SENSITIVITY ANALYSIS OF MULTIPLE EIGENVALUES (II)*1)

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

Abstract

This paper discusses the sensitivity of semisimple multiple eigenvalues and corresponding invariant subspaces of a complex (or real) $n\times n$ matrix analytically dependent on several parameters. Some results of this paper may be useful for investigating robust multiple eigenvalue assignment in control system design.

§ 1. Introduction

This paper, as a continuation of [7], discusses the sensitivity of semisimple multiple eigenvalues and corresponding invariant subspaces of a complex (or real) $n\times n$ matrix analytically dependent on several parameters. An eigenvalue of a matrix is called semisimple if the maximal degree of the elementary divisors of this eigenvalue is one.

In addition to the notation explained in [7] we use $\mathbb{C}^{m\times n}$ for the set of complex $m\times n$ matrices, $\mathbb{C}^n=\mathbb{C}^{n\times 1}$, $\mathbb{C}=\mathbb{C}^1$ and

$$\mathbb{C}_r^{m\times n}=\{A\in\mathbb{C}^{m\times n}: \operatorname{rank}(A)=r\}.$$

Let $p=(p_1, p_2, \cdots, p_N)^T\in C^N$. Suppose that $A(p)\in\mathbb{C}^{n\times n}$ is an analytic function in some neighbourhood $\mathscr{B}(p^*)$ of the point $p^*\in\mathbb{C}^N$. Without loss of generality we may assume that the point p^* is the origin of $\mathbb{C}^N$. We consider in this paper the eigenproblem

$$A(p)x(p)=\lambda(p)x(p),\quad \lambda(p)\in\mathbb{C},\ x(p)\in\mathbb{C}^n,\ p\in\mathscr{B}(0). \tag{1.1}$$

First of all we investigate an example.

Example 1.1.

$$A(p)=\begin{bmatrix}1+2p_1+2p_2 & p_2\\ 2p_1 & 1+4p_2\end{bmatrix},\quad p=(p_1, p_2)^T\in\mathbb{C}^2. \tag{1.2}$$

Obviously, the matrix $A(p)$ is an analytic function of $p\in\mathbb{C}^2$, $A(0)$ has a multiple eigenvalue 1 and the eigenvalues of $A(p)$ are

$$\lambda_1(p)=1+p_1+3p_2+\sqrt{p_1^2+p_2^2},\quad \lambda_2(p)=1+p_1+3p_2-\sqrt{p_1^2+p_2^2}. \tag{1.3}$$

Observe that by the theory of analytic function of one complex variable the function $\sqrt{z}$ for $z\in\mathbb{C}$ is defined as

$$\sqrt{z}=|z|^{1/2}e^{(i/2)\arg z},\quad \arg z\in(-\pi,\pi];$$

consequently, if we set

* Received August 11, 1986.
1) The Project Supported by National Natural Science Foundation of China.

$$\Delta_1=\left(-\frac{\pi}{2},\ \frac{\pi}{2}\right],\quad \Delta_2=\left(-\pi,\ -\frac{\pi}{2}\right]\cup\left(\frac{\pi}{2},\ \pi\right],\tag{1.4}$$

then

$$\frac{\sqrt{z^2}}{z}=\frac{e^{\frac{i}{2}\arg(z^2)}}{e^{i\arg z}}=\begin{cases}1, & \arg z\in\Delta_1,\\ -1, & \arg z\in\Delta_2.\end{cases}\tag{1.5}$$

Utilizing (1.4) and (1.5) we get

$$\left(\frac{\partial\lambda_1(p)}{\partial p_1}\right)_{p=0,\ \arg p_1\in\Delta_1}=\left(\frac{\partial\lambda_2(p)}{\partial p_1}\right)_{p=0,\ \arg p_1\in\Delta_2}=2,\tag{1.6}$$

$$\left(\frac{\partial\lambda_1(p)}{\partial p_1}\right)_{p=0,\arg p_1\in\Delta_2}=\left(\frac{\partial\lambda_2(p)}{\partial p_1}\right)_{p=0,\arg p_1\in\Delta_1}=0,\tag{1.7}$$

$$\left(\frac{\partial\lambda_1(p)}{\partial p_2}\right)_{p=0,\ \arg p_2\in\Delta_1}=\left(\frac{\partial\lambda_2(p)}{\partial p_2}\right)_{p=0,\ \arg p_2\in\Delta_2}=4\tag{1.8}$$

and

$$\left(\frac{\partial\lambda_1(p)}{\partial p_2}\right)_{p=0,\ \arg p_2\in\Delta_2}=\left(\frac{\partial\lambda_2(p)}{\partial p_2}\right)_{p=0,\ \arg p_2\in\Delta_1}=2.\tag{1.9}$$

Here we define

$$\left(\frac{\partial\lambda_s(p)}{\partial p_1}\right)_{p=0,\ \arg p_1\in\Delta_t}=\lim_{\substack{|p_1|\to 0\\ \arg p_1=\text{const.}\in\Delta_t}}\frac{\lambda_s(p_1,\ 0)-\lambda_s(0,0)}{p_1},\quad s,\ t=1,\ 2;$$

the partial derivatives $\left(\frac{\partial\lambda_s(p)}{\partial p_2}\right)_{p=0,\ \arg p_2\in\Delta_t}$ $(s,\ t=1,\ 2)$ are defined similarly.

The relations (1.6)—(1.9) show that the functions $\lambda_1(p)$ and $\lambda_2(p)$ are not derivable at $p=0$. Besides, it is worth-while to point out that the functions $\lambda_1(p)$ and $\lambda_2(p)$ are continuous at $p=0$ but not in any neighbourhood of the branch point $p=0$.

