Miskolc Mathematical Notes
Vol. 24 (2023), No. 1, pp. 457–471
HU e-ISSN 1787-2413
DOI: 10.18514/MMN.2023.2910

A MODIFIED NONMONOTONE FILTER QP-FREE METHOD

KE SU AND LELE REN

Received 22 March, 2019

Abstract. In this paper, an infeasible QP-free method without penalty function is proposed for inequality constrained optimization. We first compute a fundamental direction and then bend the search direction based on the constraint function and the Lagrange multiplier. Based on the modified nonmonotone filter technique, the acceptable criterion of trial points is relaxed and Maratos effects are avoided to a certain degree. At each iteration, only two or three systems of linear equations with the same coefficient are needed to solve to obtain the search direction. Under suitable conditions, the global convergence of the algorithm is proved without the strict complementarity conditions. In the end, some numerical results are reported.

2010 *Mathematics Subject Classification:* 90C30; 65K05

Keywords: inequality constrained optimization, QP-free method, active set, global convergence, filter

1. Introduction

Consider the nonlinear optimization with general inequality constraints:

$$\begin{aligned} &\min f(x) \\ &\text{subject to } \; g_i(x) \le 0,\; i \in I = \{1,2,\dots,m\}, \end{aligned} \tag{1.1}$$

where $f\colon R^n \to R$ and $g_i\colon R^n \to R^m$, $i \in I$ are twice continuously differentiable functions.

It is well known that sequential quadratic programming (SQP)method is one of the effective methods for solving nonlinearly constrained optimization problem and has been widely investigated by many authors [8, 10]. However, the search direction of SQP method is obtained by solving a quadratic programming subproblem in each iteration, which greatly increases the computational scale. To avoid the drawback, various QP-free approaches, also called sequential systems of linear equations (SSLE) methods, are proposed for (1.1).

This work was supported by the National Natural Science Foundation of China(No. 61572011), Hebei Province Nature Science Foundation of China(A2022201002), and the Key Research Foundation of Education Bureau of Hebei Province(No.ZD2015069).

In [15], Panier et al proposed a feasible sequential quadratic programming algorithm, the algorithm is shown to possess global convergence as well as two-step superlinear convergence under the relevant conditions, and Maratos effect can be overcome to some extent. But it requirs that the number of isolated stationary points is finite. The algorithm is later improved by Gao et al. [7]. In 2000, Qi [16] gave a new feasible QP-free algorithm based on Fischer-Burmeister function [4], which does not require strict complementarity conditions and does not need to assume the isolation of stable points. Later, Qi et al. [17] proposed an infeasible QP-free algorithm based on a continuous differentiable exact penalty function and an efficient recursive QP algorithm model,which are presented by Lucidi [14] and Facchinei [3] respectively. Tits and Bakhtiari [2] presented a simple primal-dual feasible interior point algorithm for the problem (1.1). Based on the idea in [2], Jian et al. [11] presented a primal-dual quasi interior-point algorithm, the associated system of linear equations possesses a smaller scale and requires less computational cost than that in [2]. Then Li et al. [1] and Huang et al. [12] developed QP-free methods based on the smoothing techniques and the working set techniques, but this method assumed that the Hessian estimate was positive definite. In 2017, Wang et al. [21] proposed an infeasible active-set QP-free algorithm based on filter technique without the positive definite assumption on the Hessian estimate.

Penalty function is usually used as merit function to decide whether the trial point is accepted at the new iteration, but it is well known that the chosen of penalty parameter is difficult. If the penalty parameter is too large, then any monotonic method would be forced to follow the nonlinear constraint manifold very closely, resulting in shortened Newton steps and slow convergence. On the other hand, too small a choice of the penalty parameter may result in an infeasible point, or even an unbounded increase in the penalty. Therefore, Fletcher and Leyffer [5] proposed filter technique and gave a large number of numerical experiments to prove the validity of filter method. After that, many different filter methods were proposed. For example, filter interior point approach [6], line search filter method [20], a modified trust-region filter [19], a nonmonotone filter method which used a global g-filter for global convergence [18], a nonmonotone line search multidimensional filter-SQP method [9].

Motivated by the above ideas, we propose a QP-free method with filter technique which solve two or three linear equation systems with the same coefficient matrix. Compared with the existing methods, our method has several advantages:

(1) We first obtain a fundamental direction and then bend the search direction based on the constraint function and Lagrange multiplier.
(2) The initial point is not needed to be feasible.
(3) There is no penalty function so that the penalty parameter is avoided.

(4) The strict complementarity conditions are not required, computation scale is decreased by working set, moreover, restoration phase is not required, which is needed in most of the traditional filter methods.

By the numerical results, we show that the proposed method is effective.

The remainder of this paper is organized as follows: In Section 2, we introduce our QP-free algorithm based on filter technique. The global convergence of this algorithm is established in Section 3. Some numerical experiments are shown in Section 4.

2. The algorithm

In this section, we present some related concepts and symbols. A given point $x \in R^n$ is said to be a Karush-Kuhn-Tucker (KKT) point of problem (1.1) if there exits a vector $\lambda \in R^m$ such that

$$\nabla_x L(x,\lambda) = 0, \qquad \lambda_i g_i(x) = 0, \qquad \lambda_i \geq 0, \qquad g_i(x) \leq 0 \qquad \forall i \in I \tag{2.1}$$

where the vector $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m)^T$ is the corresponding lagrangian multiplier, $L(x,\lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x)$ is the Lagrangian function of problem (1.1).

