East Asian Mathematical Journal
Vol. 28 (2012), No. 3, pp. 293–303

AN ACTIVE SET SQP-FILTER METHOD FOR SOLVING NONLINEAR PROGRAMMING

Ke Su, Yingna Yuan, and Hui An

Abstract. Sequential quadratic programming (SQP) has been one of the most important methods for solving nonlinear constrained optimization problems. Recently, filter method, proposed by Fletcher and Leyffer, has been extensively applied for its promising numerical results. In this paper, we present and study an active set SQP-filter algorithm for inequality constrained optimization. The active set technique reduces the size of quadratic programming (QP) subproblem. While by the filter method, there is no penalty parameter estimate. Moreover, Maratos effect can be overcome by filter technique. Global convergence property of the proposed algorithm are established under suitable conditions. Some numerical results are reported in this paper.

1. Introduction

In this paper, we consider the following nonlinear inequality constrained optimization problem:

$$\begin{aligned} (P) \quad & \min \quad f(x) \\ & \text{s.t.} \quad c_i(x) \geq 0,\ i \in I_0 = \{1, 2, \cdots, m\} \end{aligned} \tag{1}$$

where $x \in R^n, f : R^n \to R$ and $c_i(i \in I_0) : R^n \to R$ are assumed to be twice continuously differentiable.

It is well known that the sequential quadratic programming(SQP) method is one of the most efficient methods to solve the problem (P). Because its superlinear convergence rate, it has been widely studied by many researchers [2, 3, 8, 13, 14].

Received December 28, 2011; Revised March 15, 2012; Accepted March 15, 2012.
2000 *Mathematics Subject Classification.* 90c30.
Key words and phrases. Nonlinear programming, active set, filter method, sequential quadratic programming.
This work was financially supported by the National Natural Science Foundation of China (No.11101115), the Science Foundation of Hebei Province (No.A2010000191) and the Science Foundation of Hebei University (No.2009159).

The SQP method generates a sequence $\{x_k\}$ converging to the desired solution by means of solving the quadratic programming problem

$$\begin{aligned} \min \quad & \nabla f(x_k)^T d + \frac{1}{2} d^T B_k d \\ \text{s.t.} \quad & c_i(x_k) + \nabla c_i(x_k)^T d \geq 0, \ i \in I = \{1, 2, \cdots, m\} \end{aligned} \tag{2}$$

where $B_k \in R^{n \times n}$ is a symmetric positive definite matrix, which is supposed to be an approximate Hessian of Lagrangian

$$L(x, \lambda) = f(x) + \lambda^T c(x).$$

The iteration then has the form

$$x_{k+1} = x_k + t_k d_k$$

where d_k solves (2) and t_k is a step length chosen to reduce a merit function for (1). In majority work, the merit function is normally a penalty function such as l_1 exact penalty function. It has been proved that SQP is global convergent (see[9, 11, 13, 15, 16]).

It is obviously that for large scale problem, the memory requisite for each QP subproblem may be very large if the original problem is large (with great number of constraints). So, the active set technique is always used to tackle it so that the constraints are fewer. In direct observation, the non-active constraints have no effect on the problem in the neighborhood of the solution. Let x^* be a local solution of the original problem (P), the active set at x^* is defined by

$$A(x^*) = \{i | c_i(x^*) = 0\}.$$

There are two obvious advantages for using active set technique. One is decreasing the number of constraints in original problem, the other is the reduction of possibility of the inconsistent of the QP subproblem. Liu[11] proposed an SQP method based on active set. To get the step length, a penalty function as a merit function is introduced in [11]. But the penalty parameter estimate could be problematic to obtain, then in 2002, Fletcher and Leyffer proposed a filter method without penalty function for solving nonlinear programming and it recently attached importance to. Because of its promising numerical results, filter method has been combined with trust region method[12, 17], SQP approach[4, 5], bundle technique[6], interior point strategy[18], line search technique[19, 20] and pattern search method [1].

