Semismoothness of the Maximum Eigenvalue Function of a Symmetric Tensor and its Application *

Guoyin Li † Liqun Qi‡ Gaohang Yu §

Revised Version: September 19, 2011

Abstract

In this paper, we examine the maximum eigenvalue function of an even order real symmetric tensor. By using the variational analysis techniques, we first show that the maximum eigenvalue function is a continuous and convex function on the symmetric tensor space. In particular, we obtain the convex subdifferential formula for the maximum eigenvalue function. Next, for an mth-order n-dimensional symmetric tensor $\mathcal{A}$, we show that the maximum eigenvalue function is always ρth-order semismooth at $\mathcal{A}$ for some rational number $\rho > 0$. In the special case when the geometric multiplicity is one, we show that ρ can be set as $\frac{1}{(2m-1)^n}$. Sufficient condition ensuring the strong semismoothness of the maximum eigenvalue function is also provided. As an application, we propose a generalized Newton method to solve the space tensor conic linear programming problem which arises in medical imaging area. Local convergence rate of this method is established by using the semismooth property of the maximum eigenvalue function.

Keywords: Symmetric Tensor, Maximum eigenvalue function, Real Polynomial, Semismooth, Generalized Newton method.

1 Introduction

An mth-order n-dimensional tensor $\mathcal{A}$ consists of n^m entries in real number:

$$\mathcal{A} = (\mathcal{A}_{i_1 i_2 \cdots i_m}), \quad \mathcal{A}_{i_1 i_2 \cdots i_m} \in \mathbb{R}, \quad 1 \le i_1, i_2, \cdots, i_m \le n. \tag{1.1}$$

We say a tensor $\mathcal{A}$ is symmetric if the value of $\mathcal{A}_{i_1 i_2 \cdots i_m}$ is invariant under any permutation of its index $\{i_1, i_2, \cdots, i_m\}$. Clearly, when $m = 2$, a symmetric tensor is nothing but a symmetric matrix. A symmetric tensor uniquely defines an mth degree homogeneous polynomial function f with real coefficients: for all $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$,

$$f(x) = \mathcal{A}x^m := \sum_{i_1, \dots, i_m = 1}^{n} \mathcal{A}_{i_1 i_2 \cdots i_m} x_{i_1} \dots x_{i_m}.$$

*The authors are grateful to the referees and the editor for their valuable comments and constructive suggestions which have contributed to the final preparation of the paper. Research was partially supported by a research grant form the Council of Hong Kong, National Natural Science Foundation of China (No. 11001060) and a research grant from Australian Research Council.

†Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia. E-mail: g.li@unsw.edu.au (G. Li)

‡Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong. E-mail: maqilq@polyu.edu.hk (L. Qi)

§Jiangxi Key laboratory of numerical simulation technology, School of Mathematics and Computer Sciences, Gannan Normal University, Ganzhou, 341000, China. E-mail: maghyu@163.com (G. Yu)

Recently, to study the stability of a homogeneous polynomial dynamical system, Qi [32, 30] introduced the definition of eigenvalues of a symmetric tensor and showed that the stability of this system is tied up with the negativity of the maximum eigenvalue of the corresponding symmetric tensor. Independently, Lim [23] also gave such a definition via a variational approach and established an interesting Perron-Frobenius type theorem. Recently, numerical study on tensors has attracted a lot of researchers. In particular, various efficiently numerical schemes have been proposed to find the low rank approximations of a tensor and the eigenvalues/eigenvectors of a tensor (cf. [14, 15, 16, 17, 30, 36, 29]).

On the other hand, as a fundamental concept in matrix analysis, eigenvalues of a symmetric matrix have been well-studied by a lot of researchers via different approaches. For example, it is well-known that the maximum eigenvalue function is a continuous and convex function on the symmetric matrix space, and is differentiable almost everywhere. Moreover, recently, Sun and Sun [38] further showed that the eigenvalue function of a matrix is indeed strongly semismooth (which is a stronger differentiability result) and explained how this desired property can be used to study the local convergence rate of the generalized Newton method when it applies to solve the inverse eigenvalue problem in numerical analysis.

In this paper, we attempt to examine these desired mathematical properties for the maximum eigenvalue function of a symmetric tensor with even order. The organization of this paper is as follows. We first fix the notations and collect some basic definitions in Section 2. In Section 3, by using the variational analysis techniques, we show that the maximum eigenvalue function is continuous and convex, and hence differentiable almost everywhere. In particular, we obtain the convex subdifferential formula for the maximum eigenvalue function. In Section 4, for an mth order n-dimensional tensor $\mathcal{A}$, we show that the maximum eigenvalue function is always ρth-order semismooth at $\mathcal{A}$ for some rational number $\rho > 0$. In the special case when the geometric multiplicity is one, we show that ρ can be set as $\frac{1}{(2m-1)^n}$. Sufficient condition ensuring the strong semismoothness of the maximum eigenvalue function is also provided. In Section 5, as an application, we propose a generalized Newton method to solve a space tensor conic linear programming which arises in medical imaging. We also establish the local convergence rate of this method by using the semismooth property of the maximum eigenvalue function. Finally, we conclude our paper and present some future research topics in Section 6.

2 Preliminaries

In this section, we fix the notations and collect some basic definitions and facts which we will use later on. Let X, Y be finite dimensional inner product spaces. We use $\mathbb{B}_X$ (resp. $\mathbb{B}_Y$) to denote the unit open ball in X (resp. Y). Denote the space of all linear map from X to Y by $L(X, Y)$. Consider a locally Lipschitz function $G : X \to Y$. By the Rademacher's Theorem, G is differentiable almost everywhere on X. Let D_G be the set consisting of all the points where G is differentiable. Then, for any $x \in D_G$, the derivative of G, $\nabla G(x)$ exists. Denote

$$J_B G(x) = \{V \in L(X, Y) : V = \lim_{x_k \to x} \nabla G(x_k), x_k \in D_G\}.$$

Then, its Clarke's generalized Jacobian is defined by $J_C G(x) = \mathrm{conv} J_B G(x)$. In particular, if $Y = \mathbb{R}$ and $G = g$ where $g : X \to \mathbb{R}$ is locally Lipschitz, by identifying X^* as X, then the Clarke's generalized Jacobian reduces to the Clarke's subdifferential defined by

$$\partial_C g(x) = \{\xi \in X : \langle \xi, v \rangle_X \leq g^\circ(x; v) \text{ for all } v \in X\}.$$

where $\langle \cdot, \cdot \rangle_X$ is the inner product in X and $g^\circ(x; v)$ is the Clarke directional derivative of g at the point x in the direction v defined by

$$g^\circ(x; v) = \limsup_{y \to x,\ t \downarrow 0} \frac{g(y + tv) - g(y)}{t}.$$

The Clarke subdifferential $\partial_C g(x)$ is a nonempty, convex and compact subset of X for each $x \in X$. Moreover, if g is convex on X in the sense that

$$g(\lambda x_1 + (1-\lambda)x_2) \leq \lambda g(x_1) + (1-\lambda)g(x_2), \ \forall\, \lambda \in [0,1] \text{ and } x_1, x_2 \in X,$$

then $\partial_C g(x) = \partial g(x)$ for all $x \in X$ where $\partial g(x)$ is the usual convex subdifferential of g at x defined by $\partial g(x) := \{\xi \in X : \langle \xi, z - x\rangle_X \leq g(z) - g(x) \text{ for all } z \in X\}$. The important mean value theorem for a locally Lipschitz function in terms of Clarke subdifferential (cf. [8]) is listed as below:

Theorem 2.1. (Mean Value Theorem) *Let X be a finite dimensional inner product space and let $g : X \to \mathbb{R}$ be a Lipschitz function. Let $x_1, x_2 \in X$. Then there exist $u \in [x_1, x_2]$ and $\xi \in \partial_C g(u)$ such that*

$$g(x_2) - g(x_1) = \langle \xi, x_2 - x_1\rangle_X.$$

We are now ready to state the definitions of semismooth functions and ρth-order semismooth functions.

Definition 2.1. *Let $G : X \to Y$ be a locally Lipschitz and directionally differentiable function. Then, the function G is said to be semismooth at x if*

$$G(x + \Delta x) - G(x) - V\Delta x = o(\|\Delta x\|), \ \forall\, V \in J_C G(x + \Delta x).$$

Moreover, G is said to be ρth-order semismooth function at x for some $\rho \in (0,1]$ if

$$G(x + \Delta x) - G(x) - V\Delta x = O(\|\Delta x\|^{1+\rho}), \ \forall\, V \in J_C G(x + \Delta x).$$

In particular, if $\rho = 1$, we say G is strongly semismooth at x. We also say $G : X \to Y$ is a semismooth (resp. ρth-order semismooth, strongly semismooth) function if G is semismooth (resp. ρth-order semismooth, strongly semismooth) at x for all $x \in X$.

The concept of a semismooth function was originally given by Mifflin [26] when $Y = \mathbb{R}$. Later on, Qi and Sun [33] extended the definition to vector value functions and showed that semismooth functions play an important role in establishing the local convergence rate of the generalized Newton method for solving nonsmooth equations. From the definitions of the semismooth functions, it is clear that scalar multiplication and sums of semismooth (resp. ρth-order semismooth) functions are still semismooth (resp. ρth-order semismooth) functions. Moreover, for a locally Lipschitz function $G : X \to Y$, G is semismooth at $\overline{x}$ if G is continuously differentiable at $\overline{x}$ or G is convex in a neighbourhood of $\overline{x}$.

Next, we recall some basic definitions and facts of a tensor and its eigenvalues. Let $n \in \mathbb{N}$ and let m be an even number. Consider

$$S = \{\mathcal{A} : \mathcal{A} \text{ is an } m\text{th-order } n\text{-dimensional symmetric tensor}\}$$

Clearly, S is a vector space under the addition and multiplication defined as below: for any $t \in \mathbb{R}$, $\mathcal{A} = (\mathcal{A}_{i_1,\ldots,i_m})_{1\leq i_1,\ldots,i_m \leq n}$ and $\mathcal{B} = (\mathcal{B}_{i_1,\ldots,i_m})_{1\leq i_1,\ldots,i_m \leq n}$

$$\mathcal{A} + \mathcal{B} = (\mathcal{A}_{i_1,\ldots,i_m} + \mathcal{B}_{i_1,\ldots,i_m})_{1\leq i_1,\ldots,i_m \leq n} \text{ and } t\mathcal{A} = (t\mathcal{A}_{i_1,\ldots,i_m})_{1\leq i_1,\ldots,i_m \leq n}.$$

For each $\mathcal{A}, \mathcal{B} \in S$, we define the inner product by

$$\langle \mathcal{A}, \mathcal{B}\rangle_S = \sum_{i_1,\ldots,i_m=1}^{n} \mathcal{A}_{i_1,\ldots,i_m}\mathcal{B}_{i_1,\ldots,i_m}.$$

The corresponding norm is defined by $\|\mathcal{A}\|_S = (\langle \mathcal{A}, \mathcal{A}\rangle_S)^{1/2} = \big(\sum_{i_1,\ldots,i_m=1}^{n} \mathcal{A}_{i_1,\ldots,i_m}^2\big)^{1/2}$. The unit ball in S is denoted by $\mathbb{B}_S$. For a vector $x \in \mathbb{R}^n$, we use x_i to denotes its ith component. We use $x^{[m-1]}$ to denote a vector in $\mathbb{R}^n$ such that

$$x_i^{[m-1]} = (x_i)^{m-1}.$$

Moreover, for a vector $x \in \mathbb{R}^n$, we use x^m to denote the mth-order n-dimensional symmetric rank one tensor induced by x, i.e.,

$$(x^m)_{i_1 \dots i_m} = x_{i_1} \dots x_{i_m}, \ \forall\, i_1, \dots, i_m \in \{1, \dots, n\}.$$

Let $\mathcal{A} \in S$. By the tensor product (cf [34]), $\mathcal{A}x^m$ is a real number defined as

$$\mathcal{A}x^m = \sum_{i_1,\dots,i_m=1}^{n} \mathcal{A}_{i_1,\dots,i_m} x_{i_1} \dots x_{i_m} = \langle \mathcal{A}, x^m \rangle_S$$

and $\mathcal{A}x^{m-1}$ is a vector in $\mathbb{R}^n$ whose ith component is

$$\sum_{i_2 \cdots i_m=1}^{n} A_{i\, i_2 \cdots i_m} x_{i_2} \cdots x_{i_m}. \tag{2.2}$$

Definition 2.2. *Let $\mathcal{A}$ be an mth-order n-dimensional symmetric tensor. We say $\lambda \in \mathbb{C}$ is an eigenvalue of $\mathcal{A}$ and $x \in \mathbb{C}^n \backslash \{0\}$ is an eigenvector corresponding to λ if (x, λ) is a solution of the following homogeneous polynomial equation system:*

$$\mathcal{A}x^{m-1} = \lambda x^{[m-1]}.$$

Moreover, if λ and x are all real (i.e., $(\lambda, x) \in \mathbb{R} \times (\mathbb{R}^n \backslash \{0\})$), we say λ is an H-eigenvalue of $\mathcal{A}$.