Let (x^*, λ^*) denote a KKT point of problem (1.1). Define $\Phi\colon R^{n+m} \to R^{n+m}$:

$$\Phi(x,\lambda) = \begin{pmatrix} \nabla_x L(x,\lambda) \\ \min\{-G(x), \lambda\} \end{pmatrix},$$

where $G(x) = (g_1(x), g_2(x), \ldots, g_m(x))^T$, $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m)^T$. It is obviously that (2.1) and $\Phi(x,\lambda) = 0$ are equivalent. Then, we define another function $\varphi\colon R^{n+m} \to R$

$$\varphi(x,\lambda) = \sqrt{\|\Phi(x,\lambda)\|},$$

where $\|\cdot\|$ denotes the Euclidean norm. The function φ is non-negative and continuous. And it means that (x^*, λ^*) is a KKT point of problem (1.1) if and only if $\varphi(x^*, \lambda^*) = 0$. Let the active set

$$I(x) = \{i \in I \mid g_i(x) = 0\},$$

and two working sets

$$J_\varepsilon(x,\lambda) = \{i \in I \mid g_i(x) \geq -\varepsilon \min\{\varphi(x,\lambda), \varphi_{\max}\}\},$$
$$\overline{J}_\varepsilon(x,\lambda) = \{i \in J_\varepsilon \mid \lambda_i \geq \varepsilon \min\{\varphi(x,\lambda), \varphi_{\max}\}\},$$

where $\varphi_{\max} > 0$, $\varepsilon > 0$, $J_\varepsilon(x,\lambda)$ is an estimate of the final active set $I(x)$, and $\overline{J}_\varepsilon(x,\lambda)$ is stronger than working set $J_\varepsilon(x,\lambda)$. To simplify the presentation, we set $J_{\varepsilon_k}(x^k, \lambda^{k-1,0})$ be W_k and $\overline{J}_{\varepsilon_k}(x^k, \lambda^{k-1,0})$ be $\overline{W}_k$, where ε_k and (x^k, λ^k) are at the kth iteration.

In order to avoid computing linearly independent constraint gradients, inspired by [13], the coefficient matrix V_k of our Newton equations involve only constraints in the working set W_k,

$$V_k = \begin{pmatrix} H_k & \nabla g_{W_k}(x^k) \\ U_k \nabla g_{W_k}(x^k)^T & G_{W_k}(x^k) \end{pmatrix}, \tag{2.2}$$

where $H_k \in R^{n\times n}$ is an estimate of the Lagrangian Hessian, $U_k = \text{diag}(\mu_i^k)$, $i \in W_k$ and $G_{W_k}(x^k) = \text{diag}(g_i(x^k))$, $i \in W_k$. Note that μ_i^k which are defined as

$$\mu_i^k = \begin{cases} \theta_k + \max\{\lambda_i^{k-1,0}\}, & i \in W_k; \\ \theta_k, & i \in I \backslash W_k; \end{cases}$$

are controlled to be componentwise bounded, where $\theta_k = \nu \min\{\lambda_i^{k-1,0}, i \in \overline{W}_k\}$ if $\overline{W}_k \neq \varnothing$ and $\varphi(x^k, \lambda^{k-1,0}) > 0$; otherwise $\theta_k = \theta$, θ and ν are positive constants.

In this paper, we use a nonmonotone filter to determine a trial point is accepted or not. In order to prove the following lemma, we give some definitions as follows:

Definition 1. Define the constrained violation function $h\colon R^n \to R$ by

$$h(x) = \sum_{i\in I} \max\{g_i(x), 0\}.$$

It is easy to see that $h(x) = 0$ if and only if x is a feasible point.

Definition 2. A point x^i is dominated by x^j if and only if $h(x^i) \leq h(x^j)$ and $f(x^i) \leq f(x^j)$ for each $i \neq j$.

We give the concept of filter subsets based on the above definitions.

Definition 3. A filter set $\mathscr{F}$ is a set of pairs (h, f) such that no pair dominates any other.

So, we have a definition of whether the trial point x^i is accepted by the filter.

Definition 4. A trial point x^i is called acceptable to the filter if and only if either

$$h(x^i) \leq h(x^j) \text{ or } f(x^i) \leq f(x^j) \qquad \forall (h(x^j), f(x^j)) \in \mathscr{F},\ i \neq j.$$

In the actual calculation, some filter point pairs may fall on the boundary, resulting in convergence to the infeasible limit points where $h > 0$. In order to avoid this situation, a modified filter method is presented by adding an envelope to the current filter.

Definition 5. A trial point x^k is called acceptable to the filter if and only if either

$$h(x^k) \leq (1-\gamma)h(x^j) \text{ or } f(x^k) \leq f(x^j) - \gamma h(x^k)$$
$$\forall (h(x^j), f(x^j)) \in \mathscr{F}$$

where γ is a constant that close to 0.

So we give a nonmonotone modified filter that substitute Definition 4 and Definition 5 with the following Definition 6 in our actual algorithm.

Definition 6. A trial point x^k is called acceptable to the filter if and only if either

$$h(x^k) \leq \max_{0<j<m(k)} (1-\gamma)h(x^j) \text{ or } f(x^k) \leq \max_{0<j<m(k)} f(x^j) - \gamma h(x^k)$$

$$\forall (h(x^j), f(x^j)) \in \mathscr{F}$$

where γ is a constant that close to 0, $m(0) = 0$, $0 \le m(k) \le \min[m(k-1)+1, M]$ for $k \ge 1$, $M \ge 1$ is a given positive constant.

We add the point x^k to the filter, and update the filter set which is $\mathscr{F}_+ = \mathscr{F} \cup (h^k, f^k)$, then remove point x^j that satisfied

$$h(x^j) \ge h(x^k) \text{ and } f(x^j) - \gamma h(x^j) \ge f(x^k) - \gamma h(x^k),$$

so, the new filter set is

$$\mathscr{F}_+ = \mathscr{F} \cup (h^k, f^k) \backslash \{(h^j, f^j) | h^j \ge h^k \text{ and } f^j - \gamma h^j \ge f^k - \gamma h^k\},$$

we also refer to this operation as 'adding x^k to the filter'.

In the following algorithm, we are going to update θ_k and μ_{k+1} through the working set $\overline{W}_k$, and the algorithm is finished on the working set W_k, which reduces the calculation scale. The algorithm obtains the initial direction $d^{k,0}$ and the corresponding multiplier $\lambda^{k,0}$ by solving the first linear equation system, and then bends the search direction based on $d^{k,0}$, $\lambda^{k,0}$ and constraint functions.

In order to obtain the convergence of the algorithm, we use the nonmonotone filter technique to replace the traditional filter, thus, the acceptance criteria of trial points x^k are relaxed. We are now ready to state the algorithm.

Algorithm 1.