Now we set

$$\hat{A}(p_1)=(A(p))_{p=(p_1,0)^T},\quad \tilde{A}(p_2)=(A(p))_{p=(0,p_2)^T},$$

in which $A(p)$ is described in (1.2). It is easy to see that $\hat{A}(0)$ and $\tilde{A}(0)$ have multiple eigenvalue 1, the eigenvalues of $\hat{A}(p_1)$ are

$$\hat{\lambda}_1(p_1)=1+2p_1,\quad \hat{\lambda}_2(p_1)=1,\tag{1.10}$$

and the eigenvalues of $\tilde{A}(p_2)$ are

$$\tilde{\lambda}_1(p_2)=1+2p_2,\quad \tilde{\lambda}_2(p_2)=1+4p_2.\tag{1.11}$$

Comparing (1.10), (1.11) with (1.3) we find that

$$(\lambda_1(p))_{p=(p_1,0)^T}=\begin{cases}\hat{\lambda}_1(p_1), & \arg p_1\in\Delta_1,\\ \hat{\lambda}_2(p_1), & \arg p_1\in\Delta_2,\end{cases}$$

$$(\lambda_2(p))_{p=(p_1,0)^T}=\begin{cases}\hat{\lambda}_2(p_1), & \arg p_1\in\Delta_1,\\ \hat{\lambda}_1(p_1), & \arg p_1\in\Delta_2,\end{cases}$$

$$(\lambda_1(p))_{p=(0,p_2)^T}=\begin{cases}\tilde{\lambda}_2(p_2), & \arg p_2\in\Delta_1,\\ \tilde{\lambda}_1(p_2), & \arg p_2\in\Delta_2,\end{cases}$$

$$(\lambda_2(p))_{p=(0,p_2)^T}=\begin{cases}\tilde{\lambda}_1(p_2), & \arg p_2\in\Delta_1,\\ \tilde{\lambda}_2(p_2), & \arg p_2\in\Delta_2,\end{cases}$$

where Δ_1 and Δ_2 are defined by (1.4).

We note that the following facts are important: the functions $\hat{\lambda}_1(p_1)$ and $\hat{\lambda}_2(p_1)$

are derivable at $p_1=0$, for each $p_1\in\mathbb{C}$ the set $\{\hat{\lambda}_t(p_1)\}_{t=1}^2$ and the set $\{(\lambda_s(p))_{p=(p_1,0)^T}\}_{s=1}^2$ are just the same, $\hat{\lambda}_1(p_1)\neq\hat{\lambda}_2(p_1)$, $\forall p_1\in\mathbb{C}\setminus\{0\}$, and

$$\left\{\left(\frac{d\hat{\lambda}_t(p_1)}{dp_1}\right)_{p_1=0}\right\}_{t=1}^2=\{0,\ 2\}=\lambda\left(\left(\frac{\partial A(p)}{\partial p_1}\right)_{p=0}\right); \tag{1.12}$$

similarly, the functions $\tilde{\lambda}_1(p_2)$ and $\tilde{\lambda}_2(p_2)$ are derivable at $p_2=0$, for each $p_2\in\mathbb{C}$ the set $\{\tilde{\lambda}_t(p_2)\}_{t=1}^2$ and the set $\{(\lambda_s(p))_{p=(0,p_2)^T}\}_{s=1}^2$ are just the same, $\tilde{\lambda}_1(p_2)\neq\tilde{\lambda}_2(p_2)$, $\forall p_2\in\mathbb{C}\setminus\{0\}$, and

$$\left\{\left(\frac{d\tilde{\lambda}_t(p_2)}{dp_2}\right)_{p_2=0}\right\}_{t=1}^2=\{2,\ 4\}=\lambda\left(\left(\frac{\partial A(p)}{\partial p_2}\right)_{p=0}\right). \tag{1.13}$$

Consequently, we may define $\rho\left(\left(\frac{\partial A(p)}{\partial p_1}\right)_{p=0}\right)$ and $\rho\left(\left(\frac{\partial A(p)}{\partial p_2}\right)_{p=0}\right)$ as the sensitivities of the semisimple multiple eigenvalue 1 with respect to the parameters p_1 and p_2, respectively.

In the next section we shall prove that the above mentioned facts are of universal significance, on the basis of which we may define the sensitivities of semisimple multiple eigenvalues dependent on several parameters, and in § 3 give some formulas for computing the sensitivities.

§ 2. Some Results about Partial Derivatives

First we cite the following implicit function theorem[1, p.39].

Implicit Function Theorem. *If the complex-value functions*

$$f_j(\xi_1,\ \cdots,\ \xi_s,\ \eta_1,\ \cdots,\ \eta_t),\quad j=1,\cdots,\ s$$

are analytic functions of $s+t$ complex variables in some neighbourhood of the origin of $\mathbb{C}^{s+t}$, if $f_j(0,\ 0)=0$, $j=1,\ \cdots,\ s$, and if

$$\det\frac{\partial(f_1,\ \cdots,\ f_s)}{\partial(\xi_1,\ \cdots,\ \xi_s)}\neq 0 \text{ for } \xi_1=\cdots=\xi_s=\eta_1=\cdots=\eta_t=0,$$

then the equations

$$f_j(\xi_1,\ \cdots,\ \xi_s,\ \eta_1,\ \cdots,\ \eta_t)=0,\quad j=1,\ \cdots,\ s$$

have a unique solution

$$\xi_j=g_j(\eta_1,\ \cdots,\ \eta_t),\quad j=1,\ \cdots,\ s$$

vanishing for $\eta_1=\cdots=\eta_t=0$ and analytic in some neighbourhood of the origin of $\mathbb{C}^t$.

Then, we introduce the following definition.