This definition was introduced by Qi in [32] motivated from the study of the stability of polynomial dynamic systems. Independently, Lim [23] also gave such a definition via a variational approach. From [32] and [6], an H-eigenvalues for an even order symmetric tensor always exists. Recently, a lot of researchers have devoted themselves to the study of eigenvalue problems of symmetric tensors and have obtained many interesting results (for example see [32, 23, 5, 29]).

3 Maximum Eigenvalue Function

In this section, we examine the continuity and differentiability of the maximum eigenvalue function. To do this, we first formally define the maximum eigenvalue function. Since any symmetric tensor with even order always has an H-eigenvalue (cf [6]), it then makes sense to define the maximum eigenvalue function $\lambda_1 : S \to \mathbb{R}$ as follows:

$$\lambda_1(\mathcal{A}) = \{\lambda \in \mathbb{R} : \lambda \text{ is the largest H-eigenvalue of } \mathcal{A}\}.$$

We first establish the following simple lemma which will be useful for our later analysis.

Lemma 3.1. *Let $\mathcal{A}$ be an mth-order n-dimensional symmetric tensor where m is even. Then, we have*

$$\lambda_1(\mathcal{A}) = \max_{x \neq 0} \frac{\mathcal{A}x^m}{\|x\|_m^m} = \max_{\|x\|_m=1} \mathcal{A}x^m\,,$$

where $\|x\|_m = (\sum_{i=1}^n |x_i|^m)^{1/m}$.

Proof. Consider the following optimization problem (P)

$$\begin{array}{lll} (P) & \max_{x \in \mathbb{R}^n} & \mathcal{A}x^m \\ & \text{s.t.} & \|x\|_m^m = 1. \end{array}$$

Let $f(x) := \mathcal{A}x^m$ and $g(x) := \|x\|_m^m$. Since f is continuous and the feasible set $\{x : g(x) = 1\}$ is compact, a global maximizer of (P) exists. Denote a maximizer of (P) by x_0. Clearly, $x_0 \neq 0$. Note that g is a homogeneous polynomial with degree m, and so the Euler Identity implies that

$\nabla g(x)^T x = mg(x)$. Thus, for any x with $g(x) = 1$, $\nabla g(x) \neq 0$. So, the standard KKT theory implies that there exists $\lambda_0 \in \mathbb{R}$ such that

$$m\mathcal{A}x_0^{m-1} - m\lambda_0 x_0^{[m-1]} = \nabla f(x_0) - \lambda_0 \nabla g(x_0) = 0$$

This implies that λ_0 is a real eigenvalue of $\mathcal{A}$ with an real eigenvector x_0, and so, $\lambda_0 \leq \lambda_1(\mathcal{A})$. Note that $\lambda_0 = \mathcal{A}x_0^m = v(P)$ where $v(P)$ is the optimal value of (P). It follows that $v(P) \leq \lambda_1(\mathcal{A})$, that is,

$$\max_{\|x\|_m=1} \mathcal{A}x^m \leq \lambda_1(\mathcal{A}).$$

Finally, noting that, for any real eigenvector u corresponds to $\lambda_1(\mathcal{A})$ with $\|u\|_m = 1$, we have

$$\mathcal{A}u^m = u^T(\mathcal{A}u^{m-1}) = \lambda_1(\mathcal{A})u^T u^{[m-1]} = \lambda_1(\mathcal{A})\|u\|_m^m = \lambda_1(\mathcal{A}).$$

Thus, $\lambda_1(\mathcal{A}) = \max_{\|x\|_m=1} \mathcal{A}x^m$, and so, the conclusion follows as $\max_{\|x\|_m=1} \mathcal{A}x^m = \max_{x\neq 0} \frac{\mathcal{A}x^m}{\|x\|_m^m}$. □

Remark 3.1. *From the proof of the Lemma 3.1, we see that*

$$\{u : \mathcal{A}u^m = \lambda_1(\mathcal{A}),\ \|u\|_m = 1\} = \{u : (\lambda_1(\mathcal{A}), u) \textit{ is a real eigenpair of } \mathcal{A},\ \|u\|_m = 1\}.$$

Next, we show that the maximum eigenvalue function is continuous and convex.

Theorem 3.1. *The function λ_1 is a continuous, homogeneous and convex function on S.*

Proof. Let $\{(\lambda_1(\mathcal{A}), u\}$ be an eigenpair of $\mathcal{A}$, i.e.,

$$\sum_{i_2,\ldots,i_m=1} a_{i,i_2,\ldots,i_m} u_{i_2} \ldots u_{i_m} = (\mathcal{A}u^{m-1})_i = \lambda_1(\mathcal{A})u_i^{m-1}.$$

Let $\alpha > 0$. Then,

$$\sum_{i_2,\ldots,i_m=1} \alpha a_{i,i_2,\ldots,i_m} u_{i_2} \ldots u_{i_m} = ((\alpha\mathcal{A})u^{m-1})_i = \lambda_1(\alpha\mathcal{A})u_i^{m-1}.$$

It follows that for any $\alpha > 0$, $\{(\alpha\lambda_1(\mathcal{A}), u\}$ is an eigenpair of $\alpha\mathcal{A}$, and so, $\lambda_1(\alpha\mathcal{A}) = \alpha\lambda_1(\mathcal{A})$. Thus, λ_1 is a homogeneous function. Let $T := \{u^m : \|u\|_m = 1\} \subseteq S$. Since $\lambda_1(\mathcal{A}) = \max_{\|x\|_m=1} \mathcal{A}x^m$, we have $\lambda_1(\mathcal{A}) = \max_{\mathcal{B}\in T}\langle \mathcal{B}, \mathcal{A}\rangle_S$. Note that $\mathcal{B} \mapsto \langle \mathcal{B}, \mathcal{A}\rangle_S$ is affine and the supremum of a series of affine functions is convex. It follows that λ_1 is a finite-valued convex function on S and so, is continuous and convex. □

As λ_1 is continuous and convex, its convex subdifferential always exists where the convex subdifferential $\partial\lambda_1$ is defined by

$$\partial\lambda_1(\mathcal{A}) = \{\mathcal{B} \in S : \langle \mathcal{B}, \mathcal{A}' - A\rangle_S \leq \lambda_1(\mathcal{A}') - \lambda_1(\mathcal{A}) \text{ for all } \mathcal{A}' \in S\}.$$

We are now ready to state the subdifferential formula for the maximum eigenvalue function.

Theorem 3.2. *Let $\mathcal{A}$ be an mth-order n-dimensional symmetric tensor where m is even. Then, we have*

$$\partial\lambda_1(\mathcal{A}) = \mathrm{conv}\{u^m : (\lambda_1(\mathcal{A}), u) \textit{ is a real eigenpair of } \mathcal{A} \textit{ and } \|u\|_m = 1\}.$$

where conv(C) denotes the convex hull of the set C.

Proof. First of all, we recall that λ_1 is a convex function on S. To see the subdifferential formula, let $T := \{u^m : \|u\|_m = 1\} \subseteq S$. Since $\lambda_1(\mathcal{A}) = \max_{\|x\|_m=1} \mathcal{A}x^m$, we obtain that $\lambda_1(\mathcal{A}) = \max_{\mathcal{B}\in T}\langle \mathcal{B}, \mathcal{A}\rangle_S$. Note that $\mathcal{B} \mapsto \langle \mathcal{B}, \mathcal{A}\rangle_S$ is affine and T is a compact metric space. Thus, Danskin's theorem (cf. [4, Theorem 4.1.3]) gives us that

$$\partial\lambda_1(\mathcal{A}) = \mathrm{conv} \bigcup_{\mathcal{B}\in T(\mathcal{A})} \partial(\langle \mathcal{B}, \cdot\rangle_S)(\mathcal{A}) = \mathrm{conv} \bigcup_{\mathcal{B}\in T(\mathcal{A})} \mathcal{B} = \mathrm{conv}T(\mathcal{A}),$$

where $T(\mathcal{A})$ is defined by

$$T(\mathcal{A}) := \{\mathcal{B} \in T : \langle \mathcal{B}, \mathcal{A}\rangle_S = \lambda_1(\mathcal{A})\} = \{u^m : \mathcal{A}u^m = \lambda_1(\mathcal{A}),\ \|u\|_m = 1\}.$$

Thus, the conclusion follows. □

Remark 3.2. *If $m = 2$, our subdifferential formula for λ_1 reduces to*

$$\partial\lambda_1(\mathcal{A}) = \mathrm{conv}\{uu^T : (\lambda_1(\mathcal{A}), u) \text{ is a real eigenpair of } \mathcal{A} \text{ and } \|u\|_2 = 1\},$$

which is the classical subdifferential formula of the maximum eigenvalue function in the symmetric matrix case (cf. [19]).

Definition 3.1. *Let $\mathcal{A}$ be an mth-order n-dimensional symmetric tensor. The geometric multiplicity of $\lambda_1(\mathcal{A})$ is defined as the dimension of the eigenspace E associated with $\lambda_1(\mathcal{A})$ where $E := \{v \in \mathbb{R}^n : \mathcal{A}v^{m-1} = \lambda_1(\mathcal{A})v^{[m-1]}\}$.*

Theorem 3.3. *Let $\mathcal{A}$ be an mth-order n-dimensional symmetric tensor where m is even. Then, the maximum eigenvalue function λ_1 is locally Lipschitz, and is (Fréchet) differentiable almost everywhere. Moreover, λ_1 is differentiable at $\mathcal{A} \in S$ if and only if the geometric multiplicity of $\lambda_1(\mathcal{A})$ is one.*

Proof. From the preceding theorem, the maximum eigenvalue function λ_1 is continuous and convex. So, λ_1 is locally Lipschitz. Then, the Radamecher theorem implies that it is (Fréchet) differentiable almost everywhere. To see the last assertion, as m is even, we see that $u^m = (-u)^m$. It follows that the geometric multiplicity of $\lambda_1(\mathcal{A})$ is one if and only if the set

$$\partial\lambda_1(\mathcal{A}) = \{u^m : (\lambda_1(\mathcal{A}), u) \text{ is a real eigenpair of } \mathcal{A} \text{ and } \|u\|_m = 1\}$$

is a singleton. Note that a continuous convex function on a finite dimensional space is Fréchet differentiable if and only if its subdifferential is a singleton. Thus, the conclusion follows. □

Let $\mathcal{A}$ be an mth-order n-dimensional symmetric tensor where m is even. We define its symmetric hyperdeterminant (e.g. see [32, 7]), denoted by $\det(\mathcal{A})$, as an irreducible polynomial in $\mathcal{A}_{i_1,\cdots,i_m}$, which vanishes wherever there is an $x \neq 0$ such that $\mathcal{A}x^{m-1} = 0$. As pointed out in [32, 5], $\det(\mathcal{A})$ is nothing but the resultant of the homogeneous polynomials $\{(\mathcal{A}x^{m-1})_1, \ldots, (\mathcal{A}x^{m-1})_n\}$. Define the characteristic polynomial of $\mathcal{A}$ by $\psi(\lambda) = \det(\mathcal{A} - \lambda\mathcal{I})$ where $\mathcal{I}$ is the unit tensor where $\mathcal{I}_{i_1,\ldots,i_m} = 0$ for all $(i_1, \ldots, i_m) \notin \{(i, \ldots, i) : 1 \le i \le n\}$ and $\mathcal{I}_{i,\ldots,i} = 1$, $1 \le i \le n$. Then, we see that: λ is an eigenvalue of $\mathcal{A}$ if and only if λ is a root of the characteristic polynomial ψ. We now define the algebraic multiplicity of $\lambda_1(\mathcal{A})$ as follows.

Definition 3.2. *Let $\mathcal{A}$ be an mth-order n-dimensional symmetric tensor where m is even. Then, the algebraic multiplicity of $\lambda_1(\mathcal{A})$ is defined as the largest integer of d such that $\psi(\lambda) = (\lambda - \lambda_1(\mathcal{A}))^d\phi(\lambda)$ for some polynomial ϕ with $\phi(\lambda_1(\mathcal{A})) \neq 0$.*

From the definition, we can see that the algebraic multiplicity is greater or equal to the geometric multiplicity (see [5]). It is well known that a real symmetric matrix is diagonalizable, and so, its algebraic multiplicity equals the geometric multiplicity. However, for a general symmetric tensor, the algebraic multiplicity can be significantly larger than the geometric multiplicity. For instance, consider a diagonal tensor with distinct diagonal elements, i.e., $\mathcal{A}_{i_1,\ldots,i_m} = 0$ for all $(i_1, \ldots, i_m) \notin \{(i, \ldots, i) : 1 \le i \le n\}$ and $\mathcal{A}_{i,\ldots,i}$, $1 \le i \le n$, are not the same. As the degree of the polynomial $\psi(\lambda)$ is $n(m-1)^{n-1}$, each diagonal element is an eigenvalue with algebraic multiplicity $(m-1)^{n-1}$. On the other hand, the geometric multiplicity of them are at most n which is much smaller than $(m-1)^{n-1}$ when m is large (for more discussions and specific examples see [32, 5]).