Step 1. *Give an initial point* $x^1 \in R^n$. $t, \gamma \in (0,1)$, $h_{\max} > 1$, $\chi_1 \gg \varphi_{\max}$, $\varepsilon_1 > 0$, $\varphi_{\max} > 0$, $\lambda^{0,0} > 0$, $\omega \in (2,3)$, $\mathscr{F}_1 = \{(h_{\max}, -\infty)\}$ *with* $h(x^1) \ll h_{\max}$, $\theta > 0$, $\nu \in (0,1)$, $\rho \in [0,1]$. *If* $\overline{W}_1 \ne \varnothing$ *and* $\varphi(x^1, \lambda^{0,0}) > 0$, $\theta_1 = \nu \min\{\lambda_i^{0,0}, i \in \overline{W}_1\}$*; otherwise,* $\theta_1 = \theta$. $\mu_1 = \lambda^{0,0} + \theta_1 e$, $H_1 = \nabla^2_{xx} L(x^1, \lambda^{0,0})$. *Set* $k = 1$.

Step 2. *Compute* $d^{k,0}$ *and* $\lambda^{k,0}$ *by solving the linear system in* (d, λ)*:*

$$V_k \begin{pmatrix} d \\ \lambda \end{pmatrix} = - \begin{pmatrix} \nabla f(x^k) \\ 0 \end{pmatrix} \tag{2.3}$$

Set $\lambda_i^{k,0} = 0$, $i \in I \backslash W_k$.

Step 3. *Compute* $d^{k,1}$ *and* $\lambda^{k,1}$ *by solving the linear system in* (d, λ)*:*

$$V_k \begin{pmatrix} d \\ \lambda \end{pmatrix} = - \begin{pmatrix} \nabla f(x^k) \\ (1-\rho)\mu \|d^{k,0}\|^{\omega} + \rho \theta_k v^k_{W_k} \end{pmatrix} \tag{2.4}$$

where $v^k_{W_k} = (v_i^k, i \in W_k)$,

$$v_i^k = \begin{cases} \min\{-g_i(x^k), \lambda_i^{k,0}\}, & \lambda_i^{k,0} < 0; \\ -g_i(x^k), & \textit{otherwise.} \end{cases}$$

Set $\lambda_i^{k,0} = 0$, $i \in I \backslash W_k$. *If* $\nabla f(x^k)^T d^{k,1} = 0$ *and* $h(x^k) = 0$, *stop.*

Step 4. *Set* $l = 0$, $\alpha_{k,l} = 1$.

Step 5. *If* $x^k + \alpha_{k,l} d^{k,1}$ *is acceptable for the filter, let* $\alpha_k = \alpha_{k,l}, p^k = \alpha_k d^{k,1}$, *go to step 7.*

Step 6. *Compute* $d^{k,2}$ *and* $\lambda^{k,2}$ *by solving the linear system in* (d, λ)*:*

$$V_k \begin{pmatrix} d \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ -G_{W_k}(x^k + d^{k,1}) \end{pmatrix}$$

Set $\lambda_i^{k,2} = 0$, $i \in I \backslash W_k$. *If* $\|d^{k,2}\| > \|d^{k,1}\|$, *set* $\|d^{k,2}\| = 0$. *If* $x^k + d^{k,1} + d^{k,2}$ *is acceptable for the filter, set* $p^k = d^{k,1} + d^{k,2}$, *go to step 7. Set* $\alpha_{k,l+1} = t\alpha_{k,l}$, $l = l+1$, *go back to step 5.*

Step 7. *Set* $x^{k+1} = x^k + p^k$ *and add* x^{k+1} *to the filter.*

Step 8. *Update. If* $\|\lambda^{k,0}\|_\infty > \chi_k$, *set* $\varepsilon_{k+1} = \frac{1}{2}\varepsilon_k$ *and* $\chi_{k+1} = 2\chi_k$*; else set* $(\varepsilon_{k+1}, \chi_{k+1}) = (\varepsilon_k, \chi_k)$. *If* $\overline{W}_{k+1} \neq \phi$ *and* $\varphi(x^{k+1}, \lambda^{k,0}) > 0$, *set* $\theta_{k+1} = \nu \min\{\lambda_i^{k,0}, i \in \overline{W}_{k+1}\}$*; otherwise,* $\theta_{k+1} = \theta$. *Set*

$$\mu_i^{k+1} = \begin{cases} \theta_{k+1} + \max\{\lambda_i^{k,0}, 0\}, & i \in W_k; \\ \theta_{k+1}, & i \in I \backslash W_k. \end{cases}$$

Update H_{k+1}, *and set* $k = k+1$ *and go to step 2.*

Remark 1. We obtain a fundamental direction by solving equation (2.3), and then bend the search direction according to the constraint function or the Lagrange multiplier by equation (2.4). Moreover, we make the modified nonmonotone filter technique to avoid the Maratos effect effectively.

3. Global Convergence

In this section, we show that Algorithm 1 is global convergent to KKT points of problem (1.1). To prove, we have the following assumptions:

Assumption 1. The sequence $\{x^k\}$ and $\{x^k + p^k\}$ which are generated by Algorithm 1 are contained in a bounded set $\Omega \subseteq R^n$.

Assumption 2. The functions $f(x)$ and $g_i(x), i \in I$ are twice continuously differentiable, and their function values are bounded over $\Omega \subseteq R^n$.

Assumption 3. The vectors $\{\nabla g_i(x), i \in W_k\}$ are linearly independent for each point $x \in \Omega \subseteq R^n$.

Assumption 4. There exist $\beta_1, \beta_2 > 0$ such that for all k, $\|H_k\| \leq \beta_2$, and

$$d^T \widehat{H}_k d \geq \beta_1 \|d\|^2 \quad \forall d \in \aleph(x^k),$$

where

$$\widehat{H}_k = H_k - \sum_{i \in W_k \setminus I(x^k)} \frac{\mu_i^k}{g_i(x^k)} \nabla g_i(x^k) \nabla g_i(x^k)^T,$$

$$\aleph(x) = \{d \in R^n \mid \nabla g_i(x^k)^T d = 0, i \in W_k\}.$$

Lemma 1. *Sequences* $\{\chi_k\}$ *and* ε_k *in Algorithm 1 is changed in finite number of times.*

Proof. Suppose that χ_k and ε_k are changed infinitely many times, that is, there exists an infinite index set K such that $\chi_{k+1} = 2\chi_k$ and $\varepsilon_{k+1} = \varepsilon_k/2$ for all $k \in K$. Then we have $\{\chi_k\} \to +\infty$ and $\varepsilon_k \to 0^+$ as $k \to \infty$.

