数理解析研究所講究録
第 1544 巻 2007 年 153-162

On merit functions for the second-order cone complementarity problem

Jein-Shan Chen [1]
Department of Mathematics
National Taiwan Normal University
Taipei, Taiwan 11677
E-mail: jschen@math.ntnu.edu.tw

November 20, 2006

Abstract: There are three popular approaches, merit functions approach, nonsmooth functions approach, and smoothing methods approach, for the second-order cone complementarity problem (SOCCP). In this article, we survey recent results on the most popular approach, merit functions approach. In particular, we investigate and present several merit functions for SOCCP. We also propose some open questions for future study.

Key words. Second-order cone, complementarity, merit function, Jordan product

AMS subject classifications. 26B05, 26B35, 90C33, 65K05

1 Introduction

The second-order cone complementarity problem (SOCCP), which is a natural extension of nonlinear complementarity problem (NCP), is to find $\zeta \in \mathbb{R}^n$ satisfying

$$\langle F(\zeta), \zeta \rangle = 0, \quad F(\zeta) \in \mathcal{K}, \quad \zeta \in \mathcal{K}, \tag{1}$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product, $F : \mathbb{R}^n \to \mathbb{R}^n$ is a continuous mapping, and $\mathcal{K}$ is the Cartesian product of second-order cones (SOC), also called Lorentz cones [11]. In other words,

$$\mathcal{K} = \mathcal{K}^{n_1} \times \cdots \times \mathcal{K}^{n_m}, \tag{2}$$

where $m, n_1, \ldots, n_m \geq 1$, $n_1 + \cdots + n_m = n$, and

$$\mathcal{K}^{n_i} := \{(x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n_i - 1} \mid \|x_2\| \leq x_1\}, \tag{3}$$

with $\|\cdot\|$ denoting the Euclidean norm and $\mathcal{K}^1$ denoting the set of nonnegative reals $\mathbb{R}_+$. A special case of (2) is $\mathcal{K} = \mathbb{R}^n_+$, the nonnegative orthant in $\mathbb{R}^n$, which corresponds to $m = n$ and $n_1 = \cdots = n_m = 1$. If $\mathcal{K} = \mathbb{R}^n_+$, then (1) reduces to the nonlinear complementarity problem. Throughout this paper, we assume $\mathcal{K} = \mathcal{K}^n$ for simplicity, i.e., $\mathcal{K}$ is a single second-order cone (all the analysis can be easily carried over to the general case where $\mathcal{K}$ has the direct product structure (2)).

[1] Member of Mathematics Division, National Center for Theoretical Sciences, Taipei Office.

There have been various methods proposed for solving SOCCP. They include interior-point methods [2, 17, 19, 21, 24], non-interior smoothing Newton methods [9, 14, 15]. Recently in the papers [3, 4, 7], the author studied an alternative approach based on reformulating SOCCP as an unconstrained smooth minimization problem. For this approach, it aims to find a smooth function $\psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ such that

$$\psi(x,y) = 0 \quad \Longleftrightarrow \quad x \in \mathcal{K}^n, \quad y \in \mathcal{K}^n, \quad \langle x, y \rangle = 0. \tag{4}$$

Then SOCCP can be expressed as an unconstrained smooth (global) minimization problem:

$$\min_{\zeta \in \mathbb{R}^n} f(\zeta) := \psi(F(\zeta), \zeta). \tag{5}$$

We call such a f a *merit function* for the SOCCP.

A popular choice of ψ is the squared norm of Fischer-Burmeister function, i.e., $\psi_{\rm FB} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ associated with second-order cone given by

$$\psi_{\rm FB}(x,y) \;=\; \frac{1}{2}\|\phi_{\rm FB}(x,y)\|^2, \tag{6}$$

where $\phi_{\rm FB} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is the well-known Fischer-Burmeister function (originally proposed for NCP, see [12, 13]) defined by

$$\phi_{\rm FB}(x,y) = (x^2 + y^2)^{1/2} - x - y. \tag{7}$$

More specifically, for any $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, we define their *Jordan product* associated with $\mathcal{K}^n$ as

