Available online at www.sciencedirect.com
ScienceDirect
Operations Research Letters 36 (2008) 385–392
Operations Research Letters
www.elsevier.com/locate/orl

The SC^1 property of the squared norm of the SOC Fischer–Burmeister function

Jein-Shan Chen[a,*], Defeng Sun[b], Jie Sun[c]

[a] *Department of Mathematics, National Taiwan Normal University, Taipei, 11677, Taiwan*
[b] *Department of Mathematics, National University of Singapore, Singapore 117543, Singapore*
[c] *Department of Decision Sciences, National University of Singapore, Singapore 117592, Singapore*

Received 19 May 2007; accepted 27 August 2007
Available online 23 December 2007

Abstract

We show that the gradient mapping of the squared norm of Fischer–Burmeister function is globally Lipschitz continuous and semismooth, which provides a theoretical basis for solving nonlinear second-order cone complementarity problems via the conjugate gradient method and the semismooth Newton's method.

Keywords: Second-order cone; Merit function; Spectral factorization; Lipschitz continuity; Semismoothness

1. Introduction

A popular approach to solving the nonlinear complementarity problem (NCP) is to reformulate it as the global minimization via a certain merit function over $\mathbb{R}^n$. For this approach to be effective, the choice of the merit function is crucial. A popular choice of the merit function is the squared norm of the Fischer–Burmeister (FB) function $\Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ defined by

$$\Psi(a, b) := \frac{1}{2} \sum_{i=1}^{n} |\phi(a_i, b_i)|^2, \tag{1}$$

for all $a = (a_1, \ldots, a_n)^{\mathrm{T}} \in \mathbb{R}^n$ and $b = (b_1, \ldots, b_n)^{\mathrm{T}} \in \mathbb{R}^n$. The aforementioned Fischer–Burmeister function is denoted by $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ whose ith component function is $\Phi_i(a, b) = \phi(a_i, b_i)$ with $\phi : \mathbb{R}^2 \to \mathbb{R}$ given by

$$\phi(a_i, b_i) = \sqrt{a_i^2 + b_i^2} - a_i - b_i. \tag{2}$$

It is well known that the FB function satisfies

$$\phi(a_i, b_i) = 0 \quad \Longleftrightarrow \quad a_i \geq 0, \quad b_i \geq 0, \quad a_i b_i = 0. \tag{3}$$

It has been shown that ϕ^2 is smooth (continuously differentiable) even though ϕ is not differentiable. This merit function and its analysis were subsequently extended by Tseng [11] to the semidefinite complementarity problem (SDCP) although only differentiability, not continuous differentiability, was established. More recently, the squared norm of the FB function for SDCP was reported in [9] to be a smooth function and its gradient is Lipschitz continuous.

* Corresponding author.
E-mail addresses: jschen@math.ntnu.edu.tw (J.-S. Chen), matsundf@nus.edu.sg (D. Sun), jsun@nus.edu.sg (J. Sun).

0167-6377/$ - see front matter
doi:10.1016/j.orl.2007.08.005

The *second-order cone* (SOC), also called the Lorentz cone, in $\mathbb{R}^n$ is defined as

$$\mathcal{K}^n := \{(x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1} \mid \|x_2\| \le x_1\}, \tag{4}$$

where $\|\cdot\|$ denotes the Euclidean norm. By definition, $\mathcal{K}^1$ is the set of nonnegative reals $\mathbb{R}_+$. The second-order cone complementarity problem (SOCCP) which is to find $x, y \in \mathbb{R}^n$ satisfying

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

where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product and $F, G : \mathbb{R}^n \to \mathbb{R}^n$ are continuous (possibly nonlinear) functions. The FB function for the SOCCP is the vector-valued function $\phi_{\rm FB} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ defined by

$$\phi_{\rm FB}(x, y) := (x^2 + y^2)^{1/2} - (x + y), \tag{6}$$

and the squared norm of the FB function for the SOCCP is $\psi_{\rm FB} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ given by

$$\psi_{\rm FB}(x, y) := \frac{1}{2}\|\phi_{\rm FB}(x, y)\|^2. \tag{7}$$

Note that x^2 and y^2 in (6) mean $x \circ x$ and $y \circ y$, respectively ("$\circ$" is introduced in Section 2); and $x + y$ means 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 $(x^{1/2})^2 = x^{1/2} \circ x^{1/2} = x$. Therefore, the FB function given as in (6) is well defined for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$. Besides, it was shown in [4] that property (3) of ϕ can be extended to $\phi_{\rm FB}$. Thus, $\psi_{\rm FB}$ is a merit function for the SOCCP since the SOCCP can be expressed as an unconstrained minimization problem:

$$\min_{\zeta \in \mathbb{R}^n} \ f(\zeta) := \psi_{\rm FB}(F(\zeta), G(\zeta)). \tag{8}$$

Like in the NCP and the SDCP cases, $\psi_{\rm FB}$ is shown to be smooth, and when ∇F and $-\nabla G$ are column monotone, every stationary point of (8) solves SOCCP; see [2].