To end this section, we now show that λ_1 is analytic at $\mathcal{A}$ on S if the algebraic multiplicity of $\lambda_1(\mathcal{A})$ is one where $S = \{\mathcal{A} : \mathcal{A}$ is an mth-order n-dimensional symmetric tensor$\}$, $m, n \in \mathbb{N}$ and m is even. The proof of it is along a similar line with the matrix case and makes use of the implicit function theorem.

Proposition 3.1. *Let $p \mapsto \mathcal{A}(p)$ be an analytic map from $\mathbb{R}^N$ to S with $\mathcal{A}(p_0) = \mathcal{A}$. If the algebraic multiplicity of $\lambda_1(\mathcal{A})$ is one, then $p \mapsto \lambda_1(\mathcal{A}(p))$ is analytic at p_0.*

Proof. Without loss of generality, we may assume that $p_0 = 0$. For each $p \in \mathbb{N}$, let $\psi(\lambda, p)$ be the characteristic polynomial of $\mathcal{A}(p)$. Clearly, $(\lambda, p) \mapsto \psi(\lambda, p)$ is an analytic function. As the algebraic multiplicity of $\lambda_1(\mathcal{A})$ is one, we have

$$\psi(\lambda, 0) = (\lambda - \lambda_1(\mathcal{A}))\phi(\lambda)$$

for some polynomial ϕ with $\phi(\lambda_1(\mathcal{A})) \neq 0$. Then, $\frac{\partial \psi}{\partial \lambda}(\lambda_1(\mathcal{A}), 0) \neq 0$. By the implicit function theorem, we see that there exist $\delta_0 > 0$ and a unique continuously differentiable function $w : \mathbb{C}^N \to \mathbb{R}$ such that $w(0) = \lambda_1(\mathcal{A})$ and

$$\psi(w(p), p) = 0 \text{ for all } \|p\| \le \delta_0.$$

Moreover, as $(\lambda, p) \mapsto \psi(\lambda, p)$ is analytic, $p \mapsto w(p)$ is also analytic. It is easy to see that $w(p)$ is an eigenvalue of $\mathcal{A}(p)$. Note that $p \mapsto \lambda_1(\mathcal{A}(p))$ is a differentiable convex function (and so, is continuously differentiable) satisfying $\phi(\lambda_1(\mathcal{A}(p)), p) = 0$. By the uniqueness of the function w, we must have $w(p) = \lambda_1(\mathcal{A}(p))$. Thus, $p \mapsto \lambda_1(\mathcal{A}(p))$ is analytic at 0. $\square$

Theorem 3.4. *If the algebraic multiplicity of $\lambda_1(\mathcal{A})$ is one, then λ_1 is analytic at $\mathcal{A}$ on S.*

Proof. Let the dimension of S be N. Then, there exists an invertible linear map L such that $L(\mathbb{R}^N) = S$. Let $L^{-1}(\mathcal{A}) = p_0$. Note that L is analytic and the algebraic multiplicity of $\lambda_1(L(p_0))$ is one. Then, by the preceding proposition, the map $p \mapsto \lambda_1(L(p))$ is analytic at p_0. Therefore, λ_1 is also analytic at $L(p_0) = \mathcal{A}$. $\square$

In the special case when $m = 2$, the preceding theorem reduces to the classical fact that: for a symmetric matrix M, the maximum eigenvalue function λ_1 is analytic at M when the (algebraic) multiplicity of $\lambda_1(M)$ is one *.

4 Semismoothness of the Maximum Eigenvalue Function

In this section, we examine the semismoothness of the maximal eigenvalue function. Now, consider a function $f : S \to \mathbb{R}$ where S is the symmetric tensor space on $\mathbb{R}^n$. Note that the symmetric tensor space S can be identified as a finite dimensional space with an appropriate dimension. The definition of semismoothness of f can be translated as follows:

Definition 4.1. *Let $f : S \to \mathbb{R}$ be a locally Lipschitz and directionally differentiable function. Then, the function $f : S \to \mathbb{R}$ is said to be semismooth at $\mathcal{A} \in S$ if ,*

$$f(\mathcal{A} + \Delta\mathcal{A}) - f(\mathcal{A}) - \langle V(\Delta\mathcal{A}), \Delta\mathcal{A}\rangle_S = o(\|\Delta\mathcal{A}\|_S), \ \forall\, V(\Delta\mathcal{A}) \in \partial_C f(\mathcal{A} + \Delta\mathcal{A}).$$

*In the symmetric matrix case, there is no need to distinguish the geometric multiplicity and algebraic multiplicity

Moreover, $f : S \to \mathbb{R}$ is said to be ρth-order semismooth function for some $\rho \in (0,1]$ at $\mathcal{A} \in S$ if

$$f(\mathcal{A}+\Delta\mathcal{A}) - f(\mathcal{A}) - \langle V(\Delta\mathcal{A}), \Delta\mathcal{A}\rangle_S = O(\|\Delta\mathcal{A}\|_S^{1+\rho}),\ \forall V(\Delta\mathcal{A}) \in \partial_C f(\mathcal{A}+\Delta\mathcal{A}).$$

In particular, if $\rho = 1$, we say f is a strongly semismooth function. We also say $f : S \to \mathbb{R}$ is a semismooth (resp. ρth-order semismooth) function on S if f is semismooth (resp. ρth-order semismooth) at $\mathcal{A}$ for all $\mathcal{A} \in S$.

As we have shown in the last section, the maximum eigenvalue function is indeed a continuous and convex function, and so is always semismooth (e.g. see Proposition 7.4.4 and 7.4.5 of [9]). Next, we consider the following question: is the function $\mathcal{A} \mapsto \lambda_1(\mathcal{A})$ indeed ρth-order semismooth for some $\rho \in (0,1]$? Below, we first provide an affirmative answer for this question. To do this, we first recall some basic definition and facts from subanalytic geometry.

Following [1, 2, 18, 22, 21, 25], a set $C \subseteq \mathbb{R}^n$ is said to be
(i) semianalytic, if for any $x \in \mathbb{R}^n$, there exists a neighbourhood U of x such that

$$C \cap U = \bigcup_{i=1}^{l} \bigcap_{j=1}^{s} \{x \in U : f_{ij}(x) = 0,\ g_{ij}(x) < 0\}$$

for some integers l, s and some real analytic functions f_{ij}, g_{ij} on $\mathbb{R}^n$ $(1 \le i \le l,\ 1 \le j \le s)$;
(ii) subanalytic if for any $x \in C$, there exist a neighbourhood U of x and a bounded semianalytic set $Z \subseteq \mathbb{R}^{n+p}$ such that $C \cap U = \{x \in \mathbb{R}^n : (x, y) \in Z \text{ for some } y \in \mathbb{R}^p\}$.

Moreover, a function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be subanalytic if its graph $\mathrm{gph} f := \{(x, f(x)) : x \in \mathbb{R}^n\}$ is subanalytic and a vector-valued function $F : \mathbb{R}^n \to \mathbb{R}^n$ is said to be subanalytic if each of its component is subanalytic. It is clear that any analytic function is a subanalytic function and sum of subanalytic functions is also a subanalytic function. We summarize below some other basic properties of subanalytic sets and subanalytic functions:

- (S1) (cf. [18, (p1) and (p2) P.597]) Finite union (resp. intersection) of subanalytic sets is subanalytic. The Cartesian product (resp. complement, closure) of subanalytic sets is subanalytic.
- (S2) (cf. [2, Theorem 2.3]) Let C be a bounded subanalytic set in $\mathbb{R}^{n+p}$ (for some positive integers n, p) and let $\pi : \mathbb{R}^{n+p} \to \mathbb{R}^n$ be the projection defined by $\pi(x, y) = x$ for all $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^p$. Then $\pi(C)$ is a subanalytic set of $\mathbb{R}^n$.
- (S3) (cf. [18, (p4) P.597]) If f is subanalytic and $\lambda \in \mathbb{R}$, then $\{x : f(x) = \lambda\}$ $\{x : f(x) < \lambda\}$ and $\{x : f(x) \le \lambda\}$ are subanalytic.
- (S4) (nonsmooth Łojasiewicz's inequality, cf. [2, Theorem 2.1.1]) Let f be a continuous subanalytic function on $\mathbb{R}^n$. Let a be a critical point of f, i.e., $0 \in \partial_L f(a)$ where ∂_L is the Mordukhovich (limiting) subdifferential [27, 28] † . Then, there exist $\epsilon > 0$, $M > 0$ and a rational number $\theta \in [0, 1)$ such that

  $$\frac{|f(x) - f(a)|^\theta}{m_f(x)} \le M \text{ for all } x \in \mathbb{B}(a, \epsilon),$$

  where $m_f(x) := d(0, \partial_L f(x)) = \inf\{\|x^*\| : x^* \in \partial_L f(x)\}$.

†The Mordukhovich (limiting) subdifferential of a function f at $x \in \mathbb{R}^n$ is denoted by $\partial_L f(x)$ and is defined by $\partial_L f(x) = \{\lim_{k\to\infty} x_k^* : x_k^* \in \partial_F f(x_k)\ x_k \to x\}$ and $\partial_F f(x)$ is the Frechét subdifferential of f defined by $\partial_F f(x) = \{x^* : \liminf_{\|h\|\to 0} \frac{f(x+h)-f(x)-(x^*)^T h}{\|h\|} \ge 0\}$. It is known that the limiting subdifferential coincides with the Clarke subdifferential for continuous and convex functions as well as continuously differentiable functions. For more detailed background information and motivations, see the informative two-volume book [27, 28].

We now show that the maximum eigenvalue function is ρth-order semismooth for some $\rho \in (0, 1]$. We note that this conclusion can also be derived by using the recently established important result ([3]): any tame function (maximum eigenvalue function is a tame function) is ρth-order semismooth for some $\rho > 0$. However, for the convenience of the reader, we provide a direct and self-contained proof proof below.

Theorem 4.1. *Let $\mathcal{A}$ be an mth-order n-dimensional symmetric tensor (m is even). Then, the maximum eigenvalue function λ_1 is ρth-order semismooth at $\mathcal{A}$ for some rational number $\rho \in (0, 1]$.*

Proof. Let $\mathcal{A}$ be an arbitrary mth-order n-dimensional symmetric tensor. As λ_1 is continuous and convex, λ_1 is always directionally differentiable and the Clarke subdifferential of λ_1 coincide with the convex subdifferential of λ_1. We now show that there exists a rational number $\rho > 0$ such that, for all $w_{\Delta\mathcal{A}} \in \partial\lambda_1(\mathcal{A} + \Delta\mathcal{A})$,

$$\lim_{\|\Delta\mathcal{A}\|_S \to 0} \frac{\lambda_1(\mathcal{A} + \Delta\mathcal{A}) - \lambda_1(\mathcal{A}) - \langle \Delta\mathcal{A}, (w_{\Delta\mathcal{A}})^m \rangle_S}{\|\Delta\mathcal{A}\|_S^{1+\rho}} < +\infty. \tag{4.3}$$

Note that if (4.3) is true with $\rho = \rho_0 > 0$, then it is also true for any ρ with $\rho \leq \rho_0$ (and so, for some rational number ρ with $\rho \leq \rho_0$). Thus, to see the conclusion, we only need to show that there exists $\rho > 0$ such that, for all $w_{\Delta\mathcal{A}} \in \partial\lambda_1(\mathcal{A} + \Delta\mathcal{A})$,

$$\lim_{\|\Delta\mathcal{A}\|_S \to 0} \frac{\lambda_1(\mathcal{A} + \Delta\mathcal{A}) - \lambda_1(\mathcal{A}) - \langle \Delta\mathcal{A}, (w_{\Delta\mathcal{A}})^m \rangle_S}{\|\Delta\mathcal{A}\|_S^{1+\rho}} < +\infty. \tag{4.4}$$

Suppose that this claim is not true. Then, there exist $\rho_k \to 0^+$, $w_k \in \partial\lambda_1(\mathcal{A} + \Delta\mathcal{A}_k)$ and $\Delta\mathcal{A}_k$ with $\|\Delta\mathcal{A}_k\|_S \to 0$ such that

$$\lim_{k \to \infty} \frac{|\lambda_1(\mathcal{A} + \Delta\mathcal{A}_k) - \lambda_1(\mathcal{A}) - \langle \Delta\mathcal{A}_k, w_k^m \rangle_S|}{\|\Delta\mathcal{A}_k\|_S^{1+\rho_k}} = +\infty.$$

As $w_k \in \partial\lambda_1(\mathcal{A} + \Delta\mathcal{A}_k)$ and $\|\Delta\mathcal{A}_k\|_S \to 0$, we see that $\{w_k\}$ is bounded. So, we may assume that $w_k \to w$ for some $w \in \partial\lambda_1(\mathcal{A})$. Recall that the space of mth-order n-dimensional symmetric tensor is denoted by S. Let the dimension of S be N. Then, there is an invertible linear map L such that $L(\mathbb{R}^N) = S$. Now, consider a function $g : \mathbb{R}^N \times \mathbb{R}^n \to \mathbb{R}$ by

$$g(x, v) = \lambda_1(Lx) - \langle Lx - \mathcal{A}, v^m \rangle_S.$$

Then, we have

$$\partial_C g(x, v) = \{L^*(\partial\lambda_1(Lx) - v^m)\} \times \{m(Lx - \mathcal{A})v^{m-1}\}.$$

Thus, we see that $(L^{-1}(\mathcal{A}), w)$ is a Clarke's critical point of g, i.e., $0 \in \partial_C g(L^{-1}(\mathcal{A}), w)$.

