Applied Mathematics, 2015, 6, 1712-1726
Published Online September 2015 in SciRes. http://www.scirp.org/journal/am
http://dx.doi.org/10.4236/am.2015.610152

An Implicit Smooth Conjugate Projection Gradient Algorithm for Optimization with Nonlinear Complementarity Constraints

Cong Zhang[1*], Limin Sun[1], Zhibin Zhu[2], Minglei Fang[3]

[1]Huarui College, Xinyang Normal University, Xinyang, China
[2]School of Mathematics & Computational Science, Guilin University of Electronic Technology, Guilin, China
[3]College of Science, Anhui University of Science and Technology, Huainan, China
Email: [*]zhcopt@126.com

Received 26 July 2015; accepted 15 September 2015; published 18 September 2015

Abstract

This paper discusses a special class of mathematical programs with equilibrium constraints. At first, by using a generalized complementarity function, the discussed problem is transformed into a family of general nonlinear optimization problems containing additional variable μ. Furthermore, combining the idea of penalty function, an auxiliary problem with inequality constraints is presented. And then, by providing explicit searching direction, we establish a new conjugate projection gradient method for optimization with nonlinear complementarity constraints. Under some suitable conditions, the proposed method is proved to possess global and superlinear convergence rate.

Keywords

Mathematical Programs with Equilibrium Constraints, Conjugate Projection Gradient, Global Convergence, Superlinear Convergence

1. Introduction

Mathematical programs with equilibrium constraints (MPEC) include the bilevel programming problem as its special case and have extensive applications in practical areas such as traffic control, engineering design, and economic modeling. So many scholars are interested in this kind of problems and make great achievements, (see [1]-[10]).

[*]Corresponding author.

How to cite this paper: Zhang, C., Sun, L.M., Zhu, Z.B. and Fang, M.L. (2015) An Implicit Smooth Conjugate Projection Gradient Algorithm for Optimization with Nonlinear Complementarity Constraints. *Applied Mathematics*, **6**, 1712-1726.
http://dx.doi.org/10.4236/am.2015.610152

In this paper, we consider an important subclass of MPEC problem, which is called mathematical program with nonlinear complementarity constraints (MPCC):

$$\begin{aligned} \min \quad & f(x,y) \\ \text{s.t.} \quad & g(x,y) \le 0, \\ & 0 \le G(x,y) \perp y \ge 0 \end{aligned} \tag{1.1}$$

where $f: R^{n+m} \to R$, $g = (g_1, g_2, \cdots, g_p)^{\mathrm{T}} : R^{n+m} \to R^P$, $G = (G_1, G_2, \cdots, G_m)^{\mathrm{T}} : R^{n+m} \to R^m$ are all continuously differential functions, $(x,y,w) \in R^{n+m+m}$. $G(x,y) \perp y$ denotes orthogonality of the vectors y and $G(x,y)$, *i.e.*, $y^{\mathrm{T}} G(x,y) = 0$.

In order to eliminate the complementary constraints, which can not satisfy the standard constraint qualification [11], we introduce the generalized nonlinear complementary function

$$\phi(a,b,\mu) = a + b - \sqrt{a^2 + b^2 + 2\mu}, \ (a,b,\mu) \in R^2 \times [0,\infty).$$

Obviously, the following practical results about function ϕ hold:

- if $\phi(a,b,0) = 0$ and $a \ne b$, then

$$\begin{aligned} D_a &= \frac{\partial \phi(a,b,0)}{\partial a} \ne 0, \ D_b = \frac{\partial \phi(a,b,0)}{\partial b} = 0, \ \text{if } b > 0, \\ D_a &= \frac{\partial \phi(a,b,0)}{\partial a} = 0, \ D_b = \frac{\partial \phi(a,b,0)}{\partial b} \ne 0, \ \text{if } a > 0. \end{aligned} \tag{1.2}$$

-

$$\phi(a,b,\mu) = 0, \mu \ge 0 \Leftrightarrow a \ge 0, b \ge 0, ab = \mu. \tag{1.3}$$

By means of the function ϕ, problem (1.1) is transformed equivalently into the following standard nonlinear optimization problem

$$\begin{aligned} \min \quad & f(x,y) \\ \text{s.t.} \quad & g_j(x,y) \le 0, \ \ j \in I_1 \triangleq \{1,2,\cdots,p\}, \\ & c_j(x,y,w) = w_j - G_j(x,y) = 0, \ \ j \in I_2 \triangleq \{1,2,\cdots,m\}, \\ & \phi(y_j, w_j, \mu) = 0, \ \ j \in I_2, \\ & 1 - \mathrm{e}^{\mu} = 0. \end{aligned} \tag{1.4}$$

Similar to [12], we define the following penalty function

$$\theta_c(x,y,w,\mu) = f(x,y) - c\sum_{j=1}^{m}\left(\phi(y_j, w_j, \mu) + c_j(x,y,w)\right) + c\left(\mathrm{e}^{\mu} - 1\right),$$

where $c > 0$ is a penalty parameter. Therefore, our approach consists of solving an auxiliary inequality constrained problem which is defined by

$$\begin{aligned} \min \quad & \theta_c(x,y,w,\mu) \\ \text{s.t.} \quad & g_j(x,y) \le 0, \ \ j \in I_1, \\ & c_j(x,y,w) \le 0, \ \ j \in I_2, \\ & \phi(y_j, w_j, \mu) \le 0, \ \ j \in I_2, \\ & 1 - \mathrm{e}^{\mu} \le 0. \end{aligned} \tag{1.5}$$

2. Preliminaries and Algorithm

For the sake of simplicity, we denote

$$
\begin{aligned}
& z=(x,y,w),\ s=(x,y),\ t=(y,w),\ t_j=(y_j,w_j),\\
& dz=(dx,dy,dw),\ ds=(dx,dy),\\
& X_0=\{z \mid g_j(s)\le 0, j\in I_1, 0\le G(x,y)\perp y\ge 0\},\\
& X_1=\{(z,\mu) \mid g_j(s)\le 0, j\in I_1, c_j(z)\le 0, j\in I_2, \phi(t_j,\mu)\le 0, j\in I_2, 1-\mathrm{e}^{\mu}\le 0\},\\
& r_i=r_i(z,\mu)=\begin{cases} g_j(x,y), & i=j, j\in I_1,\\ c_j(x,y,w), & i=j+p, j\in I_2,\\ \phi(t_j,\mu), & i=j+p+m, j\in I_2,\\ 1-\mathrm{e}^{\mu}, & i=p+m+m+1.\end{cases}\\
& T=\{1,2,\cdots,p+2m+1\},\\
& h_i=h_i(z,\mu)=\nabla r_i(z,\mu),\ H_i=H_i(z,\mu)=\nabla^2 r_i(z,\mu),\ i\in T,\\
& I(z,\mu)=\{i\in T \mid r_i(z,\mu)=0\},\ J_0(z,\mu)=\{i\in I_1 \mid r_i(z,\mu)=0\}.
\end{aligned}
\tag{2.1}
$$

Throughout this paper, the following basic assumptions are assumed.

H 2.1. *The feasible set of* (1.1) *is nonempty, i.e.,* $X_0\ne\phi$.

H 2.2. *The functions* $f, g_j, G_j\ (j\in I_2)$ *are twice continuously differentiable*.

H 2.3. $\forall (z,\mu)\in X_1$, *the vectors* $\{h_i(z,\mu), i\in J_0(z,\mu)\cup(T\setminus I_1)\}$ *are linearly independent*.

The following definition and proposition can be refereed to in [13].

Definition 2.1. *Suppose that* $z^*=(x^*,y^*,w^*)\in X_0$ *satisfies the so-called nondegeneracy condition*:

$$
(y_j^*, G_j(x^*,y^*))\ne(0,0),\ j\in I_2. \tag{2.2}
$$

If there exists multipliers $(\lambda^*,u^*,\gamma^*)\in R^{p+2m}$ such that

$$
\nabla f(s^*)+\nabla g(s^*)\lambda^*+\nabla G(x^*,y^*)u^*+\begin{pmatrix}0_{n\times m}\\ E_m\end{pmatrix}\gamma^*=0, \tag{2.3}
$$

$$
0\le -g(s^*)\perp\lambda^*\ge 0;\ u_j^*=0,\ \text{ if } G_j(s^*)>0;\ \gamma_j^*=0,\ \text{ if } y_j^*>0 \tag{2.4}
$$

hold, then s^* is said to be a $K-T$ point of (1.1).

