EIGENVALUES AND EIGENVECTORS OF A MATRIX DEPENDENT ON SEVERAL PARAMETERS*

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

Abstract

This paper describes a method for investigating the analyticity and for obtaining perturbation expansions of eigenvalues and eigenvectors of a matrix dependent on several parameters. Some of results of this paper provide justification of the applications of the Newton method for inverse matrix eigenvalue problems.

§ 1. Introduction

Although investigation for the analyticity of eigenvalues and eigenvectors has a long history[4,6,8,10], relatively little attention has been paid to the analyticity and perturbation expansion of eigenvalues and eigenvectors when the matrix depends analytically on several parameters, and we feel that this problem should be discussed whenever one is trying to treat inverse matrix eigenvalue problems (ref. [2, 5, 9]). The object of this paper is to describe a method for investigating the analyticity and for obtaining perturbation expansions of eigenvalues and eigenvectors of a matrix dependent on several complex or real parameters. Our approach is on the basis of the theory of implicit functions and matrix operations. Some of our results provide justification of the applications of the Newton method for inverse matrix eigenvalue problems.

Notation. The symbol $\mathbb{C}^{m\times n}$ denotes the set of complex $m\times n$ matrices and $\mathbb{R}^{m\times n}$ the set of real $m\times n$ matrices, $\mathbb{C}^n=\mathbb{C}^{n\times 1}$, $\mathbb{C}=\mathbb{C}^1$, $\mathbb{R}^n=\mathbb{R}^{n\times 1}$ and $\mathbb{R}=\mathbb{R}^1$. $\lambda(A)$ stand for the set of eigenvalues of a metrix A. $I^{(n)}$ is the $n\times n$ identity matrix, and O is the null matrix. The superscript H is for conjugate transpose, and T for transpose. $\|x\|$ denotes the usual Euclidean vector norm of x and $\|A\|$ denotes the spectral norm of A.

Before all we cite the following implicit function theorems.

Theorem 1.1[1, p. 89]. *If the complex-value functions*

$$f_i(\xi_1, \cdots, \xi_s, \eta_1, \cdots, \eta_t), \quad i=1, \cdots, s$$

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

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

then the equations

$$f_i(\xi_1, \cdots, \xi_s, \eta_1, \cdots, \eta_t)=0, \quad i=1, \cdots, s$$

* Received January 3, 1985.

have a unique solution

$$\xi_i = g_i(\eta_1, \cdots, \eta_t), \quad i=1, \cdots, s$$

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

Theorem 1.2[3, p. 277]. *If the real-value functions*

$$f_i(\xi_1, \cdots, \xi_s, \eta_1, \cdots, \eta_t), \quad i=1, \cdots, s$$

are real analytic functions of $s+t$ *real variables in some neighbourhood of the origin of* $\mathbb{R}^{s+t}$, *if* $f_i(0, 0)=0$, $i=1, \cdots, s$, *and if*

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

then the equations

$$f_i(\xi_1, \cdots, \xi_s, \eta_1, \cdots, \eta_t) = 0, \quad i=1, \cdots, s$$

have a unique solution

$$\xi_i = g_i(\eta_1, \cdots, \eta_t), \quad i=1, \cdots, s$$

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

§ 2. Simple Eigenvalues and Associated Eigenvectors

Let $p = (p_1, \cdots, p_N)^T \in \mathbb{C}^N$ (or $\mathbb{R}^N$), and $A(p) = (\alpha_{ij}(p)) \in \mathbb{C}^{n\times n}$ (or $\mathbb{R}^{n\times n}$) be an analytic (or real analytic) function in some neighbourhood $B(0)$ of the origin. i.e.,

$$A(p) = A(0) + E(p), \quad E(p) = (\varepsilon_{ij}(p)),$$

where

$$\varepsilon_{ij}(p) = \sum_{r=1}^{\infty} \sum_{\Sigma t = r} \alpha_{t_1, \cdots, t_N}^{(i,j)} p_1^{t_1} \cdots p_N^{t_N}, \quad 1 \leqslant i, j \leqslant N, \ p \in B(0)$$

and $\sum t = t_1 + \cdots + t_N$.

Suppose that λ is an eigenvalue of $A(0)$, then there exist vectors $x, y \in \mathbb{C}^n$ (or $\mathbb{R}^n$) such that

$$A(0)x = \lambda x, \quad y^T A(0) = \lambda y^T. \tag{2.1}$$

Such vectors, x, y will be called right and left eigenvectors of $A(0)$ corresponding to the eigenvalue λ respectively.

First applying Theorem 1.1 we prove the following theorem.

Theorem 2.1. *Let* $p \in \mathbb{C}^N$, *and* $A(p) \in \mathbb{C}^{n\times n}$ *be an analytic function of* p *in some neighbourhood* $B(0)$ *of the origin. Suppose that* λ_1 *is a simple eigenvalue of* $A(0)$, *and* x_1, y_1 *are associated eigenvectors satisfying the relations* (2.1) *and* $\|x_1\| = 1$, $y_1^T x_1 = 1$. *Then*

1) *there exists a simple eigenvalue* $\lambda_1(p)$ *of* $A(p)$ *which is an analytic function of* p *in some neighbourhood* B_0 *of the origin, and* $\lambda_1(0) = \lambda_1$;

2) *the right eigenvector* $x_1(p)$ *and left eigenvector* $y_1(p)$ *of* $A(p)$ *corresponding to* $\lambda_1(p)$ *may be defined to be analytic functions of* p *in* B_0, *and* $x_1(0) = x_1$, $y_1(0) = y_1$.

Proof. By the hypotheses there exist $X_2, Y_2 \in \mathbb{C}^{n\times(n-1)}$ such that

$$X=(x_1,\ X_2),\quad Y=(y_1,\ Y_2) \tag{2.2}$$

are nonsingular matrices satisfying

$$Y^TX=I \tag{2.3}$$

and

$$Y^TA(0)X=\begin{pmatrix}\lambda_1 & 0\\ 0 & A_2\end{pmatrix},\quad \lambda_1\overline{\in}\lambda(A_2). \tag{2.4}$$

We set

$$\widetilde{A}(p)=Y^TA(p)X=\begin{pmatrix}\tilde{a}_{11}(p) & \tilde{a}_{12}(p)^T\\ \tilde{a}_{21}(p) & \widetilde{A}_{22}(p)\end{pmatrix},\quad \tilde{a}_{11}(p)\in\mathbb{C}, \tag{2.5}$$

and introduce a vector-value function

$$(f_1(z,p),\ \cdots,\ f_{n-1}(z,p))^T=\tilde{a}_{21}(p)+[\widetilde{A}_{22}(p)-\tilde{a}_{11}(p)I]z-z\tilde{a}_{12}(p)^Tz,$$
$$z=(\zeta_1,\ \cdots,\ \zeta_{n-1})^T\in\mathbb{C}^{n-1},\ p\in\mathbb{C}^N. \tag{2.6}$$

Observe that the function $(f_1(z,p),\ \cdots,\ f_{n-1}(z,p))^T$ is analytic for $z\in\mathbb{C}^{n-1}$ and $p\in B(0)$, $f_i(0,0)=0$, $i=1,\ \cdots,\ n-1$, and

$$\left(\det\frac{\partial(f_1,\ \cdots,\ f_{n-1})}{\partial(\zeta_1,\ \cdots,\ \zeta_{n-1})}\right)_{z=0,\ p=0}=\det(A_2-\lambda_1I)\neq0,$$

hence by Theorem 1.1 the equation

$$(f_1(z,p),\ \cdots,\ f_{n-1}(z,p))=(0,\ \cdots,\ 0) \tag{2.7}$$

has a unique analytic solution $z=z(p)$ in some neighbourhood $B_1(0)\ (\subseteq B(0))$ of the origin, and $z(0)=0$.