We first show that g is a subanalytic function. To see this, we only need to show that $h(x) := \lambda_1(Lx)$ is subanalytic as $(x, v) \mapsto -\langle Lx - \mathcal{A}, v^m \rangle_S$ is analytic. Now, consider the set $\{(x, r) : \lambda_1(Lx) > r\}$. Then we have

$$\begin{aligned}
\{(x, r) : \lambda_1(Lx) > r\} &= \{(x, r) : \sup_{\|a\|_m = 1,\, a \in \mathbb{R}^n} (Lx)a^m > r\} \\
&= \{(x, r) : \exists a \text{ with } \|a\|_m^m = 1, (Lx)a^m > r\} \\
&= \pi(C),
\end{aligned}$$

where $\pi : \mathbb{R}^N \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^N \times \mathbb{R}$ is the projection defined by $\pi(x, r, a) = (x, r)$ and $C := \{(x, r, a) : (Lx)a^m > r\}$. Note that $h(x, r, a) = (Lx)a^m - r$ is an analytic function, and so, property (S3) implies that C is a subanalytic set. Thus, $\{(x, r) : \lambda_1(Lx) > r\}$ is also a subanalytic set by property (S2). This together with property (S1) implies that

$$\begin{aligned}
\mathrm{gph} h &= \{(x, r) : \lambda_1(Lx) \geq r\} \cap \{(x, r) : \lambda_1(Lx) \leq r\} \\
&= \left(\overline{\{(x, r) : \lambda_1(Lx) > r\}}\right) \cap \left(\mathbb{R}^{n+1} \backslash \{(x, r) : \lambda_1(Lx) > r\}\right)
\end{aligned}$$

is a subanalytic set. Therefore, h is a subanalytic function, and so, g is also subanalytic.

As g is the sum of a convex function and an analytic function, the Mordukhovich (limiting) subdifferential and the Clarke subdifferential of g coincide (cf. [2, page 1211]). Thus, property (S4) implies that there exist $\alpha, \epsilon > 0$ and $\theta \in [0, 1)$ such that

$$|g(x,v) - g(L^{-1}(\mathcal{A}), w)|^{\theta} \le \alpha d\big((0,0), \partial_C g(x,v)\big) \text{ for all } x \in \mathbb{B}(L^{-1}(\mathcal{A}), \epsilon),\ v \in \mathbb{B}(w, \epsilon).$$

That is to say, for all $x \in \mathbb{B}(L^{-1}(\mathcal{A}), \epsilon)$, $v \in \mathbb{B}(w, \epsilon)$, we have

$$\begin{aligned} |\lambda_1(Lx) - \lambda_1(\mathcal{A}) - \langle Lx - \mathcal{A}, v^m\rangle_S|^{\theta} &\le \alpha\big(d(0, L^*(\partial\lambda_1(Lx) - v^m)) + m\|(Lx - \mathcal{A})v^{m-1}\|\big) \\ &= \alpha\big(\|L^*\| d(v^m, \partial\lambda_1(Lx)) + m\|(Lx - \mathcal{A})v^{m-1}\|\big). \end{aligned}$$

Now, take $x = L^{-1}(\mathcal{A} + \Delta\mathcal{A}_k)$ and $v = w_k$. As $w_k \in \partial\lambda_1(\mathcal{A} + \Delta\mathcal{A}_k)$, for all large k we have

$$|\lambda_1(\mathcal{A} + \Delta\mathcal{A}_k) - \lambda_1(\mathcal{A}) - \langle \Delta\mathcal{A}_k, w_k^m\rangle_S|^{\theta} \le m\alpha\|\Delta\mathcal{A}_k w_k^{m-1}\| \le M\|\Delta\mathcal{A}_k\|_S, \tag{4.5}$$

for some $M > 0$. As $\|\Delta\mathcal{A}_k\|_S \to 0$, we see that $\theta > 0$. This gives that $\frac{1}{\theta} \ge 1 + \rho_k$ for all large k. It follows from (4.5) that, for all large k ,

$$\frac{|\lambda_1(\mathcal{A} + \Delta\mathcal{A}_k) - \lambda_1(\mathcal{A}) - \langle \Delta\mathcal{A}_k, w_k^m\rangle_S|}{\|\Delta\mathcal{A}_k\|_S^{1+\rho_k}} \le \frac{|\lambda_1(\mathcal{A} + \Delta\mathcal{A}_k) - \lambda_1(\mathcal{A}) - \langle \Delta\mathcal{A}_k, w_k^m\rangle_S|}{\|\Delta\mathcal{A}_k\|_S^{\frac{1}{\theta}}} \le M.$$

This makes contradiction and so, the conclusion follows. □

In the preceding theorem, the order ρ is unknown in general. Next, we present some special cases where the order ρ can be estimated explicitly. To do this, for each mth-order n-dimensional symmetric tensor $\mathcal{H}$, we consider the following parameterized problem

$$\begin{array}{lll} (P_{\mathcal{H}}) & \max_{x \in \mathbb{R}^n} & \mathcal{H}x^m \\ & \text{s.t.} & \|x\|_m^m = 1, \end{array}$$

and study the Hölder stability of the optimal solution set $S_{\mathcal{H}}$ of the problem $P_{\mathcal{H}}$. To achieve the Hölder stability, we need the following result which gives an effective estimate for the Łojasiewicz exponent of a polynomial with real coefficients.

Lemma 4.1. *(cf. [11, Theorem 2.3] and [20, Lemma 4.3]) Let f be a polynomial with real coefficients on $\mathbb{R}^n$ with degree m. Suppose that $f(0) = 0$ and 0 is a strict local minimizer. Then, there exist $\delta > 0$ and $\beta > 0$ such that*

$$f(x) \ge \beta\|x\|^{\tau} \text{ for all } \|x\| \le \delta.$$

where τ is a natural number with $\tau \le (m-1)^n + 1$.

Below, we provide a Hölder stability result of the optimal solution set of the problem $P_{\mathcal{H}}$ at $\mathcal{H} = \mathcal{A}$ where $\mathcal{A}$ is a given mth-order n-dimensional symmetric tensor. Recall that, for an $(n \times n)$ matrix M, $M \prec 0$ (resp. $M \preceq 0$) means that M is negative definite (resp. negative semi-definite).

Proposition 4.1. *Let $\mathcal{A}$ be an mth-order n-dimensional symmetric tensor such that $P_{\mathcal{A}}$ has only finitely many solutions.*

(1) *Then, there exist $\epsilon > 0$ and $\alpha > 0$ such that for any tensor $\mathcal{B}$ with $\|\mathcal{B} - \mathcal{A}\|_S \le \epsilon$,*

$$S_{\mathcal{B}} \subseteq S_{\mathcal{A}} + \alpha\big(\|\mathcal{B} - \mathcal{A}\|_S\big)^{\frac{1}{d}}\,\mathbb{B}_{\mathbb{R}^n}, \tag{4.6}$$

where d is a natural number with $d \le (2m-1)^n$, $S_{\mathcal{A}}$ (resp. $S_{\mathcal{B}}$) is the optimal solution set of $P_{\mathcal{A}}$ (resp. $P_{\mathcal{B}}$) and $\mathbb{B}_{\mathbb{R}^n}$ is the unit ball in $\mathbb{R}^n$.

(2) *If we further assume that the following second-order condition holds:* $\forall\, u = (u_1, \ldots, u_n) \in S_{\mathcal{A}}$

$$\mathcal{A}u^{m-2} - \lambda_1(A)\mathrm{diag}(u_1^{m-2}, \ldots, u_n^{m-2}) \prec 0 \text{ on } C_u = \{h \in \mathbb{R}^n : h^T u^{[m-1]} = 0\}, \tag{4.7}$$

then the integer d *in (4.6) can be set as* 1*, i.e., there exist* $\epsilon > 0$ *and* $\alpha > 0$ *such that for any tensor* $\mathcal{B}$ *with* $\|\mathcal{B} - \mathcal{A}\|_S \le \epsilon$*,*

$$S_{\mathcal{B}} \subseteq S_{\mathcal{A}} + \alpha\|\mathcal{B} - \mathcal{A}\|_S\, \mathbb{B}_{\mathbb{R}^n}.$$

Proof. [Proof of (1)] Fix any $u \in S_{\mathcal{A}}$. Let $\gamma > 0$ and let $f(x) := \lambda_1(\mathcal{A})\|x + u\|_m^m - \mathcal{A}(x + u)^m + \gamma(\|x + u\|_m^m - 1)^2$. It can be seen that f is a real polynomial on $\mathbb{R}^n$ with degree $2m$. Moreover, it can be verified that f is nonnegative, $f(0) = 0$ and 0 is a strict local minimizer of f (otherwise, there exist $x_n \to 0$ with $x_n \neq 0$ such that $f(x_n) = 0$. This implies that $u_n := x_n + u$ are global solutions of $P_{\mathcal{A}}$. This is impossible as problem $P_{\mathcal{A}}$ has only finitely many solutions.) So, Lemma 4.1 implies that there exist $\delta_0 > 0$ and $\beta_0 > 0$ such that

$$f(x) \ge \beta_0\|x\|^\tau \text{ for all } \|x\| \le \delta_0,$$

where τ is a natural number with $\tau \le (2m - 1)^n + 1$. This implies that

$$\lambda_1(\mathcal{A})\|x\|_m^m - \mathcal{A}x^m + \gamma(\|x\|_m^m - 1)^2 = f(x - u) \ge \beta_0\|x - u\|^\tau \text{ for all } x \in \mathbb{R}^n \text{ with } \|x - u\| \le \delta_0.$$

As $S_{\mathcal{A}}$ only consists of finitely many points, we can find $\delta > 0$ and $\beta > 0$ such that for any $u \in S_{\mathcal{A}}$ and for an $x \in \mathbb{R}^n$ with $\|x - u\| \le \delta$,

$$\lambda_1(\mathcal{A})\|x\|_m^m - \mathcal{A}x^m + \gamma(\|x\|_m^m - 1)^2 \ge \beta\|x - u\|^\tau. \tag{4.8}$$

For the above δ, there exists $\epsilon > 0$ such that for any $\mathcal{B} \in S$ with $\|\mathcal{B} - \mathcal{A}\|_S \le \epsilon$

$$S_{\mathcal{B}} \subseteq S_{\mathcal{A}} + \delta\, \mathbb{B}_{\mathbb{R}^n}. \tag{4.9}$$

(Otherwise, there exists a sequence of symmetric tensors B_k with $\|\mathcal{B}_k - \mathcal{A}\|_S \to 0$ and $y_n \in S_{\mathcal{B}_k}$ such that $d(y_k, S_{\mathcal{A}}) \ge \delta$. As $\|y_k\|_m = 1$, by passing to subsequence, we may assume that $y_k \to y$ for some y with $\|y\|_m = 1$. Then, we have $d(y, S_{\mathcal{A}}) \ge \delta > 0$. Now, as λ_1 is continuous, we have $\mathcal{B}_k(y_k)^m = \lambda_1(\mathcal{B}_k) \to \lambda_1(\mathcal{A})$. So, passing to limit, we see that $\mathcal{A}y^m = \lambda_1(\mathcal{A})$. Note that $\|y\|_m = 1$. So, $y \in S_{\mathcal{A}}$ which makes contradiction.) Now, fix any arbitrary $\mathcal{B}$ with $\|\mathcal{B} - \mathcal{A}\|_S \le \epsilon$. From (4.9), we have $S_{\mathcal{B}} \subseteq S_{\mathcal{A}} + \delta\mathbb{B}_{\mathbb{R}^n}$. Let $r > 0$ be a constant such that

$$\|x^m - y^m\|_S \le r\|x - y\| \text{ for all } x, y \in K := \{x : \|x\|_m^m = 1\} \tag{4.10}$$

and let $\alpha = (\beta^{-1}r)^{\frac{1}{\tau-1}} > 0$. For any arbitrary $v \in S_{\mathcal{B}}$, take $u \in S_{\mathcal{A}}$ be such that $\|v - u\| = d(v, S_{\mathcal{A}})$. To finish the proof, note that $\tau \le (2m - 1)^n + 1$ (and so, $d = \tau - 1 \le (2m - 1)^n$), it suffices to show that

$$\|v - u\| \le \alpha\|\mathcal{B} - \mathcal{A}\|_S^{\frac{1}{\tau-1}}. \tag{4.11}$$