Proposition 2.1. *Suppose that* $z^*=(x^*,y^*,w^*)\in X_0$ *satisfies the so-called nondegeneracy condition* (2.2), *then* $(s^*,\lambda^*,u^*,\gamma^*)$ *is a* $K-T$ *point of* (1.1) *if and only if* (s^*,λ^*,u^*,v^*) *satisfies*

$$
\begin{pmatrix}\nabla_x f(s^*)\\ \nabla_y f(s^*)\\ 0_{m\times 1}\end{pmatrix}+\begin{pmatrix}\nabla_x g(s^*)\\ \nabla_y g(s^*)\\ 0_{m\times 1}\end{pmatrix}\lambda^*+\begin{pmatrix}\nabla_x G(s^*)\\ \nabla_y G(s^*)\\ -E_m\end{pmatrix}u^*+\begin{pmatrix}0_{n\times m}\\ W^*\\ Y^*\end{pmatrix}v^*=\begin{pmatrix}0_{n\times 1}\\ 0_{m\times 1}\\ 0_{m\times 1}\end{pmatrix}, \tag{2.5}
$$

$$
0\le -g(s^*)\perp\lambda^*\ge 0, \tag{2.6}
$$

where

$$
v^*=(v_j^*, j\in I_2)\in R^m, v_j^*=\begin{cases}\gamma_j^*/w_j^*, & \text{if } w_j^*=G_j(s^*)>0,\\ u_j^*/y_j^*, & \text{if } y_j^*>0,\end{cases} \tag{2.7}
$$

$Y^*=\mathrm{diag}(y_j^*, j\in I_2)$ and $W^*=\mathrm{diag}(w_j^*=G_j(s^*), j\in I_2)$.

Proposition 2.2. (1) *s is a feasible point of* (1.1) *if and only if* (z,μ) *with* $w=G(s), \mu=0$ *is a feasible point of* (1.4).

(2) s^* is a $K-T$ point of (1) if and only if (z^*,μ_*) with $w^*=G(s^*), \mu_*=0$ is a $K-T$ point of (1.4).

Proof. (1) According to the property of function ϕ, the conclusion follows immediately from (1.3).

(2) Suppose that s^* is a $K-T$ point of (1.1). If set $w^* = G(s^*)$ and $\mu_* = 0$, then, from (1), we see $(z^*, 0)$ is a feasible point of (1.4). While, it follows from proposition 2.1 that there exists vector (λ^*, u^*, v^*) such that (2.5) and (2.6) hold. Define

$$\hat{v}^* = (\hat{v}_j^*, j \in I_2) \in R^m,\ \hat{v}^* = \begin{cases} y_j^* v_j^* / D_a(t_j^*), & \text{if } y_j^* > 0, \\ w_j^* v_j^* / D_b(t_j^*), & \text{if } y_j^* = 0, \end{cases} \quad \iota^* = -\sum_{j=1}^{m} \frac{\partial \phi(t_j^*, \mu_*)}{\partial \mu} \hat{v}_j^*. \tag{2.8}$$

So, it is not difficult to prove that $(z^*, \lambda^*, u^*, \hat{v}^*, \iota^*)$ satisfies the $K-T$ system of (1.4), according to (1.3), (2.5) and (2.6).

Conversely, if (z^*, μ_*) is a $K-T$ point of (1.4), then it follows that

$$w^* = G(s^*),\ \mu_* = 0 \ \text{ and } \ \phi(t_j^*, 0) = \phi(t_j^*) = 0,$$

which shows that s^* is a feasible point of (1.1). Suppose $(\lambda^*, u^*, \hat{v}^*, \iota^*) \in R^{p+2m+1}$ is a $K-T$ multiplier corresponding to (z^*, μ_*) of (1.4). Define v^* as follows:

$$v^* = (v_j^*, j \in I_2) \in R^m,\ v_j^* = \begin{cases} D_b(t_j^*) \hat{v}_j^* / y_j^*, & \text{if } y_j^* > 0; \\ D_a(t_j^*) \hat{v}_j^* / w_j^*, & \text{if } y_j^* = 0. \end{cases} \tag{2.9}$$

Then, it is easy to see, from (1.2) and the $K-T$ system of (4) at (z^*, μ_*), that s^* with the multiplier (λ^*, u^*, v^*) satisfies (2.5) and (2.6). Therefore, we assert s^* is a $K-T$ point of (1.1) according to proposition 2.1.

Now, we present the definition of multiplier function associated with ϵ-active set [14].

Definition 2.2. *A continuous function* $\rho(z, \mu): R^{n+2m+1} \to R^{p+2m+1}$ *is said to a multiplier function, if* (z^*, μ^*) *satisfies the* $K-T$ *system of* (1.5) *with corresponding multipliers* $\rho(z^*, \mu^*)$.

Firstly, for a given point (z^k, μ_k), by using the pivoting operation, we obtain an approximate active $J_k = J(z^k, \mu_k)$.

Algorithm A:

Step 1. For the current point $(z^k, \mu_k) \in X_1$ and parameter $\varrho(z^k, \mu_k) = (\varrho_i(z^k, \mu_k), i \in T) \in R^{p+2m+1}$. Set $l = 0$, $\epsilon_l(z^k, \mu_k) = \epsilon_0$;

Step 2. If $det\left(A_l(z^k, \mu_k)^{\mathrm{T}} A_l(z^k, \mu_k)\right) \ge \epsilon_{k,l}$, let $J_k = J_{k,l}, A_k = A_l(z^k, \mu_k), l(z^k, \mu_k) = l$, stop; otherwise, goto Step 3, where

$$\begin{aligned} &J_{k,l}(z^k, \mu_k) = \left\{ i \in I_1 \middle| -\epsilon_{k,l} \left| \varrho_i(z^k, \mu_k) \right| \le r_i(z^k, \mu_k) \le 0 \right\}, \\ &A_{k,l} = \left\{ h_i(z^k, \mu_k), i \in J_k \cup (T \setminus I_1) \right\}. \end{aligned} \tag{2.10}$$

Step 3. $l = l + 1$, $\epsilon_{k,l} = \frac{1}{2} \epsilon_{k,l-1}$, go back to Step 2.

Lemma 2.1. *For any iteration index k, algorithm A terminates in finite iteration.*

For the current point (z^k, μ_k) and ϵ-active set J_k, compute

$$F(z^k, \mu_k) = (r_i(z^k, \mu_k), i \in L_k = J_k \cup (T \setminus I_1)),\ A_k = A(z^k, \mu_k) = (h_i(z^k, \mu_k), i \in L_k). \tag{2.11}$$

Now we give some notations and the explicit search direction in this paper.

$$Q_k = Q(z^k, \mu_k) = (A_k^{\mathrm{T}} B_k^{-1} A_k)^{-1} A_k^{\mathrm{T}} B_k^{-1},\ P_k = P(z^k, \mu_k) = B_k^{-1}(E_{n+2m+1} - A_k Q_k). \tag{2.12}$$

$$\begin{aligned} &\pi^k = \pi(z^k, \mu_k) = -Q(z^k, \mu_k) \nabla \theta_{c_k}(z^k, \mu_k), \\ &d_0^k = (dz_0^k, d\mu_0^k) = d_0(z^k, \mu_k) = -P_k \nabla \theta_{c_k}(z^k, \mu_k) + Q_k^{\mathrm{T}} V^k. \end{aligned} \tag{2.13}$$

$$V^k = V\left(z^k,\mu_k\right) = \left(V_i^k, i \in L_k\right),\ V_i^k = \begin{cases} -r_i(z^k,\mu_k), & \pi_i^k > 0, \\ \pi_i^k, & \pi_i^k \le 0. \end{cases} \tag{2.14}$$

$$d_1^k = -Q_k^{\mathrm{T}}\left(\left\|d_0^k\right\|^{\tau} e + F\left(z^k + dz_0^k, \mu_k + d\mu_0^k\right)\right),\ d^k = \left(dz^k, d\mu_k\right) = d_0^k + d_1^k. \tag{2.15}$$

$$\rho_k = -\nabla\theta_{c_k}\left(z^k,\mu_k\right)^{\mathrm{T}} d_0^k,\ d_2^k = \frac{-\rho_k}{1+2\left|e^{\mathrm{T}}\pi^k\right|} Q_k^{\mathrm{T}} e,\ q^k = \left(qz^k, q\mu_k\right) = \rho_k\left(d_0^k + d_2^k\right) \tag{2.16}$$

where $e = (1,\cdots,1)^{\mathrm{T}} \in R^{|L_k|}$.

According to the above analysis, the algorithm for the solution of the problem (1.1) can be stated as follows.

Algorithm B:

Step 0. Given a starting point $\left(z^1,\mu_1\right) \in X_1$, and an initial symmetric positive definite matrix $B_1 \in R^{(n+2m+1)\times(n+2m+1)}$. Choose parameters $\xi,\sigma,\upsilon,\varepsilon_0 \in (0,1), \alpha \in \left(0,\frac{1}{2}\right), \tau \in (2,3), \delta_0 > 2, \delta_1 > 0, \delta_2 > 0, c_k > 0, k = 1$.

