Vol. 6 No. 4 JOURNAL OF COMPUTATIONAL MATHEMATICS October 1988

SENSITIVITY ANALYSIS OF ZERO SINGULAR VALUES AND MULTIPLE SINGULAR VALUES*1)

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

Abstract

Some results of the author[3,4] are used to discuss the sensitivity of zero singular values and multiple singular values of a real matrix analytically dependent on several parameters.

§ 1. Preliminaries

Let $\mathbb{R}^{m\times n}$ denote the set of real $m\times n$ matrices, $\mathbb{R}^n=\mathbb{R}^{n\times 1}$ and

$$S\mathbb{R}^{n\times n}=\{A\in\mathbb{R}^{n\times n}:\ A^T=A\},$$

in which the superscript T is for transpose. The singular values of a matrix $A\in\mathbb{R}^{m\times n}$ are denoted by $\sigma_1(A),\ \cdots,\ \sigma_n(A)$, the eigenvalues of a matrix $A\in\mathbb{R}^{n\times n}$ are denoted by $\lambda_1(A),\ \cdots,\ \lambda_n(A)$, and

$$\sigma(A)=\{\sigma_j(A)\}_{j=1}^n,\quad \lambda(A)=\{\lambda_j(A)\}_{j=1}^n.$$

The symbol $\|\ \|_2$ denotes the usual Euclidean vector norm and the spectral norm.

Let $p=(p_1,\ \cdots,\ p_N)^T\in\mathbb{R}^N$. Suppose that $A(p)\in\mathbb{R}^{m\times n}$ is a real matrix-valued analytic function in some open set $\mathscr{S}\subset\mathbb{R}^N$. It is well known that for any point $p^*\in\mathscr{S}$, if $\sigma^*>0$ is a simple singular value of $A(p^*)$, then there is a real analytic function $\sigma(p)>0$ defined in some neighbourhood $\mathscr{B}(p^*)\subset\mathscr{S}$ of p^* such that $\sigma(p^*)=\sigma^*$, $\sigma(p)\in\sigma(A(p))\ \ \forall p\in\mathscr{B}(p^*)$, and one can obtain perturbation expansions of $\sigma(p)$ at the point p^* (see [1], [5]). In such a case we may define the sensitivity of the singular value σ^* with respect to the parameter p_j by

$$s_{p_j}(\sigma^*)=\left|\left(\frac{\partial\sigma(p)}{\partial p_j}\right)_{p=p^*}\right|.$$

But if σ^* is a zero singular value or a multiple singular value of $A(p^*)$, then, in general, there is no real differentiable function $\sigma(p)\geqslant 0$ defined in some neighbourhood $\mathscr{B}(p^*)\subset\mathscr{S}$ of p^* satisfying $\sigma(p^*)=\sigma^*$ such that $\sigma(p)\in\sigma(A(p))$ $\forall p\in\mathscr{B}(p^*)$. This paper will discuss the sensitivity of σ^* with respect to p_j in these cases.

Without loss of generality we may assume that the point p^* is the origin of $\mathbb{R}^N$ and $p^*\in\mathscr{S}$. For any real-valued function $f(p)$ defined in $\mathscr{S}$, we shall denote the right and left partial derivatives of $f(p)$ with respect to p_j at the origin, respectively, by

* Received December, 3, 1986.
1) The Project Supported by National Natural Scienc Foundation of China.

$$\left(\frac{\partial f(p)}{\partial p_j}\right)_{p=0,\, p_j=+0}=\lim_{p_j\to+0}\frac{f(0,\,\cdots,\,0,\,p_j,\,0,\,\cdots,\,0)-f(0,\,\cdots,\,0)}{p_j}$$

and

$$\left(\frac{\partial f(p)}{\partial p_j}\right)_{p=0,\, p_j=-0}=\lim_{p_j\to-0}\frac{f(0,\,\cdots,\,0,\,p_j,\,0,\,\cdots,\,0)-f(0,\,\cdots,\,0)}{p_j},\ j=1,\,\cdots,\,N.$$

Now we cite a result about partial derivatives of eigenvalues of a real symmetric matrix (see [3, Theorems 2.3 and 2.4], and [4, Theorem 2.1]), on the basis of which we may discuss the sensitivity of zero singular values and multiple singular values.