Combining (2.5)—(2.7) we see that $z(p)$ satisfies

$$\widetilde{A}(p)\begin{pmatrix}1\\ z(p)\end{pmatrix}=[\tilde{a}_{11}(p)+\tilde{a}_{12}(p)^Tz(p)]\begin{pmatrix}1\\ z(p)\end{pmatrix},\quad p\in B_1(0). \tag{2.8}$$

Let

$$\lambda_1(p)=\tilde{a}_{11}(p)+\tilde{a}_{12}(p)^Tz(p),\quad x_1(p)=x_1+X_2z(p), \tag{2.9}$$

then from (2.8), (2.9) and (2.5) it follows that

$$A(p)x_1(p)=\lambda_1(p)x_1(p),\quad p\in B_1(0). \tag{2.10}$$

The expressions (2.9) show that the functions $\lambda_1(p)$ and $x_1(p)$ are analytic in $B_1(0)$ and satisfy

$$\lambda_1(0)=\lambda_1,\quad x_1(0)=x_1. \tag{2.11}$$

Moreover observe that λ_1 is a simple eigenvalue of $A(0)$, by Ostrowski Theorem[10, p.63] $\lambda_1(p)$ is a simple eigenvalue of $A(p)$ provided that $\|p\|_2$ is sufficiently small. We may consider the neighbourhood $B_1(0)$ sufficient small such that $\lambda_1(p)$ is a simple eigenvalue of $A(p)$ for $p\in B_1(0)$.

Similarly, we introduce a vector-value function

$$(g_1(w,p),\ \cdots,\ g_{n-1}(w,p))=\tilde{a}_{12}(p)^T+w^T[\widetilde{A}_{22}(p)-\tilde{a}_{11}(p)I]-w^T\tilde{a}_{21}(p)w^T,$$
$$w=(\omega_1,\ \cdots,\ \omega_{n-1})^T\in\mathbb{C}^{n-1},\ p\in\mathbb{C}^N. \tag{2.12}$$

Observe that the function $(g_1(w,p),\ \cdots,\ g_{n-1}(w,p))$ is analytic for $w\in\mathbb{C}^{n-1}$ and $p\in B(0)$, $g_i(0,0)=0$, $i=1,\ \cdots,\ n-1$ and

$$\left(\det \frac{\partial(g_1, \cdots, g_{n-1})}{\partial(\omega_1, \cdots, \omega_{n-1})}\right)_{w=0,\ p=0} = \det(A_2 - \lambda_1 I) \neq 0,$$

hence by Theorem 1.1 the equation

$$(g_1(w, p), \cdots, g_{n-1}(w, p)) = (0, \cdots, 0) \tag{2.13}$$

has a unique analytic solution $w = w(p)$ in some neighbourhood $B_2(0)(\subseteq B(0))$ of the origin, and $w(0) = 0$.

Combining (2.5), (2.12) and (2.13) we see that the analytic function $w(p)$ satisfies

$$\begin{pmatrix} 1 \\ w(p) \end{pmatrix}^T \tilde{A}(p) = [\tilde{a}_{11}(p) + w(p)^T \tilde{a}_{21}(p)] \begin{pmatrix} 1 \\ w(p) \end{pmatrix}^T, \quad p \in B_2(0). \tag{2.14}$$

Let

$$\lambda(p) = \tilde{a}_{11}(p) + w(p)^T \tilde{a}_{21}(p), \quad y_1(p) = y_1 + Y_2 w(p), \tag{2.15}$$

then from (2.14), (2.15) and (2.5) it follows that

$$y_1(p)^T A(p) = \lambda(p) y_1(p)^T, \quad p \in B_2(0),$$

i.e. $y_1(p)$ is a left eigenvector of $A(p)$ corresponding to the eigenvalue $\lambda(p)$.

The expressions (2.15) show that the functions $\lambda(p)$ and $y_1(p)$ are analytic in $B_2(0)$, and satisfy $\lambda(0) = \lambda_1$ and $y_1(0) = y_1$. Observe that λ_1 is a simple eigenvalue of $A(0)$, by Ostrowski Theorem[10, p. 63] $\lambda(p)$ is a simple eigenvalue of $A(p)$ and $\lambda(p) \equiv \lambda_1(p)$ provided that $\|p\|_2$ is sufficiently small. Hence there is a neighbourhood $B_0 \subseteq B_1(0) \cap B_2(0)$ of the origin such that

$$\lambda(p) \equiv \lambda_1(p), \quad |w(p)^T z(p)| < 1, \quad p \in B_0.$$

Thus the analytic function $y_1(p)$ satisfies

$$y_1(0) = y_1, \quad y_1(p)^T x_1(p) \neq 0, \quad y_1(p)^T A(p) = \lambda_1(p) y_1(p)^T, \quad p \in B_0. \quad \blacksquare \tag{2.16}$$

Theorem 2.2. *Assume that the hypotheses of Theorem 2.1 are valid. Then there are the following formulas for the simple eigenvalue $\lambda_1(p)$ and the associated eigenvectors $x_1(p)$ and $y_1(p)$ defined by (2.9) and (2.15):*

$$\left(\frac{\partial \lambda_1(p)}{\partial p_i}\right)_{p=0} = y_1^T \left(\frac{\partial A(p)}{\partial p_i}\right)_{p=0} x_1, \tag{2.17}$$

$$\left(\frac{\partial x_1(p)}{\partial p_i}\right)_{p=0} = X_2[\lambda_1 I - Y_2^T A(0) X_2]^{-1} Y_2^T \left(\frac{\partial A(p)}{\partial p_i}\right)_{p=0} x_1, \tag{2.18}$$

$$\left(\frac{\partial y_1(p)}{\partial p_i}\right)^T_{p=0} = y_1^T \left(\frac{\partial A(p)}{\partial p_i}\right)_{p=0} X_2[\lambda_1 I - Y_2^T A(0) X_2]^{-1} Y_2^T, \tag{2.19}$$

where $i = 1, \cdots, N$ in (2.17)—(2.19), X_2 and Y_2 are defined by (2.2)—(2.4). Moreover, for $h = 2, 3, \cdots$ and for any natural numbers $i_1, \cdots, i_h \in \{1, \cdots, N\}$ (some of the numbers $i_1, \cdots, i_h$ may be equal) the following formulas are valid:

$$\left(\frac{\partial^h \lambda_1(p)}{\partial p_{i_1} \cdots \partial p_{i_h}}\right)_{p=0} = y_1^T \left(\frac{\partial^h A(p)}{\partial p_{i_1} \cdots \partial p_{i_h}}\right)_{p=0} x_1$$
$$+ \sum_{l=1}^{h-1} \sum_{r=0}^{h-l} \sum_{(i_1, \cdots, i_{h-1})_{l-1, r}} \left(\frac{\partial^r y_1(p)}{\partial p_{i_l} \cdots \partial p_{i_{l+r-1}}}\right)^T_{p=0} \left[\frac{\partial^{l-1}}{\partial p_{i_1} \cdots \partial p_{i_{l-1}}} \left(\frac{\partial A(p)}{\partial p_{i_h}}\right)\right]_{p=0} \left(\frac{\partial^{h-l-r} x_1(p)}{\partial p_{i_{l+r}} \cdots \partial p_{i_{h-1}}}\right)_{p=0}, \tag{2.20}$$

$$\left(\frac{\partial^h x_1(p)}{\partial p_{i_1}\cdots\partial p_{i_h}}\right)_{p=0} = X_2[\lambda_1 I - Y_2^T A(0) X_2]^{-1} Y_2^T$$
$$\times\left\{\sum_{l=1}^{h-1}\sum_{(t_1,\cdots,t_h)_l}\left(\frac{\partial^l}{\partial p_{t_1}\cdots\partial p_{t_l}}[A(p)-\lambda_1(p)I]\right)_{p=0}\left(\frac{\partial^{h-l}x_1(p)}{\partial p_{t_{l+1}}\cdots\partial p_{t_h}}\right)_{p=0}+\left(\frac{\partial^h A(p)}{\partial p_{i_1}\cdots\partial p_{i_h}}\right)_{p=0}x_1\right\}, \tag{2.21}$$

