and W_1 mentioned in Theorem 2.1 (see (2.2)), we define

$$d_F(Z, W)=\frac{1}{\sqrt{2}}\|P_Z-P_W\|_F, \ d_F(Z_1, W_1)=\frac{1}{\sqrt{2}}\|P_{Z_1}-P_{W_1}\|_F.$$

By MacDuffee theorem ([1], p. 23) we have $P_Z=P_{Z_1}$, $P_W=P_{W_1}$, and so

$$d_F(Z, W)=d_F(Z_1, W_1). \tag{2.6}$$

Moreover, it is easy to see that

$$d_L(Z_1, W_1)=[1-\det D(Z_1, W_1)]^{\frac{1}{2}}, \ d_F(Z_1, W_1)=\{\mathrm{tr}[I-D(Z_1, W_1)]\}^{\frac{1}{2}},$$

where

$$D(Z_1, W_1)=(Z_1^HZ_1)^{-\frac{1}{2}}Z_1^HW_1(W_1^HW_1)^{-1}W_1^HZ_1(Z_1^HZ_1)^{-\frac{1}{2}}.$$

If the non-negative square roots of the eigenvalues of $D(Z_1, W_1)$ are $\cos\theta_l$ $(1\leqslant l\leqslant k)$, then we have

$$d_L(Z_1, W_1)=\sqrt{1-\prod_{l=1}^k\cos^2\theta_l}, \ d_F(Z_1, W_1)=\sqrt{\sum_{l=1}^k\sin^2\theta_l}.$$

Obviously,

$$d_L(Z_1, W_1)\leqslant d_F(Z_1, W_1).$$

Combining this inequality with the relation (2.6), we obtain

$$d_L(Z_1, W_1) \leqslant d_F(Z, W). \tag{2.7}$$

It is worth noting that the perturbation bounds of generalized singular values given in (2.3), (2.5) and (2.7) are dependent on the non-Euclidean metrics $d_L(Z_1, W_1)$ and $d_F(Z, W)$ but not on the Euclidean metric $\|W-Z\|_F$, and in general it is not possible to compare the magnitude of $d_L(Z_1, W_1)$ (or $d_F(Z, W)$) with that of $\|W-Z\|_F$. Let us consider the following two examples.

Example 2.1. Let $\{A, B\}=\left\{\begin{pmatrix} n & 0 \\ 0 & 0 \end{pmatrix}, (0,0)\right\}$, $\{C, D\}=\left\{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, (0, 0)\right\}$, where n is a natural number. Obviously, $\{A, B\}$ and $\{C, D\} \in \mathbb{P}(2, 1, 2; 1)$, $\sigma\{A, B\}=\sigma\{C, D\}=\{(1, 0)\}$, and so $\rho_{1,1}=0$. Taking the full-rank factorizations

$$Z=\begin{pmatrix} n & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}=\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}(n, 0)=Z_1F_1^H,\quad W=\begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}=\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}(0, 1)=W_1G_1^H,$$

we have $d_L(Z_1, W_1)=d_F(Z, W)=0$, but $\|W-Z\|_F=\sqrt{n^2+1}>n$.

Example 2.2. Let $\{A, B\}=\left\{\begin{pmatrix} \varepsilon & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, (0, 0, 0)\right\}$, $\{C, D\}=\left\{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1+\varepsilon \end{pmatrix}, (0, \varepsilon, 0)\right\}$, where $0<\varepsilon\ll 1$. Obviously, $\{A, B\}$ and $\{C, D\} \in \mathbb{P}(2, 1, 3; 2)$, $\sigma\{A, B\}=\{(1, 0), (1, 0)\}$, $\sigma\{C, D\}=\{(1, 0), (0, 1)\}$, and so $\rho_{1,1}=0$, $\rho_{2,2}=1$. Taking the full-rank factorizations

$$Z=\begin{pmatrix} \varepsilon & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \varepsilon & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}=Z_1F_1^H,$$

$$W=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1+\varepsilon \\ 0 & \varepsilon & 0 \end{pmatrix}=\begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 0 & 1+\varepsilon \\ 0 & \varepsilon & 0 \end{pmatrix}=W_1G_1^H,$$

we have $d_L(Z_1, W_1)=d_F(Z, W)=1$, but $\|W-Z\|_F=\sqrt{3}\,\varepsilon\ll 1$.

