

MERIT FUNCTIONS FOR MATRIX CONE COMPLEMENTARITY PROBLEMS[†]

LI WANG*, YONG-JIN LIU AND YONG JIANG

ABSTRACT. The merit function arises from the development of the solution methods for the complementarity problems defined over the cone of non-negative real vectors and has been well extended to the complementarity problems defined over the symmetric cones. In this paper, we focus on the extension of the merit functions including the gap function, the regularized gap function, the implicit Lagrangian and others to the complementarity problems defined over the nonsymmetric matrix cone. These theoretical results of this paper suggest new solution methods based on unconstrained and/or simply constrained methods to solve the matrix cone complementarity problems (MCCP).

AMS Mathematics Subject Classification : 65K05, 90C33.

Key words and phrases : Matrix cone complementarity problem, merit function, gap function, regularized gap function, implicit Lagrangian.

1. Introduction

Matrix optimization problems have recently found many important applications, for example, in nuclear norm relaxations of affine rank minimization problems [4], see [8] for more examples. The recent work by Ding, Sun and Toh [8] has defined a class of linear conic programming (which is called matrix cone programming or MCP) involving the epigraphs of five commonly used matrix norms and studied many important properties of the corresponding metric projection. The matrix cone complementarity problem (MCCP), serving as the counterpart of MCP, deserves to be studied.

We first formally give the description of matrix cone complementarity problem (MCCP). Let $\mathfrak{R}^{m \times n}$ be the linear space of all $m \times n$ real matrices equipped

Received May 9, 2012. Revised November 27, 2012. Accepted December 10, 2012.
*Corresponding author. [†]This work was supported by the National Young Natural Science Foundation of China under project grant No. 11001180, the Program for Liaoning Excellent Talents in Universities, and the Optional Funds for Young Teachers of Shenyang Aerospace University under project No. 201010Y.

© 2013 Korean SIGCAM and KSCAM.

with the inner product $\langle X, Y \rangle = \text{Tr}(X^T Y)$ for X and Y in $\Re^{m \times n}$, where “Tr” denotes the trace, i.e., the sum of the diagonal entries of a square matrix. For each $X \in \Re^{m \times n}$, let $\|X\|_*$ denote the nuclear norm of X , i.e., the sum of the singular values of X , and $\|X\|_2$ denote the spectral norm of X , i.e., the largest singular value of X . Let \mathcal{H} be the finite dimensional Hilbert space of $\Re \times \Re^{m \times n}$ with the natural inner product being given by

$$\langle (t, X), (\tau, Y) \rangle := t\tau + \langle X, Y \rangle = t\tau + \text{Tr}(X^T Y).$$

Then, we define two convex cones in \mathcal{H} as follows

$$\mathcal{K}_*^{m,n} := \{(\tau, Y) \in \Re \times \Re^{m \times n} \mid \tau \geq \|Y\|_*\}$$

and

$$\mathcal{K}_2^{m,n} := \{(t, X) \in \Re \times \Re^{m \times n} \mid t \geq \|X\|_2\}.$$

It is well known (see, e.g. [8]) that one of these two cones is the dual cone of another, where for any subset C of \mathcal{H} , the dual cone of C , denoted by C^* , is defined as

$$C^* = \{a \in \mathcal{H} \mid \langle a, b \rangle \geq 0, \forall b \in C\}. \quad (1)$$

Without any confusion, in the remainder of this paper, we simply use \mathcal{K}_* and \mathcal{K}_2 to represent $\mathcal{K}_*^{m,n}$ and $\mathcal{K}_2^{m,n}$, respectively.

Let Ω denote a subspace of the space \mathcal{H} . Then, the matrix cone complementarity problem (MCCP) considered in this paper is to find, for given mappings $F : \Omega \rightarrow \mathcal{H}$ and $G : \Omega \rightarrow \mathcal{H}$, an $x \in \Omega$ satisfying

$$F(x) \in \mathcal{K}_2, \quad G(x) \in \mathcal{K}_*, \quad \langle F(x), G(x) \rangle = 0. \quad (2)$$

A function $f : \Omega \rightarrow [0, \infty)$ is said to be a merit function on a set $\Omega \subset \mathcal{H}$ (typically $\Omega = \mathcal{H}$ or $\Omega = G^{-1}(\mathcal{K}_*)$), provided that x satisfies (2) if and only if $f(x) = 0$. Using a merit function, we can express MCCP as the following unconstrained minimization problem:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & x \in \Omega \end{aligned}$$

and apply a feasible descent method, such as conjugate gradient methods and quasi-Newton methods, to solve this minimization problem. In this paper, we shall study merit functions for MCCP.

During the last thirty years, researchers have proposed various methods such as the interior-point methods, the semismooth Newton methods, the smoothing Newton methods, and the merit-function-based methods for solving linear/nonlinear complementarity problems (LCP/NCP)(see [5, 9, 14, 22, 23, 31] and references therein), second-order cone complementarity problems (SOCCP) (see [6, 12, 16, 21, 27, 29, 30, 34, 35] and references therein), semi-definite complementarity problems (SDCP) (see [7, 24, 30, 33, 34, 36] and references therein) and, more generally, symmetric cone complementarity problems (SCCP) (see [10, 13, 17, 18, 20, 28, 32, 38] and references therein). However, to the best of our knowledge, the solution methods for MCCP have not been well investigated due to the fact that MCCP is not included in the setting of symmetric

cone complementarity problems. Motivated by the effectiveness of the merit-function-based methods for solving LCP/NCP, our goal is to apply them to solve M CCP. For the merit-function-based methods to be effective, the choice of merit functions is crucial. This main objective of this paper is just to study various choices of merit functions for M CCP. We shall extend five kinds of merit functions which are the projection residual function, the gap function, the regularized gap function, the implicit Lagrangian and a function of Luo and Tseng to M CCP. Moreover, for each of the above five choices of merit functions, we have derived conditions for the merit function to be convex and/or differentiable, and for the stationary point of the merit function to be a solution of M CCP. These properties make it possible to apply the merit-function-based methods to solve M CCP.

Five choices for a merit function are outlined as follows. The earliest choice is the gap function

$$f(x) := \max_{\zeta \in \mathcal{K}_*} \{ \langle F(x), G(x) - \zeta \rangle \} \tag{3}$$

proposed by Auslender [3] and Hearn [14] in the context of LCP/NCP, which is a merit function on $G^{-1}(\mathcal{K}_*)$ (see Section 4).

There is a “dual” version of this gap function, given by

$$f(x) := \max_{\zeta \in \mathcal{K}_*} \{ \langle F(G^{-1}(\zeta)), G(x) - \zeta \rangle \}, \tag{4}$$

which is a merit function on $G^{-1}(\mathcal{K}_*)$ provided that F and G are relatively pseudomonotone (see Section 2 for the definition) on $G^{-1}(\mathcal{K}_*)$ and F is continuous on $G^{-1}(\mathcal{K}_*)$, and G^{-1} is defined and continuous on \mathcal{K}_* (see Section 4).

A second choice is the regularized gap function, parameterized by a scalar $\alpha > 0$,

$$f_\alpha(x) := \max_{\zeta \in \mathcal{K}_*} \{ \langle F(x), G(x) - \zeta \rangle - \frac{1}{2\alpha} \|G(x) - \zeta\|^2 \} \tag{5}$$

proposed independently by Fukushima [11] and Auchmuty [2], which is a merit function on $G^{-1}(\mathcal{K}_*)$ (see Section 5).

A third choice is the implicit Lagrangian function, parameterized by a scalar $\alpha > 1$,

$$f_\alpha(x) := \max_{\zeta \in \mathcal{K}_*, \xi \in \mathcal{K}_2} \{ \langle F(x), G(x) - \zeta \rangle - \langle \xi, G(x) \rangle - \frac{1}{2\alpha} (\|F(x) - \xi\|^2 + \|G(x) - \zeta\|^2) \} \tag{6}$$

proposed by Mangasarian and Solodov [23] in the context of NCP, which is a merit function on Ω (see Section 6).