The last hurdle to cross in applying (8) to solve (5) is to show that the gradient of $\psi_{\rm FB}$ is sufficiently smooth to warrant the convergence of appropriate computational methods. In particular, we are concerned with the conjugate gradient methods and the semismooth Newton's methods [3]. The former methods generally require the Lipschitz continuity of the gradient ($f \in LC^1$ for short since $f : \mathbb{R}^n \to \mathbb{R}$ is said to be an LC^1 function if it is continuously differentiable and its gradient is locally Lipschitz continuous), while the latter requires that the gradient is semismooth ($f \in SC^1$ for short since f is called an SC^1 function if it is continuously differentiable and its gradient is semismooth), in addition to being Lipschitz continuous.

The main purpose of this paper is to show that the gradient function of $\psi_{\rm FB}$ defined as in (7) is globally Lipschitz continuous and semismooth, which is an important property for superlinear convergence of semismooth Newton methods [8]. It should be noted that this result is not a direct implication from a similar result on function $\Psi(X, Y)$ recently published in [9]. Different analysis is necessary for the proof of Lipschitz continuity.

2. Preliminaries

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

$$x \circ y \ := \ (\langle x, y\rangle, \ y_1 x_2 + x_1 y_2). \tag{9}$$

The identity element under this product is $e := (1, 0, \ldots, 0)^{\rm T} \in \mathbb{R}^n$. We write x^2 to mean $x \circ x$ and write $x + y$ to mean the usual componentwise addition of vectors.

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

$$L_x y := \begin{bmatrix} x_1 & x_2^{\rm T} \\ x_2 & x_1 I \end{bmatrix} y.$$

It can be 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 $x \in {\rm int}(\mathcal{K}^n)$. However, $L_x^{-1} y \ne x^{-1} \circ y$, for some $x \in {\rm int}(\mathcal{K}^n)$ and $y \in \mathbb{R}^n$, i.e., $L_x^{-1} \ne L_{x^{-1}}$.

In addition, any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ can be decomposed as

$$x = \lambda_1 u^{(1)} + \lambda_2 u^{(2)}, \tag{10}$$

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

$$\lambda_i = x_1 + (-1)^i \|x_2\|, \tag{11}$$

$$u^{(i)} = \begin{cases} \frac{1}{2}\left(1, (-1)^i \frac{x_2}{\|x_2\|}\right), & \text{if } x_2 \neq 0, \\ \frac{1}{2}\left(1, (-1)^i w\right), & \text{if } x_2 = 0, \end{cases} \tag{12}$$

for $i = 1, 2$, with w being any vector in $\mathbb{R}^{n-1}$ satisfying $\|w\| = 1$.

The above spectral factorization of x, as well as x^2 and $x^{1/2}$ and the matrix L_x, have various interesting properties (cf. [4]). In particular, we have the following property about the invertibility of L_x.

Property 2.1. *If $x \in \text{int}(\mathcal{K}^n)$, then L_x is invertible with*

$$L_x^{-1} = \frac{1}{\det(x)} \begin{bmatrix} x_1 & -x_2^T \\ -x_2 & \frac{\det(x)}{x_1} I + \frac{1}{x_1} x_2 x_2^T \end{bmatrix}.$$

For any function $f : \mathbb{R} \to \mathbb{R}$, the following vector-valued function associated with $\mathcal{K}^n$ $(n \geq 1)$ was considered in [5,6]

$$f^{\text{soc}}(x) = f(\lambda_1)u^{(1)} + f(\lambda_2)u^{(2)}, \quad x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}. \tag{13}$$

For a recent treatment, see [1,4].

Since we aim to prove that the merit function ψ_{FB} defined as in (7) has a Lipschitz continuous gradient, we now write down the gradient function of ψ_{FB} as below. Let $\phi_{\text{FB}}, \psi_{\text{FB}}$ be given by (6) and (7), respectively. Then, from [2, Prop. 1], we know that $\nabla_x \psi_{\text{FB}}(0, 0) = \nabla_y \psi_{\text{FB}}(0, 0) = 0$. If $(x, y) \neq (0, 0)$ and $x^2 + y^2 \in \text{int}(\mathcal{K}^n)$, then

$$\begin{aligned} \nabla_x \psi_{\text{FB}}(x, y) &= \left(L_x L^{-1}_{(x^2+y^2)^{1/2}} - I\right) \phi_{\text{FB}}(x, y) \\ \nabla_y \psi_{\text{FB}}(x, y) &= \left(L_y L^{-1}_{(x^2+y^2)^{1/2}} - I\right) \phi_{\text{FB}}(x, y). \end{aligned} \tag{14}$$

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

$$\begin{aligned} \nabla_x \psi_{\text{FB}}(x, y) &= \left(\frac{x_1}{\sqrt{x_1^2 + y_1^2}} - 1\right) \phi_{\text{FB}}(x, y), \\ \nabla_y \psi_{\text{FB}}(x, y) &= \left(\frac{y_1}{\sqrt{x_1^2 + y_1^2}} - 1\right) \phi_{\text{FB}}(x, y). \end{aligned} \tag{15}$$

3. Main results

In this section, we will present the proof that the gradient function of ψ_{FB} is Lipschitz continuous.

