http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (17): 4082-4088, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.4082.4088

A Nonmonotone Algorithm of Moving Asymptotes for Solving Unconstrained Optimization Problems

Ping Hu and Zongyao Wang
Faculty of Mathematics and Physics, Huaiyin Institute of Technology, 223003, Huaian, China

Abstract: In this study, we aim to put forward a novel nonmonotone algorithm of moving asymptotes for solving n-variate unconstrained optimization problems. The algorithm first generates n separable subproblems by virtue of the moving asymptotes function in each iteration to determine the descent search direction and then obtain the step by new nonmonotone line search techniques. The global convergence of the proposed algorithm is established in this study. In addition, we give some numerical tests, from which it is indicates that the new algorithm is effective in solving multi-peak or large-scale optimization problems.

Key words: Unconstrained optimization problems, algorithm of moving asymptotes, nonmonotone line search, global convergence

INTRODUCTION

Consider the unconstrained optimization problem:

$$\min_{x\in R^n} f(x) \tag{1}$$

where, f(x) and the gradient $g = \nabla f(x)$ are solvable. There are numerous iterative approaches (Yuan, 2008) used to solve Eq. 1, including the Method of Moving Asymptotes (MMA) is effective in solving large-scale optimization problems (Svanberg, 1987; Wang and Ni, 2008).

In 2012, Ping Hu introduced (Hu *et al*., 2012) a new MMA by structuring the MMA subproblem:

$$\min m(x^k, s) = f(x^k) + \Phi(s) \text{s.t.,} \quad \alpha_i \le s_i \le \beta_i, i = 1, 2, \cdots, n \tag{2}$$

where $\Phi(s)$ denotes the Moving Asymptotes (MA) function of the form:

$$\Phi(s) = \sum_{i=1}^{n} \phi_i(s_i)$$

With:

$$\phi(s_i) = \begin{cases} \kappa_i s_i + \dfrac{(a_i^k)^2}{a_i^k - s_i}(g^k - \kappa_i) - a_i^k (g_i^k - \kappa_i), i \in I_+ \\ \kappa_i s_i - \dfrac{(b_i^k)^2}{b_i^k + s_i}(g^k - \kappa_i) + b_i^k (g_i^k - \kappa_i), i \in I\_ \end{cases} \tag{3}$$

where, a^k_i and $-b^k_i$ are the upper and lower bounds of the asymptotes and therefore $-b^k_i < s^k_i < a^k_i$:

$$(g_1^k, \cdots, g_n^k)^T = \nabla f(x^k) = g^k \tag{4}$$

$$a_i^k \ge \tau\|g^k\| + \eta_i,\ b_i^k \ge \tau\|g^k\| + \eta_i$$

$$\tau>0,\ \eta_i>0,\ i = 1,..., n \tag{5}$$

$$I_+ = \{i: g^k_i \ge 0\},\ I\_ = \{i: g^k_i<0\} \tag{6}$$

$$-\alpha_i = \min\{c_1, c_2 b^k_i\},\ \beta_i = \min\{c_1, c_2 a^k_i\} \tag{7}$$

With $c_i>0$ and $0<c_2<1$. The strategy of selecting the parameters involved in the subproblem Eq. 2 is presented as follows.

Put $\eta_i = K (K>1)$ if $g^k_i = 0$ and:

$$\eta_i = \frac{1 - c_2}{c_1^2 \max(|g_i^k|, 0.01)} \quad \text{if } g^k_i \ne 0.$$

Take:

$$\kappa_i = \begin{cases} g_i^k - \sigma_i \mu, i \in I_+ \\ g_i^k + \sigma_i \mu, i \in I\_ \end{cases}$$

Where:

$$\sigma_i = \frac{1 - c_2}{c_1^2(\tau\|g^k\| + \eta_i)} \text{ and } \mu>1.$$

Corresponding Author: Ping Hu, Faculty of Mathematics and Physics, Huaiyin Institute of Technology,Huaian, 223003, Jiangsu, China

In this MMA, the Eq. 1 can be reduced to be minΦ(s) by approximating f(x) via the function m(x^k, s) and therefore, as a consequence, it is not difficult to obtain the next search direction. Then, one can apply the line search techniques to determine the length a^k and thus get the next iterative point as:

$$x^{k+1} = x^k+\alpha^k s^k \tag{8}$$

In Eq. 8, the length α^k can be determined by some traditional rules such as Arimijo rule, Goldstein rule, Wolfe rule and so fort (Yuan, 2008). Each of these criteria requires the precondition that the functions sequence $\{f(x^k)\}$ are monotonically decreasing that is:

$$f(x^{k+1})\le f(x^k) \tag{9}$$

Hence, the corresponding approach is called to be nonmonotone line search technique.