Example 2.1 shows that if $\{A, B\}$, $\{C, D\} \in \mathbb{P}(m, p, n; k)$ and $d_F(Z, W)$ is very small (and so $d_L(Z_1, W_1)$ is very small too), then all of the $\rho_{i,i}$ are certainly very small, even though $\|W-Z\|_F$ is not small; Example 2.2 shows that not all of the $\rho_{i,i}$ are very small even if $\|W-Z\|_F$ is very small.

But in the case where $Z=\begin{pmatrix} A \\ B \end{pmatrix}$ is acutely perturbed, we can give an upper bound of $d_L(Z_1, W_1)$ with the aid of $W-Z$, and so we can see the dependence of the variation of $\sigma\{A, B\}$ on the perturbations in the elements of A and B.

Let $Z, W \in \mathbb{C}^{s\times t}$. We shall say that W is an acute perturbation of Z if

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

Stewart[6] has proved that W is an acute perturbation of Z if and only if

$$\operatorname{rank}(Z)=\operatorname{rank}(W)=\operatorname{rank}\,(P_ZWP_{Z^H}). \tag{2.8}$$

Under the presupposition of acute perturbation we obtain the following theorem which and (2.7) can be combined to give an upper bound of $d_L(Z_1, W_1)$.

Theorem 2.2. *Let $Z, W\in\mathbb{C}^{(m+p)\times n}$. If W is an acute perturbation of Z, then*

$$d_F(Z, W)\leqslant\omega(W)\xi(W-Z), \tag{2.9}$$

where

$$\omega(W)=\|(P_{Z^H}W^HWP_Z)^\dagger\|_2^{1/2}, \tag{2.10}$$

$$\xi(W-Z)=\|(I-P_Z)(W-Z)P_{Z^H}\|_F. \tag{2.11}$$

Proof. Suppose that $\operatorname{rank}(Z)=k$. Decomposing

$$Z=U\begin{pmatrix}Z_{11} & 0\\ 0 & 0\end{pmatrix}V^H,\ W=U\begin{pmatrix}W_{11} & W_{12}\\ W_{21} & W_{22}\end{pmatrix}V^H, \tag{2.12}$$

where U and V are unitary, $Z_{11}, W_{11}\in\mathbb{C}^{k\times k}$, rank $(Z_{11})=k$. Then it follows from (2.8) that $\operatorname{rank}(W)=\operatorname{rank}(W_{11})=k$. After some calculations we get (Ref. [5], 651—652)

$$(P_Z-P_W)^2=U\operatorname{diag}\ (X^HX(I+X^HX)^{-1},\ X(I+X^HX)^{-1}X^H)U^H,$$

where $X=W_{21}W_{11}^{-1}$. Therefrom

$$d_F^2(Z, W)=\operatorname{tr}[X^HX(I+X^HX)^{-1}]\leqslant\|(W_{11}^HW_{11}+W_{21}^HW_{21})^{-1}\|_2\|W_{21}\|_F^2. \tag{2.13}$$

Observe that

$$P_{Z^H}=V\begin{pmatrix}I^{(k)} & 0\\ 0 & 0\end{pmatrix}V^H,\ I-P_Z=U\begin{pmatrix}0 & 0\\ 0 & I^{(m+p-k)}\end{pmatrix}U^H,$$

$$(P_{Z^H}W^HWP_{Z^H})^\dagger=V\begin{pmatrix}(W_{11}^HW_{11}+W_{21}^HW_{21})^{-1} & 0\\ 0 & 0\end{pmatrix}V^H$$

and

$$(I-P_Z)(W-Z)P_{Z^H}=U\begin{pmatrix}0 & 0\\ W_{21} & 0\end{pmatrix}V^H,$$

then from (2.13) we obtain (2.9).