$$x \circ y \;:=\; (\langle x, y \rangle \;,\; y_1 x_2 + x_1 y_2\;). \tag{8}$$

The Jordan product $\circ$, unlike scalar or matrix multiplication, is not associative, which is a main source on complication in the analysis of SOCCP. The identity element under this product is $e := (1, 0, \ldots, 0)^T \in \mathbb{R}^n$. We write x^2 to mean $x \circ x$ and write $x + y$ to mean the usual componentwise addition of vectors. It is known that $x^2 \in \mathcal{K}^n$ for all $x \in \mathbb{R}^n$. Moreover, if $x \in \mathcal{K}^n$, then there exists a unique vector in $\mathcal{K}^n$, denoted by $x^{1/2}$, such that $(x^{1/2})^2 = x^{1/2} \circ x^{1/2} = x$. Thus, $\phi_{\rm FB}$ defined as (7) is well-defined for all $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$ and maps $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}^n$. It was shown in [14] that $\phi_{\rm FB}(x,y) = 0$ if and only if (x,y) satisfies (4). Therefore, $\psi_{\rm FB}$ defined as (6) induces a merit function $f_{\rm FB} := \psi_{\rm FB}(F(\zeta), \zeta))$ for the SOCCP.

The function $\psi_{\rm FB}$ given as in (6) was proved smooth with computable gradient formulas and enjoys several favorable properties, nonetheless, it does not have additional bounded level-set and error bound properties (see [7]). To conquer this, several other functions associated with second-order cone were considered [3, 4, 7]. The first one is $\psi_{\rm YF} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by

$$\psi_{\rm YF}(x,y) := \psi_0(\langle x, y \rangle) + \psi_{\rm FB}(x,y), \tag{9}$$

where $\psi_0 : \mathbb{R} \to \mathbb{R}_+$ is any smooth function satisfying

$$\psi_0(t) = 0 \quad \forall t \le 0 \quad \text{and} \quad \psi_0'(t) > 0 \quad \forall t > 0. \tag{10}$$

The function ψ_{YF} was studied by Yamashita and Fukushima [25] for SDCP (semi-definite complementarity problems) case and was extended to SOCCP case in [7]. An example of $\psi_0(t)$ is $\psi_0(t) = \frac{1}{4}(\max\{0,t\})^4$. A slight modification of ψ_{YF} yields $\widehat{\psi_{\mathrm{YF}}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by

$$\widehat{\psi_{\mathrm{YF}}}(x,y) := \frac{1}{2}\|(x \circ y)_+\|^2 + \psi_{\mathrm{FB}}(x,y), \tag{11}$$

where $(\cdot)_+$ means the orthogonal projection onto the second-order cone $\mathcal{K}^n$. The third function is $\psi_{\mathrm{LT}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by

$$\psi_{\mathrm{LT}}(x,y) := \psi_0(\langle x,y\rangle) + \widetilde{\psi}(x,y), \tag{12}$$

where $\widetilde{\psi} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ satisfies

$$\widetilde{\psi}(x,y) = 0, \ \langle x,y\rangle \le 0 \quad \Longleftrightarrow \quad x \in \mathcal{K}^n, \ y \in \mathcal{K}^n, \ \langle x,y\rangle = 0. \tag{13}$$

The function ψ_0 is the same as the above and examples of $\widetilde{\psi}$ are

$$\widetilde{\psi}_1(x,y) := \frac{1}{2}\left(\|(-x)_+\|^2 + \|(-y)_+\|^2\right) \quad \text{and} \quad \widetilde{\psi}_2(x,y) := \frac{1}{2}\|\phi_{\mathrm{FB}}(x,y)_+\|^2 \tag{14}$$

which were recently investigated in [4]. The function ψ_{LT} was proposed by Luo and Tseng for NCP case in [18] and was extended to the SDCP case by Tseng in [23]. The last function $\widehat{\psi_{\mathrm{LT}}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, a slight variant of ψ_{LT}, is defined by

$$\widehat{\psi_{\mathrm{LT}}}(x,y) := \frac{1}{2}\|(x \circ y)_+\|^2 + \widetilde{\psi}(x,y), \tag{15}$$

where $\widetilde{\psi}$ is given as in (13).

Each of the above functions naturally induces a merit function as follows:

$$\begin{array}{rcl} f_{\mathrm{YF}}(\zeta) & := & \psi_{\mathrm{YF}}(F(\zeta),\zeta), \\ \widehat{f_{\mathrm{YF}}}(\zeta) & := & \widehat{\psi_{\mathrm{YF}}}(F(\zeta),\zeta), \\ f_{\mathrm{LT}}(\zeta) & := & \psi_{\mathrm{LT}}(F(\zeta),\zeta), \\ \widehat{f_{\mathrm{LT}}}(\zeta) & := & \widehat{\psi_{\mathrm{LT}}}(F(\zeta),\zeta). \end{array} \tag{16}$$

It was shown that f_{YF} provides error bound [7, Prop. 5] if F is strongly monotone and f_{YF} has bounded level set [7, Prop. 6] if F is monotone as well as SOCCP is strictly feasible. The same results hold for $\widehat{f_{\mathrm{YF}}}$ [3, Prop. 4.1 and Prop. 4.2], for f_{LT} [4, Prop. 4.1 and Prop. 4.3], and for $\widehat{f_{\mathrm{LT}}}$ [4, Prop. 4.2 and Prop. 4.4].