A fourth choice is the function

$$f(x) := \|F(x) - \Pi_{\mathcal{K}_2}(F(x) - G(x))\|^2, \tag{7}$$

which is a merit function on Ω (see Section 3). Here $\Pi_{\mathcal{K}_2}(\cdot)$ denotes the metric projection onto \mathcal{K}_2 , i.e.,

$$\Pi_{\mathcal{K}_2}(x) = \operatorname{argmin}_{\xi \in \mathcal{K}_2} \|x - \xi\|.$$

A fifth choice is the following function

$$f(x) := \psi_0(\langle F(x), G(x) \rangle) + \psi(F(x), G(x)), \quad (8)$$

where $\psi_0 : \Re \rightarrow [0, \infty)$ satisfies $\psi_0(t) = 0$ if and only if $t \leq 0$ and $\psi : \mathcal{H} \times \mathcal{H} \rightarrow [0, \infty)$ satisfies

$$\psi(x, y) = 0, \quad \langle x, y \rangle \leq 0 \quad \text{if and only if} \quad x \in \mathcal{K}_2, y \in -\mathcal{K}_*, \langle x, y \rangle = 0. \quad (9)$$

This function, developed by Luo and Tseng [22] in the context of NCP and extended by Tseng [36] to SDCP, is a merit function on Ω (see Section 7).

2. Preliminaries

In this section, we review some preliminary results related to the Moreau-Yosida regularization of a convex function and the metric projection onto a closed convex cone.

Definition 2.1 ([26, 37]). Let \mathbb{H} be a Hilbert space and $g : \mathbb{H} \rightarrow (-\infty, +\infty]$ be a closed convex, proper and lower semicontinuous function. The proximal mapping P_g and the Moreau-Yosida regularization ψ_g are defined by

$$P_g(x) = \operatorname{argmin}_{y \in \mathbb{H}} \left\{ g(y) + \frac{1}{2} \|x - y\|^2 \right\}$$

and

$$\psi_g(x) = \min_{y \in \mathbb{H}} \left\{ g(y) + \frac{1}{2} \|x - y\|^2 \right\},$$

respectively.

We first give some well-known properties (see, e.g., [15, 25]) of P_g and ψ_g . For additional properties, see, e.g., [15, 19].

Lemma 2.2. *Let $g : \mathbb{H} \rightarrow \Re$ be a closed convex, proper and lower semicontinuous function, ψ_g be the Moreau-Yosida regularization of g , and P_g be the associated proximal mapping. Then the following hold:*

- (i): *The Moreau-Yosida regularization ψ_g is continuously differentiable, and furthermore, it holds that*

$$\nabla \psi_g(x) = x - P_g(x).$$

- (ii): *For any $x \in \mathbb{H}$ has the decomposition*

$$x = P_g(x) + P_{g^*}(x),$$

where g^ is the conjugate of the function g .*

We next give some useful properties of the metric projection onto a closed convex cone, which plays an important role in our subsequent analysis. Let \mathcal{Z} be a finite dimensional real Euclidean space equipped with an inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\| \cdot \|$. Let C be a nonempty closed convex set in \mathcal{Z} . For any $z \in \mathcal{Z}$, let $\Pi_C(z)$ denote the metric projection of z onto C , which is the unique optimal solution to following convex optimization problem:

$$\begin{aligned} \min \quad & \frac{1}{2} \|y - z\|^2 \\ \text{s.t.} \quad & y \in C. \end{aligned} \tag{10}$$

It is well known [39] that $\Pi_C(z)$ is globally Lipschitz continuous with modulus 1. When C is a closed convex cone, we summarize the following useful properties of the metric projection without proofs.

Lemma 2.3. *Let $C \subseteq \mathcal{Z}$ be a nonempty closed convex cone. Denote by $C^\circ = -C^*$ the polar of C . Then the following hold:*

(i): *The point $\bar{y} \in C$ solves (10) if and only if*

$$\begin{cases} \langle z - \bar{y}, \bar{y} \rangle = 0, \\ \langle z - \bar{y}, d \rangle \leq 0, \quad \forall d \in C. \end{cases} \tag{11}$$

(ii): *For any $z \in \mathcal{Z}$, one has*

$$z = \Pi_C(z) + \Pi_{C^\circ}(z) \tag{12}$$

and

$$\Pi_C(z) = -\Pi_{C^\circ}(-z) = -\Pi_{-C^*}(-z). \tag{13}$$

(iii): *For any $x, y \in \mathcal{Z}$, one has that $x = \Pi_C(x + y)$ or $y = \Pi_{C^\circ}(x + y)$ if and only if*

$$\langle x, y \rangle = 0, \quad x \in C, \quad y \in C^\circ. \tag{14}$$

Before closing this section, we need the following related concepts on the function F and G . We say that F and G are relatively pseudo-monotone on $\Omega \subset \mathcal{H}$ if

$$\langle F(x), G(x) - G(x') \rangle \leq 0 \quad \Rightarrow \quad \langle F(x'), G(x) - G(x') \rangle \leq 0, \quad \forall x, x' \in \Omega.$$

More restrictively, F and G are relatively monotone on Ω if

$$\langle F(x) - F(x'), G(x) - G(x') \rangle \geq 0, \quad \forall x, x' \in \Omega$$

and F and G are relatively strongly monotone on Ω if there exists a $\gamma \in (0, \infty)$ such that

$$\langle F(x) - F(x'), G(x) - G(x') \rangle \geq \gamma \|x - x'\|^2, \quad \forall x, x' \in \Omega.$$

In the case where $G \equiv I$, the above three conditions reduce to F being, respectively, pseudo-monotone, monotone, and strongly monotone. When F is differentiable (in the Fréchet sense) on $\Omega \subset \mathcal{H}$, we denote by $\nabla F(x)$ the Jacobian of F at each $x \in \Omega$, viewed as a mapping from Ω to \mathcal{H} . When a function

$f : \Omega \rightarrow \mathfrak{R}$ is differentiable (in the Fréchet sense) on Ω , we denote by ∇f the gradient of f , viewed as a mapping from Ω to \mathfrak{R} .

3. Projection Residual Function

In this section, the merit function f defined by (7) will be studied, which has a relatively simple structure and is related to the growth rate of many other merit functions.

Lemma 3.1. *For any $a, b \in \mathcal{H}$*

- (i) *We have $a \in \mathcal{K}_2, b \in \mathcal{K}_*, \langle a, b \rangle = 0$ if and only if $a = \Pi_{\mathcal{K}_2}(a - b)$;*
- (ii) *We have $a \in \mathcal{K}_2, b \in \mathcal{K}_*, \langle a, b \rangle = 0$ if and only if $b = \Pi_{\mathcal{K}_*}(b - a)$.*

Proof. (i) follows from the duality of \mathcal{K}_2 and \mathcal{K}_* and item (iii) of Lemma 2.3. (ii) follows from the duality of \mathcal{K}_2 and \mathcal{K}_* and item (iii) of Lemma 2.3. This completes the proof. \square

According to Lemma 1.2, the function f defined by (7) is a merit function on Ω . We next show that f has a quadratic growth rate.

Theorem 3.2. *Let $f : \Omega \rightarrow \mathfrak{R}$ be given by (7). Then the following hold:*

- (i): *For any $x \in \Omega$, we have $f(x) \geq 0$ with $f(x) = 0$ if and only if x satisfies (2);*
- (ii): *If F and G are Lipschitz continuous with the constants L_F, L_G and relatively strongly monotone on Ω with the constant γ , then there exists a constant $c > 0$ such that $f(x) \geq c\|x - x^*\|^2$ for all $x \in \Omega$, where x^* denotes the unique solution to (2).*

Proof. (i) follows from Lemma 3.1.

(ii) Since x^* is the unique solution to (2), we have $f(x^*) = 0$ by using (i), i.e., $F(x^*) - \Pi_{\mathcal{K}_2}(F(x^*) - G(x^*)) = 0$. For any $x \in \Omega$, we get that

$$\begin{aligned}
 f(x) &= \|F(x) - \Pi_{\mathcal{K}_2}(F(x) - G(x)) - (F(x^*) - \Pi_{\mathcal{K}_2}(F(x^*) - G(x^*)))\|^2 \\
 &\geq |(\|F(x) - F(x^*)\| - \|\Pi_{\mathcal{K}_2}(F(x) - G(x)) - \Pi_{\mathcal{K}_2}(F(x^*) - G(x^*))\|)^2| \\
 &\geq |(\|F(x) - F(x^*)\| - \|(F(x) - F(x^*)) - (G(x) - G(x^*))\|)^2| \\
 &= | - (\|F(x) - F(x^*)\| - \|(F(x) - F(x^*)) - (G(x) - G(x^*))\|)^2 | \\
 &= | - \|F(x) - F(x^*)\|^2 + 2\|F(x) - F(x^*)\| \| (F(x) - F(x^*)) - (G(x) - G(x^*)) \| \\
 &\quad - \|(F(x) - F(x^*)) - (G(x) - G(x^*))\|^2 | \\
 &\geq | - 2\|F(x) - F(x^*)\| \|G(x) - G(x^*)\| + 2\langle F(x) - F(x^*), G(x) - G(x^*) \rangle \\
 &\quad - \|G(x) - G(x^*)\|^2 | \\
 &\geq |2\gamma - L_G(2L_F + L_G)| \|x - x^*\|^2.
 \end{aligned}$$

Let $c = |2\gamma - L_G(2L_F + L_G)| > 0$, then there exists a constant $c > 0$ such that

$$f(x) \geq c\|x - x^*\|^2 \text{ for all } x \in \Omega.$$

This completes the proof. \square

4. Gap Functions

In this section, we study the gap function f given by (3) and the “dual” version of the gap function defined by (4).

