

RESEARCH

Open Access



A new filter QP-free method for the nonlinear inequality constrained optimization problem

Youlin Shang¹, Zheng-Fen Jin¹  and Dingguo Pu^{2*}

*Correspondence:

madpu@tongji.edu.cn

²Department of Mathematics,
Tongji University, Shanghai, China
Full list of author information is
available at the end of the article

Abstract

In this paper, a filter QP-free infeasible method with nonmonotone line search is proposed for minimizing a smooth optimization problem with smooth inequality constraints. This proposed method is based on the solution of nonsmooth equations, which are obtained by the Lagrangian multiplier method and the function of the nonlinear complementarity problem for the Karush–Kuhn–Tucker optimality conditions. Especially, each iteration of this method can be viewed as a perturbation of a Newton or quasi-Newton iteration on both the primal and dual variables for the solution of the Karush–Kuhn–Tucker optimality conditions. What is more, it is considered to use the function of the nonlinear complementarity problem in the filter, which makes the proposed algorithm avoid the incompatibility. Then the global convergence of the proposed method is given. And under some mild conditions, the superlinear convergence rate can be obtained. Finally, some preliminary numerical results are shown to illustrate that the proposed filter QP-free infeasible method is quite promising.

MSC: 90C20; 90C30; 90C33

Keywords: Nonlinear constrained optimization; Filter method; QP-free method; Nonmonotone line search

1 Introduction

In this paper, we mainly consider solving the nonlinear optimization problem (NLP) with the inequality constraints, where the objective function and the constrained functions are Lipschitz continuously differentiable functions. We give the Lagrangian function associated with this problem, then the *Karush–Kuhn–Tucker* (KKT) optimality conditions for our solved problem can be obtained.

It is well known that the KKT optimality conditions is a mixed nonlinear complementarity problem (NCP). And this NCP has attracted much attention due to its various applications [1–3] such as the economic equilibrium problem, the restructuring problems of electricity and gas markets, and so on. Of course, there are many efficient methods for solving the NCP, which can be seen in [4–7]. One popular way to solve the NCP is to construct a Newton method for solving the related nonlinear equations, which is a reformulation of the KKT optimality condition. Another way is to use the filter method to directly solve the NLP with the inequality constraints. Recently Pu, Li, and Xue [8] proposed a new quadratic

programming (QP)-free infeasible method for minimizing a smooth function subject to some inequality constraints. This method is based on the solution of nonsmooth equations which are obtained by the multiplier and the Fischer–Burmeister NCP function for the KKT conditions. They proved that the method had a superlinear convergence rate under some mild conditions.

Fletcher and Leyffer [9] proposed a filter method for solving the NLP problem, which was an alternative to the traditional merit functions approach. Provided that there is a sufficient decrease in the objective function or the constraints violation function, it was shown that the trial points generated from solving a sequence of trust region QP subproblems are accepted. In addition, the computational results reported in [9, 10] are also very encouraging. For more related methods, one can refer to [11–16].

Stimulated by the progress in these two aspects, in this paper, we propose a nonmonotone filter QP-free infeasible method for minimizing a smooth function subject to smooth inequality constraints. This proposed iterative method is based on the solution of nonsmooth equations, which are obtained by the multiplier and some NCP functions for the KKT first order optimality conditions. And each iteration of this method can be viewed as a perturbation of a Newton or quasi-Newton iteration on both the primal and dual variables for the solution of the KKT optimality conditions. Specifically, we use the filter on the linear search with a nonmonotone acceptance mechanism [17, 18]. Moreover, we also consider to use the NCP function in the filter. Thus our algorithm can avoid the incompatibility, which may appear in the filter SQP algorithm. We also give the global convergence and the superlinear convergence rate of the proposed method under some mild conditions. Finally, we take some numerical tests to illustrate the effectiveness of the proposed filter QP-free infeasible method.

The rest of this paper is organized as follows. In Sect. 2, we give some preliminaries and the formulation of the solved problem. Then we propose an infeasible filter QP-free method. In Sect. 3, we show that the proposed method is well defined and establish its global convergence and superlinear convergence rate under some mild conditions. Some numerical tests are given in Sect. 4. Finally, we give some brief conclusions in Sect. 5.

2 Preliminaries and algorithm

In this section, we firstly introduce the formulation of the solved problem. Then we give some preliminaries for structuring a new filter QP-free method. Finally, we present the structure of our proposed method in detail.

In this paper, we mainly consider solving the nonlinear optimization problem (NLP) with the inequality constraints, which can be formulated as

$$\min f(x), \quad \text{s.t.} \quad x \in D = \{x \in R^n \mid G(x) \leq 0\}, \tag{1}$$

where $f : R^n \rightarrow R$ and $G = (g_1, g_2, \dots, g_m)^T : R^n \rightarrow R^m$ are Lipschitz continuously differentiable functions.

The Lagrangian function associated with problem (1) is

$$L(x, \mu) = f(x) + \mu^T G(x),$$

where $\mu = (\mu_1, \mu_2, \dots, \mu_m)^T \in R^m$ is the multiplier vector. For simplicity, we use (x, μ) to denote the column vector $(x^T, \mu^T)^T$.

Then we can obtain the KKT point $(\bar{x}, \bar{\mu}) \in R^n \times R^m$ for problem (1), which satisfies the necessary optimality conditions :

$$\nabla_x L(\bar{x}, \bar{\mu}) = 0, \quad G(\bar{x}) \leq 0, \quad \bar{\mu} \geq 0, \quad \bar{\mu}_i g_i(\bar{x}) = 0, \tag{2}$$

where $1 \leq i \leq m$. We also say that $\bar{x} \in D$ is a KKT point of problem (1) if there exists $\bar{\mu} \in R^m$ such that $(\bar{x}, \bar{\mu})$ satisfies (2). It is well known that the KKT optimality condition is a mixed NCP. And the reformulation of (2) can be viewed as the following nonlinear equation:

$$\Phi(x, \mu) = 0.$$

2.1 Preliminaries

In this subsection, we give the definition of Fischer–Burmeister NCP function and some related Jacobian functions in different cases. Both theoretical results and computational experience have indicated that the nonsmooth methods based on the Fischer–Burmeister NCP function are efficient. The Fischer–Burmeister function has a very simple structure, which is defined as