Recent researches show that the convergence rate of such monotone line search techniques reduce considerably when the iteration locates in a narrow curved valley (Sun *et al*., 2002; Yu and Pu, 2008; Grippo *et al*., 1986). To overcome this problem, Grippo introduced a highly innovative method called the nonmonotone line search technique which does not require the Eq. 9. Numerical tests illus trate that this method is effective. Subsequently, it is developed by many authors (Xao *et al*., 2009; Sun and Zhou, 2007; Zhou and Sun, 2008).

In 2002, Sun put forward the so-called nonmonotone F-rule which is a general nonmonotone line search technique with the nonmontone Arimijo rule, the nonmonotone Goldstein rule and the nonmonotone Wolfe rule as its special cases. Nonmonotone F-rule requires $\alpha^k\ge 0$ such that:

$$f\left(x^k+\alpha^k s^k\right)\le \max_{0\le j\le m(k)}\left\{f\left(x^{k-j}\right)\right\}-\sigma\left(t^k\right) \tag{10}$$

where, {m(k)} is an integer sequence satisfying the following conditions:

$$m(0) = 0,\ 0\le m(k)\le \min\{m(k-1)+1, M\}$$

for some positive integer M, $t^k = -(g^k)^T S^k/\|S^k\|$, $\sigma:[0, +\infty]\to[0, +\infty]$ is a forcing function which is defined as follows: For any nonnegative sequence $\{t^k\}\subset[0, +\infty]$:

$$\lim_{k\to\infty} t^k = 0$$

holds if:

$$\lim_{k\to\infty} \sigma(t^k) = 0$$

In 2008, Yu and Pu (2008) put forward an improved nonmonotone F-rule which requires $\alpha^k\ge 0$ such that:

$$\begin{aligned} f(x^{k+1}) &= f(x^k+\alpha^k s^k) \\ &\le \max\{f(x^k), \sum_{r=0}^{m(k)-1}\lambda_{kr} f(x^{k-r})\}-\sigma(t^k) \end{aligned} \tag{11}$$

Where:

$$\sum_{r=0}^{m(k)-1}\lambda_{kr}=1$$

All other parameters are the same as that of Eq. 10.

It is seen from Eq. 10-11 that the inequality $f(x^{k+1})>f(x^k)$ may hold for some k and therefore it can play a nonmonotone search role in the above rules. However, it can also be easily concluded that $f(x^{k+1})\le f(x^0)-s(t^k)\le f(x^0)$ which means $f(x^k)\le f(x^0)$. Arimijo rule possesses the similar disadvantage. That is, x^k will be trapped and cannot escape when x^0 locates in a valley. In this case, it is difficult to search other better points.

Motivated by the above analysis, we will make an effort to improve MMA of (Hu *et al*., 2012) and the nonmonotone F-rule of (Yu and Pu, 2008) and put forward a novel algorithm by combining the MMA and the nonmonotone line search techniques. The new method will be used to solve the Eq. 1 and presented in section 2. In section 3, we will prove the convergence of the proposed algorithm based on the works (Hu *et al*., 2012; Sun *et al*., 2002; Yu and Pu, 2008; Hu and Ni, 2010). Finally, some numerical tests will be made to illustrate the main results. The rest of this study is organized as above.

NONMONOTONE ALGORITHM OF MOVING ASYMPTOTES

Algorithm of moving asymptotes for the search direction S^k: We let a_i^k = b_i^k in Eq. 5 in order to calculate and programme simply. Then we can structure the MMA subproblem:

$$\begin{aligned} &\min\ m(x^k,s)=f(x^k)+\Phi(s) \\ &\text{s.t.}\quad \alpha_i\le s_i\le\beta_i, i=1,2,\cdots,n \end{aligned} \tag{12}$$

where, $\beta_i = \min\{c_1, c_2 a^k_i\}$, other parameters are shown in the first quarter.