$$\left(\frac{\partial^h y_1(p)}{\partial p_{i_1}\cdots\partial p_{i_h}}\right)^T_{p=0} = -\left\{y_1^T\left(\frac{\partial^h A(p)}{\partial p_{i_1}\cdots\partial p_{i_h}}\right)_{p=0}\right.$$
$$\left.+\sum_{l=1}^{h-1}\sum_{(t_1,\cdots,t_h)_l}\left(\frac{\partial^{h-l}y_1(p)}{\partial p_{t_{l+1}}\cdots\partial p_{t_h}}\right)^T_{p=0}\left(\frac{\partial^l}{\partial p_{t_1}\cdots\partial p_{t_l}}[A(p)-\lambda_1(p)I]\right)_{p=0}\right\}$$
$$\times X_2[\lambda_1 I - Y_2^T A(0) X_2]^{-1} Y_2^T. \tag{2.22}$$

The meaning of the symbol $\sum_{(t_1,\cdots,t_{h-1})_{l-1,r}}$ *in* (2.20) *is that: consider the numbers* $i_2, \cdots, i_h$ *as* $h-1$ *different indexes, we select arbitrarily* $l-1$ *indexes* $t_1, \cdots, t_{l-1}$ $(1\leqslant l\leqslant h-1)$ *from the set* $\{i_2, \cdots, i_h\}$, *then select arbitrarily* r *indexes* $t_l, \cdots, t_{l+r-1}$ $(r\leqslant h-l)$ *from the set* $\{i_2, \cdots, i_h\}\setminus\{t_1, \cdots, t_{l-1}\}$, *and write the residual indexes as* $t_{l+r}, \cdots, t_{h-1}$; *the summation* $\sum_{(t_1,\cdots,t_{h-1})_{l-1,r}}$ *is taken with respect to all such combinations of the set* $\{i_2, \cdots, i_h\}$. *The meaning of the symbol* $\sum_{(t_1,\cdots,t_h)_l}$ *in* (2.21) *and* (2.22) *is that: consider the numbers* $i_1, \cdots, i_h$ *as* h *different indexes, we select arbitrarily* l *indexes* $t_1, \cdots, t_l$ $(1\leqslant l\leqslant h-1)$ *from the set* $\{i_1, \cdots, i_h\}$, *and write the residual indexes as* $t_{l+1}, \cdots, t_h$; *the summation* $\sum_{(t_1,\cdots,t_h)_l}$ *is taken with respect to all such combinations of the set* $\{i_1, \cdots, i_h\}$.

Proof.

1) By Theorem 2.1

$$\lambda_1(p) = \frac{y_1(p)^T A(p) x_1(p)}{y_1(p)^T x_1(p)}, \quad p\in B_0. \tag{2.23}$$

Utilizing the relations (2.10) and (2.16), from (2.23) it follows that

$$\frac{\partial\lambda_1(p)}{\partial p_i} = \frac{y_1(p)^T\dfrac{\partial A(p)}{\partial p_i}x_1(p)}{y_1(p)^T x_1(p)} \quad \text{for } p\in B_0,\ i=1, \cdots, N. \tag{2.24}$$

Combining with (2.11) and (2.16) we get the formula (2.17) at once.

2) From (2.10) we obtain

$$[\lambda_1(p)I - A(p)]\frac{\partial x_1(p)}{\partial p_i} = \left[\frac{\partial A(p)}{\partial p_i} - \frac{\partial\lambda_1(p)}{\partial p_i}I\right]x_1(p), \quad p\in B_0. \tag{2.25}$$

Moreover, from (2.4), (2.9) and $z(0)=0$

$$[\lambda_1 I - A(0)]\left(\frac{\partial x_1(p)}{\partial p_i}\right)_{p=0} = X\begin{pmatrix}0 & 0\\ 0 & \lambda_1 I - A_2\end{pmatrix}\begin{pmatrix}0\\ \left(\dfrac{\partial z(p)}{\partial p_i}\right)_{p=0}\end{pmatrix}. \tag{2.26}$$

Since $\lambda_1\overline{\in}\lambda(A_2)$, it follows from (2.3), (2.25) and (2.26) that

$$\begin{pmatrix}0\\ \left(\dfrac{\partial z(p)}{\partial p_i}\right)_{p=0}\end{pmatrix} = \begin{pmatrix}0 & 0\\ 0 & (\lambda_1 I - A_2)^{-1}\end{pmatrix}Y^T\left[\left(\frac{\partial A(p)}{\partial p_i}\right)_{p=0} - \left(\frac{\partial\lambda_1(p)}{\partial p_i}\right)_{p=0}I\right]x_1,$$

and thus

$$\left(\frac{\partial x_1(p)}{\partial p_i}\right)_{p=0}=X\begin{pmatrix}0\\ \left(\frac{\partial z(p)}{\partial p_i}\right)_{p=0}\end{pmatrix}=X_2(\lambda_1 I-A_2)^{-1}Y_2^T\left(\frac{\partial A(p)}{\partial p_i}\right)_{p=0}x_1.$$

Combining with $A_2=Y_2^TA(0)X_2$ (see (2.2) and (2.4)) we get the formula (2.18).

3) From (2.16) we obtain

$$\left(\frac{\partial y_1(p)}{\partial p_i}\right)^T[\lambda_1(p)I-A(p)]=y_1(p)^T\left[\frac{\partial A(p)}{\partial p_i}-\frac{\partial\lambda_1(p)}{\partial p_i}I\right],\quad p\in B_0.\quad (2.27)$$

Moreover, from (2.4), (2.15) and $w(0)=0$

$$\left(\frac{\partial y_1(p)}{\partial p_i}\right)^T_{p=0}[\lambda_1 I-A(0)]=\begin{pmatrix}0\\ \left(\frac{\partial w(p)}{\partial p_i}\right)_{p=0}\end{pmatrix}^T\begin{pmatrix}0 & 0\\ 0 & \lambda_1 I-A_2\end{pmatrix}Y^T.\quad (2.28)$$

Since $\lambda_1\overline{\in}\lambda(A_2)$, it follows from (2.3), (2.27) and (2.28) that

$$\begin{pmatrix}0\\ \left(\frac{\partial w(p)}{\partial p_i}\right)_{p=0}\end{pmatrix}^T=y_1^T\left[\left(\frac{\partial A(p)}{\partial p_i}\right)_{p=0}-\left(\frac{\partial\lambda_1(p)}{\partial p_i}\right)_{p=0}I\right]X\begin{pmatrix}0 & 0\\ 0 & (\lambda_1 I-A_2)^{-1}\end{pmatrix},$$

and thus

$$\left(\frac{\partial y_1(p)}{\partial p_i}\right)^T_{p=0}=\begin{pmatrix}0\\ \left(\frac{\partial w(p)}{\partial p_i}\right)_{p=0}\end{pmatrix}^TY^T=y_1^T\left(\frac{\partial A(p)}{\partial p_i}\right)_{p=0}X_2(\lambda_1 I-A_2)^{-1}Y_2^T.$$

Combining with $A_2=Y_2^TA(0)X_2$ we get the formula (2.19).

4) With the same argument as above, from (2.24), (2.25) and (2.27) we can deduce the formulas (2.20), (2.21) and (2.22) respectively. ∎

Now applying Theorem 1.2 we prove the following theorem similar to Theorem 2.1.