Step 1. By means of Algorithm A, compute $J_k = J\left(z^k,\mu_k\right), A_k = A\left(z^k,\mu_k\right)$ and $F\left(z^k,\mu_k\right)$.

Step 2. Compute d_0^k according to (2.13). If $d_0^k = 0$, stop; otherwise, compute d^k according to (2.14). If

$$\nabla\theta_{c_k}\left(z^k,\mu_k\right)^{\mathrm{T}} d_0^k \le \min\left\{-\xi\left\|d_0^k\right\|^{\delta_0}, -\xi\left\|d^k\right\|^{\delta_0}\right\}, \tag{2.17}$$

goto Step 3; otherwise, goto Step 4.

Step 3. Let $\lambda = 1$.

(1) If

$$\theta_{c_k}\left(z^k + \lambda dz^k, \mu_k + \lambda d\mu_k\right) \le \theta_{c_k}\left(z^k,\mu_k\right) + \alpha\lambda\nabla\theta_{c_k}\left(z^k,\mu_k\right)^{\mathrm{T}} d_0^k, \tag{2.18}$$

$$r_i\left(z^k + \lambda dz^k, \mu_k + d\mu_k\right) \le 0, i \in T. \tag{2.19}$$

Set $\lambda_k = \lambda$, goto Step 5.

(2) Let $\lambda = \frac{1}{2}\lambda$. if $\lambda < \sigma$, goto Step 4; otherwise, repeat (1).

Step 4. Obtain feasible descent direction q^k from (2.16), and compute β_k, the first number β in the sequence $\left\{1,\frac{1}{2},\frac{1}{4},\cdots\right\}$ satisfying

$$\theta_{c_k}\left(z^k + \beta qz^k, \mu_k + \beta q\mu_k\right) \le \theta_{c_k}\left(z^k,\mu_k\right) + \upsilon\beta\nabla\theta_{c_k}\left(z^k,\mu_k\right)^{\mathrm{T}} q^k, \tag{2.20}$$

$$r_i\left(z^k + \beta qz^k, \mu_k + \beta q\mu_k\right) \le 0, i \in T. \tag{2.21}$$

Let $d^k = q^k, \lambda_k = \beta_k$.

Step 5. Define $\tilde{\pi}^k = -\left(A_k^{\mathrm{T}}A_k\right)^{-1} A_k^{\mathrm{T}} \begin{pmatrix} \nabla f\left(s^k\right) \\ 0_{m\times 1} \\ 0 \end{pmatrix}$, $\bar{c}\left(z^k,\mu_k\right) = \max\left\{-\tilde{\pi}_i^k \middle| i \in T \setminus I_1\right\} + \delta_1$ and set

$$c_{k+1} = \begin{cases} \max\left\{\bar{c}\left(z^k,\mu_k\right), c_k + \delta_2\right\}, & \bar{c}\left(z^k,\mu_k\right) > c^k; \\ c_k, & \text{otherwise}. \end{cases} \tag{2.22}$$

and $\left(z^{k+1},\mu_{k+1}\right) = \left(z^k,\mu_k\right) + \lambda_k d^k$. Obtain B_{k+1} by updating the positive definite matrix B_k using some quasi-Newton formulas, and set $k = k + 1$. Go back to Step 1.

In the remainder of this section, we give some results to show that Algorithm B is correctly stated.

Lemma 2.2. (1) *If* $d_0^k \neq 0$, *then we have*

$$\nabla\theta_{c_k}\left(z^k,\mu_k\right)^{\mathrm{T}} d_0^k < 0,\quad \nabla\theta_{c_k}\left(z^k,\mu_k\right)^{\mathrm{T}} q^k \le -\frac{1}{2}\rho_k^2 < 0, \tag{2.23}$$

$$h_i\left(z^k,\mu_k\right)^{\mathrm{T}} d_0^k \le 0,\quad h_i\left(z^k,\mu_k\right)^{\mathrm{T}} q^k < 0, i\in I\left(z^k,\mu_k\right). \tag{2.24}$$

(2) If the sequence $\left\{z^k,\mu_k\right\}$ is bounded, then there exists a constant $c_0 > 0$ such that

$$\nabla\theta_{c_k}\left(z^k,\mu_k\right)^{\mathrm{T}} q^k \le c_0\left\|q^k\right\|^2. \tag{2.25}$$

Proof. (1) If $d_0^k \neq 0$, then

$$\nabla\theta_{c_k}\left(z^k,\mu_k\right)^{\mathrm{T}} d_0^k = -\nabla\theta_{c_k}\left(z^k,\mu_k\right)^{\mathrm{T}} P_k\nabla\theta_{c_k}\left(z^k,\mu_k\right) - \left(\pi^k\right)^{\mathrm{T}} V^k$$
$$= -\left(P_k\nabla\theta_{c_k}\left(z^k,\mu_k\right)\right)^{\mathrm{T}} B_k\left(P_k\nabla\theta_{c_k}\left(z^k,\mu_k\right)\right) - \sum_{\pi_i^k\le 0}\left(\pi_i^k\right)^2 + \sum_{\pi_i^k>0}\pi_i^k h_i^k < 0,$$

$$\nabla\theta_{c_k}\left(z^k,\mu_k\right)^{\mathrm{T}} q^k = \rho_k\nabla\theta_{c_k}\left(z^k,\mu_k\right)^{\mathrm{T}}\left(d_0^k + d_2^k\right) = \rho_k\left(-\rho_k + \frac{\rho_k}{1+2\left|\mathrm{e}^{\mathrm{T}}\pi^k\right|}\left(\pi^k\right)^{\mathrm{T}} e\right) \le -\frac{1}{2}\rho_k^2 < 0.$$

In view of $A_k^{\mathrm{T}} d_0^k = V^k$, we get $\left(h_i^k\right)^{\mathrm{T}} d_0^k \le 0, i\in I\left(z^k,\mu_k\right)$. Since

$$A_k^{\mathrm{T}} q^k = \rho_k A_k^{\mathrm{T}}\left(d_0^k + d_2^k\right) = \rho_k\left(V^k - \frac{\rho_k}{1+2\left|e^{\mathrm{T}}\pi^k\right|}e\right) \le \frac{-\rho_k^2}{1+2\left|e^{\mathrm{T}}\pi^k\right|}e,$$

so we have

$$\left(h_i^k\right)^{\mathrm{T}} q^k \le \frac{-\rho_k^2}{1+2\left|e^{\mathrm{T}}\pi^k\right|} < 0,\ \ i\in I\left(z^k,\mu_k\right). \tag{2.26}$$

(2) Note that the boundedness of sequence $\left\{z^k,\mu_k\right\}$ and B_k positive definite, we know that d_0^k, d_2^k are bounded. By (2.16), there exists constant $\hat{c} > 0$ such that $\rho_k \ge \hat{c}\left\|q^k\right\|$. Thus, there exists constant $c_0 > 0$ such that

$$\nabla\theta_{c_k}\left(z^k,\mu_k\right)^{\mathrm{T}} q^k \le -\frac{1}{2}\rho_k^2 \le -c_0\left\|q^k\right\|^2.$$

So, the claim holds.

According to Lemma 2.2 and the continuity of functions $\theta_{c_k}\left(z^k,\mu_k\right)$ and $r_i\left(z^k,\mu_k\right), i\in T$, the following result is true.

Lemma 2.3. *Algorithm B is well defined*.

3. Global Convergence

In this section, we consider the global convergence of the algorithm B. Firstly, we show that s^k is an exact stationary point of (1.1) if the Algorithm B terminates at the current iteration point $\left(z^k,\mu_k\right)$.

Lemma 3.1. (1) $\left(z^k,\mu_k\right)$ *is a* $K-T$ *point of* (1.5) *if and only if* $d_0^k = 0$.

(2) If $\left(z^k,\mu_k\right)$ is a $K-T$ point of (1.5), then $\left(z^k,\mu_k\right)$ with $\mu_k = 0$ is a $K-T$ point of (1.4).

Proof. (1) If $\left(z^k,\mu_k\right)$ is a $K-T$ point of (1.5), then from the definition of index set J_k, we know the $K-T$ multiplier corresponding to constraints about index $I_1\setminus J_k$ is 0. Thus, there exists vector $\chi = \left(\chi_i, i\in L_k\right)$ such that

$$\nabla\theta_{c_k}\left(z^k,\mu_k\right) + A_k\chi = 0,\ \chi_i \ge 0,\ \chi_i r_i\left(z^k,\mu_k\right) = 0,\ \ i\in L_k. \tag{3.1}$$

Note that matrix A_k is full of column rank, and B_k positive definite. Thus we have $\left(A_k^{\mathrm{T}} B_k^{-1} A_k\right)^{-1}$ exists.