Due to the finiteness of set I, we assume $W_K = W_k$ for all $k \in K$. Assume $x^k \to \bar{x}$ as $k \to \infty$ and $k \in K$. We get from the definition of W_k that $W_K \subseteq I(\bar{x})$, since $\varepsilon_k \min\{\varphi(x^k, \lambda^{k-1,0})\} \to 0$ as $k \to \infty$. It is known in the algorithm that $\|\lambda^{k,0}\|_\infty > \chi_k$, then $\|\lambda^{k,0}\|_\infty \to \infty$. Therefore, we have sequence $\{t_k\}$, with

$$t_k = \max\{\|d^{k,0}\|, \|\lambda^{k,0}_{W_K}\|_\infty, 1\}$$

tends to infinity on K. Define $\widehat{d_k} = d^{k,0}/t_k$ and $\widehat{\lambda}^k_{W_K} = \lambda_{W_K}/t_k$ for $k \in K$.

Therefore, $\max\{\|\widehat{d^k}\|, \|\widehat{\lambda}^k_{W_K}\|_\infty\} = 1$ for all $k \in K$ large enough. Then we have a non-zero vector $(\widehat{d^k}, \widehat{\lambda}^k_{W_K}) \to (\widehat{d}, \widehat{\lambda}_{W_K})$ as $k \in K_1 \subseteq K \to \infty$, where K_1 is an infinite index set.

Since $I(x^k) \subseteq W_k$, we have that $g_i(x^k) = 0, i \in I(x^k)$, it follows from equation (2.3) that

$$\mu_i^k \nabla g_i(x^k)^T d^{k,0} = -\lambda_i^{k,0} g_i(x^k) = 0,$$

therefore $d^{k,0} \in \aleph(x^k)$ as $\mu^k > 0$, that is $\nabla g_i(x^k)^T d = 0$, so from (2.3) of Algorithm 1 and Assumption 4, we have

$$\begin{aligned}
\nabla f(x^k)^T d^{k,0} &= -(d^{k,0})^T H_k d^{k,0} - \sum_{i \in I(x^k)} \lambda_i^{k,0} \nabla g_i(x^k)^T d^{k,0} \\
&= -(d^{k,0})^T (\widehat{H}_k - \sum_{i \in W_k \setminus I(x^k)} \frac{\mu_i^k}{g_i(x^k)} \nabla g_i(x^k) \nabla g_i(x^k)^T) d^{k,0} \\
&\quad - \sum_{i \in I(x^k)} \lambda_i^{k,0} \nabla g_i(x^k)^T d^{k,0} \\
&= -(d^{k,0})^T \widehat{H}_k d^{k,0} \le -\beta_1 \|d^{k,0}\|^2.
\end{aligned}$$

Let $k \in K_1 \to \infty$ yields $\widehat{d} = 0$, and then $\widehat{\lambda}_{W_K}$ is nonzero. Besides, from equation (2.3), If $k > 1$, we have

$$H_k d^{k,0} + \nabla g_{W_K}(x^k) \lambda^{k,0}_{W_K} = -\nabla f(x^k). \tag{3.1}$$

Therefore, dividing both sides of equation (3.1) by t_k and letting $k \in K_1 \to \infty$, we have $\nabla g_{W_K}(\bar{x})\widehat{\lambda}_{W_K} = 0$

This is inconsistent with Assumption 3. The conclusion follows. □

Lemma 2. *Suppose Assumption 3 holds, given any vector $x \in X$, and any nonnegative vector $\mu^k \in R^m$ such that $\mu_i^k > 0$ if $g_i(x^k) = 0$ for all $i \in W_k$, and let $H_k \in R^{n\times n}$ be a symmetric matrix satisfying (3.1), then the matrix V_k defined by equation (2.2) is non-singular.*

Proof. Suppose (d, λ) is a solution of the following equation

$$V_k \begin{pmatrix} d \\ \lambda \end{pmatrix} = 0.$$

So we have that $\lambda_i = -(\mu_i/g_i(x^k))\nabla g_i(x^k)^T d,\ i \in W_k \backslash I(x^k)$ and $\nabla g_i(x^k)^T d = 0$, $i \in I(x^k)$ therefore,

$$d^T \left(H - \sum_{W_k \backslash I(x^k)} \frac{\mu_i}{g_i(x^k)} \nabla g_i(x^k) \nabla g_i(x^k)^T \right) d = 0$$

which implies that $d = 0$. Moreover, $\nabla g_{W_k}(x^k)\lambda = 0$, $G_{W_k}\lambda = 0$. So $\nabla g_{I(x^k)}(x^k)\lambda_{I(x^k)} = 0$ and $\lambda_{W_k \backslash I(x^k)} = 0$. Since Assumption 3 implies $\lambda_{I(x^k)} = 0$, zero is the unique solution, that is V_k is non-singular. □

Lemma 3. *Under Assumption 1-4, sequences $\{\lambda^{k,0}\}$ and $\{\mu^k\}$ in Algorithm 1 are bounded.*

Proof. It follows that $\{\chi_k\}$ has an upper bound from Lemma 1, and thus $\{\lambda^{k,0}\}$ is bounded by Algorithm 1. The boundedness of $\{\mu^k\}$ follows directly from their definitions and the boundedness of $\{\lambda^{k,0}\}$. □

Lemma 4. *Under the condition of Lemma 2, denote*

$$V_k^{-1} = \begin{pmatrix} A_k & B_k \\ C_k & D_k \end{pmatrix}$$

then $C_k = U_k B_k^T$.

Proof. Similar to the proof of Lemma 3.5 in [21]. □

Lemma 5. *If $\{x^{k_j}\}$ is a subset of iterations for which $\Gamma_s^{k_j} \geq \varepsilon, s = 1, 2, 3$ with a constant ε_1 and ε_2, independent of j, such that if $h(x^{k_j}) \leq \varepsilon_1$ and $\|d^{k,0}\|^{\omega} \leq \varepsilon_2$, then $\nabla f(x^{k_j})^T d^{k_j,1} \leq -\varepsilon/2$ for all j.*

Proof. Substituting the conclusion of Lemma 4 into equation (2.3) and (2.4) yields that

$$
\begin{aligned}
&d^{k,0} = -A_k \nabla f(x^k),\ \lambda^{k,0} = -C_k \nabla f(x^k);\\
&d^{k,1} = d^{k,0} - B_k[(1-\rho)\mu^k \|d^{k,0}\|^{\omega} + \rho\theta_k v^k],\\
&\lambda^{k,1} = \lambda^{k,0} - D_k[(1-\rho)\mu^k \|d^{k,0}\|^{\omega} + \rho\theta_k v^k].
\end{aligned} \tag{3.2}
$$

