Math. Program., Ser. B 101: 95–117 (2004)

Digital Object Identifier (DOI) 10.1007/s10107-004-0538-3

Jein-Shan Chen · Xin Chen · Paul Tseng

Analysis of nonsmooth vector-valued functions associated with second-order cones

Received: February 1, 2003 / Accepted: April 5, 2004
Published online: 21 July 2004 – © Springer-Verlag 2004

Abstract. Let $\mathcal{K}^n$ be the Lorentz/second-order cone in $\mathbb{R}^n$. For any function f from $\mathbb{R}$ to $\mathbb{R}$, one can define a corresponding function $f^{\mathrm{soc}}(x)$ on $\mathbb{R}^n$ by applying f to the spectral values of the spectral decomposition of $x \in \mathbb{R}^n$ with respect to $\mathcal{K}^n$. We show that this vector-valued function inherits from f the properties of continuity, (local) Lipschitz continuity, directional differentiability, Fréchet differentiability, continuous differentiability, as well as (ρ-order) semismoothness. These results are useful for designing and analyzing smoothing methods and nonsmooth methods for solving second-order cone programs and complementarity problems.

Key words. Second-order cone – Vector-valued function – Nonsmooth analysis – Semismooth function – Complementarity

1. Introduction

Let $\mathcal{K}^n$ ($n \geq 1$) be the *second-order cone* (SOC), also called the Lorentz cone, in $\mathbb{R}^n$, defined by

$$\mathcal{K}^n := \{(x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1} \mid \|x_2\| \leq x_1\},$$

where $\|\cdot\|$ denotes the Euclidean norm. By definition, $\mathcal{K}^1$ is the set of nonnegative reals $\mathbb{R}_+$. The second-order cone has recently received much attention in optimization, particularly in the context of applications and solution methods for second-order cone programs (SOCP) [1, 2, 12, 20, 22, 27, 31]. Any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ can be decomposed as

$$x = \lambda_1 u^{(1)} + \lambda_2 u^{(2)}, \tag{1}$$

where λ_1, λ_2 and $u^{(1)}, u^{(2)}$ are the *spectral values* and the associated *spectral vectors* of x, with respect to $\mathcal{K}^n$, given by

$$\lambda_i = x_1 + (-1)^i \|x_2\|, \tag{2}$$

$$u^{(i)} = \begin{cases} \frac{1}{2}\left(1, (-1)^i \frac{x_2}{\|x_2\|}\right), & \text{if } x_2 \neq 0, \\ \frac{1}{2}\left(1, (-1)^i w\right), & \text{if } x_2 = 0, \end{cases} \tag{3}$$

J.-S. Chen, P. Tseng: Department of Mathematics, University of Washington, Seattle, WA 98195, USA. e-mail: {jchen, tseng}@math.washington.edu

X. Chen: Operations Research Center, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Building E40-149, Cambridge, MA 02139, USA. e-mail: xchen@mit.edu

Mathematics Subject Classification (1991): 26A27, 26B05, 26B35, 49J52, 90C33, 65K05

for $i = 1, 2$, with w being any vector in $\mathbb{R}^{n-1}$ satisfying $\|w\| = 1$. If $x_2 \neq 0$, the decomposition (1) is unique. In [12], for any function $f : \mathbb{R} \to \mathbb{R}$, the following vector-valued function associated with $\mathcal{K}^n$ $(n \geq 1)$ was considered:

$$f^{\text{soc}}(x) = f(\lambda_1)u^{(1)} + f(\lambda_2)u^{(2)} \quad \forall x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}. \tag{4}$$

If f is defined only on a subset of $\mathbb{R}$, then f^{soc} is defined on the corresponding subset of $\mathbb{R}^n$. The definition (4) is unambiguous whether $x_2 \neq 0$ or $x_2 = 0$. The cases of $f^{\text{soc}}(x) = x^{1/2},\ x^2,\ \exp(x)$ are discussed in the book of Faraut and Korányi [9]. The above definition (4) is analogous to one associated with the semidefinite cone $\mathcal{S}^n_+$, see [28, 30].

Our study of this function is motivated by optimization and complementarity problems whose constraints involve the direct product of second-order cones. In particular, we wish to find vectors $x, y \in \mathbb{R}^n$ and $\zeta \in \mathbb{R}^\ell$ satisfying

$$\langle x, y \rangle = 0, \quad x \in \mathcal{K}, \quad y \in \mathcal{K}, \quad F(x, y, \zeta) = 0, \tag{5}$$

where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product, $F : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^\ell \to \mathbb{R}^n \times \mathbb{R}^\ell$ is a continuously differentiable mapping, and

$$\mathcal{K} = \mathcal{K}^{n_1} \times \cdots \times \mathcal{K}^{n_m}, \tag{6}$$

with $\ell \geq 0$, $m, n_1, \ldots, n_m \geq 1$ and $n_1 + \cdots + n_m = n$. We will refer to (5), (6) as the *second-order-cone complementarity problem* (SOCCP). This problem has wide applications and, in particular, includes a large class of quadratically constrained problems as special cases [20]. It also includes as a special case the well-known nonlinear complementarity problem (NCP) [10], corresponding to $n_i = 1$ for all i, i.e., $\mathcal{K}$ is the nonnegative orthant $\mathbb{R}^n_+$. When $\ell = 0$ and the mapping F has the form

$$F(x, y, \zeta) = F_0(x) - y \tag{7}$$

for some $F_0 : \mathbb{R}^n \to \mathbb{R}^n$, the SOCCP (5) becomes

$$\langle x, F_0(x) \rangle = 0, \quad x \in \mathcal{K}, \quad F_0(x) \in \mathcal{K}, \tag{8}$$

which is a natural generalization of the ordinary NCP corresponding to $\mathcal{K} = \mathbb{R}^n_+$.

Optimization problems with SOC constraints have been the focus of several recent studies. It is known that $\mathcal{K}^n$, like $\mathbb{R}^n_+$ and the cone $\mathcal{S}^n_+$ of $n \times n$ real symmetric positive semidefinite matrices, belongs to the class of symmetric cones, to which a Jordan algebra may be associated [9]. Using this connection, interior-point methods have been developed for solving linear programs with SOC constraints [20, 22, 31] and, more generally, linear programs with symmetric cone constraints [1, 27]. An alternative approach based on reformulating SOC constraints as smooth convex constraints was studied in [2]. In [12], a non-interior smoothing approach to solving (5) was considered, for which the vector-valued function f^{soc} played a central role. For the special case of $f(\xi) = |\xi|$, $f(\xi) = \max\{0, \xi\}$, further studies of f^{soc} such as strong semismoothness and boundedness of solutions to SOCCP were made in [7, 14]. Formulas for directional derivatives and strong stability of isolated solution to SOCCP were made in [23].

In this paper, we study the continuity and differential properties of the vector-valued function f^{soc} in general. In particular, we show that the properties of continuity, strict continuity, Lipschitz continuity, directional differentiability, differentiability, continuous differentiability, and (ρ-order) semismoothness are each inherited by f^{soc} from f (see Props. 2–7). Here and throughout, differentiability means differentiability in the Fréchet sense. These results parallel those obtained recently in [6] for matrix-valued functions and are useful in the design and analysis of smoothing and nonsmooth methods for solving SOCP and SOCCP. Our ρ-order semismoothness result generalizes a recent result of Chen, Sun and Sun [7], which considers the cases of $f(\xi) = |\xi|$, $f(\xi) = \max\{0, \xi\}$, and shows that $f^{\text{soc}}(x) = (x^2)^{1/2}$, $f^{\text{soc}}(x) = [x]_+$ are strongly semismooth. Our proofs are based on an elegant relation between the vector-valued function f^{soc} and its matrix-valued counterpart (see Lemma 1). This relation enables us to apply the results from [6] for matrix-valued functions to the vector-valued function f^{soc}. Our proofs also use two lemmas from [26] and [28]. The property of semismoothness, as introduced by Mifflin [21] for functionals and scalar-valued functions and further extended by Qi and Sun [25] for vector-valued functions, is of particular interest due to the key role it plays in the superlinear convergence analysis of certain generalized Newton methods [13, 17, 24, 25, 32].

In what follows, $\mathbb{R}^n$ ($n \geq 1$) denotes the space of n-dimensional real column vectors, $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}$ is identified with $\mathbb{R}^{n_1+\cdots+n_m}$. Thus, $(x_1, \ldots, x_m) \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}$ is viewed as a column vector in $\mathbb{R}^{n_1+\cdots+n_m}$. Also, $\mathbb{R}_+$ and $\mathbb{R}_{++}$ denote the nonnegative and positive reals. For any $x, y \in \mathbb{R}^n$, the Euclidean inner product and norm are denoted by $\langle x, y\rangle = x^T y$ and $\|x\| = \sqrt{x^T x}$. For any differentiable (in the Fréchet sense) mapping $F : \mathbb{R}^n \to \mathbb{R}^m$, we denote its Jacobian at $x \in \mathbb{R}^n$ by $\nabla F(x) \in \mathbb{R}^{m\times n}$, i.e., $(F(x + u) - F(x) - \nabla F(x)u)/\|u\| \to 0$ as $u \to 0$. " := " means "define". For any linear mapping $M : \mathbb{R}^n \to \mathbb{R}^n$, we denote its operator norm $|||M||| := \max_{\|x\|=1} \|Mx\|$. For any $x \in \mathbb{R}^n$ and scalar $\gamma > 0$, we denote the γ-ball around x by $\mathcal{B}(x, \gamma) := \{y \in \mathbb{R}^n \mid \|y - x\| \leq \gamma\}$. We write $z = O(\alpha)$ (respectively, $z = o(\alpha)$), with $\alpha \in \mathbb{R}$ and $z \in \mathbb{R}^n$, to mean $\|z\|/|\alpha|$ is uniformly bounded (respectively, tends to zero) as $\alpha \to 0$.

2. Basic properties

In this section, we review some basic properties of vector-valued functions. These properties are continuity, (local) Lipschitz continuity, directional differentiability, differentiability, continuous differentiability, as well as (ρ-order) semismoothness. In what follows, we consider a function/mapping $F : \mathbb{R}^k \to \mathbb{R}^\ell$.

We say F is continuous at $x \in \mathbb{R}^k$ if

$$F(y) \to F(x) \quad \text{as} \quad y \to x;$$

and F is continuous if F is continuous at every $x \in \mathbb{R}^k$. F is strictly continuous (also called 'locally Lipschitz continuous') at $x \in \mathbb{R}^k$ [26, Chap. 9] if there exist scalars $\kappa > 0$ and $\delta > 0$ such that

$$\|F(y) - F(z)\| \leq \kappa\|y - z\| \quad \forall y, z \in \mathbb{R}^k \text{ with } \|y - x\| \leq \delta,\ \|z - x\| \leq \delta;$$

and F is strictly continuous if F is strictly continuous at every $x \in \mathbb{R}^k$. If δ can be taken to be ∞, then F is Lipschitz continuous with Lipschitz constant κ. Define the function $\mathrm{lip}F : \mathbb{R}^k \to [0, \infty]$ by

$$\mathrm{lip}F(x) := \limsup_{\substack{y,z\to x\\ y\neq z}} \frac{\|F(y) - F(z)\|}{\|y - z\|}.$$

Then F is strictly continuous at x if and only if $\mathrm{lip}F(x)$ is finite.

We say F is directionally differentiable at $x \in \mathbb{R}^k$ if

$$F'(x; h) \; := \; \lim_{t\to 0^+} \frac{F(x + th) - F(x)}{t} \quad \text{exists} \qquad \forall h \in \mathbb{R}^k;$$

and F is directionally differentiable if F is directionally differentiable at every $x \in \mathbb{R}^k$. F is differentiable (in the Fréchet sense) at $x \in \mathbb{R}^k$ if there exists a linear mapping $\nabla F(x) : \mathbb{R}^k \to \mathbb{R}^\ell$ such that

$$F(x + h) - F(x) - \nabla F(x)h \; = \; o(\|h\|).$$

We say that F is continuously differentiable if F is differentiable at every $x \in \mathbb{R}^k$ and ∇F is continuous.