Theorem 4.1. *Let $f : G^{-1}(\mathcal{K}_*) \rightarrow \mathfrak{R} \cup \infty$ be given by (3). Then the following hold:*

- (i): *For any $x \in G^{-1}(\mathcal{K}_*)$, we have $f(x) \geq 0$ with $f(x) = 0$ if and only if x satisfies (2);*
- (ii): *If F and G are affine and relatively monotone on $G^{-1}(\mathcal{K}_*)$, then f is convex on $G^{-1}(\mathcal{K}_*)$.*

Proof. (i) Fix any $x \in G^{-1}(\mathcal{K}_*)$, which implies that $G(x) \in \mathcal{K}_*$. If $F(x) \in \mathcal{K}_2$, we obtain that $\langle F(x), G(x) \rangle \geq 0$ and $\langle F(x), \zeta \rangle \geq 0$ for all $\zeta \in \mathcal{K}_*$. Otherwise, there exists a $\zeta \in \mathcal{K}_*$ such that $\langle \zeta, F(x) \rangle < 0$, implying $F(x) = \infty$. Hence, it follows from (3) that $f(x) = \langle F(x), G(x) \rangle \geq 0$ and $f(x) = 0$ if and only if $F(x) \in \mathcal{K}_2$ and $\langle F(x), G(x) \rangle = 0$.

(ii) Consider any $\zeta \in \mathcal{K}_*$ and let $f_\zeta(x) = \langle F(x), G(x) - \zeta \rangle$. For any $x, x' \in G^{-1}(\mathcal{K}_*)$ and any $t \in [0, 1]$, by using the affine property and the relative monotonicity of F and G , we have

$$\begin{aligned} f_\zeta(tx + (1-t)x') &= \langle F(tx + (1-t)x'), G(tx + (1-t)x') - \zeta \rangle \\ &= \langle tF(x) + (1-t)F(x'), t(G(x) - \zeta) + (1-t)(G(x') - \zeta) \rangle \\ &= tf_\zeta(x) + (1-t)f_\zeta(x') + t(1-t)\langle F(x) - F(x'), G(x') - G(x) \rangle \\ &\geq tf_\zeta(x) + (1-t)f_\zeta(x'). \end{aligned}$$

Then f_ζ is convex on $G^{-1}(\mathcal{K}_*)$. Hence f is also convex on $G^{-1}(\mathcal{K}_*)$ for f is the pointwise maximum of f_ζ . This completes the proof. □

Theorem 4.2. *Assume F and G are relatively pseudo-monotone on $G^{-1}(\mathcal{K}_*)$, F is continuous on $G^{-1}(\mathcal{K}_*)$, and G^{-1} is defined and continuous on \mathcal{K}_* . Let $f : G^{-1}(\mathcal{K}_*) \rightarrow \mathfrak{R} \cup \infty$ be given by (4). Then the following hold:*

- (i): *For any $x \in G^{-1}(\mathcal{K}_*)$, we have that $f(x) \geq 0$ with $f(x) = 0$ if and only if x satisfies (2);*
- (ii): *If in addition G is affine on $G^{-1}(\mathcal{K}_*)$, then f is convex on $G^{-1}(\mathcal{K}_*)$.*

Proof. (i) Fix any $x \in G^{-1}(\mathcal{K}_*)$. Then $G(x) \in \mathcal{K}_*$. From (4), we conclude that $f(x) \geq 0$. If $F(x) \in \mathcal{K}_2$ and $\langle F(x), G(x) \rangle = 0$, by using (1) we have $\langle F(x), G(x) - \zeta \rangle \leq 0$ for all $\zeta \in \mathcal{K}_*$. The relative pseudo-monotonicity of F and G would imply that $\langle F(G^{-1}(\zeta)), G(x) - \zeta \rangle \leq 0$ for all $\zeta \in \mathcal{K}_*$, which means that $f(x) \leq 0$. Hence $f(x) = 0$. Conversely, if $f(x) = 0$, we have $\langle F(G^{-1}(\zeta)), G(x) - \zeta \rangle \leq 0$ for all $\zeta \in \mathcal{K}_*$. Let $x(t) = G^{-1}(t\zeta' + (1-t)G(x))$ for any $\zeta' \in \mathcal{K}_*$ and $t \in (0, 1)$. According to the relative pseudo-monotonicity of F and G , we get that

$$\langle F(x(t)), G(x) - \zeta' \rangle = \frac{1}{t} \langle F(x(t)), G(x) - (t\zeta' + (1-t)G(x)) \rangle \leq 0.$$

Letting $t \rightarrow 0$ and using the continuity of F and G^{-1} , we obtain that $\langle F(x), G(x) - \zeta' \rangle \leq 0$, which together with (1), implies that $F(x) \in \mathcal{K}_2$ and $\langle F(x), G(x) \rangle = 0$.

(ii) Since G is affine on $G^{-1}(\mathcal{K}_*)$, the function $f_\zeta(x) = \langle F(G^{-1}(\zeta)), G(x) - \zeta \rangle$ is affine on $G^{-1}(\mathcal{K}_*)$ for each $\zeta \in \mathcal{K}_*$. Hence f , being the pointwise maximum of f_ζ and $\zeta \in \mathcal{K}_*$, is convex on $G^{-1}(\mathcal{K}_*)$. This completes the proof. \square

5. Regularized Gap Function

In this section, we study the regularized gap function $f_\alpha(x)$ given by (5). We begin with the following lemma.

Lemma 5.1. *For any $\alpha \in (0, \infty)$, define the function $\psi_\alpha : \mathcal{H} \times \mathcal{H} \rightarrow \Re$ by*

$$\psi_\alpha(a, b) = \max_{\zeta \in \mathcal{K}_*} \left\{ \langle a, b - \zeta \rangle - \frac{1}{2\alpha} \|b - \zeta\|^2 \right\}.$$

Then the following hold:

(i) For all $a \in \mathcal{H}, b \in \mathcal{K}_*$, we have

$$\psi_\alpha(a, b) \geq \frac{1}{2\alpha} \|b - \Pi_{\mathcal{K}_*}(b - \alpha a)\|^2, \quad (15)$$

and $\psi_\alpha(a, b) = 0$ if and only if in addition $a \in \mathcal{K}_2$ and $\langle a, b \rangle = 0$.

(ii) ψ_α is differentiable at every $a, b \in \mathcal{H}$ with

$$\nabla_a \psi_\alpha(a, b) = b - \Pi_{\mathcal{K}_*}(b - \alpha a), \quad \nabla_b \psi_\alpha(a, b) = a - \frac{1}{\alpha} (b - \Pi_{\mathcal{K}_*}(b - \alpha a)).$$

Proof. (i) For any $a, b \in \mathcal{H}$, let

$$g_\alpha(\zeta) = \langle a, b - \zeta \rangle - \frac{1}{2\alpha} \|b - \zeta\|^2.$$

Then, we get that

$$\begin{aligned} \psi_\alpha(a, b) &= \max_{\zeta \in \mathcal{K}_*} \left\{ \langle a, b - \zeta \rangle - \frac{1}{2\alpha} \|b - \zeta\|^2 \right\} \\ &= -\frac{1}{2\alpha} \min_{\zeta \in \mathcal{K}_*} \left\{ \|\zeta - (b - \alpha a)\|^2 - \alpha^2 \|a\|^2 \right\}, \end{aligned}$$

which implies that the maximum point of $g_\alpha(\zeta)$ is $\bar{\zeta} = \Pi_{\mathcal{K}_*}(b - \alpha a)$. Thus, for all $a \in \mathcal{H}, b \in \mathcal{K}_*$, we have the following

$$\begin{aligned} \psi_\alpha(a, b) - \frac{1}{2\alpha} \|b - \bar{\zeta}\|^2 &= g_\alpha(\bar{\zeta}) - \frac{1}{2\alpha} \|b - \bar{\zeta}\|^2 \\ &= \langle a, b - \bar{\zeta} \rangle - \frac{1}{\alpha} \|b - \bar{\zeta}\|^2 \\ &= \frac{1}{\alpha} \langle \bar{\zeta} - (b - \alpha a), b - \bar{\zeta} \rangle \\ &\geq 0, \end{aligned}$$

where the inequality follows from that $\langle b - \Pi_{\mathcal{K}_*}(b - \alpha a), b - \alpha a - \Pi_{\mathcal{K}_*}(b - \alpha a) \rangle = \langle b - \Pi_{\mathcal{K}_*}(b - \alpha a), \Pi_{\mathcal{K}_2}(b - \alpha a) \rangle \leq 0$. Hence $\psi_\alpha(a, b) \geq \frac{1}{2\alpha} \|b - \Pi_{\mathcal{K}_*}(b - \alpha a)\|^2$ for all $a \in \mathcal{H}$ and $b \in \mathcal{K}_*$.