Lemma 3.1. *Let $\omega : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ be defined by $\omega(x, y) := u(x, y) \circ v(x, y)$, where $u, v : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ are differentiable mappings. Then, ω is differentiable and*

$$\begin{aligned} \nabla_x \omega(x, y) &= \nabla_x u(x, y) L_{v(x,y)} + \nabla_x v(x, y) L_{u(x,y)}, \\ \nabla_y \omega(x, y) &= \nabla_y u(x, y) L_{v(x,y)} + \nabla_y v(x, y) L_{u(x,y)}. \end{aligned} \tag{16}$$

Proof. This is the product rule associated with Jordan product. Its proof is straightforward, so we omit it. □

Lemma 3.2. *For any $x, y \in \mathbb{R}^n$, let $z(x, y) := (x^2 + y^2)^{1/2}$, $F(x, y) := L_x L^{-1}_{z(x,y)}(x + y)$, and $G(x, y) := L_y L^{-1}_{z(x,y)}(x + y)$. Then, we have*

(a) *z is differentiable at $(x, y) \neq (0, 0) \in \mathbb{R}^n \times \mathbb{R}^n$ with $x^2 + y^2 \in \text{int}(\mathcal{K}^n)$. Moreover*

$$\nabla_x z(x, y) = L_x L^{-1}_{z(x,y)}, \qquad \nabla_y z(x, y) = L_y L^{-1}_{z(x,y)}.$$

(b) *F, G are differentiable at $(x, y) \neq (0, 0) \in \mathbb{R}^n \times \mathbb{R}^n$ with $x^2 + y^2 \in \text{int}(\mathcal{K}^n)$. Moreover, $\|\nabla F(x, y)\|$, $\|\nabla G(x, y)\|$ are uniformly bounded at such points.*

Proof. (a) That the function z is differentiable is an immediate consequence of [6]. See also [1, Prop. 4]. Since, $z^2(x,y) = x^2 + y^2$, applying Lemma 3.1 yields

$$2\nabla_x z(x,y) L_{z(x,y)} = 2L_x, \qquad 2\nabla_y z(x,y) L_{z(x,y)} = 2L_y.$$

Hence, the desired results follow.

(b) For symmetry, it is enough to show that F is differentiable at $(x,y) \neq (0,0)$ with $x^2 + y^2 \in \mathrm{int}(\mathcal{K}^n)$ and that $\|\nabla_x F(x,y)\|$, $\|\nabla_y F(x,y)\|$ are uniformly bounded. It is clear that F is differentiable at such points. The key part is to show the uniform boundedness of $\|\nabla_x F(x,y)\|$, $\|\nabla_y F(x,y)\|$. Let λ_1, λ_2 be the spectral values of $x^2 + y^2$, then

$$\begin{aligned} \lambda_1 &:= \|x\|^2 + \|y\|^2 - 2\|x_1 x_2 + y_1 y_2\|, \\ \lambda_2 &:= \|x\|^2 + \|y\|^2 + 2\|x_1 x_2 + y_1 y_2\|. \end{aligned}$$

Thus, $z(x,y) := (x^2 + y^2)^{1/2}$ has the spectral values $\sqrt{\lambda_1}, \sqrt{\lambda_2}$ and

$$z(x,y) = (z_1, z_2) = \left(\frac{\sqrt{\lambda_1} + \sqrt{\lambda_2}}{2}, \frac{\sqrt{\lambda_2} - \sqrt{\lambda_1}}{2} w_2 \right), \tag{17}$$

where $w_2 := \frac{x_1 x_2 + y_1 y_2}{\|x_1 x_2 + y_1 y_2\|}$ if $x_1 x_2 + y_1 y_2 \neq 0$ and otherwise w_2 is any vector in $\mathbb{R}^{n-1}$ satisfying $\|w_2\| = 1$.

Now, let $u := L_{z(x,y)}^{-1}(x+y)$. By applying Property 2.1, we compute u as below.

$$\begin{aligned} u &= L_{z(x,y)}^{-1}(x+y) \\ &= \frac{1}{\det(z(x,y))} \begin{bmatrix} z_1 & -z_2^T \\ -z_2 & \dfrac{\det(z(x,y))}{z_1} I + \dfrac{1}{z_1} z_2 z_2^T \end{bmatrix} \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix} \\ &= \frac{1}{\det(z(x,y))} \begin{bmatrix} (x_1 + y_1) z_1 - (x_2 + y_2)^T z_2 \\ -(x_1 + y_1) z_2 + \dfrac{\det(z)}{z_1} (x_2 + y_2) + \dfrac{(x_2 + y_2)^T z_2}{z_1} z_2 \end{bmatrix} \\ &:= \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}. \end{aligned}$$

We notice that $F(x,y) = L_x L_{z(x,y)}^{-1}(x+y) = L_x u = x \circ u$. Then by applying Lemma 3.1, we obtain

$$\nabla_x F(x,y) = L_u + \nabla_x u(x,y) L_x \quad \text{and} \quad \nabla_y F(x,y) = \nabla_y u(x,y) L_x. \tag{18}$$

To show that $\|\nabla_x F(x,y)\|$ is uniformly bounded, we shall verify that both $\|L_u\|$ and $\|\nabla_x u(x,y) L_x\|$ are uniformly bounded. We prove them as follows.

