Mathematica Aeterna, Vol. 5, 2015, no. 5, 961 - 969

Global Convergence Of A Qp-free Method Without A Penalty Function And A Filter

Ke Su

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

Bei Gao

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

Abstract

In this paper, a QP-free method without a penalty function and a filter is proposed for nonliner programming. The decrease condition of the constraint violation and the object function is require to satisfied in each step. The methods decrease condition compare with the acceptance criterion of filter may achieve more flexibility of accepting trial steps. Based on the solution of nonsmooth equations and decrease condition which the constraint violation and the objective function must satisfied the globally convergence is achieved.

Mathematics Subject Classification: xxxxx

Keywords: line search, NCP function, global convergence, QP-free

1 Introduction

We consider the following nonlinear inequality constrained optimization problem

$$\begin{aligned} &\min \ f(x) \\ &s.t. \ c(x) \leq 0 \end{aligned} \tag{1}$$

where $f(x) : R^n \to R$ and $c(x) = (c_1(x), c_2(x), \cdots, c_m(x))^T : R^n \to R^m$ are second-order continuously differentiable functions. For convenience, we denote $g(x) = \nabla f(x)$ and $A(x) = (\nabla c_1(x), \nabla_2(x), \cdots, \nabla c_m(x))$ and f_k refers to $f(x_k), g_k$ to $g(x_k), A_k$ to $A(x_k)$, etc.

The Lagrangian function of (1) is $L(x, \lambda) = f(x) + \lambda^T c(x)$, where $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_m)$ in R^m is the multiplier vector. The KKT condition is

$$\bigtriangledown_x L(x^*, \lambda^*) = 0, \quad c(x^*) \leq 0, \quad \lambda^* \geq 0, \quad \lambda^* c(x^*) = 0, \tag{2}$$

The KKT point is a point satisfied KKT condition.

Sequential quadratic programming (SQP) methods are widely used for solving inequality constrained optimization problem since late 1970s. But the high-computational costs of this methods confused most of the researchers. Then the QP-free algorithms are proposed by Panier, Tits and Herskovits[see 1]. A system of linear equation is used in compute direction in there paper, which overcome the difficulty of SQP methods. Recently a new QP-free methods for solution smooth inequality constraints was proposed by Pu, Li and Xue [see2]. These methods is based on the nonsmooth equations's solution which are obtained by the Fischer-Burmeister NCP function and the multipliers. The convergence rate under some mild condition is encouraging. Qi and Qi (see3) use the NCP function proposed a QP-free algorithm for solving the (NLP) problem. They used Two linear systems and a least-square problem to solved the search direction at each iteration in QP-free algorithm. The global convergence is provided.

In this paper, we propose a QP-free method without a penalty function and a filter for nonliner programming. The Fischer-Burmeister NCP function for the KKT first-order optimality condition and the multipliers was use to solving the nonsmooth equations in each iterative. And the flexible of the acceptance is shown by constraint violation function.

The new algorithm has the following advantages:

(1)we generalize the of constraint violation function.

(2) the scale of computation is greatly reduced ;

(3) The Maratos effect may be avoid by using self-adaption operator as constraint violation.

This paper is divided into 4 sections. The next section introduce the concept of a NCP function. In section 3 an algorithm of line search filter is given. The convergence is given in section 4.

2 Preliminaries

Definition 2.1 *(NCP pair and NCP functions) We call a pair $(a, b) \in R^2$ to be an NCP pair if $a \geq 0, b \geq 0$ and $ab = 0$.*
A function $\psi : R^2 \to R$ is called an NCP function if $(a, b) = 0$ if and only if (a, b) is an NCP pair.

The 4-1 piecewise linear NCP function φ as follows

$$\psi(a,b) = \begin{cases} k^2 a, & if b \geq k \mid a \mid; \\ 2kb - b^2 a, & if a \geq \mid b \mid /k; \\ 2k^2 a + 2kb + b^2/a, & if a < - \mid b \mid /k; \\ k^2 a + 4kb, & if b \leq -k \mid a \mid < 0. \end{cases} \tag{3}$$

where parameter $k > 0$ ψ is continuously differentiable except at the origin, but it is strongly semismooth at the origin. if $a \neq 0$ or $b \neq 0$, then ψ is continuously differentiable at $(a,b) \in R^2$,

