Far East Journal of Applied Mathematics
© 2017 Pushpa Publishing House, Allahabad, India
http://www.pphmj.com
http://dx.doi.org/10.17654/AM096010001
Volume 96, Number 1, 2017, Pages 1-16 **ISSN: 0972-0960**

AN IMPROVED ODE-TYPE FILTER METHOD FOR NONLINEAR COMPLEMENTARITY PROBLEM

Ke Su and Mingxi Lu

College of Mathematics and Computer Science
Hebei University
Baoding, 071002, P. R. China

Abstract

We propose an improved ODE-type filter method. In this method, only one linear system is required, and the penalty parameter is not needed. Compared to the traditional filter methods, the new approach is more flexible and less computational scale. Under some reasonable conditions, the global convergent result of our algorithm is presented.

1. Introduction

The nonlinear complementarity problem (NCP) is equivalent to constraint optimization problem

$$x \geq 0, \quad F(x) \geq 0, \quad x^T F(x) = 0, \tag{1.1}$$

$x \in R^n$, $F(x): R^n \to R^n$, is second-order continuously differentiable. We introduce an NCP function, and transform the complementarity problem to a optimization problem.

Received: April 20, 2016; Accepted: July 2, 2016

2010 Mathematics Subject Classification: 90C30.

Keywords and phrases: nonlinear complementarity problem, trust region, ODE, filter.

This research is supported by the National Science Foundation of China (No. 11101115, No. 61572011), the National Science Foundation of Hebei Province (No. A2014201033) and the Key Research Foundation of Education Bureau of Hebei Province (No. ZD2015069).

The function $\varphi : R^2 \to R$ is called *NCP function*, if $\varphi(a, b) = 0$, if and only if $a \geq 0, b \geq 0, ab = 0$. We use Fischer-Burmeister function $\varphi(a, b) = \sqrt{a^2 + b^2} - a - b$ as NCP function. The nonlinear complementarity problem (1) is equivalent to nonlinear system of equations

$$\phi(x) = \begin{pmatrix} \sqrt{x_1^2 + F_1^2(x)} - x_1 - F_1(x) \\ \vdots \\ \sqrt{x_n^2 + F_n^2(x)} - x_n - F_n(x) \end{pmatrix}. \tag{1.2}$$

Thus, problem (1) is equivalent to the optimization problem to solve

$$\min f(x) = \frac{1}{2}\phi(x)^T \phi(x). \tag{1.3}$$

This paper is concerned with finding a solution to the constrained nonlinear optimization problem as following:

$$\begin{aligned} &\min f(x) \\ &\text{s.t. } c_j(x) \geq 0, \ j = 1, 2, \ldots, 2n, \end{aligned} \tag{1.4}$$

where $f(x) = \| x^T F(x) \|_2^2$ and

$$c(x) = (x_1, x_2, \ldots, x_n, F_1(x), F_2(x), \ldots, F_{2n}(x))^T : R^n \to R^{2n},$$

let

$$c_1(x) = x_1, c_2(x) = x_2, \ldots, c_n(x) = x_n, c_{n+1}(x) = F_1(x),$$

$$c_{n+2}(x) = F_2(x), \ldots, c_{2n}(x) = F_n(x).$$

$F(x)$ is second-order continuously differentiable.

There are many methods for inequality constrained nonlinear programming (NLP). But up to now, most algorithms proposed are descent methods, in that they only accept the trial point as next iterate if its merit function value is strictly less than that at the current iterate. It has two drawbacks, one is that the choice of penalty parameter is very difficult if we use penalty function as a merit function, the other is that the descent methods can result on reduction of convergence rate when the iterate is trapped near a narrow curved valley. While, filter method, proposed by Fletcher and Leyffer [1], overcome the drawbacks above. In filter method, the use of a penalty function is replaced by the introduction of so-called filter. So, they have several advantages. Recently, this technique is applied to the many kinds of nonlinear problems combined with trust region method which is to be proved robust [2-4, 6].