Theorem 2.3. *Let $p\in\mathbb{R}^N$, and $A(p)=A(p)^T\in\mathbb{R}^{n\times n}$ be a real analytic function of p in some neighbourhood $B(0)$ of the origin. Suppose that λ_1 is a simple eigenvalue of $A(0)$, and x_1 is an associated eigenvector satisfying*

$$A(0)x_1=\lambda_1x_1,\quad \|x_1\|=1.$$

Then

1) *there exist a simple eigenvalue $\lambda_1(p)$ of $A(p)$ which is a real analytic function of p in some neighbourhood B_0 of the origin, and $\lambda_1(0)=\lambda_1$;*

2) *the eigenvector $x_1(p)$ of $A(p)$ corresponding to $\lambda_1(p)$ may be defined to be a real analytic function of p in B_0, and $x_1(0)=x_1$.*

Proof. By the hypotheses there exists $X_2\in\mathbb{R}^{n\times(n-1)}$ such that $X=(x_1, X_2)$ is an orthogonal matrix, and

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

Let

$$\tilde{A}(p)=X^TA(p)X=\begin{pmatrix}\tilde{a}_{11}(p) & \tilde{a}_{21}(p)^T\\ \tilde{a}_{21}(p) & \tilde{A}_{22}(p)\end{pmatrix},\quad \tilde{a}_{11}(p)\in\mathbb{R}.$$

Taking an approach similar to that in the proof of Theorem 2.1 we can derive that there is a real analytic function $z(p)\in\mathbb{R}^{n-1}$ in some neighbourhood B_0 of the origin such that $z(0)=0$, $\|z(p)\|<1$ for $p\in B_0$, and the real analytic functions

$$\lambda_1(p)=\tilde{a}_{11}(p)+\tilde{a}_{21}(p)^T z(p), \quad x_1(p)=x_1+X_2 z(p) \tag{2.29}$$

satisfy $\lambda_1(0)=\lambda_1$, $x_1(0)=x_1$ and

$$A(p)x_1(p)=\lambda_1(p)x_1(p) \quad \text{for} \quad p\in B_0. \ \blacksquare$$

By Theorem 2.3, with the same argument as that in the proof of Theorem 2.2 we obtain the following theorem.

Theorem 2.4. *Assume that the hypotheses of Theorem 2.3 are valid. Then there are the following formulas for the simple eigenvalue $\lambda_1(p)$ and the associated eigenvector $x_1(p)$ defined by (2.29):*

$$\left(\frac{\partial\lambda_1(p)}{\partial p_i}\right)_{p=0}=x_1^T\left(\frac{\partial A(p)}{\partial p_i}\right)_{p=0}x_1, \tag{2.30}$$

$$\left(\frac{\partial x_1(p)}{\partial p_i}\right)_{p=0}=X_2[\lambda_1 I-X_2^T A(0)X_2]^{-1}X_2^T\left(\frac{\partial A(p)}{\partial p_i}\right)_{p=0}x_1, \tag{2.31}$$

where $i=1, \cdots, N$ in (2.30) and (2.31), $(x_1, X_2)\in\mathbb{R}^{n\times n}$ is an orthogonal matrix. Moreover, for $h=2, 3, \cdots$ and for any natural numbers $i_1, \cdots, i_h\in\{1, \cdots, N\}$ (some of the numbers $i_1, \cdots, i_h$ may be equal) the following formulas are valid:

$$\left(\frac{\partial^h\lambda_1(p)}{\partial p_{i_1}\cdots\partial p_{i_h}}\right)_{p=0}=x_1^T\left(\frac{\partial^h A(p)}{\partial p_{i_1}\cdots\partial p_{i_h}}\right)_{p=0}x_1$$
$$+\sum_{l=1}^{h-1}\sum_{r=0}^{h-l}\sum_{(t_1,\cdots,t_{h-1})_{l-1,r}}\left(\frac{\partial^r x_1(p)}{\partial p_{t_l}\cdots\partial p_{l+r-1}}\right)_{p=0}^T\left[\frac{\partial^{l-1}}{\partial p_{t_1}\cdots\partial p_{t_{l-1}}}\left(\frac{\partial A(p)}{\partial p_{i_1}}\right)\right]_{p=0}\left(\frac{\partial^{h-l-r}x_1(p)}{\partial p_{t_{l+r}}\cdots\partial p_{t_{h-1}}}\right)_{p=0}, \tag{2.32}$$

$$\left(\frac{\partial^h x_1(p)}{\partial p_{i_1}\cdots\partial p_{i_h}}\right)_{p=0}=X_2[\lambda_1 I-X_2^T A(0)X_2]^{-1}X_2^T$$
$$\times\left\{\sum_{l=1}^{h-1}\sum_{(t_1,\cdots,t_h)_l}\left(\frac{\partial^l}{\partial p_{t_1}\cdots\partial p_{t_l}}[A(p)-\lambda_1(p)I]\right)_{p=0}\left(\frac{\partial^{h-l}x_1(p)}{\partial p_{t_{l+1}}\cdots\partial p_{t_h}}\right)_{p=0}+\left(\frac{\partial^h A(p)}{\partial p_{i_1}\cdots\partial p_{i_h}}\right)_{p=0}x_1\right\}, \tag{2.33}$$

where the symbol $\sum_{(t_2,\cdots,t_h)_{l-1,r}}$ in (2.32) and the symbol $\sum_{(t_1,\cdots,t_h)_l}$ in (2.33) have been explained in Theorem 2.2.

§ 3. Some Applications

The following inverse eigenvalue problems arise frequently in applied physics and structural design:

(I) For given complex $n\times n$ matrices $A_0, A_1, \cdots, A_N$ and $(\lambda_1^*, \cdots, \lambda_m^*)^T\in\mathbb{C}^m$, find $p^*=(p_1^*, \cdots, p_N^*)^T\in\mathbb{C}^N$ such that the composite matrix A, expressible as

$$A=A_0+p_1^*A_1+\cdots+p_N^*A_N \tag{3.1}$$

has eigenvalues $\lambda_1^*, \cdots, \lambda_m^*$ (ref. [5]).

(II) For given real symmetric $n\times n$ matrices $A_0, A_1, \cdots, A_N$ and $(\lambda_1^*, \cdots, \lambda_m^*)^T\in\mathbb{R}^m$, find $p^*=(p_1^*, \cdots, p_N^*)^T\in\mathbb{R}^N$ such that the composite matrix A, expressible as (3.1) has eigenvalues $\lambda_1^*, \cdots, \lambda_m^*$ (ref. [2], [9]).

From the theorems of § 1 we can derive the following Corollary 3.1 and Corollary 3.2 which provide justification of the Newton method for solving the above mentioned problems (I) and (II).

Corollary 3.1. Suppose that $A_0, A_1, \cdots, A_N \in \mathbb{C}^{n\times n}$, $\lambda(A_0)=\{\lambda_s\}_{s=1}^n$, $\lambda_k \neq \lambda_l$ for $k\neq l$ and $1\leqslant k, l\leqslant n$, x_s and y_s are right and left eigenvectors of A_0 corresponding to λ_s respectively, and

$$\|x_s\|=1,\quad y_s^T x_s=1,\quad x_s, y_s\in\mathbb{C}^n,\quad s=1,\cdots,n. \tag{3.2}$$