From Eq. 3, we can be easily concluded that:

$$\phi_i = 0$$

$$\phi_i'(s_i) = \begin{cases} \kappa_i + \dfrac{(a_i^k)^2}{(a_i^k - s_i)^2}(g_i^k - \kappa_i),\ i \in I_+ \\ \kappa_i + \dfrac{(a_i^k)^2}{(a_i^k + s_i)^2}(g_i^k - \kappa_i),\ i \in I_- \end{cases} \tag{13}$$

$$\phi_i''(s_i) = \begin{cases} \dfrac{2(a_i^k)^2}{(a_i^k - s_i)^3}(g_i^k - \kappa_i), i \in I_+ \\ -\dfrac{2(a_i^k)^2}{(a_i^k + s_i)^3}(g_i^k - \kappa_i), i \in I_- \end{cases}$$

According to the selection strategy of parameter κ_i, we know if $i \in I_+$, then $\kappa_i < g^k_i$, if $i \in I_-$ then $\kappa_i > g^k_i$. And because $-a^k_i < s^k_i < a^k_i$, therefore :

$$\phi_i''(s_i) > 0 \tag{14}$$

Thus, it can be ensure that the subproblem (12) is strictly convex, so there is only one global optimal solution S^k. On the other hand, we can obtain through directly computing:

$$\phi_i'(0) = g^k_i,\ i = 1,\ldots, n \tag{15}$$

Therefore, the subproblem is the first approximation of x^k in the original problem.

Because MA function $\phi(s)$ is separable, Eq. 12 becomes n separable dimensional constraint subproblems:

$$\begin{aligned} \min \quad & \phi_i(s_i) \\ \text{s.t.} \quad & -\beta_i \le s_i \le \beta_i,\ i = 1, \cdots, n \end{aligned} \tag{16}$$

Lemma 1: The optimal solution of n Eq. 16 is:

$$S^k_i = -\lambda_i g^k_i,\ i = 1,\ldots, n \tag{17}$$

Where:

$$\lambda_i = \begin{cases} \dfrac{\beta_i}{g_i^k}, & \text{if } \phi_i'(-\beta_i) \ge 0 \\ \dfrac{\beta_i}{-g_i^k}, & \text{if } \phi_i'(\beta_i) \le 0 \\ \dfrac{a_i^k\delta_i - a_i^k\varepsilon_i}{|\kappa_i| + \sqrt{\kappa_i(\kappa_i - g_i^k)}}, & \text{if } \phi_i'(-\beta_i) > 0 \text{ and } \phi_i'(\beta_i) < 0 \end{cases} \tag{18}$$

$$\delta_i = \begin{cases} 1, & g_i^k \ge 0 \\ 0, & g_i^k < 0 \end{cases},\ \varepsilon_i = \begin{cases} 0, & g_i^k \ge 0 \\ -1, & g_i^k < 0 \end{cases} \tag{19}$$

Proof: From Eq. 14 $\phi_i''(s_i)>0$, we know that $\phi_i(s_i)$ is a strictly convex function in when $-b_i \le s_i \le b_i$, So there is only one optimal solution of Eq. 16.

If $\phi_i'(-\beta_i) \ge 0$, then $\phi_i'(-s_i) \ge 0$, when $-\beta_i \le s_i \le \beta_i$. So, the Minimum value of Eq. 16 is to get at $s_i = -\beta_i$. If $\phi_i'(-\beta_i)<0$ and $\phi_i'(-\beta_i)<0$, we know there is roots of $\phi_i'(S_i) = 0$ when $-\beta_i \le s_i \le \beta_i$.

From Eq. 13, we can easily obtain the root S_i. Therefore, we have Eq. 18 when S_i is substituted into Eq. 17.

We can obtain the numerical solution directly for Eq. 12 according to lemma 1. It shows that the solving process is simplified through MA function approximation. Therefore, this algorithm can be used to solve large-scale problems.