ODE methods for minimizing a function $f(x)$ proceed by following the solution curve of a system of ordinary differentiable equations and is more reliable, accurate and efficient than conventional Newton and quasi-Newton algorithms.

Motivated by the referees [2, 5, 6], we transform the nonlinear inequality problems to a nonlinear equation, so that an improved ODE-type filter trust region method is proposed. The filter technique is employed to determine whether to accept the trial point or not. This paper is organized as follows. The next section introduces the concepts we needed. In Section 3, an improved method is put forward. The convergent properties are analyzed in Section 4. Some numerical results are reported in the last section.

2. Preliminaries

Lagrangian function $L : R^{n^2+2n} \to R$ of the problem NLP is defined by

$$L(x, \lambda) = f(x) + \sum_{j=1}^{2n} \lambda_j c_j(x), \{j = 1, 2, ..., 2n\}. \tag{2.1}$$

It is easy to obtain the KKT conditions of problem (NLP) as following:

$$\nabla_x L(x, \lambda) = 0,$$

$$\lambda_j \geq 0, c_j(x) \geq 0, \lambda_j c_j(x) = 0, j = \{1, 2, ..., 2n\}, \tag{2.2}$$

where $\lambda = (\lambda_1, \lambda_2, ..., \lambda_{2n})^T \in R^{2n}$ is multiplier vector.

Definition 2.1. Let $c : R^n \to R^n$ be locally Lipschitz continuous. Then the *B*-differential of *c* at $x \in R^n$ is defined by

$$\partial_B(c(x)) = \{V \in R^{n\times n} \mid V = \lim_{x_k \to x} \nabla c(x_k),\ x_k \in D_c\}.$$

The generalized Jacobian of *c* at *x* in the sense of Clarke is defined by

$$\partial(c(x)) = conv\partial_B(c(x)), \tag{2.3}$$

where $D_c = \{x \in R^n : c$ at x is differential$\}$. Symbol $\mathrm{conv}(x)$ denotes the convex hull of set *S*.

Definition 2.2. Let $G : R^n \to R^n$ be locally Lipschitz continuous, we call *G* at *X* semi-smooth if for all $h \in R^n$, there exists

$$\lim_{V\in\partial G(x+th')\cdot h'\to h\cdot t\downarrow 0}(Vh').$$

Definition 2.3. The function $\varphi : R^2 \to R$ is called *NCP function*, if $\varphi(a, b) = 0$ if and only if $a \geq 0,\ b \geq 0,\ ab = 0.$

One of the most popular functions is Fischer-Burmeister NCP function [7]:

$$\varphi(a, b) = \sqrt{a^2 + b^2} - a - b. \tag{2.4}$$

Lemma 2.1. *Let the function* ∇f *be semi-smooth at X on* R^{n^2}, *then* ∇f *is direction differentiable, and for* $\forall V \in \partial(\nabla f(x + h))$, *we have*

(i) $Vh - (\nabla f)'(x;\ h) = o(\| h \|);$

(ii) $\nabla f(x + h) = \nabla f(x) + (\nabla f)'(x;\ h) + o(\| h \|)$

as h decreases infinitely, where

$$(\nabla f)'(x;\, h) = \lim_{t \downarrow 0}\left(\frac{\nabla f(x + th) - \nabla f(x)}{t}\right)$$

is called directional derivative ∇f *along the direction of h at x.*

Lemma 2.2. *Let V be a neighborhood of x, and* $\varphi : R^n \to R^n$ *is a* LC^1 *function, then for* $x + d \in V$, *there exists* $\beta > 0$, *such that*

$$|\, \varphi(x + d) - \varphi(x) - \nabla\varphi(x)^T d \,| \le \frac{\beta \| d \|^2}{2}.$$