Next, we also investigate the following one-parametric class of functions, $\phi_\lambda : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ defined as

$$\phi_\lambda(x,y) := \left[(x-y)^2 + \lambda(x \circ y)\right]^{1/2} - (x+y), \tag{17}$$

where λ is a fixed parameter such that $\lambda \in (0,4)$. It can be verified that for any $x, y \in \mathbb{R}^n$

$$\begin{array}{rl} & (x-y)^2 + \lambda(x \circ y) \\ = & x^2 + (\lambda-2)(x \circ y) + y^2 \\ = & \left[x + (\frac{\lambda-2}{2})y\right]^2 + \left[1 - \frac{(\lambda-2)^2}{4}\right] y^2 \\ = & \left[x + (\frac{\lambda-2}{2})y\right]^2 + \frac{\lambda(4-\lambda)}{4} y^2 \\ \succeq_{\mathcal{K}^n} & 0, \end{array} \tag{18}$$

where the inequality holds because $\lambda \in (0,4)$. Therefore, ϕ_λ is well-defined. Furthermore, we let $\psi_\lambda : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be

$$\psi_\lambda(x,y) = \frac{1}{2}\|\phi_\lambda(x,y)\|^2. \tag{19}$$

We will see that ψ_λ is a differentiable merit function, with computable gradient formulas, for SOCCP. In other words, the SOCCP can be expressed as an unconstrained differentiable global minimization problem :

$$\min_{\zeta\in\mathbb{R}^n} f_\lambda(\zeta) := \psi_\lambda(F(\zeta), G(\zeta)). \tag{20}$$

Moreover, we will also show that every stationary point of (20) solves the SOCCP when ∇F and $-\nabla G$ are column monotone (see Prop. 3.2). Indeed, we say that $M, N \in \mathbb{R}^{n\times n}$ are column monotone if, for any $u, v \in \mathbb{R}^n$, $Mu + Nv = 0 \;\Rightarrow\; u^T v \geq 0$. In Prop. 3.2, we assume that

$$\nabla F(\zeta), -\nabla G(\zeta) \text{ are column monotone} \qquad \forall \zeta \in \mathbb{R}^n. \tag{21}$$

Notice that ϕ_λ reduces to the FB function ϕ_{FB} when $\lambda = 2$, whereas it becomes a multiple of the natural residual function ϕ_{NR} when $\lambda \to 0$. Thus, this class of merit functions covers the most two important merit functions for SOCCP so that a closer look and study of this new class of functions is worthwhile. In fact, this study is motivated by the work [16] where the function ψ_λ was considered for the NCP.

Finally, we introduce another two important merit functions for the SOCCP, which are not variants of FB function. The first one is the Implicit Lagrangian function $\psi_{\text{MS}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ defined by

$$\psi_{\text{MS}}(x,y) = \langle x, y\rangle + \frac{1}{2\alpha}\left(\|(x-\alpha y)_+\|^2 - \|x\|^2 + \|(y-\alpha x)_+\|^2 - \|y\|^2\right), \tag{22}$$

where $\alpha > 1$ and $(\cdot)_+$ is the orthogonal projection onto $\mathcal{K}^n$. The function ψ_{MS} was introduced by Mangasarian and Solodov in [20] for the NCP. The other one is based on the NCP-function proposed by Evtushenko and Purtov in [10]. It is $\psi_{\text{EP}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ defined by

$$\psi_{\text{EP}}(x,y) = \frac{1}{2}\|\phi_{\text{EP}}(x,y)\|^2, \tag{23}$$

where $\phi_{\text{EP}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is given by

$$\phi_{\text{EP}}(x,y) := -(x\circ y) + \frac{1}{2\beta_1}(x+y)_-^2 \qquad 0 < \beta_1 \leq 1. \tag{24}$$

Throughout this paper, $\mathbb{R}^n$ denotes the space of n-dimensional real column vectors and ^{T} denotes transpose. For any differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, $\nabla f(x)$ denotes the gradient of f at x. For any differentiable mapping $F = (F_1, ..., F_m)^T : \mathbb{R}^n \to \mathbb{R}^m$, $\nabla F(x) = [\nabla F_1(x) \;\cdots\; \nabla F_m(x)]$ is a $n \times m$ matrix denoting the transpose Jacobian of F at x.