To see this, we first note that $\|v - u\| \le \delta$ as $S_{\mathcal{B}} \subseteq S_{\mathcal{A}} + \delta\, \mathbb{B}_{\mathbb{R}^n}$. So, (4.8) and $\|v\|_m = 1$ (as v is a solution of $S_{\mathcal{B}}$) imply that

$$\lambda_1(\mathcal{A}) - \mathcal{A}v^m = \lambda_1(\mathcal{A})\|v\|_m^m - \mathcal{A}v^m + \gamma(\|v\|_m^m - 1)^2 \ge \beta\|v - u\|^\tau.$$

As u is an optimal solution of $P_{\mathcal{A}}$, we have $\|u\|_m = 1$ and $\mathcal{A}u^m = \lambda_1(\mathcal{A})$ and so,

$$\|v - u\|^\tau \le \beta^{-1}\big(\lambda_1(\mathcal{A}) - \mathcal{A}v^m\big) = \beta^{-1}\big(\mathcal{A}u^m - \mathcal{A}v^m\big). \tag{4.12}$$

Since v is an optimal solution of $(P_{\mathcal{B}})$, $\|v\|_m = 1$ and $\mathcal{B}u^m \leq \mathcal{B}v^m$. It follows from (4.10) that

$$\begin{aligned}\mathcal{A}u^m - \mathcal{A}v^m &= \big(\mathcal{B}u^m - \mathcal{B}v^m\big) + \big((\mathcal{A}-\mathcal{B})u^m - (\mathcal{A}-\mathcal{B})v^m\big) \\ &\leq (\mathcal{A}-\mathcal{B})u^m - (\mathcal{A}-\mathcal{B})v^m \\ &\leq \|u^m - v^m\|_S \, \|\mathcal{A}-\mathcal{B}\|_S \\ &\leq r\|u-v\| \, \|\mathcal{B}-\mathcal{A}\|_S.\end{aligned}$$

This together with (4.12) implies that

$$\|v-u\|^\tau \leq \beta^{-1}\big(\mathcal{A}u^m - \mathcal{A}v^m\big) \leq \beta^{-1} r\|u-v\| \, \|\mathcal{B}-\mathcal{A}\|_S.$$

So, we have $\|v-u\|^{\tau-1} \leq \beta^{-1} r\|\mathcal{B}-\mathcal{A}\|_S$. Note that $\alpha = (\beta^{-1}r)^{\frac{1}{\tau-1}}$. Then,

$$d(v, S_{\mathcal{A}}) = \|v-u\| \leq \alpha\|\mathcal{B}-\mathcal{A}\|_S^{\frac{1}{\tau-1}}.$$

Therefore, (4.11) holds, and so, (1) follows.

[Proof of (2)] Suppose that the second-order condition (4.7) holds. Fix an arbitrary $u \in S_{\mathcal{A}}$ and consider the minimization problem

$$\begin{array}{lll}(P_0) & \min_{x\in\mathbb{R}^n} & f(x) := -\mathcal{A}x^m \\ & \text{s.t.} & \|x\|_m^m = 1.\end{array}$$

Clearly, u satisfy the KKT condition of (P_0) with Lagrange multiplier $\lambda_1(\mathcal{A})$. Note that the usual second-order sufficient condition for this problem reduces to (4.7). So, the following second order growth condition holds at u (for example see [10, Corollary 1]): there exist $\beta > 0$ and $\delta > 0$ such that

$$-\mathcal{A}x^m + \lambda_1(\mathcal{A}) = f(x) - f(u) \geq \beta\,\|x-u\|^2 \text{ for all } x \in \mathbb{R}^n \text{ with } \|x-u\| \leq \delta.$$

Now, using the same method of the proof as in part (1), we see that the conclusion holds. $\square$

We are now ready to present the second main result of this section as follows.

Theorem 4.2. *Let $\mathcal{A}$ be an mth-order n-dimensional symmetric tensor (m is even). Suppose that the geometric multiplicity of $\lambda_1(\mathcal{A})$ is one. Then,*

(1) *The maximum eigenvalue function λ_1 is at least ρth-order semismooth at $\mathcal{A}$ with $\rho = \frac{1}{(2m-1)^n}$.*

(2) *Moreover, if we further assume that the following second-order condition holds: for all eigenvectors u associated with $\lambda_1(\mathcal{A})$*

$$\mathcal{A}u^{m-2} - \lambda_1(\mathcal{A})\mathrm{diag}(u_1^{m-2},\ldots,u_n^{m-2}) \prec 0 \text{ on } C_u = \{h \in \mathbb{R}^n : h^T u^{[m-1]} = 0\}, \tag{4.13}$$

then the maximum eigenvalue function λ_1 is strongly semismooth at $\mathcal{A}$.

Proof. [Proof of (1)] Let $\mathcal{A}$ be an arbitrary mth-order n-dimensional symmetric tensor and let $\rho = \frac{1}{(2m-1)^n}$. Let $\Delta\mathcal{A}$ be an mth-order n-dimensional symmetric tensor. As λ_1 is continuous and convex, λ_1 is always directionally differentiable. To see the conclusion, we only need to show that, for all $V(\Delta\mathcal{A}) \in \partial(\mathcal{A}+\Delta\mathcal{A})$,

$$\lambda_1(\mathcal{A}+\Delta\mathcal{A}) - \lambda_1(\mathcal{A}) - \langle V(\Delta\mathcal{A}), \Delta\mathcal{A}\rangle_S = O(\|\Delta\mathcal{A}\|_S^{1+\rho}). \tag{4.14}$$

By the mean value theorem, we have $\lambda_1(\mathcal{A}+\Delta\mathcal{A}) - \lambda_1(\mathcal{A}) = \langle U(\Delta\mathcal{A}), \Delta\mathcal{A}\rangle_S$, for some $U(\Delta\mathcal{A}) \in \partial\lambda_1(\mathcal{A}+t\Delta\mathcal{A})$ and $t \in [0,1]$. So, we see that

$$\frac{\lambda_1(\mathcal{A}+\Delta\mathcal{A}) - \lambda_1(\mathcal{A}) - \langle V(\Delta\mathcal{A}), \Delta\mathcal{A}\rangle_S}{\|\Delta\mathcal{A}\|_S^{1+\rho}} = \frac{\langle U(\Delta\mathcal{A}) - V(\Delta\mathcal{A}), \Delta\mathcal{A}\rangle_S}{\|\Delta\mathcal{A}\|_S^{1+\rho}}.$$

Denote the dimension of the symmetric tensor space S by k. Then, the Carathedory theorem and Theorem 3.2 imply that there exist $\mu_j \in [0,1]$, $j = 1, \ldots, k+1$, with $\sum_{j=1}^{k+1} \mu_j = 1$ such that $U(\Delta A) = \sum_{j=1}^{k+1} \mu_j (u_j)^m$, where u_j are all unit eigenvectors (in the $\|\cdot\|_m$ sense) associated with $\lambda_1(\mathcal{A} + t\Delta\mathcal{A})$. Similarly, we can find $\sigma_j \in [0,1]$, $j = 1, \ldots, k+1$, with $\sum_{j=1}^{k+1} \sigma_j = 1$ such that $V(\Delta A) = \sum_{j=1}^{k+1} \sigma_j (v_j)^m$, where v_j are all unit eigenvectors associated with $\lambda_1(\mathcal{A}+\Delta\mathcal{A})$. Now, for each mth-order n-dimensional symmetric tensor $\mathcal{H}$, consider the parameterized problem

$$\begin{array}{lll} (P_{\mathcal{H}}) & \max_{x \in \mathbb{R}^n} & \mathcal{H}x^m \\ & \text{s.t.} & \|x\|_m^m = 1. \end{array}$$

We see that each u_j is an optimal solution of $(P_{\mathcal{A}+t\Delta\mathcal{A}})$ and each v_j is an optimal solution of $(P_{\mathcal{A}+\Delta\mathcal{A}})$. Let C be the optimal solution of $(P_{\mathcal{A}})$. As the geometric multiplicity of $\lambda_1(\mathcal{A})$ is one, the set C only consists finitely many points. Note that $\|\mathcal{A} + t\Delta\mathcal{A} - \mathcal{A}\|_S \le \|\Delta\mathcal{A}\|_S$ and $\|\mathcal{A} + \Delta\mathcal{A} - \mathcal{A}\|_S \le \|\Delta\mathcal{A}\|_S$. Then, Proposition 4.1(1) implies that, there exist $\alpha > 0$ and $\epsilon > 0$ such that, for each $j = 1, \ldots, k+1$ and $\|\Delta\mathcal{A}\|_S \le \epsilon$

$$d(u_j, C) \le \alpha \|\Delta\mathcal{A}\|_S^\rho \text{ and } d(v_j, C) \le \alpha \|\Delta\mathcal{A}\|_S^\rho.$$

Let $\underline{c}_j, \overline{c}_j \in C$ be such that $d(u_j, C) = \|u_j - \underline{c}_j\|$ and $d(v_j, C) = \|v_j - \overline{c}_j\|$. Then,

$$\|u_j - \underline{c}_j\| \le \alpha \|\Delta\mathcal{A}\|_S^\rho \text{ and } \|v_j - \overline{c}_j\| \le \alpha \|\Delta\mathcal{A}\|_S^\rho.$$

Note that there exists $r > 0$ such that $\|x^m - y^m\|_S \le r\|x - y\|$ for all $x, y \in K := \{x : \|x\|_m^m = 1\}$. It follows that

$$\|(u_j)^m - (\underline{c}_j)^m\|_S \le r\alpha \|\Delta\mathcal{A}\|_S^\rho \text{ and } \|(v_j)^m - (\overline{c}_j)^m\|_S \le r\alpha \|\Delta\mathcal{A}\|_S^\rho.$$

As $\partial\lambda_1(\mathcal{A}) = \{w^m\}$, we see that for any $x \in C$, $x^m = w^m$, and so, $(\underline{c}_j)^m = (\overline{c}_j)^m = w^m$. Thus, we have, for each $j = 1, \ldots, k+1$ and $\|\Delta\mathcal{A}\|_S \le \epsilon$

$$\|(u_j)^m - w^m\|_S \le r\alpha \|\Delta\mathcal{A}\|_S^\rho \text{ and } \|(v_j)^m - w^m\|_S \le r\alpha \|\Delta\mathcal{A}\|_S^\rho.$$

Therefore, for each $\|\Delta\mathcal{A}\|_S \le \epsilon$

$$\begin{aligned} |\langle U(\Delta\mathcal{A}) - V(\Delta\mathcal{A}), \Delta\mathcal{A}\rangle_S| &\le |\langle U(\Delta\mathcal{A}) - w^m, \Delta\mathcal{A}\rangle_S| + |\langle w^m - V(\Delta\mathcal{A}), \Delta\mathcal{A}\rangle_S| \\ &\le \big(\|U(\Delta\mathcal{A}) - w^m\|_S + \|V(\Delta\mathcal{A}) - w^m\|_S\big)\|\Delta\mathcal{A}\|_S \\ &= \big(\sum_{j=1}^{k+1} \mu_j \|(u_j)^m - w^m\|_S + \sum_{j=1}^{k+1} \sigma_j \|(v_j)^m - w^m\|_S\big)\|\Delta\mathcal{A}\|_S \\ &\le 2r\alpha \|\Delta\mathcal{A}\|_S^{1+\rho}. \end{aligned}$$

So, there exist $r > 0$, $\alpha > 0$ and $\epsilon > 0$ such that for each $\|\Delta\mathcal{A}\|_S \le \epsilon$,

$$\left|\frac{\lambda_1(\mathcal{A} + \Delta\mathcal{A}) - \lambda_1(\mathcal{A}) - \langle V(\Delta\mathcal{A}), \Delta\mathcal{A}\rangle_S}{\|\Delta\mathcal{A}\|_S^{1+\rho}}\right| = \left|\frac{\langle U(\Delta\mathcal{A}) - V(\Delta\mathcal{A}), \Delta\mathcal{A}\rangle_S}{\|\Delta\mathcal{A}\|^{1+\rho}}\right| \le 2r\alpha.$$

Therefore, (4.14) holds, and hence λ_1 is a ρth-order semismooth function on S with $\rho = \frac{1}{(2m-1)^n}$.

