SIAM J. NUMER. ANAL.
Vol. 40, No. 6, pp. 2352–2367

STRONG SEMISMOOTHNESS OF EIGENVALUES OF SYMMETRIC MATRICES AND ITS APPLICATION TO INVERSE EIGENVALUE PROBLEMS*

DEFENG SUN† AND JIE SUN‡

Abstract. It is well known that the eigenvalues of a real symmetric matrix are not everywhere differentiable. A classical result of Ky Fan states that each eigenvalue of a symmetric matrix is the difference of two convex functions, which implies that the eigenvalues are semismooth functions. Based on a recent result of the authors, it is further proved in this paper that the eigenvalues of a symmetric matrix are strongly semismooth everywhere. As an application, it is demonstrated how this result can be used to analyze the quadratic convergence of Newton's method for solving inverse eigenvalue problems (IEPs) and generalized IEPs with multiple eigenvalues.

Key words. symmetric matrices, eigenvalues, strong semismoothness, Newton's method, inverse eigenvalue problems, quadratic convergence

AMS subject classifications. 65F15, 65F18, 65H10, 65H17

PII. S0036142901393814

1. Introduction.

The theory of semismooth functions developed in the last decade has been successful in analyzing the quadratic convergence of Newton's method for nondifferentiable (nonsmooth) equations; it is well received by the optimization community, but is perhaps not well known by researchers in numerical analysis. In this paper we take the inverse eigenvalue problem (IEP) as an example to show how this theory can be used in analyzing matrix-related equations. For applications of the IEP the interested reader is referred to the paper of Friedland, Nocedal, and Overton [10], the book of Xu [27], and the references therein. For general nonsmooth analysis involving eigenvalues of symmetric matrices and a survey on eigenvalue optimization, see Lewis [12] and Lewis and Overton [13], respectively.

Let $\mathcal{S}$ be the linear space of symmetric matrices of size n. Let $A : \Re^n \to \mathcal{S}$ be continuously differentiable. Given n real numbers $\{\lambda_i^*\}_{i=1}^n$, which are arranged in the decreasing order $\lambda_1^* \geq \cdots \geq \lambda_n^*$, the IEP is to find a vector $c^* \in \Re^n$ such that $\lambda_i(A(c^*)) = \lambda_i^*$ for $i = 1, \ldots, n$. A typical choice for $A(c)$ is

$$A(c) = A_0 + \sum_{j=1}^{n} c_j A_j, \tag{1}$$

where $A_0, A_1, \ldots, A_n \in \mathcal{S}$. In this case, $A(c)$ is an affine function of c.

*Received by the editors August 16, 2001; accepted for publication (in revised form) July 10, 2002; published electronically January 14, 2003.
http://www.siam.org/journals/sinum/40-6/39381.html

†Department of Mathematics, National University of Singapore, Singapore 117543, Republic of Singapore (matsundf@nus.edu.sg). This author's research was partially supported by grant R146-000-035-101 from the National University of Singapore.

‡School of Business and Singapore-MIT Alliance, National University of Singapore, Singapore 119620, Republic of Singapore (jsun@nus.edu.sg). This author's research was partially supported by grant R314-000-028/042-112 from the National University of Singapore and a grant from the Singapore-MIT Alliance.

Define $F : \Re^n \to \Re^n$ by

$$F(c) = \begin{bmatrix} \lambda_1(A(c)) - \lambda_1^* \\ \vdots \\ \lambda_n(A(c)) - \lambda_n^* \end{bmatrix}. \tag{2}$$

Then the IEP is equivalent to finding $c^* \in \Re^n$ to be a solution of the following equation:

$$F(c) = 0. \tag{3}$$

Of course, there are other ways to formulate the IEP as a system of equations. For instance, we may solve $F(c) = 0$, where

$$F(c) = \begin{bmatrix} \det(A(c) - \lambda_1^* I) \\ \vdots \\ \det(A(c) - \lambda_n^* I) \end{bmatrix}. \tag{4}$$

A Newton method was proposed by Biegler-König [2] for model (4), which generalizes an algorithm of Lancaster [11]. However, as analyzed by Friedland, Nocedal, and Overton [10], model (2) seems to be always preferred over model (4) both from theoretical and computational points of view. Thus, we concentrate on model (2) in this paper. The convergence theory we are going to present is based on a property of F called strong semismoothness (defined later). It is well known that for $X \in \mathcal{S}$ the eigenvalues of X, as functions of X, are not everywhere differentiable. However, we shall show that they are strongly semismooth and therefore quadratic convergence of Newton's method is a natural result when applied to equations involving eigenvalues. In doing so, we also give a constructive proof for a difficult result of Chen and Tseng [4] on upper semicontinuity of a set-valued mapping of orthogonal matrices.

The concept of semismoothness of functionals was originally studied by Mifflin [14] while strong semismoothness was introduced by Qi and Sun in [18] for vector valued functions. Recently, both concepts are further extended to matrix valued functions [24]. Generally speaking, *strong semismoothness* of an equation is tied with quadratic convergence of the Newton method applied to the equation and *semismoothness* corresponds to superlinear convergence. It was shown that smooth functions, piecewise smooth functions, and convex and concave functions are semismooth functions. They are not, however, necessarily strongly semismooth functions.

To see the motivation of this paper more clearly, let us consider the following example:

$$X = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix},$$

where x_1, x_2, and x_3 are parameters. In this case, we have

$$\lambda_1(X) = \frac{x_1 + x_3 + \sqrt{(x_1 - x_3)^2 + 4x_2^2}}{2} \quad \text{and} \quad \lambda_2(X) = \frac{x_1 + x_3 - \sqrt{(x_1 - x_3)^2 + 4x_2^2}}{2}. \tag{5}$$

Since $\lambda_1(\cdot)$ and $\lambda_2(\cdot)$ are not differentiable at X with $x_1 = x_3$ and $x_2 = 0$, a gradient-dependent numerical method (e.g., Newton's method) may get into trouble when hitting those points. In addition, theoretical analysis gets tricky without differentiability. Further inspection reveals that $\lambda_1(\cdot)$ is a convex function and $\lambda_2(\cdot)$ is a concave function. Hence, both of them are semismooth functions and a nonsmooth version of Newton's method [18] might be applied to equations containing $\lambda_1(\cdot)$ and $\lambda_2(\cdot)$. This should be not a coincidence. Let $f_m(X)$ be the sum of m largest eigenvalues of X. Then, Ky Fan's maximum principle [8, 1] says that for each $i = 1, \ldots, n$, $f_i(\cdot)$ is a convex function. This result implies that

- $\lambda_1(\cdot)$ is a convex function and $\lambda_n(\cdot)$ is a concave function; and,
- for $i = 2, \ldots, n-1$, $\lambda_i(\cdot)$ is the difference of two convex functions.

Since convex and concave functions are semismooth and the difference of two semismooth functions is still a semismooth function [14], Ky Fan's result shows that $\lambda_1(\cdot), \ldots, \lambda_n(\cdot)$ are all semismooth functions. It is therefore expected, when applying the nonsmooth Newton method to IEPs, the convergence rate is at least superlinear. A more interesting question is, Are all $\lambda_1(\cdot), \ldots, \lambda_n(\cdot)$ *strongly* semismooth functions (therefore implying quadratic convergence)? In this paper, based on a recent result of the authors [24], we will give an affirmative answer to the above question.