Furthermore, it follows from (3.1) that

$$\chi = -\left(A_k^{\mathrm{T}} B_k^{-1} A_k\right)^{-1} A_k^{\mathrm{T}} B_k^{-1} \nabla \theta_{c_k}\left(z^k, \mu_k\right) = -Q_k \nabla \theta_{c_k}\left(z^k, \mu_k\right) = \pi^k.$$

By (2.14) and (3.1), we have

$$V^k = 0,\ B_k^{-1} \nabla \theta_{c_k}\left(z^k, \mu_k\right) - B_k^{-1} A_k Q_k \nabla \theta_{c_k}\left(z^k, \mu_k\right) = 0,\ P_k \nabla \theta_{c_k}\left(z^k, \mu_k\right) = 0,$$

so $d_0^k = 0.$

On the other hand, it is easy to verify that

$$P_k A_k = 0,\ P_k B_k P_k = P_k,\ Q_k A_k = E_{|L_k|}.$$

It follows from $d_0^k = 0$ that

$$0 = A_k^{\mathrm{T}} d_0^k = V^k,\ P_k \nabla \theta_{c_k}\left(z^k, \mu_k\right) = 0.$$

From the positive definiteness of B_k and (2.12), (2.13) and (2.14), we have

$$\nabla \theta_{c_k}\left(z^k, \mu_k\right) + A_k \pi^k = 0,\ \pi_i^k \geq 0,\ \pi_i^k r_i\left(z^k, \mu_k\right) = 0,\ i \in L_k, \tag{3.2}$$

which implies that $\left(z^k, \mu_k\right)$ is a $K-T$ point of (1.5).

(2) In view of the definition of $\theta_c\left(z, \mu\right)$, we obtain from (3.2) that

$$\begin{pmatrix} \nabla f\left(s^k\right) \\ 0_{m\times 1} \\ 0 \end{pmatrix} - c_k \sum_{i\in(T\setminus I_1)} h_i^k + \sum_{i\in L_k} \pi_i^k h_i^k = 0, \tag{3.3}$$
$$\pi_i^k \geq 0,\ \pi_i^k r_i\left(z^k, \mu_k\right) = 0,\ i \in L_k.$$

Since the vectors $\left\{h_i^k, i \in L_k\right\}$ are linearly independent, we have

$$\pi_{L_k}^k = -\left(A_k^{\mathrm{T}} A_k\right)^{-1} A_k^{\mathrm{T}} \left(\begin{pmatrix} \nabla f\left(s^k\right) \\ 0_{m\times 1} \\ 0 \end{pmatrix} - c_k \sum_{i\in(T\setminus I_1)} h_i^k \right),$$

i.e.

$$\pi_{L_k}^k = \tilde{\pi}_{L_k}^k + c_k \sum_{i\in(T\setminus I_1)} \left(A_k^{\mathrm{T}} A_k\right)^{-1} A_k^{\mathrm{T}} h_i^k.$$

Thus, we deduce

$$\pi_i^k = \tilde{\pi}_i^k + c_k,\ i \in \left(T \setminus I_1\right). \tag{3.4}$$

In view of the definition of penalty parameter c_k, from (3.4), we have

$$\pi_i^k = \tilde{\pi}_i^k + c_k \geq \delta_1 > 0,\ \ i \in \left(T \setminus I_1\right). \tag{3.5}$$

Combining with (3.2) and (3.5), it holds that

$$r_i\left(z^k, \mu_k\right) = 0,\ i \in \left(T \setminus I_1\right),\ \mu_k = 0. \tag{3.6}$$

Let $\bar{\pi}^k = \left(\bar{\pi}_i^k, i \in T\right)$, where $\bar{\pi}_i^k = \pi_i^k, i \in L_k, \bar{\pi}_i^k = 0, i \in I_1 \setminus J_k$. From (3.3) and (3.6), we can easily see that $\left(z^k, \mu_k, \bar{\pi}^k\right)$ is a $K-T$ point pair of (1.4).

Theorem 3.1. *Suppose the nondegeneracy condition holds at* z^k. *If* $\left(z^k, \mu_k\right)$ *is a* $K-T$ *point of* (1.4), *then*

s^k *is a* $K-T$ *point of* (1.1).

Proof. According to the $K-T$ system of (1.4) and the relationship of index i and j in (2.1), we see that

$$\mu_k = 0, w_j^k = G_j\left(s^k\right),\ \phi\left(t_j^k\right) = 0,\ j \in I_2. \tag{3.7}$$

Then, combining with Proposition 2.1 and Proposition 2.2, we can conclude that s^k is a $K-T$ point of (1.1).

In the sequel, it is assumed that the Algorithm B generates an infinite sequence $\left\{\left(z^k, \mu_k\right)\right\}$. The following further assumption about $\left\{\left(z^k, \mu_k\right)\right\}$ is required in subsequent discussions.

H 3.1. (1) *The sequence* $\left\{\left(z^k, \mu_k\right)\right\}$ *is bounded*.

(2) The accumulation point $\left(z^*, \mu_*\right)$ of infinite sequence $\left\{\left(z^k, \mu_k\right)\right\}$ satisfies (2.2).

From H 3.1 and the fact that there are only finitely many choices for sets $J_k \subseteq I_1$, we may assume that there exists a subsequence K, such that

$$z^k \to z^*, B_k \to B_*, J_k \equiv J, k \in K, \tag{3.8}$$

where J is a constant set. Correspondingly, the following results hold:

$$A_k \to A_*, Q_k \to Q_*, P_k \to P_*, d_0^k \to d_0^*, q^k \to q^*, k \in K, k \to \infty.$$

Lemma 3.2. *Suppose* $\left(z^k, \mu_k\right) \to \left(z^*, \mu_*\right)$, *then for* $k \in K$ *large enough, we have*

(1) there exists a constant $\bar{\epsilon} > 0$ such that $\epsilon_{k,l_k} \geq \bar{\epsilon}$.

(2) there exists a constant $c > 0$ such that $c_k \equiv c$.

Proof. (1) suppose, by contradiction, that there exists an index set $K' \subseteq K$ such that $\epsilon_{k,l_k} \to 0\left(k \in K', k \to \infty\right)$. Let $J_{k'} = J_{k,l-1}$. For $k \in K'$ large enough, from Algorithm A, we have

$$\det\left(A^{\mathrm{T}}_{J'_k \cup (T\setminus I_1)} A_{J'_k \cup (T\setminus I_1)}\right) < 2\epsilon_{k,l_k},\ -\epsilon_{k,l}\left|\varrho_i\left(z^k, \mu_k\right)\right| \leq r_i\left(z^k, \mu_k\right) \leq 0,\ i \in J'_k. \tag{3.9}$$

Since there are only finite possible subsets of I_1, there must be an infinite subset $K'' \in K'$ such that for any $k \in K'', J'_k = J'$. Thus, it follows from (3.9) that

$$\begin{aligned} &\det\left(A^{\mathrm{T}}_{J' \cup (T\setminus I_1)} A_{J' \cup (T\setminus I_1)}\right) < 2\epsilon_{k,l_k} \to 0, k \in K'', k \to \infty, \\ &r_i\left(z^*, \mu_*\right) = 0, i \in J', \end{aligned} \tag{3.10}$$

which contradicts the condition H 2.3.

(2) Suppose by contradiction, there exists a subsequence $\{k_i\}$ such that $c_{k_i} > c_{k_i-1}\left(i = 1, 2, \cdots\right)$, then from the definition of c_k, we have

$$\begin{aligned} &\bar{c}\left(z^k, \mu_k\right) > c_{k_i-1}\left(i = 1, 2, \cdots\right), \\ &c_k \to \infty, k \to \infty. \end{aligned} \tag{3.11}$$

From the finite selectivity of J_k, we can suppose without loss of generality that $J_{k_i} = J\left(i = 1, 2, \cdots\right)$. By (1), we can see that $\bar{c}\left(z_{k_i}, \mu_{k_i}\right)$ is bounded, *i.e.*, $\bar{c}\left(z_{k_i}, \mu_{k_i}\right) < c^*\left(i = 1, 2, \cdots\right)$ for some c^*. Let M be such an integer that $c_{k_i-1} \geq c^*\left(i \geq M\right)$, then we have

$$c_{k_i-1} \geq c^* \geq \bar{c}\left(z_{k_i}, \mu_{k_i}\right)\left(i \geq M\right),$$

a contradiction, and the result is proved.