Theorem 1.1. *Let* $p=(p_1,\,\cdots,\,p_N)^T\in\mathbb{R}^N$, *and let* $A(p)\in S\mathbb{R}^{n\times n}$ *be a real analytic function in some neighbourhood* $\mathscr{B}(0)\subset\mathbb{R}^N$ *of the origin. Suppose that* λ_1 *is an eigenvalue of* $A(0)$ *with multiplicity* $r\geqslant 1$, *i.e., there is a real orthogonal matrix* $X\in\mathbb{R}^{n\times n}$ *such that*

$$X=(\underset{r}{X_1},\ \underset{n-r}{X_2}),\quad X^TA(0)X=\begin{pmatrix}\lambda_1 I^{(r)} & 0\\ 0 & A_2\end{pmatrix},\quad \lambda_1\overline{\in}\lambda(A_2) \tag{1.1}$$

(in the case of $r=1$, *we rewrite* $X_1=x_1$*). Then*

(i) *if* $r=1$, *there is a real analytic function* $\lambda_1(p)$ *defined in some neighbourhood* $\mathscr{B}_0\subset\mathscr{B}(0)$ *of the origin such that* $\lambda_1(0)=\lambda_1$, $\lambda_1(p)\in\lambda(A(p))$, $\forall p\in\mathscr{B}_0$, *and one has*

$$\left(\frac{\partial\lambda_1(p)}{\partial p_j}\right)_{p=0}=x_1^T\left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0}x_1,\quad j=1,\cdots,N \tag{1.2}$$

and

$$\begin{aligned}\left(\frac{\partial^2\lambda_1(p)}{\partial p_j\partial p_k}\right)_{p=0}&=x_1^T\left(\frac{\partial^2 A(p)}{\partial p_j\partial p_k}\right)_{p=0}x_1\\ &\quad+2x_1^T\left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0}X_2(\lambda_1 I-A_2)^{-1}X_2^T\left(\frac{\partial A(p)}{\partial p_k}\right)_{p=0}x_1,\\ &\qquad j,\,k=1,\,\cdots,\,N;\end{aligned} \tag{1.3}$$

(ii) *if* $r>1$, *there are real-valued functions* $\lambda_1(p),\,\cdots,\,\lambda_r(p)$ *defined in some neighbourhood* $\mathscr{B}_0\subset\mathscr{B}(0)$ *of the origin and two permutations* π *and* π' *of* $1,\,\cdots,\,r$ *such that*

$$\lambda_1(0)=\cdots=\lambda_r(0)=\lambda_1,\ \lambda_1(p),\,\cdots,\,\lambda_r(p)\in\lambda(A(p))\quad\forall p\in\mathscr{B}_0,$$

and one has

$$\left(\frac{\partial\lambda_s(p)}{\partial p_j}\right)_{p=0,\,p_j=+0}=\lambda_{\pi(s)}\left(X_1^T\left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0}X_1\right) \tag{1.4}$$

and

$$\left(\frac{\partial\lambda_s(p)}{\partial p_j}\right)_{p=0,\,p_j=-0}=\lambda_{\pi'(s)}\left(X_1^T\left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0}X_1\right),\ s=1,\,\cdots,\,r,\ j=1,\,\cdots,\,N. \tag{1.5}$$

We shall prove some formulas for partial derivatives in Section 2 and give numerical examples in Section 3.

§ 2. Some Formulas for Partial Derivatives

Theorem 2.1. *Let $p=(p_1, \cdots, p_N)^T\in\mathbb{R}^N$, and let $A(p)\in\mathbb{R}^{m\times n}(m\geqslant n)$ be a real analytic function in some neighbourhood $\mathscr{B}(0)\subset\mathbb{R}^N$ of the origin. Suppose that $\sigma_1=0$ is a singular value of $A(0)$ with multiplicity $r\geqslant 1$, i.e., there are real orthogonal matrices $U\in\mathbb{R}^{m\times m}$ and $V\in\mathbb{R}^{n\times n}$ such that*

$$U^TA(0)V=\Sigma, \tag{2.1}$$

in which

$$U=(\underset{r}{U_1}, \underset{n-r}{U_2}, \underset{m-n}{U_3}),\ V=(\underset{r}{V_1}, \underset{n-r}{V_2}) \tag{2.2}$$

and

$$\Sigma=\begin{pmatrix} 0^{(r)} & 0 \\ 0 & \Sigma_2 \end{pmatrix},\ \Sigma_2=\begin{pmatrix} \Sigma_{21} \\ 0 \end{pmatrix},\ \Sigma_{21}=\operatorname{diag}(\sigma_{r+1}, \cdots, \sigma_n),$$

$$0<\sigma_{r+1}, \cdots, \sigma_n. \tag{2.3}$$

Then there exist nonnegative functions $\sigma_1(p),\cdots, \sigma_r(p)$ defined in some neighbourhood $\mathscr{B}_0\subset\mathscr{B}(0)$ of the origin and two permutations π and π' of $1, \cdots, r$ such that

$$\sigma_1(0)=\cdots=\sigma_r(0)=0,\ \sigma_1(p),\ \cdots,\ \sigma_r(p)\in\sigma(A(p))\quad \forall p\in\mathscr{B}_0, \tag{2.4}$$

and one has

$$\left(\frac{\partial\sigma_s(p)}{\partial p_j}\right)_{p=0,\ p_j=+0}=\sigma_{\pi(s)}\left((U_1,\ U_3)^T\left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0}V_1\right) \tag{2.5}$$

