Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 11, No. 4, pp. 687–703, October 2015

ON THE INVERSE CONTINUOUS OPTIMIZATION AND ITS SMOOTHING FISCHER-BURMEISTER FUNCTION APPROACH*

Jie Gao, Hongwei Zhang, Xiantao Xiao and Liwei Zhang†

Abstract: This paper proposes a general inverse nonlinear optimization model in which parameters in both objective function and in constraints are required to be estimated. The inverse optimization model is reformulated as a mathematical programming problem with simple complementarity constraints. The tangent cone, normal cone of the feasible region of the inverse optimization problem are developed under mild conditions. First and second-order necessary optimality conditions as well as the second-order sufficient optimality conditions are derived. The smoothed Fischer-Burmeister function is used to construct a smoothing approach for solving the inverse nonlinear optimization problem. It is demonstrated that, when the positive smoothing parameter approaches to 0, the feasible set of the smoothing problem is convergent to the feasible set of the inverse problem, the global optimal value of the smoothing problem converges to that of the inverse problem, the outer limit of the solution mapping is contained in the solution set of the inverse problem, and the outer limit of the KKT-point mapping is contained in the set of Clarke stationary points associated with corresponding multipliers.

Key words: *inverse optimization, complementarity constraints, smoothing function*

Mathematics Subject Classification: *90C30*

---

1 Introduction

In an optimization problem, usually there are two parts of variables, one part consists of decision variables and the other consists of parameters. In practice, we are often facing to the instances, in which only some estimates for parameter values are known, but certain optimal solutions are available from experience, observations or experiments. An inverse optimization problem is to find values of parameters which make the known solutions optimal and which differ from the given estimates as little as possible.

Burton and Toint(1992)[4] first investigated an inverse shortest paths problem. Since then there are many important contributions to inverse optimization, and a large number of inverse combinatorial optimization problems have been studied, see the survey paper Heuberger (2004) [10] and the references Ahuja and Orlin (2001) [1],Ahuja and Orlin(2002) [2], Cai et al.(1999) [5], Zhang et al.(2000) [21], Zhang and Ma(1999)[22], etc.

*The research is supported by the National Natural Science Foundation of China under project No. 11301049, No. 91130007, No. 91330206 and Project funded by China Postdoctoral Science Foundation (2013M541217).

†Corresponding author

For continuous optimization, the first work is Zhang and Liu(1996) [19] for linear programming. After that there have been a series of papers on various types of inverse continuous optimization problems. People first studied inverse continuous optimization problems in which only parameters in the objective functions are required to be estimated. For such inverse optimization problems, Zhang and Liu(1999) [20] discusses the solution structure for some inverse linear programming problems; Iyengar and Kang(2005) [11] proposes the inverse conic programming model and discusses its applications; Xiao and Zhang (2009)[16] proposes a smoothing Newton method for solving the inverse QP problem in which the Hessian of the quadratic objective function is estimated, and for the same inverse quadratic programming problem Zhang and Zhang (2010) [23] studies the convergence properties for the augmented Lagrange method; Xiao, Zhang and Zhang (2009) [17] discusses the convergence of augmented Lagrange method for inverse semi-definite quadratic programming problems in which only the symmetric matrix in the objective function is required to be estimated; Xiao, Zhang and Zhang (2009) [18] proposes a smoothing Newton method for a type of inverse semi-definite quadratic programming problems.

Different from the above cited works, people also paid attention to the inverse optimization problems in which parameters in both objectives and constraints are required to be estimated. The first work in this direction is Zhang, Zhang and Xiao (2010) [24], in which an inexact Newton method is constructed to solve the KKT system to the smoothing dual problem for a type of inverse quadratic programming problems; Jiang et al. (2011) [8] proposes a perturbation approach for a type of inverse linear programming problems in which the smoothed Fischer-Burmeister function is employed; Zhang et al. (2013) [25] studies the similar smoothing approach for an inverse linear second-order cone programming, and Zhang et al. (2015) [26] studies a perturbation approach for an inverse quadratic programming problem over second-order cones.

In this paper, we consider the general mathematical programming problem of the form

$$(\mathrm{P}(\vartheta)) \qquad \begin{array}{ll} \min\limits_{x} & f(x,\vartheta) \\ \text{s.t.} & h(x,\vartheta)=0, \\ & g(x,\vartheta)\le 0, \end{array} \tag{1.1}$$

where $f:\Re^n\times Y\to\Re$, $h:\Re^n\times Y\to\Re^q$, $g:\Re^n\times Y\to\Re^p$ are continuously differentiable mappings, and Y is the space of parameters in problem functions, which is assumed to be a finitely dimensional Hilbert space.

Let Θ be a closed convex set of Y, which is the parameter set and ϑ is assumed to be an element of Θ.

The inverse nonlinear optimization problem is to find a vector ϑ^* solving

$$\begin{array}{ll} \min\limits_{\vartheta} & \mathcal{D}(\vartheta,\overline{\vartheta}) \\ \text{s.t.} & \overline{x}\in \mathrm{Sol}(\mathrm{P}(\vartheta)), \\ & \vartheta\in\Theta. \end{array} \tag{1.2}$$

The organization of this paper is as follows. In Section 2, the inverse optimization model is reformulated as a mathematical programming problem with simple complementarity constraints, and optimality conditions for this MPCC problem are developed. In Section 3, the convergence properties of the smoothed Fischer-Burmeister approach for for solving the inverse nonlinear optimization problem are investigated, in which it is demonstrated that, when the positive smoothing parameter approaches to 0, the outer limit of the solution mapping is contained in the solution set of the inverse problem, and the outer limit of the KKT-point mapping is contained in the set of Clarke stationary points associated with corresponding multipliers.

The inverse nonlinear optimization problem (1.2) is a bi-level problem. If Problem $\mathrm{P}(\vartheta)$ is not a convex optimization problem for $\vartheta \in \Theta$, then it is hard to characterize this inverse problem. For simplicity, we assume that Problem $\mathrm{P}(\vartheta)$ is a convex optimization problem. Under this assumption, the KKT conditions for Problem $\mathrm{P}(\vartheta)$ can be used to characterize its solutions. We propose the following assumptions for the functions in Problem (1.1) and Problem (1.2).

Assumption 1. Assume that for any $\vartheta \in \Theta$, $f(\cdot, \vartheta)$ and $g_i(\cdot, \vartheta), i = 1, \ldots, p$ are continuously differentiable functions and $h_j(\cdot, \vartheta), j = 1, \ldots, q$ are affine functions.

Assumption 2. Assume that for any $\vartheta \in \Theta$, there exist a feasible point x_ϑ to $\mathrm{P}(\vartheta)$ such that $g(x_\vartheta, \vartheta) < 0$.

For a function $p : \Re^n \to \Re$ and $q : \Re^n \to \Re^m$, $x \in \Re^n$, we use $\nabla p(x)$ and $\mathcal{J} q(x)$ to denote the gradient of p at x and the Jacobian of q at x. For a mapping $F : \mathcal{X} \to \mathcal{Y}$, where $\mathcal{X}$ and $\mathcal{Y}$ are finite dimensional Hilbert spaces, we use $\mathrm{D}F(x)$ to denote the derivative of F at x, which is a linear operator from $\mathcal{X}$ to $\mathcal{Y}$.

Under Assumption 1 and Assumption 2, $\mathrm{P}(\vartheta)$ is a convex optimization problem and Slater condition holds. In this case, $\overline{x} \in \mathrm{P}(\vartheta)$ is characterized by its KKT conditions and Problem (1.2) is equivalent to

$$\begin{array}{cl} \min\limits_{\vartheta,\mu,\lambda,z} & \mathcal{D}(\vartheta, \overline{\vartheta}) \\ \text{s.t.} & \nabla_x L(\overline{x}, \vartheta, \mu, \lambda) = 0, \\ & h(\overline{x}, \vartheta) = 0, \\ & g(\overline{x}, \vartheta) + z = 0, \\ & 0 \leq z \perp \lambda \geq 0, \\ & \vartheta \in \Theta. \end{array} \tag{1.1}$$

This is an MPEC problem because there exists a complementarity constraint $0 \leq z \perp \lambda \geq 0$.

Assumption 3. $\mathcal{D}(\vartheta, \overline{\vartheta}) \geq 0, \forall \vartheta \in \Theta$ with $\mathcal{D}(\overline{\vartheta}, \overline{\vartheta}) = 0$ and $\vartheta \to \mathcal{D}(\vartheta, \overline{\vartheta})$ is a strictly convex function.