If $a \in \mathcal{K}_2$ and $\langle a, b \rangle = 0$, for any $\alpha \in (0, \infty)$, we have $\langle \alpha a, b \rangle = 0$. It follows from Lemma 3.1 that $b = \Pi_{\mathcal{K}_*}(b - \alpha a) = \bar{\zeta}$. Thus,

$$\begin{aligned} \psi_\alpha(a, b) &= \max_{\zeta \in \mathcal{K}_*} \left\{ \langle a, b - \zeta \rangle - \frac{1}{2\alpha} \|b - \zeta\|^2 \right\} \\ &= \langle a, b - \bar{\zeta} \rangle - \frac{1}{2\alpha} \|b - \bar{\zeta}\|^2 \\ &= 0. \end{aligned}$$

Conversely, if $\psi_\alpha(a, b) = 0$, it follows from (15) that $b = \Pi_{\mathcal{K}_*}(b - \alpha a)$. According to Lemma 3.1 (ii), we deduce that $\alpha a \in \mathcal{K}_2$ and $\langle \alpha a, b \rangle = 0$ for any $\alpha \in [0, \infty)$. If $\alpha = 1$, we have $a \in \mathcal{K}_2$ and $\langle a, b \rangle = 0$.

(ii) follows from Lemma 2.2 (i).

Combing (i) and (ii), we complete the proof. \square

We will relate f_α to the norm of the projection residual functions $R_\alpha : \Omega \rightarrow \Omega$ for any $\alpha \in (0, \infty)$ defined by

$$R_\alpha(x) := G(x) - \Pi_{\mathcal{K}_*}(G(x) - \alpha F(x)).$$

By using Lemma 3.1 (ii), we have that x satisfies (2) if and only if $R_\alpha(x) = 0$. By using Lemma 5.1, we obtain the following theorem which estimates the growth rate of f_α in terms of $\|R_\alpha\|$, and gives formulas for ∇f_α and a certain descent direction for f_α at any nonglobal minimum $x \in G^{-1}(\mathcal{K}_*)$ with $\nabla G(x)$ invertible and $\nabla G(x)^{-1} \nabla F(x)$ satisfying the condition below

$$\langle a, \nabla G(x)^{-1} \nabla F(x) a \rangle > 0 \tag{16}$$

for all $a \in \mathcal{H}$.

Theorem 5.2. Fix any $\alpha \in (0, \infty)$ and let $f_\alpha : G^{-1}(\mathcal{K}_*) \rightarrow \Re$ be given by (5). Then the following hold:

(i): For all $x \in G^{-1}(\mathcal{K}_*)$, we have

$$f_\alpha(x) \geq \frac{1}{2\alpha} \|R_\alpha(x)\|^2,$$

and $f_\alpha(x) = 0$ if and only if x satisfies (2).

(ii): If F and G is differentiable on $G^{-1}(\mathcal{K}_*)$. Then, so is f_α and

$$\nabla f_\alpha(x) = \nabla F(x) R_\alpha(x) + \nabla G(x) (F(x) - \frac{1}{\alpha} R_\alpha(x))$$

for all $x \in \mathcal{K}_*$.

(iii): Assume that F and G are differentiable on $G^{-1}(\mathcal{K}_*)$. Then, for every $x \in G^{-1}(\mathcal{K}_*)$ where $\nabla G(x)$ is invertible and $\nabla G(x)^{-1} \nabla F(x)$ satisfies (16), either $f_\alpha(x) = 0$ or $\nabla f_\alpha(x) \neq 0$ with $\langle d(x), \nabla f_\alpha(x) \rangle < 0$, where

$$d(x) = -(\nabla G(x)^{-1})^* R_\alpha(x)$$

and $(\nabla G(x)^{-1})^*$ is the adjoint operator of $\nabla G(x)^{-1}$.

Proof. (i) and (ii) follow from Lemma 5.1, we omit it. Now we show that (iii) is valid. Fix any $x \in G^{-1}(\mathcal{K}_*)$ with $\nabla G(x)$ invertible and $\nabla G(x)^{-1}\nabla F(x)$ satisfying the condition (16). Then

$$\begin{aligned}\langle d(x), \nabla f_\alpha(x) \rangle &= -\langle R_\alpha(x), \nabla G(x)^{-1}\nabla F(x)R_\alpha(x) + F(x) - \frac{1}{\alpha}R_\alpha(x) \rangle \\ &\leq -\langle R_\alpha(x), \nabla G(x)^{-1}\nabla F(x)R_\alpha(x) \rangle,\end{aligned}$$

where the inequality follows from Lemma 2.3 that

$$\begin{aligned}&\langle G(x) - \Pi_{\mathcal{K}_*}(G(x) - \alpha F(x)), G(x) - \alpha F(x) - \Pi_{\mathcal{K}_*}(G(x) - \alpha F(x)) \rangle \\ &= \langle G(x) - \Pi_{\mathcal{K}_*}(G(x) - \alpha F(x)), \Pi_{-\mathcal{K}_2}(G(x) - \alpha F(x)) \rangle \\ &\leq 0.\end{aligned}$$

Since $\nabla G(x)^{-1}\nabla F(x)$ satisfies (16), then either $R_\alpha(x) = 0$ or $\langle d(x), \nabla f_\alpha(x) \rangle < 0$. We know that $R_\alpha(x) = 0$ implies that x satisfies (2), it follows from (i) that $f_\alpha(x) = 0$. This completes the proof. \square

6. Implicit Lagrangian Function

In this section, the implicit Lagrangian function f_α defined by (6) will be discussed. We establish the following lemma and theorem.

Lemma 6.1. For any $\alpha \in (0, \infty)$, define the function $\psi_\alpha : \mathcal{H} \times \mathcal{H} \rightarrow \Re$ by

$$\psi_\alpha(a, b) = \max_{\xi \in \mathcal{K}_2, \zeta \in \mathcal{K}_*} \{ \langle a, b - \zeta \rangle - \langle \xi, b \rangle - \frac{1}{2\alpha} (\|a - \xi\|^2 + \|b - \zeta\|^2) \}.$$

Then the following hold:

(i): Fix any $\alpha \in (1, \infty)$. For all $a, b \in \mathcal{H}$, we have

$$(\alpha - 1)\|b - \Pi_{\mathcal{K}_*}(b - a)\|^2 \geq \psi_\alpha(a, b) = -\psi_{\frac{1}{\alpha}}(a, b) \geq (1 - \frac{1}{\alpha})\|b - \Pi_{\mathcal{K}_*}(b - a)\|^2 \quad (17)$$

and $\psi_\alpha(a, b) = 0$ if and only if in addition $a \in \mathcal{K}_2, b \in \mathcal{K}_*$ and $\langle a, b \rangle = 0$.

(ii): Fix any $\alpha \in (0, \infty)$. ψ_α is differentiable at every $(a, b) \in \Omega \times \Omega$, with

$$\begin{aligned}\nabla_a \psi_\alpha(a, b) &= b - \Pi_{\mathcal{K}_*}(b - \alpha a) - \frac{1}{\alpha}(a - \Pi_{\mathcal{K}_2}(a - \alpha b)), \\ \nabla_b \psi_\alpha(a, b) &= a - \Pi_{\mathcal{K}_2}(a - \alpha b) - \frac{1}{\alpha}(b - \Pi_{\mathcal{K}_*}(b - \alpha a)).\end{aligned}$$