and

$$\left(\frac{\partial\sigma_s(p)}{\partial p_j}\right)_{p=0,\ p_j=-0}=-\sigma_{\pi'(s)}\left((U_1,\ U_3)^T\left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0}V_1\right),$$

$$s=1,\ \cdots,\ r,\ j=1,\ \cdots,\ N. \tag{2.6}$$

Proof. Let

$$T(p)=\begin{pmatrix} 0 & A(p) \\ A(p)^T & 0 \end{pmatrix}. \tag{2.7}$$

Obviously, $T(p)\in S\mathbb{R}^{(m+n)\times(m+n)}$ is a real analytic function in $\mathscr{B}(0)$. Let

$$W=\begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix},\quad T_1=W^TT(0)W. \tag{2.8}$$

Then from (2.1) and (2.2)

$$T_1=\begin{array}{c} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \Sigma_{21} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & \Sigma_{21} & 0 & 0 & 0 \end{pmatrix} \begin{array}{l} r \\ n-r \\ m-n \\ r \\ n-r \end{array} \\ \begin{array}{ccccc} r & n-r & m-n & r & n-r \end{array} \end{array}. \tag{2.9}$$

Moreover, let

$$Q_0=\begin{array}{cc}\begin{pmatrix} I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}}I & \frac{1}{\sqrt{2}}I \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{\sqrt{2}}I & \frac{1}{\sqrt{2}}I \end{pmatrix} & \begin{matrix} r \\ n-r \\ m-n, \\ r \\ n-r \end{matrix} \\ \begin{matrix} r & m-n & r & n-r & n-r \end{matrix} & \end{array} \quad T_0=Q_0^T T_1 Q_0. \tag{2.10}$$

Combining (2.8)—(2.10), and writing

$$X=WQ_0=\left(\begin{array}{ccc|cc} U_1 & U_3 & 0 & \frac{1}{\sqrt{2}}U_2 & \frac{1}{\sqrt{2}}U_2 \\ 0 & 0 & V_1 & -\frac{1}{\sqrt{2}}V_2 & \frac{1}{\sqrt{2}}V_2 \end{array}\right)=(\underset{2r+m-n}{X_1} \mid \underset{2n-2r}{X_2}), \tag{2.11}$$

we get

$$T_0=X^TT(0)X=\mathrm{diag}(\underset{r}{0},\ \underset{m-n}{0},\ \underset{r}{0},\ \underset{n-r}{\Sigma_{21}},\ \underset{n-r}{-\Sigma_{21}}),\quad X^TX=I. \tag{2.12}$$

Observe the following facts:

(i) If $0\leqslant\sigma(p)\in\lambda(T(p))$, then $-\sigma(p)\in\lambda(T(p))$ and $\sigma(p)\in\sigma(A(p))$; conversely, if $\sigma(p)\in\sigma(A(p))$, then $\sigma(p),\ -\sigma(p)\in\lambda(T(p))$.

(ii) There are $m-n$ functions $\tau_1(p), \cdots, \tau_{m-n}(p)$ such that

$$\tau_1(p)=\cdots=\tau_{m-n}(p)=0,\ \tau_1(p), \cdots, \tau_{m-n}(p)\in\lambda(T(p)) \quad \forall p\in\mathscr{B}(0).$$

(iii) From

$$X_1^T\left(\frac{\partial T(p)}{\partial p_j}\right)_{p=0}X_1=\begin{pmatrix} 0 & (U_1, U_3)^T\left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0}V_1 \\ V_1^T\left(\frac{\partial A(p)}{\partial p_j}\right)^T_{p=0}(U_1, U_3) & 0 \end{pmatrix}$$

we get

$$\lambda\left(X_1^T\left(\frac{\partial T(p)}{\partial p_j}\right)_{p=0}X_1\right)$$
$$=\left\{\pm\gamma_1, \cdots, \pm\gamma_r, \underbrace{0, \cdots, 0}_{m-n}: \gamma_j\in\sigma\left((U_1, U_3)^T\left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0}V_1\right),\ j=1, \cdots, r\right\}. \tag{2.13}$$

Hence, by Theorem 1.1 there are nonnegative functions $\sigma_1(p), \cdots, \sigma_r(p)$ defined in some neighbourhood $\mathscr{B}_0\subset\mathscr{B}(0)$ of the origin and two permutations π and π' of $1, \cdots, r$ such that the relations (2.4)—(2.6) are valid. ▌

From Theorem 2.1 we obtain the following corollaries.