The organization of this paper is as follows. Some basic facts on semismoothness are presented in section 2. Some nonsmooth versions of the Newton method, which we call *relative generalized Newton methods,* are introduced in section 3. Section 4 concentrates on showing the strong semismoothness of eigenvalues of a symmetric matrix. The quadratic convergence of the relative generalized Newton methods for IEPs and generalized IEPs is proved in section 5. Section 6 gives a summary and a few possible future research topics.

Some notations to be used are as follows.

- $\mathcal{S}$ is the set of real symmetric matrices; $\mathcal{O}$ is the set of all $n \times n$ orthogonal matrices.
- A superscript "T" represents the transpose of matrices and vectors. For a matrix M, $M_{i\cdot}$, and $M_{\cdot j}$ represent the ith row and jth column of M, respectively.
- Unless otherwise specified, all vector norms are 2-norms and matrix norms are Frobenius norms: $\|M\| := \text{trace}\,(M^TM)^{1/2}$.
- A diagonal matrix is written as $\text{diag}\,(\beta_1, \ldots, \beta_n)$ and a block-diagonal matrix is denoted by $\text{diag}\,(B_1, \ldots, B_s)$, where $B_1, \ldots, B_s$ are matrices.
- The eigenvalues of $X \in \mathcal{S}$ is designated by $\lambda_i(X)$, $i = 1, \ldots, n$, and $\Lambda(X) := \text{diag}\,(\lambda_1(X), \ldots, \lambda_n(X))$.
- We write $X = O(\alpha)$ (respectively, $o(\alpha)$) if $\|X\|/|\alpha|$ is uniformly bounded (respectively, tends to zero) as $\alpha \to 0$.

2. Some basic facts on semismoothness.

2.1. Semismooth functions. Let $G : \Re^n \to \Re^m$ be a locally Lipschitz continuous function. We regard the $r \times r$ symmetric matrix space as a special case of $\Re^s$ with $s = r(r+1)/2$. Hence the discussions of this subsection apply to matrix variable and/or matrix valued functions as well.

According to Rademacher's theorem, G is differentiable almost everywhere. Let D_G be the set of differentiable points of G and let G' be the Jacobian of G whenever it exists. Denote

$$\partial_B G(x) := \{V \in \Re^{m\times n} \mid V = \lim_{x^k \to x} G'(x^k),\ x^k \in D_G\}.$$

Then Clarke's generalized Jacobian [5] is

$$\partial G(x) = \text{conv}\{\partial_B G(x)\}, \tag{6}$$

where "conv" stands for the convex hull in the usual sense of convex analysis [20].

DEFINITION 2.1. *Suppose that $G : \Re^n \to \Re^m$ is a locally Lipschitz continuous function. G is said to be semismooth at $x \in \Re^n$ if G is directionally differentiable at x and for any $V \in \partial G(x + \Delta x)$*

$$G(x + \Delta x) - G(x) - V(\Delta x) = o(\|\Delta x\|).$$

G is said to be p-order ($0 < p < \infty$) semismooth at x if G is semismooth at x and

$$G(x + \Delta x) - G(x) - V(\Delta x) = O(\|\Delta x\|^{1+p}). \tag{7}$$

In particular, G is called strongly semismooth at x if G is 1-order semismooth at x.

A function G is said to be a (strongly) semismooth function if it is (strongly) semismooth everywhere on $\Re^n$. It is shown that the composition of (strongly) semismooth functions is still a (strongly) semismooth function (see [14, 9]).

The next result [24, Theorem 3.7] provides a convenient tool for proving strong semismoothness.

THEOREM 2.2. *Suppose that $G : \Re^n \to \Re^m$ is locally Lipschitzian and directionally differentiable in a neighborhood of x. Then for any $p \in (0, \infty)$ the following two statements are equivalent:*

(a) *for any $V \in \partial G(x + \Delta x)$,*

$$G(x + \Delta x) - G(x) - V(\Delta x) = O(\|\Delta x\|^{1+p});$$

(b) *for any $x + \Delta x \in D_G$,*

$$G(x + \Delta x) - G(x) - G'(x + \Delta x)(\Delta x) = O(\|\Delta x\|^{1+p}). \tag{8}$$

2.2. Generalized Newton methods. Suppose that $G : \Re^n \to \Re^n$ is locally Lipschitz continuous. Based on $\partial G(x)$, Qi and Sun [18] proposed the following Newton method for solving $G(x) = 0$.

Generalized Newton method I. Given $x^0 \in \Re^n$, for $k = 0, 1, \ldots,$

$$x^{k+1} = x^k - V_k^{-1} G(x^k), \tag{9}$$

where $V_k \in \partial G(x^k)$.

The following convergence theorem for the generalized Newton method I is established in [18].

THEOREM 2.3. *Suppose that $G(x^*) = 0$. If all $V \in \partial G(x^*)$ are nonsingular and G is semismooth at x^*, then there exists a neighborhood $N(x^*)$ of x^* such that for any $x^0 \in N(x^*)$ the generalized Newton method* I *is well defined and is Q-superlinearly convergent. Moreover, if G is strongly semismooth at x^*, then* (9) *converges Q-quadratically.*

To relax the nonsingularity assumption on $\partial G(x^*)$, Qi [17] introduced the following method based on the concept of $\partial_B G(x)$.

Generalized Newton method II. Given $x^0 \in \Re^n$, for $k = 0, 1, \ldots,$

$$x^{k+1} = x^k - V_k^{-1} G(x^k), \tag{10}$$

where $V_k \in \partial_B G(x^k)$.

The convergence theorem for the generalized Newton method II is the same as Theorem 2.3 except that ∂G is replaced by $\partial_B G$.

Now, let us consider the following composite nonsmooth equation:

$$G(x) := \Phi(\Psi(x)) = 0, \tag{11}$$

where $\Phi : \Re^n \to \Re^m$ is nonsmooth but of special structure and $\Psi : \Re^m \to \Re^n$ is continuously differentiable. It is noted that neither $\partial G(x)$ nor $\partial_B G(x)$ is easy to compute even if $\partial\Phi(y)$, $\partial_B\Phi(y)$, and $\Psi'(x)$ are available. To circumvent the difficulty in computing $\partial G(x)$ and $\partial G_B(x)$, Potra, Qi, and Sun [16] introduced the following concept of generalized Jacobian:

$$\partial_Q G(x) = \partial_B \Phi(\Psi(x))\Psi'(x),$$

where "Q" stands for "quasi." We shall see in the later discussion that $\partial_Q G(x)$ is more convenient to compute than $\partial G(x)$ and $\partial_B G(x)$ for IEPs.

Generalized Newton method III. Given $x^0 \in \Re^n$, for $k = 0, 1, \ldots,$

$$x^{k+1} = x^k - V_k^{-1} G(x^k), \tag{12}$$

where $V_k \in \partial_Q G(x^k)$.