Let

$$\hat{X}_s=(x_1,\cdots,x_{s-1},x_{s+1},\cdots,x_n),\quad \hat{Y}_s=(y_1,\cdots,y_{s-1},y_{s+1},\cdots,y_n), \tag{3.3}$$

$$\phi_s=\|\hat{X}_s\|,\ \psi_s=\|\hat{Y}_s\|,\ \eta_s=\|y_s\|,\ \delta_s=\min_{\substack{1\leqslant t\leqslant n\\ t\neq s}}|\lambda_s-\lambda_t| \tag{3.4}$$

for $s=1,\cdots,n$, and

$$\alpha_i=\|A_i\|,\quad i=1,\cdots,N. \tag{3.5}$$

Then the matrix

$$A(p)=A_0+\sum_{i=1}^N p_iA_i \quad\text{for } p_1,\cdots,p_N\in\mathbb{C}$$

has analytic eigenvalues $\{\lambda_s(p)\}_{s=1}^n$ and associated analytic eigenvectors $\{x_s(p)\}_{s=1}^n$ and $\{y_s(p)\}_{s=1}^n$ in some neighbourhood of the origin such that $\lambda_s(0)=\lambda_s$, $x_s(0)=x_s$ and $y_s(0)=y_s$ for $s=1,\cdots,n$, and they satisfy the following estimations:

$$\left|\left(\frac{\partial\lambda_s(p)}{\partial p_i}\right)_{p=0}\right|\leqslant\eta_s\alpha_i, \tag{3.6}$$

$$\left|\left(\frac{\partial x_s(p)}{\partial p_i}\right)_{p=0}\right|\leqslant\frac{\phi_s\psi_s\alpha_i}{\delta_s},\quad \left|\left(\frac{\partial y_s(p)}{\partial p_i}\right)_{p=0}\right|\leqslant\frac{\eta_s\phi_s\psi_s\alpha_i}{\delta_s}, \tag{3.7}$$

$$\left|\left(\frac{\partial^2\lambda_s(p)}{\partial p_i\partial p_j}\right)_{p=0}\right|\leqslant\frac{2\eta_s\phi_s\psi_s\alpha_i\alpha_j}{\delta_s}, \tag{3.8}$$

$$\left|\left(\frac{\partial^2 x_s(p)}{\partial p_i\partial p_j}\right)_{p=0}\right|\leqslant\frac{2(\eta_s+\phi_s\psi_s)\phi_s\psi_s\alpha_i\alpha_j}{\delta_s^2},\quad \left|\left(\frac{\partial^2 y_s(p)}{\partial p_i\partial p_j}\right)_{p=0}\right|\leqslant\frac{2\eta_s(\eta_s+\phi_s\psi_s)\phi_s\psi_s\alpha_i\alpha_j}{\delta_s^2}, \tag{3.9}$$

$$\left|\left(\frac{\partial^3\lambda_s(p)}{\partial p_i\partial p_j\partial p_k}\right)_{p=0}\right|\leqslant\frac{2\eta_s(2\eta_s+3\phi_s\psi_s)\phi_s\psi_s\alpha_i\alpha_j\alpha_k}{\delta_s^2}, \tag{3.10}$$

where $s=1,\cdots,n$, and $1\leqslant i, j, k\leqslant N$.

Proof. Let

$$X=(x_1,\cdots,x_n),\ Y=(y_1,\cdots,y_n),\ \Lambda=\operatorname{diag}(\lambda_1,\cdots,\lambda_n).$$

Then from the hypotheses it follows that X and Y are non-singular, and

$$Y^TX=I,\quad Y^TA_0X=\Lambda.$$

Thus by Theorem 2.1 there exist analytic eigenvalues $\{\lambda_s(p)\}_{s=1}^n$ and associated analytic eigenvectors $\{x_s(p)\}_{s=1}^n$ and $\{y_s(p)\}_{s=1}^n$ satisfying $\lambda_s(0)=\lambda_s$, $x_s(0)=x_s$ and $y_s(0)=y_s$ for $s=1,\cdots,n$. Moreover, by Theorem 2.2 we obtain

$$\left(\frac{\partial\lambda_s(p)}{\partial p_i}\right)_{p=0}=y_s^TA_ix_s, \tag{3.11}$$

$$\left(\frac{\partial x_s(p)}{\partial p_i}\right)_{p=0}=\hat{X}_s(\lambda_sI-\hat{\Lambda}_s)^{-1}\hat{Y}_s^TA_ix_s,\quad \left(\frac{\partial y_s(p)}{\partial p_i}\right)_{p=0}^T=y_s^TA_i\hat{X}_s(\lambda_sI-\hat{\Lambda}_s)^{-1}\hat{Y}_s^T, \tag{3.12}$$

$$\left(\frac{\partial^2\lambda_s(p)}{\partial p_i\partial p_j}\right)_{p=0}=y_s^T[A_i\hat{X}_s(\lambda_sI-\hat{\Lambda}_s)^{-1}\hat{Y}_s^TA_j+A_j\hat{X}_s(\lambda_sI-\hat{\Lambda}_s)^{-1}\hat{Y}_s^TA_i]x_s, \tag{3.13}$$

$$\left(\frac{\partial^2 x_s(p)}{\partial p_i \partial p_j}\right)_{p=0} = \hat{X}_s(\lambda_s I-\hat{\Lambda}_s)^{-1}\hat{Y}_s^T[A_i\hat{X}_s(\lambda_s I-\hat{\Lambda}_s)^{-1}\hat{Y}_s^T A_j + A_j\hat{X}_s(\lambda_s I-\hat{\Lambda}_s)^{-1}\hat{Y}_s^T A_i]x_s$$
$$-\hat{X}_s(\lambda_s I-\hat{\Lambda}_s)^{-2}\hat{Y}_s^T(y_s^T A_i x_s A_j + y_s^T A_j x_s A_i)x_s, \tag{3.14}$$

$$\left(\frac{\partial^2 y_s(p)}{\partial p_i \partial p_j}\right)^T_{p=0} = y_s^T[A_j\hat{X}_s(\lambda_s I-\hat{\Lambda}_s)^{-1}\hat{Y}_s^T A_i + A_i\hat{X}_s(\lambda_s I-\hat{\Lambda}_s)^{-1}\hat{Y}_s^T A_j]\hat{X}_s(\lambda_s I-\hat{\Lambda}_s)^{-1}\hat{Y}_s^T$$
$$-y_s^T(y_s^T A_i x_s A_j + y_s^T A_j x_s A_i)\hat{X}_s(\lambda_s I-\hat{\Lambda}_s)^{-2}\hat{Y}_s^T \tag{3.15}$$

and

$$\left(\frac{\partial^3 \lambda_s(p)}{\partial p_i \partial p_j \partial p_k}\right)_{p=0} = y_s^T\Big\{\sum_{(t_1,t_2,t_3)} A_{t_1}\hat{X}_s(\lambda_s I-\hat{\Lambda}_s)^{-1}\hat{Y}_s^T A_{t_2}\hat{X}_s(\lambda_s I-\hat{\Lambda}_s)^{-1}\hat{Y}_s^T A_{t_3}$$
$$-\sum_{(l_1,l_2)} y_s^T A_{l_1} x_s [A_i\hat{X}_s(\lambda_s I-\hat{\Lambda}_s)^{-2}\hat{Y}_s^T A_{l_2}$$
$$+A_{l_2}\hat{X}_s(\lambda_s I-\hat{\Lambda}_s)^{-2}\hat{Y}_s^T A_i]\Big\}x_s, \tag{3.16}$$

where x_s, y_s, $\hat{X}_s$ and $\hat{Y}_s$ are defined by (3.2) and (3.3), and

$$\hat{\Lambda}_s = \mathrm{diag}(\lambda_1, \cdots, \lambda_{s-1}, \lambda_{s+1}, \cdots, \lambda_n). \tag{3.17}$$

The meaning of the symbols $\sum_{(t_1,t_2,t_3)}$ and $\sum_{(l_1,l_2)}$ in (3.16) is that: consider the numbers i, j, k as three different indexes, the summation $\sum_{(t_1,t_2,t_3)}$ is taken with respect to all permutations of the set $\{i, j, k\}$; consider the numbers j, k as two different indexes, the summation $\sum_{(l_1,l_2)}$ is taken with respect to $(l_1, l_2)=(j, k)$ and $(l_1, l_2)=(k, j)$.