If F is strictly continuous, then F is almost everywhere differentiable by Rademacher's Theorem–see [8] and [26, Sec. 9J]. In this case, the generalized Jacobian $\partial F(x)$ of F at x (in the Clarke sense) can be defined as the convex hull of the generalized Jacobian $\partial_B F(x)$, where

$$\partial_B F(x) := \left\{ \lim_{x^j\to x} \nabla F(x^j) \middle| F \text{ is differentiable at } x^j \in \mathbb{R}^k \right\}.$$

The notation ∂_B is adopted from [24]. In [26, Chap. 9], the case of $\ell = 1$ is considered and the notations "$\bar{\nabla}$" and "$\bar{\partial}$" are used instead of, respectively, "∂_B" and "∂".

Assume $F : \mathbb{R}^k \to \mathbb{R}^\ell$ is strictly continuous. We say F is semismooth at x if F is directionally differentiable at x and, for any $V \in \partial F(x + h)$, we have

$$F(x + h) - F(x) - Vh = o(\|h\|).$$

We say F is ρ-order semismooth at x ($0 < \rho < \infty$) if F is semismooth at x and, for any $V \in \partial F(x + h)$, we have

$$F(x + h) - F(x) - Vh = O(\|h\|^{1+\rho}).$$

We say F is semismooth (respectively, ρ-order semismooth) if F is semismooth (respectively, ρ-order semismooth) at every $x \in \mathbb{R}^k$. We say F is strongly semismooth if it is 1-order semismooth. Convex functions and piecewise continuously differentiable functions are examples of semismooth functions. The composition of two (respectively, ρ-order) semismooth functions is also a (respectively, ρ-order) semismooth function. The property of semismoothness plays an important role in nonsmooth Newton methods [24, 25] as well as in some smoothing methods mentioned in the previous section. For extensive discussions of semismooth functions, see [11, 21, 25].

3. Results for matrix-valued functions

Let $\mathbb{R}^{n\times n}$ denote the space of $n \times n$ real matrices, equipped with the trace inner product and the Frobenious norm

$$\langle X, Y\rangle_F := \mathrm{tr}[X^T Y], \qquad \|X\|_F := \sqrt{\langle X, X\rangle},$$

where $X, Y \in \mathbb{R}^{n\times n}$ and $\mathrm{tr}[\cdot]$ denotes the matrix trace, i.e., $\mathrm{tr}[X] = \sum_{i=1}^{n} X_{ii}$. Let $\mathcal{O}$ denote the set of $P \in \mathbb{R}^{n\times n}$ that are orthogonal, i.e., $P^T = P^{-1}$. Let $\mathcal{S}^n$ denote the subspace comprising those $X \in \mathbb{R}^{n\times n}$ that are symmetric, i.e., $X^T = X$. This is a subspace of $\mathbb{R}^{n\times n}$ with dimension $n(n+1)/2$, which can be identified with $\mathbb{R}^{n(n+1)/2}$. Thus, a function mapping $\mathcal{S}^n$ to $\mathcal{S}^n$ may be viewed equivalently as a function mapping $\mathbb{R}^{n(n+1)/2}$ to $\mathbb{R}^{n(n+1)/2}$, for which the properties of Sec. 2 are all applicable. We consider such a function below.

For any $X \in \mathcal{S}^n$, its (repeated) eigenvalues $\lambda_1, \ldots, \lambda_n$ are real and it admits a spectral decomposition of the form:

$$X = P \;\mathrm{diag}\,[\lambda_1, \ldots, \lambda_n] P^T, \tag{9}$$

for some $P \in \mathcal{O}$, where $\mathrm{diag}\,[\lambda_1, \ldots, \lambda_n]$ denotes the $n \times n$ diagonal matrix with its ith diagonal entry λ_i. Then, for any function $f : \mathbb{R} \to \mathbb{R}$, we can define a corresponding function $f^{\mathrm{mat}} : \mathcal{S}^n \to \mathcal{S}^n$ [3], [16] by

$$f^{\mathrm{mat}}(X) := P \;\mathrm{diag}\,[f(\lambda_1), \ldots, f(\lambda_n)] P^T. \tag{10}$$

It is known that $f^{\mathrm{mat}}(X)$ is well defined (independent of the ordering of $\lambda_1, \ldots, \lambda_n$ and the choice of P) and belongs to $\mathcal{S}^n$, see [3, Chap. V] and [16, Sec. 6.2]. Moreover, a result of Daleckii and Krein showed that if f is continuously differentiable, then f^{mat} is differentiable and its Jacobian $\nabla f^{\mathrm{mat}}(X)$ has a simple formula–see [3, Thm. V.3.3]; also see [6, Prop. 4.3]. In [5], f^{mat} was used to develop non-interior continuation methods for solving semidefinite programs and semidefinite complementarity problems. A related method was studied in [18]. Further studies of f^{mat} in the case of $f(\xi) = |\xi|$ and $f(\xi) = \max\{0, \xi\}$ are given in [23, 28], obtaining results such as strong semismoothness, formulas for directional derivatives, and necessary/sufficient conditions for strong stability of an isolated solution to semidefinite complementarity problem (SDCP).

The following key result is from Props. 4.1, 4.2, 4.3, 4.4, 4.6, 4.10 of [6], as well as a remark at the end of Sec. 4 of [6]. It says that f^{mat} inherits from f the property of continuity (respectively, strict continuity, Lipschitz continuity, directional differentiability, differentiability, continuous differentiability, semismoothness, ρ-order semismoothness).

Proposition 1. *For any $f : \mathbb{R} \to \mathbb{R}$, the following results hold:*

(a) f^{mat} is continuous at an $X \in \mathcal{S}^n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ if and only if f is continuous at $\lambda_1, \ldots, \lambda_n$.

(b) f^{mat} is directionally differentiable at an $X \in \mathcal{S}^n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ if and only if f is directionally differentiable at $\lambda_1, \ldots, \lambda_n$

(c) f^{mat} *is differentiable at an* $X \in \mathcal{S}^n$ *with eigenvalues* $\lambda_1, \ldots, \lambda_n$ *if and only if* f *is differentiable at* $\lambda_1, \ldots, \lambda_n$.

(d) f^{mat} *is continuously differentiable at an* $X \in \mathcal{S}^n$ *with eigenvalues* $\lambda_1, \ldots, \lambda_n$ *if and only if* f *is continuously differentiable at* $\lambda_1, \ldots, \lambda_n$.

(e) f^{mat} *is strictly continuous at an* $X \in \mathcal{S}^n$ *with eigenvalues* $\lambda_1, \ldots, \lambda_n$ *if and only if* f *is strictly continuous at* $\lambda_1, \ldots, \lambda_n$.

(f) f^{mat} *is Lipschitz continuous (with respect to* $\|\cdot\|_F$*) with constant* κ *if and only if* f *is Lipschitz continuous with constant* κ.

(g) f^{mat} *is semismooth if and only if* f *is semismooth. If* $f : \mathbb{R} \to \mathbb{R}$ *is* ρ*-order semismooth* ($0 < \rho < \infty$)*, then* f^{mat} *is* $\min\{1, \rho\}$*-order semismooth.*

4. Relating vector-valued function to matrix-valued function

For any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ and $y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, we define their *Jordan product* as

$$x \cdot y = \left(x^T y,\ y_1 x_2 + x_1 y_2\right). \tag{11}$$

We will write x^2 to mean $x \cdot x$ and write $x + y$ to mean the usual componentwise addition of vectors. Then, $\cdot$, $+$, together with

$$e = (1, 0, \ldots, 0) \in \mathbb{R}^n,$$

give rise to a Jordan algebra associated with $\mathcal{K}^n$ [9, Chap. II]. If $x \in \mathcal{K}^n$, then there exists a unique vector in $\mathcal{K}^n$, which we denote by $x^{1/2}$, such that $(x^{1/2})^2 = x^{1/2} \cdot x^{1/2} = x$. For any $x \in \mathbb{R}^n$, we have $x^2 \in \mathcal{K}^n$. Hence there exists a unique vector $(x^2)^{1/2} \in \mathcal{K}^n$, which we denote by $|x|$. Clearly we have $x^2 = |x|^2$. We define $[x]_+$ to be the nearest-point (in the Euclidean norm) projection of x onto $\mathcal{K}^n$. It is shown in [12] that $|x|$ and $[x]_+$ have the form (4), corresponding to $f(\xi) = |\xi|$ and $f(\xi) = \max\{0, \xi\}$. Moreover, they are related to each other by $|x| = (x^2)^{1/2}$, $[x]_+ = (x + |x|)/2$, as in the cases of nonnegative orthant $\mathbb{R}^n_+$ and positive semidefinite cone $\mathcal{S}^n$. Further properties of $|x|$, $[x]_+$ and x^2 are investigated in [12, 14, 23].

For any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, we define the symmetric matrix

$$L_x = \begin{bmatrix} x_1 & x_2^T \\ x_2 & x_1 I \end{bmatrix}, \tag{12}$$

viewed as a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$. The matrix L_x has various interesting properties that were studied in [12]. For our purpose, the following lemma using L_x is key to relating f^{soc} to f^{mat}.

Lemma 1. *For any* $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$*, let* λ_1, λ_2 *be its spectral values given by* (2)*. Let* $z = x_2$ *if* $x_2 \neq 0$*; otherwise let* z *be any nonzero vector in* $\mathbb{R}^{n-1}$*. The following results hold:*

(a) For any $t \in \mathbb{R}$, the matrix $L_x + tM_z$ has eigenvalues of λ_1, λ_2, and $x_1 + t$ of multiplicity $n - 2$, where

$$M_z := \begin{bmatrix} 0 & 0 \\ 0 & I - zz^T/\|z\|^2 \end{bmatrix}. \tag{13}$$

(b) For any $f : \mathbb{R} \to \mathbb{R}$ and any $t \in \mathbb{R}$, we have

$$f^{\text{soc}}(x) = f^{\text{mat}}(L_x + tM_z)e.$$

Proof. It is straightforward to verify that, for any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, the eigenvalues of L_x are λ_1, λ_2, as given by (2), and x_1 of multiplicity $n-2$. The corresponding orthonormal set of eigenvectors is $\sqrt{2}u^{(1)}, \sqrt{2}u^{(2)}$ and $u^{(i)} = (0, u_2^{(i)})$, $i = 3, ..., n$, where $u^{(1)}, u^{(2)}$ are given by (3) with $w = z/\|z\|$ whenever $x_2 = 0$, and $u_2^{(3)}, ..., u_2^{(n)}$ is any orthonormal set of vectors that span the subspace of $\mathbb{R}^{n-1}$ orthogonal to z. Thus, $L_x = U\text{diag}[\lambda_1, \lambda_2, x_1, ..., x_1]U^T$, where $U := \left[\sqrt{2}u^{(1)}\ \sqrt{2}u^{(2)}\ u^{(3)} \cdots u^{(n)}\right]$. Also, it is straightforward to verify using $u^{(i)} = (0, u_2^{(i)})$, $i = 3, ..., n$, that

$$U\text{diag}[0, 0, 1, ..., 1]U^T = \begin{bmatrix} 0 & 0 \\ 0 & \sum_{i=3}^n u_2^{(i)}(u_2^{(i)})^T \end{bmatrix}.$$

Since $Q := [z/\|z\|\ u_2^{(3)} \cdots u_2^{(n)}]$ is an orthogonal matrix, we have $I = QQ^T = zz^T/\|z\|^2 + \sum_{i=3}^n u_2^{(i)}(u_2^{(i)})^T$ and hence $\sum_{i=3}^n u_2^{(i)}(u_2^{(i)})^T = I - zz^T/\|z\|^2$. This together with (13) shows that $U\text{diag}[0, 0, 1, ..., 1]U^T = M_z$. Thus,

$$L_x + tM_z = U\text{diag}[\lambda_1, \lambda_2, x_1 + t, ..., x_1 + t]U^T. \tag{14}$$

This proves (a).