Lemma 2: For any initial point x^0, the level set is defined as $\Omega = \left\{x \in R^n \,\middle|\, f(x) \le \lambda^\gamma \left|f(x^0)\right|, \text{constant } \lambda \ge 1, \gamma > 0\right\}$.

Assume f(x) is bounded and differentiable on Ω and $\nabla f(x)$ satisfies Lipschitz conditions on Ω. Let $S^k = (s^k_1,\ldots, s^k_n)^T$, where $s^k_i (i = 1,\ldots, n)$ is defined as Eq. 17. If $\|g^k\| \ge \varepsilon > 0$, then there is a constant $\varepsilon_1 > 0$ which satisfies:

$$\frac{-(g^k)^T s^k}{\|s^k\|\|g^k\|} \ge \varepsilon_1 \tag{20}$$

Proof: The proof can be completed according to lemma 4 in reference (Hu *et al*., 2012).

From Eq. 20, we have $(g^k)^T S^k \le 0$ which explains that the optimal solution S^k of subproblem 12 is a Drop direction of primitive function f(x) at x^k. From Eq. 20, we have:

$$\left|\frac{-(g^k)^T s^k}{\|s^k\|}\right| \ge \varepsilon_1 \|g^k\| = \sigma(\|g^k\|), \quad k = 0, 1, 2, \cdots, \tag{21}$$

Nonmonotone combination rule for line searches of solving search step: In order to prove simply, we let:

$$f_k = f(x^k),\ \sigma_k = \sigma(t^k),\ f_{l(k)} = \sum_{r=0}^{m(k)} \lambda_{kr} f(x^{k-r})$$

Where:

$$\sum_{r=0}^{m(k)} \lambda_{kr} = 1,\ \lambda_{kr} \ge \beta > 0, m(k) = \min[k, M-1]$$

Integer $M \ge 1$.

Nonmonotone combination rule for line searches is that search step $a^k \ge 0$ is bounded and satisfies:

$$f_{k+1} = f(x^k + \alpha^k S^k) \le \mu_k f_{l(k)} - \sigma(t^k) \tag{22}$$

Where:

$$\mu_k = \begin{cases} \lambda^{h_k}, & f_{l(k)} > 0 \\ \lambda^{-h_k}, & f_{l(k)} \le 0 \end{cases}, \lambda \ge 1,\ h_k \ge 0$$

and:

$$\sum_{k=0}^{\infty} h_k = \gamma \text{ (γ is a limited constant)} \tag{23}$$

$\sigma(t^k)$ is the F-function and $t^k = -(g^k)^T S^k/||S^k|| \geq 0$.

In Eq. 22, when M = 1, $\lambda = 1$, we have $f(x^k+\alpha^k s^k) \leq f(x^k)-\sigma(t^k)$, where $\sigma(t^k) = \alpha_k \rho(-(g^k)^T S_k) = \alpha_k t^k \rho/||S^k||$ So, combination rule becomes Arimijo rule.

If

$$f(x^{k-p}) = \max_{0 \leq r \leq m(k)} \{f(x^{k-r})\},$$

let $\lambda = 1$, $\lambda_{kp} = 1$, $\lambda_{kr} = 0$ ($r \neq p$) then Eq. 22 becomes:

$$f(x^k + \alpha^k s^k) \leq \max_{0 \leq j \leq m(k)} \{f(x^{k-j})\} - \sigma(t^k)$$

this is the nonmonotone algorithm put forward by Grippo *et al*. (1986).

Nonmonotone algorithm of moving asymptotes for solving Eq. 1 based on MA function.

Algorithm 1:

Step 1: Given initial iteration point x^0, let k = 0. According to Parameter selection strategy, we select c_1, τ, η_i, K, $\varepsilon>0$, $0<c_2<1$, $\lambda \geq 1$, integer $M \geq 1$ and:

$$\sum_{r=0}^{m(k)} \lambda_{kr} = 1,\ h_k \geq 0,\ \sum_{k=0}^{\infty} h_k \leq \gamma \text{ (γ is a limited constant)}$$

Step 2: Compute g^k according to Eq. 4. If $||g^k||<\varepsilon$ and the algorithm stops. Otherwise, the algorithm turn to step 3