By making use of the F-B NCP function φ, KKT conditions (2.2) can be reformulated to the following form:

$$H(z) = \begin{pmatrix} \nabla_x(L(x, \lambda)) \\ \phi(x, \lambda) \end{pmatrix} = 0, \tag{2.5}$$

where

$$Z = (x^T, \lambda^T) \in R^{\{n^2 + 2n\}},$$

$$\phi(x, \lambda) = (\varphi(c_1(x), \lambda_1), \varphi(c_2(x), \lambda_2), \ldots, \varphi(c_{\{2n\}}(x), \lambda_{\{2n\}}))^T.$$

By the idea of split [8], non-differentiable function $H(z)$ can be split into

$$H(z) = p(z) + q(z), \tag{2.6}$$

where $p(z)$ is a smooth function, $q(z)$ is a non-smooth function, but compared to $p(z)$, $q(z)$ is a relatively small in the sense of the norm.

Remark 1. In fact, for $\forall \varepsilon > 0$, defined function $\varphi^\varepsilon : R^2 \to R$ as follows:

$$\phi^\varepsilon(a, b) = \begin{cases} \varphi(a, b), & \text{if } \sqrt{a^2 + b^2} \ge \varepsilon, \\ \varphi(a, b) + \dfrac{(\sqrt{a^2 + b^2} - \varepsilon)^2}{2\varepsilon}, & \text{otherwise.} \end{cases}$$

Let

$$\phi^{\varepsilon}(x, \lambda) = (\varphi^{\varepsilon}(c_1(x), \lambda_1), \varphi^{\varepsilon}(c_2(x), \lambda_2), \ldots, \varphi^{\varepsilon}(c_{\{2n\}}(x), \lambda_{\{2n\}}))^T,$$

$$p(z) = \begin{pmatrix} \nabla_x(L(x, \lambda)) \\ \phi^{\varepsilon}(x, \lambda) \end{pmatrix}, \tag{2.7}$$

$$q(z) = H(z) - p(z) = \begin{pmatrix} 0 \\ \phi(x, \lambda) - \phi^{\varepsilon}(x, \lambda) \end{pmatrix} = 0 \tag{2.8}$$

it is easy to prove that $p(z)$ is continuously differentiable, and $q(z)$ is continuous but not differentiable, we have

$$\| q(z) \| \leq \frac{1}{2}\sqrt{m\varepsilon}, \ \forall z \in R^{n^2+2n}. \tag{2.9}$$

The above equation (2.9) indicates that split equation (2.6) is meaningful. In filter method, which is proposed by Fletcher and Leyffer [1], the acceptability of steps is determined by comparing the value of constraint violation and the objective function with previous iterates collected in a filter. Different from the traditional filter method, we define the objective function by $l(z)$ by $l(z) = \| \nabla_l(x, \lambda) \|$, and the violation function $\theta(z)$ by $\theta(z) = \| \phi(x, \lambda) \|$. So a trial point should either reduce the value of constraint violation $\theta(z)$ or that of the function $l(z)$.

Definition 2.4. A pair $(l(z_k), \theta(z_k))$ is said to dominate another pair $(l(z_j), \theta(z_j))$ if both $l(z_k) \leq l(z_j)$ and $\theta(z_k) \leq \theta(z_j)$. We also call this a point z_k dominates another point z_j.

Definition 2.5. A filter $\mathcal{F}$ is a list of pairs $(l(z_j), \theta(z_j))$ such that no pair dominates any other.

For convenience, we denote $z_j = (l_j, \theta_j)$. A new trial point z_k^+ is accepted if it is not dominated by any points in $\mathcal{F} \cup z_k$. Consider the convergence property of the algorithm, we call a new trial point is accepted, if and only if for $z_j \in \mathcal{F}$ it holds

$$l_k^+ \le l_j - \gamma\delta(\| H_k^+ \|, \| H_j \|) \text{ or } \theta_k^+ \le \theta_j - \gamma\delta(\| H_k^+ \|, \| H_j \|), \quad (2.10)$$

where $\delta(\| H_k^+ \|, \| H_l \|) = \min\{(\| H(z_k^+) \|, \| H(z_l) \|)\}$, and γ is a small positive number. Let $p_k^+ = p(z_k^+)$, $\theta_k^+ = \theta(z_k^+)$, and so on.