(i) To see $\|L_u\|$ is uniformly bounded, it is sufficient to argue that $|u_1|$, $\|u_2\|$ are both uniformly bounded. First, we argue that $|u_1|$ is uniformly bounded. From the above expression of u, we have

$$u_1 = \frac{1}{\det(z(x,y))}(x_1 z_1 - x_2^T z_2) + \frac{1}{\det(z(x,y))}(y_1 z_1 - y_2^T z_2).$$

Following the similar arguments as in [2, Lemma 4] yields

$$\begin{aligned} u_1 &= \frac{1}{\det(z(x,y))}(x_1 z_1 - x_2^T z_2) + \frac{1}{\det(z(x,y))}(y_1 z_1 - y_2^T z_2) \\ &= \left[O(1) + \frac{(x_1 - x_2^T w_2)}{2\sqrt{\lambda_1}} \right] + \left[O(1) + \frac{(y_1 - y_2^T w_2)}{2\sqrt{\lambda_1}} \right], \end{aligned}$$

where $O(1)$ denotes terms that are uniformly bounded with bound independent of (x,y). Moreover, by [2, Lemma 3], if $x_1 x_2 + y_1 y_2 \neq 0$ then $|x_1 - x_2^T w_2| \leq \|x_2 - x_1 w_2\| \leq \sqrt{\lambda_1}$. If $x_1 x_2 + y_1 y_2 = 0$ then $\lambda_1 = \|x\|^2 + \|y\|^2$ so that by choosing w_2 to further satisfy $x_2^T w_2 = 0$ we obtain $|x_1 - x_2^T w_2| \leq \|x_2 - x_1 w_2\| \leq \|x\| \leq \sqrt{\lambda_1}$. Similarly, it can be verified that $|y_1 - y_2^T w_2| \leq \sqrt{\lambda_1}$. Thus, $|u_1|$ is uniformly bounded.

Secondly, we argue that $\|u_2\|$ is also uniformly bounded. Again, using the expression of u and following the similar arguments as in [2, Lemma 4], we obtain

$$
\begin{aligned}
u_2 &= \frac{1}{\det(z(x,y))}\left[-x_1 z_2 + \frac{\det(z(x,y))}{z_1}x_2 + \frac{x_2^{\mathrm{T}} z_2}{z_1} z_2\right] \\
&\quad + \frac{1}{\det(z(x,y))}\left[-y_1 z_2 + \frac{\det(z(x,y))}{z_1}y_2 + \frac{y_2^{\mathrm{T}} z_2}{z_1} z_2\right] \\
&= \left[O(1) - \frac{x_1 w_2}{2\sqrt{\lambda_1}} + \frac{\frac{\sqrt{\lambda_2}}{\sqrt{\lambda_1}}(x_2^{\mathrm{T}} w_2)}{2(\sqrt{\lambda_1}+\sqrt{\lambda_2})} w_2\right] + \left[O(1) - \frac{y_1 w_2}{2\sqrt{\lambda_1}} + \frac{\frac{\sqrt{\lambda_2}}{\sqrt{\lambda_1}}(y_2^{\mathrm{T}} w_2)}{2(\sqrt{\lambda_1}+\sqrt{\lambda_2})} w_2\right] \\
&= \left[O(1) - \frac{x_1 w_2}{2(\sqrt{\lambda_1}+\sqrt{\lambda_2})} - \frac{\sqrt{\lambda_2}(x_1 - x_2^{\mathrm{T}} w_2)}{2(\sqrt{\lambda_1}+\sqrt{\lambda_2})\sqrt{\lambda_1}} w_2\right] \\
&\quad + \left[O(1) - \frac{y_1 w_2}{2(\sqrt{\lambda_1}+\sqrt{\lambda_2})} - \frac{\sqrt{\lambda_2}(y_1 - y_2^{\mathrm{T}} w_2)}{2(\sqrt{\lambda_1}+\sqrt{\lambda_2})\sqrt{\lambda_1}} w_2\right].
\end{aligned}
$$

Using the same explanations as above for u_1 yields that each term is uniformly bounded. Thus, $\|u_2\|$ is uniformly bounded. This together with $|u_1|$ being uniformly bounded implies that $\|\nabla_x F(x,y)\| = \|L_u\| = \left\|\begin{bmatrix} u_1 & u_2^{\mathrm{T}} \\ u_2 & u_1 I\end{bmatrix}\right\|$ is also uniformly bounded.