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

It is clear that this function ψ is continuously differentiable everywhere except at the origin, but it is strongly semismooth at the origin, *i.e.*, if $a \neq 0$ or $b \neq 0$, then ψ is continuously differentiable at $(a, b) \in R^2$, and

$$\nabla \psi(a, b) = \left(\frac{a}{\sqrt{a^2 + b^2}} - 1, \frac{b}{\sqrt{a^2 + b^2}} - 1 \right);$$

if $a = 0$ and $b = 0$, then the generalized Jacobian of ψ at $(0, 0)$ is (see [14])

$$\partial \psi(0, 0) = \{(\xi - 1, \eta - 1) \mid \xi^2 + \eta^2 = 1\}.$$

Let $\phi_i(x, \mu) = \psi(-g_i(x), \mu_i)$, $1 \leq i \leq m$. Given the above formulation of problem (1), we can denote $\Phi(x, \mu) = ((\nabla_x L(x, \mu))^T, (\Phi_1(x, \mu))^T)^T$, where $\Phi_1(x, \mu) = (\phi_1(x, \mu), \dots, \phi_m(x, \mu))^T$.

Clearly, the KKT optimality conditions (2) can be equivalently reformulated as the nonsmooth equations $\Phi(x, \mu) = 0$.

If $(g_i(x), \mu_i) \neq (0, 0)$, then ϕ_i is continuously differentiable at $(x, \mu) \in R^{n+m}$. In this case, we have

$$\nabla_x \phi_i = \left(\frac{-g_i(x)}{\sqrt{(g_i(x))^2 + \mu_i^2}} + 1 \right) \nabla g_i(x); \quad \nabla_\mu \phi_i = \left(\frac{\mu_i}{\sqrt{(g_i(x))^2 + \mu_i^2}} - 1 \right) e_i,$$

where $e_i = (0, \dots, 0, 1, 0, \dots, 0)^T \in R^m$ is the i th column of the unit matrix, its i th element is 1, and other elements are 0.

If $g_i(x) = 0$ and $\mu_i = 0$, $1 \leq i \leq m$, then $\phi_i(x, \mu)$ is strongly semismooth and directionally differentiable at (x, μ) . We have

$$\partial_x \phi_i(x, \mu) = \{(\xi + 1) \nabla g_i(x) \mid -1 \leq \xi \leq 1\}$$

and

$$\partial_{\mu_i} \phi_i(x, \mu) = \{(\xi - 1) \mid -1 \leq \xi \leq 1\}.$$

We may reformulate the KKT at point $(\bar{x}, \bar{\mu})$ conditions as a system of equations:

$$\Phi(\bar{x}, \bar{\mu}) = (\nabla_x L(\bar{x}, \bar{\mu}), \Phi_1(\bar{x}, \bar{\mu})) = 0,$$

where $\mu = (\mu_1, \mu_2, \dots, \mu_m)^T \in R^m$ is the multiplier vector, $\phi_j(x, \mu_j) = \psi(-g_j(x), \mu_j)$, $\Phi_1(\bar{x}, \bar{\mu}) = (\phi_1(\bar{x}, \bar{\mu}_1), \phi_2(\bar{x}, \bar{\mu}_2), \dots, \phi_m(\bar{x}, \bar{\mu}_m))^T$. To replace the violation constrained function $p(G(x))$ in the filter F of Fletcher and Leyffer method [9], we use the violation constrained function $p(G(x), \mu) = \|\Phi_1(x, \mu)\|$.

2.2 Algorithm

In this subsection, we give the process and the framework of the filter QP-free method for solving problem (1). We firstly give some closed forms for preparing the method.

If $(g(x^k), \mu^k) \neq (0, 0)$, let $\xi_j^k = \xi_j(x^k, \mu^k) = \frac{-g_j^k}{\sqrt{(g_j^k)^2 + (\mu_j^k)^2}} + 1$; $\eta_j^k = \eta_j(x^k, \mu^k) = \frac{\mu_j^k}{\sqrt{(g_j^k)^2 + (\mu_j^k)^2}} - 1$; otherwise we denote $\xi_j^k = \xi_j(x^k, \mu^k) = 1 + \sqrt{2}/2$; $\eta_j^k = \eta_j(x^k, \mu^k) = -1 + \sqrt{2}/2$. Then let

$$V^k = \begin{pmatrix} V_{11}^k & V_{12}^k \\ V_{21}^k & V_{22}^k \end{pmatrix} = \begin{pmatrix} H^k & \nabla G^k \\ \text{diag}(\xi^k)(\nabla G^k)^T & \text{diag}(\eta^k - c^k) \end{pmatrix}, \tag{3}$$

where H^k is a positive matrix, which may be modified by BFGS update. The $\text{diag}(\xi^k)$ or $\text{diag}(\eta^k - c^k)$ denotes the diagonal matrix whose j th diagonal element is ξ_j^k or $\eta_j^k - c_j^k$ respectively, and

$$c_j^k = c \min\{1, \|\Phi^k\|^\nu\},$$

where $c > 0$ and $\nu > 1$ are given parameters.

Secondly, we give the nonmonotone sequence for structuring our method. We may assume that the elements Φ^k and F^k are sorted in the decreasing order, that is, $\hat{F}^{k1} \geq \hat{F}^{k2} \geq \hat{F}^{k3} \geq \dots \geq \hat{F}^{kl}$, $\hat{\Phi}^{k1} \geq \hat{\Phi}^{k2} \geq \hat{\Phi}^{k3} \geq \dots \geq \hat{\Phi}^{kl}$. Let

$$\bar{\Phi}^k = \begin{cases} \{\Phi^k, \hat{\Phi}^{k2}, \hat{\Phi}^{k3}, \dots, \hat{\Phi}^{kl}\}, & \text{if } \|\Phi^k\| < \hat{\Phi}^{k1} \text{ and } \|\Phi^k\| > 0, \\ \{\hat{\Phi}^{k1}, \hat{\Phi}^{k2}, \hat{\Phi}^{k3}, \dots, \hat{\Phi}^{kl}\}, & \text{if } \Phi^k \geq \hat{\Phi}^{k1} \text{ or } \Phi^k = 0; \end{cases} \tag{4}$$

and

$$\bar{F}^k = \begin{cases} \{F^k, \hat{F}^{k2}, \hat{F}^{k3}, \dots, \hat{F}^{kl}\}, & \text{if } F^k < \hat{F}^{k1}, \\ \{\hat{F}^{k1}, \hat{F}^{k2}, \hat{F}^{k3}, \dots, \hat{F}^{kl}\}, & \text{if } F^k \geq \hat{F}^{k1}. \end{cases} \tag{5}$$

We denote the maximal elements in $\bar{\Phi}^k, \bar{F}^k$ by $p_{\max}^k, \bar{F}_{\max}^k$, respectively.