Let

$$G(\vartheta, \mu, \lambda, z) = \begin{bmatrix} \nabla_x L(\overline{x}, \vartheta, \mu, \lambda) \\ h(\overline{x}, \vartheta) \\ g(\overline{x}, \vartheta) + z \end{bmatrix}. \tag{1.2}$$

Then Problem (1.1) is expressed as

$$\begin{array}{cl} \min\limits_{\vartheta,\mu,\lambda.z} & \mathcal{D}(\vartheta, \overline{\vartheta}) \\ \text{s.t.} & G(\vartheta, \mu, \lambda, z) = 0, \\ & (\vartheta, \mu, \lambda, z) \in \Theta \times \Re^q \times \Omega, \end{array} \tag{1.3}$$

where

$$\Omega = \{(a, b) \in \Re^p \times \Re^p : 0 \leq a \perp b \geq 0\}.$$

We use Φ to denote the feasible set for Problem (1.3), namely

$$\Phi = \{(\vartheta, \mu, \lambda, z) \in \Theta \times \Re^q \times \Omega : G(\vartheta, \mu, \lambda, z) = 0\}.$$

Proposition 1.1. *Under Assumption 1 and Assumption 3, there is an optimal solution to Problem (1.2).*

If ϑ is a vector and Θ is a convex polyhedral set, we are able to apply the well-known first and second order optimality conditions results directly on Problem (1.1), for instance we may use the first and second order optimality conditions for MPCCs in Chapter 3 and Chapter 5 of Luo, Pang and Ralph (1996) [13], respectively. But if ϑ is not a vector, for example it is a matrix, we give the first and second-order optimality conditions for MPCC (1.1). For this purpose, we first present the first variational geometry of the feasible set of Problem (1.1), which is used naturally to develop the first order necessary optimality conditions for Problem (1.1).

The tangent cone of Φ at $(\vartheta,\mu,\lambda,z)$ denoted by $T_\Phi(\vartheta,\mu,\lambda,z)$, the regular normal cone of Φ at $(\vartheta,\mu,\lambda,z)$ denoted by $\widehat{N}_\Phi(\vartheta,\mu,\lambda,z)$ and the normal cone of Φ at $(\vartheta,\mu,\lambda,z)$ denoted by $N_\Phi(\vartheta,\mu,\lambda,z)$, are defined respectively by

$$T_\Phi(\vartheta,\mu,\lambda,z)=\left\{(d_\vartheta,d_\mu,d_\lambda,d_z):\begin{array}{c}\exists t_k\searrow 0,\exists(d_\vartheta^k,d_\mu^k,d_\lambda^k,d_z^k)\to(d_\vartheta,d_\mu,d_\lambda,d_z)\\ \text{satisfying } (\vartheta,\mu,\lambda,z)+t_k(d_\vartheta^k,d_\mu^k,d_\lambda^k,d_z^k)\in\Phi\end{array}\right\};$$
$$\widehat{N}_\Phi(\vartheta,\mu,\lambda,z)=\left\{(v_\vartheta,v_\mu,v_\lambda,v_z):\begin{array}{l}\langle(v_\vartheta,v_\mu,v_\lambda,v_z),(\vartheta',\mu',\lambda',z')-(\vartheta,\mu,\lambda,z)\rangle\\ \le o(\|(\vartheta',\mu',\lambda',z')-(\vartheta,\mu,\lambda,z)\|),\,(\vartheta',\mu',\lambda',z')\in\Phi\end{array}\right\};$$
$$N_\Phi(\vartheta,\mu,\lambda,z)=\left\{(v_\vartheta,v_\mu,v_\lambda,v_z):\begin{array}{c}\exists(\vartheta^k,\mu^k,\lambda^k,z^k)\xrightarrow{\Phi}\\ (\vartheta,\mu,\lambda,z),\exists(v_\vartheta^k,v_\mu^k,v_\lambda^k,v_z^k)\to(v_\vartheta,v_\mu,v_\lambda,v_z)\\ \text{satisfying } (v_\vartheta^k,v_\mu^k,v_\lambda^k,v_z^k)\in\widehat{N}_\Phi(\vartheta^k,\mu^k,\lambda^k,z^k)\end{array}\right\}.$$

Let $\omega=\{(\zeta_1,\zeta_2)\in\Re_+^2:\zeta_1\zeta_2=0\}$. For Ω with complementarity constraints, we have the following lemma about the variational geometry of Ω at a point $(\bar a,\bar b)\in\Omega$.

Lemma 1.2. *For $(\bar a,\bar b)\in\Omega$, the tangent cone, the regular normal cone and normal cone of Ω at $(\bar a,\bar b)$ are calculated by*

$$T_\Omega(\bar a,\bar b)=\bigotimes_{i=1}^p T_\omega(\bar a_i,\bar b_i),\ \widehat{N}_\Omega(\bar a,\bar b)=\bigotimes_{i=1}^p \widehat{N}_\omega(\bar a_i,\bar b_i)\ \textit{and}\ N_\Omega(\bar a,\bar b)=\bigotimes_{i=1}^p N_\omega(\bar a_i,\bar b_i),$$

where

$$\bigotimes_{i=1}^p T_\Omega(\bar a_i,\bar b_i)=\left\{(u,v)|\,(u_i,v_i)\in T_\omega(\bar a_i,\bar b_i),\ i=1,\ldots,p\right\},$$

$$\bigotimes_{i=1}^p \widehat{N}_\Omega(\bar a_i,\bar b_i)=\left\{(u,v)|\,(u_i,v_i)\in \widehat{N}_\omega(\bar a_i,\bar b_i),\ i=1,\ldots,p\right\},$$

$$\bigotimes_{i=1}^p N_\omega(\bar a_i,\bar b_i)=\left\{(u,v)|\,(u_i,v_i)\in N_\omega(\bar a_i,\bar b_i),\ i=1,\ldots,p\right\},$$

$$T_\omega(\bar a_i,\bar b_i)=\begin{cases}\Re\times\{0\}, & \textit{if } a_i>0,\ b_i=0,\\ \{0\}\times\Re, & \textit{if } a_i=0,\ b_i>0,\\ \omega, & \textit{if } a_i=0,\ b_i=0,\end{cases}$$

$$\widehat{N}_\omega(\bar{a}_i,\bar{b}_i)=\begin{cases}\{0\}\times\Re,, & \textit{if } a_i>0,\ b_i=0,\\ \Re\times\{0\} & \textit{if } a_i=0,\ b_i>0,\\ \Re_-\times\Re_-, & \textit{if } a_i=0,\ b_i=0,\end{cases}$$

$$N_\omega(\bar{a}_i,\bar{b}_i)=\begin{cases}\{0\}\times\Re, & \textit{if } a_i>0,\ b_i=0,\\ \Re\times\{0\}, & \textit{if } a_i=0,\ b_i>0,\\ (\Re\times\{0\})\bigcup(\{0\}\times\Re)\bigcup(\Re_-\times\Re_-), & \textit{if } a_i=0,\ b_i=0.\end{cases}$$

For deriving the tangent cone, the regular normal cone and the normal cone of Φ at $(\vartheta,\mu,\lambda,z)\in\Phi$, we need the following assumption:
Assumption 4 We say that the constraint non-degeneracy condition is satisfied at (ϑ,μ,λ) with $\vartheta\in\Theta$ if the linear operator

$$\begin{bmatrix}\mathrm{D}_\vartheta\nabla_x L(\overline{x},\vartheta,\mu,\lambda) & \nabla_x h(\overline{x},\vartheta) & \nabla_x g(\overline{x},\vartheta)\\ \mathrm{D}_\vartheta h(\overline{x},\vartheta) & 0 & 0\end{bmatrix}$$

is onto.

Assumption 4 is satisfied when $\mathrm{P}(\vartheta)$ is a linear programming problem, quadratic programming problem, linear and quadratic second-order optimization problems or linear and quadratic semi-definite optimization problems.

Proposition 1.3. *Assume Assumption 1 and Assumption 3 hold, and Assumption 4 is satisfied at (ϑ,μ,λ) with $\vartheta\in\Theta$, then $\mathrm{D}G(\vartheta,\mu,\lambda,z)$ is onto. In this case*

$$T_\Phi(\vartheta,\mu,\lambda,z)=\left\{d\in Y\times\Re^q\times\Re^p\times\Re^p:\begin{array}{l}\mathrm{D}_\vartheta\nabla_x L(\overline{x},\vartheta,\vartheta\mu,\lambda)d_\vartheta+\nabla_x h(\overline{x},\vartheta)d_\mu\\ \qquad+\nabla_x g(\overline{x},\vartheta)d_\lambda=0\\ \mathrm{D}_\vartheta h(\overline{x},\vartheta)d_\vartheta=0\\ \mathrm{D}_\vartheta g(\overline{x},\vartheta)d_\vartheta+d_z=0\\ d_\vartheta\in T_\Theta(\vartheta)\\ (d_\lambda,d_z)\in T_\Omega(\lambda,z)\end{array}\right\},\tag{1.4}$$

$$\widehat{N}_\Phi(\vartheta,\mu,\lambda,z)=\left\{\begin{pmatrix}\mathrm{D}_\vartheta\nabla_x L(\overline{x},\vartheta,\mu,\lambda)^*\eta_1+\mathrm{D}_\vartheta h(\overline{x},\vartheta)^*\eta_2\\ \qquad+\mathrm{D}_\vartheta g(\overline{x},\vartheta)^*\eta_3+\widehat{N}_\Theta(\vartheta)\\ \mathcal{J}_x h(\overline{x},\vartheta)\eta_1\\ \mathcal{J}_x g(\overline{x},\vartheta)\eta_1+\xi_a\\ \eta_3+\xi_b\end{pmatrix}:\begin{array}{l}(\eta_1,\eta_2,\eta_3)\in\Re^{n+q+p}\\ (\xi_a,\xi_b)\in\widehat{N}_\Omega(\lambda,z)\end{array}\right\}\tag{1.5}$$

and

$$N_\Phi(\vartheta,\mu,\lambda,z)=\left\{\begin{pmatrix}\mathrm{D}_\vartheta\nabla_x L(\overline{x},\vartheta,\mu,\lambda)^*\eta_1+\mathrm{D}_\vartheta h(\overline{x},\vartheta)^*\eta_2\\ \qquad+\mathrm{D}_\vartheta g(\overline{x},\vartheta)^*\eta_3+N_\Theta(\vartheta)\\ \mathcal{J}_x h(\overline{x},\vartheta)\eta_1\\ \mathcal{J}_x g(\overline{x},\vartheta)\eta_2+\xi_a\\ \eta_3+\xi_b\end{pmatrix}:\begin{array}{l}(\eta_1,\eta_2,\eta_3)\in\Re^{n+q+p}\\ (\xi_a,\xi_b)\in N_\Omega(\lambda,z)\end{array}\right\}\tag{1.6}$$