Definition 2.1. *Let $A\in\mathbb{C}^{n\times n}$. A subspace $\mathscr{X}\subset\mathbb{C}^n$ is an invariant subspace of the matrix A if $A\mathscr{X}\subset\mathscr{X}$.*

By [5], an l-dimensional subspace $\mathscr{X}\subset\mathbb{C}^n$ is an invariant subspace of a matrix $A\in\mathbb{C}^{n\times n}$ if and only if there are $X_1\in\mathbb{C}_r^{n\times r}$ and $A_1\in\mathbb{C}^{r\times r}$ such that the set of column vectors of X_1 spans the subspace $\mathscr{X}$ and

$$AX_1=X_1A_1. \tag{2.1}$$

Moreover, the relation (2.1) implies $\lambda(A_1)\subset\lambda(A)$. If $\lambda_1,\ \cdots,\ \lambda_r$ are the eigenvalues of A_1, we may call $\mathscr{X}$ the invariant subspace of A corresponding to the eigenvalues $\lambda_1,\ \cdots,\ \lambda_r$.

The following theorem is the main result of this section.

Theorem 2.1. *Let $p=(p_1,\ \cdots,\ p_N)^T\in\mathbb{C}^N$, and let $A(p)\in\mathbb{C}^{n\times n}$ be an analytic* *function of p in some neighbourhood* $\mathscr{B}(0)$ *of the origin of* $\mathbb{C}^N$. *Suppose that there are nonsingular matrices* $X,Y\in\mathbb{C}^{n\times n}$ *satisfying*

$$X=(X_1, X_2),\quad Y=(Y_1, Y_2),\quad X_1, Y_1\in\mathbb{C}^{n\times r},$$
$$Y^TX=I \tag{2.2}$$

and

$$Y^TA(0)X=\begin{pmatrix}\lambda_1 I^{(r)} & 0\\ 0 & A_2\end{pmatrix},\quad \lambda_1\overline{\in}\lambda(A_2). \tag{2.3}$$

Then

1) *the eigenproblem* (1.1) *has r eigenvalues* $\lambda_1(p), \cdots, \lambda_r(p)$, *which are continuous at* $p=0$ *and satisfy*

$$\lambda_s(0)=\lambda_1,\quad s=1, \cdots, r \tag{2.4}$$

2) *for any fixed index* $j\in\{1, \cdots, N\}$, *there exist r functions* $\hat{\lambda}_1(p_j), \cdots, \hat{\lambda}_r(p_j)$, *a neighbourhood* $\hat{\mathscr{B}}_0$ *of the origin and a straight line* $\mathscr{L}_0$ *from the origin to infinity in the* p_j*-plane, and a permutation* π *of* $1, \cdots, r$ *dependent on the index j, such that the functions* $\hat{\lambda}_1(p_j), \cdots, \hat{\lambda}_r(p_j)$ are regular in $\hat{\mathscr{B}}_0\backslash\mathscr{L}_0$, *the values of* $\{\hat{\lambda}_t(p_j)\}_{t=1}^r$ *constitute* $\{(\lambda_s(p))_{p=(0,\cdots,0,p_j,0,\cdots,0)^T}\}_{s=1}^r$ *for each* $p_j\in\hat{\mathscr{B}}_0$, $\hat{\lambda}_{t_1}(p_j)\neq\hat{\lambda}_{t_2}(p_j)$, $\forall p_j\in\hat{\mathscr{B}}_0\setminus\{0\}$ *provided that* $\hat{\lambda}_{t_1}(p_j)\not\equiv\hat{\lambda}_{t_2}(p_j)$ *for* $p_j\in\hat{\mathscr{B}}_0$, *and*

$$\left(\frac{d\hat{\lambda}_t(p_j)}{dp_j}\right)_{p_j=0}=\lambda_{\pi(t)}\left(Y_1^T\left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0}X_1\right),\quad t=1, \cdots, r. \tag{2.5}$$

3) *there exists an analytic function* $X_1(p)\in\mathbb{C}_r^{n\times r}$ *whose column vectors span the invariant subspace of* $A(p)$ *corresponding to the eigenvalues* $\lambda_1(p), \cdots, \lambda_r(p)$ *in some neighbourhood* $\mathscr{B}_0\subset\mathscr{B}(0)$ *of the origin, such that*

$$X_1(0)=X_1 \tag{2.6}$$

and

$$\left(\frac{\partial X_1(p)}{\partial p_j}\right)_{p=0}=X_2(\lambda_1 I-A_2)^{-1}Y_2^T\left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0}X_1,\quad j=1, \cdots, N. \tag{2.7}$$

Proof. 1° Let

$$\tilde{A}(p)=Y^TA(p)X=\begin{pmatrix}\tilde{A}_{11}(p) & \tilde{A}_{12}(p)\\ \tilde{A}_{21}(p) & \tilde{A}_{22}(p)\end{pmatrix},\quad \tilde{A}_{11}(p)\in\mathbb{C}^{r\times r} \tag{2.8}$$

and

$$F(Z, p)=\tilde{A}_{21}(p)-Z\tilde{A}_{11}(p)+\tilde{A}_{22}(p)Z-Z\tilde{A}_{12}(p)Z, \tag{2.9}$$

in which

$$Z=(\zeta_{jk})\in\mathbb{C}^{(n-r)\times r},\quad F(Z, p)=(f_{jk}(Z, p))\in\mathbb{C}^{(n-r)\times r}. \tag{2.10}$$