[Proof of (2)] Using the same method of proof as in part (1) and using Proposition 4.1(2) (instead of Proposition 4.1(1)), we see that the conclusion follows. $\square$

Remark 4.1. *It is worth noting that the condition (4.13) holds automatically in the symmetric matrix case (i.e., when $m = 2$) if the multiplicity of the maximum eigenvalue is one. Indeed, let*

$m = 2$ and consider an $(n \times n)$ symmetric matrix M. To see the desire conclusion, it suffices to show that: for all unit eigenvectors u associated with $\lambda_1(M)$,

$$M - \lambda_1(M)I_n \prec 0 \text{ on } C_u = \{h \in \mathbb{R}^n : h^T u = 0\}, \tag{4.15}$$

where I_n is the $(n \times n)$ identity matrix. To see this, let u be an unit eigenvector associated with $\lambda_1(M)$ and let $h \in \mathbb{R}^n \backslash \{0\}$ with $h^T u = 0$. As the multiplicity of $\lambda_1(M)$ is one, we can find $\{v_2, \ldots, v_n\}$ which are unit eigenvectors associated with $\{\lambda_2(\mathcal{A}), \ldots, \lambda_n(M)\}$ such that $\{u, v_2, \ldots, v_n\}$ forms an orthonormal basis of $\mathbb{R}^n$ where $\lambda_1(M) > \lambda_2(M) \geq \ldots \geq \lambda_n(\mathcal{A})$ (counting the multiplicity). As $\{v_2, \ldots, v_n\}$ span the space $\{x \in \mathbb{R}^n : x^T u = 0\}$, we have $h = \sum_{i=2}^n \alpha_i v_i$ for some $\alpha_i \in \mathbb{R}$ with α_i are not all 0, $i = 2, \ldots, n$. So, it follows from the fact that $\{u, v_2, \ldots, v_n\}$ is an orthonormal basis of $\mathbb{R}^n$ and $\{v_2, \ldots, v_n\}$ are unit eigenvectors associated with $\{\lambda_2(\mathcal{A}), \ldots, \lambda_n(M)\}$ that

$$h^T\big(M - \lambda_1(M)I\big)h = (\sum_{i=2}^n \alpha_i v_i)^T\big(M - \lambda_1(M)I_n\big)(\sum_{i=2}^n \alpha_i v_i) = \sum_{i=2}^n \alpha_i^2\big(\lambda_i(M) - \lambda_1(M)\big) < 0.$$

Thus, (4.13) holds when $m = 2$.

5 Nonsmooth Newton Methods

Let $S(4,3)$ be the space consisting of all the 4th-order 3-dimensional symmetric tensors. It is shown [36] that $S(4,3)$ is of dimension 15, and so, there exists a one-to-one mapping $L : \mathbb{R}^{15} \to S(4,3)$ (indeed the mapping L can be explicitly constructed see [36, 35] for details). Let n be a natural number, $\mathcal{A}_i \in S(4,3)$, $i = 0, 1, \ldots, n$ and $b_i \in \mathbb{R}$, $i = 1, \ldots, n$. Consider the space tensor conic linear programming (STCLP) problem which was proposed and studied in [35]:

$$\begin{array}{lll} (STCLP) & \min\limits_{\mathcal{X} \in S(4,3)} & \langle \mathcal{A}_0, \mathcal{X} \rangle_S \\ & \text{s.t.} & \langle \mathcal{A}_i, \mathcal{X} \rangle_S \leq b_i,\ i = 1, \ldots, n, \\ & & \mathcal{X} \in -\mathcal{C}(4,3), \end{array}$$

where $\mathcal{C}(4,3)$ is the cone of all negative semidefinite 4th-order 3-dimensional symmetric tensor, i.e., $\mathcal{C}(4,3) := \{\mathcal{A} \in S(4,3) : \mathcal{A}x^4 \leq 0, \forall\, x \in \mathbb{R}^3\}$. The problem (STCLP) arises from the medical imaging area where a high order tensor is used to describe the non-Gaussian diffusion feature [36].

Note that $\mathcal{C}(4,3) = \{\mathcal{A} \in S(4,3) : \lambda_1(\mathcal{A}) \leq 0\}$ and λ_1 is convex. We see that problem (STCLP) is a convex programming problem. By assuming the Slater constraint qualification, solving (STCLP) is equivalent to solving its KKT system. As shown in [37] (see also [35]), $\mathcal{C}(4,3)$ is a self-dual cone and $\big(\mathcal{C}(4,3)\big)^* = \mathcal{C}(4,3) = \mathcal{U}(4,3)$ where $\mathcal{U}(4,3)$ is the rank one tensor space in $S(4,3)$ defined by $\mathcal{U}(4,3) = \{\mathcal{A} \in S(4,3) : \mathcal{A} = \sum_{j=1}^r (a_j)^4,\ a_j \in \mathbb{R}^3, r \in \mathbb{N}\}$, where a^4 is the 4th-order 3-dimensional symmetric rank one tensor defined by $(a^4)_{i_1 i_2 i_3 i_4} = a_{i_1} a_{i_2} a_{i_3} a_{i_4}$ for each $i_1, i_2, i_3, i_4 \in \{1,2,3\}$. So, its KKT system is to find $y_1, \ldots, y_n \in \mathbb{R}$ and $\mathcal{X} \in S(4,3)$ such that

$$(KKT)\ \left\{\begin{array}{ll} \mathcal{X} \in -\mathcal{C}(4,3) & \text{(Primal Feasible)} \\ -\mathcal{A}_0 - \sum_{i=1}^n y_i \mathcal{A}_i \in \mathcal{C}(4,3) & \text{(Dual Feasible)} \\ 0 \leq y_i \perp \big(\langle \mathcal{A}_i, \mathcal{X} \rangle_S - b_i\big) \leq 0 & \text{(Complementary Slackness I)} \\ \langle \mathcal{A}_0 + \sum_{i=1}^n y_i \mathcal{A}_i, \mathcal{X} \rangle_S = 0 & \text{(Complementary Slackness II)} \end{array}\right. . \tag{5.16}$$

The following proposition establishes a useful observation: solving the KKT problem is equivalent to solving the nonsmooth equation $F(x) = 0$ where $F : \mathbb{R}^{n+15} \to \mathbb{R}^{n+2}$ is defined by

$$F(x) = \begin{pmatrix} \max\{\lambda_1(-\mathcal{A}_0 - \sum_{i=1}^n y_i \mathcal{A}_i), \lambda_1(-Lz)\} \\ \langle \mathcal{A}_0 + \sum_{i=1}^n y_i \mathcal{A}_i, Lz \rangle_S \\ \max\{-y_1, \langle \mathcal{A}_1, Lz \rangle_S - b_1\} \\ \vdots \\ \max\{-y_n, \langle \mathcal{A}_n, Lz \rangle_S - b_n\} \end{pmatrix},\ x = (y,z) \in \mathbb{R}^n \times \mathbb{R}^{15} \tag{5.17}$$

where $L : \mathbb{R}^{15} \to S(4,3)$ is the one-to-one linear mapping between $\mathbb{R}^{15}$ and $S(4,3)$.

Proposition 5.1. *Let $y \in \mathbb{R}^n$, $\mathcal{X} \in S(4,3)$ and $x := (y, L^{-1}(\mathcal{X})) \in \mathbb{R}^n \times \mathbb{R}^{15}$. Then, $(y, \mathcal{X}) \in \mathbb{R}^n \times S(4,3)$ solves the KKT system (5.16) if and only if $F(x) = 0$.*

Proof. [$\Rightarrow$] Suppose that $(y, \mathcal{X}) \in \mathbb{R}^n \times S(4,3)$ solves the KKT system. Note that $\mathcal{A} \in C(4,3)$ if and only if $\lambda_1(\mathcal{A}) \leq 0$. So, the primal and the dual feasibility conditions imply that $\max\{\lambda_1(-\mathcal{A}_0 - \sum_{i=1}^n y_i \mathcal{A}_i), \lambda_1(-\mathcal{X})\} \leq 0$. Note that the Complementary Slackness condition I is equivalent to

$$\max\{-y_i, \langle \mathcal{A}_i, \mathcal{X}\rangle_S - b_i\} = 0, \ i = 1, \ldots, n.$$

To finish the proof, it suffices to show that $\max\{\lambda_1(-\mathcal{A}_0 - \sum_{i=1}^n y_i \mathcal{A}_i), \lambda_1(-\mathcal{X})\} = 0$. To see this, we proceed by the method of contradiction and assume that $\max\{\lambda_1(-\mathcal{A}_0 - \sum_{i=1}^n y_i \mathcal{A}_i), \lambda_1(-\mathcal{X})\} < 0$. As $-\mathcal{A}_0 - \sum_{i=1}^n y_i \mathcal{A}_i \in \mathcal{C}(4,3)$ and $\mathcal{C}(4,3) = \mathcal{U}(4,3)$, we can decompose $-\mathcal{A}_0 - \sum_{i=1}^n y_i \mathcal{A}_i$ as the sum of finitely many rank one 4th-order 3-dimensional tensors, say

$$-\mathcal{A}_0 - \sum_{i=1}^n y_i \mathcal{A}_i = \sum_{j=1}^r (a_j)^4, \ a_j \in \mathbb{R}^3.$$

Then, from the Complementary Slackness condition II, we have

$$0 = \langle \mathcal{X}, -\mathcal{A}_0 - \sum_{i=1}^n y_i \mathcal{A}_i \rangle_S = \sum_{j=1}^r \langle \mathcal{X}, (a_j)^4 \rangle_S = \sum_{j=1}^r \mathcal{X}(a_j)^4$$

This together with $\mathcal{X} \in -\mathcal{C}(4,3)$ gives that $\mathcal{X}(a_j)^4 = 0$ for each $j = 1, \ldots, r$. Clearly, not all a_j, $j = 1, \ldots, r$, are 0 (otherwise, $-\mathcal{A}_0 - \sum_{i=1}^n y_i \mathcal{A}_i = \sum_{j=1}^r (a_j)^4 = 0$ and so, $\max\{\lambda_1(-\mathcal{A}_0 - \sum_{i=1}^n y_i \mathcal{A}_i), \lambda_1(-\mathcal{X})\} \geq 0$ which is impossible). This implies that 0 is an eigenvalue of $-\mathcal{X}$ and so, $\lambda_1(-\mathcal{X}) \geq 0$. Thus, $\max\{\lambda_1(-\mathcal{A}_0 - \sum_{i=1}^n y_i \mathcal{A}_i), \lambda_1(-\mathcal{X})\} \geq 0$ which contradicts our assumption.

[$\Leftarrow$] This part is direct. $\square$

From the proceeding proposition, to obtain a solution of the KKT system, it suffices to solve the nonsmooth underdetermined equation $F(x) = 0$ where $F : \mathbb{R}^{n+15} \to \mathbb{R}^{n+2}$. It is known that the nonsmooth Newton method [33] (see also [13] for recent development and further improvement) is now considered a powerful tool for solving nonsmooth equation where the dimension of the domain space and the range space are the same. To solve the space tensor conic linear problem, this motivates us to establish a nonsmooth Newton method for under-determined equations.

Nonsmooth Newton Method for underdetermined equation

Consider a general nonsmooth equation $G(x) = 0$ where $G : \mathbb{R}^m \to \mathbb{R}^l$ with $m \geq l$. The general algorithm of the nonsmooth Newton method for a underdetermined equation is stated as follows:

Algorithm-0

Step 0. Choose $x^{(0)} \in \mathbb{R}^m$. If $G(x^{(0)}) \neq 0$, then set $k := 0$ and go to Step 1. Otherwise, output $x^{(0)}$.

Step 1. Compute a $V_k \in \mathbb{R}^{l \times m}$ such that $V_k \in J_C G(x^{(k)})$.

Step 2. Let $x^{(k+1)} = x^{(k)} + \Delta x^{(k)}$ where

$$\Delta x^{(k)} = -V_k^T (V_k V_k^T)^{-1} G(x^{(k)}).$$

Step 3. If $G(x^{(k+1)}) \neq 0$, then replace k by $k+1$ and go back to Step 1. Otherwise, output $x^{(k+1)} \in \mathbb{R}^m$.

We now introduce a definition of regularity which extends the classical definition of CD-regular solution.