Proof. From the definition of G, one has that

$$\mathrm{D}G(\vartheta,\mu,\lambda,z)=\begin{bmatrix}\mathrm{D}_\vartheta\nabla_x L(\overline{x},\vartheta,\mu,\lambda) & \nabla_x h(\overline{x},\vartheta) & \nabla_x g(\overline{x},\vartheta) & 0\\ \mathrm{D}_\vartheta h(\overline{x},\vartheta) & 0 & 0 & 0\\ \mathrm{D}_\vartheta g(\overline{x},\vartheta) & 0 & 0 & I\end{bmatrix}.\tag{1.7}$$

Let $(\xi_1, \xi_2, \xi_3) \in \Re^n \times \Re^q \times \Re^p$. Since Assumption 4 is satisfied at $(\vartheta, \mu, \lambda)$, there exists $(z_1, z_2, z_3) \in Y \times \Re^q \times \Re^p$ such that

$$\begin{bmatrix} \mathrm{D}_\vartheta \nabla_x L(\overline{x}, \vartheta, \mu, \lambda) & \nabla_x h(\overline{x}, \vartheta) & \nabla_x g(\overline{x}, \vartheta) \\ \mathrm{D}_\vartheta h(\overline{x}, \vartheta) & 0 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}.$$

Let $z_4 = -\mathrm{D}_\vartheta g(\overline{x}, \vartheta) z_1$, then one has $(z_1, z_2, z_3, z_4) \in Y \times \Re^q \times \Re^p \times \Re^p$ such that

$$\mathrm{D}G(\vartheta, \mu, \lambda, z)(z_1, z_2, z_3, z_4) = (\xi_1, \xi_2, \xi_3),$$

which means that $\mathrm{D}G(\vartheta, \mu, \lambda, z)$ is onto because of the arbitrariness of (ξ_1, ξ_2, ξ_3).

Since $\mathrm{D}G(\vartheta, \mu, \lambda, z)$ is onto, we now prove the following equality: (the similar result like 6.7 Exercise of [14])

$$T_\Phi(\vartheta, \mu, \lambda, z) = \left\{ d \in T_\Theta(\vartheta) \times \Re^q \times T_\Omega(\lambda, z) : \mathrm{D}G(\vartheta, \mu, \lambda, z) d = 0 \right\}. \tag{1.8}$$

It is obvious that the set in the left hand-side is contained in the right hand-side, so we only need to prove the opposite inclusion. For any $d = (d_\vartheta, d_\mu, d_\lambda, d_z)$ satisfying $d \in T_\Theta(\vartheta) \times \Re^q \times T_\Omega(\lambda, z), \mathrm{D}G(\vartheta, \mu, \lambda, z) d = 0$, one has that there exist $d^k = (d^k_\vartheta, d^k_\mu, d^k_\lambda, d^k_z) \to d$ and $t_k \searrow 0$ such that $(\vartheta, \mu, \lambda, z) + t_k d^k \in \Theta \times \Re^q \times \Omega$. It follows from Lemma 1.2 that $[d_\lambda]_i [d_z]_i = 0$ for $i = 1, \ldots, p$. Let

$$\alpha = \{i : \lambda_i > 0, z_i = 0\}, \beta = \{i : \lambda_i = z_i = 0\}, \gamma = \{i : \lambda_i = 0, z_i > 0\}$$

and

$$\begin{aligned} \beta_a &= \{i \in \beta : [d_\lambda]_i > 0, [d_z]_i = 0\}, \\ \beta_b &= \{i \in \beta : [d_\lambda]_i = [d_z]_i = 0\}, \\ \beta_c &= \{i \in \beta : [d_\lambda]_i = 0, [d_z]_i > 0\}. \end{aligned}$$

Let

$$\Gamma_d = \left\{ (\lambda, z) \in \Re^p \times \Re^p : \begin{array}{l} (\lambda_{\alpha \cup \beta_a}, z_{\alpha \cup \beta_a}) \in \Re_+^{|\alpha| + |\beta_a|} \times \{0_{|\alpha| + |\beta_a|}\} \\ (\lambda_{\beta_c \cup \gamma}, z_{\beta_c \cup \gamma}) \in \{0_{|\beta_c| + |\gamma|}\} \times \Re_+^{|\beta_c| + |\gamma|} \\ (\lambda_{\beta_b}, z_{\beta_b}) = (0_{|\beta_b|}, 0_{|\beta_b|}) \end{array} \right\}.$$

Then Γ_d is a convex set and $\Gamma_d \subset \Omega$. Since $\mathrm{D}G(\vartheta, \mu, \lambda, z)$ is onto, it follows from Theorem 2.87 of [3] that there exist a neighborhood $\mathcal{V}$ of $(\vartheta, \mu, \lambda, z)$ and a positive constant κ such that

$$\begin{aligned} &\mathrm{dist}\left((\vartheta', \mu', \lambda', z'), [\Theta \times \Re^q \times \Gamma_d] \cap G^{-1}(0)\right) \\ &\qquad \le \kappa \| G(\vartheta', \mu', \lambda', z'), \Pi_{\Theta \times \Re^q \times \Gamma_d}(\vartheta', \mu', \lambda', z') \|, \; (\vartheta', \mu', \lambda', z') \in \mathcal{V}. \end{aligned}$$

Noticing that for $(\vartheta^k, \mu^k, \lambda^k, z^)) = (\vartheta, \mu, \lambda, z) + t_k d^k$, $\vartheta^k \in \Theta$, and

$$\begin{aligned} &(\lambda^k_{\alpha \cup \beta_a}, z^k_{\alpha \cup \beta_a}) \in \Re_+^{|\alpha| + |\beta_a|} \times \{0_{|\alpha| + |\beta_a|}\} \\ &(\lambda^k_{\beta_c \cup \gamma}, z^k_{\beta_c \cup \gamma}) \in \{0_{|\beta_c| + |\gamma|}\} \times \Re_+^{|\beta_c| + |\gamma|} \\ &(\lambda^k_{\beta_b}, z^k_{\beta_b}) = (t_k [d^k_\lambda]_{\beta_b}, t_k [d^k_z]_{\beta_b}) = \mathrm{o}(t_k), \end{aligned}$$

we have that

$$
\begin{aligned}
&\operatorname{dist}((\vartheta,\mu,\lambda,z)+t_k d^k,\Phi)=\operatorname{dist}((\vartheta,\mu,\lambda,z)+t_k d^k,[\Theta\times\Re^q\times\Omega]\cap G^{-1}(0))\\
&=\operatorname{dist}((\vartheta,\mu,\lambda,z)+t_k d^k,[\Theta\times\Re^q\times\Gamma_d]\cap G^{-1}(0))\\
&\le\kappa[\|(t_k[d^k_\lambda]_{\beta_b},t_k[d^k_z]_{\beta_b})\|+\|G(\vartheta^k,\mu^k,\lambda^k,z^k)\|]\\
&=\kappa\|G(\vartheta,\mu,\lambda,z)+t_k\mathrm{D}G(\vartheta,\mu,\lambda,z)d^k+\textstyle\int_0^1[\mathrm{D}G((\vartheta,\mu,\lambda,z)+st_kd^k)-\mathrm{D}G(\vartheta,\mu,\lambda,z)]\mathrm{d}st_kd^k]\|\\
&\quad+\kappa\|(t_k[d^k_\lambda]_{\beta_b},t_k[d^k_z]_{\beta_b})\|=\mathrm{o}(t_k),
\end{aligned}
$$

which implies that $d\in T_\Phi(\vartheta,\mu,\lambda,z)$. Therefore we obtain equality (1.8).

Combining with (1.7) and (1.8), we obtain (1.4).