(b) Using (14), we have

$$\begin{aligned} f^{\text{mat}}(L_x + tM_z)e &= U\text{diag}[f(\lambda_1), f(\lambda_2), f(x_1 + t), ..., f(x_1 + t)]U^T e \\ &= f(\lambda_1)u^{(1)} + f(\lambda_2)u^{(2)} = f^{\text{soc}}(x), \end{aligned}$$

where the second equality uses the special form of U and (3). This proves (b). □

Of particular interest is the choice of $t = \pm\|x_2\|$, for which $L_x + tM_{x_2}$ has eigenvalues of λ_1, λ_2. More generally, for any $f, g : \mathbb{R} \to \mathbb{R}_+$, any $h : \mathbb{R}_+ \to \mathbb{R}$ and any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, we have

$$h^{\text{soc}}(f^{\text{soc}}(x) + g(\mu)e) = h^{\text{mat}}(f^{\text{mat}}(L_x) + g(\mu)I)e.$$

In particular, the spectral values of $f^{\text{soc}}(x)$ and $g(\mu)e$ are nonnegative, as are the eigenvalues of $f^{\text{mat}}(L_x)$ and $g(\mu)I$, so both sides are well defined. In particular, for

$$f(\xi) = \xi^2, \quad g(\mu) = \mu^2, \quad h(\xi) = \xi^{1/2},$$

we obtain that

$$(x^2 + \mu^2 e)^{1/2} = (L_x^2 + \mu^2 I)^{1/2} e.$$

It was shown in [29] that $(X, \mu) \mapsto (X^2 + \mu^2 I)^{1/2}$ is strongly semismooth. Then, it follows from the above equation that $(x, \mu) \mapsto (x^2 + \mu^2 e)^{1/2}$ is strongly semismooth. This provides a shorter proof of Theorem 4.2 in [7].

5. Continuity and differential properties of vector-valued function

In this section, we use the results from Secs. 3, 4 to show that if $f : \mathbb{R} \to \mathbb{R}$ has the property of continuity (respectively, strict continuity, Lipschitz continuity, directional differentiability, differentiability, continuous differentiability, semismoothness, ρ-order semismoothness), then so does the vector-valued function f^{soc} defined by (1)–(4).

We begin with the following continuity result for f^{soc}, based on Lemma 1 and Prop. 1(a) on continuity properties of f^{mat}.

Proposition 2. *For any $f : \mathbb{R} \to \mathbb{R}$, the following results hold:*

(a) f^{soc} is continuous at an $x \in \mathcal{S}$ with spectral values λ_1, λ_2 if and only if f is continuous at λ_1, λ_2.
(b) f^{soc} is continuous if and only if f is continuous.

Proof. (a) Suppose f is continuous at λ_1, λ_2. If $x_2 = 0$, then $x_1 = \lambda_1 = \lambda_2$ and, by Lemma 1(a), L_x has eigenvalue of $\lambda_1 = \lambda_2$ of multiplicity n. Then, by Prop. 1(a), f^{mat} is continuous at L_x. Since L_x is continuous in x, Lemma 1(b) yields that $f^{\text{soc}}(x) = f^{\text{mat}}(L_x)e$ is continuous at x. If $x_2 \neq 0$, then, by Lemma 1(a), $L_x + \|x_2\| M_{x_2}$ has eigenvalue of λ_1 of multiplicity 1 and λ_2 of multiplicity $n-1$. Then, by Prop. 1(a), f^{mat} is continuous at $L_x + \|x_2\| M_{x_2}$. Since $x \mapsto L_x + \|x_2\| M_{x_2}$ is continuous at x, Lemma 1(b) yields that $x \mapsto f^{\text{soc}}(x) = f^{\text{mat}}(L_x + \|x_2\| M_{x_2})e$ is continuous at x.

Suppose f^{soc} is continuous at x with spectral values λ_1, λ_2 and spectral vectors $u^{(1)}, u^{(2)}$. For any $\mu_1 \in \mathbb{R}$, let

$$y := \mu_1 u^{(1)} + \lambda_2 u^{(2)}.$$

Then $y \to x$ as $\mu_1 \to \lambda_1$. Since f^{soc} is continuous at x, we have

$$f(\mu_1)u^{(1)} + f(\lambda_2)u^{(2)} = f^{\text{soc}}(y) \to f^{\text{soc}}(x) = f(\lambda_1)u^{(1)} + f(\lambda_2)u^{(2)}.$$

Since $u^{(1)} \neq 0$, this implies $f(\mu_1) \to f(\lambda_1)$ as $\mu_1 \to \lambda_1$. Thus f is continuous at λ_1. A similar argument shows that f is continuous at λ_2.

(b) is an immediate consequence of (a). □

The "if" direction of Prop. 2(a) can alternatively be proved using the Lipschitzian property of the spectral values (Lemma 2) and an upper Lipschitzian property of the spectral vectors. However, this alternative proof is more complicated. If f has a power series expansion, then so does f^{soc}, with the same coefficients of expansion; see [12, Prop. 3.1].

By using Lemma 1 and Prop. 1(b), we have the following directional differentiability result for f^{soc}, together with a computable formula for the directional derivative of f^{soc}. In the special case of $f(\cdot) = \max\{0, \cdot\}$, for which $f^{\text{soc}}(x)$ corresponds to the projection of x onto $\mathcal{K}^n$, an alternative formula expressing the directional derivative as the unique solution to a certain convex program is given in [23, Prop. 13].

Proposition 3. *For any $f : \mathbb{R} \to \mathbb{R}$, the following results hold:*

(a) f^{soc} is directionally differentiable at an $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ with spectral values λ_1, λ_2 if and only if f is directionally differentiable at λ_1, λ_2; Moreover, for any nonzero $h = (h_1, h_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, we have

$$(f^{\text{soc}})'(x; h) = f'(x_1; h_1)e$$

if $x_2 = 0$ and $h_2 = 0$;

$$(f^{\text{soc}})'(x; h) = \frac{1}{2} f'(x_1; h_1 - \|h_2\|) \left(1, \frac{-h_2}{\|h_2\|}\right) + \frac{1}{2} f'(x_1; h_1 + \|h_2\|) \left(1, \frac{h_2}{\|h_2\|}\right) \tag{15}$$

if $x_2 = 0$ and $h_2 \neq 0$; otherwise

$$\begin{aligned}(f^{\text{soc}})'(x; h) &= \frac{1}{2} f'\left(\lambda_1; h_1 - \frac{x_2^T h_2}{\|x_2\|}\right)\left(1, \frac{-x_2}{\|x_2\|}\right) - \frac{f(\lambda_1)}{2\|x_2\|} M_{x_2} h \\ &\quad + \frac{1}{2} f'\left(\lambda_2; h_1 + \frac{x_2^T h_2}{\|x_2\|}\right)\left(1, \frac{x_2}{\|x_2\|}\right) + \frac{f(\lambda_2)}{2\|x_2\|} M_{x_2} h. \end{aligned} \tag{16}$$

(b) f^{soc} is directionally differentiable if and only if f is directionally differentiable.

Proof. (a) Suppose f is directionally differentiable at λ_1, λ_2. If $x_2 = 0$, then $x_1 = \lambda_1 = \lambda_2$ and, by Lemma 1(a), L_x has eigenvalue of x_1 of multiplicity n. Then, by Prop. 1(b), f^{mat} is directionally differentiable at L_x. Since L_x is differentiable in x, Lemma 1(b) yields that $f^{\text{soc}}(x) = f^{\text{mat}}(L_x)e$ is directionally differentiable at x. If $x_2 \neq 0$, then, by Lemma 1(a), $L_x + \|x_2\| M_{x_2}$ has eigenvalue of λ_1 of multiplicity 1 and λ_2 of multiplicity $n - 1$. Then, by Prop. 1(b), f^{mat} is directionally differentiable at $L_x + \|x_2\| M_{x_2}$. Since $x \mapsto L_x + \|x_2\| M_{x_2}$ is differentiable at x, Lemma 1(b) yields that $x \mapsto f^{\text{soc}}(x) = f^{\text{mat}}(L_x + \|x_2\| M_{x_2})e$ is directionally differentiable at x.

Fix any nonzero $h = (h_1, h_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$. Below we calculate $(f^{\text{soc}})'(x; h)$. Suppose $x_2 = 0$. Then $\lambda_1 = \lambda_2 = x_1$ and the spectral vectors $u^{(1)}, u^{(2)}$ sum to $e = (1, 0)$. If $h_2 = 0$, then for any $t > 0$, $x + th$ has the spectral values $\mu_1 = \mu_2 = x_1 + th_1$ and its spectral vectors $v^{(1)}, v^{(2)}$ sum to $e = (1, 0)$. Thus

$$\begin{aligned}\frac{f^{\text{soc}}(x + th) - f^{\text{soc}}(x)}{t} &= \frac{1}{t}\left(f(\mu_1)v^{(1)} + f(\mu_2)v^{(2)} - f(\lambda_1)u^{(1)} - f(\lambda_2)u^{(2)}\right) \\ &= \frac{f(x_1 + th_1) - f(x_1)}{t} e \\ &\to f'(x_1; h_1)e \quad \text{as } t \to 0^+. \end{aligned}$$

If $h_2 \neq 0$, then for any $t > 0$, $x + th$ has the spectral values $\mu_i = (x_1 + th_1) + (-1)^i t\|h_2\|$ and spectral vectors $v^{(i)} = \frac{1}{2}(1, (-1)^i h_2/\|h_2\|)$, $i = 1, 2$. Moreover, since $x_2 = 0$, we can choose $u^{(i)} = v^{(i)}$ for $i = 1, 2$. Thus

$$\begin{aligned}&\frac{f^{\mathrm{soc}}(x+th)-f^{\mathrm{soc}}(x)}{t}\\&=\frac{1}{t}\left(f(\mu_1)v^{(1)}+f(\mu_2)v^{(2)}-f(\lambda_1)v^{(1)}-f(\lambda_2)v^{(2)}\right)\\&=\frac{f(x_1+t(h_1-\|h_2\|))-f(x_1)}{t}v^{(1)}+\frac{f(x_1+t(h_1+\|h_2\|))-f(x_1)}{t}v^{(2)}\\&\to f'(x_1;h_1-\|h_2\|)v^{(1)}+f'(x_1;h_1+\|h_2\|)v^{(2)}\quad\text{as } t\to 0^+.\end{aligned}$$

This together with $v^{(i)}=\frac{1}{2}(1,(-1)^i h_2/\|h_2\|)$, $i=1,2$, yields (15). Suppose $x_2\neq 0$. Then $\lambda_i=x_1+(-1)^i\|x_2\|$ and the spectral vectors are $u^{(i)}=\frac{1}{2}(1,(-1)^i x_2/\|x_2\|)$, $i=1,2$. For any $t>0$ sufficiently small so that $x_2+th_2\neq 0$, $x+th$ has the spectral values $\mu_i=x_1+th_1+(-1)^i\|x_2+th_2\|$ and spectral vectors $v^{(i)}=\frac{1}{2}(1,(-1)^i(x_2+th_2)/\|x_2+th_2\|)$, $i=1,2$. Thus

$$\begin{aligned}&\frac{f^{\mathrm{soc}}(x+th)-f^{\mathrm{soc}}(x)}{t}\\&=\frac{1}{t}\left(f(\mu_1)v^{(1)}+f(\mu_2)v^{(2)}-f(\lambda_1)u^{(1)}-f(\lambda_2)u^{(2)}\right)\\&=\frac{1}{t}\left(\frac{1}{2}f(x_1+th_1-\|x_2+th_2\|)(1,-\frac{x_2+th_2}{\|x_2+th_2\|})-\frac{1}{2}f(\lambda_1)(1,-\frac{x_2}{\|x_2\|})\right.\\&\quad\left.+\frac{1}{2}f(x_1+th_1+\|x_2+th_2\|)(1,\frac{x_2+th_2}{\|x_2+th_2\|})-\frac{1}{2}f(\lambda_2)(1,\frac{x_2}{\|x_2\|})\right).\end{aligned}\tag{17}$$