Based on the above given information, we now give the framework of the nonmonotone filter QP-free infeasible method (NFQPIM) for minimizing a smooth function subject to smooth inequality constraints as follows in Algorithm 1.

Algorithm 1 NFQPIM

Step 0. Initialization. Given an initial guess $x^0 \in R^n$, $\tau \in (0, 1)$, $\bar{\mu} > 0$, $0 < \theta_1 < \theta_1 < 1$, and a positive definite matrix H^0 . Given initial $(f(x^0), \bar{\mu}) \in F^0$. Let $p_{\max}^0 = \{K_p, K_p \cdot K_p\}$, $\bar{F}_{\max}^0 = \{K_f, K_f \cdot K_f\}$ and $k = 0$, where $K_p > 0$ and $K_f > 0$ are sufficiently large constant numbers.

Step 1. Computation of the search direction.

If $\Phi^k \neq 0$, then compute d^{k0} and $\bar{\lambda}^{k0}$ by solving the following linear system in (d, λ) :

$$V^k \begin{pmatrix} d \\ \lambda \end{pmatrix} = \begin{pmatrix} -\nabla f^k \\ 0 \end{pmatrix}. \tag{6}$$

If $\eta_j^k \neq 0$, then let $\lambda_j^{k0} = \eta_j^k \bar{\lambda}_j^{k0} / (-\eta_j^k + c_j^k)$, otherwise let $\lambda_j^{k0} = \bar{\lambda}_j^{k0}$. Compute d^{k1} and $\bar{\lambda}^{k1}$ by solving the following linear system in (d, λ) :

$$V^k \begin{pmatrix} d \\ \lambda \end{pmatrix} = \begin{pmatrix} -\nabla L^k \\ -\Phi_1^k \end{pmatrix}. \tag{7}$$

If $\eta_j^k \neq 0$, then let $\lambda_j^{k1} = \eta_j^k \bar{\lambda}_j^{k1} / (-\eta_j^k + c_j^k)$, otherwise let $\lambda_j^{k1} = \bar{\lambda}_j^{k1}$.

Step 2. Line search with filter.

1) If

$$\|\Phi(x^k + d^{k1}, \mu^k + \lambda^{k1})\|^2 \leq \theta_1 \|\Phi^k\|^2, \tag{8}$$

and (9) or (10), at least one holds, then let $x^{k+1} = x^k + d^{k1}$ and $\mu^{k+1} = \mu^k + \lambda^{k1}$, go to **Step 3**.

2) Line research:

Find x^{k+1} and μ^{k+1} to be acceptable for the filter test: let $\bar{x}^{k+1} = x^k + \alpha_k d^{k0}$, $\bar{\mu}^{k+1} = \mu^k + \alpha_k \lambda^{k0}$, $\hat{x}^{k+1} = x^k + \alpha_k d^{k1}$, and $\hat{\mu}^{k+1} = \mu^k + \alpha_k \lambda^{k1}$, where $\alpha_k = \tau^j$ and j is the smallest nonnegative integer satisfying either

$$\|\Phi_1(\hat{x}^{k+1}, \hat{\mu}^{k+1})\| \leq \theta_1 \max\{\|\Phi_1(x^k, \mu^k)\|, p_{\max}^j\} \tag{9}$$

or

$$f(\bar{x}^{k+1}) - \max\{f^k, \bar{F}_{\max}^j\} \leq -\alpha_k \theta_1 \|\Phi_1^{k+1}\| \tag{10}$$

for all $(f(x^j), \|\Phi_1^j\|^2) \in F^k$. If (9) holds, then $x^{k+1} = \hat{x}^{k+1}$, $\mu^{k+1} = \hat{\mu}^{k+1}$ and call it Φ -step; if (10) holds, then $x^{k+1} = \bar{x}^{k+1}$, $\mu^{k+1} = \bar{\mu}^{k+1}$ and call it f -step.

If there are no such x^{k+1} and μ^{k+1} or α_k too small, we use the backtracking technology or use the feasibility restoration phase to find x^{k+1} and μ^{k+1} so that it is an acceptable filter and the $QP(x^{k+1})$ subproblem is compatible. Go to **Step 1**.

Step 3. Update.

If x^{k+1} is a KKT point, then stop, otherwise if $\mu_i^{k+1} \leq \bar{\mu}$, then $\mu_i^{k+1} = \mu_i^{k+1}$; otherwise let $\mu_i^{k+1} = \bar{\mu}$, give H^{k+1} by BFGS update, $\bar{f}^{k+1} = 1/2(\bar{f}^k + f^{k+1})$, $F^{k+1} = F^k \cup (\bar{f}(x^{k+1}), \|\Phi_1^k\|)$ and delete all pairs $(\bar{f}(x^l), \|\Phi_1^l\|)$ which are dominated by $(F(x^{k+1}), \mu^{k+1})$ in F^{k+1} , obtain F^{k+1} . Let $k = k + 1$ and go to **Step 1**.

Remark 1 Let $\Phi(x, \mu) = ((\nabla_x L, H(x)), (\Phi_1(x, \mu))^T)^T$, the above proposed NFQPIM can also be used to solve the following constrained NLP:

$$\begin{aligned} & \min f(x) \\ & \text{s.t. } G(x) \leq 0, \quad H(x) = 0, \quad x \in R^n, \end{aligned}$$

where $f : R^n \rightarrow R$ and $G(x) = (g_1(x), g_2(x), \dots, g_m(x))^T : R^n \rightarrow R^m$ and $H(x) = (h_1(x), h_2(x), \dots, h_p(x))^T : R^n \rightarrow R^p$ are Lipschitz continuously differentiable functions.