(ii) Now, it comes to show that $\|\nabla_x u(x,y) L_x\|$ is uniformly bounded. From the definition of $u := L_{z(x,y)}^{-1}(x+y)$, we know that $z(x,y) \circ u = x + y$. Applying Lemma 3.1 gives

$$\nabla_x z(x,y) L_u + \nabla_x u(x,y) L_{z(x,y)} = I,$$

which leads to

$$
\begin{aligned}
&\nabla_x u(x,y) L_{z(x,y)} = I - \nabla_x z(x,y) L_u = I - (L_x L_{z(x,y)}^{-1}) L_u \\
&\Rightarrow \nabla_x u(x,y) = \left(I - L_x L_{z(x,y)}^{-1} L_u\right) L_{z(x,y)}^{-1} \\
&\Rightarrow \nabla_x u(x,y) L_x = \left(I - L_x L_{z(x,y)}^{-1} L_u\right) L_{z(x,y)}^{-1} L_x \\
&\Rightarrow \nabla_x u(x,y) L_x = L_{z(x,y)}^{-1} L_x - L_x L_{z(x,y)}^{-1} L_u L_{z(x,y)}^{-1} L_x \\
&\Rightarrow \nabla_x u(x,y) L_x = (L_x L_{z(x,y)}^{-1})^{\mathrm{T}} - (L_x L_{z(x,y)}^{-1}) L_u (L_x L_{z(x,y)}^{-1})^{\mathrm{T}}.
\end{aligned}
$$

Therefore,

$$\|\nabla_x u(x,y) L_x\| \le \|(L_x L_{z(x,y)}^{-1})^{\mathrm{T}}\| + \|L_x L_{z(x,y)}^{-1}\| \cdot \|L_u\| \cdot \|(L_x L_{z(x,y)}^{-1})^{\mathrm{T}}\|.$$

By [2, Lemma 4], $\|L_x L_{z(x,y)}^{-1}\|$ is uniformly bounded, so is $\|(L_x L_{z(x,y)}^{-1})^{\mathrm{T}}\|$. This together with $\|L_u\|$ being uniformly bounded shown as above yields that $\|\nabla_x u(x,y) L_x\|$ is uniformly bounded.

From (i) and (ii), we conclude that $\|\nabla_x F(x,y)\|$ is uniformly bounded. Similar arguments apply to $\|\nabla_y F(x,y)\|$; and hence, $\|\nabla F(x,y)\|$ is uniformly bounded. Thus, we complete the proof. □

Lemma 3.3. *Let ψ_{FB} be defined as (7). Then, $\nabla\psi_{\mathrm{FB}}$ is continuously differentiable everywhere except for $(x,y) = (0,0)$. Moreover, $\|\nabla^2\psi_{\mathrm{FB}}(x,y)\|$ is uniformly bounded for all $(x,y) \neq (0,0)$.*

Proof. For any $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$, let $z := (x^2 + y^2)^{1/2}$. We prove this lemma by considering the following two cases.

(i) Consider all points $(x,y) \neq (0,0)$ with $x^2 + y^2 \in \mathrm{int}(\mathcal{K}^n)$. Since

$$
\begin{aligned}
\nabla_x \psi_{\mathrm{FB}}(x,y) &= \left(L_x L_z^{-1} - I\right)\phi_{\mathrm{FB}}(x,y) = x - L_x L_z^{-1}(x+y) - \phi_{\mathrm{FB}}(x,y), \\
\nabla_y \psi_{\mathrm{FB}}(x,y) &= \left(L_y L_z^{-1} - I\right)\phi_{\mathrm{FB}}(x,y) = y - L_y L_z^{-1}(x+y) - \phi_{\mathrm{FB}}(x,y),
\end{aligned}
$$

we can compute $\nabla^2\psi_{\mathrm{FB}}(x,y)$ as follows:

$$
\begin{aligned}
\nabla_{xx}^2 \psi_{\mathrm{FB}}(x,y) &= I - \nabla_x\left(L_x L_z^{-1}(x+y)\right) - \left(L_x L_z^{-1} - I\right), \\
\nabla_{xy}^2 \psi_{\mathrm{FB}}(x,y) &= -\nabla_y\left(L_x L_z^{-1}(x+y)\right) - \left(L_y L_z^{-1} - I\right), \\
\nabla_{yx}^2 \psi_{\mathrm{FB}}(x,y) &= -\nabla_x\left(L_y L_z^{-1}(x+y)\right) - \left(L_x L_z^{-1} - I\right), \\
\nabla_{yy}^2 \psi_{\mathrm{FB}}(x,y) &= I - \nabla_y\left(L_y L_z^{-1}(x+y)\right) - \left(L_y L_z^{-1} - I\right).
\end{aligned}
\tag{19}
$$

The continuity of $\nabla^2\psi_{\rm FB}$ at (x,y) thus follows. It is easy to see that $\|L_xL_z^{-1}\|$, $\|L_yL_z^{-1}\|$ are uniformly bounded by [2, Lemma 4] ($\|\cdot\|$ and $\|\cdot\|_F$ are equivalent in $\mathbb{R}^{n\times n}$). Let $F(x,y) := L_xL_z^{-1}(x+y)$ and $G(x,y) := L_yL_z^{-1}(x+y)$. By Lemma 3.2, we know that $\left\|\nabla_x\left(L_xL_z^{-1}(x+y)\right)\right\| = \|\nabla_x F(x,y)\|$ is uniformly bounded. Likewise, we have that $\left\|\nabla_y\left(L_xL_z^{-1}(x+y)\right)\right\|$, $\left\|\nabla_x\left(L_yL_z^{-1}(x+y)\right)\right\|$, $\left\|\nabla_y\left(L_yL_z^{-1}(x+y)\right)\right\|$ are all uniformly bounded. Thus, we can conclude that $\|\nabla^2_{xx}\psi_{\rm FB}(x,y)\|$, $\|\nabla^2_{xy}\psi_{\rm FB}(x,y)\|$, $\|\nabla^2_{yx}\psi_{\rm FB}(x,y)\|$, $\|\nabla^2_{yy}\psi_{\rm FB}(x,y)\|$ are all uniformly bounded which implies that $\|\nabla^2\psi_{\rm FB}(x,y)\|$ is also uniformly bounded.