Definition 5.1. *For a nonlinear equation $G(x) = 0$ where $G : \mathbb{R}^m \to \mathbb{R}^l$ with $m \geq l$. We say x^* is a regular solution of this equation if $G(x^*) = 0$ and there exists a neighbourhood N of x^* such that*

(a) $\mathrm{rank}(V) = l$ *for all $V \in J_C G(x)$ and $x \in N \cap \{x : G(x) \neq 0\}$.*

(b) *$R(V^T(VV^T)^{-1}) \equiv P$ for all $V \in J_C G(x)$ and $x \in N \cap \{x : G(x) \neq 0\}$ where $R(A)$ denotes the range of a $(m \times l)$ matrix A which is defined by $R(A) = \{Ax : x \in \mathbb{R}^l\} \subseteq \mathbb{R}^m$ and P is some vector space in $\mathbb{R}^m$.*

Moreover, the vector space P in (b) *is called the regular space associated with x^*.*

Recall that, when $m = l$, x^* is said to be CD regular (cf [31, 33]) if $G(x^*) = 0$ and there exists a neighbourhood N of x^* such that $\mathrm{rank}(V) = m$ for all $V \in J_C G(x)$ and $x \in N$ (and so, $R(V^T(VV^T)^{-1}) \equiv \mathbb{R}^m$ for all $V \in J_C G(x)$ and $x \in N$). Thus, it is easy to see that, x^* is CD-regular implies x^* is regular with regular space $P = \mathbb{R}^m$. As shown in [31], a CD regular solution is a locally unique solution. However, we now present examples showing that a regular solution need not to be locally unique even when $m = l$ (and so, the definition of regular is strictly weaker than CD regular).

Example 5.1. *Let $m = l = 1$ and $G(x) = \max\{x, 0\}$. Consider $x^* = 0$. Clearly, x^* is not a locally unique solution and is not CD-regular (as $0 \in J_C G(0)$). On the other hand, $J_C G(x) = 1$ for any $x \in \mathbb{R}$ with $G(x) \neq 0$. So, x^* is regular with regular space $P = \mathbb{R}$.*

Example 5.2. *Let $m = 2$, $l = 1$ and $G(x_1, x_2) = \max\{x_1, 0\} + |x_2|$. It is easy to see that the solution set $\{x : G(x) = 0\} = \{(x_1, x_2) : x_1 \leq 0, x_2 = 0\}$. So $x^* = (-2, 0)$ is a solution which is not locally unique. It can be verified that for any $x \in x^* + \mathbb{B}_{\mathbb{R}^2}$ with $G(x) \neq 0$*

$$J_C G(x) = \begin{cases} (0,1), & \text{if} \quad x_2 > 0, \\ (0,-1), & \text{if} \quad x_2 < 0, \end{cases} \qquad \textit{where } x = (x_1, x_2).$$

So, x^* is regular with $P = \{0\} \times \mathbb{R}$.

We first establish the local convergence result of a nonsmooth Newton method for underdetermined equations.

Theorem 5.1. *Consider a underdetermined equation $G(x) = 0$ where $G : \mathbb{R}^m \to \mathbb{R}^l$ with $m \geq l$. Let x^* be a regular solution and P be the regular space associated with x^*.*

(a) *If G is semismooth at x^*, then there exists a neighbourhood N_0 of x^* such that the Algorithm-0 is well defined for any initial point $x^{(0)} \in N_0 \cap (x^* + P)$ and Algorithm-0 either terminates in finitely many iterations or generates a sequence $\{x^{(k)}\}$ such that $x^{(k)}$ converges to x^* Q-superlinearly, i.e.,*

$$\lim_{k \to \infty} \frac{\|x^{(k+1)} - x^*\|}{\|x^{(k)} - x^*\|} = 0.$$

(b) *If we further assume that G is ρth-order semismooth for some $\rho \in (0, 1]$. Then, Algorithm-0 either terminates in finitely many iterations or generates a sequence $\{x^{(k)}\}$ such that $x^{(k)}$ converges to x^* with order $(1 + \rho)$, i.e.,*

$$\lim_{k \to \infty} \frac{\|x^{(k+1)} - x^*\|}{\|x^{(k)} - x^*\|^{1+\rho}} < +\infty.$$

Proof. [Proof of (a)] If Algorithm-0 terminates in finitely many iterations, then by the construction of the algorithm it must return a solution of $G(x) = 0$. So, without loss of generality, we may assume that Algorithm-0 does not terminate in finitely many iterations, and so, generates a sequence $\{x^{(k)}\}_{k=1}^{+\infty}$. By the construction of the algorithm, we have $G(x^{(k)}) \neq 0$ for all $k \in \mathbb{N} \cup \{0\}$.

As x^* is regular, there exist a bounded neighbourhood N and a subspace P such that $\mathrm{rank}(V) = l$ for all $V \in J_C G(x)$ and $x \in N \cap \{x : G(x) \neq 0\}$ and $R(V^T(VV^T)^{-1}) \equiv P$ for all $V \in J_C G(x)$ and $x \in N \cap \{x : G(x) \neq 0\}$. As G is locally Lipschitz and semismooth at x^*, there exist $\delta \in (0, 1)$, $M > 0$ such that $x^* + \delta \mathbb{B}_{\mathbb{R}^m} \subseteq N$, and for all $x \in (x^* + \delta \mathbb{B}_{\mathbb{R}^m}) \cap \{x : G(x) \neq 0\}$ and for all $V \in J_C G(x)$,

$$\|(V^T(VV^T)^{-1}\| \leq M \text{ and } \|G(x^*) - G(x) + V(x - x^*)\| \leq \frac{\|x - x^*\|}{2M}. \tag{5.18}$$

Let $N_0 = x^* + \delta \mathbb{B}_{\mathbb{R}^m}$ and let $x^{(0)} \in N_0 \cap (x^* + P)$. We shall prove by mathematical induction that $\{x^{(k)}\}_{k=1}^{\infty} \subseteq (x^* + \frac{\delta}{2^k} \mathbb{B}_{\mathbb{R}^m}) \cap (x^* + P)$. To see this, we first note that

$$x^{(1)} = x^{(0)} - V_0^T(V_0V_0^T)^{-1}G(x^{(0)})$$

where $V^{(0)} \in J_C G(x^0)$. Note that $x^{(0)} \in N_0 \subseteq N$, and so, $R(V_0^T(V_0V_0^T)^{-1}) = P$. This together with $x^{(0)} - x^* \in P$ implies that

$$x^{(1)} \in x^* + P.$$

Moreover, as $x^{(0)} - x^* \in P = R(V_0^T(V_0V_0^T)^{-1})$, there exist a_0 such that $x^{(0)} - x^* = V_0^T(V_0V_0^T)^{-1}a_0$. It follows that

$$V_0^T(V_0V_0^T)^{-1}V_0(x^{(0)} - x^*) = V_0^T(V_0V_0^T)^{-1}V_0V_0^T(V_0V_0^T)^{-1}a_0 = V_0^T(V_0V_0^T)^{-1}a_0 = x^{(0)} - x^*.$$

So, we have

$$\begin{aligned}
x^{(1)} - x^* &= x^{(0)} - x^* - V_0^T(V_0V_0^T)^{-1}G(x^{(0)}) \\
&= V_0^T(V_0V_0^T)^{-1}V_0(x^{(0)} - x^*) - V_0^T(V_0V_0^T)^{-1}G(x^{(0)}) \\
&= V_0^T(V_0V_0^T)^{-1}\big(V_0(x^{(0)} - x^*) - G(x^{(0)})\big) \\
&= V_0^T(V_0V_0^T)^{-1}\big(G(x^*) - G(x^{(0)}) + V_0(x^{(0)} - x^*)\big)
\end{aligned}$$

This together with (5.18) implies that

$$\begin{aligned}
\|x^{(1)} - x^*\| &\leq \|V_0^T(V_0V_0^T)^{-1}\| \cdot \|G(x^*) - G(x^{(0)}) + V_0(x^{(0)} - x^*)\| \\
&\leq M\,\frac{\|x^{(0)} - x^*\|}{2M} \\
&= \frac{\|x^{(0)} - x^*\|}{2}.
\end{aligned}$$

So, $x^{(1)} \in (x^* + P) \cap (x^* + \frac{\delta}{2}\mathbb{B}_{\mathbb{R}^m})$ and hence the claim is true for $k = 1$. Now, suppose the the case is true when $k = s$ and consider the case when $k = s + 1$. Using the same method of proof, one can easily see that the claim is also true for $k = s + 1$. Therefore, we have $\{x^{(k)}\}_{k=1}^{\infty} \subseteq (x^* + P) \cap (x^* + \frac{\delta}{2^k}\mathbb{B}_{\mathbb{R}^m})$. In particular, we have $x^{(k)} \to x^*$. Now, fix an arbitrary $k \in \mathbb{N}$. As $x^{(k)} - x^* \in P = R(V_k^T(V_kV_k^T)^{-1})$, there exists a_k such that $x^{(k)} - x^* = V_k^T(V_kV_k^T)^{-1}a_k$. It follows that

$$V_k^T(V_kV_k^T)^{-1}V_k(x^{(k)} - x^*) = V_k^T(V_kV_k^T)^{-1}V_kV_k^T(V_kV_k^T)^{-1}a_k = V_k^T(V_kV_k^T)^{-1}a_k = x^{(k)} - x^*.$$

So, from the algorithm, we have

$$\begin{aligned}
x^{(k+1)} - x^* &= x^{(k)} - x^* - V_k^T(V_kV_k^T)^{-1}G(x^{(k)}) \\
&= V_k^T(V_kV_k^T)^{-1}V_k(x^{(k)} - x^*) - V_k^T(V_kV_k^T)^{-1}G(x^{(k)}) \\
&= V_k^T(V_kV_k^T)^{-1}\big(V_k(x^{(k)} - x^*) - G(x^{(k)})\big) \\
&= V_k^T(V_kV_k^T)^{-1}\big(G(x^*) - G(x^{(k)}) + V_k(x^{(k)} - x^*)\big).
\end{aligned}$$

Note that $\|V_k^T(V_kV_k^T)^{-1}\| \le M$ for all $k \in \mathbb{N}$ (as $V_k \in J_C G(x^{(k)})$, $x^{(k)} \in x^* + \frac{\delta}{2^k}\mathbb{B}_{\mathbb{R}^m} \subseteq N$ and (5.18)). As G is semismooth at x^*, it follows that

$$\frac{\|x^{(k+1)} - x^*\|}{\|x^{(k)} - x^*\|} \le \frac{\|V_k^T(V_kV_k^T)^{-1}\| \cdot \|G(x^*) - G(x^{(k)}) + V_k(x^{(k)} - x^*)\|}{\|x^{(k)} - x^*\|} \to 0.$$

[Proof of (b)] The proof of part (b) is similar to the proof of part (a). □

From the preceding theorem, we see that, to solve a underdetermined equation $G(x) = 0$ where $G : \mathbb{R}^m \to \mathbb{R}^l$ with $m \ge l$, an initial point need to be chosen in the intersection of the neighbourhood N_0 and a l-dimensional vector space $x^* + P$. In particular, if $m = l$, this reduces to the following classical local convergence result.

Corollary 5.1. *(cf [31]) Consider a nonsmooth equation $G(x) = 0$ where $G : \mathbb{R}^m \to \mathbb{R}^m$. Let x^* be a CD-regular solution.*

(a) *If G is semismooth at x^*, then there exists a neighbourhood N_0 of x^* such that Algorithm-0 is well defined for any initial point $x^{(0)} \in N_0$ and Algorithm-0 either terminates in finitely many iterations or generates a sequence $\{x^{(k)}\}$ such that $x^{(k)}$ converges to x^* Q-superlinearly, i.e.,*

$$\lim_{k\to\infty} \frac{\|x^{(k+1)} - x^*\|}{\|x^{(k)} - x^*\|} = 0.$$

(b) *If we further assume that G is ρth-order semismooth with $\rho \in (0, 1]$. Then, Algorithm-0 either terminates in finitely many iterations or generates a sequence $\{x^{(k)}\}$ such that $x^{(k)}$ converges to x^* with order $(1 + \rho)$, i.e.,*

$$\lim_{k\to\infty} \frac{\|x^{(k+1)} - x^*\|}{\|x^{(k)} - x^*\|^{1+\rho}} < +\infty.$$

Proof. Let $l = m$. Note that x^* is CD-regular implies x^* is regular with regular space $P = \mathbb{R}^m$. Thus, the conclusion immediately follows from the preceding theorem. □

Let F be the vector valued function defined as in (5.17). Now, we state a nonsmooth Newton method for solving the space tensor conic linear problem (STCLP):

Algorithm-1

Step 0. Choose $(y^{(0)}, \mathcal{X}^{(0)}) \in \mathbb{R}^n \times S(4, 3)$. Compute $z^{(0)} = L^{-1}(\mathcal{X}^{(0)})$ and let $x^{(0)} = (y^{(0)}, z^{(0)})$. If $F(x^{(0)}) \ne 0$, then set $k := 0$. Otherwise, output $(y^{(0)}, \mathcal{X}^{(0)})$.

Step 1. Compute a $V_k \in \mathbb{R}^{(n+2)\times(n+15)}$ such that $V_k \in J_C F(x^{(k)})$. †

Step 2. Let $x^{(k+1)} = x^{(k)} + \Delta x^{(k)}$ where $\Delta x^{(k)} = -(V_k^T V_k)^{-1} V_k^T F(x^{(k)})$.