$$\nabla\psi(a,b) = \begin{cases} \begin{pmatrix} k^2 \\ 0 \end{pmatrix} & if b \geq k \mid a \mid; \\ \begin{pmatrix} b^2/a^2 \\ 2k - 2b/a \end{pmatrix} & if a \geq \mid b \mid /k; \\ \begin{pmatrix} 2k^2 - b^2/a^2 \\ 2k + 2b/a \end{pmatrix} & if a < - \mid b \mid /k; \\ \begin{pmatrix} k^2 \\ 4k \end{pmatrix} & if b \leq -k \mid a \mid < 0. \end{cases} \tag{4}$$

and

$$A_\psi = \partial\psi(0,0) = \left\{ \begin{pmatrix} k^2 t^2 \\ 2k(1-t) \end{pmatrix} \bigcup \begin{pmatrix} 2k^2(1-t^2) \\ 2k(1-t) \end{pmatrix} \middle| |t| \leq 1 \right\}. \tag{5}$$

Let

$$\phi_i(x,\lambda) = \psi(-g_i(x), \lambda_i), 1 \leq i \leq m. \ \Phi(x,\lambda) = (\phi_1(x,\lambda), \cdots, \phi_m(x,\lambda))$$

If $(g_i(x), \lambda_i) \neq (0,0)$, then ϕ_i is continuously differentiable at $(x,\lambda) \in R^{n+m}$. We have

$$\nabla\phi_i(x,\lambda) = \begin{cases} \begin{pmatrix} -k^2 \nabla g_i(x) \\ 0 \end{pmatrix} & if \lambda_i \geq k \mid g_i(x) \mid; \\ \begin{pmatrix} -\lambda_i^2 \nabla g_i(x)/g_i(x)^2 \\ (2k - 2\lambda_i/g_i)e_i \end{pmatrix} & if - g_i(x) \geq \mid \lambda_i \mid /k; \\ \begin{pmatrix} (-2k + \lambda_i^2/g_i(x)^2)\nabla g_i(x) \\ (2k - 2\lambda_i/g_i(x))e_i \end{pmatrix} & if - g_i(x) < - \mid \lambda_i \mid /k; \\ \begin{pmatrix} -k^2 \nabla g_i(x) \\ 4ke_i \end{pmatrix} & if \lambda_i \leq -k \mid g_i(x) \mid < 0. \end{cases} \tag{6}$$

If $g_i(x) = 0$ and $\lambda_i = 0$,$1 \leq i \leq m$, then we have

$$\partial\phi_i(x,\lambda) = \left\{ \begin{pmatrix} -k^2 t^2 \nabla g_i(x) \\ 2k(1-t)e_i \end{pmatrix} \bigcup \begin{pmatrix} -2k^2(1 - t^2 \nabla g_i(x)) \\ (2k-2t)e_i \end{pmatrix} \middle| |t| \leq 1 \right\}$$

where $e_i = (0, \cdots, 0, 1, 0, \cdots, 0)^T \in R^m$ is the ith column of the unite matrix. We take $k = 1$ in this paper.
If $(-g_j(x^k), \lambda_k) = (0, 0)$, let $\xi_j^k = -2, \eta_j^k = 2$, otherwise, let $(-\xi_j^k, \eta_j^k) = \nabla\psi(a, b)|_{a=-g_i(x)}, b = \lambda_j^k$, We obtain

$$(\xi_j^k \nabla g_j(x^k), \eta_j^k e_j) = \nabla\phi_j(x^k, \lambda^k).$$

Clearly $\xi_j^k \leq 0$ and $\eta_j^k \geq 0$. Let

$$V_k = \begin{pmatrix} H_k & \nabla G_k \\ diag(\xi^k)(\nabla G^k)^T & diag(\eta^k + c^k) \end{pmatrix}$$

where H_k is a positive matrix which may be modified by BFGS update. $diag(\xi^k)$ or $diag(\eta^k + c^k)$ denotes the diagonal matrix whose j diagonal element is ξ_j^k or $\eta^k + c^k$, respectively, and $c_j^k = c\min\{1, \|\Phi^k\|^v\}$, where $\Phi^k = \Phi^k(x^k, \lambda^k), c > 0$ and $v > 1$are given paramenters.