2 Jordan product and spectral factorization

For any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, its *determinant* is defined by $det(x) := x_1^2 - \|x_2\|^2$. In general, $det(x \circ y) \neq det(x)det(y)$ unless $x_2 = y_2$. A vector $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ is said to be *invertible* if $det(x) \neq 0$. If x is invertible, then there exists a unique $y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ satisfying $x \circ y = y \circ x = e$. We call this y the inverse of x and denote it by x^{-1}. In fact, we have $x^{-1} = \dfrac{1}{x_1^2 - \|x_2\|^2}(x_1 \ , \ -x_2)$. Therefore, $x \in \text{int}(\mathcal{K}^n)$ if and only if $x^{-1} \in \text{int}(\mathcal{K}^n)$. Moreover, if $x \in \text{int}(\mathcal{K}^n)$, then $x^{-k} = (x^k)^{-1}$ is also well-defined. For any $x \in \mathcal{K}^n$, it is known that there exists a unique vector in $\mathcal{K}^n$ denoted by $x^{1/2}$ such that $(x^{1/2})^2 = x^{1/2} \circ x^{1/2} = x$. More specifically, $x^{1/2} = \left(s \ , \ \frac{x_2}{2s}\right)$, where $s = \sqrt{\frac{1}{2}\left(x_1 + \sqrt{x_1^2 - \|x_2\|^2}\right)}$. In the above formula, the term x_2/s is defined to be the zero vector if $x_2 = 0$ and $s = 0$, i.e., $x = 0$.

For any $x \in \mathbb{R}^n$, we always have $x^2 \in \mathcal{K}^n$ (i.e., $x^2 \succeq_{\mathcal{K}^n} 0$). Hence, there exists a unique vector $(x^2)^{1/2} \in \mathcal{K}^n$ denoted by $|x|$. It is easy to verify that $|x| \succeq_{\mathcal{K}^n} 0$ and $x^2 = |x|^2$ for any $x \in \mathbb{R}^n$. It is also known that $|x| \succeq_{\mathcal{K}^n} x$ and that $|x|$ and x are related to each other just like the cases of nonnegative orthant $\mathbb{R}^n_+$ and positive semi-definite cone $\mathcal{S}^n_+$. For any $x \in \mathbb{R}^n$, we define $[x]_+$ to be the nearest point (in Euclidean norm, since Jordan product does not induce a norm) projection of x onto $\mathcal{K}^n$, which is the same definition as in $\mathbb{R}^n_+$. In other words, $[x]_+$ is the optimal solution of the parametric SOCP: $[x]_+ = \text{argmin}\{ \ \|x - y\| \mid y \in \mathcal{K}^n\}$. It is well known that $[x]_+ = \frac{1}{2}(x + |x|)$.

Now, for any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, we define a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$ as

$$\begin{array}{rccl} L_x: & \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ & y & \longrightarrow & L_x y := \begin{bmatrix} x_1 & x_2^T \\ x_2 & x_1 I \end{bmatrix} y \ . \end{array}$$

It is easily verified that $x \circ y = L_x y$, $\forall y \in \mathbb{R}^n$, and L_x is positive definite (and hence invertible) if and only if $x \in \text{int}(\mathcal{K}^n)$. However, $L_x^{-1} y \neq x^{-1} \circ y$, for some $x \in \text{int}(\mathcal{K}^n)$ and $y \in \mathbb{R}^n$, i.e., $L_x^{-1} \neq L_{x^{-1}}$. From the above definition, we have the following: $L_{x+y} = L_x + L_y$; $x \succeq_{\mathcal{K}^n} 0 \iff L_x \succeq O$; $x \succeq_{\mathcal{K}^n} y \iff L_x \succeq L_y$ as well as L_x, L_{x^2} commute. General speaking, $L_x^2 = L_x L_x \neq L_{x^2}$, $L_x^{-1} \neq L_{x^{-1}}$ and $L_{x^{1/2}} \neq L_x^{1/2}$.