We now focus on the individual terms in (17). Since

$$\begin{aligned}\frac{\|x_2+th_2\|-\|x_2\|}{t}&=\frac{\|x_2+th_2\|^2-\|x_2\|^2}{(\|x_2+th_2\|+\|x_2\|)t}\\&=\frac{2x_2^Th_2+t\|h_2\|^2}{\|x_2+th_2\|+\|x_2\|}\to\frac{x_2^Th_2}{\|x_2\|}\quad\text{as } t\to 0^+,\end{aligned}$$

we have

$$\begin{aligned}&\frac{1}{t}\left(f(x_1+th_1-\|x_2+th_2\|)-f(\lambda_1)\right)\\&=\frac{1}{t}\left(f\left(\lambda_1+t\left(h_1-\frac{\|x_2+th_2\|-\|x_2\|}{t}\right)\right)-f(\lambda_1)\right)\\&\to f'\left(\lambda_1;h_1-\frac{x_2^Th_2}{\|x_2\|}\right)\quad\text{as } t\to 0^+.\end{aligned}$$

Similarly, we find that

$$\begin{aligned}&\frac{1}{t}\left(f(x_1+th_1+\|x_2+th_2\|)-f(\lambda_2)\right)\\&\to f'\left(\lambda_2;h_1+\frac{x_2^Th_2}{\|x_2\|}\right)\quad\text{as } t\to 0^+.\end{aligned}$$

Also, letting $\Phi(x_2) = x_2/\|x_2\|$, we have that

$$\frac{1}{t}\left(\frac{x_2 + th_2}{\|x_2 + th_2\|} - \frac{x_2}{\|x_2\|}\right) = \frac{\Phi(x_2 + th_2) - \Phi(x_2)}{t} \to \nabla\Phi(x_2)h_2 \quad \text{as } t \to 0^+.$$

Combining the above relations with (17) and using a product rule, we obtain that

$$\begin{aligned}
&\lim_{t\to 0^+} \frac{f^{\text{soc}}(x+th) - f^{\text{soc}}(x)}{t} \\
&= \frac{1}{2}\left(f'\left(\lambda_1; h_1 - \frac{x_2^T h_2}{\|x_2\|}\right)\left(1, \frac{-x_2}{\|x_2\|}\right) - f(\lambda_1)(0, \nabla\Phi(x_2)h_2)\right) \\
&\quad + \frac{1}{2}\left(f'\left(\lambda_2; h_1 + \frac{x_2^T h_2}{\|x_2\|}\right)\left(1, \frac{x_2}{\|x_2\|}\right) + f(\lambda_2)(0, \nabla\Phi(x_2)h_2)\right)
\end{aligned}$$

Using $\nabla\Phi(x_2)h_2 = \frac{1}{\|x_2\|}\left(I - \frac{x_2 x_2^T}{\|x_2\|^2}\right)h_2$ so that $(0, \nabla\Phi(x_2)h_2) = \frac{1}{\|x_2\|}M_{x_2}h$ yields (16).

Suppose f^{soc} is directionally differentiable at x with spectral eigenvalues λ_1, λ_2 and spectral vectors $u^{(1)}, u^{(2)}$. For any direction $d_1 \in \mathbb{R}$, let

$$h := d_1 u^{(1)}.$$

Since $x = \lambda_1 u^{(1)} + \lambda_2 u^{(2)}$, this implies $x + th = (\lambda_1 + td_1)u^{(1)} + \lambda_2 u^{(2)}$, so that

$$\frac{f^{\text{soc}}(x+th) - f^{\text{soc}}(x)}{t} = \frac{f(\lambda_1 + td_1) - f(\lambda_1)}{t}u^{(1)}.$$

Since f^{soc} is directionally differentiable at x, the above difference quotient has a limit as $t \to 0^+$. Since $u^{(1)} \neq 0$, this implies that

$$\lim_{t\to 0^+} \frac{f(\lambda_1 + td_1) - f(\lambda_1)}{t} \quad \text{exists.}$$

Hence f is directionally differentiable at λ_1. A similar argument shows f is directionally differentiable at λ_2.

(b) is an immediate consequence of (a). □

Proposition 4. *For any $f : \mathbb{R} \to \mathbb{R}$, the following results hold:*

(a) f^{soc} is differentiable at an $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ with spectral values λ_1, λ_2 if and only if f is differentiable at λ_1, λ_2. Moreover,

$$\nabla f^{\text{soc}}(x) = f'(x_1)I \tag{18}$$

if $x_2 = 0$, and otherwise

$$\nabla f^{\text{soc}}(x) = \begin{bmatrix} b & c\, x_2^T/\|x_2\| \\ c\, x_2/\|x_2\| & aI + (b-a)x_2x_2^T/\|x_2\|^2 \end{bmatrix}, \tag{19}$$

where

$$a = \frac{f(\lambda_2) - f(\lambda_1)}{\lambda_2 - \lambda_1}, \quad b = \frac{1}{2}\left(f'(\lambda_2) + f'(\lambda_1)\right), \quad c = \frac{1}{2}\left(f'(\lambda_2) - f'(\lambda_1)\right). \tag{20}$$

(b) f^{soc} *is differentiable if and only if* f *is differentiable.*

Proof. (a) The proof of the "if" direction is identical to the proof of Prop. 3, but with "directionally differentiable" replaced by "differentiable" and with Prop. 1(b) replaced by Prop. 1(c). The formula for $\nabla f^{\text{soc}}(x)$ is from [12, Prop. 5.2].

To prove the "only if" direction, suppose f is differentiable at λ_1, λ_2. Then, for each $i \in \{1, 2\}$,

$$\frac{f^{\text{soc}}(x + tu^{(i)}) - f^{\text{soc}}(x)}{t} = \frac{f(\lambda_i + t) - f(\lambda_i)}{t} u^{(i)}$$

has a limit as $t \to 0$. Since $u^{(i)} \neq 0$, this implies that

$$\lim_{t \to 0} \frac{f(\lambda_i + t) - f(\lambda_i)}{t} \text{ exists.}$$

Hence f is differentiable at λ_i.

(b) is an immediate consequence of (a). □

We next have the following continuous differentiability result for f^{soc} based on Prop. 1(d) and Lemma 1.

Proposition 5. *For any* $f : \mathbb{R} \to \mathbb{R}$*, the following results hold:*

(a) f^{soc} *is continuously differentiable at an* $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ *with spectral values* λ_1, λ_2 *if and only if* f *is continuously differentiable at* λ_1, λ_2.

(b) f^{soc} *is continuously differentiable if and only if* f *is continuously differentiable.*

Proof. (a) The proof of the "if" direction is identical to the proof of Prop. 2, but with "continuous" replaced by "continuously differentiable" and with Prop. 1(a) replaced by Prop. 1(d). Alternatively, we note that (19) is continuous at any x with $x_2 \neq 0$. The case of $x_2 = 0$ can be checked by taking $y = (y_1, y_2) \to x$ and considering the two cases: $y_2 = 0$ or $y_2 \neq 0$.

Conversely, suppose f^{soc} is continuously differentiable at an $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ with spectral values λ_1, λ_2. Then, by Prop. 4, f is differentiable in neighborhoods around λ_1, λ_2. If $x_2 = 0$, then $\lambda_1 = \lambda_2 = x_1$ and (18) yields $\nabla f^{\text{soc}}(x) = f'(x_1)I$. For any $h_1 \in \mathbb{R}$, let $h := (h_1, 0)$. Then $\nabla f^{\text{soc}}(x+h) = f'(x_1+h_1)I$. Since ∇f^{soc} is continuous at x, then $\lim_{h_1 \to 0} f'(x_1+h_1)I = f'(x_1)I$, implying $\lim_{h_1 \to 0} f'(x_1+h_1) = f'(x_1)$. Thus, f' is continuous at x_1. If $x_2 \neq 0$, then $\nabla f^{\text{soc}}(x)$ is given by (19) with a, b, c given by (20). For any $h_1 \in \mathbb{R}$, let $h := (h_1, 0)$. Then $x + h = (x_1 + h_1, x_2)$ has spectral values $\mu_1 := \lambda_1 + h_1$, $\mu_2 := \lambda_2 + h_1$. By (19),

$$\nabla f^{\text{soc}}(x + h) = \begin{bmatrix} \beta & \chi\, x_2^T/\|x_2\| \\ \chi\, x_2/\|x_2\| & \alpha I + (\beta - \alpha)x_2 x_2^T/\|x_2\|^2 \end{bmatrix},$$

where

$$\alpha = \frac{f(\mu_2) - f(\mu_1)}{\mu_2 - \mu_1}, \quad \beta = \frac{1}{2}\left(f'(\mu_2) + f'(\mu_1)\right), \quad \chi = \frac{1}{2}\left(f'(\mu_2) - f'(\mu_1)\right).$$

Since ∇f^{soc} is continuous at x so that $\lim_{h\to 0} \nabla f^{\text{soc}}(x+h) = \nabla f^{\text{soc}}(x)$ and $x_2 \neq 0$, we see from comparing terms that $\beta \to b$ and $\chi \to c$ as $h \to 0$. This means that

$$\begin{aligned} f'(\mu_2) + f'(\mu_1) &\to f'(\lambda_2) + f'(\lambda_1) \quad \text{and} \\ f'(\mu_2) - f'(\mu_1) &\to f'(\lambda_2) - f'(\lambda_1) \ \text{ as } h_1 \to 0. \end{aligned}$$

Adding and subtracting the above two limits and we obtain

$$f'(\mu_1) \to f'(\lambda_1) \quad \text{and} \quad f'(\mu_2) \to f'(\lambda_2) \ \text{ as } h_1 \to 0.$$

Since $\mu_1 = \lambda_1 + h_1$, $\mu_2 = \lambda_2 + h_1$, this shows that f' is continuous at λ_1, λ_2.

(b) is an immediate consequence of (a). □

In the case where $f = g'$ for some differentiable g, Prop. 1(d) is a special case of Thm. 4.2 in [19]. This raises the question of whether an SOC analog of the second derivative results in [19] holds.

We now study the strict continuity and Lipschitz continuity properties of f^{soc}. The proof is similar to that of [6, Prop. 4.6], but with a different estimation of $\nabla (f^\nu)^{\text{soc}}$. We begin with the following lemma, which is analogous to a result of Weyl for eigenvalues of symmetric matrices, e.g., [3, p. 63], [15, p. 367].

Lemma 2. *For any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ with spectral values λ_1, λ_2 and any $y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ with spectral values μ_1, μ_2, we have*

$$|\lambda_i - \mu_i| \le \sqrt{2}\|x - y\|, \quad i = 1, 2.$$

Proof. We have

$$\begin{aligned} |\lambda_1 - \mu_1| &= |x_1 - \|x_2\| - y_1 + \|y_2\|| \\ &\le |x_1 - y_1| + |\|x_2\| - \|y_2\|| \\ &\le |x_1 - y_1| + \|x_2 - y_2\| \\ &\le \sqrt{2}(|x_1 - y_1|^2 + \|x_2 - y_2\|^2)^{1/2} \\ &= \sqrt{2}\|x - y\|, \end{aligned}$$

where the second inequality uses $\|x_2\| \le \|x_2 - y_2\| + \|y_2\|$ and $\|y_2\| \le \|x_2 - y_2\| + \|x_2\|$; the last inequality uses the relation between the 1-norm and the 2-norm. A similar argument applies to $|\lambda_2 - \mu_2|$. □

We also need the following result of Rockafellar and Wets [26, Thm. 9.67].