In this paper, motivated by the above ideas, we propose an active SQP-filter method by combining the subproblem proposed in Liu[11] and the filter technique. The method has the following merits: starts from an arbitrary initial point; requires to solve only one QP problem with only one subset of the constraints; and need not to consider the penalty parameter. In the end, under some conditions, we obtain the global convergence and prove that the algorithm either terminates at a Karush-Kuhn-Tucker(KKT) point within finite steps or generates an infinite sequential whose every accumulation point is a KKT point.

This paper is organized as follows. In section 2, the filter method is introduced. A new SQP-filter method is given in section 3. In section 4, the global convergence theory for the method is presented, and some numerical examples are given in the last section.

2. The notion of a filter

To avoid using the classical merit function with penalty term, in which the penalty parameter is difficult to choose, we adopt the filter technique. The acceptability of steps is determined by comparing the constraint violation and objective function value with previous iterates collected in a filter. The new iterate is acceptable to the filter if it is feasible or the objective function value is sufficiently improved in compared to all iterates bookmarked in the current filter. The promising numerical results led to a growing interest in filter methods in recent years.

In this work, define the violation function $h(c(x))$ by

$$h(c(x)) = \|c^{(-)}(x)\|_\infty \tag{3}$$

where $c_i^{(-)}(x) = \min\{0, c_i(x) : \ i \in I\}$.

To balance the objective function and the constrained function, we substitute $p(x)$ for $f(x)$ as following:

$$p(x) = f(x) + \sigma h(x)$$

where σ is a constant and the value of σ is not required to be very large. If $\sigma = 0$, it is the traditional filter method. If $\sigma < 0$, the accepted conditions are relaxed. So, with the appropriate choice of σ, the Maratos effect can be overcome.

It is easy to see that $h(x) = 0$ if and only if x is a feasible point. So, a trial step should reduce either the constraint value h or the function value p. To ensure sufficient decrease of at least one of the two criteria, we say that a point x_1 dominates a point x_2 whenever

$$h_1 \leq h_2 \text{ and } p_1 \leq p_2 \tag{4}$$

where $h_i = h(c(x_i)), p_i = p(x_i), for\ i = 1, 2.$

All we need to do is to remember iterates that are not dominated by any other iterates using a structure called a filter. A filter is a set F of points in R^n such that no point dominates any other.

In practical computation, we do not wish to accept $x_k + d_k$ if it is arbitrarily close to that of x_k or that of a point already in the filter. Thus we set a small "margin" around the border of the dominate point of the (h, p) space in which we shall also reject trial points. Formally, we say that a point x is acceptable to the filter if and only if

$$h(c(x)) \leq \beta h_j \text{ or } p(x) \leq p_j - \gamma h_j \tag{5}$$

for all $x_j \in F$, where $0 < \gamma < \beta < 1$ is close to zero. As the algorithm progresses, we may want to add a point x to the filter. If a iteration x_k is acceptable for F, we do this by adding the point x_k to the filter and removing from it every other point x_j such that both

$$h_j \geq h_k \text{ and } p_j - \gamma h_j \geq p_k - \gamma h_k. \tag{6}$$

We also refer to this operation as "adding x_k to the filter". We note that if a point x_k is in the filter or is acceptable for the filter, then any other point x such that

$$h(c(x)) \leq \beta h_k \text{ and } p(x) \leq p_k - \gamma h_k \tag{7}$$

is also acceptable for the filter and x_k.

3. An active set SQP-filter algorithm

Let $x \in R^n$ be the current iteration point and $\lambda = (\lambda^{(1)}, \lambda^{(2)}, \cdots, \lambda^{(m)}) \in R^m$ be an approximate multiplier. Define $z = (x, \lambda)$, let $\epsilon > 0$ is a scalar. Define the ϵ-active set at x corresponding to λ as following:

$$I(z, \epsilon) = \{i : c_i(x) \leq \lambda^{(i)} + \epsilon\}. \tag{8}$$

The QP problem that we use as a subproblem is defined by $Q(z, B)$:

$$\begin{aligned} Q(z,B): \quad & \min \quad \nabla f(x)^T d + \frac{1}{2} d^T B d \\ & \text{s.t.} \quad c_i(x) + \nabla c_i(x)^T d \geq 0 \;\; i \in I(z, \epsilon). \end{aligned} \tag{9}$$