If the trial point z_k is accepted in the sense of (2.10), then we add the pair (l_k^+, θ_k^+) into the filter, that is $\mathcal{F} = \mathcal{F} \cup (l_k^+, \theta_k^+)$. And removed those points which are dominated from the filter. For convenience, we call this the update of the filter.

3. An Improved ODE-type Filter Algorithm

For non-smooth linear equation (2.5), at the *k*th iteration, we use ODE trust region method [9] to obtain the search direction d_k. That is to solve the following system of linear equation

$$\left[h_k(\nabla p(z_k)\nabla p(z_k)^T + B_k) + \frac{I}{h_k}\right]d = -\nabla p(z_k)H(z_k), \qquad (3.1)$$

or the equivalent form

$$[(\nabla p(z_k)\nabla p(z_k)^T + B_k) + I]d = -h_k\nabla(z_k)H(z_k), \qquad (3.2)$$

where $h_k > 0$ is the integral step, $\nabla p(z_k)$ is the gradient of the smooth function p at the point z_k.

The matrix B_k is the $n^2 + 2n$ order symmetric matrix, which can be updated by SR1 correction [3], i.e.,

$$B_{\{k+1\}} = B_k + \frac{(y_k - B_k d_k)(y_k - B_k d_k)^T}{(y_k - B_k d_k)^T d_k}, \qquad (3.3)$$

where

$$y_k = \nabla p(z_{\{k+1\}})H(z_{\{k+1\}}) - \nabla p(z_k)H(z_k) - \nabla p(z_{\{k+1\}})\, p(z_{\{k+1\}})^T d_k.$$

Remark 2. The computational scale of this method to obtain d_k is much less than that of solving quadratic subproblem, meanwhile, the adjustment of the step h_k is much easier in the parameter space.

Algorithm

Step 0. Initialization, choose $z_1,\ h_1 > 0,\ \bar{\varepsilon} \ge 0,\ 0 < \eta < l,\ 0 < \gamma < 1.$ Initial split control value $\varepsilon_1 > 0,$ let $k = 1.$

Step 1. Compute $H(z) = p(z) + q(z)$ and $\nabla p(z_k).$

Step 2. If $\| \nabla p(z_k) H(z_k) \| \le \bar{\varepsilon},$ stop.

If $h_k(\nabla p(z_k)\nabla p(z_k)^T + B_k) + I$ is positive definite, go to Step 3. Otherwise, let m_k is the smallest integer with which the symmetric matrix $2^{-m_k} h_k(\nabla p(z_k)\nabla p(z_k)^T + B_k) + I$ is positive definite, let $h_k = 2^{-m_k} h_k.$

Step 3. Solve (3.2) to get $d_k.$ Let $z_k^+ = z_k + d_k,$ compute $r_k = \dfrac{M(z_k) - M(z_k^+)}{M(z_k) - q_k(z_k)},$ where $M(z) = \dfrac{1}{2}\|H(z)\|^2,\ q(d) = \dfrac{1}{2}\| H(z) + \nabla p(z_k)^T d \|^2 + \dfrac{1}{2} d^T B_k d.$

Step 4. Denote $l_k^+ = l(z_k^+),\ \theta_k^+ = \theta(z_k^+).$ If $(l_k^+,\ \theta_k^+)$ is not accepted by the filter, go to Step 5. Otherwise $z_{\{k+1\}} = z_k^+,$ go to Step 6.

Step 5. If $r_k \ge \eta,\ z_{\{k+1\}} = z_k^+,\ h_{\{k+1\}} = 2h_k,$ go to Step 7. Otherwise, $h_k = \dfrac{1}{2} h_k,\ \varepsilon_k = \| d_k \|^2$ go to Step 3 (inner loop).