(ii) Consider all points $(x,y)\neq(0,0)$ with $x^2+y^2\notin \mathrm{int}(\mathcal{K}^n)$. Since

$$\nabla_x\psi_{\rm FB}(x,y) = \left(\frac{x_1}{\sqrt{x_1^2+y_1^2}}-1\right)\phi_{\rm FB}(x,y) = x - \frac{x_1}{\sqrt{x_1^2+y_1^2}}(x+y)-\phi_{\rm FB}(x,y),$$

$$\nabla_y\psi_{\rm FB}(x,y) = \left(\frac{y_1}{\sqrt{x_1^2+y_1^2}}-1\right)\phi_{\rm FB}(x,y) = y - \frac{y_1}{\sqrt{x_1^2+y_1^2}}(x+y)-\phi_{\rm FB}(x,y),$$

we can compute $\nabla^2\psi_{\rm FB}(x,y)$ as follows:

$$\nabla^2_{xx}\psi_{\rm FB}(x,y) = I - \left(\frac{x_1}{\sqrt{x_1^2+y_1^2}}I + \frac{x_1y_1^2+y_1^3}{(x_1^2+y_1^2)^{3/2}}\begin{bmatrix}1 & 0\\ 0 & \mathbf{0}\end{bmatrix}\right) - \left(\frac{x_1}{\sqrt{x_1^2+y_1^2}}-1\right)I, \tag{20}$$

$$\nabla^2_{xy}\psi_{\rm FB}(x,y) = -\left(\frac{x_1}{\sqrt{x_1^2+y_1^2}}I - \frac{x_1^2y_1+x_1y_1^2}{(x_1^2+y_1^2)^{3/2}}\begin{bmatrix}1 & 0\\ 0 & \mathbf{0}\end{bmatrix}\right) - \left(\frac{y_1}{\sqrt{x_1^2+y_1^2}}-1\right)I,$$

$$\nabla^2_{yx}\psi_{\rm FB}(x,y) = -\left(\frac{y_1}{\sqrt{x_1^2+y_1^2}}I - \frac{x_1^2y_1+x_1y_1^2}{(x_1^2+y_1^2)^{3/2}}\begin{bmatrix}1 & 0\\ 0 & \mathbf{0}\end{bmatrix}\right) - \left(\frac{x_1}{\sqrt{x_1^2+y_1^2}}-1\right)I,$$

$$\nabla^2_{yy}\psi_{\rm FB}(x,y) = I - \left(\frac{y_1}{\sqrt{x_1^2+y_1^2}}I + \frac{x_1^3+x_1^2y_1}{(x_1^2+y_1^2)^{3/2}}\begin{bmatrix}1 & 0\\ 0 & \mathbf{0}\end{bmatrix}\right) - \left(\frac{y_1}{\sqrt{x_1^2+y_1^2}}-1\right)I,$$

where $\mathbf{0}$ denotes the $(n-1)\times(n-1)$ zero matrix.

Now we provide a sketch proof to verify that $\nabla_{xx}\psi_{\rm FB}$ is continuous. Let $(a,b)\neq(0,0)$ and $a^2+b^2\notin\mathrm{int}(\mathcal{K}^n)$. We want to prove that

$$\nabla_{xx}\psi_{\rm FB}(x,y)\to\nabla_{xx}\psi_{\rm FB}(a,b), \quad \text{as} \quad (x,y)\to(a,b). \tag{21}$$

Due to the neighborhood of such (a,b), we have to consider two subcases: (1) $(x,y)\neq(0,0)$ with $x^2+y^2\in\mathrm{int}(\mathcal{K}^n)$ and (2) $(x,y)\neq(0,0)$ with $x^2+y^2\notin\mathrm{int}(\mathcal{K}^n)$. It is clear that (21) holds in subcase (2) because the formula given in (20) is continuous. In subcase (1), we have

$$\begin{aligned}\nabla_{xx}\psi_{\rm FB}(x,y) &= I - \nabla_x\left(L_xL_z^{-1}(x+y)\right) - \left(L_xL_z^{-1}-I\right)\\ &= I - \left[L_u + \left(L_xL_z^{-1}\right)^{\rm T} - \left(L_xL_z^{-1}\right)(L_u)\left(L_xL_z^{-1}\right)^{\rm T}\right] - \left(L_xL_z^{-1}-I\right).\end{aligned} \tag{22}$$

In view of (19), (20) and (22), it suffices to show the following three statements for (21) to be held in this subcase (1):

(a) $L_xL_z^{-1}\to\frac{a_1}{\sqrt{a_1^2+b_1^2}}I$, as $(x,y)\to(a,b)$.