We note that if W is an acute perturbation of Z as W approaches Z, then

$$\omega(W)=\|(W_{11}^HW_{11}+W_{21}^HW_{21})^{-1}\|_2^{1/2}=O(1),\ \xi(W-Z)\to 0. \tag{2.14}$$

Hence, (2.3), (2.7), (2.9)—(2.11) and (2.14) show that the singular values of any $\{A, B\}\in\mathbf{P}(m, p, n; k)$ are insensitive to perturbations in the elements of A and B if $\begin{pmatrix}A\\B\end{pmatrix}$ is acute perturbed.

§ 3. The Weyl Type Theorem

In this section we develop a uniform upper bounds for differences of corresponding singular values of two $(m, p, n; k)-MP$.

For each point $\alpha+i\beta\in\mathbb{C}$ satisfying $\alpha, \beta\geqslant 0$ and $(\alpha, \beta)\neq(0, 0)$ we define $\theta(\alpha, \beta)$ as the angle subtended by the negative real axis $\{t:t\leqslant 0\}$ and $\{t(\alpha+i\beta):t\geqslant 0\}$ measured clockwise.

Let $\{A, B\}\in\mathbb{P}(m, p, n; k)$. We take a full-rank factorization of $Z=\begin{pmatrix}A\\B\end{pmatrix}$: $Z=Z_1F_1^H$, $Z_1=\begin{pmatrix}A_1\\B_1\end{pmatrix}_p^m$. By Theorem 1.3, $\sigma\{A, B\}=\sigma\{A_1, B_1\}$. If $(\alpha_i, \beta_i)\in\sigma\{A_1, B_1\}$, then from Definition 1.2, $(\alpha_i^2, \beta_i^2)\in\lambda(A_1^HA_1, B_1^HB_1)$. Since $(A_1^HA_1, B_1^HB_1)$ is a definite matrix-pair (Ref. [6]), there exists a non-zero $x_i\in\mathbb{C}^k$ such that

$$\beta_i^2A_1^HA_1x_i=\alpha_i^2B_1^HB_1x_i,\ (A_1x_i, B_1x_i)\neq(0, 0).$$

Now we define the singular angles

$$\theta_i=\theta(\|A_1x_i\|_2, \|B_1x_i\|_2),\ 1\leqslant i\leqslant k, \tag{3.1}$$

and assume that $\theta_1\leqslant\cdots\leqslant\theta_k$. From $\beta_i\|A_1x_i\|_2=\alpha_i\|B_1x_i\|_2$ there are $t_i>0$ for $1\leqslant i\leqslant k$ such that

$$(\|A_1x_i\|_2, \|B_1x_i\|_2)=(t_i\alpha_i, t_i\beta_i),\ \text{i. e. } \theta_i=\theta(\alpha_i, \beta_i);$$

hence every singular angle defined by (3.1) is uniquely determined, i. e. the singular angles θ_i $(1\leqslant i\leqslant k)$ are independent of the selection of the full-rank factorization of $\begin{pmatrix}A\\B\end{pmatrix}$.

The following lemma about unitary-invariant norm is useful for our discussion. Let $\|\ \|$ be a unitary-invariant norm on $\mathbb{C}^{n\times n}$. For any $A\in\mathbb{C}^{n\times q}(q<n)$ we define

$$\|A\|=\|\hat{A}\|,\ \hat{A}=(\underset{q}{A}, \underset{n-q}{0}). \tag{3.2}$$

Lemma 3.1. *Suppose that* $X\in\mathbb{C}^{n\times k}$, $Y\in\mathbb{C}^{n\times l}$, $k, l\leqslant n$. *If*

$$XX^H\leqslant YY^H, \tag{3.3}$$

then

$$\|X\|\leqslant\|Y\| \tag{3.4}$$

for every unitary-invariant norm $\|\ \|$ *on* $\mathbb{C}^{n\times n}$.

Proof. Let $\hat{X}=(X, 0)$, $\hat{Y}=(Y, 0)\in\mathbb{C}^{n\times n}$. From (3.3), $\hat{X}\hat{X}^H\leqslant\hat{Y}\hat{Y}^H$. Suppose that the singular values of $\hat{X}$ and $\hat{Y}$ are $0\leqslant\sigma_1\leqslant\cdots\leqslant\sigma_n$ and $0\leqslant\tau_1\leqslant\cdots\leqslant\tau_n$, respectively, then by the minimax theorem for Hermitian matrices we can deduce that $\sigma_i\leqslant\tau_i$ for $i=1, \cdots, n$. Hence utilizing Lemma 1 of [4], we have $\|\hat{X}\|\leqslant\|\hat{Y}\|$. Combining with (3.2) we get (3.4).