2.3 Implementation

In this subsection, we give the implementation of the proposed NFQPIM. Firstly, we suppose that the following assumptions A1–A3 hold.

- A1. The level set $\{x|f(x) \leq f(x^0)\}$ is bounded, and for sufficiently large k , $\|\mu^k + \lambda^{k0} + \lambda^{k1}\| < \bar{\mu}$.
- A2. f and g_i are Lipschitz continuously differentiable, and for all $y, z \in R^{n+m}$,

$$\|\nabla L(y) - \nabla L(z)\| \leq m_0 \|y - z\|, \quad \|\Phi(y) - \Phi(z)\| \leq m_0 \|y - z\|,$$

where $m_0 > 0$ is the Lipschitz constant.

- A3. H^k is positive definite and there exist positive numbers m_1 and m_2 such that $m_1 \|d\|^2 \leq d^T H^k d \leq m_2 \|d\|^2$ for all $d \in R^n$ and all k .

Lemma 1 *If $\Phi^k \neq 0$, then V^k is nonsingular.*

Proof Assume $\Phi^k \neq 0$. If $V^k(u, v) = 0$ for some $(u, v) \in R^{n+m}$, where $u = (u_1, \dots, u_n)^T$, $v = (v_1, \dots, v_m)^T$, then

$$H^k u + \nabla G^k v = 0 \tag{11}$$

and

$$\text{diag}(\xi^k)(\nabla G^k)^T u + \text{diag}(\eta^k - c^k)v = 0. \tag{12}$$

From the definitions of ξ_j^k and η_j^k , we know that $\xi_j^k \geq 0$ and $\eta_j^k - c^k \neq 0$ for all j . So $\text{diag}(\eta^k - c^k)$ is nonsingular. We have

$$v = -(\text{diag}(\eta^k - c^k))^{-1} \text{diag}(\xi^k)(\nabla G^k)^T u. \tag{13}$$

Taking (13) into (11), we have

$$u^T (H^k u + \nabla G^k v) = u^T H^k u - u^T \nabla G^k \text{diag}(\xi^k)(\text{diag}(\eta^k - c^k))^{-1} (\nabla G^k)^T u = 0.$$

The fact that $-\nabla G^k \text{diag}(\xi^k)(\text{diag}(\eta^k - c^k))^{-1} (\nabla G^k)^T$ is positive semidefinite implies $u = 0$, and then $v = 0$ by (13). V^k is nonsingular. This lemma holds. □

Lemma 2 $d^{k0} = 0$ if and only if $\nabla f^k = 0$, and $d^{k0} = 0$ implies $\bar{\lambda}^{k0} = 0$ and $\lambda^{k0} = 0$.

If (x^*, μ^*) is an accumulation point of $\{(x^k, \mu^k)\}$, then $d^{*0} = 0$, and $(d^{*0}, \lambda^{*0})^T$ is the solution of the following equations:

$$V^* \begin{pmatrix} d \\ \lambda \end{pmatrix} = \begin{pmatrix} -\nabla f^* \\ 0 \end{pmatrix} \tag{14}$$

and $\nabla L(x^*, \mu^*) = 0$.

It is clear that the following lemma holds, with reference to [8].

Lemma 3 If $d^{k0} \neq 0$, then

$$(d^{k0})^T H^k d^{k0} \leq -(d^{k0})^T \nabla f^k.$$

Proof (14) implies

$$H^k d^{k0} + \nabla G^k \lambda^{k0} = -\nabla f^k, \tag{15}$$

and

$$\text{diag}(\xi^k)(\nabla G^k)^T d^{k0} + \text{diag}(\eta^k - c^k)\hat{\lambda}^{k0} = 0. \tag{16}$$

We have

$$\hat{\lambda}^{k0} = -(\text{diag}(\eta^k - c^k))^{-1} \text{diag}(\xi^k)(\nabla G^k)^T d^{k0}. \tag{17}$$

Taking (17) into (15), we have

$$\begin{aligned} &(d^{k0})^T (H^k d^{k0} + \nabla G^k \lambda^{k0}) \\ &= (d^{k0})^T H^k d^{k0} - (d^{k0})^T \nabla G^k \text{diag}(\xi^k)(\text{diag}(\eta^k - c^k))^{-1}(\nabla G^k)^T d^{k0} \\ &= -(d^{k0})^T \nabla f^k. \end{aligned} \tag{18}$$

$(d^{k0})^T \nabla G(x^k) \text{diag}(\xi^k)(\text{diag}(\eta^k - c^k))^{-1}(\nabla G^k)^T d^{k0} \leq 0$ implies

$$(d^{k0})^T H^k d^{k0} \leq -(d^{k0})^T \nabla f^k. \tag{19}$$

The lemma holds. □

Lemma 4 There exists $m_3 > 0$ such that, for any $0 < t \leq 1$,

$$\|\Phi_1(x^k + td^{k0}, \mu^k + t\lambda^{k0})\|^2 - \|\Phi_1\|^2 \leq m_3 t^2.$$

Proof If $\Phi_1^k = 0$, then there exists $m_4 > 0$ such that, for any $0 < t \leq 1$,

$$\|\Phi_1(x^k + td^{k0}, \mu^k + t\lambda^{k0})\|^2 = \|\Phi_1(x^k + td^{k0}, \mu^k + t\lambda^{k0}) - \Phi_1^k\|^2 \leq t^2 m_2^2 \|(d^{k0}, \lambda^{k0})\|^2.$$

The lemma holds for $\Phi_1^k = 0$. □

We define that if $(g_i^k, \mu_i^k) \neq (0, 0)$, then $(\bar{\xi}_i^{k0}, \bar{\eta}_i^{k0}) = (\xi_i^k, \eta_i^k)$, otherwise $\bar{\xi}_i^{k0}(\nabla g_i^k)^T d^{k0} + \bar{\eta}_i^{k0}\lambda_i^{k0} = \phi'_i((x^k, \mu^k), (d^{k0}, \lambda^{k0}))$, where $\phi'_i((x^k, \mu^k), (d^{k0}, \lambda^{k0}))$ is the direction derivative of $\phi_i(x, \mu)$ at (x^k, μ^k) in the direction (d^{k0}, λ^{k0}) .