We use the following signals as that in [11]:

$$S(z, \epsilon) = \{d : c_i(x) + \nabla c_i(x)^T d \geq 0 \;\; i \in I(z, \epsilon)\}, \tag{10}$$

$$S_0(z, \epsilon) = \{d : c_i(x) + \nabla c_i(x)^T d \geq 0 \;\; i \in I_0\}, \tag{11}$$

where $I_0 = \{1, 2, \cdots, m\}$. Then $S_0(x) \subset S(z, \epsilon)$ since $I(z, \epsilon) \subset I_0$. We can see that if there exists a vector $v \in R^n$ such that

$$\nabla c_i(x)^T v > 0, \;\; i \in I(z, \epsilon). \tag{12}$$

Then $S(z, \epsilon) \neq \emptyset$ follows. If z is a KKT point of the problem (P) and $\epsilon = 0$, then it is precisely the Mangasaria-Fromovitz constraint qualification (MFCQ) at x. Moreover, we can assume that $\nabla c_i(x)$ are linear independent, which is stronger than (12).

Under the above assumption, if B is positive definite, the convex programming $Q(z, B)$ has an unique solution d. Let $\lambda_I = \{\lambda^{(i)} \; : \; i \in I(z, \epsilon)\}$ is a multiplier vector corresponding to d, and $\lambda^{(i)} = 0$ for $i \in I \backslash I(z, \epsilon)$, then we call $\lambda = \{\lambda^{(i)} \; : \; i \in I_0\}$ the multiplier corresponding to problem $Q(z, B)$.

Lemma 3.1. ([11]) *For any $z = (x, \lambda) \in R^{n+m}$, let $p(z, \epsilon) = \min\{c_i(x)/\|\nabla c_i(x)\|_2 : \; i \in I_0 \setminus I(z, \epsilon)\}$ and define the sets*

$$\hat{S}(z, \epsilon) = \{d : \;\; c_i(x) + \nabla c_i(x)^T d \geq 0, \;\; i \in I(z, \epsilon) \text{ and } \|d\|_2 \leq p(z, \epsilon)\}, \tag{13}$$

$$\hat{S}_0(z, \epsilon) = \{d : \;\; c_i(x) + \nabla c_i(x)^T d \geq 0, \;\; i \in I_0 \text{ and } \|d\|_2 \leq p(z, \epsilon)\}. \tag{14}$$

Then we have $\hat{S}(z,\epsilon) = \hat{S}_0(z,\epsilon)$.

Lemma 3.2. ([11]) *Suppose that* B *is positive definite and* d_+ *is the unique solution of problem* $Q(z,B)$. *There always exists a positive constant* δ *such that for* $0 \le \tau \le \delta$,

$$\begin{aligned} c_i(x) + \nabla c_i(x)^T(\tau d_+) &\ge 0 \quad i \in I_0 \setminus I(z,\epsilon), &(15)\\ c_i(x) + \nabla c_i(x)^T(\tau d_+) &\ge (1-\tau)c_i(x) \quad i \in I(z,\epsilon). &(16)\end{aligned}$$

Furthermore,

$$h(c(x) + \nabla c(x)^T d) \le (1-\tau)h(c(x)). \tag{17}$$

In above Lemma, let $\bar{I}(z,\epsilon) = \{i \in I_0 \setminus I(z,\epsilon) :\ \nabla c_i(x)^T d_+ < 0\}$, then

$$\delta = \min\{\min_{i\in \bar{I}(z,\epsilon)} \frac{-c_i(x)}{\nabla c_i(x)^T d_+}; 1\}$$

is satisfied (15)(16) and (17).

Let $d = \tau d_+$, then

$$\begin{aligned} c_i(x) + \nabla c_i(x)^T d &\ge 0 \quad i \in I_0 \setminus I(z,\epsilon), &(18)\\ \tau c_i(x) + \nabla c_i(x)^T d &\ge 0 \quad i \in I(z,\epsilon). &(19)\end{aligned}$$

Lemma 3.3. ([5]) *Consider sequences* $\{h_k\}$ *and* $\{p_k\}$ *such that* $h_k \ge 0$ *and* p_k *is monotonically decreasing and bounded below. Let constants* β *and* γ *satisfied* $0 < \gamma < \beta < 1$ *for all* k,

$$\text{either } h_{k+1} \le \beta h_k \text{ or } p_{k+1} \le p_k - \gamma h_k \tag{20}$$

then $h_k \to 0$.