Lemma 3.3. *Suppose that* $\left(z^k, \mu_k\right) \to \left(z^*, \mu_*\right)\left(k \in K\right)$, *and* $\left(z^{k+1}, \mu_{k+1}\right) = \left(z^k, \mu_k\right) + \lambda_k d^k\left(k \in K\right)$ *which is generated by Step* 4 *and Step* 5. *If* $\left(z^*, \mu_*\right)$ *is not a* $K-T$ *point of* (1.5), *then we have*

(1) $\nabla\theta_c\left(z^*, \mu_*\right)^{\mathrm{T}} q^* < 0,\ h_i\left(z^*, \mu_*\right)^{\mathrm{T}} q^*, i \in I\left(z^*, \mu_*\right)$,

(2) $\beta_k \geq \beta_* = inf\left\{\beta_k, k \in K\right\} > 0, k \in K$.

Proof. (1) Suppose $B_k \to B_*, d_0^k \to d_0^*, q^k \to q^*, k \in K$. Since $\left(z^*, \mu_*\right)$ is not a $K-T$ of (1.5), so we have $d_0^* \neq 0$ and

$$\nabla\theta_c\left(z^*,\mu_*\right)^{\mathrm{T}} d_0^* < 0, \nabla\theta_c\left(z^*,\mu_*\right)^{\mathrm{T}} q^* < 0,$$
$$h_i\left(z^*,\mu_*\right)^{\mathrm{T}} q^* < 0, i \in I\left(z^*,\mu_*\right).$$

Therefore, for $k \in K$ large enough, we obtain

$$\begin{aligned} &-\rho_k \le \frac{1}{2}\nabla\theta_c\left(z^*,\mu_*\right)^{\mathrm{T}} d_0^* < 0, \\ &\nabla\theta_c\left(z^k,\mu_k\right)^{\mathrm{T}} q^k \le \nabla\frac{1}{2}\theta_c\left(z^*,\mu_*\right)^{\mathrm{T}} d_0^* < 0, \\ &h_i\left(z^k,\mu_k\right)^{T} q^k < 0, i \in I\left(z^*,\mu_*\right). \end{aligned} \tag{3.12}$$

(2) For (2.20), denote

$$\begin{aligned} a &\triangleq \theta_c\left(z^k+\beta q z^k,\mu_k+\beta q\mu_k\right)-\theta_c\left(z^k,\mu_k\right)-\upsilon\beta\nabla\theta_c\left(z^k,\mu_k\right)^{\mathrm{T}} q^k \\ &=\left(1-\upsilon\right)\beta\nabla\theta_c\left(z^k,\mu_k\right)^{\mathrm{T}} q^k+o\left(\beta\right). \end{aligned}$$

From (3.12), for $k \in K$ large enough and $\beta > 0$ small enough, it holds that $s \le 0$.

For (2.21), when $i \in T \setminus I\left(z^*,\mu_*\right)$, the fact $\left(z^k,\mu_k\right) \to (z^*,\mu_*), r_i\left(z^*,\mu_*\right) < 0, k \in K$ and the continuity of r_i imply that (4.5) holds. When $i \in I\left(z^*,\mu_*\right)$, it holds that $r_i\left(z^*,\mu_*\right) = 0$. From (3.12), for $k \in K$ large enough and $\beta > 0$ small enough, we have

$$\begin{aligned} r_i\left(z^k+\beta q z^k,\mu_k+\beta q\mu_k\right) &= r_i\left(z^k,\mu_k\right)+\beta h_i\left(z^k,\mu_k\right)^{\mathrm{T}} q^k+o\left(\beta\right) \\ &\le \beta h_i\left(z^k,\mu_k\right)^{\mathrm{T}} q^k+o\left(\beta\right) \le 0. \end{aligned}$$

According to the analysis above, the result is true.

Lemma 3.4. *Algorithm B generates infinite sequence* $\left\{\left(z^k,\mu_k\right)\right\}$, *whose any accumulation points* $\left(z^*,\mu_*\right)$ *are* $K-T$ *points of* (1.1).

Proof. Suppose that $\left\{\left(z^k,\mu_k\right)\right\} \to \left(z^*,\mu_*\right), k \in K$. From (2.17), (2.18), (2.20) and Lemma 2.2, we know that $\left\{\theta_c\left(z^k,\mu_k\right)\right\}$ is a descent sequence. While, for $k \in K, k \to \infty$, it is obvious that $\theta_c\left(z^k,\mu_k\right) \to \theta_c\left(z^*,\mu_*\right)$. So

$$\theta_c\left(z^k,\mu_k\right) \to \theta_c\left(z^*,\mu_*\right), k \to \infty. \tag{3.13}$$

Now we consider the following two cases:

(1) Suppose there exists an infinite subset $K_1 \subseteq K$ such that

$$\left(z^{k+1},\mu_{k+1}\right) = \left(z^k,\mu_k\right)+\lambda_k d^k,$$

which is obtained by Step 3 and Step 5. In view of $\lambda_k \ge \sigma, k \in K_1$ in Step 3, it follows from (2.17) and (2.18) that

$$\lim_{k\in K_1}\left(\theta_c\left(z^{k+1},\mu_{k+1}\right)-\theta_c\left(z^k,\mu_k\right)\right) \le \lim_{k\in K_1}\alpha\lambda\nabla\theta_c\left(z^k,\mu_k\right)^{\mathrm{T}} d_0^k \le \lim_{k\in K_1}\left(-\alpha\varepsilon\xi\left\|d_0^k\right\|^{\delta_0}\right) \le 0.$$

Obvious, $d_0^k \to 0, k \to K_1$. Again, $d_0^k \to d_0^*, k \in K$, so we have $d_0^* = 0$. Imitating the proof of Lemma 3.1, it is easy to see that $\left(z^*,\mu_*\right)$ is a $K-T$ point of (1.5).

(2) Assume the iteration $\left(z^{k+1},\mu_{k+1}\right) = \left(z^k,\mu_k\right)+\lambda_k d^k, \forall k \in K$ is generated by Step 4 and Step 5. Suppose by contradiction that $\left(z^*,\mu_*\right)$ is not a $K-T$ point of (1.5). Then, from (3.12) and Lemma 3.3, we have

$$0 = \lim_{k\in K}\left(\theta_c\left(z^{k+1},\mu_{k+1}\right)-\theta\left(z^k,\mu_k\right)\right) \le \lim_{k\in K}\upsilon\beta_k\nabla\theta_c\left(z^k,\mu_k\right)^{\mathrm{T}} q^k \le \frac{1}{2}\upsilon\beta_*\nabla\theta_c\left(z^*,\mu_*\right)^{\mathrm{T}} q^* < 0,$$

which is a contradiction. Thus, the claim holds.

Theorem 3.2. *The* $K-T$ *point* $\left(z^*, \mu_*\right)$ *of* (1.5) *must be the one of* (1.4), *where* $\mu_* = 0$.

Proof. If $\left(z^*, \mu_*\right)$ is a $K-T$ point of (1.5), then there exists multiplier π^* such that

$$\begin{pmatrix} \nabla f\left(s^*\right) \\ 0_{m\times 1} \\ 0 \end{pmatrix} - c_* \sum_{i\in(T\setminus I_1)} h_i^* + \sum_{i\in\left(J\bigcup(T\setminus I_1)\right)} \pi_i^* h_i^* = 0, \tag{3.14}$$

$$\pi_i^* \geq 0,\ \pi_i^* r_i\left(z^*, \mu_*\right) = 0,\ i \in J \bigcup \left(T \setminus I_1\right).$$

Set

$$A_* = \left(h_i^* \middle| i \in J \bigcup \left(T \setminus I_1\right)\right),\ \tilde{\pi}^* = -\left(A_*^{\mathrm{T}} A_*\right)^{-1} A_*^{\mathrm{T}} \begin{pmatrix} \nabla f\left(s^*\right) \\ 0_{m\times 1} \\ 0 \end{pmatrix}.$$

Obvious, $\tilde{\pi}^k_{J\bigcup(T\setminus I_1)} \to \tilde{\pi}^*_{J\bigcup(T\setminus I_1)}$. While, from (3.14) we get

$$\pi^*_{J\bigcup(T\setminus I_1)} = -\left(A_*^{\mathrm{T}} A_*\right)^{-1} A_*^{\mathrm{T}} \left(\begin{pmatrix} \nabla f\left(s^*\right) \\ 0_{m\times 1} \\ 0 \end{pmatrix} - c_* \sum_{i\in(T\setminus I_1)} h_i^* \right)$$

i.e.

$$\pi^*_{J\bigcup(T\setminus I_1)} = \tilde{\pi}^*_{J\bigcup(T\setminus I_1)} + c_* \sum_{i\in(T\setminus I_1)} \left(A_*^{\mathrm{T}} A_*\right)^{-1} A_*^{\mathrm{T}} h_i^*.$$

Thereby,

$$\pi_i^* = \tilde{\pi}_i^* + c_*,\ \ i \in \left(T \setminus I_1\right). \tag{3.15}$$

According to the definition of c_*, it is clear that

$$\pi_i^* = \tilde{\pi}_i^* + c_* \geq \delta_1 > 0,\ \ i \in \left(T \setminus I_1\right). \tag{3.16}$$

In addition, combining with (3.2) (3.16), we obtain

$$r_i\left(z^*, \mu_*\right) = 0,\ i \in \left(T \setminus I_1\right),\ \mu_* = 0. \tag{3.17}$$

Let $\bar{\pi}^* = \left(\bar{\pi}_i^*, i \in T\right)$, where $\bar{\pi}_i^* = \pi_i^*, i \in J \bigcup \left(T \setminus I_1\right), \bar{\pi}_i^* = 0, i \in I_1 \setminus J$. It follows from (3.14) and (3.17) that $\left(z^*, 0, \bar{\pi}^*\right)$ is a $K-T$ point pair of (1.4).