Step 6. If $r_k \ge \eta,\ h_{\{k+1\}} = 2h_k$ go to Step 7. Otherwise, add the pair $(l_k^+,\ \theta_k^+)$ to the filter $\mathcal{F}$ and update, $h_k = \dfrac{1}{2} h_k,$ go to Step 7.

Step 7. Let $\varepsilon_{\{k+1\}} = \| d_k \|^2,$ update B_k to $B_{k+1},\ k = k + 1,$ go to Step 1 (outside loop).

4. Convergence Properties

In this section, in order to present a proof of global convergence of algorithm, we always assume that following conditions hold.

Assumptions

A1. The iterate x^k remains in a closed, bounded convex subset $S \in R^n$.

A2. There exist two constants $0 < a \le b$ such that $a\| d \|^2 \le d^T B_k d \le b\| d \|^2$, for all k and for all $d \in R^n$.

Suppose that $Ared_k = M(z) - M(z_k^+)$, $Pred_k = M(z_k) - q_k(d_k)$, then $r_k = \frac{Ared_k}{Pred_k}$. By Lemma 2.2, we obtain some results as following.

Lemma 4.1. $| Ared_k - Pred_k | = O(\| d_k \|^2)$.

Lemma 4.2. $Pred_k \ge \frac{1}{2h_k} \| d_k \|^2$.

Proof. By (3.2), we have

$$\begin{aligned} Pred_k &= M(z_k) - q_k(d_k) \\ &= \frac{1}{2}\| H(z) \|^2 - \frac{1}{2} d^T B_k d_k - \frac{1}{2} \| h(z_k) + \nabla p(z_k)^T d_k \|^2 \\ &= -(\nabla p(z_k)^T H(z_k) d_k - \frac{1}{2} d_k^T [\nabla p(z_k) \nabla p(z_k)^T + B_k] d_k) \\ &= d_k^T \left[\nabla p(z_k) \nabla p(z_k)^T + B_k + \frac{1}{h_k} I \right] d_k \\ &\quad - \frac{1}{2} d_k^T [\nabla p(z_k) \nabla p(z_k)^T + B_k] d_k^T \\ &= \frac{1}{2} d_k^T \left[\nabla p(z_k) \nabla p(z_k)^T + B_k + \frac{1}{h_k} I \right] d_k + \frac{1}{2h_k} d_k^T d_k \\ &\ge \frac{1}{2h_k} \| d_k \|^2. \end{aligned} \tag{4.1}$$

Theorem 4.3. *The inner loop of the Algorithm terminates in finite number of times.*

Proof. Assume the conclusion is not true, we have $r_k < \eta$ by the algorithm as k increases infinitely, and $h_k \to 0$. By (3.2), it holds $d_k \to 0$, so

$$|r_k - 1| = \left| \frac{Ared_k - Pred_k}{Pred_k} \right| \le \frac{O(\| d_k \|^2)}{\frac{1}{2h_k} \| d_k \|^2} \to 0. \qquad (4.2)$$

It shows that $r_k \ge \eta$ as k increases infinitely and the desired conclusion holds.

In order to analyse the convergence properties of the algorithm, we define some sets as following:

Let $A = \{k \mid (l_k^+, \theta_k^+)\}$ is accepted by $\{F_k\}$ represent the index set which is accepted by the filter.

$N = \{k \mid z_{\{k+1\}} = z_k^+\}$ represent the index set which successful iteration.

$S = \{k \mid (l_k^+, \theta_k^+)$ add to the filter$\}$, represent the index set which filter update.