From the formulas (3.11)—(3.16) we get the estimations (3.6)—(3.10) at once. ▮

Corollary 3.2. Suppose that the symmetric matrices $A_0, A_1, \cdots, A_N \in \mathbb{R}^{n\times n}$, $\lambda(A_0)=\{\lambda_s\}_{s=1}^n$, $\lambda_k \neq \lambda_l$ for $k\neq l$ and $1\leq k, l\leq n$, and the associated unit eigenvectors of A_0 are $x_1, \cdots, x_n$.

Let

$$\hat{X}_s = (x_1, \cdots, x_{s-1}, x_{s+1}, \cdots, x_n), \quad \phi_s = \|\hat{X}_s\|, \quad \delta_s = \min_{\substack{1\leq t\leq n \\ t\neq s}} |\lambda_s - \lambda_t| \tag{3.18}$$

for $s=1, \cdots, n$, and

$$\alpha_i = \|A_i\|, \quad i=1, \cdots, N. \tag{3.19}$$

Then the matrix

$$A(p) = A_0 + \sum_{i=1}^N p_i A_i \quad \text{for} \quad p_1, \cdots, p_N \in \mathbb{R}$$

has real analytic eigenvalues $\{\lambda_s(p)\}_{s=1}^n$ and associated real analytic eigenvectors $\{x_s(p)\}_{s=1}^n$ in some neighbourhood of the origin such that $\lambda_s(0)=\lambda_s$ and $x_s(0)=x_s$ for $s=1, \cdots, n$, and they satisfy the following estimations:

$$\left|\left(\frac{\partial \lambda_s(p)}{\partial p_i}\right)_{p=0}\right| \leq \alpha_i, \quad \left|\left(\frac{\partial x_s(p)}{\partial p_i}\right)_{p=0}\right| \leq \frac{\alpha_i}{\delta_s}, \tag{3.20}$$

$$\left|\left(\frac{\partial^2 \lambda_s(p)}{\partial p_i \partial p_j}\right)_{p=0}\right| \leq \frac{2\alpha_i\alpha_j}{\delta_s}, \quad \left|\left(\frac{\partial^2 x_s(p)}{\partial p_i \partial p_j}\right)_{p=0}\right| \leq \frac{4\alpha_i\alpha_j}{\delta_s^2}, \tag{3.21}$$

$$\left|\left(\frac{\partial^3 \lambda_s(p)}{\partial p_i \partial p_j \partial p_k}\right)_{p=0}\right| \leq \frac{10\alpha_i\alpha_j\alpha_k}{\delta_s^2}, \tag{3.22}$$

where $s=1, \cdots, n$ and $1\leq i, j, k\leq N$.

Proof. Let

$$X=(x_1,\cdots,x_n),\quad \Lambda=\mathrm{diag}(\lambda_1,\cdots,\lambda_n).$$

Then from the hypotheses it follows that X satisfies

$$X^TX=I,\quad X^TA_0X=\Lambda.$$

Thus by Theorem 2.3 there exist real analytic eigenvalues $\{\lambda_s(p)\}_{s=1}^n$ and associated real analytic eigenvectors $\{x_s(p)\}_{s=1}^n$ satisfying $\lambda_s(0)=\lambda_s$ and $x_s(0)=x_s$ for $s=1,\cdots,n$. Moreover, by Theorem 2.4 we obtain formulas for $\left(\frac{\partial\lambda_s(p)}{\partial p_i}\right)_{p=0}$, $\left(\frac{\partial x_s(p)}{\partial p_i}\right)_{p=0}$, $\left(\frac{\partial^2\lambda_s(p)}{\partial p_i\partial p_j}\right)_{p=0}$, $\left(\frac{\partial^2 x_s(p)}{\partial p_i\partial p_j}\right)_{p=0}$ and $\left(\frac{\partial^3\lambda_s(p)}{\partial p_i\partial p_j\partial p_k}\right)_{p=0}$ which are similar to those in (3.11)—(3.14) and (3.16), and then we get the estimations (3.20)—(3.22). ∎

Remark 3.1. For the eigenvalues $\{\lambda_s(p)\}_{s=1}^n$ described in Corollary 3.2, Bohte[2] has obtained

$$\left|\left(\frac{\partial^2\lambda_s(p)}{\partial p_i\partial p_j}\right)_{p=0}\right|\leqslant\frac{2(n-1)M^2}{\delta},\quad s=1,\cdots,n,\ 1\leqslant i,j\leqslant N, \tag{3.23}$$

where

$$M=\max_{1\leqslant r\leqslant n}\|A_r\|,\quad \delta=\min_{\substack{1\leqslant s,t\leqslant n\\ t\neq s}}|\lambda_s-\lambda_t|. \tag{3.24}$$

But from (3.21) we obtain

$$\left|\left(\frac{\partial^2\lambda_s(p)}{\partial p_i\partial p_j}\right)_{p=0}\right|\leqslant\frac{2M^2}{\delta},\quad s=1,\cdots,n,\ 1\leqslant i,j\leqslant N. \tag{3.25}$$

Obviously, our estimations (3.25) are much better than Bohte's estimations (3.23).

Besides the above mentioned applications we can derive perturbation expansions for simple eigenvalues and associated eigenvectors of a matrix.

Corollary 3.3. Let $A\in\mathbb{C}^{n\times n}$. Suppose that λ_1 is a simple eigenvalue of A, and x_1, y_1 are associated right and left eigenvectors satisfying $\|x_1\|=1$ and $y_1^Tx_1=1$. Let

$$\varepsilon=(\varepsilon_{11},\cdots,\varepsilon_{1n},\varepsilon_{21},\cdots,\varepsilon_{2n},\cdots,\varepsilon_{n1},\cdots,\varepsilon_{nn})^T\in\mathbb{C}^{n^2}$$

and

$$A(\varepsilon)=A+E,\quad E=(\varepsilon_{ij}).$$

Then there are an analytic simple eigenvalue $\lambda_1(\varepsilon)$ and associated analytic eigenvectors $x_1(\varepsilon)$ and $y_1(\varepsilon)$ of $A(\varepsilon)$ in some neighbourhood of the origin, and

$$\lambda_1(\varepsilon)=\lambda_1+y_1^TEx_1+y_1^TEX_2(\lambda_1I-A_2)^{-1}Y_2^TEx_1+O(\|E\|^3), \tag{3.26}$$

$$\begin{aligned}x_1(\varepsilon)=x_1&+X_2(\lambda_1I-A_2)^{-1}Y_2^TEx_1\\&+X_2(\lambda_1I-A_2)^{-1}(Y_2^TEX_2-y_1^TEx_1I)(\lambda_1I-A_2)^{-1}Y_2^TEx_1+O(\|E\|^3)\end{aligned} \tag{3.27}$$

and

$$\begin{aligned}y_1(\varepsilon)^T=y_1^T&+y_1^TEX_2(\lambda_1I-A_2)^{-1}Y_2^T\\&+y_1^TEX_2(\lambda_1I-A_2)^{-1}(Y_2^TEX_2-y_1^TEx_1I)(\lambda_1I-A_2)^{-1}Y_2^T+O(\|E\|^3),\end{aligned} \tag{3.28}$$

where X_2, Y_2 and A_2 are determined as follows:

$$X=(x_1,X_2),\ Y=(y_1,Y_2),\ Y^TX=I,\ Y^TAX=\begin{pmatrix}\lambda_1&0\\0&A_2\end{pmatrix}. \tag{3.29}$$