Theorem 3.3. *Suppose* (2.2) *holds at* z^*. *If* $\left(z^*, 0\right)$ *is a* $K-T$ *point of* (1.4), *then* s^* *is a* $K-T$ *point of* (1.1).

Proof. According to Theorem 3.2 and (2.1), Proposition 2.1 and Proposition 2.2 imply s^* is a $K-T$ point of (1.1).

4. Superlinear Convergence

Now we discuss the convergence rate of the Algorithm B, and prove that the sequence $\left(z^k, \mu_k\right)$ generated by the Algorithm B is one-step superlinearly convergent. For this purpose, we add some stronger regularity assumptions.

H 4.1. *The bounded sequence* $\left\{\left(z^k, \mu_k\right)\right\}$ *possesses an accumulation point* $\left(z^*, \mu_*\right)$, *at which second-order sufficiency condition and strict complementary slackness hold, where* $\zeta^* = \left(\zeta_i^*, i \in T\right)$ *is the corresponding multiplier of* $\left(z^*, \mu_*\right)$.

Lemma 4.1. *Under H* 2.1-*H* 4.2, *we have that*

$$\lim_{k\to\infty}\left\|z^{k+1}-z^k\right\|=0,\ \lim_{k\to\infty}\left(\mu_{k+1}-\mu_k\right)=0.$$

Proof. For $\left(z^{k+1},\mu_{k+1}\right)=\left(z^k,\mu_k\right)+\lambda_k d^k$ generated by Step 3 and Step 5, from (2.17) and (2.18), it holds that

$$\theta_c\left(z^k+\lambda_k dz^k,\mu_k+\lambda_k d\mu_k\right)\le\theta_c\left(z^k,\mu_k\right)+\alpha\lambda_k\nabla\theta_c\left(z^k,\mu_k\right)^{\mathrm{T}}d_0^k\le\theta_c\left(z^k,\mu_k\right)+\alpha\xi\lambda_k\left\|d^k\right\|^{\delta_0}.$$

While, for $\left(z^{k+1},\mu_{k+1}\right)=\left(z^k,\mu_k\right)+\lambda_k d^k$ generated by Step 4 and Step 5, from (2.17), (2.20) and Lemma 2.2, we have

$$\theta_c\left(z^k+\lambda_k dz^k,\mu_k+\lambda_k d\mu_k\right)\le\theta_c\left(z^k,\mu_k\right)+\upsilon\lambda_k\nabla\theta_c\left(z^k,\mu_k\right)^{\mathrm{T}}d_0^k\le\theta_c\left(z^k,\mu_k\right)-c_0\upsilon\lambda_k\left\|d^k\right\|^2.$$

So

$$\theta_c\left(z^{k+1},\mu_{k+1}\right)\le\theta_c\left(z^k,\mu_k\right)-\lambda_k\left\|d^k\right\|^2\min\left\{c_0\upsilon,\alpha\xi\left\|d^k\right\|^{\delta_0-2}\right\},\forall k.$$

Passing to the limit $k\to\infty$ and from (3.13), we obtain

$$\lim_{k\to\infty}\lambda_k\left\|d^k\right\|^2\min\left\{c_0\upsilon,\alpha\xi\left\|d^k\right\|^{\delta_0-2}\right\}=0.$$

Thereby

$$\lim_{k\to\infty}\left\|z^{k+1}-z^k\right\|=0,\ \lim_{k\to\infty}\left(\mu_{k+1}-\mu_k\right)=0.$$

Theorem 4.1. *The entire sequence* $\left(z^k,\mu_k\right)$ *converges to* $\left(z^*,0\right)$, *i.e.*, $\left(z^k,\mu_k\right)\to\left(z^*,0\right),k\to\infty$.

In order to obtain the superlinear convergence rate, we make the following assumption.

H 4.2. $B_k\to B_*,k\to\infty,B_*$ *positive definite*.

Lemma 4.2. *If H* 2.1-*H* 4.2 *hold*, *then we get that*

(1) for k large enough, $L_k=I\left(z^*,\mu_*\right)=I_*$.

(2) $\lim_{k\to\infty}d_0^k=0,\ \lim_{k\to\infty}\pi^k=\left(\zeta_i^*,i\in(I_*\cup(T\setminus I_1)\right)$.

Proof. (1) On one hand, by Lemma 3.2, for k large enough, there exists a constant $\bar\epsilon>0$ such that $\bar\epsilon\le\epsilon_{k,l_k}<1$ in Algorithm A. It follows from H 4.1 and the fact $\varrho\left(z^k,\mu_k\right)\to\zeta^*$ that, for k large enough, $I_*\subseteq L_k$.

On the other hand, we assert that $L_k\subseteq I_*$. Otherwise, there exists some index t and infinite subset K such that

$$t\in L_k\setminus I_*,\ r_t\left(z^*,0\right)<0,\ r_t\left(z^k,\mu_k\right)\ge-\epsilon_{k,l_k}\left|\mu_t\left(z^k,\mu_k\right)\right|\ge-\left|\mu_t\left(z^k,\mu_k\right)\right|,\ \forall k\in K.$$

Let $k\in K,k\to\infty$, then

$$0>r_t\left(z^*,0\right)\ge-\left|\mu_t\left(z^*,0\right)\right|=-\zeta_t^*,\ \zeta_t^*>0.$$

It is a contradiction with the complementary slackness condition, which shows that $L_k\subseteq I_*$, *i.e.*, $L_k=I_*$.

(2) According to $\left(z^k,\mu_k\right)\to\left(z^*,0\right)$ and $B_k\to B_*$, the fact $L_k=I_*$ implies $d_0^k\to d_0^*,\pi^k\to\pi^*,k\to\infty$. Again, since $\left(z^*,0\right)$ is a $K-T$ point of (1.5), imitating the proof of Lemma 3.1, we get that

$$d_0^*=0,\quad\nabla\theta_c\left(z^*,\mu_*\right)+A_*\pi^*=0,\quad\pi_i^*r_i=0,\quad\pi_i^*\ge0,\quad i\in I_*.$$

So the uniqueness of $K-T$ multiplier shows $\lim_{k\to\infty}\pi^k=\zeta^*$.

Lemma 4.3. *Under H* 2.1-*H* 4.2, *for k large enough*, d_0^k *with the corresponding multiplier* $\zeta^k=\pi^k+\left(A_k^{\mathrm{T}}B_k^{-1}A_k\right)^{-1}F\left(z^k,\mu_k\right)$ *is a* $K-T$ *point of the following quadratic program*

$$\begin{aligned} &\min \quad \nabla\theta_c\left(z^k,\mu_k\right)^{\mathrm{T}} d+\frac{1}{2}d^{\mathrm{T}}B_k d \\ &\text{s.t.} \quad r_i\left(z^k,\mu_k\right)+\left(h_i^k\right)^{\mathrm{T}} d=0,\ i\in I_*. \end{aligned} \tag{4.1}$$

Proof. Suppose that (d,ζ) is a $K-T$ point pair of (4.1). From (2.12), (2.14) and (4.1), it holds that

$$d=-P_k\nabla\theta_c\left(z^k,\mu_k\right)-Q_k^{\mathrm{T}}F\left(z^k,\mu_k\right),\quad \zeta=\pi^k+\left(A_k^{\mathrm{T}}B_k^{-1}A_k\right)^{-1}F\left(z^k,\mu_k\right)=\zeta^k.$$

In addition, for k large enough, $\pi_i^k>0, i\in I_*$ holds from fact $\lim_{k\to\infty}\pi^k=\zeta^*$ and strict complementarity condition. While, from the definition of V^k, it holds that $d_0^k=-P_k\nabla\theta_c\left(z^k,\mu_k\right)-Q_k^{\mathrm{T}}F\left(z^k,\mu_k\right)=d$. So the claim holds.