Proof. (i) Fix any $\alpha \in (1, \infty)$. For any $a, b \in \mathcal{H}$, we first verify that

$$\psi_\alpha(a, b) = -\psi_{\frac{1}{\alpha}}(a, b).$$

By careful calculation, we have

$$\begin{aligned}\psi_\alpha(a, b) &= \max_{\xi \in \mathcal{K}_2, \zeta \in \mathcal{K}_*} \{ \langle a, b - \zeta \rangle - \langle \xi, b \rangle - \frac{1}{2\alpha} (\|a - \xi\|^2 + \|b - \zeta\|^2) \} \\ &= \max_{\xi \in \mathcal{K}_2} \{ -\frac{1}{2\alpha} \|a - \xi\|^2 - \langle \xi, b \rangle \} + \max_{\zeta \in \mathcal{K}_*} \{ -\frac{1}{2\alpha} \|b - \zeta\|^2 - \langle a, \zeta \rangle \} + \langle a, b \rangle \\ &= \max_{\xi \in \mathcal{K}_2} \{ -\frac{1}{2\alpha} \|\xi - (a - \alpha b)\|^2 + \frac{1}{2\alpha} \|a - \alpha b\|^2 - \frac{1}{2\alpha} \|a\|^2 \} \quad (18)\end{aligned}$$

$$\begin{aligned}
 & + \max_{\zeta \in \mathcal{K}_*} \left\{ -\frac{1}{2\alpha} \|\zeta - (b - \alpha a)\|^2 - \frac{1}{2\alpha} \|b - \alpha a\|^2 - \frac{1}{2\alpha} \|b\|^2 \right\} + \langle a, b \rangle \\
 & = \frac{1}{2\alpha} \|\Pi_{\mathcal{K}_2}(a - \alpha b)\|^2 - \frac{1}{2\alpha} \|a\|^2 + \frac{1}{2\alpha} \|\Pi_{\mathcal{K}_*}(b - \alpha a)\|^2 - \frac{1}{2\alpha} \|b\|^2 + \langle a, b \rangle
 \end{aligned}$$

and

$$\begin{aligned}
 -\psi_{\frac{1}{\alpha}}(a, b) & = -\max_{\xi \in \mathcal{K}_2, \zeta \in \mathcal{K}_*} \{ \langle a, b - \zeta \rangle - \langle \xi, b \rangle - \frac{\alpha}{2} (\|a - \xi\|^2 + \|b - \zeta\|^2) \} \\
 & = \min_{\xi \in \mathcal{K}_2, \zeta \in \mathcal{K}_*} \left\{ \frac{\alpha}{2} (\|a - \xi\|^2 + \|b - \zeta\|^2) - \langle a, b \rangle + \langle a, \zeta \rangle + \langle \xi, b \rangle \right\} \\
 & = \min_{\xi \in \mathcal{K}_2} \left\{ \frac{\alpha}{2} \|a - \xi\|^2 + \langle \xi, b \rangle \right\} + \min_{\zeta \in \mathcal{K}_*} \left\{ \frac{\alpha}{2} \|b - \zeta\|^2 + \langle a, \zeta \rangle \right\} - \langle a, b \rangle \\
 & = \min_{\xi \in \mathcal{K}_2} \left\{ \frac{\alpha}{2} \|\xi - (a - \alpha^{-1}b)\|^2 - \frac{\alpha}{2} \|a - \alpha^{-1}b\|^2 + \frac{\alpha}{2} \|a\|^2 \right\} \tag{19} \\
 & \quad + \min_{\zeta \in \mathcal{K}_*} \left\{ \frac{\alpha}{2} \|\zeta - (b - \alpha^{-1}a)\|^2 - \frac{\alpha}{2} \|b - \alpha^{-1}a\|^2 + \frac{\alpha}{2} \|b\|^2 \right\} - \langle a, b \rangle \\
 & = -\frac{1}{2\alpha} \|\Pi_{\mathcal{K}_2}(\alpha a - b)\|^2 + \frac{\alpha}{2} \|a\|^2 - \frac{1}{2\alpha} \|\Pi_{\mathcal{K}_*}(\alpha b - a)\|^2 + \frac{\alpha}{2} \|b\|^2 - \langle a, b \rangle.
 \end{aligned}$$

According to Lemma 2.3, (18) and (19), we have

$$\begin{aligned}
 & \psi_{\alpha}(a, b) - (-\psi_{\frac{1}{\alpha}}(a, b)) \\
 & = \frac{1}{2\alpha} \|\Pi_{\mathcal{K}_2}(a - \alpha b)\|^2 + \frac{1}{2\alpha} \|\Pi_{\mathcal{K}_*}(b - \alpha a)\|^2 + \frac{1}{2\alpha} \|\Pi_{\mathcal{K}_2}(\alpha a - b)\|^2 + \frac{1}{2\alpha} \|\Pi_{\mathcal{K}_*}(\alpha b - a)\|^2 \\
 & \quad - \frac{1}{2\alpha} \|a\|^2 - \frac{\alpha}{2} \|a\|^2 - \frac{1}{2\alpha} \|b\|^2 - \frac{\alpha}{2} \|b\|^2 + 2\langle a, b \rangle \\
 & = \frac{1}{2\alpha} (\|\Pi_{\mathcal{K}_2}(a - \alpha b)\|^2 + \|\Pi_{\mathcal{K}_*}(\alpha b - a)\|^2) + \frac{1}{2\alpha} (\|\Pi_{\mathcal{K}_2}(\alpha a - b)\|^2 + \|\Pi_{\mathcal{K}_*}(b - \alpha a)\|^2) \\
 & \quad - \frac{1}{2\alpha} \|a - \alpha b\|^2 - \frac{1}{2\alpha} \|b - \alpha a\|^2 \\
 & = 0,
 \end{aligned}$$

where the second equality follows from that $\|\Pi_{\mathcal{K}_2}(a - \alpha b)\|^2 + \|\Pi_{\mathcal{K}_*}(\alpha b - a)\|^2 = \|a - \alpha b\|^2$ and $\|\Pi_{\mathcal{K}_2}(\alpha a - b)\|^2 + \|\Pi_{\mathcal{K}_*}(b - \alpha a)\|^2 = \|b - \alpha a\|^2$. This shows that $\psi_{\alpha}(a, b) = -\psi_{\frac{1}{\alpha}}(a, b)$.

We show below that $\psi_{\alpha}(a, b) \geq (1 - \frac{1}{\alpha})\|b - \Pi_{\mathcal{K}_*}(b - a)\|^2$. For any $a, b \in \mathcal{H}$, let

$$\xi_0 = \Pi_{\mathcal{K}_2}(a - b), \quad \zeta_0 = \Pi_{\mathcal{K}_*}(b - a)$$

and

$$g_{\alpha}(\xi, \zeta) = \langle a, b - \zeta \rangle - \langle \xi, b \rangle - \frac{1}{2\alpha} (\|a - \xi\|^2 + \|b - \zeta\|^2).$$

Then, from the definition of ψ_{α} , we know that

$$\psi_{\alpha}(a, b) \geq g_{\alpha}(\xi_0, \zeta_0). \tag{20}$$

Next, we compute $g_{\alpha}(\xi_0, \zeta_0)$. Recalling that

$$\begin{aligned}
 b - a & = \Pi_{\mathcal{K}_*}(b - a) + \Pi_{-\mathcal{K}_2}(b - a), & a - b & = \Pi_{\mathcal{K}_2}(a - b) + \Pi_{-\mathcal{K}_*}(a - b), \\
 \langle \Pi_{-\mathcal{K}_2}(b - a), \Pi_{\mathcal{K}_*}(b - a) \rangle & = 0, & \Pi_{-\mathcal{K}_2}(b - a) & = -\Pi_{\mathcal{K}_2}(a - b),
 \end{aligned}$$