3 Description of the algorithm

We aim the problem as a bi-objective optimization problem with two goals,i.e. minimizing the objective function $f(x)$ and the constraint violation $h(c(x)) = \Sigma\max\{0, c_i(x)\}$.
First of all, we consider the decrease condition for the constraint violation function. We denote $P_k = \|A_k\lambda_k + G_k\|$. A slack variable M_k is needed in order to resuce $h(c_k)$ and P_k evenly.

$$n_0 > 0,\ n_k = \frac{n_0}{k+1}(j \geq 1),\ n_j \to 0(j \to 0),\ and\ \frac{1}{2} \leq \frac{n_{k+1}}{n_k} < 1.$$

For constraint violation function, we adopt a parameter M_k to relax the criterion of iterates and enhance the flexible. let

$$Q_{l,k} = \max_{k-l+1 \leq i \leq k-1} h(c_i)$$

If the constraint violation $h(c_k) < \min\{\eta_1 n_k, \eta_2 P_k\}, 0 < \eta_1, \eta_2 < \frac{1}{2}$, we set

$$M_k = \min\{n_{jk}, P_k\} \tag{7}$$

if slack variable $M_k \geq Q_{l,k}$ then we said that a trial step is accepted as a new iterate, as long as the follow formula is satisfied.

$$M_k - h(c_{k+1}) \geq \alpha\eta M_k, \quad \eta \in (0, \frac{1}{2})$$

otherwise if $M_k \leq Q_{l,k}$then we said that a trial step is accepted as a new iterate, as long as the follow formula is satisfied.

$$Q_{l,k} - h(c_{k+1}) \geq \alpha\eta Q_{l,k}, \quad \eta \in (0, \frac{1}{2})$$

If the constraint violation function $h(c_k) \geq \min\{\eta_1 n_k, \eta_2 P_k\}, 0 < \eta_1, \eta_2 < \frac{1}{2}$, set

$$M_k = h(c_k) \tag{8}$$

then similar to the description of the above, if slack variable $M_k \geq Q_{l,k}$ then we said that a trial step is accepted as a new iterate, as long as the follow formula is satisfied.

$$M_k - h(c_{k+1}) \geq \alpha\eta M_k, \quad \eta \in (0, \frac{1}{2})$$

otherwise if $M_k \leq Q_{l,k}$then we said that a trial step is accepted as a new iterate, as long as the follow formula is satisfied.

$$Q_{l,k} - h(c_{k+1}) \geq \alpha\eta Q_{l,k}, \quad \eta \in (0, \frac{1}{2})$$

For convenience, we denote $T_k = \max(M_k, Q_{l,k})$, then the constraint violation can be expression as

$$T_k - h(c_{k+1}) \geq \alpha\eta Q_{l,k} \tag{9}$$

Now we consider the decrease condition for the objective function. we define

$$\triangle f_k = f(x_k - f(x_{k+1}))$$

If

$$g_k^T d_k \leq -\beta d_k^T B_k d_k \ \ and\ h(c_k) \leq \zeta_1 \|d_k\|^{\zeta_2} \tag{10}$$

where $\beta \in (0, \frac{1}{2}), \zeta_1 > 0, \zeta_2 \in (2, 3)$, then the sufficient decrease condition of objective function is

$$\triangle f_k \geq -g_k^T d_k \sigma\alpha$$

QP-free algorithm

The improved algorithm is presented as following.

Step0. Give a starting point $x^0 \in R^n$ and a initial positive definite matrix $H_0, \sigma \in (0, \frac{1}{2}), n_0 > 0, \eta \in (0, \frac{1}{2}), \theta > 0, \beta \in (0, \frac{1}{2}), \eta_1\ , \eta_2 \in (0, \frac{1}{2}), \zeta_1 > 0, \zeta_2 \in (2, 3)$, $t \in (0, 1), c > 0, k = 0$.

Step1. compute d_{k_0} and λ_{k_0} by solving the following linear system in (d, λ) :

$$V_k \begin{pmatrix} d \\ \lambda \end{pmatrix} = \begin{pmatrix} -\nabla f_k \\ 0 \end{pmatrix}. \tag{11}$$

where $\nabla f_k = \nabla f(x_k)$.
If $d_{k_0} = 0$, then stop
otherwise, if $\eta_k^j \neq= 0$, then let $\lambda_j^{k_0} = \eta_k^j \overline{\lambda}_j^{k_0} / (-\eta_j^{k_0} + c_j^k)$, otherwise let $\lambda_j^{k_0} = \overline{\lambda}_j^{k_0}$
compute $(d_{k^1}, \overline{\lambda}_{k_1})$ by solving the following linear system in (d, λ):

$$V_k \begin{pmatrix} d \\ \lambda \end{pmatrix} = \begin{pmatrix} -\nabla L_k \\ -\Phi_k \end{pmatrix}. \tag{12}$$

where $\nabla L_k = \nabla L(x_k, \lambda_k)$ and $\Phi_k = \Phi(x_k, \lambda_k)$. If $\eta_j^k \neq 0$, then let $\lambda_j^{k_1} = \eta_k^j \overline{\lambda}_j^{k_1} / (-\eta_j^{k_1} + c_j^k)$, otherwise let $\lambda_j^{k_1} = \overline{\lambda}_j^{k_1}$