Lemma 4.4. (1) *For k large enough, there exist constants* $b,\eta>0$ *such that*

$$\sum_{i\in I_*\cup(T\setminus I_1)}\zeta_i^k r_i^k\le\eta\left\|F\left(z^k,\mu_k\right)\right\|,\quad \nabla\theta_c\left(z^k,\mu_k\right)^{\mathrm{T}} d_0^k\le -b\left\|d_0^k\right\|^2. \tag{4.2}$$

(2) $d^k=d_0^k+d_1^k$ obtained by (2.15) satisfies

$$\left\|d^k\right\|\sim\left\|d_0^k\right\|,\quad \left\|d_1^k\right\|\sim O\left(\left\|d_0^k\right\|^2\right). \tag{4.3}$$

Proof. (1) Since $\left(z^k,\mu_k\right)\to\left(z^*,\mu_*\right)$, and for k large enough, $L_k=I_*$, it is easy to see

$$F\left(z^k,\mu_k\right)\to\left(r_i\left(z^*,\mu_*\right), i\in I_*\right)=0,\quad \zeta^k\to\pi^*>0,\quad k\to\infty. \tag{4.4}$$

Obviously, for k large enough, $\zeta_i^k>0, i\in I_*$. Thereby, there exists a constant $\eta>0$ such that

$$\sum_{i\in I_*}\zeta_i^k r_i^k=-\sum_{i\in I_*}\zeta_i^k\left|r_i^k\right|\le-\eta\left\|F\left(z^k,\mu_k\right)\right\|.$$

In addition, from Lemma 4.3, we see

$$\nabla\theta_c\left(z^k,\mu_k\right)+B_k d_0^k+A_k\zeta^k=0,\quad F\left(z^k,\mu_k\right)+A_k^{\mathrm{T}}d_0^k=0. \tag{4.5}$$

So

$$\begin{aligned} \nabla\theta_c\left(z^k,\mu_k\right)^{\mathrm{T}} d_0^k &=-\left(d_0^k\right)^{\mathrm{T}}B_k d_0^k-\left(A_k^{\mathrm{T}}d_0^k\right)^{\mathrm{T}}\zeta^k=-\left(d_0^k\right)^{\mathrm{T}}B_k d_0^k+\sum_{i\in I_*}\zeta_i^k r_i^k \\ &\le -b\left\|d_0^k\right\|^2-\eta\left\|F\left(z^k,\mu_k\right)\right\|\le -b\left\|d_0^k\right\|^2. \end{aligned} \tag{4.6}$$

(2) Since

$$r_i\left(z^k+dz_0^k,\mu_k+d\mu_0^k\right)=r_i\left(z^k\right)+\left(h_i^k\right)^{\mathrm{T}}d_0^k+O\left(\left\|d_0^k\right\|^2\right),\quad \forall i\in I_*,$$

we know

$$\left\|F\left(z^k+dz_0^k,\mu_k+d\mu_0^k\right)\right\|=O\left(\left\|d_0^k\right\|^2\right).$$

From $\tau\in(2,3), Q_k\to Q_*$ and the boundedness of Q_*, it follows that

$$\left\|d^k\right\|\sim\left\|d_0^k\right\|,\ \left\|d_1^k\right\|\sim O\left(\left\|d_0^k\right\|^2\right).$$

So, the result is true.

In order to obtain the superlinear convergence rate, we make another assumption.

H 4.3. *The sequence of symmetric matrices* $\{B_k\}$ *satisfies*

$$\left\|\bar{P}_k\left(B_k-\nabla^2 L\left(z^k,\mu_k,\zeta^k\right)\right)d_0^k\right\|=o\left(\left\|d_0^k\right\|\right)\Leftrightarrow\left\|\bar{P}_k\left(B_k-\nabla^2 L\left(z^*,0,\zeta^*\right)\right)d_0^k\right\|=o\left(\left\|d_0^k\right\|\right),$$

where

$$\bar{P}_k=E_{n+2m+1}-A_k\left(A_k^{\mathrm{T}}A_k\right)^{-1}A_k^{\mathrm{T}},$$
$$\nabla^2 L\left(z^k,\mu_k,\zeta^k\right)=\nabla^2\theta_{c_k}\left(z^k,\mu_k\right)+\sum_{i\in I_*}\zeta_i^k H_i^k,$$
$$\nabla^2 L\left(z^*,0,\zeta^*\right)=\nabla^2\theta_c\left(z^*,0\right)+\sum_{i\in T}\zeta_i^* H_i^*.$$

Lemma 4.5. *For k large enough, Algorithm B is not implemented on Step* 4, *and*

$$\lambda\equiv 1,\quad\left(z^{k+1},\mu_{k+1}\right)=\left(z^k,\mu_k\right)+\lambda d^k$$

holds in Step 3.

Proof. According to $d_0^k\to 0$ and Lemma 4.4, we have

$$\left\|d_0^k\right\|\sim\left\|d^k\right\|,\quad\nabla\theta_c\left(z^k,\mu_k\right)^{\mathrm{T}}d_0^k\le -b\left\|d_0^k\right\|^2,$$

which shows (2.17) hold. Now we prove that, the arc search (2.19) and (2.18) eventually accept unit step, *i.e.*, $\lambda_k=1$, for *k* large enough.

Firstly, for (2.19), when $i\in T\setminus I_*$, the fact that $d_0^k\to 0, z^k\to z^*, \mu_k\to 0, r_i\left(z^*,0\right)<0$ and the continuity of r_i imply

$$r_i\left(z^k+dz^k,\mu_k+d\mu_k\right)\le 0$$

when $i\in I_*$, using Taylor expansion, we get

$$\begin{aligned}&r_i\left(z^k+dz_0^k+dz_1^k,\mu_k+d\mu_0^k+d\mu_1^k\right)\\&=r_i\left(z^k+dz_0^k,\mu_k+d\mu_0^k\right)+h_i\left(z^k+dz_0^k,\mu_k+d\mu_0^k\right)^{\mathrm{T}}d_1^k+O\left(\left\|d_1^k\right\|^2\right)\\&=r_i\left(z^k+dz_0^k,\mu_k+d\mu_0^k\right)+h_i\left(z^k,\mu_k\right)^{\mathrm{T}}d_1^k+O\left(\left\|d_0^k\right\|^3\right).\end{aligned}\tag{4.7}$$

Again, from

$$A_k^{\mathrm{T}}d_1^k=-\left\|d_0^k\right\|^{\tau}e-F\left(z^k+dz_0^k,\mu_k+d\mu_0^k\right),$$

we see

$$h_i\left(z^k,\mu_k\right)^{\mathrm{T}}d_1^k=-\left\|d_0^k\right\|^{\tau}-r_i\left(z^k+dz_0^k,\mu_k+d\mu_0^k\right),\ i\in I_*.$$

Thus, (4.7) yields

$$r_i\left(z^k+dz^k,\mu_k+d\mu^k\right)=-\left\|d_0^k\right\|^{\tau}+O\left(\left\|d_0^k\right\|^3\right).\tag{4.8}$$

In view of $\tau\in(2,3)$, (2.19) obviously holds when $\lambda_k=1$.

Secondly, we prove that, for *k* large enough, (2.18) holds for $\lambda_k=1$. Denote

$$\begin{aligned}\varphi&=\theta_{c_k}\left(z^k+\lambda dz^k,\mu_k+\lambda d\mu_k\right)-\theta_{c_k}\left(z^k,\mu_k\right)-\alpha\nabla\theta_{c_k}\left(z^k,\mu_k\right)^{\mathrm{T}}d_0^k,\\&=\nabla\theta_{c_k}\left(z^k,\mu_k\right)^{\mathrm{T}}d^k+\frac{1}{2}\left(d_0^k\right)^{\mathrm{T}}\nabla^2\theta_{c_k}\left(z^k,\mu_k\right)^{\mathrm{T}}d_0^k-\alpha\nabla\theta_{c_k}\left(z^k,\mu_k\right)^{\mathrm{T}}d_0^k+\left(\left\|d_0^k\right\|^2\right).\end{aligned}\tag{4.9}$$

From (4.5), we have

$$\nabla\theta_{c_k}\left(z^k,\mu_k\right)^{\mathrm{T}} d^k = -\left(d_0^k\right)^{\mathrm{T}} \nabla^2\theta_{c_k}\left(z^k,\mu_k\right)^{\mathrm{T}} d_0^k - \sum_{i\in I_*}\zeta_i^k\left(h_i^k\right)^{\mathrm{T}} d^k + o\left(\left\|d_0^k\right\|^3\right).$$

Also, by (4.8), it holds that

$$r_i^k + \left(h_i^k\right)^{\mathrm{T}} d^k + \frac{1}{2}\left(d_0^k\right)^{\mathrm{T}} H_i^k\left(d_0^k\right) + o\left(\left\|d_0^k\right\|^2\right) = -\left\|d_0^k\right\|^{\tau} + O\left(\left\|d_0^k\right\|^3\right),\ i\in I_*.$$