Corollary 2.1. Under the hypotheses of Theorem 2.1, if $r=1$ (i.e., $\sigma_1=0$ is a simple singular value of $A(0)$) and $U_1=u_1$, $V_1=v_1$, then there is a nonnegative function $\sigma_1(p)$ defined in some neighbourhood $\mathscr{B}_0\subset\mathscr{B}(0)$ of the origin such that

$$\sigma_1(0)=0,\ \sigma_1(p)\in\sigma(A(p)) \quad \forall p\in\mathscr{B}_0,$$

and one has

$$\left(\frac{\partial\sigma_1(p)}{\partial p_j}\right)_{p=0,\, p_j=+0}=\left\|(u_1,\ U_3)^T\left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0} v_1\right\|_2 \tag{2.14}$$

and

$$\left(\frac{\partial\sigma_1(p)}{\partial p_j}\right)_{p=0,\, p_j=-0}=-\left\|(u_1,\ U_3)^T\left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0} v_1\right\|_2,\quad j=1,\ \cdots,\ N. \tag{2.15}$$

Corollary 2.2. Let $A=(\alpha_{ij})\in\mathbb{R}^{m\times n}$, in which $\alpha_{ij}(1\leqslant i\leqslant m,\ 1\leqslant j\leqslant n)$ are regarded as parameters. Suppose that there exist real orthogonal matrices $U\in\mathbb{R}^{m\times m}$, $V\in\mathbb{R}^{n\times n}$ such that $U^TAV=\Sigma$, where U, V and Σ are represented by (2.2) and (2.3), and

$$(U_1,\ U_3)=(\hat{u}_1,\ \cdots,\ \hat{u}_m)^T,\ V_1=(\hat{v}_1,\ \cdots,\ \hat{v}_n)^T,\ \hat{u}_i\in\mathbb{R}^{m-n+r},\ \forall i,\ \hat{v}_j\in\mathbb{R}^r\ \forall j. \tag{2.16}$$

Then the right and left partial derivatives of the r zero singular values with respect to α_{ij} are respectively

$$\underbrace{0,\ \cdots,\ 0}_{r-1},\ \|\hat{u}_i\|_2\|\hat{v}_j\|_2,\quad i=1,\ \cdots,m,\ j=1,\ \cdots,\ n, \tag{2.17}$$

and

$$\underbrace{0,\ \cdots,\ 0}_{r-1},\ -\|\hat{u}_i\|_2\|\hat{v}_j\|_2,\quad i=1,\ \cdots,\ m,\ j=1,\ \cdots,\ n. \tag{2.18}$$

Proof. It is sufficient to point out that we have

$$(U_1,\ U_3)^T\frac{\partial A}{\partial\alpha_{ij}}V_1=\hat{u}_i\hat{v}_j^T$$

and the singular values of $\hat{u}_i\hat{v}_j^T$ are zero with multiplicity $r-1$ and $\|\hat{u}_i\|_2\|\hat{v}_j\|_2$. ∎

By Theorem 2.1 we may introduce the following definition.

Definition 2.1. *Under the hypotheses of Theorem* 2.1, *the quantity*

$$s_{p_j}(0)=\left\|(U_1,\ U_3)^T\left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0} V_1\right\|_2 \tag{2.19}$$

is called the sensitivity of the zero singular values of $A(p)$ at $p=0$ with respect to p_j.

According to Definition 2.1 we see that for the matrix A described in Corollary 2.2 we have

$$s_{\alpha_{ij}}(0)=\|\hat{u}_i\|_2\|\hat{v}_j\|_2,\quad i=1,\ \cdots,\ m,\ j=1,\ \cdots,\ n. \tag{2.20}$$

Theorem 2.2. *Let p, $A(p)$ and $\mathscr{B}(0)$ be as in Theorem* 2.1. *Suppose that $\sigma_1>0$ is a multiple singular value with multiplicity $r>1$, i.e., there are real orthogonal matrices U, V represented by* (2.2) *such that*

$$U^TA(0)V=\Sigma, \tag{2.21}$$

in which

$$\Sigma=\begin{pmatrix}\sigma_1I^{(r)} & 0\\ 0 & \Sigma_2\end{pmatrix},\ \Sigma_2=\begin{pmatrix}\Sigma_{21}\\ 0\end{pmatrix},\ \Sigma_{21}=\mathrm{diag}(\sigma_{r+1},\ \cdots,\ \sigma_n),$$

$$\sigma_1\neq\sigma_j\geqslant 0,\ j=r+1,\ \cdots,\ n. \tag{2.22}$$

Then there are nonnegative functions $\sigma_1(p)$, $\cdots$, $\sigma_r(p)$ defined in some neighbourhood $\mathscr{B}_0\subset\mathscr{B}(0)$ of the origin and two permutations π and π' of 1, $\cdots$, r *such that*

$$\sigma_1(0)=\cdots=\sigma_r(0)=\sigma_1,\ \sigma_1(p),\ \cdots,\ \sigma_r(p)\in\sigma(A(p))\quad\forall p\in\mathscr{B}_0,$$

and one has

$$\left|\left(\frac{\partial\sigma_s(p)}{\partial p_j}\right)_{p=0,\,p_j=+0}\right|=\sigma_{\pi(s)}\left(U_1^T\left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0}V_1\right) \tag{2.23}$$

and

$$\left|\left(\frac{\partial\sigma_s(p)}{\partial p_j}\right)_{p=0,\,p_j=-0}\right|=\sigma_{\pi'(s)}\left(U_1^T\left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0}V_1\right),\quad s=1,\ \cdots,\ r,\ j=1,\ \cdots,\ N. \tag{2.24}$$