Lemma 3. *Suppose $f : \mathbb{R}^k \to \mathbb{R}$ is strictly continuous. Then there exist continuously differentiable functions $f^\nu : \mathbb{R}^k \to \mathbb{R}$, $\nu = 1, 2, \ldots$, converging uniformly to f on any compact set C in $\mathbb{R}^k$ and satisfying*

$$\nabla f^\nu(x) \le \sup_{y \in C} \operatorname{lip} f(y) \qquad \forall x \in C, \ \forall \nu.$$

Lemma 3 is slightly different from the original version given in [26, Thm. 9.67]. In particular, the second part of Lemma 3 is not contained in [26, Thm. 9.67], but is implicit in its proof. This second part is needed to show that strict continuity and Lipschitz continuity are inherited by f^{soc} from f. We note that Prop, 1(e),(f) and Lemma 1 can be used to give a short proof of strict continuity and Lipschitz continuity of f^{soc}, but the Lipschitz constant would not be sharp. In particular, the constant would be off by a multiplicative factor of $\sqrt{n}$ due to $\|L_x\|_F \le \sqrt{n}\|x\|$ for all $x \in \mathbb{R}^n$. Also, spectral vectors do not behave in a (locally) Lipschitzian manner, so we cannot use (4) directly.

Proposition 6. *For any $f : \mathbb{R} \to \mathbb{R}$, the following results hold:*

(a) f^{soc} is strictly continuous at an $x \in \mathbb{R}^n$ with spectral values $\lambda_1, \ldots, \lambda_n$ if and only if f is strictly continuous at $\lambda_1, \ldots, \lambda_n$.
(b) f^{soc} is strictly continuous if and only if f is strictly continuous.
(c) f^{soc} is Lipschitz continuous (with respect to $\|\cdot\|$) with constant κ if and only if f is Lipschitz continuous with constant κ.

Proof. (a) Fix any $x \in \mathbb{R}^n$ with spectral values λ_1, λ_2 given by (2).

"if" Suppose f is strictly continuous at λ_1, λ_2. Then there exist $\kappa_i > 0$ and $\delta_i > 0$ for $i = 1, 2$, such that

$$|f(\xi) - f(\zeta)| \le \kappa_i|\xi - \zeta| \quad \forall \xi, \zeta \in [\lambda_i - \delta_i, \lambda_i + \delta_i] \quad i = 1, 2.$$

Let $\bar{\delta} := \min\{\delta_1, \delta_2\}$ and

$$C := [\lambda_1 - \bar{\delta}, \lambda_1 + \bar{\delta}] \,\cup\, [\lambda_2 - \bar{\delta}, \lambda_2 + \bar{\delta}] \;.$$

We define $\tilde{f} : \mathbb{R} \to \mathbb{R}$ to be the function that coincides with f on C; and is linearly extrapolated at the boundary points of C on $\mathbb{R} \setminus C$. In other words,

$$\tilde{f}(\xi) = \begin{cases} f(\xi) & \text{if } \xi \in C, \\ (1-t)f(\lambda_1 + \bar{\delta}) & \text{if } \lambda_1 + \bar{\delta} < \lambda_2 - \bar{\delta} \text{ and, for some } t \in (0,1), \\ \quad +tf(\lambda_2 - \bar{\delta}) & \xi = (1-t)(\lambda_1 + \bar{\delta}) + t(\lambda_2 - \bar{\delta}), \\ f(\lambda_1 - \bar{\delta}) & \text{if } \xi < \lambda_1 - \bar{\delta}, \\ f(\lambda_2 + \bar{\delta}) & \text{if } \xi > \lambda_2 + \bar{\delta}. \end{cases}$$

From the above, we see that $\tilde{f}$ is Lipschitz continuous, so that there exists a scalar $\kappa > 0$ such that $\operatorname{lip} \tilde{f}(\xi) \le \kappa$ for all $\xi \in \mathbb{R}$. Since C is compact, by Lemma 3, there exist continuously differentiable functions $f^\nu : \mathbb{R} \to \mathbb{R}$, $\nu = 1, 2, \ldots,$ converging uniformly to $\tilde{f}$ and satisfying

$$|(f^\nu)'(\xi)| \le \kappa \quad \forall \xi \in C, \;\forall \nu \;. \tag{21}$$

Let $\delta := \frac{1}{\sqrt{2}}\bar{\delta}$, so by Lemma 2, C contains all spectral values of $y \in \mathcal{B}(x, \delta)$. Moreover, for any $w \in \mathcal{B}(x, \delta)$ with spectral factorization

$$w = \mu_1 u^{(1)} + \mu_2 u^{(2)} \;,$$

we have $\mu_1, \mu_2 \in C$ and

$$\|(f^\nu)^{\text{soc}}(w) - f^{\text{soc}}(w)\|^2 = \|(f^\nu(\mu_1) - f(\mu_1))u^{(1)} + (f^\nu(\mu_2) - f(\mu_2))u^{(2)}\|^2$$
$$= \frac{1}{2}|f^\nu(\mu_1) - f(\mu_1)|^2 + \frac{1}{2}|f^\nu(\mu_2) - f(\mu_2)|^2 \; , \qquad (22)$$

where we use $\|u^{(i)}\|^2 = 1/2$ for $i = 1, 2$, and $(u^{(1)})^T u^{(2)} = 0$. Since $\{f^\nu\}_{\nu=1}^\infty$ converges uniformly to f on C, equation (22) shows that $\{(f^\nu)^{\text{soc}}\}_{\nu=1}^\infty$ converges uniformly to f^{soc} on $\mathcal{B}(x, \delta)$. Moreover, for all $w = (w_1, w_2) \in \mathcal{B}(x, \delta)$ and all ν, we have from Prop. 4 that $\nabla(f^\nu)^{\text{soc}}(w) = (f^\nu)'(w_1)I$ if $w_2 = 0$, in which case $\nabla(f^\nu)^{\text{soc}}(w) = |(f^\nu)'(w_1)| \le \kappa$. Otherwise $w_2 \ne 0$ and

$$\nabla(f^\nu)^{\text{soc}}(w) = \begin{bmatrix} b & c\, w_2^T/\|w_2\| \\ c\, w_2/\|w_2\| & aI + (b-a)w_2 w_2^T/\|w_2\|^2 \end{bmatrix},$$

where a, b, c are given by (20) but with λ_1, λ_2 replaced by μ_1, μ_2, respectively. If $c = 0$, the above matrix has the form $bI + (a-b)M_{w_2}$. Since M_{w_2} has eigenvalues of 0 and 1, this matrix has eigenvalues of b and a. Thus, $\nabla(f^\nu)^{\text{soc}}(w) = \max\{|a|, |b|\} \le \kappa$. If $c \ne 0$, the above matrix has the form $\frac{c}{\|w_2\|}L_z + (a-b)M_{w_2} = \frac{c}{\|w_2\|}\big(L_z + (a-b)\|w_2\|\, c^{-1}M_{w_2}\big)$, where $z = (b\|w_2\|/c, w_2)$. By Lemma 1, this matrix has eigenvalues of $b \pm c$ and a. Thus, $\nabla(f^\nu)^{\text{soc}}(w) = \max\{|b+c|, |b-c|, |a|\} \le \kappa$. In all cases, we have

$$|||\nabla(f^\nu)^{\text{soc}}(w)||| \le \kappa. \qquad (23)$$

Fix any $y, z \in \mathcal{B}(x, \delta)$ with $y \ne z$. Since $\{(f^\nu)^{\text{soc}}\}_{\nu=1}^\infty$ converges uniformly to f^{soc} on $\mathcal{B}(x, \delta)$, for any $\epsilon > 0$ there exists an integer ν_0 such that for all $\nu \ge \nu_0$ we have

$$\|(f^\nu)^{\text{soc}}(w) - f^{\text{soc}}(w)\| \le \epsilon\|y - z\| \quad \forall w \in \mathcal{B}(x, \delta) \; .$$

Since f^ν is continuously differentiable, then Prop. 5 shows that $(f^\nu)^{\text{soc}}$ is also continuously differentiable for all ν. Thus, by inequality (23) and the mean value theorem for continuously differentiable functions, we have

$$\begin{aligned}
&\|f^{\text{soc}}(y) - f^{\text{soc}}(z)\| \\
&\quad = \|f^{\text{soc}}(y) - (f^\nu)^{\text{soc}}(y) + (f^\nu)^{\text{soc}}(y) - (f^\nu)^{\text{soc}}(z) + (f^\nu)^{\text{soc}}(z) - f^{\text{soc}}(z)\| \\
&\quad \le \|f^{\text{soc}}(y) - (f^\nu)^{\text{soc}}(y)\| + \|(f^\nu)^{\text{soc}}(y) - (f^\nu)^{\text{soc}}(z)\| + \|(f^\nu)^{\text{soc}}(z) - f^{\text{soc}}(z)\| \\
&\quad \le 2\epsilon\|y - z\| + \Big\|\int_0^1 \nabla(f^\nu)^{\text{soc}}(z + \tau(y-z))(y-z)d\tau\Big\| \\
&\quad \le (\kappa + 2\epsilon)\|y - z\| \; .
\end{aligned}$$

Since $y, z \in \mathcal{B}(x, \delta)$ and ϵ is arbitrary, this yields

$$\|f^{\text{soc}}(y) - f^{\text{soc}}(z)\| \le \kappa\|y - z\| \quad \forall y, z \in \mathcal{B}(x, \delta) \; . \qquad (24)$$

Hence, f^{soc} is strictly continuous at x.

"only if" Suppose instead that f^{soc} is strictly continuous at x with spectral values λ_1, λ_2 and spectral vectors $u^{(1)}, u^{(2)}$. Then, there exist scalars $\kappa > 0$ and $\delta > 0$ such that (24) holds. For any $i \in \{1, 2\}$ and any $\psi, \zeta \in [\lambda_i - \delta, \lambda_i + \delta]$, let

$$y := x + (\psi - \lambda_i)u^{(i)}, \qquad z := x + (\zeta - \lambda_i)u^{(i)}.$$

Then, $\|y - x\| = |\psi - \lambda_i|/\sqrt{2} \le \delta$ and $\|z - x\| = |\zeta - \lambda_i|/\sqrt{2} \le \delta$, so it follows from (4) and (24) that

$$\begin{aligned} |f(\psi) - f(\zeta)| &= \sqrt{2}\|f^{\text{soc}}(y) - f^{\text{soc}}(z)\| \\ &\le \sqrt{2}\kappa\|y - z\| \\ &= \kappa|\psi - \zeta|. \end{aligned}$$

This shows that f is strictly continuous at λ_1, λ_2.

(b) is an immediate consequence of (a).

(c) Suppose f is Lipschitz continuous with constant $\kappa > 0$. Then $\text{lip} f(\xi) \le \kappa$ for all $\xi \in \mathbb{R}$. Fix any $x \in \mathbb{R}^n$ with spectral values λ_1, λ_2. For any scalar $\delta > 0$, let

$$C := [\lambda_1 - \delta, \lambda_1 + \delta] \ \cup \ [\lambda_2 - \delta, \lambda_2 + \delta] \ .$$

Then, as in the proof of part (a), we obtain that (24) holds. Since the choice of $\delta > 0$ was arbitrary and κ is independent of δ, this implies that

$$\|f^{\text{soc}}(y) - f^{\text{soc}}(z)\| \le \kappa\|y - z\| \quad \forall y, z \in \mathbb{R}^n \ .$$

Hence, f^{soc} is Lipschitz continuous with Lipschitz constant κ.