So

$$-\sum_{i\in I_*}\zeta_i^k\left(h_i^k\right)^{\mathrm{T}} d^k = -\sum_{i\in I_*}\zeta_i^k r_i^k + \frac{1}{2}\left(d_0^k\right)^{\mathrm{T}}\left(\sum_{i\in I_*}\zeta_i^k H_i^k\right) d_0^k + o\left(\left\|d_0^k\right\|^2\right).$$

Thus, (4.6) yields

$$\begin{aligned}\varphi &= (\alpha-1)\left(d_0^k\right)^{\mathrm{T}} B_k d_0^k + \frac{1}{2}\left(d_0^k\right)^{\mathrm{T}} \nabla^2 L\left(z^k,\mu_k,\zeta^k\right) d_0^k + \sum_{i\in I_*}(1-\alpha)\zeta_i^k r_i^k + o\left(\left\|d_0^k\right\|^2\right)\\ &\le (\alpha-1)\left(d_0^k\right)^{\mathrm{T}} B_k d_0^k + \frac{1}{2}\left(d_0^k\right)^{\mathrm{T}}\left(\nabla^2 L\left(z^k,\mu_k,\zeta^k\right) - B_k\right) d_0^k - (1-\alpha)\eta\left\|F\left(z^k,\mu_k\right)\right\| + o\left(\left\|d_0^k\right\|^2\right).\end{aligned}$$

Denote $\bar{P}_* = E_{n+2m+1} - A_*\left(A_*^{\mathrm{T}}A_*\right)^{-1}A_*^{\mathrm{T}}$, then $\bar{P}_k \to \bar{P}_*$. Set

$$d_0^k = \bar{P}_* d_0^k + y,\quad y = A_*\left(A_*^{\mathrm{T}}A_*\right)^{-1}A_*^{\mathrm{T}} d_0^k. \tag{4.10}$$

Clearly,

$$\begin{aligned} y &= A_*\left(A_*^{\mathrm{T}}A_*\right)^{-1}\left(A_* - A_k\right)^{\mathrm{T}} d_0^k + A_*\left(A_*^{\mathrm{T}}A_*\right)^{-1}A_k^{\mathrm{T}} d_0^k\\ &= o\left(\left\|d_0^k\right\|\right) - A_*\left(A_*^{\mathrm{T}}A_*\right)^{-1}F\left(z^k,\mu_k\right)\end{aligned}$$

while, from (2) and (10), it holds that

$$\|y\| = O\left(\left\|d_0^k\right\|\right),\ \|y\| = o\left(\left\|d_0^k\right\|\right) + O\left(\left\|F\left(z^k,\mu_k\right)\right\|\right).$$

So

$$\begin{aligned}\varphi &\le b\left(\alpha-\frac{1}{2}\right)\left(\left(d_0^k\right)^{\mathrm{T}}\bar{P}_* + y^{\mathrm{T}}\right)\left(\nabla^2 L\left(z^k,\mu_k,\zeta^k\right) - B_k\right) d_0^k - (1-\alpha)\eta\left\|F\left(z^k,\mu_k\right)\right\| + o\left(\left\|d_0^k\right\|^2\right)\\ &= b\left(\alpha-\frac{1}{2}\right)\left\|d_0^k\right\|^2 + o\left(\left\|d_0^k\right\|^2\right) - (1-\alpha)\eta\left\|F\left(z^k,\mu_k\right)\right\| + o\left(\left\|F\left(z^k,\mu_k\right)\right\|\right) \le 0,\end{aligned}$$

which implies the theorem hold.

According to Lemma 4.3, Lemma 4.4 and Lemma 4.5, combining with Theorem 12.3.3 in [15], the following state holds.

Theorem 4.2. *The Algorithm B is superlinearly convergent, i.e.,*

$$\left\|\begin{matrix} z^{k+1}-z^* \\ \mu_{k+1}\end{matrix}\right\| = o\left(\left\|\begin{matrix} z^k - z^* \\ \mu_k\end{matrix}\right\|\right).$$

5. Conclusion

By means of perturbed technique and generalized complementarity function, we, using implicit smoothing strategy, equivalently transform the original problem into a family of general optimization problems. Based on the idea of penalty function, the discussed problem is transformed an associated problem with only inequality constraints containing parameter. And then, by providing explicit searching direction, a new variable metric gradient projection method for MPCC is established. The smoothing factor μ regarded as a variable ensures that we can obtain an exact stationary point of original problem once the algorithm terminates in finite iteration. What's more, the proposed algorithm adjusts penalty parameter automatically. Under some mild conditions, the global convergence is obtained as well as the superlinear convergence rate.

Acknowledgements

The authors are indebted to the anonymous referees for valuable comments and remarks that helped them improve the original version of the paper.

Funding

This work was supported in part by the National Natural Science Foundation (No. 11361018), the Natural Science Foundation of Guangxi Province (No. 2014GXNSFFA118001), the Key Program for Science and Technology in Henan Education Institution (No. 15B110008) and Huarui College Science Foundation (No. 2014qn35) of China.

References

[1] Kocvara, M. and Outrata, J. (1994) On Optimization Systems Govern by Implicit Complementarity Problems. *Numerical Functional Analysis and Optimization*, **15**, 869-887. http://dx.doi.org/10.1080/01630569408816597

[2] Fletcher, R., Leyffer, S., Ralph, D. and Scholtes, S. (2006) Local Convergence of SQP Methods for Mathematical Programs with Equilibrium Constraints. *SIAM*: *SIAM Journal on Optimization*, **17**, 259-286. http://dx.doi.org/10.1137/S1052623402407382

[3] Luo, Z.Q., Pang, J.S. and Ralph, D. (1996) Mathmetical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge. http://dx.doi.org/10.1017/CBO9780511983658

[4] Jiang, H. (2000) Smooth SQP Methods for Mathematical Programs with Nonlinear Complementarity Constaints. *SIAM Journal of Optimization*, **10**, 779-808. http://dx.doi.org/10.1137/S1052623497332329

[5] Kocvara, M. and Outrata, J. (1995) A Nonsmmoth Approach to Optimization Problems with Equilibrium Constraints. In: Ierns, M.C. and Pang, J.D., Eds., *Proceedings of the International Conference on Complementatity Problems*, SIAM Publications, Baltimore, 148-164.

[6] Outrata, J. (1990) On the Numberical Solution of a Class of Stachelberg Problems. *Zeitschrift for Operations Research*, **4**, 255-278.

[7] Zhang, C., Zhu, Z.B., Chen, F.H. and Fang, M.L. (2010) Sequential System of Linear Equations Algorithm for Optimization with Complementary Constraints. *Mathematics Modelling and Applied Computing*, **1**, 71-80.

[8] Zhang, C., Zhu, Z.B. and Fang, M.L. (2010) A Superlinearly Covergent SSLE Algorithm for Optimization Problems with Linear Complementarity Constraints. *Journal of Mathematical Science*: *Advance and Application*, **6**, 149-164.

[9] Huang, Z.H., Lin, G.H. and Xiu, N.H. (2014) Several Developments of Variational Inequalities and Complementarity Problems, Bilevel Programming and MPEC. *Operations Research Transactions*, **18**, 113-133.

[10] Zhang, C., Sun, L.M., Zhu, Z.B. and Fang, M.L. (2015) Levenberg-Marquardt Method for Mathematical Programs with Linearly Complementarity Constraints. *American Journal of Computational Mathematics*, **5**, 239-242. http://dx.doi.org/10.4236/ajcm.2015.53020

[11] Outrata, J., Kocvara, M. and Zowe, J. (1998) Nonsmmoth Approach to Aptimization Problems with Equilibrium Constraints. Kluwer Academic Publishers, The Netherland. http://dx.doi.org/10.1007/978-1-4757-2825-5

[12] Gao, Z.Y., He, G.P. and Wu, F. (2004) Sequential Systems of Linear Equations Algorithm for Nonlinear Optimization Problems—General Constrained Problems. *Applied Mathematics and Computation*, **147**, 211-226. http://dx.doi.org/10.1016/S0096-3003(02)00662-8

[13] Jian, J.B., Qin, Y. and Liang, Y.M. (2007) A Generalized Strongly Sub-Feasible Algorithm for Mathematical Problems with Nonliear Complementarity Constraints. *Numerical Mathematics*: *A Journal of Chinese Universities*, **29**, 15-27.

[14] Zhu, Z.B. and Zhang, K.C. (2004) A New Conjugate Projuction Gradient Method and Superlinear Convergence. *Acta Mathematicae Applicatae Sinica*, **27**, 149-161.

[15] Yuan, Y.X. and Sun, W.Y. (1997) Optimization Theory and Method. Science Press, Beijing.