Let $\text{diag}(\bar{\xi}^{k0})$ or $\text{diag}(\bar{\eta}^{k0})$ denote the diagonal matrix whose j th diagonal element is $\bar{\xi}_j^{k0}$ or $\bar{\eta}_j^{k0}$, respectively. Then $\phi_i(0, 0) = 0$ implies

$$(\Phi_1^k)^T (\text{diag}(\bar{\xi}^{k0})(\nabla G^k)^T, \text{diag}(\bar{\eta}^{k0})) = (\Phi_1^k)^T (\text{diag}(\xi^k)(\nabla G^k)^T, \text{diag}(\eta^k)).$$

Then

$$\begin{aligned} & \|\Phi_1^k + t(\text{diag}(\bar{\xi}^{k0})(\nabla G^k)^T d^{k0} + \text{diag}(\bar{\eta}^{k0})\lambda^{k0})\|^2 \\ &= \|\Phi_1^k\|^2 + t^2 \|\text{diag}(\bar{\xi}^{k0})(\nabla G^k)^T d^{k0} + \text{diag}(\bar{\eta}^{k0})\lambda^{k0}\|^2. \end{aligned} \tag{20}$$

It is clear that

$$\|\Phi_1(x^k + td^{k0}, \mu^k + t\lambda^{k0})\|^2 = \|\Phi_1^k\|^2 + O(t^2).$$

This lemma holds.

Lemma 5 *If $\Phi_1^k \neq 0$, then given any $\varepsilon > 0$ there is $\bar{t} > 0$ such that, for any $0 < t \leq \bar{t}$,*

$$\|\Phi_1^k\|^2 - \|\Phi_1(x^k + td^{k1}, \mu^k + t\lambda^{k1})\|^2 \geq (2 - \varepsilon)t\|\Phi_1^k\|^2.$$

Proof If $\Phi_1^k \neq 0$, (7) implies

$$\text{diag}(\xi^k)(\nabla G^k)^T d^{k1} + \text{diag}(\eta^k - c^k)\lambda^{k1} = -\Phi_1^k. \tag{21}$$

We define that if $(g_i^k, \mu_i^k) \neq (0, 0)$ then $(\bar{\xi}_i^{k1}, \bar{\eta}_i^{k1}) = (\xi_i^k, \eta_i^k)$, otherwise $(\bar{\xi}_i^{k1}\nabla g_i^k, \bar{\eta}_i^{k1})(d^{k1}, \lambda^{k1}) = \phi'_i((x^k, \mu^k), (d^{k1}, \lambda^{k1}))$, where $\phi'_i((x^k, \mu^k), (d^{k1}, \lambda^{k1}))$ is the direction derivative of $\phi_i(x, \mu)$ at (x^k, μ^k) in the direction (d^{k1}, λ^{k1}) . Let $\text{diag}(\bar{\xi}^{k1})$ or $\text{diag}(\bar{\eta}^{k1})$ denote the diagonal matrix whose i th diagonal element is $\bar{\xi}_i^{k1}$ or $\bar{\eta}_i^{k1}$, respectively. Clearly, for all i ,

$$\phi_i(x^k + td^{k1}, \mu^k + t\lambda^{k1}) - \phi_i^k \leq t(\bar{\xi}_i^{k1}(\nabla g_i^k)^T d^{k1} + (\bar{\eta}_i^{k1})). \tag{22}$$

Since $c_i^k \neq 0$, it follows by the definitions of c_i^k and η_i^k that $\eta_i^k = 0, g_i^k = 0, \mu_i^k \geq 0$, and $\phi_i^k = 0$. We have

$$\begin{aligned} & \|\Phi_1^k + t(\text{diag}(\bar{\xi}^{k1})(\nabla G^k)^T d^{k1} + \text{diag}(\bar{\eta}^{k1})\lambda^{k1})\|^2 \\ &= (1 - 2t)\|\Phi_1^k\|^2 + t\|\text{diag}(\bar{\xi}^{k1})(\nabla G^k)^T d^{k1} + \text{diag}(\bar{\eta}^{k1})\lambda^{k1}\|^2. \end{aligned} \tag{23}$$

It follows from (22) and (23) that, given any $\varepsilon > 0$, there is $\bar{t} > 0$ such that, for any $0 < t \leq \bar{t}$,

$$\|\Phi_1^k\|^2 - \|\Phi_1(x^k + t^2d^{k1}, \mu^k + t\lambda^{k1})\|^2 \geq (2 - \varepsilon)t\|\Phi_1^k\|^2.$$

Hence this lemma holds. □

Lemma 6 $d^{k0} = 0$ if and only if $\nabla f^k = 0$, and $d^{k0} = 0$ implies $\bar{\lambda}^{k0} = 0$ and $\lambda^{k0} = 0$.

Proof If $\nabla f^k = 0$, then $(x^k, \bar{\lambda}^{k0}) = \hat{V}^k(0, 0) = (0, 0)$. If $d^{k0} = 0$, then (14) implies

$$\text{diag}(\xi^k)(\nabla G^k)^T d^{k0} + \text{diag}(\eta^k - c^k)\bar{\lambda}^{k0} = \text{diag}(\eta^k - c^k)\bar{\lambda}^{k0} = 0. \tag{24}$$

Clearly, $\bar{\lambda}^{k0} = 0, \lambda^{k0} = 0$, and $(\nabla f^k, 0) = (V^k)^{-1}(0, 0) = (0, 0)$. □

From Lemmas 3–6, we know that, if $\Phi_1^k \neq 0$, then (d^k, λ^k) is the decreasing direction of $\|\Phi^k\|$; if $d^{k0} \neq 0$, then d^k is the decreasing direction of f^k . If $\Phi_1^k = 0$ and $d^{k0} = 0$, then (x^k, μ^k) is a KKT point. We consider four cases for linear searches.

Case 1. $k - 1$ iteration has a Φ -step and $\Phi_1^k = 0$. In this case, $p_{\max}^k = p_{\max}^{k-1}$ and $\min\{p_{\max}^j | j \in F^k > 0\}$. Clearly, we can find α_k such that \hat{x}^{k+1} satisfies (9).