Suppose instead that f^{soc} is Lipschitz continuous with constant $\kappa > 0$. Then, for any $\xi, \zeta \in \mathbb{R}$ we have

$$\begin{aligned} |f(\xi) - f(\zeta)| &= \|f^{\text{soc}}(\xi e) - f^{\text{soc}}(\zeta e)\| \\ &\le \kappa\|\xi e - \zeta e\| \\ &= \kappa|\xi - \zeta|, \end{aligned}$$

so f is Lipschitz continuous with constant κ. □

Suppose $f : \mathbb{R} \to \mathbb{R}$ is strictly continuous. Then, by Prop. 6, f^{soc} is strictly continuous. Hence $\partial_B f^{\text{soc}}(x)$ is well defined for all $x \in \mathbb{R}^n$. The following lemma studies the structure of this generalized Jacobian.

Lemma 4. *Let $f : \mathbb{R} \to \mathbb{R}$ be strictly continuous. Then, for any $x \in \mathbb{R}^n$, the generalized Jacobian $\partial_B f^{\text{soc}}(x)$ is well defined and nonempty. Moreover, if $x_2 \neq 0$, then $\partial_B f^{\text{soc}}(x)$ equals the following set*

$$\left\{ \begin{bmatrix} b & c\, x_2^T/\|x_2\| \\ c\, x_2/\|x_2\| & aI + (b-a)x_2x_2^T/\|x_2\|^2 \end{bmatrix} \middle| \, a = \frac{f(\lambda_2) - f(\lambda_1)}{\lambda_2 - \lambda_1}, \begin{array}{l} b + c \in \partial_B f(\lambda_2) \\ b - c \in \partial_B f(\lambda_1) \end{array} \right\}, \tag{25}$$

where λ_1, λ_2 are the spectral values of x. If $x_2 = 0$, then $\partial_B f^{\text{soc}}(x)$ is a subset of the following set

$$\left\{ \begin{bmatrix} b & c\, w^T \\ c\, w & aI + (b-a)ww^T \end{bmatrix} \middle| \, a \in \partial f(x_1),\ b \pm c \in \partial_B f(x_1),\ \|w\| = 1 \right\}. \tag{26}$$

Proof. Suppose $x_2 \neq 0$. For any sequence $\{x^k\}_{k=1}^{\infty} \to x$ with f^{soc} differentiable at x^k, we have from Prop. 4 that $\{\lambda_i^k\}_{k=1}^{\infty} \to \lambda_i$ with f differentiable at λ_i^k, $i = 1, 2$, where λ_1^k, λ_2^k are the spectral values of x^k. Since any cluster point of $\{f'(\lambda_i^k)\}_{k=1}^{\infty}$ is in $\partial_B f(\lambda_i)$, it follows from the gradient formula (19)–(20) that any cluster point of $\{\nabla f^{\text{soc}}(x^k)\}_{k=1}^{\infty}$ is an element of (25). Conversely, for any b, c with $b - c \in \partial_B f(\lambda_1)$, $b + c \in \partial_B f(\lambda_2)$, there exist $\{\lambda_1^k\}_{k=1}^{\infty} \to \lambda_1$, $\{\lambda_2^k\}_{k=1}^{\infty} \to \lambda_2$ with f differentiable at λ_1^k, λ_2^k and $\{f'(\lambda_1^k)\}_{k=1}^{\infty} \to b - c$, $\{f'(\lambda_2^k)\}_{k=1}^{\infty} \to b + c$. Since $\lambda_2 > \lambda_1$, by taking k large, we can assume that $\lambda_2^k \geq \lambda_1^k$ for all k. Let

$$x_1^k = \frac{1}{2}(\lambda_2^k + \lambda_1^k), \qquad x_2^k = \frac{1}{2}(\lambda_2^k - \lambda_1^k)\frac{x_2}{\|x_2\|}, \qquad x^k = (x_1^k, x_2^k).$$

Then, $\{x^k\}_{k=1}^{\infty} \to x$ and, by Prop. 4, f^{soc} is differentiable at x^k. Moreover, the limit of $\{\nabla f^{\text{soc}}(x^k)\}_{k=1}^{\infty}$ is an element of (25) associated with the given b, c. Thus $\partial_B f^{\text{soc}}(x)$ equals (25).

Suppose $x_2 = 0$. Consider any sequence $\{x^k\}_{k=1}^{\infty} = \{(x_1^k, x_2^k)\}_{k=1}^{\infty} \to x$ with f^{soc} differentiable at x^k for all k. By passing to a subsequence, we can assume that either $x_2^k = 0$ for all k or $x_2^k \neq 0$ for all k. If $x_2^k = 0$ for all k, Prop. 4 yields that f is differentiable at x_1^k and $\nabla f^{\text{soc}}(x^k) = f'(x_1^k)I$. Hence any cluster point of $\{\nabla f^{\text{soc}}(x^k)\}_{k=1}^{\infty}$ is an element of (26) with $a = b \in \partial_B f(x_1) \subseteq \partial f(x_1)$ and $c = 0$. If $x_2^k \neq 0$ for all k, by further passing to a subsequence, we can assume without loss of generality that $\{x_2^k/\|x_2^k\|\}_{k=1}^{\infty} \to w$ for some w with $\|w\| = 1$. Let λ_1^k, λ_2^k be the spectral values of x^k and let a^k, b^k, c^k be the coefficients given by (20) corresponding to λ_1^k, λ_2^k. We can similarly prove that $b \pm c \in \partial_B f(x_1)$, where (b, c) is any cluster point of $\{(b^k, c^k)\}_{k=1}^{\infty}$. Also, by a mean-value theorem of Lebourg [8, Prop. 2.3.7],

$$a^k = \frac{f(\lambda_2^k) - f(\lambda_1^k)}{\lambda_2^k - \lambda_1^k} \in \partial f(\hat{\lambda}^k)$$

for some $\hat{\lambda}^k$ in the interval between λ_2^k and λ_1^k. Since f is strictly continuous so that ∂f is upper semicontinuous [8, Prop. 2.1.5] or, equivalently, outer semicontinuous [26, Prop. 8.7], this together with $\lambda_i^k \to x_1$, $i = 1, 2$, implies that any cluster point of $\{a^k\}_{k=1}^{\infty}$ belongs to $\partial f(x_1)$. Then, the gradient formula (19)–(20) yields that any cluster point of $\{\nabla f^{\text{soc}}(x^k)\}_{k=1}^{\infty}$ is an element of (26). $\square$

Below we refine Lemma 4 to characterize $\partial_B f^{\text{soc}}(x)$ completely for two special cases of f. In the first case, the directional derivative of f has a one-sided continuity property, and our characterization is analogous to [6, Prop. 4.8] for the matrix-valued function f^{mat}. However, despite Lemma 1, our characterization cannot be deduced from [6, Prop. 4.8] and hence is proved directly. The second case is an example from [26, p. 304]. Our analysis shows that the structure of $\partial_B f^{\text{soc}}(x)$ depends on f in a complicated way. In particular, in both cases, $\partial_B f^{\text{soc}}(x)$ is a proper subset of (26) when $x_2 = 0$.

In what follows we denote the right- and left-directional derivative of $f : \mathbb{R} \to \mathbb{R}$ by

$$f'_+(\xi) := \lim_{\zeta \to \xi^+} \frac{f(\zeta) - f(\xi)}{\zeta - \xi}, \qquad f'_-(\xi) := \lim_{\zeta \to \xi^-} \frac{f(\zeta) - f(\xi)}{\zeta - \xi}.$$

Lemma 5. *Suppose* $f:\mathbb{R}\to\mathbb{R}$ *is strictly continuous and directionally differentiable function with the property that*

$$\lim_{\substack{\zeta,\nu\to\xi^\sigma\\ \zeta\neq\nu}}\frac{f(\zeta)-f(\nu)}{\zeta-\nu}=\lim_{\substack{\zeta\to\xi^\sigma\\ \zeta\in D_f}}f'(\zeta)=f'_\sigma(\xi),\quad \forall\xi\in\mathbb{R},\ \sigma\in\{-,+\}. \tag{27}$$

where $D_f=\{\xi\in\mathbb{R}|f$ *is differentiable at* $\xi\}$. *Then, for any* $x=(x_1,0)\in\mathbb{R}\times\mathbb{R}^{n-1}$, $\partial_B f(x_1)=\{f'_-(x_1),\ f'_+(x_1)\}$, *and* $\partial_B f^{\text{soc}}(x)$ *equals the following set*

$$\left\{\begin{bmatrix} b & c\,w^T\\ c\,w & aI+(b-a)ww^T\end{bmatrix}\middle|\begin{array}{l}\text{either } a=b\in\partial_B f(x_1),\ c=0\\ \text{or } a\in\partial f(x_1),\ b-c=f'_-(x_1),\ b+c=f'_+(x_1)\end{array},\ \|w\|=1\right\}. \tag{28}$$

Proof. By (27), $\partial_B f(x_1)=\{f'_-(x_1),\ f'_+(x_1)\}$. Consider any sequence $\{x^k\}_{k=1}^\infty\to x$ with f^{soc} differentiable at $x^k=(x_1^k,x_2^k)$ for all k. By passing to a subsequence, we can assume that either $x_2^k=0$ for all k or $x_2^k\neq 0$ for all k.

If $x_2^k=0$ for all k, Prop. 4 yields that f is differentiable at x_1^k and $\nabla f^{\text{soc}}(x^k)=f'(x_1^k)I$. Hence any cluster point of $\{\nabla f^{\text{soc}}(x^k)\}_{k=1}^\infty$ is an element of (28) with $a=b\in\partial_B f(x_1)$ and $c=0$.

If $x_2^k\neq 0$ for all k, by passing to a subsequence, we can assume without loss of generality that $\{x_2^k/\|x_2^k\|\}_{k=1}^\infty\to w$ for some w with $\|w\|=1$. Let λ_1^k,λ_2^k be the spectral values of x^k. Then $\lambda_1^k<\lambda_2^k$ for all k and $\lambda_i^k\to x_1$, $i=1,2$. By further passing to a subsequence if necessary, we can assume that either (i) $\lambda_1^k<\lambda_2^k\le x_1$ for all k or (ii) $x_1\le\lambda_1^k<\lambda_2^k$ for all k or (iii) $\lambda_1^k<x_1<\lambda_2^k$ for all k. Let a^k,b^k,c^k be the coefficients given by (20) corresponding to λ_1^k,λ_2^k. By Prop. 4, f is differentiable at λ_1^k,λ_2^k and $f'(\lambda_1^k)=b^k-c^k$, $f'(\lambda_2^k)=b^k+c^k$. Let (a,b,c) be any cluster point of $\{(a^k,b^k,c^k)\}_{k=1}^\infty$. In case (i), we see from (27) that $b\pm c=a=f'_-(x_1)$, which implies $b=f'_-(x_1)$ and $c=0$. In case (ii), we obtain similarly that $a=b=f'_+(x_1)$ and $c=0$. In case (iii), we obtain that $b-c=f'_-(x_1)$, $b+c=f'_+(x_1)$. Also, the directional differentiability of f implies that

$$a^k=\frac{f(\lambda_2^k)-f(\lambda_1^k)}{\lambda_2^k-\lambda_1^k}=\frac{\lambda_2^k-x_1}{\lambda_2^k-\lambda_1^k}\frac{f(\lambda_2^k)-f(x_1)}{\lambda_2^k-x_1}+\frac{x_1-\lambda_1^k}{\lambda_2^k-\lambda_1^k}\frac{f(x_1)-f(\lambda_1^k)}{x_1-\lambda_1^k},$$

which yields in the limit that

$$a=(1-\omega)f'_+(x_1)+\omega f'_-(x_1),$$

for some $\omega\in[0,1]$. Thus $a\in\partial f(x_1)$. This shows that $\partial_B f^{\text{soc}}(x)$ is a subset of (28).