(b) $L_u\to\frac{a_1+b_1}{\sqrt{a_1^2+b_1^2}}I$, as $(x,y)\to(a,b)$.

(c) $L_u-(L_xL_z^{-1})(L_u)(L_xL_z^{-1})^{\rm T}\to\frac{a_1^2(a_1+b_1)}{(a_1^2+b_1^2)^{3/2}}I$, as $(x,y)\to(a,b)$.

First, we know from [2, Prop. 2] that there holds

$$L_xL_z^{-1}(x+y)\to\frac{a_1}{\sqrt{a_1^2+b_1^2}}(a+b) \quad \text{as } (x,y)\to(a,b),$$

which implies that $L_xL_z^{-1}\to\frac{a_1}{\sqrt{a_1^2+b_1^2}}I$, as $(x,y)\to(a,b)$ since both $(x+y)$ and $L_xL_z^{-1}$ are continuous and $(x+y)\to(a+b)$ when $(x,y)\to(a,b)$. Secondly, if we look into the entries of L_u and compare them with the entries of $L_xL_z^{-1}$ (see [2, eq. (27)]), then it is clear that $L_u \to \frac{a_1+b_1}{\sqrt{a_1^2+b_1^2}} I$, as $(x, y) \to (a, b)$. Finally, part(c) follows immediately from part (a) and (b). Thus, we complete the verifications of (21). The other cases can be argued similarly for $\nabla_{xy}\psi_{\rm FB}$, $\nabla_{yx}\psi_{\rm FB}$, and $\nabla_{yy}\psi_{\rm FB}$. In addition, it is also clear that each term in the above expressions (20) is uniformly bounded. Thus, we obtain that $\nabla^2\psi_{\rm FB}$ is continuously differentiable near (x, y) and $\|\nabla^2\psi_{\rm FB}(x, y)\|$ is uniformly bounded. □

Theorem 3.1. *Let $\psi_{\rm FB}$ be defined as (7). Then, $\nabla\psi_{\rm FB}$ is globally Lipschitz continuous, i.e., there exists a constant C such that for all $(x, y),\ (a, b) \in \mathbb{R}^n \times \mathbb{R}^n$,*

$$\begin{aligned} &\|\nabla_x\psi_{\rm FB}(x, y) - \nabla_x\psi_{\rm FB}(a, b)\| \le C\|(x, y) - (a, b)\|, \\ &\|\nabla_y\psi_{\rm FB}(x, y) - \nabla_y\psi_{\rm FB}(a, b)\| \le C\|(x, y) - (a, b)\| \end{aligned} \tag{23}$$

and is semismooth everywhere.

Proof. Owing to symmetry, we only need to show that the first part of (23) holds. For any $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, let $z := (x^2 + y^2)^{1/2}$.

(i) First, we prove that $\nabla_x\psi_{\rm FB}$ is Lipschitz continuous at (0, 0). We have to discuss three subcases for completing the proof of this part.

If $(x, y) = (0, 0)$, it is obvious that (23) is satisfied.

If $(x, y) \ne (0, 0)$ with $x^2 + y^2 \in \mathrm{int}(\mathcal{K}^n)$, then

$$\|\nabla_x\psi_{\rm FB}(x, y) - \nabla_x\psi_{\rm FB}(0, 0)\| = \|\nabla_x\psi_{\rm FB}(x, y)\| = \|x - L_x L_z^{-1}(x + y) - \phi_{\rm FB}(x, y)\|.$$

It is already known that x and $\phi_{\rm FB}(x, y)$ are Lipschitz continuous (see [10, Cor. 3.3]). In addition, Theorem 3.2.4 of [7, pp. 70] says that the uniform boundedness of $\nabla\left(L_x L_z^{-1}(x + y)\right)$ (by Lemma 3.2) yields the Lipschitz continuity. Thus, (23) is satisfied for this subcase. If $(x, y) \ne (0, 0)$ with $x^2 + y^2 \notin \mathrm{int}(\mathcal{K}^n)$, then

$$\|\nabla_x\psi_{\rm FB}(x, y) - \nabla_x\psi_{\rm FB}(0, 0)\| = \|\nabla_x\psi_{\rm FB}(x, y)\| = \left\| x - \frac{x_1}{\sqrt{x_1^2 + y_1^2}}(x + y) - \phi_{\rm FB}(x, y) \right\|.$$

Since $\left|\frac{x_1}{\sqrt{x_1^2+y_1^2}}\right| \le 1$ and both $(x + y)$, $\phi_{\rm FB}(x, y)$ are known Lipschitz continuous, the desired result follows.