Observe that the function $F(Z, p)$ is analytic for $Z\in\mathbb{C}^{(n-r)\times r}$ and $p\in\mathscr{B}(0)$,

$$f_{jk}(0, 0)=0,\quad j=1, \cdots, n-r,\quad k=1, \cdots, r$$

and

$$\left(\det\frac{\partial(f_{11}, \cdots, f_{1r}, f_{21}, \cdots, f_{2r}, f_{n-r,1}, \cdots, f_{n-r,r})}{\partial(\zeta_{11}, \cdots, \zeta_{1r}, \zeta_{21}, \cdots, \zeta_{2r}, \zeta_{n-r,1}, \cdots, \zeta_{n-r,r})}\right)_{Z=0,p=0}$$
$$=\det(I^{(r)}\otimes A_2-\lambda_1 I^{(r)}\otimes I^{(n-r)})=\det(A_2-\lambda_1 I)^r\neq0,$$

where $\otimes$ denotes the Kronecker product. Hence, by the Implicit Function Theorem the equation

$$F(Z, p)=0 \tag{2.11}$$

has a unique analytic solution $Z=Z(p)$ in some neighbourhood $\mathscr{B}_1(\subset\mathscr{B}(0))$ of the origin of $\mathbb{C}^N$ with $Z(0)=0$, and thus we have

$$\begin{pmatrix} I & 0 \\ Z(p) & I \end{pmatrix}^{-1} \widetilde{A}(p)\begin{pmatrix} I & 0 \\ Z(p) & I \end{pmatrix}=\begin{pmatrix} A_1(p) & \widetilde{A}_{12}(p) \\ 0 & A_2(p) \end{pmatrix}, \tag{2.12}$$

in which

$$A_1(p)=\widetilde{A}_{11}(p)+\widetilde{A}_{12}(p)Z(p)\in\mathbb{C}^{r\times r},\ A_2(p)=\widetilde{A}_{22}(p)-Z(p)\widetilde{A}_{12}(p). \tag{2.13}$$

From (2.12)

$$\widetilde{A}(p)\begin{pmatrix} I \\ Z(p) \end{pmatrix}=\begin{pmatrix} I \\ Z(p) \end{pmatrix}A_1(p).$$

Combining it with (2.3) and (2.8) and writing

$$X_1(p)=X\begin{pmatrix} I \\ Z(p) \end{pmatrix}, \tag{2.14}$$

we get

$$A(p)X_1(p)=X_1(p)A_1(p) \tag{2.15}$$

and

$$A_1(0)=\lambda_1 I^{(r)},\quad X_1(0)=X_1. \tag{2.16}$$

2°. It follows from a similar argument that the equation

$$\widetilde{A}_{12}(p)-\widetilde{A}_{11}(p)W^T+W^T\widetilde{A}_{22}(p)-W^T\widetilde{A}_{21}(p)W^T=0$$

has a unique analytic solution $W=W(p)$ in some neighbourhood $\mathscr{B}_0(\subset\mathscr{B}_1)$ of the origin of $\mathbb{C}^N$ with $W(0)=0$, and thus we have

$$\begin{pmatrix} I & W(p)^T \\ 0 & I \end{pmatrix}\widetilde{A}(p)\begin{pmatrix} I & W(p)^T \\ 0 & I \end{pmatrix}^{-1}=\begin{pmatrix} A_1'(p) & 0 \\ A_{21}(p) & A_2'(p) \end{pmatrix}, \tag{2.17}$$

in which

$$A_1'(p)=\widetilde{A}_{11}(p)+W(p)^T\widetilde{A}_{21}(p),\quad A_2'(p)=\widetilde{A}_{22}(p)-\widetilde{A}_{21}(p)W(p)^T. \tag{2.18}$$

From (2.17)

$$\begin{pmatrix} I \\ W(p) \end{pmatrix}^T \widetilde{A}(p)=A_1'(p)\begin{pmatrix} I \\ W(p) \end{pmatrix}^T.$$

Combining it with (2.3) and (2.8), and writing

$$Y_1(p)=Y\begin{pmatrix} I \\ W(p) \end{pmatrix}, \tag{2.19}$$

we get

$$Y_1(p)^TA(p)=A_1'(p)Y_1(p)^T \tag{2.20}$$

and

$$A_1'(0)=\lambda_1 I^{(r)},\quad Y_1(0)=Y_1. \tag{2.21}$$

Moreover, according to the continuity of $X_1(p)$ and $Y_1(p)$, from (2.16), (2.21) and $Y_1^TX_1=I$ we know that the matrix $Y_1(p)^TX_1(p)$ is nonsingular $\forall p\in\mathscr{B}_0$ provided that the neighbourhood $\mathscr{B}_0$ is sufficiently small. In the following we shall assume that $\mathscr{B}_0$ is so small that

$$\operatorname{rank}(Y_1(p)^TX_1(p))=r,\quad \forall p\in\mathscr{B}_0. \tag{2.22}$$

3°. From (2.15) and (2.22)

$$A_1(p)=(Y_1(p)^T X_1(p))^{-1}Y_1(p)^T A(p)X_1(p),\quad p\in\mathscr{B}_0. \tag{2.23}$$

According to (2.2), (2.3), (2.14), (2.16), (2.19) and (2.21), the first partial derivative of Eq. (2.23) with respect to p_j at $p=0$ may be written in the form

$$\left(\frac{\partial A_1(p)}{\partial p_j}\right)_{p=0}=Y_1^T\left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0}X_1,\quad j=1,\cdots,N. \tag{2.24}$$

Let

$$\lambda(A_1(p))=\{\lambda_s(p)\}_{s=1}^{r}. \tag{2.25}$$

By [5] from (2.15) we know that

$$\lambda_s(p)\in\lambda(A(p)),\quad \lambda_s(0)=\lambda_1,\quad s=1,\cdots,r \tag{2.26}$$

and that the eigenvalues $\lambda_1(p),\cdots,\lambda_r(p)$ are sufficiently near λ_1 provided that the point p belongs to a sufficient small neighbourhood $\mathscr{B}_0$ of the origin.