The following convergence theorem for the generalized Newton method III for solving (11) is proved in [16, Theorem 5.3].

THEOREM 2.4. *Suppose that G is defined by* (11) *and $G(x^*) = 0$. If all $V \in \partial_Q G(x^*)$ are nonsingular and Φ is semismooth at $\Psi(x^*)$, then there exists a neighborhood $N(x^*)$ of x^* such that for any $x^0 \in N(x^*)$ the generalized Newton method* III *is well defined and is Q-superlinearly convergent. Moreover, if Φ is strongly semismooth at $\Psi(x^*)$ and Ψ' is Lipschitz continuous around x^*, then* (12) *converges Q-quadratically.*

3. Relative generalized Newton methods. It should be noted that, apart from the semismoothness, another key assumption for the superlinear convergence of the generalized Newton methods I–III is the nonsingularity of $\partial G(x^*)$, $\partial_B G(x^*)$, or $\partial_Q G(x^*)$. However, this may not be satisfied in general for IEPs with multiple eigenvalues. In order to weaken the nonsingularity assumption on the generalized Jacobians, we shall introduce the concept of *relative* generalized Jacobians and the corresponding generalized Newton methods based on the concept of relative generalized gradient introduced by Clarke [5, p. 231].

Let S be a subset of $\Re^n$. For instance, in the context of matrix functions, S could represent the set of all nonsingular matrices. The S-relative generalized Jacobian $\partial|_S G(x)$ of G at x is defined by

$$\partial|_S G(x) := \{V \mid V \text{ is a limit of } V_i \in \partial G(y_i),\ y_i \in S, y_i \to x\}.$$

The following result can be proved in an analogous way to [5, Proposition 6.2.1]. We omit the details.

LEMMA 3.1. *Let G be Lipschitz continuous near x. Then we have the following:*

(a) *$\partial|_S G(x)$ is a compact subset of $\partial G(x)$.*

(b) *$\partial|_S G(x) = \partial G(x)$ if x lies in the interior part of S; $\partial|_S G(x) = \emptyset$ if $(x+\varepsilon B) \cap S = \emptyset$ for some $\varepsilon > 0$; and $\partial|_S G(x)$ is nonempty if $x \in \mathrm{cl}(S)$, the closure of S.*

(c) *$\partial|_S G(\cdot)$ is upper semicontinuous at x.*

Now, we can introduce our first relative generalized Newton method for solving $G(x) = 0$.

Relative generalized Newton method I. Given $x^0 \in \Re^n$, for $k = 0, 1, \ldots,$ and $x^k \in S$,

$$x^{k+1} = x^k - V_k^{-1} G(x^k), \tag{13}$$

where $V_k \in \partial|_S G(x^k)$.

In the following analysis, we assume that the relative generalized Newton method I does not find a solution of $G(x) = 0$ in a finite number of steps.

THEOREM 3.2. *Suppose that $G(x^*) = 0$ and $x^* \in \mathrm{cl}(S)$. If all $V \in \partial|_S G(x^*)$ are nonsingular and G is semismooth at x^*, then there exists a neighborhood $N(x^*)$ of x^* such that for any $x^0 \in N(x^*) \cap S$ the relative generalized Newton method* I *either stops in a finite number of steps with some $x^k \notin S$ or generates an infinite sequence $\{x^k\} \in N(x^*) \cap S$ and the whole sequence converges Q-superlinearly to x^*. Moreover, if G is strongly semismooth at x^*, then the rate of convergence is Q-quadratic.*

Proof. By using Lemma 3.1, there exist a neighborhood $N(x^*)$ of x^* and a positive number κ such that for any $x \in N(x^*) \cap S$, all $V \in \partial|_S G(x)$ are nonsingular and

$$\|V^{-1}\| \le \kappa. \tag{14}$$

Since G is semismooth at x^*, by shrinking $N(x^*)$ if necessary, we have for all $x \in N(x^*) \cap S$ and $V \in \partial|_S G(x)$,

$$\|G(x) - G(x^*) - V(x - x^*)\| \le \frac{1}{2\kappa}\|x - x^*\|. \tag{15}$$

By using (14) and (15), we have for $k = 0, 1, \ldots$ that

$$\begin{aligned}\|x^{k+1} - x^*\| &= \|x^k - V_k^{-1} G(x^k) - x^*\| \\ &= \|V_k^{-1}[G(x^k) - G(x^*) - V_k(x - x^*)]\| \\ &\le \|V_k^{-1}\| \|G(x^k) - G(x^*) - V_k(x - x^*)\| \\ &\le \tfrac{1}{2}\|x^k - x^*\|,\end{aligned}$$

which implies that if (13) does not stop at some step with $x^k \notin S$, then $\{x^k\} \in N(x^*) \cap S$ and the whole sequence converges to x^* linearly.

Next, suppose that (13) does not stop at some step with $x^k \notin S$. Since G is semismooth at x^* and $x^k \to x^*$, we have

$$G(x^k) - G(x^*) - V_k(x^k - x^*) = o(\|x^k - x^*\|),$$

which, together with (13), implies that

$$\begin{aligned}\|x^{k+1} - x^*\| &= \|x^k - V_k^{-1} G(x^k) - x^*\| \\ &= \|V_k^{-1}[G(x^k) - G(x^*) - V_k(x - x^*)]\| \\ &= O(\|G(x^k) - G(x^*) - V_k(x - x^*)\|) \\ &= o(\|x^k - x^*\|).\end{aligned}$$

This proves the superlinear convergence of $\{x^k\}$.

By the above argument, we can see that if G is strongly semismooth at x^*, then (13) either stops in finitely many steps with some $x^k \notin S$ or generates an infinite sequence $\{x^k\} \in N(x^*) \cap S$ and the whole sequence converges Q-quadratically to x^*. This completes the proof. $\square$

The proof of Theorem 3.2 might serve as an example to show the simplicity of the analysis of Newton's method by using the concept of (strong) semismoothness. Parallel to the definition of $\partial_B G(x)$ and $\partial_Q G(x)$, we define

$$\partial_B|_S G(x) := \{V|\ V \text{ is a limit of } V_i \in \partial_B G(y_i),\ y_i \in S,\ y_i \to x\}$$

and

$$\partial_Q|_S G(x) := \{V|\ V \text{ is a limit of } V_i \in \partial_Q G(y_i),\ y_i \in S,\ y_i \to x\}.$$

Similar to Lemma 3.1, we have the following lemma.

LEMMA 3.3. *Let G be Lipschitz continuous near x. Then we have the following:*

(a) *$\partial_B|_S G(x)$ and $\partial_Q|_S G(x)$ are compact subsets of $\partial_B G(x)$ and $\partial_Q G(x)$, respectively.*

(b) *$\partial_B|_S G(x) = \partial_B G(x)$ and $\partial_Q|_S G(x) = \partial_Q G(x)$, if x lies in the interior part of S; $\partial_B|_S G(x) = \partial_Q|_S G(x) = \emptyset$ if $(x + \varepsilon B) \cap S = \emptyset$ for some $\varepsilon > 0$; both $\partial_B|_S G(x)$ and $\partial_Q|_S G(x)$ are nonempty if $x \in \mathrm{cl}(S)$, the closure of S.*

(c) *$\partial_B|_S G(\cdot)$ and $\partial_Q|_S G(\cdot)$ are upper semicontinuous at x.*

Analogously, we define the second and third relative generalized Newton methods.