Since $\mathrm{D}G(\vartheta,\mu,\lambda,z)$ is onto, formula (1.5) comes from the equality

$$\widehat{N}_\Phi(\vartheta,\mu,\lambda,z)=\mathrm{D}G(\vartheta,\mu,\lambda,z)^*(\Re^n\times\Re^q\times\Re^p)+\widehat{N}_{\Theta\times\Re^q\times\Omega}(\vartheta,\mu,\lambda,z)$$

and

$$\widehat{N}_{\Theta\times\Re^q\times\Omega)}(\vartheta,\mu,\lambda,z))=\widehat{N}_\Theta(\vartheta)\times 0_q\times\widehat{N}_\Omega(\lambda,z).$$

Formula (1.6) can be established in the same way, as when $\mathrm{D}G(\vartheta,\mu,\lambda,z)$ is onto one has

$$N_\Phi(\vartheta,\mu,\lambda,z)=\mathrm{D}G(\vartheta,\mu,\lambda,z)^*(\Re^n\times\Re^q\times\Re^p)+N_{\Theta\times\Re^q\times\Omega}(\vartheta,\mu,\lambda,z))$$

and

$$N_{\Theta\times\Re^q\times\Omega}(\vartheta,\mu,\lambda,z))=N_\Theta(\vartheta)\times 0_q\times N_\Omega(\lambda,z).$$

The proof is completed. □

From the above lemma, we can easily develop the necessary optimality conditions for a local minimizer of Problem (1.3).

Theorem 1.4. *Let $(\vartheta^*,\mu^*,\lambda^*,z^*)$ be a local minimizer of Problem (1.3). Let Assumptions 1, 2, 3 hold and Assumption 4 be satisfied at $(\vartheta^*,\mu^*,\lambda^*)$. Then $z^*=-g(\overline{x},\vartheta^*)$ and there exist $\eta_1\in\Re^n,\eta_2\in\Re^q,[\eta_3]_{\beta\cup\gamma}\in\Re^{|\beta|+|\gamma|}$ such that*

$$
\begin{aligned}
&\mathrm{D}_\vartheta\mathcal{D}(\vartheta^*,\overline{\vartheta})+\mathrm{D}_\vartheta\nabla_xL(\overline{x},\vartheta^*,\mu^*,\lambda^*)^*\eta_1+\mathrm{D}_\vartheta h(\overline{x},\vartheta^*)^*\eta_2+\mathrm{D}_\vartheta g_{\beta\cup\gamma}(\overline{x},\vartheta^*)^*[\eta_3]_{\beta\cup\gamma}=0,\\
&\mathcal{J}_xh(\overline{x},\vartheta^*)\eta_1=0,\\
&\mathcal{J}_xg_\gamma(\overline{x},\vartheta^*)\eta_2=0_{|\gamma|},\\
&\mathcal{J}_xg_\beta(\overline{x},\vartheta^*)\eta_2\ge 0_{|\beta|},\\
&[\eta_3]_\beta\ge 0_{|\beta|},
\end{aligned}
\tag{1.9}
$$

where

$$\alpha=\{i:\lambda^*_i=0<z^*_i\},\ \beta=\{i:\lambda^*_i=0=z^*_i\},\ \gamma=\{i:\lambda^*_i>0=z^*_i\}.$$

Proof. It comes from the inclusion

$$0\in\mathrm{D}_{\vartheta,\mu,\lambda,z}\mathcal{D}(\vartheta^*,\overline{\vartheta})+\widehat{N}_\Phi(\vartheta^*,\mu^*,\lambda^*,z^*),$$

where $\widehat{N}_\Phi(\vartheta^*,\mu^*,\lambda^*,z^*)$ is from Lemma 1.2. □

Remark 1.5. Let

$$F(\vartheta,\mu,\lambda,z)=\left[\begin{array}{c}\nabla_x L(\overline{x},\vartheta,\mu,\lambda)\\ h(\overline{x},\vartheta)\\ g(\overline{x},\vartheta)+z\\ \min(\lambda,z)\end{array}\right]. \tag{1.10}$$

Then Problem (1.1) is expressed as

$$\begin{array}{ll}\min\limits_{\vartheta} & \mathcal{D}(\vartheta,\overline{\vartheta})\\ \text{s.t.} & F(\vartheta,\mu,\lambda,z)=0,\\ & \vartheta\in\Theta.\end{array} \tag{1.11}$$

Noticing that F is a Lipschitz continuous mapping, Problem (1.11) is a Lipschitz continuous optimization problem. So we may use the optimality conditions for Lipschitz continuous optimization developed in Clarke (1983). This leads to the so-called C-stationary point. We say that the point $(\vartheta^*,\mu^*,\lambda^*,z^*)$ is a C-stationary point if there exist $\eta_1\in\Re^n,\eta_2\in\Re^q,[\eta_3]_{\beta\cup\gamma}\in\Re^{|\beta|+|\gamma|}$ such that

$$\begin{array}{l}\mathrm{D}_\vartheta\mathcal{D}(\vartheta^*,\overline{\vartheta})+\mathrm{D}_\vartheta\nabla_x L(\overline{x},\vartheta^*,\mu^*,\lambda^*)^*\eta_1+\mathrm{D}_\vartheta h(\overline{x},\vartheta^*)^*\eta_2+\mathrm{D}_\vartheta g_{\beta\cup\gamma}(\overline{x},\vartheta^*)^*[\eta_3]_{\beta\cup\gamma}=0,\\ \mathcal{J}_x h(\overline{x},\vartheta^*)\eta_1=0,\\ \mathcal{J}_x g_\gamma(\overline{x},\vartheta^*)\eta_2=0_{|\gamma|},\\ \mathcal{J}_x g_i(\overline{x},\vartheta^*)\eta_2[\eta_3]_i\geq 0 \text{ for } i\in\beta,\end{array} \tag{1.12}$$

We say that the point $(\vartheta^*,\mu^*,\lambda^*,z^*)$ is an M-stationary point if there exist $\eta_1\in\Re^n,\eta_2\in\Re^q,[\eta_3]_{\beta\cup\gamma}\in\Re^{|\beta|+|\gamma|}$ such that

$$\begin{array}{l}\mathrm{D}_\vartheta\mathcal{D}(\vartheta^*,\overline{\vartheta})+\mathrm{D}_\vartheta\nabla_x L(\overline{x},\vartheta^*,\mu^*,\lambda^*)^*\eta_1+\mathrm{D}_\vartheta h(\overline{x},\vartheta^*)^*\eta_2+\mathrm{D}_\vartheta g_{\beta\cup\gamma}(\overline{x},\vartheta^*)^*[\eta_3]_{\beta\cup\gamma}=0,\\ \mathcal{J}_x h(\overline{x},\vartheta^*)\eta_1=0,\\ \mathcal{J}_x g_\gamma(\overline{x},\vartheta^*)\eta_2=0_{|\gamma|},\\ \mathcal{J}_x g_i(\overline{x},\vartheta^*)\eta_2[\eta_3]_i=0 \text{ or } \mathcal{J}_x g_i(\overline{x},\vartheta^*)\eta_2>0 \text{ and } [\eta_3]_i>0 \text{ for } i\in\beta\end{array} \tag{1.13}$$

It follows from Theorem 1.4, under Assumption 4, the point $(\vartheta^*,\mu^*,\lambda^*,z^*)$ is a strong stationary point of Problem (1.3). Thus, $(\vartheta^*,\mu^*,\lambda^*,z^*)$ is an M-stationary point, also a C-stationary point of Problem (1.3).

Now we discuss the second-order optimality conditions. Let $\omega\subset\beta$, $\omega^c=\beta\setminus\omega$, $\alpha(\omega)=\alpha\cup\omega$ and $\gamma(\omega)=\omega^c\cup\gamma$, consider the following problem

$$\begin{array}{ll}\min\limits_{\vartheta,\mu,\lambda.z} & \mathcal{D}(\vartheta,\overline{\vartheta})\\ \text{s.t.} & G(\vartheta,\mu,\lambda,z)=0,\\ & \vartheta\in\Theta,\\ & \left\{\begin{array}{l}\lambda_{\alpha(\omega)}=0,\\ z_{\alpha(\omega)}\geq 0,\end{array}\right.\\ & \left\{\begin{array}{l}\lambda_{\gamma(\omega)}\geq 0,\\ z_{\gamma(\omega)}=0,\end{array}\right.\end{array} \tag{1.14}$$

Define

$$H_\omega(\overline{x},\vartheta,\mu,\lambda_{\gamma(\omega)})=\nabla_x f(\overline{x},\vartheta)+\nabla_x h(\overline{x},\vartheta)\mu+\nabla_x g_{\gamma(\omega)}(\overline{x},\vartheta)\lambda_{\gamma(\omega)},$$

then Problem (1.14) is equivalent to

$$\begin{array}{rl}\min\limits_{\vartheta,\mu,\lambda_{\gamma(\omega)}} & \mathcal{D}(\vartheta,\overline{\vartheta})\\ \text{s.t.} & H_\omega(\overline{x},\vartheta,\mu,\lambda_{\gamma(\omega)})=0,\\ & g_{\alpha(\omega)}(\overline{x},\vartheta)\le 0,\\ & g_{\gamma(\omega)}(\overline{x},\vartheta)=0,\\ & \vartheta\in\Theta,\ \lambda_{\gamma(\omega)}\ge 0.\end{array} \tag{1.15}$$

It follows from Subsection 2.3.4 of [3], that Robinson constraint qualification for Problem (1.15), denoted by CQ(ω), at $(\vartheta^*,\mu^*,\lambda^*_{\gamma(\omega)})$, can be written as

(i) The mapping $\mathrm{D}_{\vartheta,\mu,\lambda_{\gamma(\omega)}}\begin{bmatrix} H_\omega(\overline{x},\vartheta^*,\mu^*,\lambda^*_{\gamma(\omega)})\\ g_{\gamma(\omega)}(\overline{x},\vartheta^*)\end{bmatrix}$ is onto.