In addition, we recall from [14] that each $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ admits a spectral factorization, associated with $\mathcal{K}^n$, of the form

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

where λ_1, λ_2 and $u^{(1)}, u^{(2)}$ are the spectral values and the associated spectral vectors of x given by

$$\begin{aligned} \lambda_i &= x_1 + (-1)^i \|x_2\|, \\ u^{(i)} &= \begin{cases} \frac{1}{2}\left(1, \ (-1)^i \dfrac{x_2}{\|x_2\|}\right) & \text{if } x_2 \neq 0; \\ \frac{1}{2}\left(1, \ (-1)^i w_2\right) & \text{if } x_2 = 0, \end{cases} \end{aligned}$$

for $i = 1, 2$, with w_2 being any vector in $\mathbb{R}^{n-1}$ satisfying $\|w_2\| = 1$. If $x_2 \neq 0$, the factorization is unique.

3 Recent results

Proposition 3.1 *[5, Prop. 3.2] Let ϕ_{FB}, ϕ_λ be given by (7) and (17), respectively. Then ψ_λ given by (19) is differentiable at every $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$. Moreover, $\nabla_x\psi_\lambda(0,0) = \nabla_y\psi_\lambda(0,0) = 0$. Let $z := (x-y)^2 + \lambda(x \circ y)$. If $(x,y) \neq (0,0)$ and $(x-y)^2 + \lambda(x \circ y) \in \mathrm{int}(\mathcal{K}^n)$, then*

$$\begin{aligned} \nabla_x\psi_\lambda(x,y) &= \left[\left(L_x + \frac{\lambda-2}{2}L_y\right)L_{z^{1/2}}^{-1} - I\right]\phi_\lambda(x,y), \\ \nabla_y\psi_\lambda(x,y) &= \left[\left(L_y + \frac{\lambda-2}{2}L_x\right)L_{z^{1/2}}^{-1} - I\right]\phi_\lambda(x,y). \end{aligned} \tag{25}$$

If $(x,y) \neq (0,0)$ and $(x-y)^2 + \lambda(x \circ y) \notin \mathrm{int}(\mathcal{K}^n)$, then $x_1^2 + y_1^2 + (\lambda-2)x_1y_1 \neq 0$ and

$$\begin{aligned} \nabla_x\psi_\lambda(x,y) &= \left[\frac{x_1 + \frac{\lambda-2}{2}y_1}{\sqrt{x_1^2 + y_1^2 + (\lambda-2)x_1y_1}} - 1\right]\phi_\lambda(x,y), \\ \nabla_y\psi_\lambda(x,y) &= \left[\frac{y_1 + \frac{\lambda-2}{2}x_1}{\sqrt{x_1^2 + y_1^2 + (\lambda-2)x_1y_1}} - 1\right]\phi_\lambda(x,y). \end{aligned} \tag{26}$$

In particular, when $\lambda = 2$, the above farmulas for gradient of ψ_λ reduce to

$$\begin{aligned} \nabla_x\psi_\lambda(x,y) &= \left[L_x L_{(x^2+y^2)^{1/2}}^{-1} - I\right]\phi_{\text{FB}}(x,y), \\ \nabla_y\psi_\lambda(x,y) &= \left[L_y L_{(x^2+y^2)^{1/2}}^{-1} - I\right]\phi_{\text{FB}}(x,y), \end{aligned} \tag{27}$$

for $(x,y) \neq (0,0)$ with $x^2 + y^2 \in \mathrm{int}(\mathcal{K}^n)$; and reduce to

$$\begin{aligned} \nabla_x\psi_\lambda(x,y) &= \left[\frac{x_1}{\sqrt{x_1^2 + y_1^2}} - 1\right]\phi_{\text{FB}}(x,y), \\ \nabla_y\psi_\lambda(x,y) &= \left[\frac{y_1}{\sqrt{x_1^2 + y_1^2}} - 1\right]\phi_{\text{FB}}(x,y), \end{aligned} \tag{28}$$

for $(x,y) \neq (0,0)$ with $x^2 + y^2 \notin \mathrm{int}(\mathcal{K}^n)$.