Proof. Let $T(p)$, W and T_1 be defined by (2.7) and (2.8). Then from (2.21) and (2.22)

$$T_1=\begin{array}{cc}\begin{pmatrix}0&0&0&\sigma_1I&0\\0&0&0&0&\Sigma_{21}\\0&0&0&0&0\\\sigma_1I&0&0&0&0\\0&\Sigma_{21}&0&0&0\end{pmatrix}&\begin{matrix}r\\n-r\\m-n.\\r\\n-r\end{matrix}\\\begin{matrix}r&n-r&m-n&r&n-r\end{matrix}&\end{array} \tag{2.25}$$

Moreover, let

$$Q=\begin{array}{cc}\begin{pmatrix}I&0&0&0&0\\0&0&0&\frac{1}{\sqrt{2}}I&\frac{1}{\sqrt{2}}I\\0&0&I&0&0\\0&I&0&0&0\\0&0&0&-\frac{1}{\sqrt{2}}I&\frac{1}{\sqrt{2}}I\end{pmatrix}&\begin{matrix}r\\n-r\\m-n,\\r\\n-r,\end{matrix}\\\begin{matrix}r&r&m-n&n-r&n-r\end{matrix}&\end{array}\quad T=Q^TT_1Q. \tag{2.26}$$

Combining (2.8), (2.25) and (2.26), and writing

$$X=WQ=\left(\begin{array}{cc|ccc}U_1&0&U_3&\frac{1}{\sqrt{2}}U_2&\frac{1}{\sqrt{2}}U_2\\&&&&\\0&V_1&0&-\frac{1}{\sqrt{2}}V_2&\frac{1}{\sqrt{2}}V_2\end{array}\right)=(\underset{2r}{X_1}\mid\underset{m+n-2r}{X_2}), \tag{2.27}$$

we get

$$T=X^TT(0)X=\mathrm{diag}(\underset{r}{\sigma_1I},\ \underset{r}{-\sigma_1I},\ \underset{m-n}{0},\ \underset{n-r}{\Sigma_{21}},\ \underset{n-r}{-\Sigma_{21}}),\quad X^TX=I. \tag{2.28}$$

Observe that if $\sigma(p)\geqslant 0$ and $\sigma(p)\in\lambda(T(p))$, then $-\sigma(p)\in\lambda(T(p))$ and $\sigma(p)\in\sigma(A(p))$; conversely, if $\sigma(p)\in\sigma(A(p))$, then $\sigma(p)$, $-\sigma(p)\in\lambda(T(p))$. Hence, utilizing Theorem 1.1 there are nonnegative functions $\sigma_1(p)$, $\cdots$, $\sigma_r(p)$ defined in some neighbourhood $\mathscr{B}_0\subset\mathscr{B}(0)$ of the origin such that

$$\sigma_1(0)=\cdots=\sigma_r(0)=\sigma_1,\ \pm\sigma_1(p),\ \cdots,\ \pm\sigma_r(p)\in\lambda(T(p))\quad\forall p\in\mathscr{B}_0,$$

and if we let

$$\mathfrak{M}_j^+=\left\{\pm\mu:\mu=\left(\frac{\partial\sigma_s(p)}{\partial p_j}\right)_{p=0,\,p_j=+0},\ s=1,\ \cdots,\ r\right\},$$

$$\mathfrak{M}_j^- = \left\{ \pm\mu:\ \mu = \left(\frac{\partial \sigma_s(p)}{\partial p_j}\right)_{p=0,\ p_j=-0},\ s=1,\ \cdots,\ r \right\}$$

and

$$\mathfrak{N}_j = \left\{ \nu: \nu \in \lambda\left(X_1^T \left(\frac{\partial T(p)}{\partial p_j}\right)_{p=0} X_1 \right) \right\},$$

then we have

$$\mathfrak{M}_j^+ = \mathfrak{M}_j^- = \mathfrak{N}_j,\quad j=1,\cdots,N.$$

Moreover, there are one-to-one correspondences between the elements of the sets $\mathfrak{M}_j^+$ and $\mathfrak{N}_j$ and between the elements of the sets $\mathfrak{M}_j^-$ and $\mathfrak{N}_j$. But from

$$X_1^T \left(\frac{\partial T(p)}{\partial p_j}\right)_{p=0} X_1 = \begin{pmatrix} 0 & U_1^T \left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0} V_1 \\ V_1^T \left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0}^T U_1 & 0 \end{pmatrix}$$

it follows that

$$\mathfrak{N}_j = \left\{ \pm\nu:\ \nu \in \sigma\left(U_1^T \left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0} V_1 \right) \right\},$$

therefrom we get the relations (2.23) and (2.24). ∎

Consequently, by Theorem 2.2 we may introduce the following definition.