Step 3: Compute search direction. Update the boundary a^k_i of moving asymptotes to satisfy Eq. 5, Determine β_i via., Eq. 7 to yield the sub-problem Eq. 16 and then get the solution, saying S^k, based on Eq. 17

Step 4: Ascertain line search step α^k

Step 4.1: Select initial value. Given $\rho \in (0, 1/2)$, let $\alpha = 1$

Step 4.2: Check conditions. Let m(k) = min{k, M-1}, μ_k satisfies Eq. 23. If $f(x^k+\alpha s^k) \leq \mu_k f_{l(k)}+,\rho\alpha(g^k)^T s^k$, then $\alpha^k = \alpha$, turn to step 5. Otherwise, turn to step 4.3

Step 4.3: Shorten step. Let $\alpha = w\alpha$, $w \in [0.01, 0.99]$, turn to step 4.2

Step 5: Let $x^{k+1} = x^k+\alpha^k s^k$, k = k+1, turn to step 2

GLOBAL CONVERGENCE

Here, we will prove the global convergence of nonmonotone algorithm of moving asymptotes.

Lemma 3: μ_k satisfies Eq. 23, obviously:

$$\sum_{n=0}^{k} h_n$$

is does not reduce the sequence about k and:

$$\lambda^{-\gamma} \leq \lambda^{-\sum_{n=0}^{k} h_n} \leq \lambda^{-h_k} \leq \mu_k \leq \lambda^{h_k} \leq \lambda^{\sum_{n=0}^{k} h_n} \leq \lambda^{\gamma},\ k = 0, 1, 2, \cdots$$

Theorem 1: If a_k satisfies the combination rule Eq. 22, then:

$$f_{k+1} \leq |f_0| \lambda^{\sum_{n=0}^{k} h_n} - \lambda^{-\sum_{n=0}^{k} h_n} \beta \sum_{i=0}^{k-1} \sigma_i - \sigma_k,\ k = 1, 2, \cdots \tag{24}$$

Proof: We will use mathematical induction to prove this theorem:

Notice: $0 < \lambda^{-\sum_{n=0}^{k} h_n} \leq 1,\ 0 < \beta \leq \lambda_{kr} \leq 1$

For the case of k = 1.

If M = 1, then m(k) = 0, from Eq. 22, we have $f_2 \leq \mu_1 \lambda_{10} f_1 - \sigma_1 = \mu_1 f_1 - \sigma_1$.

Because $f_1 \leq \mu_0 \lambda_{00} f_0 - \sigma_0 = \mu_0 f_0 - \sigma_0$, we have:

$$f_2 \leq \mu_1 (\mu_0 f_0 - \sigma_0) - \sigma_1 = \mu_1 \mu_0 f_0 - \mu_1 \sigma_0 - \sigma_1 \leq |f_0| \lambda^{\sum_{n=0}^{1} h_n} - \lambda^{-\sum_{n=0}^{1} h_n} \beta \sigma_0 - \sigma_1$$

If M>1, from Eq. 22, we have:

$$\begin{aligned} f_2 &\leq \mu_1 \{\lambda_{10} f_1 + \lambda_{11} f_0\} - \sigma_1 \leq \mu_1 \{\lambda_{10} (\mu_0 f_0 - \sigma_0) + \lambda_{11} f_0\} \\ &-\sigma_1 \leq \lambda^{h_0} \lambda^{h_1} (\lambda_{10} + \lambda_{11}) |f_0| - \mu_1 \lambda_{10} \sigma_0 - \sigma_1 \\ &\leq |f_0| \lambda^{\sum_{n=0}^{1} h_n} - \lambda^{-\sum_{n=0}^{1} h_n} \beta \sigma_0 - \sigma_1 \end{aligned}$$

That is Eq. 24 holds for k = 1.