Proposition 3.2 *[5, Prop. 4.2] Let ϕ_λ, ψ_λ be given by (17) and (19), respectively. Let f_λ be given by (20), where F and G are differentiable mappings from $\mathbb{R}^n$ to $\mathbb{R}^n$ satisfying (21). Then, for every $\zeta \in \mathbb{R}^n$, either (i) $f_\lambda(\zeta) = 0$ or (ii) $\nabla f_\lambda(\zeta) \neq 0$. In case (ii), if $\nabla G(\zeta)$ is invertible, then $\langle d(\zeta), \nabla f_\lambda(\zeta)\rangle < 0$, where*

$$d(\zeta) := -(\nabla G(\zeta)^{-1})^T \nabla_x\psi_\lambda(F(\zeta), G(\zeta)).$$

Proposition 3.3 *[7, Prop. 4] Let ϕ_{FB} be given by (7), let ψ_{FB} be given by (6), and let ψ_{YF} be given by (9), with $\psi_0 : \mathbb{R} \to [0, \infty)$ being any smooth function satisfying (10). Let f_{YF} be given by (16), where F and G are differentiable mappings from $\mathbb{R}^n$ to $\mathbb{R}^n$ satisfying (21). Then, for every $\zeta \in \mathbb{R}^n$, either (i) $f_{\text{YF}}(\zeta) = 0$ or (ii) $\nabla f_{\text{YF}}(\zeta) \neq 0$. In case (ii), if $\nabla G(\zeta)$ is invertible, then $\langle d_{\text{YF}}(\zeta), \nabla f_{\text{YF}}(\zeta)\rangle < 0$, where*

$$d_{\text{YF}}(\zeta) := -(\nabla G(\zeta)^{-1})^T \Big(\psi_0'(\langle F(\zeta), G(\zeta)\rangle)G(\zeta) + \nabla_x \psi_{\text{FB}}(F(\zeta), G(\zeta))\Big).$$

Proposition 3.4 *[4, Prop. 3.3] Let $f_{\text{LT}} : \mathbb{R}^n \to \mathbb{R}_+$ be given as (12)-(16) with ψ_0 satisfying (10) and $\widetilde{\psi}$ satisfying (13). Then, the following results hold.*

(a) *For all $\zeta \in \mathbb{R}^n$, we have $f_{\text{LT}}(\zeta) \geq 0$ and $f_{\text{LT}}(\zeta) = 0$ if and only if ζ solves the SOCCP.*

(b) *If $\psi_0, \widetilde{\psi}$ and F, G are differentiable, then so is f_{LT} and*

$$\begin{aligned} \nabla f_{\text{LT}}(\zeta) &= \psi_0'(\langle F(\zeta), G(\zeta)\rangle)\Big[\nabla F(\zeta)G(\zeta) + \nabla G(\zeta)F(\zeta)\Big] \\ &\quad + \nabla F(\zeta)\nabla_x \widetilde{\psi}(F(\zeta), G(\zeta)) \\ &\quad + \nabla G(\zeta)\nabla_y \widetilde{\psi}(F(\zeta), G(\zeta)). \end{aligned}$$

(c) *Assume F, G are differentiable on $\mathbb{R}^n$ and $\widetilde{\psi}$ belongs to Ψ_+ (respectively, Ψ_{++}). Then, for every $\zeta \in \mathbb{R}^n$ where $\nabla G(\zeta)^{-1}\nabla F(\zeta)$ is positive definite (respectively, positive semi-definite), either (i) $f_{\text{LT}}(\zeta) = 0$ or (ii) $\nabla f_{\text{LT}}(\zeta) \neq 0$ with $\langle d(\zeta), \nabla f_{\text{LT}}(\zeta)\rangle < 0$, where*

$$d(\zeta) := -(\nabla G(\zeta)^{-1})^T \Big[\psi_0'(\langle F(\zeta), G(\zeta)\rangle)G(\zeta) + \nabla_x \widetilde{\psi}(F(\zeta), G(\zeta))\Big].$$

Proposition 3.5 *[4, Prop. 3.4] Let $\widehat{f_{\text{LT}}} : \mathbb{R}^n \to \mathbb{R}_+$ be given as (15)-(16). Then, the following results hold.*

(a) *For all $x \in \mathbb{R}^n$, we have $\widehat{f_{\text{LT}}}(\zeta) \geq 0$ and $\widehat{f_{\text{LT}}}(\zeta) = 0$ if and only if ζ solves the SOCCP.*

(b) *If $\widetilde{\psi}$ and F, G are differentiable, then so is $\widehat{f_{\text{LT}}}$ and*

$$\begin{aligned} \nabla \widehat{f_{\text{LT}}}(\zeta) &= \Big[\nabla F(\zeta)L_{G(\zeta)} + \nabla G(\zeta)L_{F(\zeta)}\Big](F(\zeta) \circ G(\zeta))_+ \\ &\quad + \nabla F(\zeta)\nabla_x \widetilde{\psi}(F(\zeta), G(\zeta)) \\ &\quad + \nabla G(\zeta)\nabla_y \widetilde{\psi}(F(\zeta), G(\zeta)). \end{aligned}$$