Let ε be an arbitrary positive number, and let j be any fixed index from $1,\cdots,N$. By the Jordan canonical form theorem there is a matrix $Q_j\in\mathbb{C}_r^{r\times r}$ such that

$$Q_j^{-1}\left(\frac{\partial A_1(p)}{\partial p_j}\right)_{p=0}Q_j=\operatorname{diag}(J_{\mu_1^{(j)}},\cdots,J_{\mu_q^{(j)}}), \tag{2.27}$$

in which $\mu_k^{(j)}\neq\mu_l^{(j)}$ for $k\neq l$, and

$$\begin{cases} J_{\mu_k^{(j)}}=\operatorname{diag}(J_{\mu_k^{(j)}}^{(1)},\cdots,J_{\mu_k^{(j)}}^{(m_k)})\in\mathbb{C}^{r_k\times r_k},\\[2mm] J_{\mu_k^{(j)}}^{(m)}=\begin{pmatrix}\mu_k^{(j)} & \frac{\varepsilon}{2} & & 0\\ & \ddots & \ddots & \\ & & \ddots & \frac{\varepsilon}{2}\\ 0 & & & \mu_k^{(j)}\end{pmatrix}, \quad m=1,\cdots,m_k,\ k=1,\cdots,q. \end{cases} \tag{2.28}$$

For simplicity we write

$$Q_j^{-1}\left(\frac{\partial A_1(p)}{\partial p_j}\right)_{p=0}Q_j=\begin{pmatrix}\delta_1^{(j)} & \varepsilon_1^{(j)} & & 0\\ & \ddots & \ddots & \\ & & \ddots & \varepsilon_{r-1}^{(j)}\\ 0 & & & \delta_r^{(j)}\end{pmatrix}, \tag{2.29}$$

where $\delta_1^{(j)}=\mu_1^{(j)},\cdots,\delta_r^{(j)}=\mu_q^{(j)}$, and the subdiagonal elements $\varepsilon_l^{(j)}\,(1\leqslant l\leqslant r-1)$ are equal to $\frac{\varepsilon}{2}$ or zero. Since

$$Q_j^{-1}\left(\frac{\partial A_1(p)}{\partial p_j}\right)_{p=0}Q_j=\left(\frac{\partial}{\partial p_j}(Q_j^{-1}A_1(p)Q_j)\right)_{p=0}, \tag{2.30}$$

we have

$$(Q_j^{-1}A_1(p)Q_j)_{p=(0,\cdots,0,p_j,0,\cdots,0)^T}=(\theta_{kl}(p_j))_{1<k,l<r}, \tag{2.31}$$

in which the functions $\theta_{kl}(p_j)$ are analytic and so may be written as convergent power series:

$$\theta_{kl}(p_j)=\theta_{kl}^{(j,0)}+\theta_{kl}^{(j,1)}p_j+\theta_{kl}^{(j,2)}p_j^2+\theta_{kl}^{(j,3)}p_j^3+\cdots,\quad k,l=1,\cdots,r. \tag{2.32}$$

From (2.29) and

$$(Q_j^{-1}A_1(p)Q_j)_{p=0}=\lambda_1 I^{(r)}$$

it follows that

$$\theta_{kl}^{(j,0)}=\begin{cases}\lambda_1, & \text{if } k=l,\\ 0, & \text{if } k\neq l,\end{cases}\qquad \theta_{kl}^{(j,1)}=\begin{cases}\delta_k^{(j)}, & \text{if } k=l,\\ \frac{\varepsilon}{2} \text{ or } 0, & \text{if } k=l-1,\\ 0, & \text{otherwise.}\end{cases}$$

Therefore we have

$$\theta_{kl}(p_j)=\begin{cases}\lambda_1+\delta_k^{(j)}p_j+\theta_{kk}^{(j,2)}p_j^2+\theta_{kk}^{(j,3)}p_j^3+\cdots, & \text{if } k=l,\\ \frac{\varepsilon}{2}p_j+\theta_{k,k+1}^{(j,2)}p_j^2+\theta_{k,k+1}^{(j,3)}p_j^3+\cdots & \\ \quad\text{or } \theta_{k,k+1}^{(j,2)}p_j^2+\theta_{k,k+1}^{(j,3)}p_j^3+\cdots, & \text{if } k=l-1,\\ \theta_{kl}^{(j,2)}p_j^2+\theta_{kl}^{(j,3)}p_j^3+\cdots, & \text{otherwise.}\end{cases} \tag{2.33}$$

Let

$$\hat{A}(p_j)=(A_1(p))_{p=(0,\cdots,0,p_j,0,\cdots,0)^T} \tag{2.34}$$

and

$$\lambda(\hat{A}(p_j))=\{\hat{\lambda}_t(p_j)\}_{t=1}^{r}. \tag{2.35}$$

By (2.16), λ_1 is a semisimple multiple eigenvalue of $\hat{A}(0)$; moreover, $\hat{A}(p_j)$ is an analytic function of p_j in some neighbourhood $\mathscr{B}_0$ of the origin of $\mathbb{C}$. Therefore the eigenvalues $\hat{\lambda}_t(p_j)$ of $\hat{A}(p_j)$ may be expressed as the following convergent Puiseux series[4, p.237, p.249]:

$$\hat{\lambda}_t(p_j)=\lambda_1+\varphi_t^{(j,1)}p_j+\sum_{m>r'}\varphi_t^{(j,m)}(p^{1/r'})^m,\quad t=1,\cdots,r, \tag{2.36}$$