Relative generalized Newton method II (III). Given $x^0 \in \Re^n$, for $k = 0, 1, \ldots,$ and $x^k \in S$,

$$x^{k+1} = x^k - V_k^{-1} G(x^k), \tag{16}$$

where $V_k \in \partial_B|_S G(x^k)$ ($V_k \in \partial_Q|_S G(x^k)$ in method III).

The following theorem can be similarly proved by using Lemma 3.3 and the approach of proving Theorems 3.2. We omit the details.

THEOREM 3.4. *Suppose that $G(x^*) = 0$ and $x^* \in \mathrm{cl}(S)$. If all $V \in \partial_B|_S G(x^*)$ ($V \in \partial_Q|_S G(x^*)$ in method* III*) are nonsingular and G is semismooth at x^* (Φ is semismooth at $\Psi(x^*)$ in method* III*), then there exists a neighborhood $N(x^*)$ of x^* such that for any $x^0 \in N(x^*) \cap S$ the relative generalized Newton methods* II *and* III *either stop in a finite number of steps with some $x^k \notin S$ or generate an infinite sequence $\{x^k\} \in N(x^*) \cap S$ and the whole sequence converges Q-superlinearly to x^*. Moreover, if G (Φ in method* III*) is strongly semismooth at x^* (at $\Psi(x^*)$ and Ψ' is Lipschitz continuous around x^* in method* III*), then the rate of convergence is Q-quadratic.*

4. Strong semismoothness of eigenvalues. As a building block for applying relative generalized Newton methods, we shall prove the strong semismoothness of eigenvalues of symmetric matrices in this section. Suppose $X \in \mathcal{S}$. Then, there exists an orthogonal matrix $Q \in \mathcal{O}$ such that X satisfies

$$Q^T X Q = \Lambda(X) := \operatorname{diag}(\lambda_1(X), \ldots, \lambda_n(X)), \tag{17}$$

where $\lambda_1(X) \geq \cdots \geq \lambda_n(X)$.

We define a "configuration vector" K to distinguish different eigenvalues. Let

$$K := \{k_0, k_1, \ldots, k_l\} \tag{18}$$

with $1 = k_0 < k_1 < \cdots < k_l = n + 1$ such that there is a change of eigenvalues at k_i. Namely for $t = 1, \ldots, l$,

$$\lambda_s(X) = \lambda_{k_{t-1}}(X), \qquad s \in [k_{t-1}, k_t - 1], \tag{19}$$

where we use the simple notation $[k_{t-1}, k_t - 1]$ to represent the index set $\{k_{t-1}, k_{t-1} + 1, \ldots, k_t - 1\}$.

Let $H \in \mathcal{S}$ and let P (depending on H) be an orthogonal matrix such that

$$P^T(\Lambda(X) + H)P = \Lambda(Y) := \operatorname{diag}(\lambda_1(Y), \ldots, \lambda_n(Y)), \tag{20}$$

where $\lambda_1(Y) \geq \cdots \geq \lambda_n(Y)$ and $Y := \Lambda(X) + H$.

After the above preparation, we can state the following result, which was essentially proved in the derivation of Lemma 4.2 of [24].

LEMMA 4.1. *For any $H \in \mathcal{S}$ and $H \to 0$, we have*

$$P_{ij} = O(\|H\|), \qquad i,j = 1, \ldots, n,\ (i,j) \notin \bigcup_{t=1}^{l} \{[k_{t-1}, k_t - 1] \times [k_{t-1}, k_t - 1]\}. \tag{21}$$

Proof. It has been proved in the proof of Lemma 4.2 of [24] that (21) is true for any $H \in \mathcal{S}$ such that $\Lambda(X) + H$ is nonsingular and $H \to 0$.

Next, we prove that (21) is also true for the case that $\Lambda(X) + H$ is singular and $H \to 0$. It is easy to check that the conclusion of this lemma holds if $H = 0$. Hence, we can assume $H \neq 0$. Define

$$\lambda_{\min}(|Y|) = \min_{\lambda_i(Y) \neq 0} |\lambda_i(Y)| \quad \text{and} \quad \tilde{\Lambda} = \operatorname{diag}(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n),$$

where $|Y| := (Y^2)^{\frac{1}{2}}$ and for $i = 1, \ldots, n$

$$\tilde{\lambda}_i = \begin{cases} \lambda_i(Y) & \text{if } \lambda_i(Y) \neq 0, \\ \lambda_{\min}(|Y|) \min\{\frac{1}{2}, ||H||^2\} & \text{otherwise.} \end{cases}$$

Denote

$$\tilde{H} = P\tilde{\Lambda}P^T - \Lambda(X).$$

Hence, $P^T[\Lambda(X) + \tilde{H}]P = \tilde{\Lambda}$ is nonsingular. By noting the fact $\tilde{H} = H + O(\|H\|^2)$, it follows that (21) also holds for the case that $\Lambda(X) + H$ is singular and $H \to 0$. This completes the proof. □

Define a "truncated" matrix $W \in \Re^{n \times n}$ as follows:

$$W_{ij} = \begin{cases} P_{ij} & \text{if } (i,j) \in \bigcup_{t=1}^{l} \{[k_{t-1}, k_t - 1] \times [k_{t-1}, k_t - 1]\}, \\ 0 & \text{otherwise,} \end{cases} \qquad i,j = 1, \ldots, n. \tag{22}$$

Hence, from Lemma 4.1, we know that for any $H \to 0$,

$$W = P + O(\|H\|). \tag{23}$$

It is noted, however, that W may not be an orthogonal matrix but has a block-diagonal structure with each block corresponding to a set of identical eigenvalues of X. That is,

$$W = \operatorname{diag}(W_1, \ldots, W_l),$$

where

$$W_t = \left(P_{ij}\right)_{i,j=k_{t-1}}^{k_t - 1}, \qquad \text{for } t = 1, \ldots, l.$$

Since $P \in \mathcal{O}$, by using Lemma 4.1 and (22), for $t = 1, \ldots, l$ and $i, j = 1, \ldots, k_t - k_{t-1}$, we have for any $H \to 0$,

$$\|(W_t)_{\cdot j}\|^2 = 1 + O(\|H\|^2) \quad \text{and} \quad \langle (W_t)_{\cdot j}, (W_t)_{\cdot i} \rangle = O(\|H\|^2), \qquad i \neq j. \tag{24}$$

It is obvious from (24) that for any $H \in \mathcal{S}$ sufficiently close to 0 the columns of W_t are independent because

$$\sum_j \beta_j (W_t)_{\cdot j} = 0 \ \Rightarrow \ \beta_j [1 + O(\|H\|^2)] = O(\|H\|^2) \ \Rightarrow \ \beta_j = 0 \quad \forall\, j.$$