Conversely, for any $a=b\in\partial_B f(x_1)$, $c=0$ and any $w\in\mathbb{R}^{n-1}$ with $\|w\|=1$, we can find a sequence $x_1^k\in D_f$, $k=1,2,\ldots$, such that $x_1^k\to x_1$ and $f'(x_1^k)\to a$. Then $x^k=(x_1^k,0)\to x$ and the preceding analysis shows that $\{\nabla f^{\text{soc}}(x^k)\}_{k=1}^\infty$ converges to the element of (28) corresponding to the given a,b,c,w. For any a,b,c with $b-c=f'_-(x_1)$, $b+c=f'_+(x_1)$, $a\in\partial f(x_1)$, and any $w\in\mathbb{R}^{n-1}$ with $\|w\|=1$, we have that $a = (1-\omega)f_+(x_1) + \omega f_-(x_1)$ for some $\omega \in [0, 1]$. Since D_f is dense in $\mathbb{R}$, for any integer $k \geq 1$,

$$D_f \cap \left[x_1 - \omega\frac{1}{k} - \frac{1}{k^2}, x_1 - \omega\frac{1}{k}\right] \neq \emptyset, \qquad D_f \cap \left[x_1 + (1-\omega)\frac{1}{k}, x_1 + (1-\omega)\frac{1}{k} + \frac{1}{k^2}\right] \neq \emptyset.$$

Let λ_1^k be any element of the first set and let λ_2^k be any element of the second set. Then $x^k = \left(\frac{\lambda_2^k + \lambda_1^k}{2}, \frac{\lambda_2^k - \lambda_1^k}{2} w\right) \to x$ and x^k has spectral values $\lambda_1^k < \lambda_2^k$ which satisfy

$$\lambda_1^k < x_1 < \lambda_2^k \;\; \forall k, \qquad \frac{\lambda_2^k - x_1}{\lambda_2^k - \lambda_1^k} \to 1 - \omega, \qquad \frac{x_1 - \lambda_1^k}{\lambda_2^k - \lambda_1^k} \to \omega.$$

The preceding analysis shows that $\{\nabla f^{\text{soc}}(x^k)\}_{k=1}^{\infty}$ converges to the element of (28) corresponding to the given a, b, c, w. □

The assumptions of Lemma 5 are satisfied if f is piecewise continuously differentiable, e.g., $f(\cdot) = |\cdot|$ or $f(\cdot) = \max\{0, \cdot\}$. If f is differentiable, but not continuously differentiable, then $\partial_B f^{\text{soc}}(x)$ is more complicated as is shown in the following lemma.

Lemma 6. *Suppose $f : \mathbb{R} \to \mathbb{R}$ is defined by*

$$f(\xi) = \begin{cases} \xi^2 \sin(1/\xi) & \text{if} \xi \neq 0, \\ 0 & \text{else.} \end{cases}$$

Then, for any $x = (x_1, 0) \in \mathbb{R} \times \mathbb{R}^{n-1}$, we have that $\partial_B f(x_1) = [-1, 1]$, and $\partial_B f^{\text{soc}}(x) = \{f'(x_1)I\}$ if $x_1 \neq 0$ and otherwise equals the following set

$$\left\{ \begin{bmatrix} b & c\,w^T \\ c\,w & aI + (b-a)ww^T \end{bmatrix} \;\middle|\; \begin{array}{l} b - c = -\cos(\theta_1),\; b + c = -\cos(\theta_2),\; \|w\| = 1, \\ a = \dfrac{\sin(\theta_1) - \sin(\theta_2)}{\theta_1 - \theta_2 + 2\kappa\pi},\; \kappa \in \{0, 1, ..., \infty\},\; \theta_1, \theta_2 \in [0, 2\pi], \\ \qquad\qquad\qquad\qquad \theta_1 \geq \theta_2 \text{ if } \kappa = 0 \end{array} \right\}, \tag{29}$$

with the convention that $a = 0$ if $\kappa = \infty$ and $a = \cos(\theta_1)$ if $\kappa = 0$ and $\theta_1 = \theta_2$.

Proof. f is differentiable everywhere, with

$$f'(\xi) = \begin{cases} 2\xi \sin(1/\xi) - \cos(1/\xi) & \text{if} \xi \neq 0, \\ 0 & \text{else.} \end{cases} \tag{30}$$

Thus $\partial_B f(x_1) = [-1, 1]$. Consider any sequence $\{x^k\}_{k=1}^{\infty} \to x$ with f^{soc} differentiable at $x^k = (x_1^k, x_2^k)$ for all k. By passing to a subsequence, we can assume that either $x_2^k = 0$ for all k or $x_2^k \neq 0$ for all k. Let $\lambda_1^k = x_1^k - \|x_2^k\|$, $\lambda_2^k = x_1^k + \|x_2^k\|$ be the spectral values of x^k.

If $x_2^k = 0$ for all k, Prop. 4 yields that f is differentiable at x_1^k and $\nabla f^{\text{soc}}(x^k) = f'(x_1^k)I$. Hence any cluster point of $\{\nabla f^{\text{soc}}(x^k)\}_{k=1}^{\infty}$ is of the form bI for some $b \in \partial_B f(x_1)$. If $x_1 \neq 0$, then $b = f'(x_1)$. If $x_1 = 0$, then $b \in [-1, 1]$, i.e., $b = \cos(\theta_1)$

for some $\theta \in [0, 2\pi]$. Then bI has the form (29) with $a = b$, $c = 0$, corresponding to $\theta_1 = \theta_2$, $\kappa = 0$.

If $x_2^k \neq 0$ for all k, by passing to a subsequence, we can assume without loss of generality that $\{x_2^k/\|x_2^k\|\}_{k=1}^\infty \to w$ for some w with $\|w\| = 1$. By Prop. 4, f is differentiable at λ_1^k, λ_2^k and $f'(\lambda_1^k) = b^k - c^k$, $f'(\lambda_2^k) = b^k + c^k$, where a^k, b^k, c^k are the coefficients given by (20) corresponding to λ_1^k, λ_2^k. If $x_1 \neq 0$, then $a^k \to f'(x_1)$, $b^k \to f'(x_1)$ and $c^k \to 0$, so any cluster point of $\{\nabla f^{\text{soc}}(x^k)\}_{k=1}^\infty$ equals $f'(x_1)I$. Suppose $x_1 = 0$. Then $\lambda_1^k < \lambda_2^k$ tend to zero. By further passing to a subsequence if necessary, we can assume that either (i) both are nonzero for all k or (ii) $\lambda_1^k = 0$ for all k or (iii) $\lambda_2^k = 0$ for all k. In case (i),

$$\frac{1}{\lambda_1^k} = \theta_1^k + 2\nu_k\pi, \qquad \frac{1}{\lambda_2^k} = \theta_2^k + 2\mu_k\pi \tag{31}$$

for some $\theta_1^k, \theta_2^k \in [0, 2\pi]$ and integers ν_k, μ_k tending to ∞ or $-\infty$. By further passing to a subsequence if necessary, we can assume that $\{(\theta_1^k, \theta_2^k)\}_{k=1}^\infty$ converges to some $(\theta_1, \theta_2) \in [0, 2\pi]^2$. Then (30) yields

$$f'(\lambda_i^k) = 2\lambda_i^k \sin(\theta_i^k) - \cos(\theta_i^k) \ \to\ -\cos(\theta_i), \quad i = 1, 2,$$

$$\begin{aligned} a^k \ &= \ \frac{f(\lambda_2^k) - f(\lambda_1^k)}{\lambda_2^k - \lambda_1^k} = \frac{(\lambda_2^k)^2 \sin(\theta_2^k) - (\lambda_1^k)^2 \sin(\theta_1^k)}{\lambda_2^k - \lambda_1^k} \\ &= (\lambda_2^k + \lambda_1^k)\sin(\theta_2^k) + \frac{\sin(\theta_2^k) - \sin(\theta_1^k)}{(\theta_1^k - \theta_2^k + 2(\nu_k - \mu_k)\pi)\lambda_2^k/\lambda_1^k}. \end{aligned}$$

If $|\nu_k - \mu_k|$ is bounded as $k \to \infty$, then $\lambda_2^k/\lambda_1^k \to 1$ and, by (31) and $\lambda_1^k < \lambda_2^k$, $\nu_k \geq \mu_k$. In this case, any cluster point (a, b, c) of $\{(a^k, b^k, c^k)\}_{k=1}^\infty$ would satisfy

$$b - c = -\cos(\theta_1), \qquad b + c = -\cos(\theta_2), \qquad a = \frac{\sin(\theta_2) - \sin(\theta_1)}{\theta_1 - \theta_2 + 2\kappa\pi} \tag{32}$$

for some integer $\kappa \geq 0$. Here, we use the convention that $a = \cos(\theta_1)$ if $\kappa = 0$, $\theta_1 = \theta_2$. Moreover, if $\kappa = 0$, then $\nu_k = \mu_k$ for all k sufficiently large along the corresponding subsequence, so (31) and $\lambda_1^k < \lambda_2^k$ yields $\theta_1^k > \theta_2^k > 0$, implying furthermore that $\theta_1 \geq \theta_2$. If $|\nu_k - \mu_k| \to \infty$ and $|\mu_k/\nu_k|$ is bounded away from zero, then $|\nu_k - \mu_k||\mu_k/\nu_k| \to \infty$. If $|\nu_k - \mu_k| \to \infty$ and $|\mu_k/\nu_k| \to 0$, then $|\nu_k - \mu_k||\mu_k/\nu_k| = |\mu_k(1 - \mu_k/\nu_k)| \to \infty$ due to $|\mu_k| \to \infty$. Thus, if $|\nu_k - \mu_k| \to \infty$, we have $|\nu_k - \mu_k||\lambda_2^k/\lambda_1^k| \to \infty$ and the above equation yields $a^k \to 0$, corresponding to (32) with $\kappa = \infty$. In case (ii), we have $f'(\lambda_1^k) = 0$ and $a^k = f(\lambda_2^k)/\lambda_2^k = \lambda_2^k \sin(1/\lambda_2^k)$ for all k, so any cluster point (a, b, c) of $\{(a^k, b^k, c^k)\}_{k=1}^\infty$ satisfies $b - c = 0$, $b + c = -\cos(\theta_2)$, $a = 0$. This corresponds to (32) with $\theta_1 = \frac{\pi}{2}$, $\kappa = \infty$. In case (iii), we obtain similarly (32) with $\theta_2 = \frac{\pi}{2}$, $\kappa = \infty$. This and (19)–(20) show that any cluster point of $\{\nabla f^{\text{soc}}(x^k)\}_{k=1}^\infty$ is in the set (29).