Case 2. $k - 1$ iteration has a Φ -step and $\Phi_1^k \neq 0$. In this case, it follows from Lemma 5 that, given any $\varepsilon > 0$, there is $\bar{t} > 0$ such that, for any $0 < t \leq \bar{t}$,

$$\|\Phi_1^k\|^2 - \|\Phi_1(x^k + td^{k1}, \mu^k + t\lambda^{k1})\|^2 \geq (2 - \varepsilon)t\|\Phi_1^k\|^2.$$

$p_{\max}^k > 0$ is monotonically nonincreasing. So, we can find α_k such that \hat{x}^{k+1} satisfies (9).

Case 3. $k - 1$ iteration has an f -step and $d^{k0} \neq 0$. In this case, it follows from Lemma 5 that, if $d^{k0} \neq 0$, then

$$(d^{k0})^T H^k d^{k0} \leq -(d^{k0})^T \nabla f^k,$$

where f_{\max}^k is monotonically nonincreasing. We can find α_k such that \bar{x}^{k+1} satisfies (10).

Case 4. The $(k - 1)$ iteration has an f -step and $d^{k0} = 0$. In this case, if $\Phi_1^k = 0$, then (x^k, μ^k) is a KKT point, otherwise x^k may be an infeasible stationary point.

If there are no such x^{k+1} and μ^{k+1} or α_k too small, we use the backtracking technology or use the feasibility restoration phase to find x^{k+1} and μ^{k+1} so that it is acceptable that the filter and the $QP(x^{k+1})$ subproblem are compatible.

3 Convergence

In this section, we discuss the global and superlinear convergence rate of the proposed method. We give the following A4 and suppose that the assumptions A1–A4 hold in this section.

A4. For all k and some $\alpha_{\min} > 0, \alpha_k > \alpha_{\min} > 0$.

It implies from (3) and (4) that $p_{\max}^k > 0$ is monotonically nonincreasing and, if $\|\Phi_1(x^k)\| \rightarrow 0$, then $p_{\max}^k \rightarrow 0$.

Lemma 7 Consider the sequence $\{\|\Phi_1(x^k)\|^2\}$ and $\{f^k\}$ such that $\{\|\Phi_1(x^k)\|^2 \geq 0\}$ and $\{f^k\}$ is monotonically decreasing and bounded below. Let a constant θ satisfy, for all k and $l \in F^k$, that

$$\|\Phi_1(\hat{x}^{k+1}, \hat{\mu}^{k+1})\| \leq \theta_1 \max\{\|\Phi_1(x^k, \mu^k)\|, p_{\max}^j\} \tag{25}$$

or

$$f(\bar{x}^{k+1}) - \max\{f^k, \bar{F}_{\max}^j\} \leq -\alpha_k \theta_1 \|\Phi_1^{k+1}\|, \tag{26}$$

where $\alpha_k \geq \alpha_{\min} > 0$ is the step length, θ is a given positive number. Then $p_{\max}^k \rightarrow 0$.

Proof Suppose that the theorem is not true, then $\Phi_1(x^k) \not\rightarrow 0$, and there exist $\varepsilon > 0$ and infinitely many members of index set K such that $\|\Phi_1(x^{k+1}, \mu^{k+1})\| \geq \varepsilon > 0$, $p_{\max}^k \geq \varepsilon > 0$, and $\|\Phi_1(x^{k+1}, \mu^{k+1})\| \geq \theta \|\Phi_1(x^k, \mu^k)\|$ for any $k \in K$. We have

$$f(x^k) - f(x^{k+1}) \geq \alpha_k \theta \|\Phi_1(x^{k+1}, \mu^{k+1})\| > \alpha_{\min} \theta \varepsilon. \tag{27}$$

Because $\{f^k\}$ is monotonically decreasing, (27) implies $f(x^k) \rightarrow -\infty$ as $k \rightarrow +\infty$, which is contravention of $\{f^k\}$ being bounded below. This lemma holds. □

Lemma 8 Consider an infinite sequence iterations on which $\{f^k, \|\Phi_1(x^k)\|^2\}$ entered into the filter, where $\|\Phi_1(x^k)\| > 0$ and $\{f^k\}$ is bounded below. It follows that $\Phi_1(x^k) \rightarrow 0$.

Theorem 1 If (x^*, μ^*) is an accumulation point of $\{(x^k, \mu^k)\}$, then x^* is a KKT point of problem (1).

Proof It is obvious that Lemmas 8 and 2 imply that Theorem 1 holds. □

Next we consider the superlinear convergence of the method and firstly give the following assumptions we need.

- A5. The Mangasarian–Fromovitz (M-F) qualification condition is satisfied at x^* , i.e., $\{\nabla g_i(x^*)\}$ are linear independent for all $i \in I = \{i | g_i(x^*) = 0\}$, and there exists a direction such that $d^T \nabla g_i(x^*) < 0$, $i \in I = \{i | g_i(x^*) = 0\}$, where $i \in I = \{i | g_i(x^*) = 0\}$, where x^* is an accumulation point of $\{x^k\}$ and a KKT point of problem (1).
- A6. The sequence of $\{H^k\}$ satisfies

$$\frac{\|(H^k - \nabla_x^2 L(x^k, \mu^k))d^{k1}\|}{\|d^{k1}\|} \rightarrow 0.$$

- A7. The strict complementarity condition holds at each KKT point (x^*, μ^*) . It follows that ϕ^k is differentiable at each KKT point (x^*, μ^*) . Assumption A7 implies that Φ is continuously differentiable at each KKT point (x^*, μ^*) . As Lemma 1, we have that the following lemmas hold.

Lemma 9 Assume A1–A7 hold, then $\{\|(V^k)^{-1}\|\}$ and $\{\|(\hat{V}^k)^{-1}\|\}$ are bounded. Furthermore, if V^* is an accumulation matrix of $\{V^k\}$, then V^* is nonsingular.

Proof By Theorem 1, $\Phi^* = 0$ and $c^k \rightarrow 0$. Without loss of generality, we may assume that $(x^k, \mu^k) \rightarrow (x^*, \mu^*)$, $H^k \rightarrow H^*$, $\text{diag}(\xi^k) \rightarrow \text{diag}(\xi^*)$, and $\text{diag}(\eta^k) \rightarrow \text{diag}(\eta^*)$. By the definitions of ξ_i^k and η_i^k , we know that $(\xi_i^*)^2 + (\eta_i^*)^2 \neq 0$. $H^k \rightarrow H^*$ implies that H^* is positive definite.