Theorem 3.1. *Suppose that* $\{A, B\}$, $\{C, D\}\in\mathbb{P}(m, p, n; k)$, $\sigma\{A, B\}=\{(\alpha_i, \beta_i)\}_{i=1}^k$, $\sigma\{C, D\}=\{(\gamma_i, \delta_i)\}_{i=1}^k$, *and the corresponding singular angles are* $\theta_1\leqslant\cdots\leqslant\theta_k$ *and* $\varphi_1\leqslant\cdots\leqslant\varphi_k$, *respectively. Then for any full-rank factorizations of* $Z=\begin{pmatrix}A\\B\end{pmatrix}$ *and* $W=\begin{pmatrix}C\\D\end{pmatrix}$:

$$Z=Z_1F_1^H,\ W=W_1G_1^H,\ Z_1=\begin{pmatrix}A_1\\B_1\end{pmatrix}_p^m,\ W_1=\begin{pmatrix}C_1\\D_1\end{pmatrix}_p^m, \tag{3.5}$$

we have

$$\rho_{i,i}\leqslant r(\{A_1, B_1\}, \{C_1, D_1\}),\ 1\leqslant i\leqslant k, \tag{3.6}$$

where

$$\rho_{i,i}=\rho((\alpha_i, \beta_i), (\gamma_i, \delta_i)),\ 1\leqslant i\leqslant k$$

and

$$r(\{A_1, B_1\}, \{C_1, D_1\}) \equiv \max_{\|x\|_2=1}\{\rho((\|A_1x\|_2, \|B_1x\|_2), (\|C_1x\|_2, \|D_1x\|_2))\}. \tag{3.7}$$

Moreover, if W is an acute perturbation of Z, then there exist full-rank factorizations (3.5) of Z and W such that Z_1 and W_1 satisfy

$$\|W_1 - Z_1\| \leqslant \|W - Z\| \tag{3.8}$$

for every unitary-invariant norm $\|\ \ \|$ on $\mathbb{C}^{(m+p)\times(m+p)}$.

Proof. From the assumptions, $\{A_1, B_1\}, \{C_1, D_1\} \in \mathbb{P}(m, p, k; k)$, $\sigma\{A_1, B_1\} = \sigma\{A, B\}$ and $\sigma\{C_1, D_1\} = \sigma\{C, D\}$. Hence by Theorem 2 of [7] we get the inequality (3.6) at once.

Now suppose that W is an acute perturbation of Z. Decomposing Z and W as in (2.12), then we know that rank $(W_{11}) = k$, and so we may take

$$Z_1 = U\begin{pmatrix} Z_{11} \\ 0 \end{pmatrix}, \ F_1^H = (I, 0)V^H, \ W_1 = U\begin{pmatrix} W_{11} \\ W_{21} \end{pmatrix}, \ G_1^H = (I, Q_1)V^H \tag{3.9}$$

in (3.5), where $Q_1 \in \mathbb{C}^{k\times(n-k)}$. By Lemma 3.1, the inequality holds for every unitary-invariant norm on $\mathbb{C}^{(m+p)\times(m+p)}$.

Observe that

$$r(\{A_1, B_1\}, \{C_1, D_1\}) \leqslant \max_{\|x\|_2=1}\left\{\sqrt{\frac{\|(C_1-A_1)x\|_2^2 + \|(D_1-B_1)x\|_2^2}{\|A_1x\|_2^2 + \|B_1x\|_2^2}}\right\}$$
$$\leqslant \frac{\sigma_{\max}(W_1 - Z_1)}{\sigma_{\min}(Z_1)}, \tag{3.10}$$

hence, if W is an acute perturbation of Z and we take the full-rank factorizations (3.9), then from (3.10),

$$r(\{A_1, B_1\}, \{C_1, D_1\}) \leqslant \frac{\sigma_{\max}(W-Z)}{\sigma_{\min}^{+}(Z)}.$$

Therefore according to Theorem 3.1 we obtain a weaker but more intuitive result.