From equations (2.3) and (3.2), we have

$$
\begin{aligned}
\nabla f(x^{k_j})^T d^{k_j,1} &= \nabla f(x^{k_j})^T \{d^{k_j,0} - B_{k_j}[(1-\rho)\mu^{k_j}\|d^{k,0}\|^{\omega} + \rho\theta_k v^{k_j}]\}\\
&= -(d^{k_j,0})^T \widehat{H}_{k_j} d^{k_j,0}\\
&\quad - \nabla f(x^{k_j})^T B_{k_j}[(1-\rho)\mu^{k_j}\|d^{k,0}\|^{\omega} + \rho\theta_k v^{k_j}]\\
&= -(d^{k_j,0})^T \widehat{H}_{k_j} d^{k_j,0}\\
&\quad + (\lambda^{k_j,0})^T_{W_{k_j}} U_{k_j}^{-1}[(1-\rho)\mu^{k_j}\|d^{k,0}\|^{\omega} + \rho\theta_k v^{k_j}],
\end{aligned} \tag{3.3}
$$

Case I: $\rho = 1$.

$$
\begin{aligned}
\nabla f(x^{k_j})^T d^{k_j,1} &= -(d^{k_j,0})^T \widehat{H}_{k_j} d^{k_j,0} + (\lambda^{k_j,0})^T_{W_{k_j}} U_{k_j}^{-1} \theta_{k_j} v^{k_j}\\
&= -(d^{k_j,0})^T \widehat{H}_{k_j} d^{k_j,0} - \sum_{i\in \overline{W}_k^-} \frac{\lambda_i^{k_j,0}\theta_{k_j}}{\mu_i^{k_j}} g_i(x^{k_j}) - \sum_{i\in \overline{W}_k^+} \frac{\lambda_i^{k_j,0}\theta_{k_j}}{\mu_i^{k_j}} g_i(x^{k_j})\\
&\quad + \sum_{i\in W_k\setminus \overline{W}_k} \frac{\lambda_i^{k_j,0}\theta_{k_j}}{\mu_i^{k_j}} \min\{-g_i(x^{k_j}), \lambda_i^{k_j,0}\}\\
&= -\Gamma_1^{k_j} - \sum_{i\in \overline{W}_k^-} \frac{\lambda_i^{k_j,0}\theta_{k_j}}{\mu_i^{k_j}} g_i(x^{k_j}).
\end{aligned}
$$

Therefore, it follows with $\Gamma_1^{k_j} \geq \varepsilon$ and for all $c > 0$, such that

$$
\nabla f(x^{k_j})^T d^{k_j,1} = -\Gamma_1^{k_j} - \sum_{i\in \overline{W}_k^-} \frac{\lambda_i^{k_j,0}\theta_k}{\mu_i^{k_j}} g_i(x^{k_j}) \leq -\varepsilon + ch(x^{k_j}).
$$

If $h(x^{k_j}) \leq \varepsilon_1 = \varepsilon/(2c)$, then $\nabla f(x^{k_j})^T d^{k_j,1} \leq -\varepsilon_2 = -\varepsilon/2$.

Case II: $\rho = 0$.

$$
\begin{aligned}
\nabla f(x^{k_j})^T d^{k_j,1} &= -(d^{k_j,0})^T \widehat{H}_{k_j} d^{k_j,0} + (\lambda^{k_j,0})^T_{W_{k_j}} U_{k_j}^{-1} \mu^k \|d^{k,0}\|^{\omega}\\
&= -(d^{k_j,0})^T \widehat{H}_{k_j} d^{k_j,0} + \sum_{i\in \overline{W}_k^-} \lambda_i^{k_j,0}\|d^{k,0}\|^{\omega} + \sum_{i\in \overline{W}_k^+} \lambda_i^{k_j,0}\|d^{k,0}\|^{\omega}\\
&\quad + \sum_{i\in W_k\setminus \overline{W}_k} \lambda_i^{k_j,0}\|d^{k,0}\|^{\omega}
\end{aligned}
$$

$$= -\Gamma_2^{k_j} + \sum_{i\in\overline{W}_k^-} \lambda_i^{k_j,0}\|d^{k,0}\|^\omega.$$

There exists a scalar $c > 0$ such that $\sum_{i\in\overline{W}_k^+} \lambda_i^{k_j,0}\|d^{k,0}\|^\omega < c\|d^{k,0}\|^\omega$, combing with $\Gamma_2^{k_j} \geq \varepsilon$, we have

$$\nabla f(x^{k_j})^T d^{k_j,1} = -\Gamma_2^{k_j} + \sum_{i\in\overline{W}_k^-} \lambda_i^{k_j,0}\|d^{k,0}\|^\omega \leq -\varepsilon + c\|d^{k,0}\|^\omega.$$

If $\|d^{k,0}\|^\omega \leq \varepsilon_1 = \varepsilon/(2c)$, then $\nabla f(x^{k_j})^T d^{k_j,1} \leq -\varepsilon_2 = -\varepsilon/2$.

Case III: $0 < \rho < 1$.

$$\begin{aligned}\nabla f(x^{k_j})^T d^{k_j,1} &= -(d^{k_j,0})^T \widehat{H}_{k_j} d^{k_j,0} + (\lambda^{k_j,0})^T_{W_{k_j}} U_{k_j}^{-1}[(1-\rho)\mu^{k_j}\|d^{k,0}\|^\omega + \rho\theta_{k_j} v^{k_j}]\\
&= -(d^{k_j,0})^T \widehat{H}_{k_j} d^{k_j,0} + \sum_{i\in\overline{W}_k^-} \frac{\lambda_i^{k_j,0}}{\mu_i^{k_j}}[(1-\rho)\mu_i^{k_j}\|d^{k,0}\|^\omega - \rho\theta_{k_j} g_i(x^{k_j})]\\
&\quad + \sum_{i\in\overline{W}_k^+} \frac{\lambda_i^{k_j,0}}{\mu_i^{k_j}}[(1-\rho)\mu_i^{k_j}\|d^{k,0}\|^\omega - \rho\theta_{k_j} g_i(x^{k_j})]\\
&\quad + \sum_{i\in W_k\setminus\overline{W}_k} \frac{\lambda_i^{k_j,0}}{\mu_i^{k_j}}[(1-\rho)\mu_i^{k_j}\|d^{k,0}\|^\omega - \rho\theta_{k_j} \min\{-g_i(x^{k_j}), \lambda_i^{k_j,0}\}]\\
&= -\Gamma_3^{k_j} + \sum_{i\in\overline{W}_k^+} \lambda_i^{k_j,0}(1-\rho)\|d^{k,0}\|^\omega - \sum_{i\in\overline{W}_k^+} \frac{\lambda_i^{k_j,0}}{\mu_i^{k_j}}\rho\theta_{k_j} g_i(x^{k_j}).\end{aligned}$$