Definition 2.2. *Under the hypotheses of Theorem 2.2, the quantity*

$$s_{p_j}(\sigma_1) = \left\| U_1^T \left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0} V_1 \right\|_2 \tag{2.29}$$

is called the sensitivity of the multiple singular value $\sigma_1>0$ of $A(p)$ at $p=0$ with respect to p_j.

Now we use Theorem 1.1 to deduce a second order Taylor expansion of the simple zero singular value of a matrix $A\in\mathbb{R}^{m\times n}$ which has been obtained by Stewart[2] in a different way.

Theorem 2.3. *Assume that $\sigma_1=0$ is a simple singular value of a matrix $A\in\mathbb{R}^{m\times n}$ $(m\geqslant n)$, i.e., there are real orthogonal matrices $U\in\mathbb{R}^{m\times m}$, $V\in\mathbb{R}^{n\times n}$ such that*

$$U^T A V = \Sigma, \tag{2.30}$$

in which

$$U = (\underset{1}{u_1},\ \underset{n-1}{U_2},\ \underset{m-n}{U_3}),\quad V = (\underset{1}{v_1},\ \underset{n-1}{V_2}),\quad \Sigma = \begin{pmatrix} 0 & 0 \\ 0 & \Sigma_2 \end{pmatrix} \tag{2.31}$$

and

$$\Sigma_2 = \begin{pmatrix} \Sigma_{21} \\ 0 \end{pmatrix},\ \Sigma_{21} = \operatorname{diag}(\sigma_2,\ \cdots,\ \sigma_n),\ 0<\sigma_2,\ \cdots,\ \sigma_n. \tag{2.32}$$

Let

$$s = (s_{11},\ \cdots,\ s_{1n},\ s_{21},\ \cdots,\ s_{2n},\ \cdots,\ s_{m1},\ \cdots,\ s_{mn})^T \in \mathbb{R}^{mn}$$

and

$$A(s) = A + E,\quad E = (s_{ij})_{1\leqslant i\leqslant m,\,1\leqslant j\leqslant n}.$$

Then there is a simple singular value $\sigma_1(s)$ of $A(s)$ in some neighbourhood $\mathscr{B}(0)$ of the origin of $\mathbb{R}^{mn}$ such that $\sigma_1(0)=0$ and the function $\sigma_1^2(s)$ has a second order Taylor *expansion*

$$\sigma_1(\varepsilon)=\|(u_1, U_3)^T E v_1\|_2^2+O(\|E\|_2^3), \quad \varepsilon\in\mathscr{B}(0). \tag{2.33}$$

Proof. Consider the matrix

$$S(\varepsilon)=A(\varepsilon)^T A(\varepsilon)=A^T A+E^T A+A^T E+E^T E. \tag{2.34}$$

By hypotheses the eigenvalues $\lambda_j(j=1, \cdots, n)$ of $S(0)$ satisfy

$$0=\lambda_1<\lambda_2, \cdots, \lambda_n, \quad \lambda_j=\sigma_j^2, \quad j=1, \cdots, n.$$

Hence, by Theorem 1.1 there is a real analytic simple eigenvalue $\lambda_1(\varepsilon)$ of $S(\varepsilon)$ in some neighbourhood $\mathscr{B}(0)\subset\mathbb{R}^{mn}$ of the origin satisfying $\lambda_1(0)=0$. Let $\sigma_1(\varepsilon)=\sqrt{\lambda_1(\varepsilon)}$. Obviously, $\sigma_1(\varepsilon)\in\sigma(A(\varepsilon))$ and $\sigma_1(0)=0$.

Let $e_i^{(n)}$ denote the i-th column of the identity $I^{(n)}$, and let

$$E_{jk}=e_j^{(m)}e_k^{(n)T}, \quad j=1, \cdots, m, \quad k=1, \cdots, n.$$

Then

$$E=\sum_{j=1}^{m}\sum_{k=1}^{n}\varepsilon_{jk}E_{jk}.$$

Utilizing the formulas (1.2), (1.3) and the relations (2.30)—(2.32) we obtain

$$\left(\frac{\partial\lambda_1(\varepsilon)}{\partial\varepsilon_{jk}}\right)_{\varepsilon=0}=v_1^T\left(\frac{\partial S(\varepsilon)}{\partial\varepsilon_{jk}}\right)_{\varepsilon=0}v_1=v_1^T(E_{jk}^T A+A^T E_{jk})v_1=0 \tag{2.35}$$