(ii) There exists $d^0=(d^0_\vartheta,d^0_\mu,d^0_{\lambda_{\gamma(\omega)}})$ such that

$$\mathrm{D}_{\vartheta,\mu,\lambda_{\gamma(\omega)}}\begin{bmatrix} H_\omega(\overline{x},\vartheta^*,\mu^*,\lambda^*_{\gamma(\omega)})\\ g_{\gamma(\omega)}(\overline{x},\vartheta^*)\end{bmatrix}d^0=0$$

and

$$g_{\alpha(\omega)}(\overline{x},\vartheta^*)+\mathrm{D}_\vartheta g_{\alpha(\omega)}(\overline{x},\vartheta^*)d^0_\vartheta<0,\vartheta^*+d^0_\vartheta\in\operatorname{int}\Theta,\lambda^*_{\gamma(\omega)}+d^0_{\lambda_{\gamma(\omega)}}<0.$$

Let Φ_ω be the feasible set of Problem (1.15). If CQ(ω) holds at $(\vartheta^*,\mu^*,\lambda^*_{\gamma(\omega)})$, then the tangent cone of Φ_ω at $(\vartheta^*,\mu^*,\lambda^*_{\gamma(\omega)})$ is expressed as

$$T_{\Phi_\omega}(\vartheta^*,\mu^*,\lambda^*_{\gamma(\omega)})=\left\{d:\begin{array}{l}\mathrm{D}_{\vartheta,\mu,\lambda_{\gamma(\omega)}}H_\omega(\overline{x},\vartheta^*,\mu^*,\lambda^*_{\gamma(\omega)})d=0\\ \mathrm{D}_\vartheta g_{\gamma(\omega)}(\overline{x},\vartheta^*)d_\vartheta=0\\ \mathrm{D}_\vartheta g_{\alpha(\omega)}(\overline{x},\vartheta^*)d_\vartheta\in T_{\Re_-^{|\alpha(\omega)|}}(g_{\alpha(\omega)}(\overline{x},\vartheta^*))\\ d_\vartheta\in T_\Theta(\vartheta^*),\ d_{\lambda_{\gamma(\omega)}}\in T_{\Re_+^{|\gamma(\omega)|}}(\lambda^*_{\gamma(\omega)})\end{array}\right\}.$$

The critical cone of Problem (1.15) at $(\vartheta^*,\mu^*,\lambda^*_{\gamma(\omega)})$ is

$$\mathcal{C}_\omega(\vartheta^*,\mu^*,\lambda^*_{\gamma(\omega)})=\{d\in T_{\Phi_\omega}(\vartheta^*,\mu^*,\lambda^*_{\gamma(\omega)}):\mathrm{D}_\vartheta\mathcal{D}(\vartheta^*,\overline{\vartheta})d_\vartheta\le 0\}.$$

Let $L_\omega: Y\times\Re^q\times\Re^{|\gamma(\omega)|}\times\Re^n\times\Re^{|\alpha(\omega)|}\times\Re^{|\gamma(\omega)|}\times\Re^{|\gamma(\omega)|}\times Y\to\Re$ be the Lagrangian for Problem (1.15):

$$\begin{aligned}&L_\omega(\vartheta,\mu,\lambda_{\gamma(\omega)},\zeta,\xi,)\\ &\quad=\mathcal{D}(\vartheta,\overline{\vartheta})+H_\omega(\overline{x},\vartheta,\mu,\lambda_{\gamma(\omega)})\zeta_1+g_{\alpha(\omega)}(\overline{x},\vartheta)^T\zeta_2+g_{\gamma(\omega)}(\overline{x},\vartheta)^T\zeta_3-\lambda^T_{\gamma(\omega)}\zeta_4+\langle\vartheta,\xi\rangle.\end{aligned}$$

where $\zeta=(\zeta_1,\zeta_2,\zeta_3,\zeta_4)\in\Re^n\times\Re^{|\alpha(\omega)|}\times\Re^{|\gamma(\omega)|}\times\Re^{|\gamma(\omega)|}$.

As an example, we consider the second-order optimality conditions for Problem (1.3) when $\Theta=\mathcal{S}^l_+$.

Theorem 1.6. *Let $\Theta=\mathcal{S}^l_+$. Let $(\vartheta^*,\mu^*,\lambda^*,z^*)$ be a local minimizer of Problem (1.3) and Assumptions 1, 2, 3 hold. Suppose that, for every $\omega\subset\beta$, CQ(ω), at $(\vartheta^*,\mu^*,\lambda^*_{\gamma(\omega)})$. Then, for each $\omega\subset\beta$, the set of Lagrange multipliers of (1.15)*

$$\Lambda_\omega(\vartheta^*,\mu^*,\lambda^*_{\gamma(\omega)}) = \{(\zeta,\xi) : \mathrm{D}_{\vartheta,\mu,\lambda_{\gamma(\omega)}} L_\omega(\vartheta^*,\mu^*,\lambda^*_{\gamma(\omega)},\zeta,\xi) = 0,$$
$$0 \le \zeta_2 \perp g_{\alpha(\omega)}(\overline{x},\vartheta^*), 0 \le \zeta_4 \perp \lambda^*_{\gamma(\omega)}, \mathcal{S}^l_+ \ni \xi \perp \vartheta^*\}$$

is nonempty and compact. And for $\forall d \in \mathcal{C}_\omega(\vartheta^*,\mu^*,\lambda^*_{\gamma(\omega)})$*,*

$$\sup_{(\zeta,\xi)\in\Lambda_\omega(\vartheta^*,\mu^*,\lambda^*_{\gamma(\omega)})} \left\{\mathrm{D}^2_{\vartheta,\mu,\lambda_{\gamma(\omega)}} L_\omega(\vartheta^*,\mu^*,\lambda^*_{\gamma(\omega)},\zeta,\xi)(d,d) + 2\langle\xi, d_\vartheta[\vartheta^*]^\dagger d_\vartheta\rangle\right\} \quad \ge \quad 0,$$

Theorem 1.7. *Let* $\Theta = \mathcal{S}^l_+$*. Let* $(\vartheta^*,\mu^*,\lambda^*,z^*)$ *be a feasible point of Problem (1.3) and Assumptions 1, 2, 3 hold. Suppose that, for every* $\omega \subset \beta$*, the set of Lagrange multipliers of (1.15)* $\Lambda_\omega(\vartheta^*,\mu^*,\lambda^*_{\gamma(\omega)})$ *is nonempty. And for* $\forall d \in \mathcal{C}_\omega(\vartheta^*,\mu^*,\lambda^*_{\gamma(\omega)}) \setminus \{0\}$*,*

$$\sup_{(\zeta,\xi)\in\Lambda_\omega(\vartheta^*,\mu^*,\lambda^*_{\gamma(\omega)})} \left\{\mathrm{D}^2_{\vartheta,\mu,\lambda_{\gamma(\omega)}} L_\omega(\vartheta^*,\mu^*,\lambda^*_{\gamma(\omega)},\zeta,\xi)(d,d) + 2\langle\xi, d_\vartheta[\vartheta^*]^\dagger d_\vartheta\rangle\right\} > 0.$$

Then the second-order growth condition holds at $(\vartheta^*,\mu^*,\lambda^*)$*.*

2 The smoothed Fischer-Burmeister function approach

Problem (1.1) is an MPEC problem, for such a problem, it is not suitable to treat it as a traditional NLP problem because, as explained in [13, Example 3.1.1 and Example 3.1.2], even the basic constraint qualification (namely the tangent cone is equal to the linearized cone at an optimal solution) does not hold. To overcome this difficulty, various relaxation approaches have been proposed dealing with the complementarity constraints. Facchinei et al.(1999) [7] and Fukushima and Pang(1999) [9] used $\psi_\varepsilon(a,b) = 0$ to approximate the complementarity relation $0 \le a, 0 \le b, ab = 0$, where $\psi_\varepsilon(a,b)$ is the smoothed Fischer-Burmeister function

$$\psi_\varepsilon(a,b) = a + b - \sqrt{a^2+b^2+2\varepsilon^2}. \tag{2.1}$$

Scholtes(2001)[15] used

$$a \ge 0, b \ge 0, ab \le \varepsilon,$$

and Lin and Fukushima (2005) [12] used

$$(a+\varepsilon)(b+\varepsilon) \ge \varepsilon^2 \text{ and } ab \le \varepsilon^2$$

to relax the complementarity relationship of a and b.