For each $t = 1, \ldots, l$, let $\tilde{P}_t$ be a matrix of the same order of W_t and be obtained by applying the Gram–Schmidt orthogonalization algorithm to each W_t; i.e., for $j = 1, \ldots, k_t - k_{t-1}$, let

$$(\tilde{W}_t)_{\cdot j} = (W_t)_{\cdot j} - \sum_{i=1}^{j-1} \langle (\tilde{P}_t)_{\cdot i}, (W_t)_{\cdot j} \rangle (\tilde{P}_t)_{\cdot i} \ \text{ and } \ (\tilde{P}_t)_{\cdot j} = (\tilde{W}_t)_{\cdot j} / \|(\tilde{W}_t)_{\cdot j}\|. \tag{25}$$

By (24) and (25), for $i, j = 1, \ldots, k_t - k_{t-1}$, $t = 1, \ldots, l$, we have for any $H \to 0$ that

$$\|(\tilde{P}_t)_{\cdot j}\|^2 = 1, \ \ (\tilde{P}_t)_{\cdot j} = (W_t)_{\cdot j} + O(\|H\|^2) \quad \text{and} \quad \langle (\tilde{P}_t)_{\cdot j}, (\tilde{P}_t)_{\cdot i} \rangle = 0, \quad i \neq j. \tag{26}$$

Denote

$$\tilde{P} = \operatorname{diag}(\tilde{P}_1, \ldots, \tilde{P}_l). \tag{27}$$

Then, we have the following lemma.

LEMMA 4.2. *For any $H \in \mathcal{S}$ sufficiently small, the matrix $\tilde{P}$ defined by* (27) *and* (25) *is an orthogonal matrix and satisfies*

$$\tilde{P}^T \Lambda(X) \tilde{P} = \Lambda(X). \tag{28}$$

Furthermore, for any $H \to 0$,

$$P = \tilde{P} + O(\|H\|). \tag{29}$$

Proof. By (26), we know that each $\tilde{P}_t$, $t = 1, \ldots, l$, is an orthogonal matrix. Since $\lambda_{k_{t-1}}(X) = \cdots = \lambda_{k_t - 1}(X)$, $t = 1, \ldots, l$, we have

$$\tilde{P}_t^{\,T} \operatorname{diag}(\lambda_{k_{t-1}}(X), \ldots, \lambda_{k_t - 1}(X)) \tilde{P}_t = \operatorname{diag}(\lambda_{k_{t-1}}(X), \ldots, \lambda_{k_t - 1}(X)).$$

Hence, $\tilde{P}$ is an orthogonal matrix and satisfies (28). By using (23) and (26), we directly obtain (29). This completes the proof. $\square$

For any $\Delta X \in \mathcal{S}$, let $U \in \mathcal{O}$ (depending on X and ΔX) be any orthogonal matrix such that

$$U^T (X + \Delta X) U = \Lambda(X + \Delta X) := \operatorname{diag}(\lambda_1(X + \Delta X), \ldots, \lambda_n(X + \Delta X)), \tag{30}$$

where $\lambda_1(X + \Delta X) \geq \cdots \geq \lambda_n(X + \Delta X)$.

By using the above lemma, we have the following result.

LEMMA 4.3. *For any $\Delta X \in \mathcal{S}$ sufficiently small and U satisfying* (30), *there exists a $V \in \mathcal{O}$ such that*

$$V^T X V = \Lambda(X) \quad \text{and} \quad U = V + O(\|\Delta X\|). \tag{31}$$

Proof. Let $P = Q^T U$ and $H = Q^T \Delta X Q$, where Q is defined in (17). Then, by Lemma 4.2, for any such defined P, there exists $\tilde{P} \in \mathcal{O}$ such that

$$\tilde{P}^T \Lambda(X) \tilde{P} = \Lambda(X)$$

and

$$P = \tilde{P} + O(\|H\|) = \tilde{P} + O(\|\Delta X\|).$$

Let $V = Q\tilde{P}$. Then $V \in \mathcal{O}$,

$$V^T X V = \tilde{P}^T Q^T X Q \tilde{P} = \tilde{P}^T \Lambda(X) \tilde{P} = \Lambda(X),$$

and for any $\Delta X \to 0$

$$U = V + O(\|\Delta X\|).$$

This completes the proof. ◻

A similar result to Lemma 4.3 has also been proved in [4] based on a so-called $\sin(\Theta)$ theorem in [21, Theorem 3.4]. The proof provided here is due to a direct comparison between entries of P and $\tilde{P}$ and it indeed furnishes an algorithm for computing V.

One direct result of Lemma 4.3 is that the (normalized) eigenvectors of symmetric matrices, though not continuous, are upper Lipschitz continuous. To see this, for any $Z \in \mathcal{S}$, let

$$\mathcal{U}(Z) := \{U \in \mathcal{O} |\ U^T Z U \text{ is diagonal}\},$$

and let

$$\mathcal{E} := \{M \in \mathcal{S} \mid\ |M_{i,j}| \leq 1,\ i,j = 1, \ldots, n\}.$$

PROPOSITION 4.4. *For any $X \in \mathcal{S}$, there exists a constant $\mu > 0$ such that*

$$\mathcal{U}(X + \Delta X) \subseteq \mathcal{U}(X) + \mu ||\Delta X|| \mathcal{E} \tag{32}$$

for all ΔX sufficiently small.

Proof. For any $U \in \mathcal{U}(X + \Delta X)$, there exists a diagonal matrix $D(X + \Delta X)$ such that

$$U^T (X + \Delta X) U = D(X + \Delta X).$$

Let $R \in \Re^{n \times n}$ be a permutation matrix such that

$$R D(X + \Delta X) R^T = \Lambda(X + \Delta X)$$

with $\lambda_1(X + \Delta X) \geq \cdots \geq \lambda_n(X + \Delta X)$. Let $\tilde{U} = U R^T$. Then we obtain $\tilde{U}^T(X + \Delta X)\tilde{U} = \Lambda(X + \Delta X)$. Hence, by Lemma 4.3, there exists a $\tilde{V} \in \mathcal{O}$ such that $\tilde{V}^T X \tilde{V} = \Lambda(X)$ and

$$\tilde{U} = \tilde{V} + O(\|\Delta X\|),$$

i.e.,

$$U = \tilde{V} R + O(\|\Delta X\|)$$

because $R^T = R^{-1}$ and $\|R^T\| = \sqrt{n}$. Let $V = \tilde{V}R$. Then

$$V^TV = R^T\tilde{V}^T\tilde{V}R = R^TR = I, \text{ and } V^TXV = (\tilde{V}R)^TX\tilde{V}R = R^T\Lambda(X)R$$

is a diagonal matrix. Hence, we have proved $V \in \mathcal{U}(X)$ and for $\Delta X \to 0$

$$U = V + O(\|\Delta X\|).$$

This implies that there exists a $\mu > 0$ such that (32) holds. □

In section 1 we have seen from an example of a two by two matrix that the eigenvalues are not differentiable if X has multiple eigenvalues. This can be easily extended to the general case: $\Lambda(\cdot)$ is not differentiable at X if X has multiple eigenvalues. On the other hand, by [26, pp. 66–68] and [25, Theorem 2.3] we know that if X has distinct eigenvalues, then $\Lambda(\cdot)$ is analytic in a neighborhood of X. Hence, we have the following lemma.