Based on Algorithm, we have $A \subseteq N$, and there are several cases as following:

(1) $|A| < \infty$, $|N| < \infty$ (it must hold $|S| < \infty$),

(2) $|A| < \infty$, $|N| = \infty$ (it must hold $|S| < \infty$),

(3) $|A| = \infty$, $|N| < \infty$ (it must hold $|S| = \infty$),

(4) $|S| < \infty$ (it must hold $|A| = \infty$, $|N| = \infty$).

Then we discuss the convergence properties of the algorithm according to these four cases, respectively.

Lemma 4.4. *Suppose there are finitely many points added to the filter* $(|A| < \infty)$, *and there are finitely times successful iteration* $(|N| < \infty)$, *then the algorithm can terminate in finite times. It holds* $\nabla p(z_k)H(z_k) = 0$.

Proof. Suppose by contradiction that there exists a constant $\varepsilon > 0$, such that $\| \nabla p(z_k)H(z_k) \| \geq \varepsilon$. Suppose k_0 be the last successful iteration, then for $\forall j$, we have $z_{k_0+1} = z_{k_0+j}$ and $z_{k_0+j} < \eta$. Based on the algorithm, it holds $h_k \to 0\,(k \to \infty)$. Similar to Theorem 4.3, we obtain the contradiction and the result follows.

Lemma 4.5. *Suppose there are finitely many points added into the filter* $(|A| < \infty)$, *and there are infinitely times successful iteration* $(|N| = \infty)$, *then*

$$\liminf_{k\to\infty} \| \nabla p(z_k)H(z_k) \| = 0.$$

Proof. By algorithm, if $|A| < \infty, |N| = \infty$, that means (l_k^+, θ_k^+) is not accepted by filter as k increases infinitely. By Theorem 4.3, it holds $r_k \geq \eta$, and $\exists h > 0$ such that $h_k \geq h$ as k increases infinitely.

Suppose by contradiction that there exists a constant $\varepsilon > 0$, and a positive index set k_0 such that $\| \nabla p(z_k)H(z_k) \| \geq \varepsilon$ as $k \geq k_0$. By the algorithm and Lemma 4.4, we have

$$Ared_k \geq \eta Pred_k \geq \frac{\eta}{2} d_k^T \left[\nabla p(z_k) \nabla p(z_k)^T + B_k + \frac{1}{h_k} I \right] d_k$$

$$= -\frac{\eta}{2} (\nabla p(z_k)H(z_k))^T d_k \geq 0. \tag{4.3}$$

Thus, the sequence $\{h_k\}$ is monotonically decreasing, and by A_1, it is lower bounded. So $h_k - h_{k+1} \to 0\,(k \to \infty)$, hence

$$(\nabla p(z_k)H(z_k))^T d_k \to 0\,(k \to \infty). \tag{4.4}$$

On the other hand

$$
\begin{aligned}
&|\,(\nabla p(z_k)H(z_k))^T d_k\,| \\
&= (\nabla p(z_k)H(z_k))\left(\nabla p(z_k)\nabla p(z_k)^T + B_k + \frac{1}{h_k}I\right)^{\{-1\}}(\nabla p(z_k)H(z_k)) \\
&\geq \frac{\|\nabla p(z_k)H(z_k)\|^2}{\left\|\nabla p(z_k)\nabla p(z_k)^T + B_k + \frac{1}{h_k}I\right\|} \\
&\geq \frac{\varepsilon^2}{C^2 + M + \frac{1}{h}} \\
&> 0 \qquad (4.5)
\end{aligned}
$$

which contradicts to the (4.4). Hence the desired result follows.

Lemma 4.6. *If there exists an infinite sequence of points is accepted by the filter* $(|A| = \infty)$, *and finitely many points added to the filter* $(|S| < \infty)$, *it holds*

$$\liminf_{k\to\infty}\|\nabla p(z_k)H(z_k)\| = 0.$$

By algorithm, it is similar to Lemma 4.5.