In this section, we use $\psi_\varepsilon(a,b) = 0$ to approximate the complementarity relation $0 \le a, 0 \le b, ab = 0$, where $\phi_\varepsilon(a,b)$ is the smoothed Fischer-Burmeister function defined by (2.1). Define

$$\Psi_\varepsilon(\lambda,z) = \begin{bmatrix} \psi_\varepsilon(\lambda_1,z_1) \\ \vdots \\ \psi_\varepsilon(\lambda_p,z_p) \end{bmatrix} \tag{2.2}$$

and

$$\Omega(\varepsilon) := \left\{(\lambda,z) \in \Re^p \times \Re^p : \Psi_\varepsilon(\lambda,z) = 0\right\}. \tag{2.3}$$

Then if $(\lambda,z) \in \Omega(\varepsilon)$ we have

$$\lambda > 0, z > 0 \text{ and } \lambda_i z_i = \varepsilon^2, i = 1,\ldots,p.$$

Obviously $\psi_0(a,b)=0$ if and only if $0\le a, 0\le b, ab=0$. Therefore $\Omega(0)=\Omega$.

For any $(\lambda,z)\in\Re^{2p}$, we have

$$\mathcal{J}_{\lambda,z}\Psi_\varepsilon(\lambda,z)=[\mathcal{J}_\lambda\Psi_\varepsilon(\lambda,z)\quad \mathcal{J}_z\Psi_\varepsilon(\lambda,z)]$$

where

$$\mathcal{J}_\lambda\Psi_\varepsilon(\lambda,z)=\begin{bmatrix}1-\dfrac{\lambda_1}{\sqrt{\lambda_1^2+z_1^2+2\varepsilon^2}} & & \\ & \ddots & \\ & & 1-\dfrac{\lambda_p}{\sqrt{\lambda_p^2+z_p^2+2\varepsilon^2}}\end{bmatrix}$$

and

$$\mathcal{J}_z\Psi_\varepsilon(\lambda,z)=\begin{bmatrix}1-\dfrac{z_1}{\sqrt{\lambda_1^2+z_1^2+2\varepsilon^2}} & & \\ & \ddots & \\ & & 1-\dfrac{z_p}{\sqrt{\lambda_p^2+z_p^2+2\varepsilon^2}}\end{bmatrix}.$$

Let $(\lambda,z)\in\Omega_\varepsilon$, then, for $i=1,\dots,p$,

$$\lambda_i+z_i-\sqrt{\lambda_i^2+z_i^2+2\varepsilon^2}=0,$$

we have $\lambda_i>0, z_i>0$ and $\lambda_i z_i=\varepsilon^2$. Thus

$$\begin{aligned}1-\frac{\lambda_i}{\sqrt{\lambda_i^2+z_i^2+2\varepsilon^2}}&=1-\frac{\lambda_i}{\sqrt{\lambda_i^2+z_i^2+2\lambda_i z_i}}\\&=1-\frac{\lambda_i}{\lambda_i+z_i}\\&=\frac{z_i}{\lambda_i+z_i},\end{aligned}$$

and in turn we obtain

$$1-\frac{\lambda_i}{\sqrt{\lambda_i^2+z_i^2+2\varepsilon^2}}=\frac{z_i}{\lambda_i+z_i},\ 1-\frac{z_i}{\sqrt{\lambda_i^2+z_i^2+2\varepsilon^2}}=\frac{\lambda_i}{\lambda_i+z_i}. \tag{2.4}$$

Obviously for any $\varepsilon>0$, both $\mathcal{J}_\lambda\Psi_\varepsilon(\lambda,z)$ and $\mathcal{J}_z\Psi_\varepsilon(\lambda,z)$ are nonsingular matrices, we can easily obtain the following conclusion.

Lemma 2.1. *Let $\varepsilon>0$. Then for any $(\lambda,z)\in\Omega(\varepsilon)$ the linear independence constraint qualification (LICQ) holds and the tangent cone of $\Omega(\varepsilon)$ at (λ,z) is*

$$T_{\Omega(\varepsilon)}(\lambda,z)=\left\{(\triangle\lambda,\triangle z)\in\Re^{2m}:\mathcal{J}_{\lambda,z}\Psi_\varepsilon(\lambda,z)(\triangle\lambda,\triangle z)=0\right\}, \tag{2.5}$$

and the normal cone of $\Omega(\varepsilon)$ at (λ,z) is

$$N_{\Omega(\varepsilon)}(\lambda,z)=\widehat{N}_{\Omega(\varepsilon)}(\lambda,z)=\mathcal{J}_{\lambda,z}\Psi_\varepsilon(\lambda,z)^T\Re^p. \tag{2.6}$$

We use the following problem, denoted by P_ε, to approximate Problem (1.1):

$$\begin{array}{cl}\min\limits_{\vartheta,\mu,\lambda.z} & \mathcal{D}(\vartheta,\overline{\vartheta}) \\ \text{s.t.} & G(\vartheta,\mu,\lambda,z)=0, \\ & (\vartheta,\mu,\lambda,z)\in\Theta\times\Re^q\times\Omega(\varepsilon),\end{array} \tag{2.7}$$

where $\Omega(\varepsilon)$ is defined by (2.3).

We use $\Phi(\varepsilon)$ to denote the feasible set for Problem (2.7), namely

$$\Phi(\varepsilon)=\{(\vartheta,\mu,\lambda,z)\in\Theta\times\Re^q\times\Omega(\varepsilon):G(\vartheta,\mu,\lambda,z)=0\}. \tag{2.8}$$

Define

$$F_\varepsilon(\vartheta,\mu,\lambda,z)=\left[\begin{array}{c}G(\vartheta,\mu,\lambda,z)\\ \Psi_\varepsilon(\lambda,z)\end{array}\right]. \tag{2.9}$$

Then $\Phi(\varepsilon)$ is expressed as

$$\Phi(\varepsilon)=\{(\vartheta,\mu,\lambda,z)\in\Theta\times\Re^q\times\Re^p\times\Re^p:F_\varepsilon(\vartheta,\mu,\lambda,z)=0\}.$$

Similar to the proof of Proposition 1.3, we can establish the following result.

Proposition 2.2. *Assume Assumption 1 and Assumption 3 hold, and Assumption 4 is satisfied at (ϑ,μ,λ) with $\vartheta\in\Theta$. Then*

$$\begin{array}{l}T_{\Phi(\varepsilon)}(\vartheta,\mu,\lambda,z)\\ =\left\{d\in Y\times\Re^q\times\Re^p\times\Re^p:\begin{array}{l}\mathrm{D}_\vartheta\nabla_x L(\overline{x},\vartheta,\vartheta\mu,\lambda)d_\vartheta+\nabla_x h(\overline{x},\vartheta)d_\mu+\nabla_x g(\overline{x},\vartheta)d_\lambda=0\\ \mathrm{D}_\vartheta h(\overline{x},\vartheta)d_\vartheta=0\\ \mathrm{D}_\vartheta g(\overline{x},\vartheta)d_\vartheta+d_z=0\\ d_\vartheta\in T_\Theta(\vartheta)\\ (d_\lambda,d_z)\in T_{\Omega_\varepsilon}(\lambda,z)\end{array}\right\}\\ =\left\{d\in T_\Theta(\vartheta)\times\Re^q\times\Re^p\times\Re^p:\mathrm{D}F_\varepsilon(\vartheta,\mu,\lambda,z)d=0\right\}.\end{array} \tag{2.10}$$

and

$$\begin{array}{l}N_{\Phi(\varepsilon)}(\vartheta,\mu,\lambda,z)=\widehat{N}_{\Phi_\varepsilon}(\vartheta,\mu,\lambda,z)\\ =\left\{\left(\begin{array}{l}\mathrm{D}_\vartheta\nabla_x L(\overline{x},\vartheta,\mu,\lambda)^*\eta_1\\ \quad+\mathrm{D}_\vartheta h(\overline{x},\vartheta)^*\eta_2+\mathrm{D}_\vartheta g(\overline{x},\vartheta)^*\eta_3+N_\Theta(\vartheta)\\ \mathcal{J}_x h(\overline{x},\vartheta)\eta_1\\ \mathcal{J}_x g(\overline{x},\vartheta)\eta_1+\xi_a\\ \eta_3+\xi_b\end{array}\right):\begin{array}{l}(\eta_1,\eta_2,\eta_3)\in\Re^{n+q+p}\\ (\xi_a,\xi_b)\in N_{\Omega_\varepsilon}(\lambda,z)\end{array}\right\}\\ =\mathrm{D}F_\varepsilon(\vartheta,\mu,\lambda,z)^*\Re^{n+q+p}+N_\Theta(\vartheta)\times\{0_q\}\times\{0_p\}\times\mathcal{J}_{\lambda,z}\Psi_\varepsilon(\lambda,z)^T\Re^p.\end{array} \tag{2.11}$$