we obtain by direct calculation that

$$\begin{aligned}
g_\alpha(\xi_0, \zeta_0) &= \langle a, b - \Pi_{\mathcal{K}_*}(b - a) \rangle - \langle \Pi_{\mathcal{K}_2}(a - b), b \rangle \\
&\quad - \frac{1}{2\alpha} (\|a - \Pi_{\mathcal{K}_2}(a - b)\|^2 + \|b - \Pi_{\mathcal{K}_*}(b - a)\|^2) \\
&= \langle b - \Pi_{\mathcal{K}_*}(b - a) - \Pi_{-\mathcal{K}_2}(b - a), b - \Pi_{\mathcal{K}_*}(b - a) \rangle \\
&\quad - \langle \Pi_{\mathcal{K}_2}(a - b), b \rangle - \frac{1}{2\alpha} (\|a - \Pi_{\mathcal{K}_2}(a - b)\|^2 + \|b - \Pi_{\mathcal{K}_*}(b - a)\|^2) \\
&= \|b - \Pi_{\mathcal{K}_*}(b - a)\|^2 - \langle \Pi_{-\mathcal{K}_2}(b - a), b - \Pi_{\mathcal{K}_*}(b - a) \rangle \\
&\quad - \langle \Pi_{\mathcal{K}_2}(a - b), b \rangle - \frac{1}{2\alpha} (\|a - \Pi_{\mathcal{K}_2}(a - b)\|^2 + \|b - \Pi_{\mathcal{K}_*}(b - a)\|^2) \\
&= \|b - \Pi_{\mathcal{K}_*}(b - a)\|^2 - \langle \Pi_{-\mathcal{K}_2}(b - a), b \rangle - \langle \Pi_{\mathcal{K}_2}(a - b), b \rangle \\
&\quad - \frac{1}{2\alpha} (\|a - \Pi_{\mathcal{K}_2}(a - b)\|^2 + \|b - \Pi_{\mathcal{K}_*}(b - a)\|^2) \\
&= \|b - \Pi_{\mathcal{K}_*}(b - a)\|^2 + \langle \Pi_{\mathcal{K}_2}(a - b), b \rangle - \langle \Pi_{\mathcal{K}_2}(a - b), b \rangle \\
&\quad - \frac{1}{2\alpha} (\|a - \Pi_{\mathcal{K}_2}(a - b)\|^2 + \|b - \Pi_{\mathcal{K}_*}(b - a)\|^2) \\
&= \|b - \Pi_{\mathcal{K}_*}(b - a)\|^2 - \frac{1}{2\alpha} (\|a - \Pi_{\mathcal{K}_2}(a - b)\|^2 + \|b - \Pi_{\mathcal{K}_*}(b - a)\|^2) \\
&= \|b - \Pi_{\mathcal{K}_*}(b - a)\|^2 - \frac{1}{\alpha} \|b - \Pi_{\mathcal{K}_*}(b - a)\|^2 \\
&= (1 - \frac{1}{\alpha}) \|b - \Pi_{\mathcal{K}_*}(b - a)\|^2,
\end{aligned}$$

which, together with (20), implies that $\psi_\alpha(a, b) \geq (1 - \frac{1}{\alpha}) \|b - \Pi_{\mathcal{K}_*}(b - a)\|^2$. Similarly, we can prove that $\psi_{\frac{1}{\alpha}}(a, b) \geq (1 - \alpha) \|b - \Pi_{\mathcal{K}_*}(b - a)\|^2$, i.e., $-\psi_{\frac{1}{\alpha}}(a, b) \leq (\alpha - 1) \|b - \Pi_{\mathcal{K}_*}(b - a)\|^2$. Therefore, the relation (17) holds.

We are ready to prove that the remainder conclusion of item (i). Suppose that $a \in \mathcal{K}_2, b \in \mathcal{K}_*$ and $\langle a, b \rangle = 0$. Then, from Lemma 3.1, we know that $b = \Pi_{\mathcal{K}_*}(b - a)$. This, together with (17), shows that $\psi_\alpha(a, b) = 0$.

Conversely, suppose that $\psi_\alpha(a, b) = 0$ holds for any $\alpha \in (1, \infty)$. Then, from (17), we obtain that $\|b - \Pi_{\mathcal{K}_*}(b - a)\|^2 = 0$, which means that $b = \Pi_{\mathcal{K}_*}(b - a)$. It follows from Lemma 3.1 that $a \in \mathcal{K}_2, b \in \mathcal{K}_*$ and $\langle a, b \rangle = 0$.

(ii) follows from Lemma 2.2 (i). This completes the proof. \square

Next we define the projection residual function $S_\alpha : \Omega \rightarrow \Omega$ for any $\alpha \in (0, \infty)$ by

$$S_\alpha(x) = F(x) - \Pi_{\mathcal{K}_2}(F(x) - \alpha G(x)).$$

According to Lemma 6.1, we obtain the following theorem which estimates the growth rate of f_α in terms of $\|R_1\|$, and give formulas for ∇f_α and a certain descent direction for f_α at any nonglobal minimum x with $\nabla G(x)$ invertible and $\nabla G(x)^{-1}F(x)$ satisfying (16).

Theorem 6.2. Fix any $\alpha \in (1, \infty)$ and let $f_\alpha : \Omega \rightarrow \Re$ be given by (6). Then the following hold:

(i): For all $x \in \Omega$, we have

$$(\alpha - 1)\|R_1(x)\|^2 \geq f_\alpha(x) = -f_{\frac{1}{\alpha}}(x) \geq (1 - \frac{1}{\alpha})\|R_1(x)\|^2$$

and $f_\alpha(x) = 0$ if and only if x satisfies (2).

(ii): If F is differentiable on Ω , Then, so is f_α and

$$\nabla f_\alpha(x) = \nabla F(x)(R_\alpha(x) - \frac{1}{\alpha}S_\alpha(x)) + \nabla G(x)(S_\alpha(x) - \frac{1}{\alpha}R_\alpha(x))$$

for all $x \in \Omega$.

(iii): Assume that F and G are differentiable on Ω . Then, for every $x \in \Omega$ where $\nabla G(x)$ is invertible and $\nabla G(x)^{-1}\nabla F(x)$ satisfies (16), either $f_\alpha(x) = 0$ or $\nabla f_\alpha(x) \neq 0$

with $\langle d(x), \nabla f_\alpha(x) \rangle \leq -\langle d(x), \nabla G(x)^{-1}\nabla F(x)d(x) \rangle$, where

$$d(x) = -(\nabla G(x)^{-1})^*(R_\alpha(x) - \frac{1}{\alpha}S_\alpha(x))$$

and $(G(x)^{-1})^*$ is the adjoint operator of $G(x)^{-1}$.

Proof. (i) and (ii) follow from Lemma 6.1.

(iii) Fix any $x \in \Omega$ with $\nabla G(x)$ invertible and $\nabla G(x)^{-1}\nabla F(x)$ satisfying (16). For simplicity, we omit (x) below. By the definition of R_α and S_α , we have that

$$\begin{aligned} R_\alpha - \frac{1}{\alpha}S_\alpha &= G - \Pi_{\mathcal{K}_*}(G - \alpha F) - \frac{1}{\alpha}(F - \Pi_{\mathcal{K}_2}(F - \alpha G)) \\ &= -\Pi_{\mathcal{K}_*}(G - \alpha F) - \frac{1}{\alpha}[F - \alpha G - \Pi_{\mathcal{K}_2}(F - \alpha G)], \\ S_\alpha - \frac{1}{\alpha}R_\alpha &= F - \Pi_{\mathcal{K}_2}(F - \alpha G) - \frac{1}{\alpha}(G - \Pi_{\mathcal{K}_*}(G - \alpha F)) \\ &= -\Pi_{\mathcal{K}_2}(F - \alpha G) - \frac{1}{\alpha}[G - \alpha F - \Pi_{\mathcal{K}_*}(G - \alpha F)]. \end{aligned}$$

It is easy to see that

$$\langle R_\alpha - \frac{1}{\alpha}S_\alpha, S_\alpha - \frac{1}{\alpha}R_\alpha \rangle \geq 0.$$

Consequently, we obtain that

$$\begin{aligned} \langle d, \nabla f_\alpha \rangle &= -\langle R_\alpha - \frac{1}{\alpha}S_\alpha, \nabla G^{-1}\nabla F(R_\alpha - \frac{1}{\alpha}S_\alpha) + S_\alpha - \frac{1}{\alpha}R_\alpha \rangle \\ &\leq -\langle R_\alpha - \frac{1}{\alpha}S_\alpha, \nabla G^{-1}\nabla F(R_\alpha - \frac{1}{\alpha}S_\alpha) \rangle \\ &= -\langle d, \nabla G^{-1}\nabla Fd \rangle. \end{aligned}$$

If $\nabla f_\alpha(x) = 0$, then $d(x) = 0$, which means that $R_\alpha - \frac{1}{\alpha}S_\alpha = 0$. Thus, from (ii) and the nonsingularity of $\nabla G(x)$, we conclude that $S_\alpha - \frac{1}{\alpha}R_\alpha = 0$. Since $\alpha \neq 1$, the latter two equations would yield $R_\alpha(x) = S_\alpha(x) = 0$, which shows that $f_\alpha(x) = 0$. This completes the proof. \square

7. A Function of Luo and Tseng

In this section, we study the merit function f given by (8) with ψ_0 satisfying $\psi_0(t) = 0$ if and only if $t \leq 0$ and ψ satisfying (9). In the subsequent analysis, we will further restrict the choice of ψ . Let ψ_+ denote the collection of $\psi : \mathcal{H} \times \mathcal{H} \rightarrow [0, \infty)$ satisfying (9) that are differentiable and satisfying the following conditions:

$$\langle \nabla_a \psi(a, b), \nabla_b \psi(a, b) \rangle \geq 0, \langle a, \nabla_a \psi(a, b) \rangle + \langle b, \nabla_b \psi(a, b) \rangle \geq 0, \forall a, b \in \mathcal{H}. \quad (21)$$

The theorem below provides one choice of ψ belong to ψ_+ . Moreover, the choice of ψ is convex.