Lemma 4.7. *Suppose there exist infinitely many points added to the filter* $(|S| = \infty)$. *Then*

$$\liminf_{k\to\infty}\|\nabla p(z_k)H(z_k)\| = 0.$$

Proof. Suppose by contradiction that there exists a constant $\varepsilon > 0$ such that $\|\nabla p(z_k)H(z_k)\| \geq \varepsilon$. Consider the iteration in S. Suppose there exists a subsequence $\{k_i\}$ such that $S = \{k_i\}$, then $z_{k_i} = z^+_{k_i - 1}$. It follows that there exists a subsequence $\{k_l\} \subseteq \{k_i\}$ such that

$$\liminf_{l\to\infty}(\nabla p(z_{k_l})H(z_{k_l})) = \nabla p(z_\infty)H(z_\infty) \text{ and } \|\nabla p(z_\infty)H(z_\infty)\| \geq \varepsilon. \quad (4.6)$$

By the definition of $\{k_i\}$, z_{k_l} is accepted by the filter for $\forall l$. Then as l increases infinitely, it holds

$$h_{k_l} \leq h_{k_l-1} - \gamma\min\{\|H_{k_l}\|, H_{k_l-1}\} \text{ or } \theta_{k_l} \leq \theta_{k_l-1} - \gamma\min\{\|H_{k_l}\|, H_{k_l-1}\}. \quad (4.7)$$

By the assumptions, there exists a number $\bar{\delta} > 0$ such that $\min\{\|H_{k_l}\|, H_{k_l-1}\} \geq \bar{\delta}$. Then

$$h_{k_l} - h_{k_l-1} \leq -\gamma\bar{\delta} < 0 \text{ or } \theta_{k_l} - \theta_{k_l-1} \leq -\gamma\bar{\delta} < 0. \quad (4.8)$$

By (4.6), the left of inequality tends to 0, it is a contradiction. So we have

$$\liminf_{i\to\infty}\|\nabla p(z_{k_i})H(z_{k_i})\| = 0. \quad (4.9)$$

Now consider $l \nsubseteq \{k_i\}$, let $\{k_{i(l)}\}$ be the last iteration before l make $z_{k_{i(l)}}$ add to the filter. By algorithm, $r_k \geq \eta$ for $l \not\subset \{k_i\}$. Then it holds

$$\|\nabla p(z_{k_l})H(z_{k_l})\| \leq \|\nabla p(z_{k_{i(l)}})H(z_{k_{i(l)}})\|.$$

Together with (4.9) and $k_{i(l)} \in \{k_i\}$, the desired conclusion follows.

By Lemma 4.4 and Lemma 4.7, we obtain the convergence conclusion.

Theorem 4.8. *Suppose that Assumptions hold, and* $\nabla p(z)$ *is nonsingular for all* $z \in S$. *Then the sequence* $\{z_k\}$ *generated by the algorithm satisfies two cases as following*:

(1) *Iteration terminated at the KKT point of the original problem* (*NLP*).

(2) *Every accumulation point is a KKT point of the original problem* (*NLP*).

Theorem 4.9. *Suppose* $z_k \to z^*$, $h_k \to \infty$, *and* $\nabla p(z_k)H(z_k)$ *is semi-smooth in* ∇z^*, *there exist constants* μ_1, μ_2 *for* $\nabla\forall k$ *and* $V_k \in \partial(\nabla p(z_k)H(z_k))$ *it holds*

$$d^T\left(\theta_k + \frac{1}{h_k}\right)d \geq \mu_1 \| d \|^2 \text{ and } \| V_k - Q_k \| \leq \frac{\mu_2}{h_k}, \tag{4.10}$$

where $Q_k = \nabla p(z_k)\nabla p(z_k)^T + B_k$. *If* $\forall z \in S$ *and* $\nabla p(z)$ *is nonsingular, then*