If $\Phi_k = 0$ then let $b_k = 1$ and $\rho_k = 0$, otherwise if $d_{k_0} = 0$ then let $b_k = 0$ and $\rho_k = 1$, otherwise denote $b_k = (1 - \rho_k)$ and

$$\rho_k = \begin{cases} 1 & if (d_{k_1})^T \nabla f_k \leq \theta (d_{k_0})^T \nabla f_k \\ (1-\theta) \frac{(d_{k_0})^T \nabla f_k}{(d_{k_0} - d_{k_1})^T \nabla f_k} & otherwise \end{cases}$$

and let

$$\begin{pmatrix} d_k \\ \lambda_k \end{pmatrix} = b_k \begin{pmatrix} d_{k_0} \\ \lambda_{k_0} \end{pmatrix} + \rho_k \begin{pmatrix} d_{k_1} \\ \lambda_{k_1} \end{pmatrix}, \tag{13}$$

Step2 set $\alpha = 1$ evaluate function at x_k. compute $h_{x_k}, f_{x_k}, g_{x_k}, A_{x_k}$ If the KKT condition of problem (1) are satisfied stop, otherwise update M_k by (7) or (8).

Step3. If (3.3) holds go to 3.1 else go to 3.2

3.1 If $T - h(c_{k+1}) < \alpha \eta Q_{l,k}$ or $\triangle f_k < -g_k^T d_k \sigma \alpha$, then set $\alpha = t\alpha$,, go to step 5, otherwise set $\alpha_k = \alpha, x_{k+1} = x_k + \alpha_k d_k$, go to step 6.

3.2 If $T - h(c_{k+1}) \geq \alpha \eta Q_{l,k}$ then set $\alpha = \alpha$,$x_{k+1} = x_k + \alpha_k d_k$, go to step 6.

Step4. give H_k by BFGS update, let $k = k + 1$ go to step 1.

4 The Convergence Properties

To present a proof of global convergence of algorithm, In this section, we always assume that the following conditions hold.

A 1 The level set $\{x|f(x) \leq f(x_0)\}$ is bounded, and for sufficiently large k, $\|\mu_k + \lambda_{k_0} + \lambda_{k_1}\| < \overline{\mu}$

A 2 *f and g_i* are twice Lipschitz continuously differentiable, and for all $y, z \in R^{n+m}$,

$$\|\nabla L(y) - \nabla L(z)\| \leq m_3\|y - z\|, \quad \|\Phi(y) - \Phi(z)\| \leq m_3\|y - z\|,$$

where $m_3 > 0$ is the Lipschitz constant.

A 3 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 and the lagrange multiplier λ_k is bounded for all k.

We suppose that the assumptions A1-A3 hold.

Lemma 4.1 *If $\Phi^k \neq 0$ then V^k and V^* are nonsingular.*

Proof: If

$$V_k \begin{pmatrix} u \\ \vartheta \end{pmatrix} = 0,$$

for some $(u, \vartheta) \in R^n$, where $\vartheta = (\vartheta_1, \cdots, \vartheta)^T$, $u = (u_1, \cdots, u_n)^T$, then we have

$$H^k u + \nabla c^k v = 0 \tag{14}$$

and

$$diag(\xi^k)(\nabla c^k)^T u + diag(\eta^k) v = 0$$

From the definition of ξ_j^k and η_j^k, we know that $\xi_j^k \geq 0$ and $\eta_j^k \neq 0$ for all j. So, diag η^k is nonsingular. We have

$$v = -(diag(\eta^k))^{-1} diag(\xi^k)(\nabla c^k)^T u \tag{15}$$

Putting (4.3) into (4.1), we have

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

The fact that $-\nabla c^k diag(\xi^k)(diag(\eta^k))^{-1}(\nabla c^k)^T$ is positive semidefinite implies $u = 0$, and then $v = 0$ by(4.3). V^k is nonsingular. This lemma holds.