Theorem 7.1. *Let $\psi : \mathcal{H} \times \mathcal{H} \rightarrow [0, \infty)$ be given by*

$$\psi(a, b) := \frac{1}{2}(\|\Pi_{-\mathcal{K}_*}(a)\|^2 + \|\Pi_{-\mathcal{K}_2}(b)\|^2). \quad (22)$$

Then the following hold:

- (i): ψ satisfies (9).
- (ii): ψ is convex and differentiable at every $a, b \in \mathcal{H}$ with $\nabla_a \psi(a, b) = \Pi_{-\mathcal{K}_*}(a)$ and $\nabla_b \psi(a, b) = \Pi_{-\mathcal{K}_2}(b)$.
- (iii): For every $a, b \in \mathcal{H}$, we have $\langle \nabla_a \psi(a, b), \nabla_b \psi(a, b) \rangle \geq 0$.
- (iv): For every $a, b \in \mathcal{H}$, we have $\langle a, \nabla_a \psi(a, b) \rangle + \langle b, \nabla_b \psi(a, b) \rangle = \|\Pi_{-\mathcal{K}_*}(a)\|^2 + \|\Pi_{-\mathcal{K}_2}(b)\|^2$.

Proof. (i) and (ii). By Lemma 2.3, we have $a = \Pi_{\mathcal{K}_2}(a) + \Pi_{-\mathcal{K}_*}(a)$ for $a \in \mathcal{H}$. Hence, the following holds:

$$\|\Pi_{-\mathcal{K}_*}(a)\|^2 = \|a - \Pi_{\mathcal{K}_2}(a)\|^2 = \min_{w \in \mathcal{K}_2} \|a - w\|^2. \quad (23)$$

It follows from Lemma 2.2 that $\|\Pi_{-\mathcal{K}_*}(a)\|^2$ is differentiable and convex in a and $\nabla_a \psi(a, b) = a - \Pi_{\mathcal{K}_2}(a) = \Pi_{-\mathcal{K}_*}(a)$. From (23), we know that $\|\Pi_{-\mathcal{K}_*}(a)\|^2 = 0$ if and only if $a \in \mathcal{K}_2$. Similarly, we can prove that $b \in \mathcal{K}_*$ and $\nabla_b \psi(a, b) = \Pi_{-\mathcal{K}_2}(b)$. Hence, $\psi(a, b)$ is differentiable convex in (a, b) and equals 0 if and only if $a \in \mathcal{K}_2$ and $b \in \mathcal{K}_*$. Since $a \in \mathcal{K}_2$ and $b \in \mathcal{K}_*$ implies that $\langle a, b \rangle \geq 0$, it follows that (9) holds.

(iii) and (iv). By (ii) and Lemma 2.3, we have

$$\langle \nabla_a \psi(a, b), \nabla_b \psi(a, b) \rangle = \langle \Pi_{-\mathcal{K}_*}(a), \Pi_{-\mathcal{K}_2}(b) \rangle = \langle \Pi_{\mathcal{K}_2}(-a), \Pi_{\mathcal{K}_*}(-b) \rangle \geq 0.$$

Also, we have

$$\langle a, \nabla_a \psi(a, b) \rangle = \langle a, \Pi_{-\mathcal{K}_*}(a) \rangle = \langle \Pi_{\mathcal{K}_2}(a) + \Pi_{-\mathcal{K}_*}(a), \Pi_{-\mathcal{K}_*}(a) \rangle = \|\Pi_{-\mathcal{K}_*}(a)\|^2.$$

Similarly, we can show that $\langle b, \nabla_b \psi(a, b) \rangle = \|\Pi_{-\mathcal{K}_2}(b)\|^2$. This completes the proof. \square

Next, we consider a further restriction on ψ . Let ψ_{++} denote the collection of $\psi \in \psi_+$ satisfying the following condition:

$$\psi(a, b) = 0 \quad \forall a, b \in \mathcal{H} \quad \text{with} \quad \langle \nabla_a \psi(a, b), \nabla_b \psi(a, b) \rangle = 0. \quad (24)$$

The following theorem shows that f defined by (8) and (9) is a merit function on \mathcal{H} , and gives formulas for ∇f . In the following, we assume that $\psi \in \psi_+$ and a certain descent direction for f at any non global minimum x with $\nabla G(x)$ invertible and $\nabla G(x)^{-1}\nabla F(x)$ satisfying (16). Furthermore, we suppose that $\psi \in \psi_{++}$ and a certain descent direction for f at any non global minimum x with $\nabla G(x)$ invertible and $\nabla G(x)^{-1}\nabla F(x)$ satisfying the following

$$\langle a, \nabla G(x)^{-1}\nabla F(x)a \rangle \geq 0 \tag{25}$$

for all $a \in \mathcal{H}$.

Theorem 7.2. *Let $f : \mathcal{H} \rightarrow \Re$ be given by (8) with $\psi_0 : \Re \rightarrow [0, \infty)$ satisfying $\psi_0(t) = 0$ if and only if $t \leq 0$ and $\psi : \mathcal{H} \times \mathcal{H} \rightarrow [0, \infty)$ satisfying (9). Then the following hold:*

(i): *For all $x \in \Omega$, we have $f(x) \geq 0$ and $f(x) = 0$ if and only if x satisfies (2).*

(ii): *If ψ_0, ψ and F, G are differentiable, then so is f and*

$$\begin{aligned} \nabla f(x) = & \nabla\psi_0(\langle F(x), G(x) \rangle)(\nabla F(x)G(x) + \nabla G(x)F(x)) \\ & + \nabla F(x)\nabla_a\psi(F(x), G(x)) + \nabla G(x)\nabla_b\psi(F(x), G(x)) \end{aligned}$$

for all $x \in \Omega$.

(iii): *If ψ_0, ψ are convex and F and G are affine and relatively monotone, then f is convex.*

(iv): *Assume that F and G are differentiable on Ω and $\psi \in \psi_+$ (respectively, ψ_{++}) and ψ_0 is differentiable and strictly increasing on $[0, \infty)$. Then for every $x \in \Omega$ where $\nabla G(x)$ is invertible and $\nabla G(x)^{-1}\nabla F(x)$ satisfies (16)(respectively, (25)), either $f(x) = 0$ or $\nabla f(x) \neq 0$ with $\langle d(x), \nabla f(x) \rangle < 0$, where*

$$d(x) := -(\nabla G(x)^{-1})^*(\nabla\psi_0(\langle F(x), G(x) \rangle)G(x) + \nabla_a\psi(F(x), G(x)))$$

and $(G(x)^{-1})^*$ is the adjoint operator of $G(x)^{-1}$.

Proof. (i) follows from (8) and the assumptions on ψ_0, ψ . (ii) follows from the chain rule. (iii) follows from the observations that, under the given hypothesis, $x \rightarrow \langle F(x), G(x) \rangle$ is convex (see the proof of Theorem 4.1 (ii)) and ψ_0 is convex and nondecreasing, so their composition is convex. Also, $x \rightarrow (F(x), G(x))$ with the convex function ψ , is convex.

(iv) Consider the case $\psi \in \psi_{++}$ and fix any $x \in \mathcal{H}$ with $\nabla G(x)$ invertible and $\nabla G(x)^{-1}\nabla F(x)$ satisfying (25). Let $\beta := \nabla\psi_0(\langle F(x), G(x) \rangle)$. For simplicity, we drop (x) in the following.

$$\begin{aligned} \langle d, \nabla f \rangle = & -\langle \beta G + \nabla_a\psi(F, G), \nabla G^{-1}\nabla F(\beta G + \nabla_a\psi(F, G)) + \beta F + \nabla_b\psi(F, G) \rangle \\ \leq & -\langle \beta G + \nabla_a\psi(F, G), \beta F + \nabla_b\psi(F, G) \rangle \\ = & -\beta^2 \langle F, G \rangle - \beta(\langle G, \nabla_b\psi(F, G) \rangle + \langle F, \nabla_a\psi(F, G) \rangle) - \langle \nabla_a\psi(F, G), \nabla_b\psi(F, G) \rangle \\ \leq & -\beta^2 \langle F, G \rangle - \langle \nabla_a\psi(F, G), \nabla_b\psi(F, G) \rangle, \end{aligned}$$

where the last inequality follows from $\beta \geq 0$ and (21). Since ψ_0 is strictly increasing on $[0, \infty)$, $t\nabla\psi_0(t) > 0$ if and only if $t > 0$. Hence $-\beta^2\langle F, G \rangle \leq 0$ and $\beta^2\langle F, G \rangle = 0$ only if $\langle F, G \rangle \leq 0$. From (21) and (24), we know that $-\langle \nabla_a\psi(F, G), \nabla_b\psi(F, G) \rangle \leq 0$ and $\langle \nabla_a\psi(F, G), \nabla_b\psi(F, G) \rangle = 0$ only if $\psi(F, G) = 0$. Thus, $\langle d(x), \nabla f(x) \rangle < 0$ unless $\langle F(x), G(x) \rangle = 0$, which implies x satisfies (2) or $f(x) = 0$. The case of $\psi \in \psi_+$ and $\nabla G(x)^{-1}\nabla F(x)$ satisfying (16) goes in a similar fashion. This completes the proof. \square

8. Conclusions

In this paper, we have discussed the merit functions such as the projection residual function, the gap function, the regularized gap function, the implicit Lagrangian function and the function of Luo and Tseng for solving MCCP. For each of the above five choices of merit functions, we have derived conditions for the merit function to be convex and/or differentiable, and for the stationary point of the merit function to be a solution of MCCP. These results pave the way for the task of using the optimization methods based on merit functions to solve MCCP. Our next step is to carry out implementations and empirical comparison of the algorithms based on the above five choices.