Lemma 2.3. *For $\Omega(\varepsilon)$ defined by (2.3), we have*

$$\lim_{\varepsilon\searrow 0}\Omega(\varepsilon)=\Omega(0). \tag{2.12}$$

Proof. For any $(\lambda,z)\in\limsup\limits_{\varepsilon\searrow 0}\Omega(\varepsilon)$, there exist $\varepsilon_k\searrow 0$ and $(\lambda^k,z^k)\in\Omega(\varepsilon_k)$ such that $(\lambda^k,z^k)\to(\lambda,z)$. The inclusion $(\lambda^k,z^k)\in\Omega(\varepsilon_k)$ implies

$$\lambda^k+z^k-\sqrt{(\lambda^k)^2+(z^k)^2+2\varepsilon_k^2}=0.$$

Then, letting $k \to \infty$, we have

$$\lambda + z - \sqrt{\lambda^2 + z^2} = 0,$$

namely $\psi_0(\lambda, z) = 0$ and $(\lambda, z) \in \Omega(0)$. Therefore we have

$$\limsup_{\varepsilon \searrow 0} \Omega(\varepsilon) \subset \Omega(0).$$

For any $(\lambda, z) \in \Omega(0)$, let

$$I_+ = \{i : \lambda_i > 0\}, J_+ = \{i : z_i > 0\}, I_0 = \{1, \ldots, m\} \setminus \Big(I_+ \cup J_+\Big).$$

For any $\varepsilon > 0$ defined $(\lambda(\varepsilon), z(\varepsilon))$ by

$$(\lambda_i(\varepsilon), z_i(\varepsilon)) = \begin{cases} (\lambda_i, \varepsilon^2/\lambda_i) & \text{if } i \in I_+, \\ (\varepsilon^2/z_i, z_i) & \text{if } i \in J_+, \\ (\varepsilon, \varepsilon) & \text{if } i \in I_0, \end{cases} \tag{2.13}$$

Then $\psi_\varepsilon(\lambda_i(\varepsilon), z_i(\varepsilon)) = 0$ for $i = 1, \ldots, m$ or equivalently $\Psi_\varepsilon(\lambda(\varepsilon), z(\varepsilon)) = 0$ or $(\lambda(\varepsilon), z(\varepsilon) \in \Omega(\varepsilon)$. Obviously $(\lambda(\varepsilon), z(\varepsilon) \to (\lambda, z)$ and this implies that

$$\liminf_{\varepsilon \searrow 0} \Omega(\varepsilon) \supset \Omega(0).$$

Therefore $\Omega(\varepsilon) \to \Omega(0)$ as $\varepsilon \searrow 0$. □

Corollary 2.4. *Let $\Phi(\varepsilon)$ be defined by (2.8), then*

$$\Phi(\varepsilon) \to \Phi \text{ as } \varepsilon \searrow 0.$$

Proof. In terms of Lemma 2.12, the result can be obtained by noting that $\Phi(\varepsilon)$ and Φ can be expressed as

$$\Phi(\varepsilon) = \{(\vartheta, \mu, \lambda, z) \in \Theta \times \Re^q \times \Re^p \times \Re^p : G(\vartheta, \mu, \lambda, z) = 0\} \cap Y \times \Re^q \times \Omega(\varepsilon)$$

and

$$\Phi = \{(\vartheta, \mu, \lambda, z) \in \Theta \times \Re^q \times \Re^p \times \Re^p : G(\vartheta, \mu, \lambda, z) = 0\} \cap Y \times \Re^q \times \Omega,$$

respectively. □

We denote the optimal value and the (global) solution set of Problem P_ε by $\kappa(\varepsilon)$ and $S(\varepsilon)$, respectively, namely

$$\kappa(\varepsilon) := \inf\{\mathcal{D}(\vartheta, \overline{\vartheta}) \,|\, (\vartheta, \mu, \lambda, z) \in \Omega(\varepsilon)\},$$

$$S(\varepsilon) := \mathrm{Argmin}\{\mathcal{D}(\vartheta, \overline{\vartheta}) \,|\, (\vartheta, \mu, \lambda, z) \in \Omega(\varepsilon)\}.$$

Theorem 2.5. *Let P_ε is defined by (2.7), and $\kappa(\varepsilon)$ and $S(\varepsilon)$ be its optimal value and solution set, respectively. Then the function $\kappa(\varepsilon)$ is continuous at 0 with respect to $\Re_+$ and the set-valued mapping $S(\varepsilon)$ is outer semi-continuous at 0 with respect to $\Re_+$.*

Proof. As $\mathcal{D}(\vartheta, \overline{\vartheta})$ is strictly convex and $\mathcal{D}$ is level-bounded, we have $\kappa(\varepsilon)$ is finite and $S(\varepsilon) \neq \emptyset$ for any $\varepsilon \geq 0$.

Let

$$\widehat{\mathcal{D}}_\varepsilon(\vartheta, \mu, \lambda, z) = \mathcal{D}(\vartheta, \overline{\vartheta}) + \delta_{\Omega(\varepsilon)}(\vartheta, \mu, \lambda, z),$$

where $\delta_{\Omega(\varepsilon)}$ is the indicator function of $\Omega(\varepsilon)$. From Lemma 2.3, $\Omega(\varepsilon) \to \Omega(0)$ as $\varepsilon \searrow 0$, $\widehat{\mathcal{D}}_\varepsilon$ epi-converges to $\widehat{\mathcal{D}}_0$. The level-boundedness of $\widehat{\mathcal{D}}_\varepsilon$ is easily verified for $\varepsilon \geq 0$. Therefore, we have from Theorem 7.41 of Rockafellar and Wets (1998) that the function $\kappa(\varepsilon)$ is continuous at 0 with respect to $\Re_+$ and the set-valued mapping $S(\varepsilon)$ is outer semi-continuous at 0 with respect to $\Re_+$. The proof is completed. □

We say that the point $(\vartheta, \mu, \lambda, z) \in \Phi(\varepsilon)$ is a stationary point of P_ε such that

$$0 \in \mathrm{D}_{\vartheta,\mu,\lambda,z}\mathcal{D}(\vartheta, \overline{\vartheta}) + N_{\Phi(\varepsilon)}(\vartheta, \mu, \lambda, z). \tag{2.14}$$

The following theorem is about the convergence of the stationary points for P_ε, which shows that a cluster point of stationary points for P_ε is related to the C-stationary conditions for Problem (1.1) when $\varepsilon \searrow 0$.

Theorem 2.6. *Let Assumption 1- Assumption 3 and Assumption 4, at every $(\vartheta, \mu, \lambda)$ with $\vartheta \in \Theta$, be satisfied. Let $(\vartheta(\varepsilon), \mu(\varepsilon), \lambda(\varepsilon), z(\varepsilon))$ be a stationary point for P_ε for $\varepsilon > 0$, with multipliers $\eta(\varepsilon) = (\eta_1(\varepsilon), \eta_2(\varepsilon), \eta_3(\varepsilon)) \in \Re^{n+q+p}$ and $\xi(\varepsilon) = (\xi_a(\varepsilon), \xi_b(\varepsilon)) \in N_{\Omega_\varepsilon}(\lambda(\varepsilon), z(\varepsilon))$, then any point $(\vartheta^*, \mu^*, \lambda^*, z^*, \eta^*, \xi^*)$ in the set*

$$\limsup_{\varepsilon \searrow 0}\{(\vartheta(\varepsilon), \mu(\varepsilon), \lambda(\varepsilon), z(\varepsilon), \eta(\varepsilon), \xi(\varepsilon))\}$$

satisfies the C-stationary conditions for Problem (1.1).

Proof. Let $(\vartheta^*, \mu^*, \lambda^*, z^*, \eta^*, \xi^*) \in \limsup_{\varepsilon \searrow 0}\{(\vartheta(\varepsilon), \mu(\varepsilon), \lambda(\varepsilon), z(\varepsilon), \eta(\varepsilon), \xi(\varepsilon))\}$. Then there exists a sequence $\varepsilon_k \searrow 0$ and $(\vartheta^k, \mu^k, \lambda^k, z^k, \eta^k, \xi^k)$ such that $(\vartheta^k, \mu^k, \lambda^k, z^k, \eta^k, \xi^k) \to (\vartheta^*, \mu^*, \lambda^*, z^*, \eta^*, \xi^*)$ with

$$\begin{aligned}
&0 = \mathrm{D}\mathcal{D}(\vartheta^k, \overline{\vartheta}^k) + \mathrm{D}_\vartheta \nabla_x L(\overline{x}, \vartheta^k, \mu^k, \lambda^k)^*\eta_1^k + \mathrm{D}_\vartheta h(\overline{x}, \vartheta^k)^*\eta_2^k + \mathrm{D}_\vartheta g(\overline{x}, \vartheta^k)^*\eta_3^k + v^k,\\
&0 = \mathcal{J}_x h(\overline{x}, \vartheta^k)\eta_1^k,\\
&0 = \mathcal{J}_x g(\overline{x}, \vartheta^k)\eta_1^k + \xi_a^k,\\
&0 = \eta_3^k + \xi_b^k,
\end{aligned} \tag{2.15}$$

where

$$\eta^k = (\eta_1^k, \eta_2^k, \eta_3^k) \in \Re^{n+q+p},\ \xi^k = (\xi_a^k, \xi_b^k) \in N_{\Omega(\varepsilon_k)}(\lambda^k, z^k).$$