Let us now assume

$$f_k \leq |f_0| \lambda^{\sum_{n=0}^{k-1} h_n} - \lambda^{-\sum_{n=0}^{k-1} h_n} \beta \sum_{i=0}^{k-2} \sigma_i - \sigma_{k-1}$$

Let us now assume:

$$f_k \le |f_0|\lambda^{\sum_{n=0}^{k-1} h_n} - \lambda^{-\sum_{n=0}^{k-1} h_n}\beta\sum_{i=0}^{k-2}\sigma_i - \sigma_{k-1}$$

It follows that:

$$f_{k+1} \le \mu_k \sum_{r=0}^{m(k)} \lambda_{kr} f_{k-r} - \sigma_k$$

$$\le \lambda^{h_k}\sum_{r=0}^{m(k)}\lambda_{kr}\left(|f_0|\lambda^{\sum_{n=0}^{k-r-1} h_n} - \lambda^{-\sum_{n=0}^{k-r-1} h_n}\beta\sum_{i=0}^{k-r-2}\sigma_i - \sigma_{k-r-1}\right)$$

$$-\sigma_k \le \lambda^{h_k}\sum_{r=0}^{m(k)}\lambda_{kr}$$

$$\left(|f_0|\lambda^{\sum_{n=0}^{k-1} h_n} - \lambda^{-\sum_{n=0}^{k-1} h_n}\beta\sum_{i=0}^{k-m(k)-2}\sigma_i - \sigma_{k-r-1}\right) - \sigma_k$$

$$\le |f_0|\lambda^{\sum_{n=0}^{k} h_n}\sum_{r=0}^{m(k)}\lambda_{kr} - \lambda^{-\sum_{n=0}^{k} h_n}\beta\left(\sum_{i=0}^{k-m(k)-2}\sigma_i\right)\sum_{r=0}^{m(k)}\lambda_{kr}$$
$$-\lambda^{h_k}\sum_{r=0}^{m(k)}\lambda_{kr}\sigma_{k-r-1} - \sigma_k$$

$$= |f_0|\lambda^{\sum_{n=0}^{k} h_n} - \lambda^{-\sum_{n=0}^{k} h_n}\beta\left(\sum_{i=0}^{k-m(k)-2}\sigma_i\right) - \lambda^{h_k}\sum_{i=0}^{m(k)}\lambda_{kr}\sigma_{k-r-1} - \sigma_k$$

$$\le |f_0|\lambda^{\sum_{n=0}^{k} h_n} - \lambda^{-\sum_{n=0}^{k} h_n}\beta\left(\sum_{i=0}^{k-m(k)-2}\sigma_i\right) - \lambda^{-\sum_{n=0}^{k} h_n}\beta\sum_{i=0}^{m(k)}\sigma_{k-r-1} - \sigma_k$$

$$\le |f_0|\lambda^{\sum_{n=0}^{k} h_n} - \lambda^{-\sum_{n=0}^{k} h_n}\beta\left(\sum_{i=0}^{k-m(k)-2}\sigma_i - \sum_{i=k-m(k)-1}^{k-1}\sigma_i\right) - \sigma_k$$

$$= |f_0|\lambda^{\sum_{n=0}^{k} h_n} - \lambda^{-\sum_{n=0}^{k} h_n}\beta\sum_{i=0}^{k-1}\sigma_i - \sigma_k$$

This means that Eq. 24 holds.

Next, we will prove the global convergence of the nonmonotone algorithm of moving asymptotes.

Theorem 2: Assume lemma 2 hold, the search direction S^k satisfies the algorithm of moving asymptotes, line search length α^k satisfies nonmonotone rule Eq. 22, $\{x^k\}$ is generate sequence, then $\{x^k\}\epsilon\Omega$ and $\lim_{k\to\infty}\|g^k\| = 0$.

Proof: From theorem 1 we have:

$$f_{k+1} \le |f_0|\lambda^{\sum_{n=0}^{k} h_n} - \lambda^{-\sum_{n=0}^{k} h_n}\beta\sum_{i=0}^{k-1}\sigma_i - \sigma_k \le \lambda^{\gamma}|f_0|$$

according to the definition of Ω, we know that $\{x^k\}\epsilon\Omega$.