REFERENCES

1. F. Alizadeh and S. Schmieta, *Symmetric cones, potential reduction methods*, in Handbook of Semidefinite Programming, H. Wolkowicz, R. Saigal, and L. Vandenberghe, eds., Kluwer, Boston, (2000), 195-233.
2. G. Auchmuty, *Variational principles for variational inequalities*, Numerical Functional Analysis and Optimization, **10** (1989), 863-874.
3. A. Auslender, *Optimisation: Méthodes Numériques*, Masson, Paris, 1976.
4. E.J. Candès and B. Recht, *Exact matrix completion via convex optimization*, Foundations of Computational Mathematics, **9** (2008) 717-772.
5. C. Chen and O.L. Mangasarian, *A class of smoothing functions for nonlinear and mixed complementarity problems*, Computational Optimization and Applications, **5** (1996), 97-138.
6. J.-S. Chen and P. Tseng, *An unconstrained smooth minimization reformulation of the second-order cone complementarity problem*, Mathematical Programming, **104** (2005), 297-327.
7. X. Chen, H.-D. Qi and P. Tseng, *Analysis of nonsmooth symmetric matrix functions with applications to semidefinite complementarity problems*, SIAM Journal on Optimization, **13** (2003), 960-985.
8. C. Ding, D. Sun and K.-C. Toh, *An introduction to a class of matrix cone programming*, Preprint available at http://www.optimization-online.org/DB_HTML/2010/09/2746.html.
9. F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Volumes I and II, Springer-Verlag, New York, 2003.
10. L. Faybusovich, *Euclidean Jordan algebras and interior-point algorithm*, Positivity, **1** (1997), 331-357.
11. M. Fukushima, *Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems*, Mathematical Programming, **53** (1992), 99-110.

12. M. Fukushima, Z.-Q. Luo, and P. Tseng, *Smoothing functions for second-order cone complementarity problems*, SIAM Journal on Optimization, **12** (2002), 436-460.
13. D.R. Han, *On the coerciveness of some merit functions for complementarity problems over symmetric cones*, Journal of Mathematical Analysis and Applications, **336** (2007), 727-737.
14. D.W. Hearn, *The gap function of a convex program*, Operations Research Letters, **1** (1982), 67-71.
15. J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms*, Vols. 1 and 2, Springer-Verlag, Berlin, Heidelberg, New York, 1993.
16. C. Kanzow, I. Ferenczi and M. Fukushima, *On the local convergence of semismooth Newton methods for linear and nonlinear second-order cone programs without strict complementarity*, SIAM Journal on Optimization, **20** (2009), 297-320.
17. L.C. Kong, J. Sun and N.H. Xiu, *A regularized smoothing Newton method for symmetric cone complementarity problems*, SIAM Journal on Optimization, **19** (2009), 1028-1047.
18. L.C. Kong, L. Tunçel and N.H. Xiu, *Vector-valued implicit Lagrangian for symmetric cone complementarity problems*, Asia-Pacific Journal of Operational Research, **26** (2009), 199-233.
19. C. Lemaréchal and C. Sagastizábal, *Practical aspects of the Moreau-Yosida regularization I: theoretical preliminaries*, SIAM Journal on Optimization, **7** (1997), 367-385.
20. Y.-J. Liu, L.W. Zhang and Y.H. Wang, *Some properties of a class of merit functions for symmetric cone complementarity problems*, Asia-Pacific Journal of Operational Research, **23** (2006), 473-495.
21. M.S. Lobo, L. Vandenbergh, S. Boyd and H. Lebret, *Applications of second-order cone programming*, Linear Algebra and its Applications, **284** (1998), 193-228.
22. Z.-Q. Luo and P. Tseng, *A new class of merit functions for the nonlinear complementarity problem*, in: M.C. Ferris, J.-S. Pang (Eds.), *Complementarity and Variational Problems: State of the Art*, SIAM, Philadelphia, 1997, 204-225.
23. O.L. Mangasarian and M.V. Solodov, *Nonlinear complementarity as unconstrained and constrained minimization*, Mathematical Programming, **62** (1993), 277-297.
24. F. Meng, D. Sun and G. Zhao, *Semismoothness of solutions to generalized equations and the Moreau-Yosida regularization*, Mathematical Programming, **104** (2005), 561-581.
25. J.J. Moreau, *Décomposition orthogonale d'un espace hilbertien selon deux cônes mutuellement polaires*, Comptes Rendus de l'Académie des Sciences, **255** (1962), 238-240.
26. J.J. Moreau, *Proximité et dualité dans un espace hilbertien*, Bulletin de la Société Mathématique de France, **93** (1965), 273-299.
27. 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, **88** (2000), 61-83.
28. S.H. Pan, Y.-L. Chang, and J.-S. Chen, *Stationary point conditions for the FB merit function associated with symmetric cones*, Operations Research Letters, **38** (2010), 372-377.
29. S.-H. Pan and J.-S. Chen, *A damped Gauss-Newton method for the second-order cone complementarity problem*, Applied Mathematics and Optimization, **59** (2009), 293-318.
30. J.-S. Pang, S. Sun, and J. Sun, *Semismooth homeomorphisms and strong stability of semi-definite and Lorentz complementarity problems*, Mathematics of Operations Research, **28** (2003), 39-63.
31. L. Qi, D. Sun and G.L. Zhou, *A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities*, Mathematical Programming, **87** (2000), 1-35.
32. S. Schmieta and F. Alizadeh, *Associative and Jordan algebras, and polynomial time interior-point algorithms for symmetric cones*, Mathematics of Operations Research, **26** (2001), 543-564.

33. D. Sun and J. Sun, *Semismooth matrix valued functions*, Mathematics of Operations Research, **27** (2002), 150-169.
34. D. Sun and J. Sun, *Strong semismoothness of Fischer-Burmeister SDC and SOC function*, Mathematical Programming, **103** (2005), 575-581.
35. T. Tsuchiya, *A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming*, Optimization Methods and Software, **11** (1999), 141-182.
36. P. Tseng, *Merit functions for semidefinite complementarity problems*, Mathematical Programming, **83** (1998), 159-185.
37. K. Yosida, *Functional Analysis*, Springer Verlag, Berlin, 1964.
38. A. Yoshise, *Complementarity problems over symmetric cones: a survey of recent developments in several aspects*, Handbook on Semidefinite, Conic and Polynomial Optimization, **166** (2012), 339-375.
39. E.H. Zarantonello, *Projections on convex sets in Hilbert space and spectral theory I and II*, Contributions to Nonlinear Functional Analysis (E. H. Zarantonello, ed.), Academic Press, New York, 1971, 237-424.

Li Wang received her Ph.D. degree in 2011 from Dalian University of Technology. Her research interests include theory and applications of conic constrained optimization problems and conic complementarity problems.

School of Science, Shenyang Aerospace University, Shenyang, 110136, P.R. China.
e-mail: liwang211@163.com

Yong-Jin Liu received Ph.D. degree in Operation Research and Control Theory from Dalian University of Technology in 2004. He now is a professor of School of Science, Shenyang Aerospace University. His research interests cover theory, numerical computation and applications of conic constrained optimization problems and matrix programming problems.

School of Science, Shenyang Aerospace University, Shenyang, 110136, P.R. China.
e-mail: yjliu@sau.edu.cn

Yong Jiang got his Ph.D. degree from Dalian University of Technology in 2011. His research interests are theory and applications of risk optimization and conic optimization problems.

School of Science, Shenyang Aerospace University, Shenyang, 110136, P.R. China.
e-mail: jy5127@163.com