LEMMA 4.5. *$\Lambda(\cdot)$ is analytic in a neighborhood of X if and only if $X \in \mathcal{S}$ has distinct eigenvalues.*

Next, we cite a useful formula for the derivative of $\Lambda(X)$ when $X \in \mathcal{S}$ has distinct eigenvalues.

LEMMA 4.6 (see [21, p. 185, Corollary 2.4]). *For any $X \in \mathcal{S}$, if X has distinct eigenvalues, then $\Lambda(\cdot)$ is continuously differentiable at X and for any $\Delta X \in S$*

$$\lambda_i'(X)(\Delta X) = q_i(X)^T\Delta X q_i(X), \quad i = 1, \ldots, n. \tag{33}$$

For any $X \in \mathcal{S}$, let $Q(X) \in \mathcal{O}$ be such that $Q(X)^TXQ(X) = \Lambda(X)$ with $\lambda_1(X) \geq \cdots \geq \lambda_n(X)$. Define

$$q_i(X) = (Q(X))_{\cdot i}, \quad i = 1, \ldots, n.$$

The following result is our main theorem of this section.

THEOREM 4.7. *$\Lambda(\cdot)$ is a strongly semismooth function.*

Proof. By Ky Fan [8] and Mifflin [14], $\Lambda(\cdot)$ is a semismooth function. Thus, we only have to prove (8) with $p = 1$. Let $D_\Lambda = \{Y \in \mathcal{S} | \ Y \text{ has distinct eigenvalues}\}$. By Lemma 4.5, D_Λ is the subset of $\mathcal{S}$ on which Λ is continuously differentiable. Clearly, D_Λ is dense in $\mathcal{S}$.

Suppose that $X_0 \in \mathcal{S}$ is a given matrix. For any $X \in D_\Lambda$, denote $\Delta X = X - X_0$. For $i = 1, \ldots, n$ from $Xq_i(X) = \lambda_i(X)q_i(X)$, we have

$$q_i(X)^TX_0q_i(X) + q_i(X)^T\Delta X q_i(X) = \lambda_i(X). \tag{34}$$

By Lemma 4.3, there exists a $\mu > 0$ such that for any X sufficiently close to X_0 there exists a matrix $Q(X_0) \in \mathcal{O}$ (depending on the choice of X) such that $Q(X_0)^TX_0Q(X_0) = \Lambda(X_0)$ and

$$\|q_i(X) - q_i(X_0)\| \leq \mu\|X - X_0\|, \tag{35}$$

where $q_i(X_0) := (Q(X_0))_{\cdot i}$, $i = 1, \ldots, n$. Hence, from (34), (35), and the local

Lipschitz continuity of $\Lambda(\cdot)$, for $i = 1, \ldots, n$, $X \in D_\Lambda$, and $\Delta X \to 0$, we have

$$\begin{aligned}
\lambda_i(X) &= q_i(X)^T \left[X_0 q_i(X_0) + X_0(q_i(X) - q_i(X_0))\right] + q_i(X)^T \Delta X q_i(X) \\
&= \lambda_i(X_0) q_i(X)^T q_i(X_0) + q_i(X)^T X \left[q_i(X) - q_i(X_0)\right] + O(\|\Delta X\|^2) \\
&\quad + q_i(X)^T \Delta X q_i(X) \\
&= \lambda_i(X_0) q_i(X)^T q_i(X_0) + \lambda_i(X) q_i(X)^T \left[q_i(X) - q_i(X_0)\right] \\
&\quad + q_i(X)^T \Delta X q_i(X) + O(\|\Delta X\|^2) \\
&= \lambda_i(X_0) q_i(X)^T q_i(X_0) + \left[\lambda_i(X_0) + O(\|\Delta X\|)\right] q_i(X)^T \left[q_i(X) - q_i(X_0)\right] \\
&\quad + q_i(X)^T \Delta X q_i(X) + O(\|\Delta X\|^2) \\
&= \lambda_i(X_0) q_i(X)^T q_i(X) + q_i(X)^T \Delta X q_i(X) + O(\|\Delta X\|^2) \\
&= \lambda_i(X_0) + q_i(X)^T \Delta X q_i(X) + O(\|\Delta X\|^2), \qquad (36)
\end{aligned}$$

which, according to Lemma 4.6, implies

$$\lambda_i(X) - \lambda_i(X_0) - \lambda_i'(X)(\Delta X) = O(\|\Delta X\|)^2, \quad i = 1, \ldots, n.$$

This, together with Theorem 2.2, implies that for $X \to X_0$ and $V \in \partial\Lambda(X)$,

$$\Lambda(X) - \Lambda(X_0) - V(X - X_0) = O(\|X - X_0\|^2).$$

Hence, (8), and therefore the strong semismoothness of $\Lambda(\cdot)$, is proved. □

5. Newton's method for inverse eigenvalue problems. In this section, we shall show how the strong semismoothness of eigenvalues of symmetric matrices can be used to analyze the quadratic convergence of Newton's method for solving IEPs. Unless stated otherwise, $A : \Re^n \to \mathcal{S}$ is assumed to be continuously differentiable everywhere and $F : \Re^n \to \Re^n$ is defined by (2), i.e.,

$$F(c) = \begin{bmatrix} \lambda_1(A(c)) - \lambda_1^* \\ \vdots \\ \lambda_n(A(c)) - \lambda_n^* \end{bmatrix},$$

where $\{\lambda^*\}_{i=1}^n$ are given n numbers and arranged in the decreasing order. Then the IEP is equivalent to finding $c^* \in \Re^n$ such that $F(c^*) = 0$.

For any $c \in \Re^n$, let $\mathcal{Q}(c) \subseteq \mathcal{O}$ be a subset of $\Re^{n\times n}$ such that for any $Q(c) \in \mathcal{Q}(c)$ we have

$$Q(c)^T A(c) Q(c) = \Lambda(A(c))$$

with $\lambda_1(A(c)) \geq \cdots \geq \lambda_n(A(c))$. For any $Q(c) \in \mathcal{Q}(c)$, define

$$q_i(c) = (Q(c))_{\cdot i}, \qquad i = 1, \ldots, n.$$

Let $\partial A(c)/\partial c_j$ be the partial derivative of $A(c)$ with respect to c_j, $j = 1, \ldots, n$. Then for any $c \in \Re^n$

$$\partial_Q F(c) = \partial_B \Lambda(A(c))(A'(c))$$

is well defined. By using Lemmas 4.5 and 4.6 and [24, Theorem 2.5], we have the following result.

PROPOSITION 5.1.

(a) *For any* $c \in \Re^n$, $V \in \partial_Q F(c)$ *if and only if there exists a* $Q(c) \in \mathcal{Q}(c)$ *such that*

$$V_{i\cdot} = \left[q_i(c)^T(\partial A(c)/\partial c_1)q_i(c), \ldots, q_i(c)^T(\partial A(c)/\partial c_n)q_i(c)\right]. \tag{37}$$

(b) *If* $c \in \Re^n$ *is such that* $A(c)$ *has distinct eigenvalues, then* F *is continuously differentiable at* c *and for any* $Q(c) \in \mathcal{Q}(c)$

$$F_i'(c) = \left[q_i(c)^T(\partial A(c)/\partial c_1)q_i(c), \ldots, q_i(c)^T(\partial A(c)/\partial c_n)q_i(c)\right]. \tag{38}$$

Hence, according to Proposition 5.1, a generalized Newton method for solving the IEP can be described as follows.