From Eq. 24 we have:

$$f_{k+1} \le |f_0|\lambda^{\sum_{n=0}^{k} h_n} - \lambda^{-\sum_{n=0}^{k} h_n}\beta\sum_{i=0}^{k-1}\sigma_i - \sigma_k \le |f_0|\lambda^{\sum_{n=0}^{k} h_n} - \lambda^{-\sum_{n=0}^{k} h_n}\beta\sum_{i=0}^{k}\sigma_i$$

That is:

$$0 \le \lambda^{-\sum_{n=0}^{k} h_n}\beta\sum_{i=0}^{k}\sigma_i \le |f_0|\lambda^{\sum_{n=0}^{k} h_n} - f_{k+1} \tag{25}$$

From assumption 1 we know that f_{k+1} are bounded in the Ω, so according to Eq. 25, when $k\to\infty$, we have:

$$0 \le \lambda^{-\gamma}\beta\sum_{i=0}^{\infty}\sigma_i < \infty$$

then:

$$\lim_{k\to\infty}\sigma(t^k) = 0$$

From Definition of F-function, we have:

$$\lim_{k\to\infty} t^k = \lim_{k\to\infty}(-(g^k)^T s^k / \|s^k\|) = 0$$

according to the conclusion Eq. 21 of Lemma 2,we have:

$$\lim_{k\to\infty}\sigma(\|g^k\|) = 0$$

Therefore:

$$\lim_{k\to\infty}\sigma(\|g^k\|) = 0$$

because of the Definition of F-function.

NUMERICAL EXPERIMENTS

Here, we use the nonmonotone MMA algorithm to test a few standard test problems. Algorithm is operated in PC whose operating system is windows 7 version and the compiler is C^{++} version.

In algorithm 1, parameter are $c_1 = 2$, $c_2 = 0.5$, $\tau = 300$, $K = 7$, $\mu = 40$. Other parameters are determined according to selection strategy. In nonmonotone line search, select:

$$\mu_k = \lambda^{\frac{1}{(k+1)^2}\mathrm{sign}(f_{l(k)})}$$

$\rho = 0.003$, $w = 0.4$, Conditions for the termination is $\|g^k\|_\infty<\varepsilon$, $\varepsilon = 10^{-3}$. If the iteration of the algorithm outnumber 2000, then the algorithm will be forced to end.

We use M and l for debugging variables. When M = 1 and l = 1, algorithm 1 become MMA algorithm of monotone line search. When l = 1 and M>1, algorithm 1 become nonmonotone MMA algorithm based on F-rule. In the tables below, n is the dimension of test function, n_g is search iterations, $f(X^*)$ is functional value of approximate solution X^*. The following two problems are selected to be tested.

Function 1: Multimodal function:

$$f(X)=2x_1^2-1.05x_1^4+\frac{1}{6}x_1^6-x_1x_2+x_2^2,\ X^0=(1.5,0.5)^T$$

As is seen from Fig. 1, there are 3 relative minimum points A,B,C and 1 global minimum point A, $f_{min}=0$.

As is seen from Table 1, when M = 1 and l = 1, it is monotone search and does not obtain the point A of global optimal value. Keep M remain unchanged. Let λ increase gradually, then $\mu_k(\lambda)$ increase along with λ. When λ increases to a certain amount, $f(X^{k+1})$ which satisfies $f(X^{k+1})\le\mu_k f_{l(k)}$ can go beyond $f(X^0)$. Therefore, iteration point X^k can jump out the lows B and search to the global minimum point A.

As is seen from this example, the result of monotone line search method is not necessarily the globally optimal. Nonmonotone algorithm based on F-rule is to search the next iteration point by combining the function values of the first m(k) iteration points. Note that the associated result is not necessarily optimal globally when the combination is small enough. Algorithm 1 can enlarge the function combination of the several points that appear before the current iteration to achieve relaxation effect. Therefore, the globally optimal solution can be obtained by virtue of very few iterations based on algorithm 1.

Table 1: Experimental result of Function 1

M	λ	f(X*)	M	λ	n_g	F(X*)
1	11	0.298638	30	1	125	0.298638
1	72	0.298638	30	3	219	9.84325e-8
1	775	0.298638	30	5	219	9.84325e-8
1	70	0.298638	30	7	219	9.84325e-8
1	17	2.89662e-8	30	9	219	9.49873e-8

As is seen from Table 2, when λ = 1 then μ_k = 1, $f(X^{k+1})$ can't go beyond $f(X^0)$. Once X^0 is located near the B or C, it is impossible to find out the optimal solution point A. When λ remains fixed and M varies,the solution is almost the same. When M≤5 and λ = 100, the optimal solution point A is finded out with less iteration number than λ = 1.