If $V^*(u, v) = 0$, where $(u, v) \in R^{n+m}$ and $u = \{(u_1, \dots, u_n)^T\}$, $v = \{(v_1, \dots, v_m)^T\}$, then we have

$$H^*u + \nabla G^*v = 0 \tag{28}$$

and

$$\text{diag}(\xi^*)(\nabla G^*)^T u + \text{diag}(\eta^*)v = 0. \tag{29}$$

From (29) and the definitions of ξ_j^* and η_j^* , we know that if $\xi_j^* = 0$ then $\eta_j^* \neq 0$. If $\xi_j^* \neq 0$ then

$$u^T \nabla g_j^* = -\frac{\eta_j^*}{\xi_j^*} v_j. \tag{30}$$

Putting (29) and (30) into (28), we have

$$\begin{aligned} &u^T (H^*u + \nabla G^*v) \\ &= u^T H^*u + \sum_{j:\xi_j^* \neq 0} -\frac{\eta_j^*}{\xi_j^*} v_j^2 = 0. \end{aligned} \tag{31}$$

$\eta_j^*/\xi_j^* \leq 0$ implies $u = 0$, and if $\eta_j^* \neq 0$ then $v_j = 0$. Let $I = \{j | g_j^* = 0\}$, because $g_j^* \neq 0$ implies $\eta_j^* \neq 0$ and $v_j = 0$, we have

$$\sum_{j \in I} \nabla g_j^* v_j = 0, \tag{32}$$

and $v_j = 0$ ($j \in I$) by A4, i.e., $(u, v) = 0$ and V^* is nonsingular.

On the other hand, suppose to the contrary that there exists a subsequence $\{(x^{k(i)}, \lambda^{k(i)})\}$ such that $\|(V^{k(i)})^{-1}\| \rightarrow \infty$ as $k(i) \rightarrow \infty$ and $(x^{k(i)}, \lambda^{k(i)}) \rightarrow (x^*, \lambda^*)$. We can choose $k(i)$ properly such that $V^{k(i)} \rightarrow V^*$ including $\xi^{k(i)} \rightarrow \xi^*$ and $\eta^{k(i)} \rightarrow \eta^*$. Clearly, $(\xi_j^*)^2 + (\eta_j^*)^2 \geq 3 - 2\sqrt{2} > 0$ and $V^* \in \partial \Phi^*$. But V^* is nonsingular by the above proof, which contradicts the assumption $\|(V^{k(i)})^{-1}\| \rightarrow \infty$. Hence, $\{\|(V^{k(i)})^{-1}\|\}$ is bounded. $\Phi^k \rightarrow 0$ implies $\lim_{k \rightarrow \infty} V^k = \lim_{k \rightarrow \infty} \hat{V}^k$, we can also obtain that $\{\|(\hat{V}^k)^{-1}\|\}$ is bounded. This lemma holds. □

Assumption A5 shows that (x^k, μ^k) is a Newton direction with a high order perturbation. We obtain the following lemma.

Lemma 10 For sufficiently large k , $x^{k+1} = x^k + d^{k1}$ and $\mu^{k+1} = \mu^k + \lambda^{k1}$.

Furthermore, Lemma 10 implies that the following theorem holds.

Theorem 2 Assume A1–A7 hold. Let Algorithm 1 (NFQPIM) be implemented to generate a sequence $\{(x^k, \mu^k)\}$, (x^*, μ^*) be an accumulation point of $\{(x^k, \lambda^k)\}$. Then (x^*, μ^*) is a KKT point of problem (1), and (x^k, μ^k) converges to (x^*, μ^*) superlinearly.

4 Numerical tests

We use Algorithm 1 (NFQPIM) for the constrained optimization problems (see [19]): H^k is updated by the BFGS method. The termination criterion is $\|\phi\| \leq 10^{-5}$. The parameters are chosen as follows: $c = 0.1, \nu = 2, \tau = 0.7, \theta_1 = 0.8, \theta = 0.6, \bar{\mu} = 10,000$. In the “NIT/NG” entry of the table below, NIT is the number of iterations, NF represents the number of function evaluations, NG denotes the number of gradient evaluations. The numerical results can be seen in the Table 1. We test the proposed NFQPIM for solving almost 100 optimization problems. And the numerical results illustrate that the proposed method is efficient and promising.

Table 1 Numerical results on the NFQPIM for some constrained optimization problems

Problem	n	m	NIT	NG	NF
hs001	2	1	65	43	40
hs002	2	1	15	19	18
hs003	2	1	3	5	4
hs004	2	2	4	6	5
hs005	2	4	9	11	9
hs006	2	1	5	9	8
hs007	2	1	16	25	21
hs008	2	2	3	6	5
hs009	2	1	10	13	12
hs010	2	1	27	43	33
hs011	2	1	13	23	15
hs012	2	1	13	19	16
hs013	2	1	3	6	4
hs014	2	2	4	7	4
hs015	2	2	7	11	7
hs016	2	5	5	8	7
hs017	2	5	14	19	16
hs018	2	6	20	24	23
hs019	2	6	4	7	6
hs020	2	5	5	8	6
hs021	2	5	4	7	5
hs022	2	2	5	8	5
hs023	2	9	3	4	4
hs024	2	5	7	13	10
hs025	3	6	4	6	6
hs026	3	1	24	33	27
hs027	3	1	28	34	31
hs028	3	1	6	12	9
hs029	3	1	21	37	31
hs030	3	7	9	11	10
hs031	3	7	12	17	14
hs032	3	5	8	14	12
hs033	3	6	11	16	24
hs034	3	8	8	12	10
hs035	3	4	7	11	10
hs036	3	7	10	13	11
hs037	3	8	13	18	15
hs038	4	8	83	111	98
hs039	4	2	21	35	32
hs040	4	3	11	16	14
hs041	4	9	13	17	15
hs042	4	2	11	14	12
hs043	4	3	15	23	20
hs044	4	10	16	23	21
hs045	5	10	5	7	7
hs046	5	10	29	37	33
hs047	5	3	26	33	30

Table 1 (Continued)