Lemma 3.4. ([5]) *Consider an infinite sequence of iterations on which* (h_k, p_k) *is entered into the filter, where* $h_k > 0$ *and* $\{p_k\}$ *is bounded below. It follows that* $h_k \to 0$.

Algorithm A

Step 0: Initialization:

Given $x_0 \in R^n$, $\lambda_0 \in R^m, \lambda_0 \ge 0$. B_0 is a symmetric positive definite matrix. $\epsilon, \epsilon_0 > 0, 0 < \gamma < \beta < 1, k = 0$;

Step 1: Solve $Q(z_k, B_k)$ to get the solution d_k'. If $\|d_k'\| \le \epsilon$, then stop;

Step 2: Let $\bar{I}_k = \{i \in I_0 \setminus I(z_k,\epsilon_k) : \nabla c_i(x_k)^T d_k' < 0\}$, compute $\delta = \min\{\min_{i\in \bar{I}(z,\epsilon)} \frac{-c_i(x)}{\nabla c_i(x)^T d_+}; 1\}$ and set $d_k = \delta_k d_k'$;

Step 3: Let $l = 0,\ \alpha_{k,l} = 1$;

Step 4: $\bar{x}_k = x_k + \alpha_{k,l} d_k$, if $\bar{x}_k$ is acceptable to the filter, then set $\alpha_k = \alpha_{k,l}, x_{k+1} = \bar{x}_k$ and go to step 6;

Step 5: $\alpha_{k,l+1} = \dfrac{\alpha_{k,l}}{2}, l = l+1$ and go to step 4;

Step 6: If $g_k^T d_k > -\frac{1}{2} d_k^T B_k d_k$, then add x_{k+1} to the filter. Update B_k to $B_{k+1}, \lambda_{k+1} = \lambda_k^{'}$, $\epsilon_{k+1} = \epsilon_k/2$, set $k = k+1$ and go to step 1.

We call step4-step5-step4 the inner loop and the whole cycle outer loop.

4. The global convergence properties

Just as in [1, 4, 5], our analysis of the algorithm are based on the standard assumptions as follows.

Assumptions:

A1. The objective function f and the constraint functions c_i $(i \in I_0)$ are twice continuously differentiable.

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

A3. All points that are sampled by the algorithm lie in a nonempty closed and bounded set $S \subset R^n$.

A4. If $x_k \to x^*(k \in K)$, then there exists a vector $v \in R^n$ such that

$$\nabla c_i(x^*)^T v > 0 \ i \in \Im(x^*, \epsilon)$$

where $\Im(x^*, \epsilon) = \{i : i \in I(z_k, \epsilon) \text{ for infinitely many } k \in K\}$, K is an infinite index set and ϵ is a positive scalar.

The assumption (A1) and (A3) are the standard assumptions. (A4) is the sufficient condition for that $Q(z_k, B_k)$ is solvable, which is a weak condition compared to the MFCQ conditions. Fletcher et.al [7] have showed that (A1) and (A3) together directly ensure that, for all k

$$f(x_k) \ge f_{min} \text{ and } 0 \le h(c(x_k)) \le h_{max}$$

for some constants f_{min} and $h_{max} > 0$. Thus we see that the sequence $\{p(x_k)\}$ is bounded below, and we can also assume there exists $\rho > 0$, such that $h(c(x_k)) \le \rho$. Also, let us assume, by (A1) and (A3), without loss the generality, that $f(x), \nabla f(x), c(x), \nabla c(x)$ are bounded on S.