Proof The existence of the matrices X_2 and Y_2 satisfying (3.29) is clear. Let e_i denote the i-th column of the identity $I^{(n)}$, and

$$A_{ij}=e_ie_j^T,\quad 1\leqslant i,\ j\leqslant n.$$

Then

$$E=\sum_{i,j=1}^{n}\varepsilon_{ij}A_{ij}. \tag{3.30}$$

By Theorem 2.1 there are an analytic simple eigenvalue $\lambda_1(\varepsilon)$ of $A(\varepsilon)$ and associated analytic eigenvectors $x_1(\varepsilon)$ and $y_1(\varepsilon)$ satisfying $\lambda_1(0)=\lambda_1$, $x_1(0)=x_1$ and $y_1(0)=y_1$ provided that $\|\varepsilon\|$ is sufficiently small. Utilizing the formulas (2.17)—(2.19) and (2.20) with $h=2$ we obtain

$$\left(\frac{\partial\lambda_1(\varepsilon)}{\partial\varepsilon_{ij}}\right)_{\varepsilon=0}=y_1^TA_{ij}x_1 \tag{3.31}$$

and

$$\left(\frac{\partial^2\lambda_1(\varepsilon)}{\partial\varepsilon_{ij}\partial\lambda_{kl}}\right)_{\varepsilon=0}=y_1^T[A_{kl}X_2(\lambda_1I-A_2)^{-1}Y_2^TA_{ij}+A_{ij}X_2(\lambda_1I-A_2)^{-1}Y_2^TA_{kl}]x_1. \tag{3.32}$$

Substituting (3.31) and (3.32) into the second order expansion of $\lambda_1(\varepsilon)$

$$\lambda_1(\varepsilon)=\lambda_1+\sum_{i,j=1}^{n}\left(\frac{\partial\lambda_1(\varepsilon)}{\partial\varepsilon_{ij}}\right)_{\varepsilon=0}\varepsilon_{ij}+\frac{1}{2}\sum_{i,j,k,l=1}^{n}\left(\frac{\partial^2\lambda_1(\varepsilon)}{\partial\varepsilon_{ij}\partial\varepsilon_{kl}}\right)_{\varepsilon=0}\varepsilon_{ij}\varepsilon_{kl}+O(\|\varepsilon\|^3),$$

and combining with (3.30) we get (3.26).

From the formulas (2.17)—(2.19) and (2.21) with $h=2$ we obtain

$$\left(\frac{\partial x_1(\varepsilon)}{\partial\varepsilon_{ij}}\right)_{\varepsilon=0}=X_2(\lambda_1I-A_2)^{-1}Y_2^TA_{ij}x_1 \tag{3.33}$$

and

$$\left(\frac{\partial^2x_1(\varepsilon)}{\partial\varepsilon_{ij}\partial\varepsilon_{kl}}\right)_{\varepsilon=0}=X_2(\lambda_1I-A_2)^{-1}Y_2^TA_{i,j,k,l}x_1, \tag{3.34}$$

where

$$\begin{aligned}A_{i,j,k,l}=&(A_{ij}-y_1^TA_{ij}x_1I)X_2(\lambda_1I-A_2)^{-1}Y_2^TA_{kl}\\&+(A_{kl}-y_1^TA_{kl}x_1I)X_2(\lambda_1I-A_2)^{-1}Y_2^TA_{ij}.\end{aligned} \tag{3.35}$$

Substituting (3.33)—(3.35) into the second order expansion of $x_1(\varepsilon)$ we get (3.27).

With the similar argument we can obtain the expansion (3.28). ∎

Similarly, from Theorem 2.3 and Theorem 2.4 we obtain the following corollary.

Corollary 3.4. Let $A\in\mathbb{R}^{n\times n}$ be symmetric. Suppose that λ_1 is a simple eigenvalue of A, and x_1 is an associated unit eigenvector. Let

$$\varepsilon=(\varepsilon_{11},\cdots,\varepsilon_{1n},\varepsilon_{22},\cdots,\varepsilon_{2n},\cdots,\varepsilon_{nn})^T\in\mathbb{R}^{n(n+1)/2}$$

and

$$A(\varepsilon)=A+E,\quad E=(\varepsilon_{ij}),\quad \varepsilon_{ji}=\varepsilon_{ij}\quad\text{for}\quad 1\leqslant i<j\leqslant n.$$

Then there are a real analytic simple eigenvalue $\lambda_1(\varepsilon)$ and associated real analytic eigenvector $x_1(\varepsilon)$ of $A(\varepsilon)$ in some neighbourhood of the origin, and

$$\lambda_1(\varepsilon)=\lambda_1+x_1^TEx_1+x_1^TEX_2(\lambda_1I-A_2)^{-1}X_2^TEx_1+O(\|E\|^3) \tag{3.36}$$

and

$$x_1(\varepsilon)=x_1+X_2(\lambda_1 I-A_2)^{-1}X_2^T E x_1 + X_2(\lambda_1 I-A_2)^{-1}(X_2^T E X_2-x_1^T E x_1 I)(\lambda_1 I-A_2)^{-1}X_2^T E x_1+O(\|E\|^3), \tag{3.37}$$

where X_2 and A_2 are determined as follows:

$$X=(x_1,\ X_2),\quad X^TX=I,\quad X^TAX=\begin{pmatrix}\lambda_1 & 0\\ 0 & A_2\end{pmatrix}.$$

§ 4. On Multiple Eigenvalues

We consider a simple example (see [9, p. 606], [2, p. 386]):

$$A(p)=\begin{pmatrix}1+2p_1+2p_2 & p_2\\ p_2 & 1+2p_2\end{pmatrix},\quad p=(p_1,\ p_2)^T\in\mathbb{R}^2. \tag{4.1}$$

$A(0)$ has eigenvalue 1 with multiplicity 2, and the eigenvalues of $A(p)$ are

$$\lambda_1(p)=1+p_1+2p_2+\sqrt{p_1^2+p_2^2},\quad \lambda_2(p)=1+p_1+2p_2-\sqrt{p_1^2+p_2^2}. \tag{4.2}$$

It is easy to verify that the functions $\lambda_1(p)$ and $\lambda_2(p)$ are not differentiable at $p_1=p_2=0$, but it is worth-while to point out that we have

$$\lambda\Big(\Big(\frac{\partial A(p)}{\partial p_i}\Big)_{p=0}\Big)=\Big\{\Big(\frac{\partial\lambda_s(p)}{\partial p_i}\Big)_{p=0,\,p_i=+0}\Big\}_{s=1}^2\cup\Big\{\Big(\frac{\partial\lambda_s(p)}{\partial p_i}\Big)_{p=0,\,p_i=-0}\Big\}_{s=1}^2,\ i=1,\ 2. \tag{4.3}$$

The following theorem clarifies that the relations (4.3) are of universal significance.

Theorem 4.1. *Let $p=(p_1,\ \cdots,\ p_N)^T\in\mathbb{R}^N$, $A(p)$ be a real analytic function of p in some neighbourhood $B(0)$ of the origin, and $A(p)^T=A(p)$. Suppose that there is an orthogonal matrix $X\in\mathbb{R}^{n\times n}$ such that*

$$X=(\underset{r}{X_1},\ \underset{n-r}{X_2}),\quad X^TA(0)X=\begin{pmatrix}\lambda_1 I^{(r)} & 0\\ 0 & A_2\end{pmatrix},\ \lambda_1\overline{\in}\lambda(A_2). \tag{4.4}$$

Then $A(p)$ has eigenvalues $\lambda_1(p),\ \cdots,\ \lambda_r(p)$ satisfying

$$\lambda_s(0)=\lambda_1,\quad s=1,\ \cdots,\ r \tag{4.5}$$

and

$$\lambda\Big(X_1^T\Big(\frac{\partial A(p)}{\partial p_i}\Big)_{p=0}X_1\Big)=\Big\{\Big(\frac{\partial\lambda_s(p)}{\partial p_i}\Big)_{p=0,\,p_i=+0}\Big\}_{s=0}^r\cup\Big\{\Big(\frac{\partial\lambda_s(p)}{\partial p_i}\Big)_{p=0,\,p_i=-0}\Big\}_{s=0}^r,\quad i=1,\cdots,N. \tag{4.6}$$