Problem	n	m	NIT	NG	NF
hs048	5	2	10	15	13
hs049	5	2	36	46	40
hs050	5	3	35	43	39
hs051	5	3	9	13	11
hs052	5	3	7	14	12
hs053	5	3	5	7	7
hs054	6	13	9	13	10
hs055	6	14	8	14	11
hs056	7	4	12	14	12
hs057	2	3	6	9	8
hs059	2	7	28	33	30
hs060	3	7	13	23	21
hs061	3	2	59	68	58
hs062	3	7	24	33	28
hs063	3	5	15	23	20
hs064	3	4	36	43	37
hs065	3	7	21	29	27
hs066	3	8	13	23	19
hs067	3	41	65	53	47
hs099	7	16	65	43	40
hs100	7	4	65	43	40
hs101	7	20	65	43	40
hs102	7	20	65	43	40
hs103	7	20	65	43	40
hs104	8	22	65	43	40
hs105	8	17	65	43	40
hs106	8	22	65	43	40
hs107	9	14	65	43	40
hs108	9	14	33	30	30
hs110	10	20	25	43	40
hs111	10	23	61	73	68
hs112	10	13	51	73	67
hs113	10	8	55	65	61
hs114	10	31	56	73	68
hs116	13	41	123	143	140
hs117	15	20	511	63	60
hs118	15	59	67	81	77
hs119	16	40	68	83	78

5 Conclusions

In this paper, we developed a nonmonotone filter QP-free infeasible method for minimizing a smooth optimization problem with inequality constraints. This proposed method is based on the solution of nonsmooth equations which are obtained by the multiplier and some NCP functions for the KKT first-order optimality conditions. At each iteration of the proposed method, it was a perturbation of a Newton or quasi-Newton iteration on both the primal and dual variables for the solution of the KKT optimality conditions. Moreover, we used the filter on linear searches with a nonmonotone acceptance mechanism. We also showed that the proposed method had a global convergence and a superlinear convergence rate. Finally, the numerical results illustrated that the proposed method was efficient. However, how to apply this method to the real optimal problem will be studied in the near future.

Acknowledgements

We would like to thank the anonymous referees and the associate editor for their useful comments and suggestions which improved this paper greatly.

Funding

The work of Y. Shang is supported by the National Natural Science Foundation of China Grant No.11471102. The work of D. Pu is supported by the National Natural Science Foundation of China Grant No. 11371281. The work of Z.-F. Jin is supported by the National Natural Science Foundation of China Grant No. 61772174, Plan for Scientific Innovation Talent of Henan Province Grant No. 174200510011.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors contributed equally to this work. All authors read and approved the final manuscript.

Author details

¹School of Mathematics and Statistics, Henan University of Science and Technology, Luoyang, China. ²Department of Mathematics, Tongji University, Shanghai, China.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 18 April 2018 Accepted: 12 September 2018 Published online: 11 October 2018

References

1. Chen, X.: Smoothing methods for complementarity problems and their applications: a survey. *J. Oper. Res. Soc. Jpn.* **43**, 32–47 (2000)
2. Arrow, K.J., Debreu, G.: Existence of an equilibrium for a competitive economy. *Econometrica*, **22**, 265–290 (1954)
3. Smeers, Y.: Computable equilibrium models and the restructuring of the European electricity and gas markets. *Energy J.* **0**(4), 1–31 (1997)
4. Tian, B., Yang, X.: Smoothing power penalty method for nonlinear complementarity problems. *Pac. J. Optim.* **12**(2), 461–484 (2016)
5. Xie, S.L., Xu, H.R., Zeng, J.P.: Two-step modulus-based matrix splitting iteration method for a class of nonlinear complementarity problems. *Linear Algebra Appl.* **494**, 1–10 (2016)
6. Otero, R.G., Iusem, A.: A proximal method with logarithmic barrier for nonlinear complementarity problems. *J. Glob. Optim.* **64**(4), 663–678 (2016)
7. Hao, Z., Wan, Z., Chi, X., Chen, J.: A power penalty method for second-order cone nonlinear complementarity problems. *J. Comput. Appl. Math.* **290**, 136–149 (2015)
8. Pu, D.G., Li, K.D., Xue, J.: Convergence of QP-free infeasible methods for nonlinear inequality constrained optimization problems. *Tongji Daxue Xuebao Ziran Kexue Ban* **33**, 525–529 (2005)
9. Fletcher, R., Leyffer, S., Toint, P.: On the global convergence of a filter-SQP algorithm. *SIAM J. Optim.* **13**, 44–59 (2002)
10. Ulbrich, M., Ulbrich, S.: Non-monotone trust region methods for nonlinear equality constrained optimization without a penalty function. *Math. Program.* **95**, 103–135 (2003)
11. Yang, Y., Shang, Y.: A new filled function method for unconstrained global optimization. *Appl. Math. Comput.* **173**, 501–512 (2006)
12. Qi, H.D., Qi, L.: A new QP-free, globally convergent, locally superlinearly convergent algorithm for inequality constrained optimization. *SIAM J. Optim.* **11**, 113–132 (2000)
13. Liu, F.T., Fan, Y.H., Yin, J.H.: The use of QP-free algorithm in the limit analysis of slope stability. *J. Comput. Appl. Math.* **235**, 3889–3897 (2011)
14. Pu, D.G., Zhou, Y., Zhang, H.Y.: A QP free feasible method. *J. Comput. Math.* **22**(5) 651–660 (2004)
15. Jian, J.B., Guo, C.H., Tang, C.M., Bai, Y.Q.: A new superlinearly convergent algorithm of combining QP subproblem with system of linear equations for nonlinear optimization. *J. Comput. Appl. Math.* **273**, 88–102 (2015)
16. Zhang, H., Li, G., Zhao, H.: A kind of QP-free feasible method. *J. Comput. Appl. Math.* **224**, 230–241 (2009)
17. Huang, M., Pu, D.: Line search SQP method with a flexible step acceptance procedure. *Appl. Numer. Math.* **92**, 98–110 (2015)
18. Huang, M., Pu, D.: A line search SQP method without a penalty or a filter. *Comput. Appl. Math.* **34**, 741–753 (2015)
19. Dolan, E.D., Moré, J.J.: Benchmarking optimization software with performance profiles. *Math. Program.* **91**, 201–213 (2002)