There exists a scalar $c_1 > 0$ such that $\sum_{i\in\overline{W}_k^+} \lambda_i^{k_j,0}(1-\rho)\|d^{k,0}\|^\omega < c_1\|d^{k,0}\|^\omega$ and combing with $\Gamma_3^{k_j} \geq \varepsilon$, for all $c_2 > 0$, such that

$$\begin{aligned}\nabla f(x^{k_j})^T d^{k_j,1} &\leq -\Gamma_3^{k_j} + \sum_{i\in\overline{W}_k^+} \lambda_i^{k_j,0}(1-\rho)\|d^{k,0}\|^\omega - \sum_{i\in\overline{W}_k^+} \frac{\lambda_i^{k_j,0}\theta_{k_j}}{\mu_i^{k_j}}\rho g_i(x^{k_j})\\
&\leq -\varepsilon + c_1\|d^{k,0}\|^\omega + c_2 h(x^{k_j}).\end{aligned}$$

If $\|d^{k,0}\|^\omega \leq \varepsilon_2 = \varepsilon/(4c)$, $h(x^{k_j}) \leq \varepsilon_1 = \varepsilon/(4c)$, then $\nabla f(x^{k_j})^T d^{k_j,1} \leq -\varepsilon_3 = -\varepsilon/2$. Therefore, the conclusion holds.

□

Lemma 6. *The inner loop terminates in finite iterations.*

Proof. Suppose that the inner loop run infinitely, then the filter rejects the trial point $x^k + \alpha_{k,l} d^{k,1}$ and $\lim_{l\to\infty} \alpha_{k,l} = 0$. If $h(x^k) = 0$, from the the definition of $h(x^k)$, we have

$$h(x^k+\alpha_{k,l}d^{k,1}) = \sum_{i\in I}\max\{g_i(x^k+\alpha_{k,l}d^{k,1}),0\}$$
$$\leq \sum_{i\in I}\max\{g_i(x^k)+\alpha_{k,l}\nabla g_i(x^k)^T d^{k,1}+o(\|\alpha_{k,l}d^{k,1}\|^2),0\},$$

it is obvious that $\nabla g_i(x^k)^T d^{k,1}<0$ for all $i\in W_k$, so there exists a constant γ, such that

$$h(x^k+\alpha_{k,l}d^{k,1}) \leq \sum_{i\in I}\max\{g_i(x^k)\} \leq \sum_{i\in I}(1-\gamma)\max\{g_i(x^k)\}$$
$$=(1-\gamma)h(x^k) \leq \max_{0\leq k\leq m(k)}(1-\gamma)h(x^k)$$

From Lemma 5, we have

$$f(x^k+\alpha_{k,l}d^{k,1}) \leq f(x^k)+\alpha_{k,l}\nabla f(x^k)^T d^{k,1}+O(\|\alpha_{k,l}d^{k,1})\|^2)$$
$$\leq f(x^k) \leq \max_{0\leq k\leq m(k)} f(x^k)$$

In view of equations (2.3) and (2.4), we know that $x^k+\alpha_{k,l}d^{k,1}$ must be acceptable for the filter and x^k, which is a contradiction. □

Lemma 7. *Suppose Algorithm 1 dose not terminate finitely and the assumptions hold, then* $\lim_{k\to 0} h(x^k)=0$.

Proof. Suppose there exists a subsequence $\{x^{k_j}\}$ such that $\lim_{j\to\infty} h(x^{k_j})=\varepsilon$ for some constant $\varepsilon>0$. Without loss of generality, for all j, we assume that $((1-\gamma)/2)^{1/2}\varepsilon \leq h(x^{k_j}) \leq ((1-\gamma)/2)^{-1/2}\varepsilon$.

From Lemma 6 , we know that there exists a constant K such that $x^{k_{j+1}}$ must be accepted by $\mathscr{F}_{k_{j+1}}$ for all $k_{j+1}\geq K$.

Therefore,

$$h(x^{k_{j+1}}) \leq \max_{0\leq k_j\leq m(k)}(1-\gamma)h(x^k) \text{ or } f(x^{k_{j+1}}) \leq \max_{0\leq k_j\leq m(k)} f(x^{k_j})-\gamma h(x^{k_{j+1}}).$$

So, we have

$$f(x^{k_{j+1}}) \leq \max_{0\leq k_j\leq m(k)} f(x^{k_j})-\gamma h(x^{k_{j+1}}) \leq \max_{0\leq k_j\leq m(k)} f(x^{k_j})-\gamma((1-\gamma)/2)^{1/2}\varepsilon.$$

Let $j\to\infty$ and the above inequality implies $f(x^{k_j})\to-\infty$ in contradiction to the assumption that f is bounded below. □

Lemma 8. *If Algorithm 1 dose not terminate finitely, then, there exists an index set K, such that* $\lim_{k\to\infty,k\in K}\nabla f(x^k)^T d^{k,1}=0$.

Proof. Since $\nabla f(x^k)^T d^{k,1}\leq 0$, there exists a constant $\xi_k>0$, which $\{\xi\}$ is uniformly bounded and $\liminf_k \xi_k>0$, such that$f(x^k+\alpha_k d^{k,1})-\max_{0\leq k\leq m(k)} f(x^k)\leq$

$f(x^k+\alpha_k d^{k,1})-f(x^k)\le \xi_k\alpha_k\nabla f(x^k)^T d^{k,1}\le 0$. It follows from Assumption 1 that $\{x^k\}_{k\in K}\to x^*$ for some index set K. Combing with continuity of f yields $f(x^k)\to f(x^*), k\in K, k\to+\infty$.