Proposition 3.6 *[6, Prop. 3.2] Let $\psi_{\text{MS}}, \psi_{\text{EP}}$ be defined as in (22) and (23), respectively. Then, the following results hold:*

(a) ψ_{MS} *is continuously differentiable everywhere with*

$$\begin{aligned}\nabla_x\psi_{\mathrm{MS}}(x,y) &= y+\frac{1}{\alpha}\Big[(x-\alpha y)_+ - x\Big] - (y-\alpha x)_+,\\ \nabla_y\psi_{\mathrm{MS}}(x,y) &= x+\frac{1}{\alpha}\Big[(y-\alpha x)_+ - y\Big] - (x-\alpha y)_+.\end{aligned}$$

(b) ψ_{EP} *is continuously differentiable everywhere. Moreover,*

$$\begin{aligned}\nabla_x\psi_{\mathrm{EP}}(x,y) &= \nabla_x\phi_{\mathrm{EP}}(x,y)\cdot\phi_{\mathrm{EP}}(x,y),\\ \nabla_y\psi_{\mathrm{EP}}(x,y) &= \nabla_y\phi_{\mathrm{EP}}(x,y)\cdot\phi_{\mathrm{EP}}(x,y),\end{aligned}$$

where

$$\begin{aligned}\nabla_x\phi_{\mathrm{EP}}(x,y) &= -L_y+\frac{1}{2\beta_1}\begin{bmatrix}2(x_1+y_1)_- & 0\\ 0 & 0\end{bmatrix}, \quad \text{if } x_2+y_2=0;\\ \nabla_y\phi_{\mathrm{EP}}(x,y) &= -L_x+\frac{1}{2\beta_1}\begin{bmatrix}2(x_1+y_1)_- & 0\\ 0 & 0\end{bmatrix}, \quad \text{if } x_2+y_2=0;\end{aligned}$$

and otherwise

$$\begin{aligned}\nabla_x\phi_{\mathrm{EP}}(x,y) &= -L_y+\frac{1}{2\beta_1}\begin{bmatrix} b & \dfrac{c(x_2+y_2)^T}{\|x_2+y_2\|}\\ \dfrac{c(x_2+y_2)}{\|x_2\|} & aI+(b-a)\dfrac{(x_2+y_2)(x_2+y_2)^T}{\|x_2+y_2\|^2}\end{bmatrix},\\ \nabla_y\phi_{\mathrm{EP}}(x,y) &= -L_x+\frac{1}{2\beta_1}\begin{bmatrix} b & \dfrac{c(x_2+y_2)^T}{\|x_2+y_2\|}\\ \dfrac{c(x_2+y_2)}{\|x_2\|} & aI+(b-a)\dfrac{(x_2+y_2)(x_2+y_2)^T}{\|x_2+y_2\|^2}\end{bmatrix},\end{aligned}$$

with $a=\frac{(\lambda_2)_-^2-(\lambda_1)_-^2}{\lambda_2-\lambda_1}$, $b=(\lambda_2)_-+(\lambda_1)_-$, $c=(\lambda_2)_- -(\lambda_1)_-$ *and* $\lambda_i=(x_1+y_1)+(-1)^i\|x_2+y_2\|$ *being spectral values of* $x+y$.

4 Open questions

There are several unresolved questions related to these merit functions introduced in this paper. We propose them as future research topics.

Q1. When $\lambda=2$, ψ_λ reduces ψ_{FB} and it was shown [7] that ψ_{FB} is smooth everywhere. Is ψ_λ, $\lambda\in(0,4)$ smooth everywhere?

Q2. In SDCP case, the gradient of ψ_{FB} was shown Lipschitz continuous in [22]. Is it still true for the SOCCP case?

Q3. Which merit function performs best in numerical implementations for the merit function approach? How about for the other approaches?

Q4. Are there weaker conditions for properties of bounded level lets and error bounds?

References

[1] F. Alizadeh and D. Goldfarb, *Second-order cone programming*, Mathematical Programming, vol. 95, pp. 3-51, 2003.

[2] E. D. Andersen, C. Roos, and T. Terlaky, *On implementing a primal-dual interior-point method for conic quadratic optimization*, Mathematical Programming Ser. B, vol. 95, pp. 249–277, 2003.