Lemma 4.1. *Suppose the standard assumptions hold. If (d_k', λ_{k+1}) is a KKT point of $Q(z_k, B_k)$, then $\|d_k'\|$ and λ_{k+1} are bounded.*

Proof. Since $\{x_k\}$ lie in a bounded set, there exists a point x^* such that $x_k \to x^*$ $(k \in K)$, where K is an infinite index set. By (A4), it follows

$$c_i(x^*) + \nabla c_i(x^*)^T d^* \ge 0 \ i \in \Im(x^*, \epsilon) \tag{21}$$

for some $d^* \in R^n$.

Note that the functions $c(x), \nabla c(x)$ are continuous, we thus obtain that, there exists a $k_0 > 0$, for $k > k_0$,

$$c_i(x_k) + \nabla c_i(x_k)^T d^* \ge 0 \ i \in \Im(x^*, \epsilon). \tag{22}$$

By the definition of $\Im(x^*, \epsilon)$, we can see there is a constant $k_1 > 0$, such that $\Im(z_k, \epsilon_k) \subset \Im(x^*, \epsilon)$ for $k > k_1$. Thus d^* is a feasible point of $Q(z_k, B_k)$ for all $k \ge \max\{k_1, k_0\}$.

The $\|d'_k\|$ is bounded follows by the assumption (A2).

Since $\lambda_{k+1}^{(i)} = 0$ for $i \notin \Im(x^*, \epsilon)$, we just need to show $\lambda_{k+1}^{(i)}$ $(i \in I(z_k, \epsilon),\ k \in K)$ is bounded in order to prove the conclusion.

By the KKT condition of $Q(z_k, B_k)$, we then obtain

$$g_k + B_k^T d'_k - \lambda_{k+1}^T A_k = 0,\ \lambda_{k+1}^T (c(x_k) + A_k d'_k) = 0 \tag{23}$$

where $A_k = (\nabla c_i(x_k))^T\ :\ (i \in I(z_k, \epsilon))$.

We then get the desired result by the standard assumptions and the above explanation. □

Without loss of generality, according to the above illustration and Lemmas, we can suppose there exist $M, \rho > 0$ such that $\|\lambda_{k+1}\| \le M, \|f(x)\| \le M, \|\nabla f(x)\| \le M, \|c(x)\| \le M, \|\nabla c(x)\| \le M$ and $\|d_k\| \le \rho$.

Lemma 4.2. *Under the standard assumptions, the inner loop terminate in finite times.*

Proof. Suppose the point x_k is the last point that entered into the filter. We will show that $x_k + \alpha d_k$ will be accepted by the filter for sufficiently small $\alpha > 0$ in the following two cases.

Case I: $h(c(x_k)) = 0$

It means that x_k is a feasible point of the problem (P) and hence $d_k = 0$ is the solution of $Q(z_k, B_k)$, so by the definition of $h(c(x_k))$, we have

$$\begin{aligned} h(c(x_k + \alpha d_k)) &= \max\{0, -c_i(x_k + \alpha d_k)\} \\ &= \max\{0, -c_i(x_k) - \alpha \nabla c_i(x_k)^T d_k + o(\alpha)\} \\ &= \max\{0, -\alpha(c_i(x_k) + \nabla c_i(x_k)^T d_k) - (1-\alpha) c_i(x_k) + o(\alpha)\} \\ &= 0. \end{aligned} \tag{24}$$

We thus obtain $h(c(x_k + \alpha d_k)) \le \beta h(c(x_k))$.

Together with (24), $x_k + \alpha d_k$ will be therefore accepted by filter because the definition of d_k and

$$\begin{aligned} p(x_k) - p(x_k + \alpha d_k) &= f(x_k) - f(x_k + \alpha d_k) && (25) \\ &= -\alpha \nabla f(x_k)^T d_k + o(\alpha) \\ &\ge \frac{\alpha}{2} d_k^T B_k d_k + o(\alpha) > 0. && (26) \end{aligned}$$

Case II: $h(c(x_k)) > 0$

By the intermediate value of Taylor's theorem, we have

$$c_i(x_k + \alpha d_k) = c_i(x_k) + \alpha \nabla c_i(x_k)^T d_k + \frac{\alpha^2}{2} d_k^T \nabla c_i^2(y) d_k^T \tag{27}$$

where y denotes some point on the line segment from x_k to $x_k + \alpha d_k$. It follows from the definition of $h(c(x_k))$ and assumption (A3), that

$$h(c(x_k + \alpha d_k)) \le \frac{\alpha^2}{2} \|d_k\|^2 M \le \frac{\alpha^2}{2} \rho^2 M. \tag{28}$$