Therefore,

$$0=\lim_{k\to\infty,k\in K}[f(x^k+\alpha_k d^{k,1})-\max_{0\le k\le m(k)} f(x^k)] \quad \le \lim_{k\to\infty,k\in K}\xi_k\alpha_k\nabla f(x^k)^T d^{k,1}\le 0.$$

Since $\liminf_k \xi_k\alpha_k>0$, so $\lim_{k\to\infty,k\in K}\nabla f(x^k)^T d^{k,1}=0$. □

Lemma 9. *If $h(x^k)=0$ and $\nabla f(x^k)^T d^{k,1}=0$ hold, then x^k is a KKT point of problem* (1.1).

Proof. Since $h(x^k)=0$, then $g_i(x^k)\le 0, i\in I$, and then $\overline{W}_k^+=\varnothing$. And since $\nabla f(x^k)^T d^{k,1}=0$, it follows from equation (3.3),

$$\begin{aligned}\nabla f(x^k)^T d^{k,1} &= -(d^{k,0})^T\widehat{H}_k d^{k,0}-(\lambda^{k,0})^T_{W_k}U_k^{-1}[(1-\rho)\mu^k\|d^{k,0}\|^{\omega}+\rho\theta_k v^k]\\ &= -(d^{k,0})^T\widehat{H}_k d^{k,0}-\sum_{i\in\overline{W}_k^-}\frac{\lambda_i^{k,0}}{\mu^k}[(1-\rho)\mu^k\|d^{k,0}\|^{\omega}-\rho\theta_k g_i(x^k)]\\ &\quad -\sum_{i\in W_k\setminus\overline{W}_k}\frac{\lambda_i^{k,0}}{\mu^k}[(1-\rho)\mu^k\|d^{k,0}\|^{\omega}-\rho\theta_k\min\{-g_i(x^k),\lambda_i^{k,0}\}]=0.\end{aligned}$$

therefore,

$$\begin{aligned}(d^{k,0})^T\widehat{H}_k d^{k,0}=0,\\ \sum_{i\in\overline{W}_k^-}\lambda_i^{k,0}[(1-\rho)\|d^{k,0}\|^{\omega}-\frac{\rho\theta_k}{\mu^k}g_i(x^k)]=0,\\ \sum_{i\in W_k\setminus\overline{W}_k}\lambda_i^{k,0}[(1-\rho)\|d^{k,0}\|^{\omega}-\frac{\rho\theta_k}{\mu^k}\min\{-g_i(x^k),\lambda_i^{k,0}\}]=0.\end{aligned}$$

Since $I(x^k)\subseteq W_k$, hence it follows from equation (2.3) that,

$$\mu_i^k\nabla g_i(x^k)^T d^{k,0}=-g_i(x^k)\lambda_i^{k,0}=0\forall i\in I(x^k).$$

We have $\mu^k>0$ by step 8 of Algorithm 1, so from Assumption 4, we know that $d^{k,0}=0$, $\lambda_i^{k,0}g_i(x^k)=0, i\in I$, And because $(\rho\theta_k)/\mu^k>0$, so combing these with equation (2.3), the KKT condition is established. □

Theorem 1. *Suppose the Assumption 1-4 hold, and the sequence $\{(x^k,\lambda^k)\}$ which is generated by Algorithm 1 is infinite, then every accumulation point of the sequence $\{(x^k,\lambda^k)\}$ is a KKT pair of problem* (1.1).

4. NUMERICAL RESULTS

Algorithm 1 is implemented in the environment of MATLAB R2016a. We give our preliminary results on the some test problems from Hock and Schittkowski [22], and compare with the algorithm in [21] and that in Matlab.

The results are summarized in Table 1. The details about the implementation are described as follows.

(a) The parameter values as chosen as follows: $\gamma = 10^{-4}$, $h_{\max} = 10^6$, $\nu = 0.5$, $\rho = 0.5$, $\chi_1 = 10$, $\varphi_{\max} = 0.5$, $\varepsilon = 5$, $\omega = 2.5$.

(b) The meanings of some notations in Table 1 are described as follows:
- No: the problem number given in Hock and Schittkowski [22];
- n: the number of variables;
- m: the number of constraints;
- NIT: the number of iterations;
- NF: the number of evaluations for $f(x)$;
- NG: the number of evaluations for $g(x)$.

(c) A stops if $|\nabla f(x^k)^T d^{k,1}|/(\|f(x^k)\|+1) \le 10^{-6}$ and $h(x^k) \le 10^{-6}$.

(d) H_k is updated by the damped BFGS formula.

Table 1 Numerical results

			Algorithm 1			Algorithm in [21]			Matlab
No	n	m	NIT	NF	NG	NIT	NF	NG	NIT - NF
$HS1$	2	1	7	13	9	23	48	54	$27-95$
$HS3$	2	1	5	9	9	13	19	23	$4-15$
$HS4$	2	2	5	9	9	2	4	2	$2-6$
$HS5$	2	4	12	75	75	10	12	31	$10-34$
$HS6$	2	1	3	5	3	7	16	17	$6-28$
$HS11$	2	1	3	5	3	23	35	35	$7-25$
$HS12$	2	1	16	54	53	10	12	10	$8-25$
$HS15$	2	3	8	37	33	14	28	16	$3-9$
$HS16$	2	5	7	57	53	10	27	11	$4-12$
$HS17$	2	5	8	15	9	4	7	4	$14-43$
$HS18$	2	6	9	17	12	7	14	9	$9-28$
$HS21$	2	5	7	13	7	7	11	11	$3-9$
$HS22$	2	2	8	15	10	21	45	49	$4-15$
$HS26$	3	1	6	11	6	10	36	41	$6-27$
$HS27$	3	1	6	11	8	13	46	17	$44-303$
$HS28$	3	1	8	15	9	17	33	40	$7-29$
$HS30$	3	7	9	18	18	6	11	20	$11-44$
$HS33$	3	6	5	9	9	3	13	22	$5-20$
$HS35$	3	4	8	15	15	14	43	59	$6-24$
$HS43$	4	3	6	11	6	12	22	17	$12-63$

$HS46$ 5 2	13	32	25	13	57	83	$12-75$
$HS48$ 5 2	10	21	19	14	24	14	$8-50$
$HS49$ 5 2	27	69	51	103	198	213	$19-117$

In the table, we first give the result of algorithm 1. For comparison, we have included the corresponding results obtained by Wang et al. [21] and the optimization code in Matlab (column 'MATLAB'). Compared with [21] and the code in Matlab, algorithm 1 has a relatively small iteration number both in NIT and in NF/NG. Therefore, our algorithm is effective and has numerical promising.