Theorem 3.2. *Assume the hypotheses of Theorem 3.1. If $W = \begin{pmatrix} C \\ D \end{pmatrix}$ is an acute perturbation of $Z = \begin{pmatrix} A \\ B \end{pmatrix}$, then*

$$\rho_{i,i} \leqslant \frac{\sigma_{\max}(W-Z)}{\sigma_{\min}^{+}(Z)}, \ 1 \leqslant i \leqslant k.$$

Theorem 3.1 shows further that the singular values of any $\{A, B\} \in \mathbb{P}(m, p, n; k)$ are insensitive to perturbations in the elements of A and B if $\begin{pmatrix} A \\ B \end{pmatrix}$ is acutely perturbed.

§ 4. Final Remark

Finally, we note that in the case $k < n$ the singular values of $\{A, B\} \in \mathbb{P}(m, p, n; k)$ are not certain to be insensitive to perturbations in the elements of A and B. Consider for example a $(1; 1, 2; 1)-MP$ with

$$A = (2, 0), \ B = (-1, 0).$$

We have

$$A^HA=\begin{pmatrix}4 & 0\\ 0 & 0\end{pmatrix},\quad B^HB=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix},$$

and thus

$$\lambda(A^HA,\ B^HB)=\{(4,\ 1)\},\ \sigma\{A,\ B\}=\{(2,\ 1)\}.$$

Consider the neighboring problem with

$$\widetilde{A}=(2,\ \varepsilon),\ \widetilde{B}=(-1+\eta,\ 0),$$

where $\varepsilon\eta\neq 0$ and $|\varepsilon|,\ |\eta|\ll 1$. We have

$$\widetilde{A}^H\widetilde{A}=\begin{pmatrix}4 & 2\varepsilon\\ 2\varepsilon & \varepsilon^2\end{pmatrix},\quad \widetilde{B}^H\widetilde{B}=\begin{pmatrix}(1-\eta)^2 & 0\\ 0 & 0\end{pmatrix},$$

and thus

$$\lambda(\widetilde{A}^H\widetilde{A},\ \widetilde{B}^H\widetilde{B})=\{(0,\ 1),\ (1,\ 0)\},\ \sigma\{\widetilde{A},\ \widetilde{B}\}=\{(0,\ 1),\ (1,\ 0)\}.$$

This means, somewhat disappointingly, that in the case $k<n$ the singular values of $\{A,\ B\}\in\mathbb{P}(m,\ p,\ n;\ k)$ may be considerably changed by a small perturbation in the elements of A and B.

References

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

[2] Q. K. Lu, The elliptic geometry of extended spaces, *Acta Math. Sinica*, 13 (1963), 49—62.

[3] L. Mirsky, Symmetric gauge functions and unitarily invariant norms, *Quart. J. Math. Oxford*, 11 (1960), 50—59.

[4] C. C. Paige, M. A. Saunders, Towards a generalized singular value decomposition, *SIAM J. Numer. Anal.*, 18 (1981), 398—405.

[5] G. W. Stewart, On the perturbation of pseudo-inverses, projections and linear least squares problems, *SIAM Rev.*, 19 (1977), 634—662.

[6] G. W. Stewart, Perturbation bounds for the definite generalized eigenvalue problem, *Linear Algebra and Appl.*, 23 (1979), 69—83.

[7] J. G. Sun, On the perturbation of generalized singular values, *Math. Numer. Sinica*, 4 (1982), 229—233.

[8] J. G. Sun, Perturbation analysis for the generalized singular value problem, *SIAM J. Numer. Anal.* (to appear)

[9] J. G. Sun, Orthogonal projections and the perturbation of the eigenvalues of singular pencils, *J. Comp. Math.*, 1:1 (1983), 63—74.

[10] Charles F. Van Loan, Generalizing the singular value decomposition, *SIAM J. Numer. Anal.*, 13 (1976), 76—83.