(1) z^* *is the KKT point of the original problem.*

(2) $\{z_k\}$ *converges to* z^* *superlinearly.*

Proof. The former part of the theorem follows the Theorem 4.8. Now we turn to prove the second part of the theorem. $\nabla p(z_k)H(z_k)$ is semi-smooth in z^*, by Lemma 2.1, as h decreases infinitely, it holds

$$\nabla p(z_k + h)H(z_k + h) = \nabla p(z_k)H(z_k) + V_k + o(\| z^* - z_k \|). \tag{4.11}$$

Based on that $z_k \to z^*$ as k increase infinitely, it holds

$$\nabla p(z^*)H(z^*) = \nabla p(z_k)H(z_k) + V_k(z_k - z^*) + o(\| z^* - z_k \|). \tag{4.12}$$

And z^* is KKT point of the original problem, that is $H(z^*) = 0$, then we have

$$\nabla p(z_k)H(z_k) = V_k(z_k - z^*) + o(\| z_k - z^* \|). \tag{4.13}$$

By (4.11)-(4.13) and $h_k \to \infty$, we obtain

$$\| z_{k+1} - z^* \|$$

$$= \| z_k + d_k - z^* \| = \left\| z_k - z^* - \left(Q_k + \frac{1}{h_k} I\right)^{-1} \nabla p(z_k)H(z_k) \right\|$$

$$\geq \left\| \left(Q_k + \frac{1}{h_k} I \right)^{-1} \right\| \left\| \left(Q_k + \frac{1}{h_k} I \right)(z_k - z^*) - \nabla p(z_k) H(z_k) \right\|$$

$$\geq \left\| \left(Q_k + \frac{1}{h_k} I \right)^{-1} \right\| \left(\| Q_k(z_k - z^*) - \nabla p(z_k) H(z_k) \| + \frac{\| z_k - z^* \|}{h_k} \right)$$

$$\geq \left\| \left(Q_k + \frac{1}{h_k} I \right)^{-1} \right\|$$

$$\times \left(\| V_k(z_k - z^*) - \nabla p(z_k) H(z_k) \| + \| V_k - Q_k \| \| z_k - z^* \| + \frac{\| z_k - z^* \|}{h_k} \right)$$

$$= o(\| z_k - z^* \|) \tag{4.14}$$

which yields the desired conclusion.

References

[1] R. Fletcher and S. Leyffer, Nonlinear programming without a penalty function, Math. Program. 91(2) (2002), 239-269.

[2] N. I. M. Gould, S. Leyffer and Ph. L. Toint, A multidimensional filter algorithm for nonlinear equation and nonlinear least-squares, SIAM J. Control Optimization 15(1) (2004), 17-38.

[3] Y. X. Yuan and W. Y. Sun, Optimization Theory and Methods, Science Press, Beijing, China, 1997.

[4] W. Sun, R. J. B. Sampaio and J. Yuan, Quasi Newton trust region algorithm for nonlinear least squares problems, Appl. Math. Comput. 105(2) (1999), 183-194.

[5] Y. G. Ou, A trust region method for constrained optimization problems with nonlinear inequalities, Appl. Math. 19(1) (2006), 80-85.

[6] N. I. M. Gould and Toint Ph L. Sainvitu, A filter trust-region method for unconstrained optimization, SIAM J. Optimization 16(2) (2006), 341-357.

[7] B. Chen, X. Chen and C. A. Kancow, Fischer-Bufmeister NCP-function, Math. Program. 88(1) (2000), 211-216.

[8] Tong Xiaojiao and Zhou Shuzi, A trust-region algorithm for inequality constrained optimization, J. Compu. Math. 21(2) (2003), 207-220.

[9] Y. G. Ou and D. P. Hou, A superlinearly convergent ODE-type trust region algorithm for LC1 optimization *n* problems, China Quart. J. Math. 18(2) (2003), 140-145.

[10] W. Hock and Schittknowskik, Test examples for nonlinear programming codes, M. Beakma and H. P. Kunzi, Lecture Notes in Economics and Mathematical Systems 187, Springer Verlag, Berlin, 1981.