[3] J.-S. Chen, *A new merit function and its related properties for the second-order cone complementarity problem*, Pacific Journal of Optimization, vol. 2, pp. 167-179, 2006.

[4] J.-S. Chen, *Two classes of merit functions for the second-order cone complementarity problem*, Mathematical Methods of Operations Research, vol. 64, pp. 495-519, 2006.

[5] J.-S. Chen, *A one-parametric class of merit functions for the second-order cone complementarity problem*, submitted to Mathematical Programming, 2006.

[6] J.-S. Chen and S. Pan, *The implicit Lagrangian merit function and two merit functions for the second-order cone complementarity problem*, submitted to SIAM Journal on Optimization, 2006.

[7] J.-S. Chen and P. Tseng, *An unconstrained smooth minimization reformulation of the second-order cone complementarity problem*, Mathematical Programming, vol. 104, pp. 293-327, 2005.

[8] J.-S. Chen, X. Chen, and P. Tseng, *Analysis of nonsmooth vector-valued functions associated with second-order cone*, Mathematical Programming, vol.101, pp. 97-117, 2004.

[9] X.-D. Chen, D. Sun, and J. Sun, *Complementarity functions and numerical experiments for second-order cone complementarity problems*, Computational Optimization and Applications, vol. 25, pp. 39-56, 2003.

[10] Y.G. Evtushenko and V.A. Purtov (1984), *Sufficient conditions for a minimum for nonlinear programming problems*, Soviet Mathematics Doklady, vol. 30, pp. 313-316.

[11] U. Faraut and A. Korányi, *Analysis on Symmetric Cones*, Oxford Mathematical Monographs, Oxford University Press, New York, 1994.

[12] A. Fischer, *A special Newton-type optimization methods*, Optimization, vol. 24, pp. 269-284, 1992.

[13] A. Fischer, *Solution of the monotone complementarity problem with locally Lipschitzian functions*, Mathematical Programming, vol. 76, pp. 513-532, 1997.

[14] M. Fukushima, Z.-Q. Luo, and P. Tseng, *Smoothing functions for second-order cone complementarity problems*, SIAM Journal on Optimization, vol. 12, pp. 436-460, 2002.

[15] S. HAYASHI, N. YAMASHITA, AND M. FUKUSHIMA, *A combined smoothing and regularization method for monotone second-order cone complementarity problems*, SIAM Journal of Optimization, vol. 15, pp. 593-615, 2005.

[16] C. KANZOW AND H. KLEINMICHEL, *A new class of semismooth Newton-type methods for nonlinear complementarity problems*, Computational Optimization and Applications, vol. 11 pp. 227-251, 1998.

[17] M. S. LOBO, L. VANDENBERGHE, S. BOYD, AND H. LEBRET, *Application of second-order cone programming*, Linear Algebra and its Applications, vol. 284, pp. 193-228, 1998.

[18] Z.-Q. LUO AND P. TSENG, *A new class of merit functions for the nonlinear complementarity problem*, in Complementarity and Variational Problems: State of the Art, edited by M. C. Ferris and J.-S. Pang, SIAM, Philadelphia, pp. 204–225, 1997.

[19] H. D. MITTELMANN, *An independent benchmarking of SDP and SOCP solvers*, Mathematical Programming, vol. 95, pp. 407-430, 2003.

[20] O. L. MANGASARIAN AND M. V. SOLODOV, *Nonlinear complementarity as unconstrained and constrained minimization*, Mathematical Programming, vol. 62, pp. 277-297, 1993.

[21] R. D. C. MONTEIRO AND T. TSUCHIYA, *Polynomial convergence of primal-dual algorithms for the second-order cone programs based on the MZ-family of directions*, Mathematical Programming, vol. 88, pp. 61-83, 2000.

[22] C.-K. SIM, J. SUN, AND D. RALPH, *A note on the Lipschitz continuity of the gradient of the squared norm of the matrix-valued Fischer-Burmeister function*, Mathematical Programming, vol. 107, pp. 547-553, 2006.

[23] P. TSENG, *Merit function for semidefinite complementarity problems*, Mathematical Programming, vol. 83, pp. 159-185, 1998.

[24] T. TSUCHIYA, *A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming*, Optimization Methods and Software, vol. 11, pp. 141–182, 1999.

[25] N. YAMASHITA AND M. FUKUSHIMA, *A new merit function and a descent method for semidefinite complementarity problems*, in *Reformulation – Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods* edited by M. Fukushima and L. Qi, Kluwer Academic Publishers, Boston, 405–420, 1999.