Conversely, if $x_1 \neq 0$, since $\partial_B f^{\text{soc}}(x)$ is a nonempty subset of $\{f'(x_1)I\}$, the two must be equal. If $x_1 = 0$, then for any integer $\kappa \geq 0$ and any $\theta_1, \theta_2 \in [0, 2\pi]$ satisfying $\theta_1 \geq \theta_2$ whenever $\kappa = 0$, and any $w \in \mathbb{R}^{n-1}$ with $\|w\| = 1$, we let, for each integer $k \geq 1$,

$$\lambda_1^k = \frac{1}{\theta_1 + 2(k+\kappa)\pi + 1/k}, \qquad \lambda_2^k = \frac{1}{\theta_2 + 2k\pi}.$$

Then $0 < \lambda_1^k < \lambda_2^k, x^k = \left(\frac{\lambda_2^k + \lambda_1^k}{2}, \frac{\lambda_2^k - \lambda_1^k}{2} w\right) \to x$ and x^k has spectral values λ_1^k, λ_2^k which satisfy (31) with $\nu_k = k + \kappa$, $\mu_k = k$, $\theta_1^k = \theta_1 + 1/k$, $\theta_2^k = \theta_2$. The preceding analysis shows that $\{\nabla f^{\text{soc}}(x^k)\}_{k=1}^{\infty}$ converges to the element of (28) corresponding to the given $\theta_1, \theta_2, \kappa, w$ with a given by (32). The case of $a = 0$ can be obtained similarly by taking κ to go to ∞ with k. □

The following lemma, proven by Sun and Sun [28, Thm. 3.6] using the definition of generalized Jacobian,[1] enables one to study the semismooth property of f^{soc} by examining only those points $x \in \mathbb{R}^n$ where f^{soc} is differentiable and thus work only with the Jacobian of f^{soc}, rather than the generalized Jacobian.

Lemma 7. *Suppose $F : \mathbb{R}^k \to \mathbb{R}^k$ is strictly continuous and directionally differentiable in a neighborhood of $x \in \mathbb{R}^k$. Then, for any $0 < \rho < \infty$, the following two statements (where $O(\cdot)$ depends on F and x only) are equivalent:*

(a) *For any $h \in \mathbb{R}^k$ and any $V \in \partial F(x + h)$,*

$$F(x + h) - F(x) - Vh = o(\|h\|) \quad (\textit{respectively, } O(\|h\|^{1+\rho})).$$

(b) *For any $h \in \mathbb{R}^k$ such that F is differentiable at $x + h$,*

$$F(x + h) - F(x) - \nabla F(x + h)h = o(\|h\|) \quad (\textit{respectively, } O(\|h\|^{1+\rho})).$$

By using Lemmas 1, 7 and Props. 1, 3, 6, 4, we can now state and prove the last result of this section, on the semismooth property of f^{soc}. This result generalizes [7, Thm. 4.2] for the cases of $f(\xi) = |\xi|$, $f(\xi) = \max\{0, \xi\}$.

Proposition 7. *For any $f : \mathbb{R} \to \mathbb{R}$, the vector-valued function f^{soc} is semismooth if and only if f is semismooth. If f is ρ-order semismooth ($0 < \rho < \infty$), then f^{soc} is $\min\{1, \rho\}$-order semismooth.*

Proof. Suppose f is semismooth. Then f is strictly continuous and directionally differentiable. By Props. 3 and 6, f^{soc} is strictly continuous and directionally differentiable. By Lemma 1(b), $f^{\text{soc}}(x) = f^{\text{mat}}(L_x)e$ for all x. By Prop. 1(g), f^{mat} is semismooth. Since L_x is continuously differentiable in x, $f^{\text{soc}}(x) = f^{\text{mat}}(L_x)e$ is semismooth in x. If f is ρ-order semismooth ($0 < \rho < \infty$), then, by Prop. 1(g), f^{mat} is $\min\{1, \rho\}$-order semismooth. Since L_x is continuously differentiable in x, $f^{\text{soc}}(x) = f^{\text{mat}}(L_x)e$ is $\min\{1, \rho\}$-order semismooth in x.

Suppose f^{soc} is semismooth. Then f^{soc} is strictly continuous and directionally differentiable. By Props. 3 and 6, f is strictly continuous and directionally differentiable. For any $\xi \in \mathbb{R}$ and any $\eta \in \mathbb{R}$ such that f is differentiable at $\xi + \eta$, Prop. 4 yields that f^{soc} is differentiable at $x + h$, where we denote $x := \xi e$ and $h := \eta e$. Since f^{soc} is semismooth, it follows from Lemma 7 that

$$f^{\text{soc}}(x + h) - f^{\text{soc}}(x) - \nabla f^{\text{soc}}(x + h)h = o(\|h\|),$$

[1] Sun and Sun did not consider the case of $o(\|h\|)$ but their argument readily applies to this case.

which, by (4) and (18), is equivalent to

$$f(\xi + \eta) - f(\xi) - f'(\xi + \eta)\eta = o(|\eta|).$$

Then Lemma 7 yields that f is semismooth. □

Notice that, for each of the preceding global results there is a corresponding local result. Some of our results, namely Props. 2, 4, 5 and Lemma 2, had appeared in the unpublished Master thesis by the first author [4]. However, the proofs in [4] did not make use of Lemma 1 and hence were more complex in some cases.

6. Applications to SOCCP

Consider the SOCCP (5), i.e., for a given mapping $F : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^\ell \to \mathbb{R}^n \times \mathbb{R}^\ell$, find an $(x, y, \zeta) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^\ell$ satisfying

$$\langle x, y\rangle = 0, \quad x \in \mathcal{K}, \quad y \in \mathcal{K}, \quad F(x, y, \zeta) = 0, \tag{33}$$

where $\mathcal{K}$ is given by (6). We assume that F is continuously differentiable. It is known [12] that $(x, y, \zeta) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^\ell$ solves SOCCP if and only if it solves the equations

$$H(x, y, \zeta) := \begin{pmatrix} x - [x - y]_+ \\ F(x, y, \zeta) \end{pmatrix} = 0, \tag{34}$$

where $[\cdot]_+ : \mathbb{R}^n \to \mathcal{K}$ denotes the nearest-point projection onto $\mathcal{K}$, i.e.,

$$[x]_+ := \arg\min\{\|x - y\| \mid y \in \mathcal{K}\}.$$

The function H is nonsmooth due to the nonsmoothness of the projection operator $[\cdot]_+$. Chen, Sun and Sun [7] showed that $[\cdot]_+$ is strongly semismooth, so that H is semismooth. This result also follows from Prop. 7 with $f(\cdot) = \max\{0, \cdot\}$, for which $f^{\text{soc}}(\cdot) = [\cdot]_+$. [Here, f is applied to the spectral decomposition associated with each $\mathcal{K}^{n_i}$, $i = 1, ..., m$.] More generally, the results of Sec. 5 can be used to design and analyze smoothing or nonsmooth Newton-type methods for solving $H(x, y, \zeta) = 0$, such as was done in [7] for SOCCP and in [6] for SDCP when F has the form (7). In particular, it appears that the analysis in [6, Sec. 5] can be adapted to the above SOCCP. For simplicity, we omit the details.

Acknowledgements. We thank the three referees for their careful reading of our paper and their helpful suggestions.

References

1. Alizadeh, F., Schmieta, S.: Symmetric cones, potential reduction methods. In: H. Wolkowicz, R. Saigal, L. Vandenberghe (eds.), Handbook of Semidefinite Programming, Kluwer, Boston, 2000, pp. 195–233
2. Benson, H.Y., Vanderbei, R.J.: Solving problems with semidefinite and related constraints using interior-point methods for nonlinear programming. Math. Program. **95**, 279–302 (2003)
3. Bhatia, R.: Matrix Analysis. Springer-Verlag, New York, 1997

4. Chen, J.-S.: Merit function and nonsmooth functions for second-order cone complementarity problems. Master thesis, Department of Mathematics, University of Washington, Seattle, April 2001
5. Chen, Xin, Tseng, P.: Non-interior continuation methods for solving semidefinite complementarity problems. Math. Program. **95**, 431–474 (2003)
6. Chen, Xin, Qi, Houduo, Tseng, P.: Analysis of nonsmooth symmetric-matrix-valued functions with applications to semidefinite complementarity problems. SIAM J. Optim. **13**, 960–985 (2003)
7. Chen, X.D., Sun, D., Sun, J.: Complementarity functions and numerical experiments for second-order cone complementarity problems. Comput. Optim. Appl. **25**, 39–56 (2003)
8. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley, New York, 1983
9. Faraut, U., Korányi, A.: Analysis on Symmetric Cones. Oxford Mathematical Monographs, Oxford University Press, New York, 1994
10. Ferris, M.C., Pang, J.-S., eds.: Complementarity and Variational Problems: State of the Art. SIAM, Philadelphia, 1997
11. Fischer, A.: Solution of monotone complementarity problems with locally Lipschitzian functions. Math. Program. **76**, 513–532 (1997)
12. Fukushima, M., Luo, Z.-Q., Tseng, P.: Smoothing functions for second-order-cone complementarity problems. SIAM J. Optim. **12**, 436–460 (2001)
13. Fukushima, M., Qi, L., eds.: Reformulation–Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods. Kluwer, Boston, 1999
14. Hayashi, S., Yamashita, N., Fukushima, M.: On the coerciveness of merit functions for the second-order cone complementarity problem. Report, Department of Applied Mathematics and Physics, Kyoto University, Kyoto, Japan, June 2001
15. Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press, Cambridge, 1985
16. Horn, R.A., Johnson, C.R.: Topics in Matrix Analysis. Cambridge Press, Cambridge, United Kingdom, 1991
17. Jiang, H., and Ralph, D.: Global and local superlinear convergence analysis of Newton-type methods for semismooth equations with smooth least squares. In: M. Fukushima, L. Qi (eds.), Reformulation - Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, Kluwer Academic Publishers, Boston, 1999, pp. 181–209
18. Kanzow, C., Nagel, C.: Semidefinite programs: new search directions, smoothing-type methods, and numerical results. SIAM J. Optim. **13**, 1–23 (2002)
19. Lewis, A.S., Sendov, H.S.: Twice differentiable spectral functions. SIAM J. Matrix Anal. Appl. **23**, 368–386 (2001)
20. Lobo, M.S., Vandenberghe, L., Boyd, S., Lebret, H.: Applications of second-order cone programming. Lin. Algeb. Appl. **284**, 193–228 (1998)
21. Mifflin, R.: Semismooth and semiconvex functions in constrained optimization. SIAM J. Control Optim. **15**, 959-972 (1977)
22. Monteiro, R.D.C., Tsuchiya, T.: Polynomial convergence of primal-dual algorithms for the second-order cone programs based on the MZ-family of directions. Math. Program. **88**, 61–83 (2000)
23. Pang, J.-S., Sun, D., Sun, J.: Semismooth homeomorphisms and strong stability of semidefinite and Lorentz cone complementarity problems. Math. Oper. Res. **28**, 39–63 (2003)
24. Qi, L.: Convergence analysis of some algorithms for solving nonsmooth equations. Math. Oper. Res. **18**, 227–244 (1993)
25. Qi, L., Sun, J.: A nonsmooth version of Newton's method. Math. Program. **58**, 353–367 (1993)
26. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Springer-Verlag, Berlin, 1998
27. Schmieta, S., Alizadeh, F.: Associative and Jordan algebras, and polynomial time interior-point algorithms for symmetric cones. Math. Oper. Res. **26**, 543–564 (2001)
28. Sun, D., Sun, J.: Semismooth matrix valued functions. Math. Oper. Res. **27**, 150–169 (2002)
29. Sun, J., Sun, D., Qi, L.: Quadratic convergence of a squared smoothing Newton method for nonsmooth matrix equations and its applications in semidefinite optimization problems. submitted to SIAM J. Optim.
30. Tseng, P.: Merit functions for semi-definite complementarity problems. Math. Program. **83**, 159–185 (1998)
31. Tsuchiya, T.: A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming. Optim. Methods Softw. **11**, 141–182 (1999)
32. Yamashita, N., Fukushima, M.: Modified Newton methods for solving semismooth reformulations of monotone complementarity problems. Math. Program. **76**, 469–491 (1997)
33. Yamashita, N., Fukushima, M.: A new merit function and a descent method for semidefinite complementarity problems. In: M. Fukushima, L. Qi (eds.), Reformulation - Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, Kluwer Academic Publishers, Boston, 1999, pp. 405–420