where the natural number $r'\leqslant r$. On the other hand, by the Gerschgorin Theorem, from (2.27)—(2.29), (2.31) and (2.33) we know that there are precisely q circular disks $D_1,\cdots,D_q$ with centers $\lambda_1+\mu_1^{(j)}p_j,\cdots,\lambda_1+\mu_q^{(j)}p_j$ and with radii of magnitude $\frac{\varepsilon}{2}|p_j|+O(|p_j|^2)$, respectively. Observe that the disks $D_1,\cdots,D_q$ are mutually disjoint provided that ε is sufficiently small and p_j belong to a sufficient small $\hat{\mathscr{B}}_0$, and in such a case every disk D_k contains exactly r_k eigenvalues $\hat{\lambda}_{r_1+\cdots+r_{k-1}+1}(p_j),\cdots,\hat{\lambda}_{r_1+\cdots+r_{k-1}+r_k}(p_j)$ which may be written as a convergent Puiseux series

$$\hat{\lambda}_{r_1+\cdots+r_{k-1}+l}(p_j)=\lambda_1+\varphi_{r_1+\cdots+r_{k-1}+l}^{(j,1)}p_j+\sum_{m>r'}\varphi_{r_1+\cdots+r_{k-1}+l}^{(j,m)}p_j^{m/r'},$$
$$l=1,\cdots,r_k,\quad k=1,\cdots,q,\ p_j\in\hat{\mathscr{B}}_0, \tag{2.37}$$

here $r_0=0$. From

$$|\varphi_{r_1+\cdots+r_{k-1}+l}^{(j,1)}p_j-\mu_k^{(j)}p_j|\leqslant\frac{\varepsilon}{2}|p_j|+O(|p_j|^{1+\frac{1}{r'}}),\quad p_j\in\hat{\mathscr{B}}_0,$$

we get

$$|\varphi_{r_1+\cdots+r_{k-1}+l}^{(j,1)}-\mu_k^{(j)}|\leqslant\frac{\varepsilon}{2}+O(|p_j|^{\frac{1}{r'}})<\varepsilon,\quad l=1,\cdots,r_j$$

provided that the neighbourhood $\hat{\mathscr{B}}_0$ is sufficiently small. Since ε is an arbitrary positive number, , we have

$$\varphi_{r_1+\cdots+r_{k-1}+l}^{(j,1)}=\mu_k^{(j)},\quad l=1,\cdots,r_k,\ k=1,\cdots,q. \tag{2.38}$$

Therefore, from (2.37), (2.38) and (2.29) we obtain

$$\hat{\lambda}_t(p_j)=\lambda_1+\delta_t^{(j)}p_j+\sum_{m>r'}\varphi_t^{(j,m)}(p^{\frac{1}{r'}})^m,\quad t=1,\cdots,r,\quad p_j\in\hat{\mathscr{B}}_0. \tag{2.39}$$

It is worth-while to point out that if the p_j-plane is cut by a suitable straight line $\mathscr{L}_0$ from the origin to infinity, then the functions $\hat{\lambda}_1(p_j), \cdots, \hat{\lambda}_r(p_j)$, as the eigenvalues of the matrix $(A_1(p))_{p=(0,\cdots,0,p_j,0,\cdots,0)^T}$, are regular in $\hat{\mathscr{B}}_0 \setminus \mathscr{L}_0$, and for any two different indexes $t_1, t_2 \in \{1, \cdots, r\}$ we have

$$\hat{\lambda}_{t_1}(p_j) \neq \hat{\lambda}_{t_2}(p_j), \quad \forall p_j \in \hat{\mathscr{B}}_0 \setminus \{0\} \tag{2.40}$$

provided that $\hat{\lambda}_{t_1}(p_j) \not\equiv \hat{\lambda}_{t_2}(p_j)$ for $p_j \in \hat{\mathscr{B}}_0$ and that $\hat{\mathscr{B}}_0$ is sufficiently small ([2, p. 74—75]). Moreover, for each $p_j \in \hat{\mathscr{B}}_0$ the set $\{\hat{\lambda}_t(p_j)\}_{t=1}^r$ and the set $\{\lambda_s(0, \cdots, 0, p_j, 0, \cdots, 0)\}_{s=1}^r$ are just the same, in which the eigenvalues $\lambda_1(p), \cdots, \lambda_r(p)$ of the eigenproblem (1.1) are continuous at $p=0$ and $\lambda_s(0) = \lambda_1$, $s=1, \cdots, r$.

From (2.39)

$$\left(\frac{d\hat{\lambda}_t(p_j)}{dp_j}\right)_{p_j=0} = \delta_t^{(j)}, \quad t=1, \cdots, r. \tag{2.41}$$

and from (2.29) and (2.24) we have

$$\{\delta_t^{(j)}\}_{t=1}^r = \lambda\left(Q_j^{-1}\left(\frac{\partial A_1(p)}{\partial p_j}\right)_{p=0} Q_j\right) = \lambda\left(Y_1^T\left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0} X_1\right). \tag{2.42}$$

Combining (2.41) with (2.42), we get (2.5).

4° The relations (2.15) and (2.25) show that the set of column vectors of $X_1(p)$ spans an invariant subspace of $A(p)$ corresponding to the eigenvalues $\lambda_1(p), \cdots, \lambda_r(p)$ in some neighbourhood of the origin of $\mathbb{C}^N$, and $X_1(0) = X_1$. From (2.14) and (2.15) we get

$$(\lambda_1 I - A(0)) X \begin{pmatrix} 0 \\ \dfrac{\partial Z(p)}{\partial p_j} \end{pmatrix}_{p=0} = \left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0} X_1 - X_1\left(\frac{\partial A_1(p)}{\partial p_j}\right)_{p=0}. \tag{2.43}$$