Step 3. If $F(x^{(k+1)}) \ne 0$, then replace k by $k + 1$ and go back to Step 1. Otherwise, let $x^{(k+1)} = (y^{(k+1)}, z^{(k+1)}) \in \mathbb{R}^n \times \mathbb{R}^{15}$ and output $\big(y^{(k+1)}, L^{-1}(z^{(k+1)})\big) \in \mathbb{R}^n \times S(4, 3)$.

Below, we present the local convergence result of our Algorithm-1:

† For a 4th-order 3-dimensional tensor, one can efficiently find an eigenvector associated with its maximum eigenvalue (e.g. see [36]), and so, one can also efficiently compute a member of the (Clarke) generalized Hessian $J_C F(x^{(k)}) \subseteq \mathbb{R}^{(n+2)\times(n+15)}$.

Theorem 5.2. *Let $(y^*, \mathcal{X}^*) \in \mathbb{R}^n \times S(4,3)$ be a solution of the KKT system of (STCLP). Let $x^* = \big(y^*, L^{-1}(\mathcal{X}^*)\big) \in \mathbb{R}^n \times \mathbb{R}^{15}$. Let x^* be a regular solution and let P be the regular space associated with x^*. Then there exists a neighbourhood N of x^* such that the Algorithm-1 is well defined for any initial point $x^{(0)} \in N_0 \cap (x^* + P)$ and Algorithm-1 either terminates in finitely many iterations or generates a sequence $\{x^{(k)}\}$ such that $x^{(k)}$ converges to x^* with order $1+\rho$ for some $\rho > 0$, i.e.,*

$$\lim_{k\to\infty} \frac{\|x^{(k+1)} - x^*\|}{\|x^{(k)} - x^*\|^{1+\rho}} < +\infty. \tag{5.19}$$

Moreover, let $\mathcal{A} = -\mathcal{A}_0 - \sum_{i=1}^n y_i^* \mathcal{A}_i$. If we further assume that the following conditions hold:

(i) *the geometric multiplicity of $\lambda_1(\mathcal{A})$ and $\lambda_1(-\mathcal{X}^*)$ both equals one;*

(ii) *for any eigenvector u associated with $\lambda_1(\mathcal{A})$, we have*

$$\mathcal{A}u^{m-2} - \lambda_1(\mathcal{A})\,\mathrm{diag}(u_1^{m-2}, \ldots, u_n^{m-2}) \prec 0 \text{ on } C_u = \{h \in \mathbb{R}^n : h^T u^{[m-1]} = 0\};$$

(iii) *for any eigenvector v associated with $\lambda_1(-\mathcal{X}^*)$*

$$(-\mathcal{X}^*)v^{m-2} - \lambda_1(-\mathcal{X}^*)\,\mathrm{diag}(v_1^{m-2}, \ldots, v_n^{m-2}) \prec 0 \text{ on } C_v = \{h \in \mathbb{R}^n : h^T v^{[m-1]} = 0\},$$

then Algorithm-1 either terminates in finitely many iterations or generates a sequence $\{x^{(k)}\}$ such that $x^{(k)}$ converges to x^* with order 2 (Q-quadratically), i.e.,

$$\lim_{k\to\infty} \frac{\|x^{(k+1)} - x^*\|}{\|x^{(k)} - x^*\|^{2}} < +\infty.$$

Proof. Clearly, F is locally Lipschitz. Moreover, we note that $(a,b) \mapsto \max\{a,b\}$ is strongly semismooth, any continuous differentiable function with locally Lipschitz gradient is strongly semismooth and composition of (ρth-order) semismooth function is still (ρth-order) semismooth. Then, Theorem 4.2(1) implies that $\mathcal{X} \mapsto \lambda_1(\mathcal{X})$ is a ρth-order semismooth function for some $\rho > 0$. It follows that F is a vector valued function where each of its coordinate is a ρth-order semismooth function. Thus, F is also ρth-order semismooth, and so, Theorem 5.1 (b) gives us that

$$\lim_{k\to\infty} \frac{\|x^{(k+1)} - x^*\|}{\|x^{(k)} - x^*\|^{1+\rho}} < +\infty.$$

Now, let us assume that the conditions (i), (ii) and (iii) hold. Then, Theorem 4.2(2) implies that the maximum eigenvalue function λ_1 is strongly semismooth at $\mathcal{A} := -\mathcal{A}_0 - \sum_{i=1}^n y_i^* \mathcal{A}_i$ and at $-\mathcal{X}^*$. Since the composition function of two strongly semismooth functions is still strongly semismooth, we see that F is a strongly semismooth function. So, the conclusion follows from Theorem 5.1 (b) with $\rho = 1$. $\square$

6 Conclusion and Remarks

In this paper, we examined the maximum eigenvalue function of an even order real symmetric tensor. Using the variational analysis techniques, we showed that the maximum eigenvalue function is continuous and convex. We also showed that, for an mth-order n-dimensional tensor $\mathcal{A}$, the maximum eigenvalue function is always ρth-order semismooth for some $\rho > 0$. Sufficient condition ensuring the strong semismoothness of the maximum eigenvalue function was also provided. As an application, we proposed a generalized Newton method for solving a space tensor conic linear programming problem which arises in medical imaging area, and established the local convergence rate of this method.

Below, we present a few open questions and remarks:

(a) Let $\rho = \frac{1}{(2m-1)^n}$. We achieve the ρth-order semismoothness of λ_1 at a given tensor $\mathcal{A}$ by assuming the geometric multiplicity of the maximum eigenvalue is one at $\mathcal{A}$. Is the ρth-order semismoothness conclusion still true without assuming this condition?

(b) We only establish a local convergence result of our nonsmooth Newton method. How to globalize the nonsmooth Newton method is still an open question. Moreover, it would be interesting to see the numerical performance of our nonsmooth Newton algorithm in solving the space tensor problem.

(c) Existing literature on matrix eigenvalue problem suggests that it is important to study the projection mapping onto the positive semidefinite matrix cone (or equivalently the negative semidefinite matrix cone). What can we say about the projection mapping onto the negative semidefinite tensor cone $\mathcal{C}(4,3)$? Is it semismooth? Can we establish the subdifferential formula for this mapping? These properties are crucial for developing further efficient algorithms for solving the space tensor conic linear problems.

These will be our future research topics and will be examined in a forthcoming paper.

Acknowledgement: The authors would like to express their sincere thanks to Professor Boris Mordukhovich and Professor Chong Li for many valuable suggestions.

References

[1] E. Bierstone and P.D. Milman, Subanalytic geometry. Model theory, algebra, and geometry, 151-172, Math. Sci. Res. Inst. Publ., **39**, Cambridge Univ. Press, Cambridge, 2000.

[2] J. Bolte, A. Daniilidis and A.S. Lewis, The Łojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems, *SIAM J. Optim.* Volume 17, Issue 4, pp. 1205-1223 (2007).

[3] J. Bolte, A. Daniilidis and A.S. Lewis, Tame functions are semismooth, *Mathematical Programming* 117 (2009) 5-19.

[4] J.F. Bonnans and A. Shapiro, Perturbation Analysis of Optimization Problems, Springer, New York, 2000.

[5] K.C. Chang, K. Pearson and T. Zhang. Perron-Frobenius Theorem for nonnegative tensors, *Communications in Mathematical Sciences*, 2008(6), pp. 507-520.

[6] K.C. Chang, K. Pearson and T. Zhang, On eigenvalue problems of real symmetric tensors, *Journal of Mathematical Analysis and Applications*, 350 (2009) 416-422.

[7] D. Cox, J. Little, and D. O'Shea, Using Algebraic Geometry. Springer-Verlag, New York, 1998.

[8] F. H. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983.

[9] F. Facchinei and J. S. Pang, Finite-Dimensional Variational Inequalities and Complementarity Problems Volume II. Springer publication, 2003.

[10] A.V. Fiacco, Second order sufficient conditions for weak and strict constrained minima. *SIAM J. Appl. Math.*, 1968(16), 105-108.

[11] J. Gwoździewicz, The Łojasiewicz exponent of an analytic function at an isolated zero, *Comment. Math. Helv.*, 74 (1999), pp. 364-375.

[12] J.B. Hiriart-Urruty and C. Lemaréchal, Convex analysis and minimization algorithms. I. Fundamentals. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 305. Springer-Verlag, Berlin, 1993. xviii+417 pp.

[13] T. Hoheisel, C. Kanzow B.S. Mordukhovich and H. Phan, Generalized Newton's method for nonsmooth equations based on graphical derivatives, to appear in *Nonl. Anal.*.

[14] E. Kofidis and Ph. Regalia, On the best rank-1 approximation of higher-order symmetric tensors, *SIAM J. Matrix Anal. Appl.* 23 (2002), pp. 863-884.

[15] T.G. Kolda and B.W. Bader, Tensor decompositions and applications. *SIAM Rev.* 51 (2009), no. 3, 455–500.

[16] L. De Lathauwer and B. De Moor, *From matrix to tensor: Multilinear algebra and signal processing*. In: J. McWhirter, Editor, Mathematics in Signal Processing IV, Selected papers presented at 4th IMA Int. Conf. on Mathematics in Signal Processing, Oxford University Press, Oxford, United Kingdom (1998), pp. 115.

[17] L. De Lathauwer, B. De Moor and J. Vandewalle, On the best rank-1 and rank-(R1,R2,,RN) approximation of higher-order tensor, *SIAM J. Matrix Anal. Appl.* 21 (2000), pp. 1324-1342.

[18] F. Facchinei and J.S. Pang, Finite-Dimensional Variational Inequalities and Complementarity Problems Volume I, Springer publication, 2003.

[19] A.S. Lewis, Nonsmooth analysis of eigenvalues, *Mathematical Programming* 84 (1999), 1-24.

[20] G. Li, On the asymptotically well behaved functions and global error bound for convex polynomials, *SIAM J. Optim. 20 (2010)*, No. 4, 1923-1943.

[21] G. Li, Global error bounds for piecewise convex polynomials, *Mathematical Programming*, DOI: 10.1007/s10107-011-0481-z.

[22] G. Li and K.F. Ng, Error bounds of generalized D-gap functions for nonsmooth and nonmonotone variational inequality problems , *SIAM Journal on Optimization*, 20(2009), no. 2, 667-690.

[23] L.H. Lim, Singular values and eigenvalues of tensors, A variational approach, *Proc. 1st IEEE International workshop on computational advances of multi-tensor adaptive processing*, 2005, 129-132.

[24] G. Li, C,.M. Tang and Z.X. Wei, Error bound results for generalized D-gap functions of nonsmooth variational inequality problems, *J. Comp. and Appl. Math.* 233(2010), no. 11, 2795-2806.

[25] Z.Q. Luo, J.S. Pang and D. Ralph, Mathematical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge, 1996.

[26] R. Mifflin, Semismooth and semiconvex functions in constrained optimization. *SIAM J. Control Optimization*, 15 (1977), no. 6, 959–972.

[27] B.S. Mordukhovich, Variational analysis and generalized differentiation. I. Basic theory. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 330. Springer-Verlag, Berlin, 2006. xxii+579 pp.

[28] B.S. Mordukhovich, Variational analysis and generalized differentiation. II. Applications. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 331. Springer-Verlag, Berlin, 2006 ixxii+610 pp.

[29] M. Ng, L. Qi and G. Zhou, Finding the largest eigenvalue of a nonnegative tensor, *SIAM J. Matrix Anal. & Appl.*, 31 (2009) 1090-1099.

[30] Q. Ni, L. Qi and F. Wang, An eigenvalue method for the positive definiteness identification problem, *IEEE Transactions on Automatic Control*, 53 (2008), 1096-1107.

[31] J.S. Pang and L.Q. Qi, Nonsmooth equations: motivation and algorithms, *SIAM J. Optim.*, 3 (1993), no. 3, 443-465.

[32] L. Qi, Eigenvalues of a real symmetric tensor, *J. Symbolic Computation*, 40 (2005) 1302-1324.

[33] L. Qi and J. Sun, A nonsmooth version of Newton's method, *Mathematical Programming*, Vol 78, 1993, pp. 353-368.

[34] L. Qi and K.L. Teo, Multivariate polynomial minimization and its application in signal processing, *Journal of Global Optimization*, 46 (2003) 419-433.

[35] L. Qi and Y. Ye, Space tensor conic programming, preprint.

[36] L. Qi, G. Yu and E.X. Wu, Higher order positive semi-definite diffusion tensor imaging, *SIAM Journal on Imaging Sciences*, 3 (2010) 416-433.

[37] B. Reznick, Some concrete aspects of Hilbert's 17th Problem. *Real algebraic geometry and ordered structures* (Baton Rouge, LA, 1996), 251–272, Contemp. Math., 253, Amer. Math. Soc., Providence, RI, 2000.

[38] D.F. Sun and J. Sun, Strong semismoothness of eigenvalues of symmetric matrices and its application to inverse eigenvalue problems. *SIAM J. Numer. Anal.*, 40 (2002), no. 6, 2352–2367.