ALGORITHM 5.1 (a generalized Newton method).

Step 0. *Choose a starting point value* c^0. $k := 0$.

Step 1. *Compute a* $Q(c^k) \in \mathcal{Q}(c^k)$ *and form* $V_k \in \partial_Q F(c^k)$ *according to Proposition* 5.1.

Step 2. *Set* $c^{k+1} := c^k + \Delta c^k$, *where* Δc^k *is computed by* $F(c^k) + V_k \Delta c^k = 0$.

Step 3. *Replace* k *by* $k+1$ *and go to Step* 1.

In the above generalized Newton method, at the kth step one needs to compute eigenvectors $Q(c^k)$ and eigenvalues $\Lambda(A(c^k))$. Once they are computed, $F(c^k)$ and $V^k \in \partial_Q F(c^k)$ can be formulated easily. If $A(c)$ takes form (1) and at each step $A(c^k)$ has distinct eigenvalues, Algorithm 5.1 reduces to the Newton method considered by many authors, e.g., see [15, 10] and references therein.

THEOREM 5.2. *Suppose that* F *is defined by* (2) *and* $F(c^*) = 0$. *If all* $V \in \partial_Q F(c^*)$ *are nonsingular and* A' *is Lipschitz continuous around* c^*, *then there exists a neighborhood* $N(c^*)$ *of* c^* *such that for any* $c^0 \in N(c^*)$ *Algorithm* 5.1 *is well defined and the iterates* $\{c^k\}$ *converge to* c^* *Q-quadratically.*

Proof. From Theorem 4.7, we know that $\Lambda(\cdot)$ is strongly semismooth everywhere. Hence, by Theorem 2.4 we obtain the conclusion of this theorem. □

Theorem 5.2 contains a very general convergence result for the quadratic convergence of Newton's method for solving IEPs. However, the nonsingularity assumption on $\partial_Q F(c^*)$ is too strong for IEPs when $A(c^*)$ has multiple eigenvalues. To relax this condition, let $S \subseteq \Re^n$ be defined by

$$S = \{c \in \Re^n |\ A(c) \text{ has distinct eigenvalues}\}. \tag{39}$$

Then, by Lemma 4.5 and Proposition 5.1 for any $c \in S$, $F(\cdot)$ is continuously differentiable at c and

$$\partial_B F(c) = \partial_Q F(c) = \partial F(c) = \{F'(c)\}.$$

THEOREM 5.3. *Suppose that* F *is defined by* (2), $F(c^*) = 0$, *and* S *is defined by* (39). *If* (i) *for each* k, $c^k \in S$ *and* $c^* \in \mathrm{cl}S$; (ii) *all* $V \in \partial_B|_S F(c^*)$ *are nonsingular; and* (iii) A' *is Lipschitz continuous around* c^*, *then there exists a neighborhood* $N(c^*)$ *of* c^* *such that, for any* $c^0 \in N(c^*)$, *Algorithm* 5.1 *is well defined and the iterates* $\{c^k\}$ *converge to* c^* *Q-quadratically.*

Proof. By using Theorems 3.4 and 4.7, we obtain this theorem. □

In Theorem 5.3, we need only the nonsingularity of $\partial_B|_S F(c^*)$ rather than $\partial_Q F(c^*)$. The price to pay is that all the iterates must stay in S, where S is defined by (39).

Since $\Re^n \backslash S$ is usually a null set, this condition is reasonable for IEPs [10, pp. 647–648]. For illustration, let us consider the following IEP with

$$F(c) = \begin{bmatrix} \lambda_1(A(c)) - \lambda_1^* \\ \lambda_2(A(c)) - \lambda_2^* \end{bmatrix},$$

$$A(c) = c_1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_2 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

and $\lambda_1^* = \lambda_2^* = 1$. Then, $\lambda_1(A(c)) = c_1 + |c_2|$, $\lambda_2(A(c)) = c_1 - |c_2|$, and $S = \{c \in \Re^2 |\ c_2 \neq 0\}$. The function F has a unique solution at $c^* = (1,0)$. Note that $A(c^*)$ has a multiple eigenvalues at c^* and

$$\partial_B|_S F(c^*) = \left\{ \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right\}.$$

Therefore, all $V \in \partial_B|_S F(c^*)$ are nonsingular.

It was probably Nocedal and Overton [15] who first discussed the quadratic convergence of Newton's method for solving IEPs with multiple eigenvalues. In their proof, a theorem of Rellich [19] on analytic matrix functions was invoked. In [10], by using the eigenprojector, Friedland, Nocedal, and Overton presented a different elegant proof on the quadratic convergence of Newton's method for solving IEPs with multiple eigenvalues. The latter did not use Rellich's theorem. Our results in this paper could be thought of as a generalization of their method I by explicitly exploring the strong semismoothness of the eigenvalue functions.

Before we finish this section, let us consider the generalized inverse eigenvalue problem (GIEP). Let $C : \Re^n \to \mathcal{S}$ and $D : \Re^n \to \mathcal{S}$ be continuously differentiable and $D(c)$ be positive definite whenever $c \in \Omega$, an open subset of $\Re^n$. Given n real numbers $\{\lambda_i^*\}_{i=1}^n$, which are arranged in the decreasing order $\lambda_1^* \geq \cdots \geq \lambda_n^*$, the GIEP is to find a vector $c^* \in \Omega$ such that the symmetric generalized eigenvalue problem $C(c^*)x = \lambda D(c^*)x$ has the prescribed eigenvalues $\lambda_1^*, \ldots, \lambda_n^*$. If $D(c) \equiv I$, then the GIEP is the IEP considered above. It is readily seen that the GIEP can be converted into the form of solving $F(c) = 0$ with

$$F(c) = \begin{bmatrix} \lambda_1(A(c)) - \lambda_1^* \\ \vdots \\ \lambda_n(A(c)) - \lambda_n^* \end{bmatrix}, \quad c \in \Omega, \tag{40}$$

where $A(c) = D(c)^{-\frac{1}{2}} C(c) D(c)^{-\frac{1}{2}}$.

Dai and Lancaster [7] and Dai [6] considered a special case of the GIEP, i.e., $C(c)$ and $D(c)$ are defined by

$$C(c) = C_0 + \sum_{i=1}^n c_i C_i, \qquad D(c) = D_0 + \sum_{i=1}^n c_i D_i, \tag{41}$$

where $C_0, C_1, \ldots, C_n, D_0, D_1, \ldots D_n \in \mathcal{S}$ and $D(c)$ is positive definite whenever $c \in \Omega$.

When $C(c)$ and $D(c)$ take the form (41), Dai and Lancaster [7] proposed the following Newton method for solving the GIEP.

ALGORITHM 5.2 (a Newton method of Dai and Lancaster [7]).