Combining (2.43) with (2.3) we have

$$\begin{pmatrix} 0 & 0 \\ 0 & \lambda_1 I - A_2 \end{pmatrix}\begin{pmatrix} 0 \\ \dfrac{\partial Z(p)}{\partial p_j} \end{pmatrix}_{p=0} = Y^T\left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0} X_1 - \begin{pmatrix} I^{(r)} \\ 0 \end{pmatrix}\left(\frac{\partial A_1(p)}{\partial p_j}\right)_{p=0}. \tag{2.44}$$

Since $\lambda_1 \bar{\in} \lambda(A_2)$, it follows from (2.44) that

$$\left(\frac{\partial Z(p)}{\partial p_j}\right)_{p=0} = (\lambda_1 I - A_2)^{-1} Y_2^T \left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0} X_1. \tag{2.45}$$

Substituting (2.45) into

$$\left(\frac{\partial X_1(p)}{\partial p_j}\right)_{p=0} = X_2\left(\frac{\partial Z(p)}{\partial p_j}\right)_{p=0}$$

we obtain the relations (2.7). ▮

§ 3. Applications

3.1. Sensitivity of Semisimple Multiple Eigenvalues

According to Theorem 2.1 we may introduce the following definition.

Definition 3.1. *Let $p = (p_1, \cdots, p_N)^T \in \mathbb{C}^N$, and let $A(p) \in \mathbb{C}^{n\times n}$ be an analytic function of p in some neighbourhood $\mathscr{B}(p^*)$ of the point $p^* \in \mathbb{C}^N$. Suppose that there are nonsingular matrices $X, Y \in \mathbb{C}^{n\times n}$ satisfying*

$$Y^T X = I, \quad X = (X_1, X_2), \quad Y = (Y_1, Y_2), \quad X_1, Y_1 \in \mathbb{C}^{n\times r} \tag{3.1}$$

and

$$Y^T A(0) X = \begin{pmatrix} \lambda_1 I^{(r)} & 0 \\ 0 & A_2 \end{pmatrix}, \quad \lambda_1 \overline{\in} \lambda(A_2).$$

Then the quantity

$$s_{p_j}(\lambda_1) = \rho\left(Y_1^T \left(\frac{\partial A(p)}{\partial p_j}\right)_{p=0} X_1\right) \tag{3.2}$$

is called the sensitivity of the semisimple multiple eigenvalue λ_1 with respect to the parameter p_j; the quantity

$$s_{p_{j_1},\cdots,p_{j_m}}(\lambda_1) = \sqrt{\sum_{k=1}^{m} s^2_{p_{j_k}}(\lambda_1)} \tag{3.3}$$

is called the sensitivity of the semisimple eigenvalue λ_1 with respect to the parameters $p_{j_1}, \cdots, p_{j_m}$;

$$s_p(\lambda_1) = \sqrt{\sum_{j=1}^{N} s^2_{p_j}(\lambda_1)} \tag{3.4}$$

is called the sensitivity of the semisimple multiple eigenvalue λ_1.

Example 3.1. The matrix $A(p)$ of Example 1.1 has a semisimple multiple eigenvalue $\lambda_1 = 1$ at $p = 0 \in \mathbb{C}^2$. By Definition 3.1 we have

$$s_{p_1}(\lambda_1) = 2, \quad s_{p_2}(\lambda_1) = 4, \quad s_p(\lambda_1) = 2\sqrt{5}.$$

3.2. Determination of Sensitive Elements

Let $A = (\alpha_{jk}) \in \mathbb{C}^{n\times n}$. Assume that there are nonsingular matrices $X, Y \in \mathbb{C}^{n\times n}$ satisfying condition (3.1) and

$$Y^T A X = \begin{pmatrix} \lambda_1 I^{(r)} & 0 \\ 0 & A_2 \end{pmatrix}, \quad \lambda_1 \overline{\in} \lambda(A_2).$$

We regard the elements α_{jk} as parameters. By Definition 3.1 the sensitivity of the semisimple multiple eigenvalue λ_1 with respect to α_{jk} is

$$s_{\alpha_{jk}}(\lambda_1) = \rho\left(Y_1^T \frac{\partial A}{\partial \alpha_{jk}} X_1\right), \quad j, k = 1, \cdots, n. \tag{3.5}$$

Let

$$X_1 = \begin{pmatrix} x_1'^T \\ \vdots \\ x_n'^T \end{pmatrix}, \quad Y_1 = \begin{pmatrix} y_1'^T \\ \vdots \\ y_n'^T \end{pmatrix}, \quad x_j', y_j' \in \mathbb{C}^r \quad \forall j.$$

From (3.5) we get

$$s_{\alpha_{jk}}(\lambda_1) = \rho(y_j' x_k'^T) = |y_j'^T x_k'|, \quad j, k = 1, \cdots, n. \tag{3.6}$$

Moreover, by Definition 3.1 the sensitivity of the semisimple multiple eigenvalue λ_1 with respect to A is

$$s_A(\lambda_1) = \sqrt{\sum_{j,k=1}^{n} |y_j'^T x_k'|^2} = \sqrt{\operatorname{tr}(Y_1^T Y_1 X_1^T X_1)}. \tag{3.7}$$

3.3. Measures of Robustness in Control System Design

A closed loop linear system with coefficient matrix $M \in \mathbb{R}^{n\times n}$ is said to be robust if its eigenvalues are as insensitive to perturbations in M as possible [3].