Lemma 4.2 *If $d^{k_0} = 0$, then $\nabla f(x^k) = 0$. and x^k is KKT point of problem(NLP). The lemma 2 hold (see [6] Lemma 2)*

Lemma 4.3 *If the Algorithm cannot terminate finitely, then* $\lim_{k \to +\infty} h(c_k) = 0$

Proof: We assue that $h(c_k) \to 0$. there exists a suffciently large inter $j > 0$, such that for $k \geq j$,

$$T_k = \max(h(c_k), Q_{l,k})$$

It implies with $Q_{l,k} = \max_{k-l+1 \leq i \leq k-1} h(c_i)$ that

$$T_k = \max(h(c_k), Q_{l,k}) = \max_{k-l+1 \leq i \leq k-1} h(c_i) \tag{16}$$

According to the update of M_k, (9)and(16), we have

$$\begin{cases} h(c_{k+1}) < (1 - \alpha_k \eta) \max_{k-l+1 \leq i \leq k-1} h(c_i) \\ h(c_{k+2}) < (1 - \alpha_{k+1} \eta) \max_{k-l+2 \leq i \leq k+1} h(c_i) \\ \cdots \\ h(c_{k+l}) < (1 - \alpha_{k+l-1} \eta) \max_{k \leq i \leq k+l-1} h(c_i) \end{cases} \tag{17}$$

and for $k \geq j$

$$\max_{k-l+1 \leq i \leq k-1} h(c_i) \geq \max_{k-l+2 \leq i \leq k+1} h(c_i) \geq \cdots \geq \max_{k \leq i \leq k+l-1} h(c_i) \tag{18}$$

Since$h(c_k) \to 0$, it follows that there exists a positive constant $\epsilon_1 > 0$ and an infinite subsequence $h(c_{k_i})$, such that $h(c_{k_i}) \geq \epsilon_1$ for all $k_i > j$. Then, for any $k_i > j$ there exists a positive integer i_0, such that $k_{i_0} > k$, So it follows with (18) that

$$\max_{k \leq i \leq k+l-1} h(c_i) \geq \max_{k_{i_0} \leq i \leq k_{i_0}+l-1} h(c_i) \geq \epsilon_1 \tag{19}$$

According to (16), we have that there exists a step size $\alpha_{\min} > 0$, such that $\alpha_k > \alpha_{\min}$ for all $k > j$. thus,

$$\max_{k \leq i \leq k+l-1} h(c_i) \leq (1 - \alpha_{\min}\eta)^{(\frac{k-k_1}{l})} \max_{k_l-l+1 \leq i \leq k_l} h(c_i) \tag{20}$$

where α denotes the maximal integer less than α.

So we have that

$$\lim_{k \to \infty} \max_{k \leq i \leq k+l-1} h(c_i) = 0$$

It implies that

$$\lim_{i \to \infty} h(c_i) = 0$$

which contradicts the assumption that $h(c_{k_i}) \geq \epsilon_1$ for all $k_i > j$. The conclusion follows.

Lemma 4.4 *Suppose assumption A1-A3 hold then*

$$\|d_k\| = o(P_k)$$

Proof: It follows from KKT condition of QP subproblem that

$$d_k = -B_k^{-1}(g_k + A_k\lambda_k)$$

Since $B_k$is uniformly positive sefinite and uniformly bounded, we obtain that B_k^{-1} is also positive definite and bounded for all k. Therefore $\|d_k\| = o(\|g_k + A_k\lambda_k\|)o(P_k)$.

Theorem 4.5 *Suppose assumptions A1-A3 hold. Then one of the following two situations occurs; (i)Algorithm terminates at a KKT point of problem(1). (ii)There exists at least one accumulation point, which is a KKT point.*

ACKNOWLEDGEMENTS. We thank the Editor and the referee for their comments. This work is supported by the National Science Foundation of China (No.11101115),the Natural Science Foundation of Hebei Provience (No.A2014201033)and the Key Research Foundation of Educational Bureau of Hebei province (No.ZD2015069)

References

[1] Panier E.R.,Tits A.L. and Herskovits J.N.A QP-free, A QP-free, globally, locally super convergent method for the inequality constrained optimization problems, SIAM Journal on Control and Optimization, 1988, 788-811.

[2] Pu D.G., Li K.D. and Xue W.J, Convergence of QP-free infeasible methods for nonlinear inequality constrained optimization problems, J. of Tongji University,2001,525-529.

[3] Qi H. and Qi L, A New QP-free , globally V, locally superlinear convergent feasible method for the inequality constrained optimization problems, SIAM Journal on Optimization,2000, 113-132.

[4] Y. Wang. and D. Pu, A new QP-free feasible method for inequality constrained optimization, In OR Transacitions,2007, 31-43.

[5] Pu D., Liu A. and Sang Y, QP-free infeasible method without a penalty function and a filter, Operations research transactions,2013, 106-116.

[6] K. Su, A new globally and superlinearly convergent QP-free method for inequality constrained optimization, Journal of Tongji University,2008, 265-272.

Received: October, 2015