From the outer continuity of N_Θ, we have from (2.15) that

$$\begin{aligned}
&0 \in \mathrm{D}\mathcal{D}(\vartheta^*, \overline{\vartheta}) + \mathrm{D}_\vartheta \nabla_x L(\overline{x}, \vartheta^*, \mu^*, \lambda^*)^*\eta_1^* + \mathrm{D}_\vartheta h(\overline{x}, \vartheta^*)^*\eta_2^* + \mathrm{D}_\vartheta g(\overline{x}, \vartheta^*)^*\eta_3^* + N_\Theta(\vartheta^*),\\
&0 = \mathcal{J}_x h(\overline{x}, \vartheta^*)\eta_1^*,\\
&0 = \mathcal{J}_x g(\overline{x}, \vartheta^*)\eta_1^* + \xi_a^*,\\
&0 = \eta_3^* + \xi_b^*,
\end{aligned} \tag{2.16}$$

where

$$(\xi_a^*, \xi_b^*) = \lim_{k\to\infty} (\xi_a^k, \xi_b^k) \in \limsup_{k\to\infty} N_{\Omega(\varepsilon_k)}(\lambda^k, z^k). \tag{2.17}$$

It follows from Lemma 2.3 that $(\lambda^*, z^*) \in \Omega$. Define

$$\alpha = \{i : \lambda_i^* = 0 < z_i^*\},\ \beta = \{i : \lambda_i^* = 0 = z_i^*\},\ \gamma = \{i : \lambda_i^* > 0 = z_i^*\}.$$

Let

$$D^k = \begin{bmatrix} \dfrac{z_1^k}{z_1^k + \lambda_1^k} & & \\ & \ddots & \\ & & \dfrac{z_p^k}{z_p^k + \lambda_p^k} \end{bmatrix},$$

then it follows from (2.6) that

$$N_{\Omega(\varepsilon_k)}(\lambda^k, z^k) = \begin{bmatrix} D^k \\ I_p - D^k \end{bmatrix} \Re^p.$$

so that there exists $y^k \in \Re^p$ such that

$$\xi_a^k = D^k y^k,\ \xi_b^k = (I_p - D^k) y^k.$$

Noting that

$$D^k \to \begin{bmatrix} I_{|\alpha|} & 0 & 0 \\ 0 & D_{|\beta|} & 0 \\ 0 & 0 & 0_{|\gamma|} \end{bmatrix} \text{ and } I_p - D^k \to \begin{bmatrix} 0_{|\alpha|} & 0 & 0 \\ 0 & I_{|\beta|} - D_{|\beta|} & 0 \\ 0 & 0 & I_{|\gamma|} \end{bmatrix},$$

we have that $y_\alpha^k \to [\xi_a^*]_\alpha$ and $y_\gamma^k \to [\xi_b^*]_\gamma$. Therefore we have

$$[\xi_a^*]_\gamma = 0_{|\gamma|} \text{ and } [\xi_b^*]_\alpha = 0_{|\alpha|}.$$

For $i \in \beta$, one has

$$\lambda_i^k z_i^k = \frac{z_i^k \lambda_i^k}{(z_i^k + \lambda_i^k)^2} (y_i^k)^2 \geq 0$$

which implies that $[\xi_a^*]_i [\xi_b^*]_i = \lim_{k\to\infty} \lambda_i^k z_i^k \geq 0$. Therefore, $(\vartheta^*, \mu^*, \lambda^*, z^*, \eta^*, \xi^*)$ satisfies the C-stationary conditions for Problem (1.1). The proof is completed. □

Acknowledgments

The authors would like to thank the anonymous reviewers for their careful reading of the manuscript and insightful comments.

References

[1] R.K. Ahuja and J.B. Orlin, Inverse optimization, *Oper. Res.* 49 (2001) 771–783.

[2] R.K. Ahuja and J.B. Orlin, Combinatorial algorithms for inverse network flow problems, *Networks* 40 (2002) 181–187.

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

[4] W.R. Burton and PhL. Toint, On an instance of the inverse shortest paths problem,*Math. Program.* 53 (1992) 45–61.

[5] M.C. Cai, X.G. Yang and J. Zhang, The complexity analysis of the inverse center location problem, *J. Global Optim.* 15 (1999) 213–218.

[6] F.H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley and Sons, New York, 1983.

[7] F. Facchinei, H. Jiang and L. Qi, A smoothing method for mathematical programs with equilibrium constraints, *Math. Program.* 85 (1999) 107–134.

[8] Y. Jiang, X.T. Xiao, L.W. Zhang and J. Z. Zhang, A perturbation approach for a type of inverse linear programming problems, *Inter. J. Computer Math.* 88 (2011) 508–516.

[9] M. Fukushima and J.S. Pang, Convergence of a smoothing continuation method for mathematical problems with complementarity constraints, Ill-posed Variational Problems and Regularization Techniques, In *Lecture Notes in Economics and Mathematical Systems*, Vol. 477, Théra, M. and Tichatschke, R.(eds.), Springer-Verlag, Berlin/Heidelberg, 1999, pp.105–116.

[10] C. Heuberger,Inverse combinatorial optimization: a survey on problems, methods and results,*J. Comb. Optim.* 8 (2004) 329–361.

[11] G. Iyengar and W. Kang, Inverse conic programming with applications, *Oper. Res. Lett.* 33 (2005) 319–330.

[12] G.H. Lin and M. Fukushima, A modified relaxation scheme for mathematical programs with complementarity constraints, *Ann. Oper. Res.* 133 (2005) 63–84.

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

[14] R.T. Rockafellar and R.J.B. Wets, *Variational Analysis*, Springer-Verlag, Berlin-Heidelberg, 1998.

[15] S. Scholtes, Convergence properties of a regularization scheme for mathematical programs with complementarity constraints, *SIAM J. Optim.* 11 (2001) 918–936.

[16] X.T. Xiao and L.W. Zhang,Solving a class of inverse QP problems by a smoothing Newton method, *J. Comput. Math.* 27 (2009) 787–801.

[17] X.T. Xiao, L.W. Zhang and J.Z. Zhang, On convergence of augmented Lagrange method for inverse semi-definite quadratic programming problems, *J. Indus. Manag. Optim.* 5 (2009) 319–339.

[18] X.T. Xiao, L.W. Zhang and J.Z. Zhang, A smoothing Newton method for a type of inverse semi-definite quadratic programming problems, *J. Comput. Appl. Math.* 223 (2009) 485–498.

[19] J. Zhang and Z. Liu, Calculating some inverse linear programming problems, *J. Comput. Appl. Math.* 72 (1996) 261–273.

[20] J. Zhang and Z. Liu, A further study on inverse linear programming problems, *J. Comput. Appl. Math.* 106 (1999) 345–359.

[21] J. Zhang, Z. Liu and Z. Ma, Some Reverse Location Problems, *European J. Oper. Res.* 124 (2000) 77–88.

[22] J. Zhang and Z. Ma, Solution structure of some inverse combinatorial optimization problems, *J. Comb. Optim.* 3 (1999) 127–139.

[23] J. Zhang and L. Zhang, An augmented Lagrangian method for a class of inverse quadratic programming problems, *Appl. Math. Optim.* 21 (2010) 57–83.

[24] J.Z. Zhang, L.W. Zhang and X.T. Xiao, A perturbation approach for an inverse quadratic programming problem, *Math. Meth. Oper. Res.* 72 (2010) 379–404.

[25] Y. Zhang, Y. Jiang, L.W. Zhang and J.Z. Zhang, A perturbation approach for an inverse linear second-order cone programming, *J. Indus. Manag. Optim.* 9 (2013) 171–189.

[26] Y. Zhang,L.W. Zhang,J. Wu and J.Z. Zhang, A perturbation approach for an inverse quadratic programming problem over second-order cones, *Mathematics of Computation* 84 (2015) 209–236.

Manuscript received 15 May 2014
revised 1 January 2015
accepted for publication 20 March 2015

JIE GAO
School of Mathematical Sciences
Dalian University of Technology, Dalian 116024, China
E-mail address: `jiegao @mail.dlut.edu.cn`

HONGWEI ZHANG
School of Mathematical Sciences, Dalian University of Technology
Dalian 116024, China
E-mail address: `hwzhang@dlut.edu.cn`

XIANTAO XIAO
School of Mathematical Sciences, Dalian University of Technology
Dalian 116024, China
E-mail address: `xtxiao@dlut.edu.cn`

LIWEI ZHANG
School of Mathematical Sciences, Dalian University of Technology
Dalian 116024, China
E-mail address: `lwzhang @dlut.edu.cn`