Assume that there are $X \in \mathbb{R}_n^{n\times n}$ and

$$\Lambda = \operatorname{diag}(\lambda_1 I^{(n_1)}, \cdots, \lambda_q I^{(n_q)}) \in \mathbb{R}^{n\times n}, \quad \lambda_j \neq \lambda_k \text{ for } j \neq k \tag{3.8}$$

such that

$$M=X\Lambda X^{-1}. \tag{3.9}$$

Let

$$X=(x_1,\cdots,x_n)=(\underset{n_1}{X_1},\cdots,\underset{n_q}{X_q}),\ Y=X^{-T}=(y_1,\cdots,y_n)=(\underset{n_1}{Y_1},\cdots,\underset{n_q}{Y_q}). \tag{3.10}$$

Then the relation (3.9) may be rewritten as

$$Y^TMX=\Lambda. \tag{3.11}$$

By Definition 3.1 we may define

$$\nu_s=\sqrt{\sum_{j=1}^{q}s_M^2(\lambda_j)} \tag{3.12}$$

as a measure of robustness of M. From (3.7) and (3.10)—(3.12) we get

$$\nu_s=\sqrt{\sum_{j=1}^{q}\operatorname{tr}(Y_j^TY_jX_j^TX_j)}. \tag{3.13}$$

Remark 3.1. If the matrix X in (3.10) satisfies $X_j^TX_j=I^{(n_j)}\ \forall j$, then

$$\nu_s=\|Y\|_F \tag{3.14}$$

which coincides with $\sqrt{n}\ \nu_3(I)$, here $\nu_3(I)$ is a measure of robustness introduced in [3].

Remark 3.2. For the vectors $x_1,\cdots,x_n$ and $y_1,\cdots,y_n$ described in (3.10), we set

$$c_j=\|x_j\|_2\|y_j\|_2\quad\forall j,\quad c=(c_1,\cdots,c_n)^T\in\mathbb{R}^n. \tag{3.15}$$

The formula (3.13) shows that if $\max_j n_j=1$ then

$$\nu_s=\|c\|_2; \tag{3.16}$$

on the other hand, if $\max_j n_j>1$, then from (3.13) we can deduce

$$\nu_s=\Big[\sum_{j=1}^{q}\Big(\sum_{l=n_1+\cdots+n_{j-1}+1}^{n_1+\cdots+n_{j-1}+n_j}\|x_l\|^2\|y_l\|^2+\sum_{\substack{l,m=n_1+\cdots+n_{j-1}+1\\ l\neq m}}^{n_1+\cdots+n_{j-1}+n_j}x_l^Tx_my_l^Ty_m\Big)\Big]^{\frac{1}{2}}$$

$$\leqslant\sqrt{\sum_{j=1}^{q}\Big(n_j\sum_{l=n_1+\cdots+n_{j-1}+1}^{n_1+\cdots+n_{j-1}+n_j}c_l^2\Big)}\leqslant\max_{1<j<q}\sqrt{n_j}\cdot\|c\|_2. \tag{3.17}$$

Hence, in the case of multiple eigenvalues we may use the quantity $\|c\|_2$ (or $\|c\|_2/\sqrt{n}$, see [6]) as a measure of robustness in control system design.

Remark 3.3. From (3.7) we see that for the matrix M described in 3.3 and for an arbitrary j $(1\leqslant j\leqslant q)$ we have

$$s_M(\lambda_j)=\sqrt{\operatorname{tr}(Y_j^TY_jX_j^TX_j)}\leqslant n_jc_j', \tag{3.18}$$

where

$$c_j'=\max_{1<l<n_j}c_{n_1+\cdots+n_{j-1}+l}. \tag{3.19}$$

Usually we use n_jc_j' as a condition number of the semisimple multiple eigenvalue λ_j (see [8, Chapter 3, § 2]).

Example 3.2[6]. Suppose that the matrices X, Y and Λ in (3.11) are as follows:

$$X=\left(\begin{array}{cc|c}0.883222 & -0.503560 & -1.129710\\ 0.468608 & 0.127883 & 0.356507\\ 0.114293 & 0.861696 & 0.221624\end{array}\right)=(X_1\mid X_2),$$

$$Y=\begin{pmatrix} 0.426276 & 0.096469 & \vline & -0.594922 \\ 1.317480 & -0.496596 & \vline & 1.251380 \\ 0.053581 & 1.290580 & \vline & -0.533375 \end{pmatrix}=(Y_1 \mid Y_2),$$

$$\Lambda=\mathrm{diag}(-0.2,\ -0.2,\ -10.0)=\mathrm{diag}(\lambda_1,\ \lambda_1,\ \lambda_2).$$

By (3.7) and (3.13) we have

$$s_M(\lambda_1)=\sqrt{\mathrm{tr}(Y_1^TY_1X_1^TX_1)}=2.04982,$$

$$s_M(\lambda_2)=\sqrt{Y_2^TY_2X_2^TX_2}=1.78933$$

and

$$\nu_s=\sqrt{s_M^2(\lambda_1)+s_M^2(\lambda_2)}=2.72093.$$

On the other hand, by (3.15) we have

$$c_1=1.39456,\ c_2=1.39478,\ c_3=1.78934$$

and

$$\|c\|_2=2.66307.$$

References

[1] Bochner S.; Martin W. T.: Several Complex Variables, Princeton, 1948.

[2] Kato T.: A Short Introduction to Perturbation Theory for Linear Operators, Springer-Verlag, New York, Heidelberg, Berlin, 1982.

[3] Kautsky J.; Nichols N. K.: Robust multiple eigenvalue assignment by state feedback in linear system, *Numerical Analysis Report* 7/84 (1984), Department of Mathematics, University of Reading.

[4] Lancaster P.: Theory of Matrices, Academic Press, New York and London, 1969.

[5] Sun J. G.: Invariant subspaces and generalized invariant subspaces I. Existence and uniqueness theorems, *Math. Numer. Sinica*, **2** (1980), 1—14.

[6] Sun J. G.: On numerical methods for robust pole assignment in control system design, *J. Comp. Math.*, **5**: 2 (1987), 119—134.

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

[8] Sun J. G.: Matrix Perturbation Analysis, Science Press, Beijing, 1986.