Define $\tau_k = \min_{x_j \in F} h(c(x_j))$. It is convenient to get $\tau_k > 0$ by the construction of the algorithm. So, if $\alpha^2 < \frac{2\beta\tau_k}{\rho^2 M}$, we obtain

$$h(c(x_k + \alpha d_k)) \le \beta\tau_k \le \beta h(c(x_j)) \tag{29}$$

for all points x_j in the filter. Hence, the Lemma holds for sufficiently small α. □

We are now in a position to state the global convergence of our algorithm.

Theorem 4.3. *Suppose there are infinitely many points entered into the filter. Then* $\lim_{k\to\infty} h(c(x_k)) = 0$.

Proof. If $h(x_k) = 0$, then x_k is a feasible point. By the algorithm, we have $g_k^T d_k + \frac{1}{2} d_k^T B_k d_k \le 0$. It follows that x_k is not entered into the filter. So, $h^{(k)} > 0$. Hence, by Lemma 2 and together with the fact that $\{p(x_k)\}$ is bounded below, we obtain $\lim_{k\to\infty} h(x_k) = 0$. □

Theorem 4.4. *Suppose there are finitely many points entered into the filter. Then* $h(c(x_k)) = 0$.

Proof. The result is obvious from the algorithm. □

Theorem 4.5. *Assume the standard assumptions hold and* $\{x_k\}$ *is an infinite sequence generated by algorithm. Then any accumulation point of* $\{x_k\}$ *is a KKT point of the problem (P).*

Proof. Because $\{x_k\}$ lie in a bounded set S, there must exits x^*, such that $x_k \to x^*$ $(k \in K)$, which K is an infinite index set. By the algorithm, we prove the theorem in the following two possible cases:

Case I: There are infinite many points entered into the filter.

In this case, by the algorithm, we have $K_1 = \{k \in K | \nabla f(x_k)^T d_k > -\frac{1}{2} d_k^T B_k d_k\}$ is an infinite index set. Also from Theorem 4.3, we get $h(c(x_k)) \to 0$, $(k \in K_1)$. So, x^* is a feasible point. Suppose by contradiction that x^* is not a KKT point, and if we assume there exists a set $K_2 \subset K_1$ such that $\|d_k\| \to 0$ $(k \in K_2)$, then it is easy to see x^* is a KKT point. Hence, without loss of generality, we suppose that $\|d_k\| > \epsilon$ for some constant $\epsilon > 0$. By the definition of d_k, $\|d_k\| = \|\delta_k d'_k\| \le \|d'_k\|$, where $0 < \delta_k < 1$, d'_k is the solution of $Q(z_k, B_k)$, then $\|d'_k\| > \epsilon$. Following from $h(c(x_k)) \to 0$, we can assume $\exists k_0$, for $k > k_0, k \in K_1$, it holds

$$h(c(x_k)) \le \frac{a(2-\delta_k)\epsilon^2}{2M} \le \frac{a\|d'_k\|^2(2-\delta_k)}{2M} \le \frac{2-\delta_k}{2M} d_k^T B_k d_k. \tag{30}$$

While by KKT condition of $Q(z_k, B_k)$, it follows that

$$g_k + B_k d'_k + \lambda_{k+1}^T A_k = 0,\ \lambda_{k+1}^T (c(x_k) + A_k d'_k) = 0, \tag{31}$$

where $A_k = \nabla c(x_k)^T$. Together with (29), we obtain that for all $k \in K_1, k > k_0$, it holds

$$\begin{aligned} g_k^T d_k' &= -(d_k')^T B_k d_k' - \lambda_{k+1}^T A_k d_k' \\ &= -(d_k')^T B_k d_k' - \lambda_{k+1}^T c(x_k) \\ &\leq -(d_k')^T B_k d_k' + M h(c(x_k)) \\ &\leq -(d_k')^T B_k d_k' + \frac{2-\delta_k}{2}(d_k')^T B_k d_k \\ &\leq -\frac{\delta_k}{2} d_k^T B_k d_k. \end{aligned} \tag{32}$$

So, $g_k^T(\delta_k d_k') \leq -\frac{1}{2}(\delta_k d_k')^T B_k(\delta_k d_k')$, then $g_k^T d_k \leq -\frac{1}{2} d_k^T B_k d_k$ $(k \in K_1)$. Which contradicts the definition of K_1. It follows that x^* is a KKT point.