ACKNOWLEDGEMENT

The author thanks the anonymous reviewers for their careful reading of the manuscript and their many insightful comments and suggestions.

REFERENCES

[1] J. L. L. abd J. Lv and J. B. Jian, "A globally and superlinearly convergent primal-dual interior point method for general constrained optimization." *Numer. Math. Theor. Meth. Appl.*, vol. 8, no. 3, pp. 313–335, 2015, doi: 10.4208/nmtma.2015.m1338.

[2] S. Bakhtiari and A. L. Tits, "A simple primal-dual feasible interior-point method for nonlinear programming with monotone descent." *Comput. Optim. Appl.*, vol. 25, no. 1-3, pp. 17–38, 2003, doi: 10.1023/A:1022944802542.

[3] F. Facchinei, "Robust recursive quadratic programming algorithm model with global and superlinear convergence properties." *J. Optim. Theory Appl.*, vol. 92, no. 3, pp. 543–579, 1997, doi: 10.1023/A:1022655423083.

[4] A. Fischer, "A Solution of the monotone complementarity problem with locally Lipschitzian functions." *Math. Program.*, vol. 76, no. 3, pp. 513–532, 1997, doi: 10.1007/bf02614396.

[5] R. Fletcher and S. Leyffer, "Nonlinear programming without a penalty function." *Math. Program.*, vol. 91, no. 2, pp. 239–269, 2002, doi: 10.1007/s101070100244.

[6] R. Fletcher, S. Leyffer, and P. L. Toint, "On the global convergence of a filter SQP algorithm." *SIAM J. Optim.*, vol. 13, no. 1, pp. 44–59, 2002, doi: 10.1137/s105262340038081x.

[7] Z. Y. Gao, G. P. He, and F. Wu, "Sequential systems of linear equations algorithm for nonlinear optimization problems with general constraints." *J. Optim. Theory Appl.*, vol. 95, no. 2, pp. 371–397, 1997, doi: 10.1023/A:1022639306130.

[8] P. E. Gill and W. Murray, "Saunders MA SNOPT: an SQP algorithm for large-scale constrained optimization." *SIAM Rev.*, vol. 47, no. 1, pp. 99–131, 2005, doi: 10.2307/20453604.

[9] C. Gu and D. Zhu, "A non-monotone line search multidimensional filter-SQP method for general nonlinear programming." *Numer. Algorithms*, vol. 56, no. 4, pp. 537–559, 2011, doi: 10.1007/s11075-010-9403-z.

[10] S. P. Han, "Superlinearly convergent variable metric algorithms for general nonlinear programming problems." *Math. Program.*, vol. 11, no. 1, pp. 263–282, 1976, doi: 10.1007/bf01580395.

[11] J. B. Jian, H. Q. Pan, C. M. Tang, and J. L. Li, "A strongly sub-feasible primal-dual quasi interior-point algorithm for nonlinear inequality constrained optimization." *Appl. Math. Comput.*, vol. 266, pp. 560–578, 2015, doi: 10.1016/j.amc.2015.05.091.

[12] J. L. Li, R. S. Huang, and J. B. Jian, "A superlinearly convergent QP-free algorithm for mathematical programs with equilibrium constraints." *Appl. Math. Comput.*, vol. 269, pp. 885–903, 2015, doi: 10.1016/j.amc.2015.07.081.

[13] J. L. Li and Z. P. Yang, "A QP-free algorithm without a penalty function or a filter for nonlinear general-constrained optimization." *Appl. Math. Comput.*, vol. 316, pp. 52–72, 2018, doi: 10.1016/j.amc.2017.08.013.

[14] S. Lucidi, "New results on a continuously differentiable exact penalty function." *SIAM J. Optim.*, vol. 2, no. 4, pp. 558–574, 1992, doi: 10.1137/0802027.

[15] E. R. Panier, A. L. Tits, and J. N. Herskovits, "A QP-free, globally convergent, locally superlinearly convergent algorithm for inequality constrained optimization." *SIAM J. Control Optim.*, vol. 26, no. 4, pp. 788–811, 1988, doi: 10.1137/0326046.

[16] H. D. Qi and L. Q. Qi, "A new QP-free, globally convergent, locally superlinearly convergent algorithm for inequality constrained optimization." *SIAM J. Optim.*, vol. 11, no. 1, pp. 113–132, 2000, doi: 10.1137/S1052623499353935.

[17] L. Qi and Y. F. Yang, "Globally and superlinearly convergent qp-free algorithm for nonlinear constrained optimization." *J. Optim. Theory Appl.*, vol. 113, no. 2, pp. 297–323, 2002, doi: 10.1023/a:1014882909302.

[18] C. G. Shen, S. Leyffer, and R. Fletcher, "A nonmonotone filter method for nonlinear optimization." *Comput. Optim. Appl.*, vol. 52, no. 3, pp. 583–607, 2012, doi: 10.1007/s10589-011-9430-2.

[19] S. Ulbrich, "On the superlinear local convergence of a filter-SQP method." *Math. Program.*, vol. 100, no. 1, pp. 217–245, 2004, doi: 10.1007/s10107-003-0491-6.

[20] A. Wächter and L. T. Biegler, "Line search filter methods for nonlinear programming: Motivation and global convergence." *SIAM J. Optim.*, vol. 16, no. 1, pp. 32–48, 2005, doi: 10.1137/s1052623403426544.

[21] H. Wang, F. Liu, C. Gu, and D. G. Pu, "An infeasible active-set QP-free algorithm for general nonlinear programming." *Comput. Math.*, vol. 94, no. 5, pp. 884–901, 2017, doi: 10.1080/00207160.2016.1154947.

[22] W.Hock and K.Schittkowski, *Test examples for nonlinear programming codes, Lecture Notes in Economics and Mathematics System.* New York: Springer Verlag, 1981.

Authors' addresses

Ke Su
College of Mathematics and Information Science, Hebei University, Key Laboratory of Machine Learning and Computational Intelligence, Baoding, 071002, P.R.China
E-mail address: pigeonsk@163.com

Lele Ren
(**Corresponding author**) College of Mathematics and Information Science, Hebei University, Key Laboratory of Machine Learning and Computational Intelligence, Baoding, 071002, P.R.China
E-mail address: rlle6171878@163.com