Proof. Let

$$\widetilde{A}(p)=X^TA(p)X=\begin{pmatrix}\widetilde{A}_{11}(p) & \widetilde{A}_{21}(p)^T\\ \widetilde{A}_{21}(p) & \widetilde{A}_{22}(p)\end{pmatrix},\quad \widetilde{A}_{11}(p)\in\mathbb{R}^{r\times r}. \tag{4.7}$$

We introduce a matrix-value function

$$F(Z,\ p)=\widetilde{A}_{21}(p)+\widetilde{A}_{22}(p)Z-Z\widetilde{A}_{11}(p)-Z\widetilde{A}_{21}(p)^TZ,\quad Z=(\zeta_{ij})\in\mathbb{R}^{(n-r)\times r},\quad p=(p_1,\ \cdots,\ p_N)^T\in\mathbb{R}^N. \tag{4.8}$$

Observe that the function $F(Z,\ p)=(f_{ij}(Z,\ p))$ is analytic for $Z\in\mathbb{R}^{(n-r)\times r}$ and

$p\in B(0)$ satisfying

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

and

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

where $\otimes$ denotes the Kronecker product symbol (see [7, p. 8—9]). Hence by Theorem 1.2 the equation

$$F(Z,p)=0 \tag{4.9}$$

has a unique real analytic solution $Z=Z(p)$ in some neighbourhood of the origin, and $Z(0)=0$.

Combining (4.7)—(4.9) we can verify that the real analytic function $Z(p)$ satisfies

$$\tilde{A}(p)\begin{pmatrix} I^{(r)}\\ Z(p)\end{pmatrix}[I+Z(p)^T Z(p)]^{-\frac{1}{2}}=\begin{pmatrix} I^{(r)}\\ Z(p)\end{pmatrix}[I+Z(p)^T Z(p)]^{-\frac{1}{2}}A_1(p), \tag{4.10}$$

where

$$A_1(p)=[I+Z(p)^T Z(p)]^{-\frac{1}{2}}\tilde{A}_1(p)[I+Z(p)^T Z(p)]^{-\frac{1}{2}}, \tag{4.11}$$

and

$$\tilde{A}_1(p)=\tilde{A}_{11}(p)+Z(p)^T\tilde{A}_{21}(p)+\tilde{A}_{21}(p)^T Z(p)+Z(p)^T\tilde{A}_{22}(p)Z(p).$$

Let

$$X_1(p)=X\begin{pmatrix} I^{(r)}\\ Z(p)\end{pmatrix}[I+Z(p)^T Z(p)]^{-\frac{1}{2}}, \tag{4.12}$$

then from (4.7) and (4.10)—(4.12) it follows that

$$A(p)X_1(p)=X_1(p)A_1(p), \tag{4.13}$$

where $A_1(p)$ and $X_1(p)$ are real analytic functions in some neighbourhood of the origin satisfying

$$X_1(p)^T X_1(p)\equiv I^{(r)},\quad X_1(0)=X_1,\quad A_1(0)=\lambda_1 I^{(r)} \tag{4.14}$$

and

$$\lambda(A_1(p))=\{\lambda_1(p),\cdots,\lambda_r(p)\}\subseteq\lambda(A(p)).$$

By (4.13)

$$A_1(p)=X_1(p)^T A(p)X_1(p). \tag{4.15}$$

From (4.13)—(4.15) we obtain

$$\left(\frac{\partial A_1(p)}{\partial p_i}\right)_{p=0}=X_1^T\left(\frac{\partial A(p)}{\partial p_i}\right)_{p=0}X_1,\quad i=1,\cdots,N. \tag{4.16}$$

Let i be an any fixed index from $1,\cdots,N$. since $\left(\frac{\partial A_1(p)}{\partial p_i}\right)^T_{p=0}=\left(\frac{\partial A_1(p)}{\partial p_i}\right)_{p=0}$ there is a real orthogonal matrix $Q_i\in\mathbb{R}^{r\times r}$ such that

$$\Omega_i=Q_i^T\left(\frac{\partial A_1(p)}{\partial p_i}\right)_{p=0}Q_i=\operatorname{diag}(\omega_i^{(1)},\cdots,\omega_i^{(r)}),\quad \omega_i^{(s)}\in\mathbb{R},\ s=1,\cdots,r. \tag{4.17}$$

Hence we have

$$[Q_i^T A_1(p) Q_i]_{p=(0,\cdots,0,p_i,0,\cdots,0)^T} = \begin{pmatrix} \lambda_1+\omega_i^{(1)}p_i & & 0 \\ & \ddots & \\ 0 & & \lambda_1+\omega_i^{(r)}p_i \end{pmatrix},$$

and so

$$\{\lambda_1+\omega_i^{(s)}p_i\}_{s=1}^r = \{[\lambda_t(p)]_{p=(0,\cdots,0,p_i,0,\cdots,0)^T}\}_{t=1}^r.$$

Therefore, for any $\lambda_t(p)$ $(1\leqslant t\leqslant r)$ there are $\omega_i^{(t')}$ and $\omega_i^{(t'')}$ (t' and t'' may be the same index) such that

$$\lambda_t((0,\cdots,0,p_i,0,\cdots,0)^T) = \begin{cases} \lambda_1+\omega_i^{(t')}p_i, & p_i>0, \\ \lambda_1+\omega_i^{(t'')}p_i, & p_i<0, \end{cases}$$

thus

$$\left(\frac{\partial\lambda_t(p)}{\partial p_i}\right)_{p=0} = \begin{cases} \omega_i^{(t')}, & p_i=+0, \\ \omega_i^{(t'')}, & p_i=-0. \end{cases} \tag{4.18}$$

On the other hand, for any $\omega_i^{(s)}$ there are $\lambda_{s'}(p)$ and $\lambda_{s''}(p)$ (s' and s'' may be the same index) such that

$$\lambda_1+\omega_i^{(s)}p_i = \begin{cases} \lambda_{s'}((0,\cdots,0,p_i,0,\cdots,0)^T), & p_i>0, \\ \lambda_{s''}((0,\cdots,0,p_i,0,\cdots,0)^T), & p_i<0, \end{cases}$$

thus

$$\omega_i^{(s)} = \begin{cases} \left(\frac{\partial\lambda_{s'}(p)}{\partial p_i}\right)_{p=0,\,p_i=+0}, \\ \left(\frac{\partial\lambda_{s''}(p)}{\partial p_i}\right)_{p=0,\,p_i=-0}. \end{cases} \tag{4.19}$$

Combining (4.16)—(4.19) we obtain the relations (4.6). ▮

References

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

[2] Z. Bohte, Numerical solution of the inverse algebraic eigenvalue problem, *Comput. J.*, **10** (1968), 385—388.

[3] J. Dieudonné, Éléments d'Analyse, 1. Fondements de l'Analyse Moderne, Gauthier-Villars, Paris, 1968.

[4] T. Kato, Perturbation Theory for Linear Operators, Springer-Verlag, New York, 1966.

[5] F. W. Biegler-König, A Newton iteration process for inverse eigenvalue problems, *Numer. Math.*, **37** (1981), 349—354.

[6] P. Lancaster, On eigenvalues of matrices dependent on a parameter, *Numer. Math.*, **6** (1964), 377—387.

[7] M. Marcus, H. Minc, A Survey of Matrix Theory and Matrix Inequalities, Allyn and Bacon, Boston, 1964.

[8] F. Rollich, Störungstheorie der Spektralzerlegung, *Math. Ann.*, **113** (1937), 600—619.

[9] J. Y. Wang, B. S. Garbow, A numerical method for solving inverse real symmetric eigenvalue problems, *SIAM J. Sci. Stat. Comput.*, **4** (1983), 45—51.

[10] J. H. Wilkinson, The Algebraic Eigenvalue Problem, Clarendon Press, Oxford, 1965