Case II: There are finite many points entered into the filter.

That means it holds $g_k^T d_k \leq -\frac{1}{2} d_k^T B_k d_k < 0$ for k sufficiently large, and K_1 is a finite index set.

There must exists $\bar{\alpha} > 0$ such that

$$p(x_k) - p(x_k + \alpha d_k) \geq -\alpha g_k^T d_k + o(\alpha) \geq \frac{\bar{\alpha}}{2} d_k^T B_k d_k.$$

Because p is bounded below, for some integer i_0, we have

$$\infty > \sum_{k=i_0}^{\infty} (p(x_k) - p(x_k + \alpha d_k)) \geq \sum_{k=i_0}^{\infty} \frac{a\alpha}{2} \|d_k\|^2.$$

Then

$$\sum_{k=i_0}^{\infty} \|d_k\|^2 < +\infty.$$

That means $\|d_k\| \to 0$. Hence x^* is a KKT point of the problem (P). □

Lemma 4.6. ([11]) *Under the assumption of Theorem 4.5, suppose that* $x_k \to x^*$ $(k \to \infty), \nabla c_i(x^*)$ $(i \in I^*(x^*))$ *are linearly independent,* λ^* *is the multiplier associated with* x^**. If* $\epsilon_k \to 0$*, the strict complementarity condition holds at* z^**, then* $I(z_k, \epsilon_k) = I^*(z^*)$ *for all sufficiently large* k*.*

5. Some numerical experiments

In this section, we discuss further refinements of the algorithm proposed above to accommodate practical calculations, and give some numerical experiments to show the success of proposed method. All examples are chosen from [10].

(1) Updating of B_k is done by

$$B_{k+1} = \begin{cases} B_k \text{ if } s_k^T y_k \leq 0, \\ B_k + \dfrac{y_k^T y_k}{y_k^T s_k} - \dfrac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} \text{ if } s_k^T y_k > 0. \end{cases} \tag{33}$$

(2) The stop criteria is $\|d_k\|$ sufficiently small.
(3) An equality constraint $c(x) = 0$ exists in the original problem, it is most easily handle as two corresponding inequalities $c(x) \leq 0$ and $c(x) \geq 0$, and we can apply the above algorithm.
(4) The algorithm parameters were set as follows: $\gamma = 0.05, \beta = 0.95, H_0 = I \in R^{n\times n}, \epsilon = 1e\text{-}06$. The program is written in Matlab.
(5) Our method has no demand on the initial point. It can be either feasible(HS3,5,31,33,35,44,113) or infeasible(HS15,23,41,45,53).

Numerical results for the algorithm are listed in Table 1.

Table 1

No.	n	m	NI	Filter-NI	FV
HS3	2	1	5	10	0.0000
HS5	2	4	8	8	-1.9132
HS15	2	3	3	21	306.5000
HS23	2	9	7	8	0.0000
HS31	3	7	3	28	6.0000
HS33	3	6	2	5	-4.5858
HS35	3	7	7	7	0.1111
HS41	4	9	8	7	1.9259
HS44	4	14	6	6	-15.0000
HS45	5	10	2	8	1.0000
HS53	5	13	8	8	4.0930
HS113	10	8	16	36	24.3062

For each test problem, No. is the number of the test problem in [10], for example, HS3 refers to the problem 3 in [10]. n refers to the number of variables, m the number of inequality constraints, NI the number of iterates for our algorithm, Filter-NI the number iterates for the traditional filter method, FV the final value of the objective function.

The result in Table 1 indicate that our algorithm is quit effective compared to the traditional filter method.