(ii) Secondly, we prove that $\nabla_x\psi_{\rm FB}$ is Lipschitz continuous at $(a, b) \ne (0, 0)$. Let $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, we wish to show that (23) is satisfied. In fact, if the line segment $[(a, b), (x, y)]$ does not contain the origin, then we can write

$$\|\nabla_x\psi_{\rm FB}(x, y) - \nabla_x\psi_{\rm FB}(a, b)\| \le \left\| \int_0^1 \nabla^2\psi_{\rm FB}[(a, b) + t((x, y) - (a, b))]\mathrm{d}t \right\| \le C\|(x, y) - (a, b)\|,$$

where the first inequality is from the Mean Value Theorem (see [7, Theorem 3.2.3]), and the second inequality is by Lemma 3.3. On the other hand, if the line segment $[(a, b), (x, y)]$ contains the origin, we can construct a sequence $\{(x^k, y^k)\}$ converging to (x, y) but for each k, the line segment $[(a, b), (x^k, y^k)]$ does not contain the origin and apply the above inequalities to get

$$\|\nabla_x\psi_{\rm FB}(x^k, y^k) - \nabla_x\psi_{\rm FB}(a, b)\| \le C\|(x^k, y^k) - (a, b)\|,$$

which, by the continuity, implies

$$\|\nabla_x\psi_{\rm FB}(x, y) - \nabla_x\psi_{\rm FB}(a, b)\| \le C\|(x, y) - (a, b)\|.$$

Thus, (23) is satisfied.

To complete the proof of this theorem, we only need to show that $\nabla\psi_{\rm FB}$ is semismooth at the origin as, by Lemma 3.3, $\nabla\psi_{\rm FB}$ is continuously differentiable near any $(0, 0) \ne (x, y) \in \mathbb{R}^n \times \mathbb{R}^n$. From (14) and (15), we know that for any $t \in \mathbb{R}_+$ and $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ we have

$$\nabla\psi_{\rm FB}(tx, ty) = t\nabla\psi_{\rm FB}(x, y).$$

Thus, $\nabla\psi_{\rm FB}$ is directionally differentiable at the origin and for any $(0, 0) \ne (x, y) \in \mathbb{R}^n \times \mathbb{R}^n$

$$\nabla^2\psi_{\rm FB}(x, y)(x, y) = (\nabla\psi_{\rm FB})'((x, y); (x, y)) = \nabla\psi_{\rm FB}(x, y).$$

This means that for any $(0, 0) \ne (x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ converging to (0, 0),

$$\nabla\psi_{\rm FB}(x, y) - \nabla\psi_{\rm FB}(0, 0) - \nabla^2\psi_{\rm FB}(x, y)(x, y) = \nabla\psi_{\rm FB}(x, y) - 0 - \nabla\psi_{\rm FB}(x, y) = 0,$$

which, together with the Lipschitz continuity of $\nabla\psi_{\rm FB}$ and the directional differentiability of $\nabla\psi_{\rm FB}$ at the origin ($\nabla\psi_{\rm FB}$ is, however, not differentiable at the origin), shows that $\nabla\psi_{\rm FB}(x, y)$ is (strongly) semismooth at the origin. The proof is completed. □

From Theorem 3.1, we immediately obtain that the function $\psi_{\rm FB}$ defined as in (7) is an SC^1 function as well as an LC^1 function.

Acknowledgements

Jein-Shan Chen is a member of Mathematics Division, National Center for Theoretical Sciences, Taipei Office and his research is partially supported by the National Science Council of Taiwan. The research of Defeng Sun is partially supported by the Academic Research Fund under Grant No. R-146-000-104-112 of the National University of Singapore. The research of Jie Sun is partially supported by Singapore-MIT Alliance.

References

[1] J.-S. Chen, X. Chen, P. Tseng, Analysis of nonsmooth vector-valued functions associated with second-order cones, Mathematical Programming 101 (2004) 95–117.
[2] J.-S. Chen, P. Tseng, An unconstrained smooth minimization reformulation of the second-order cone complementarity problem, Mathematical Programming 104 (2005) 293–327.
[3] X.-D. Chen, D. Sun, J. Sun, Complementarity functions and numerical experiments for second-order cone complementarity problems, Computational Optimization and Applications 25 (2003) 39–56.
[4] M. Fukushima, Z.-Q. Luo, P. Tseng, Smoothing functions for second-order cone complementarity problems, SIAM Journal on Optimization 12 (2002) 436–460.
[5] M. Koecher, The Minnesota Notes on Jordan Algebras and their Applications Edited and annotated by A. Brieg, S.Walcher, Springer, Berlin, 1999.
[6] A. Korányi, Monotone functions on formally real Jordan algebras, Mathematische Annalen 269 (1984) 73–76.
[7] J. Ortega, W. Rheinboldt, Iterative solution of nonlinear equations in several variables, SIAM Classics in Applied Mathematics (2000).
[8] L. Qi, J. Sun, A nonsmooth version of Newton's method, Mathematical Programming 58 (1993) 353–367.
[9] C.-K. Sim, J. Sun, D. Ralph, A note on the Lipschitz continuity of the gradient of the squared norm of the matrix-valued Fischer–Burmeister function, Mathematical Programming 107 (2006) 547–553.
[10] D. Sun, J. Sun, Strong semismoothness of the Fischer–Burmeister SDC and SOC complementarity functions, Mathematical Programming 103 (2005) 575–581.
[11] P. Tseng, Merit function for semidefinite complementarity problems, Mathematical Programming 83 (1998) 159–185.