and

$$\left(\frac{\partial^2\lambda_1(\varepsilon)}{\partial\varepsilon_{jk}\partial\varepsilon_{st}}\right)_{\varepsilon=0}=v_1^T\left(\frac{\partial^2 S(\varepsilon)}{\partial\varepsilon_{jk}\partial\varepsilon_{st}}\right)_{\varepsilon=0}v_1-2v_1^T\left(\frac{\partial S(\varepsilon)}{\partial\varepsilon_{jk}}\right)_{\varepsilon=0}V_2\Sigma_{21}^{-2}V_2^T\left(\frac{\partial S(\varepsilon)}{\partial\varepsilon_{st}}\right)_{\varepsilon=0}v_1, \tag{2.36}$$

where

$$\left(\frac{\partial^2 S(\varepsilon)}{\partial\varepsilon_{jk}\partial\varepsilon_{st}}\right)_{\varepsilon=0}=E_{jk}^T E_{st}+E_{st}^T E_{jk} \tag{2.37}$$

and

$$\left(\frac{\partial S(\varepsilon)}{\partial\varepsilon_{jk}}\right)_{\varepsilon=0}=E_{jk}^T A+A^T E_{jk}, \quad \left(\frac{\partial S(\varepsilon)}{\partial\varepsilon_{st}}\right)_{\varepsilon=0}=E_{st}^T A+A^T E_{st}. \tag{2.38}$$

Substituting (2.35)—(2.38) into the Taylor expansion of $\lambda_1(\varepsilon)$ at $\varepsilon=0$ and utilizing

$$Av_1=0, \quad AV_2=U_2\Sigma_{21},$$

we get

$$\begin{aligned}\lambda_1(\varepsilon)&=v_1^T E^T E v_1-v_1^T E^T A V_2\Sigma_{21}^{-2}V_2^T A^T E v_1+O(\|\varepsilon\|_2^3)\\&=v_1^T E^T E v_1-v_1^T E^T U_2 U_2^T E v_1+O(\|\varepsilon\|_2^3)\\&=v_1^T E^T(u_1u_1^T+U_3U_3^T)Ev_1+O(\|\varepsilon\|_2^3)\\&=\|(v_1, U_3)^T E v_1\|_2^2+O(\|\varepsilon\|_2^3), \quad \varepsilon\in\mathscr{B}(0).\end{aligned}$$

Hence, the formula (2.33) is valid. ∎

§ 3. Numerical Examples

Now we give two examples to illustrate some of the above results.

Example 3. 1. $m=n=N=2$.

$$A(p)=\begin{pmatrix}1+p_1 & 1-3p_2\\ 1-3p_2 & 1+2p_1\end{pmatrix},\ p\in\mathscr{B}(0)=\left\{p\in\mathbb{R}^2:\ \|p\|_2<\frac{1}{3}\right\}.$$

Let

$$U=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}=(u_1,\ u_2).$$

Obviously, U is a real orthogonal matrix and we have

$$U^TA(0)U=\begin{pmatrix}0 & 0\\ 0 & 2\end{pmatrix}.$$

Hence $\sigma_1=0$ is a simple singular value of $A(0)$. By Theorem 2.1 there is a nonnegative function $\sigma_1(p)$ defined in some neighbourhood $\mathscr{B}_0\subset\mathbb{R}^2$ of the origin such that

$$\sigma_1(0)=0,\ \sigma_1(p)\in\sigma(A(p))\quad \forall p\in\mathscr{B}_0,$$

and we have

$$\left(\frac{\partial\sigma_1(p)}{\partial p_1}\right)_{p=0,\ p_1=+0}=\left|u_1^T\left(\frac{\partial A(p)}{\partial p_1}\right)_{p=0}u_1\right|=\frac{3}{2},\quad \left(\frac{\partial\sigma_1(p)}{\partial p_1}\right)_{p=0,\ p_1=-0}=-\frac{3}{2}\quad (3.1)$$

and

$$\left(\frac{\partial\sigma_1(p)}{\partial p_2}\right)_{p=0,\ p_2=+0}=\left|u_1^T\left(\frac{\partial A(p)}{\partial p_2}\right)_{p=0}u_1\right|=3,\quad \left(\frac{\partial\sigma_1(p)}{\partial p_2}\right)_{p=0,\ p_2=-0}=-\frac{3}{2}.\quad (3.2)$$

Consequently, the sensitivities of the simple zero singular value of $A(0)$ with respect to p_1 and p_2 are

$$s_{p_1}(0)=\frac{3}{2},\quad s_{p_2}(0)=3.$$

It is worth while to point out that the singular values of $A(p)$ are

$$\sigma_1(p)=\frac{|2+3p_1-\sqrt{p_1^2+4(1-3p_2)^2}\,|}{2},\quad \sigma_2=\frac{2+3p_1+\sqrt{p_1^2+4(1-3p_2)^2}}{2},$$
$$p\in\mathscr{B}(0).\quad (3.3)$$

From (3.3) we see that $\sigma_1(0)=0$, $\sigma_2(0)=2$ and

$$\sigma_1(p_1,\ 0)=\frac{|2+3p_1-\sqrt{p_1^2+4}\,|}{2},\ \sigma_1(0,\ p_2)=3|p_2|,\quad p\in\mathscr{B}(0).\quad (3.4)$$

Equalities (3.1) and (3.2) can also be obtained from (3.4).