Step 0. *Choose a starting point value* c^0. $k := 0$.

Step 1. *Compute* $C(c^k) = C_0 + \sum_{j=1}^n c_j^k C_j,\ D(c^k) = D_0 + \sum_{j=1}^n c_j^k D_j$.
Step 2. *Set* $c^{k+1} := c^k + \Delta c^k$, *where* Δc^k *is computed by* $F(c^k) + F'(c^k)\Delta c^k = 0$.
Step 3. *Replace* k *by* $k+1$ *and go to Step* 1.

The following theorem gives an affirmative answer to a conjecture made in [7, p. 11] on the quadratic convergence of Algorithm 5.2, which was supported by numerical experiments.

THEOREM 5.4. *Suppose that* $c^* \in \Omega$ *such that* $F(c^*) = 0$. *If* (i) *for each* k, $A(c^k)$ *has distinct eigenvalues and* $F'(c^k)$ *is invertible; and* (ii) $\limsup_{k\to\infty} \|F'(c^k)^{-1}\| < \infty$, *then there exists a neighborhood* $N(c^*)$ *of* c^* *such that for any* $c^0 \in N(c^*)$ *the iterates* $\{c^k\}$ *generated by Algorithm* 5.2 *converge to* c^* *Q-quadratically.*

Proof. Since $A(c^k)$ has distinct eigenvalues, F is continuously differentiable at c^k. Note that Algorithm 5.2 is a special case of Algorithm 5.1. By using Theorems 4.7 and 3.4 with $S = \{c^0, c^1, \ldots\}$, we get the conclusion of the theorem. □

6. Summary and possible future research topics. In this paper we review basic concepts of semismoothness and Newton's method for semismooth equations. We show the strong semismoothness of eigenvalues of symmetric matrices and demonstrate how this result can be used to provide a unified analysis for the quadratic convergence of the Newton-type methods for IEPs and GIEPs.

We feel that several topics could be further investigated. First, it would be interesting to look at the strong semismoothness of the functions arising from other IEPs, e.g., the least square IEPs [10]. Second, we could develop nonsmooth quasi-Newton [22] methods, rather than Newton's method, for IEPs and GIEPs. Chan and Tseng [3] provided such an approach for IEPs with distinct eigenvalues. The problem is still unsolved in the case of multiple eigenvalues. Third, it is desirable to have a "smoothing" version of the Newton method discussed in this paper; namely, we find a parameterized function $H(\varepsilon, x)$ for a strongly semismooth function $F(x)$ such that $H(\varepsilon, y) \to F(x)$ as $(\varepsilon, y) \to (0^+, x)$ and that $H(\varepsilon, x)$ is differentiable for $\varepsilon \neq 0$. It is proved in [23] that any nonsmooth function has approximate smoothing functions, but the proof does not give any concrete smoothing functions for IEPs. It is then interesting to ask what smoothing function could be used for IEPs.

Acknowledgment. The authors are grateful to the referees for their very constructive comments.

REFERENCES

[1] R. BHATIA, *Matrix Analysis*, Springer–Verlag, New York, 1997.
[2] F. W. BIEGLER-KÖNIG, *A Newton iteration process for inverse eigenvalue problems*, Numer. Math., 37 (1981), pp. 349–354.
[3] R. H. CHAN, S.-F. XU, AND H.-M. ZHOU, *On the convergence rate of a quasi-Newton method for inverse eigenvalue problems*, SIAM J. Numer. Anal., 36 (1999), pp. 436–441.
[4] X. CHEN AND P. TSENG, *Non-interior continuation methods for solving semidefinite complementarity problems*, Math. Program., to appear.
[5] F. CLARKE, *Optimization and Nonsmooth Analysis*, John Wiley, New York, 1983.
[6] H. DAI, *An algorithm for symmetric generalized inverse eigenvalue problems*, Linear Algebra Appl., 296 (1999), pp. 79–98.
[7] H. DAI AND P. LANCASTER, *Newton's method for a generalized inverse eigenvalue problem*, Numer. Linear Algebra Appl., 4 (1997), pp. 1–21.
[8] K. FAN, *On a theorem of Weyl concerning eigenvalues of linear transformations*. I., Proc. Nat. Acad. Sci. U.S.A., 35 (1949), pp. 652–655.
[9] A. FISCHER, *Solution of monotone complementarity problems with locally Lipschitzian functions*, Math. Program., 76 (1997), pp. 513–532.

[10] S. FRIEDLAND, J. NOCEDAL, AND M. L. OVERTON, *The formulation and analysis of numerical methods for inverse eigenvalue problems*, SIAM J. Numer. Anal., 24 (1987), pp. 634–667.
[11] P. LANCASTER, *On eigenvalues of matrices dependent on a parameter*, Numer. Math., 6 (1964), pp. 377–387.
[12] A. S. LEWIS, *Nonsmooth analysis of eigenvalues*, Math. Program., 84 (1999), pp. 1–24.
[13] A. S. LEWIS AND M. L. OVERTON, *Eigenvalue optimization*, Acta Numer., 5 (1996), pp. 149–190.
[14] R. MIFFLIN, *Semismooth and semiconvex functions in constrained optimization*, SIAM J. Control Optim., 15 (1977), pp. 959–972.
[15] J. NOCEDAL AND M. L. OVERTON, *Numerical methods for solving inverse eigenvalue problems*, in Lecture Notes in Math. 1005, V. Pereya and A. Reinoza, eds., Springer–Verlag, New York, Berlin, 1983, pp. 212–226.
[16] F. POTRA, L. QI, AND D. SUN, *Secant methods for semismooth equations*, Numer. Math., 80 (1998), pp. 305–304.
[17] L. QI, *Convergence analysis of some algorithms for solving nonsmooth equations*, Math. Oper. Res., 18 (1993), pp. 227–244.
[18] L. QI AND J. SUN, *A nonsmooth version of Newton's method*, Math. Program., 58 (1993), pp. 353–367.
[19] F. RELLICH, *Perturbation Theory of Eigenvalue Problems*, Gordon and Breach, New York, London, Paris, 1969.
[20] R. T. ROCKAFELLAR, *Convex Analysis*. Princeton University Press, Princeton, NJ, 1970.
[21] G. W. STEWART AND J.-G. SUN, *Matrix Perturbation Theory*, Academic Press, New York, 1990.
[22] D. SUN AND J. HAN, *Newton and quasi-Newton methods for a class of nonsmooth equations and related problems*, SIAM J. Optim., 7 (1997), pp. 463–480.
[23] D. SUN AND L. QI, *Solving variational inequality problems via smoothing-nonsmooth reformulations*, J. Comput. Appl. Math., 129 (2001), pp. 37–62.
[24] D. SUN AND J. SUN, *Semismooth matrix valued functions*, Math. Oper. Res., 27 (2002), pp. 150–169.
[25] J.-G. SUN, *Eigenvalues and eigenvectors of a matrix dependent on several parameters*, J. Comput. Math., 3 (1985), pp. 351–364.
[26] J. H. WILKINSON, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.
[27] S.-F. XU, *An Introduction to Inverse Algebraic Eigenvalue Problems*, Beijing University Press, Beijing, 1998.