References

[1] C. Audet and J. E. Dennis, *A patter search filter method for nonlinear programming without dereivatives*, SIAM J. Optim. (2004), no. 1, 980–1010.
[2] J. F. Bonnons, E. R. Painer, A. L. Titts and J. L. Zhou, *Avoiding the Maratos effect by means of nonmontone linesearch, Inequality constrained problems-feasible iterates*, SIAM J. Numer. Anal. (1992), no. 29, 1187–1202.
[3] P. T. Boggs, J. W. Tolle and P. Wang, *On the local convergence of quasi-newton methods for constrained optimization*, SIAM J. Contr. Optim. (1982), no. 20, 161–171.
[4] R. Fletcher and S. Leyffer, *Nonlinear programming without a penalty function*, Math.Program. (2002), no. 91, 239–269.
[5] R. Fletcher, S. Leyffer and P. L. Toint, *On the global convergence of a filter-SQP algorithm*, SIAM J. Optim. (2002), no. 13, 44–59.

[6] R. Fletcher and S. Leyffer, *A boundle filter method for nonsmooth nonlinear optimization*, Technical Report NA/195, Department of Mathematics, University of Dundee, Scotland, December, 1999.
[7] R. Fletcher, N. I. M. Gould, S. Leyffer, P. L. Toint and A. Wachter, *Global convergence of a trust region SQP-filter algorithm for general nonlinear programming*, SIAM J. Optim. (2002) no. 13, 635–660.
[8] S. P. Han, *Superlinearly convergence variable metric algorithm for general nonlinear programming problems*, Math. Program. (1976) no. 11, 263–282.
[9] S. P. Han, *A globally convergent method for nonlinear programming*, J. Optim. Theory Appl. (1977) no. 22, 297–309.
[10] W. Hock and K. Schittkowski, *Test examples for nonlinear programming codes*, Lecture notes in econom. and math. systems, 187, Springer-Verlag, Berlin, 1981.
[11] X. W. Liu, *Global convergence on an active set SQP for inequality constrained optimization*, J. Comput. Appl. Math. (2005) no. 180, 201–211.
[12] P. Y. Nie and C. F. Ma, *A trust region filter method for general nonlinear programming*, Appl. Math. Comput. (2006) no. 172, 1000–1017.
[13] M. J. D. Powell, *A fast algorithm for nonlinear constrained optimization calculations*, in Waston, G. A. (Ed.). Numerical Analysis, Springer-Verlag,Berlin. pp.144–157, 1978.
[14] M. J. D. Powell, *Variable metric methods for constrained optimization*, in Bachen et al.(Eds.). Mathematcal programming-The state of art, Springer-Verlag, Berline. 1982.
[15] M. J. D. Powell, *The convergence of variable metric methods for nonlinear constrained optimiation calculations*, in: O. L. Mangasarian, R. R. Meyer, S. M. Robinson(Eds.) Nonlinear Programming **3**, Academic Press, New York, pp.27–63, 1978.
[16] J. Stoer, *Principles of sequential quadratic programming methods for solving nonlinear programs*, in: K.Schittkawsk;(Ed.) NATO ASI Series, vol. F15, Computational Mathematical Programming, Springer, Berlin.
[17] M. Ulbrich and S. Ulbrich, *Nonmonotone trust region method for nonlinear equality constrained optimization without a penalty function*, Math. Program., SerB (2003). 103–113.
[18] M. Ulbrich, S. Ulbrich and L. N. Vicente, *A global convergent primal-dual interior filter method for nonconvex nonlinear programming*, Math Program. (2004) no. 100, 379–410.
[19] A. Watcher and L. T. Biogle, *Line search filter methods for nonlinear programming:motivation and global convergence*, SAIM J. Optim. (2005) no. 16, 1–31.
[20] A. Wathcer and L. T. Bigole, *Line search filter methods for nonlinear programming: local convergence*, SIAM J. Optim. (2005), no. 16, 32–48.

College of Mathematics and Computer Sciencs, Hebei University, Baoding, 071002, P.R.China
E-mail address: pigeonsk@163.com