Example 3. 2. [6, p.149]. $m=8$, $n=5$.

$$A=(a_{ij})_{1\le i\le 8,\ 1\le j\le 5}=\begin{pmatrix}22 & 10 & 2 & 3 & 7\\ 14 & 7 & 10 & 0 & 8\\ -1 & 13 & -1 & -11 & 3\\ -3 & -2 & 13 & -2 & 4\\ 9 & 8 & 1 & -2 & 4\\ 9 & 1 & -7 & 5 & -1\\ 2 & -6 & 6 & 5 & 1\\ 4 & 5 & 0 & -2 & 2\end{pmatrix}.\quad (3.5)$$

We have.

$$\sigma(A)=\{0,\ 0,\ 19.5958,\ 19.9999,\ 35.3270\}.$$

Utilizing the formula (2.20) we get the sensitivities $s_{\alpha_{ij}}(0)$ of the zero singular values of A with respect to α_{ij}, which are given in Table 1.

Table 1

$s_{\alpha_{ij}}(0)$ i \ j	1	2	3	4	5
1	0.2792	0.4048	0.2349	0.4788	0.6010
2	0.3161	0.4583	0.2660	0.5421	0.6865
3	0.2343	0.3397	0.1971	0.4018	0.5044
4	0.2888	0.4188	0.2430	0.4954	0.6219
5	0.3864	0.5602	0.3251	0.6627	0.8319
6	0.3411	0.4946	0.2870	0.5851	0.7344
7	0.3614	0.5240	0.3041	0.6199	0.7781
8	0.4071	0.5903	0.3426	0.6983	0.8765

From (3.5) and Table 1 we see that $\alpha_{13}=\alpha_{85}=2$, but $s_{\alpha_{13}}(0)=0.2349$, $s_{\alpha_{85}}(0)=0.8765$, and so $s_{\alpha_{85}}(0)/s_{\alpha_{13}}(0)\approx 3.73$. Let

$$A_{13}(\delta)=(\hat{\alpha}_{ij}(\delta)),\ \hat{\alpha}_{ij}(\delta)=\begin{cases}2+\delta, & i=1,\ j=3,\\ \alpha_{ij}, & \text{otherwise}\end{cases}$$

and

$$A_{85}(\delta)=(\tilde{\alpha}_{ij}(\delta)),\ \tilde{\alpha}_{ij}(\delta)=\begin{cases}2+\delta, & i=8,\ j=5,\\ \alpha_{ij}, & \text{otherwise.}\end{cases}$$

Taking different values of the parameter δ, we obtain the smallest two singular values σ_1, σ_2 of $A_{13}(\delta)$ and $A_{85}(\delta)$, which are given in Table 2.

Table 2

δ	-1.0	-0.5	-0.1	-0.05	-0.01	0.01	0.05	0.1	0.5	1.0
$\sigma_1(A_{13}(\delta))$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\sigma_2(A_{13}(\delta))$	0.2343	0.1173	0.0235	0.0117	0.0024	0.0024	0.0117	0.0235	0.1176	0.2353
$\sigma_1(A_{85}(\delta))$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\sigma_2(A_{85}(\delta))$	0.8778	0.4386	0.0877	0.0438	0.0088	0.0088	0.0438	0.0876	0.4379	0.8750

The results listed in Tables 1 and 2 were obtained on the L-340 computer.

References

[1] MacFarlane, A. G. J.; Hung, Y. S.: Analytic properties of the singular values of a rational matrix, *Int. J. Control*, **37**: 2 (1983), 221—234.

[2] Stewart, G. W.: A second order perturbation expansion for small singular values, *Linear Algebra and Appl.*, **56** (1984), 231—235.

[3] Sun, J. G.: Eigenvalues and eigenvectors of a matrix dependent on several parameters, *J. Comp. Math.*, **3** (1985), 351—364.

[4] Sun, J. G.: Sensitivity analysis of multiple eigenvalues (I), *J. Comp. Math.*, **6**: 1(1988), 28—38.

[5] Sun, J. G.: A note on simple non-zero singular values, *J. Comp. Math.*, **6**: 3 (1988), 258—266.

[6] Wilkinson, J. H.; Reinsch, C.: Linear Algebra, Springer-Verlag, New York, Heidelberg, Berlin, 1971.