Function 2: Oren function (Grippo *et al.*, 1989):

$$f(X)=\left[\sum_{i=1}^{n} i x_i^2\right]^2$$

$$X^0=(1,\cdots,1)^T,\ X^*=(0,\cdots,0)^T,\ f(X^*)=0$$

The comparisons between algorithm 1 and the TNNL algorithm are made and presented in Table 3, where n_{gTNNL}

Table 2: Experimental result of function 1

M	λ	n_g	f(X*)	M	λ	n_g	F(X*)
2	1	24	0.298638	2	100	21	5.82805e-8
3	1	32	0.298638	3	100	27	2.68312e-8
4	1	36	0.298638	4	100	33	5.34422e-8
5	1	37	0.298638	5	100	38	8.43806e-8
6	1	49	0.298638	6	100	51	5.38465e-8

Table 3: Experimental result of oren function

n	M	λ	n_g	N_{gTNNL}
10	2	5	14	17
50	2	5	12	21
100	2	5	17	23
1000	2	5	25	-
10000	2	5	27	-

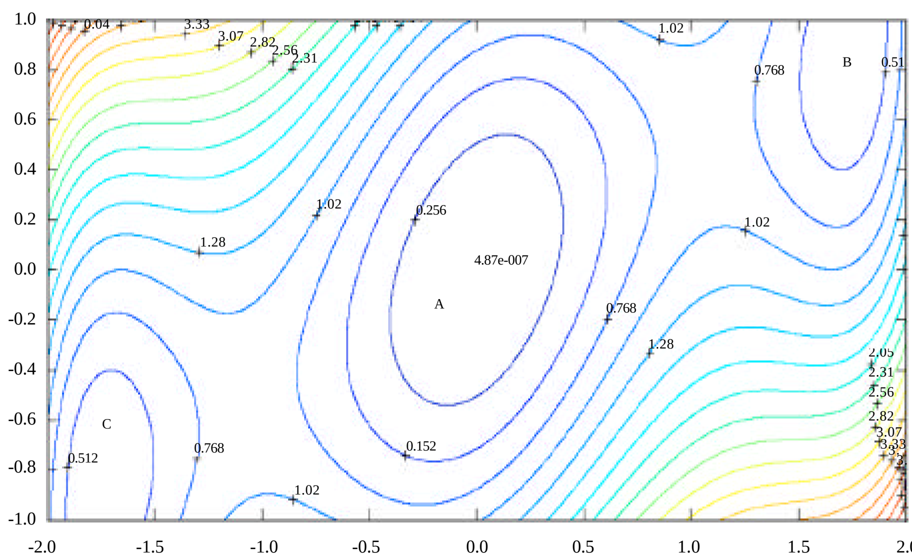

Fig. 1: Contour of function 1

refers to as the iterations given in Sun and Zhou (2007). From the table, it is seen that n_{gTNNL} is far larger than n_g for each situation.

The test of numerical example indicates that nonmonotone MMA algorithm we have proposed can solve large-scale optimization problems. Further, the comparisons given above show that the proposed new nonmonotone MMA algorithm possesses certain advantages, especially in multiple peak problems. It is mentioned here that the selection of the parameters involved in algorithm 1 can greatly influence on the efficiency of the algorithm. In order to keep the parameter value the same, so the iterations n_g in the table above are not the least. Different functions have their own better parameter combination. How to choose proper parameters to reduce the iterations is a problem of further research.

CONCLUSION

For unconstrained optimization problems, we improve MMA algorithm which is not necessary to adjust the trust region radius for reducing the search time. We propose a new nonmonotone combination rule for line searches. We have proved the global convergence of this new algorithm and numerical experiment results show that the new nonmonotonic MMA algorithm is effective to solve the optimization problems of multi peak and large-scale.

From nonmonotone rule for line searches Eq. 22, we know $f(x^{k+1}) \le \mu_k f_{l(k)}$. When $f_l(k)>0$, then $\mu_k \le 1$. So, $f(x^{k+1}) \ge f(x^k)$ is possible. When $f_{l(k)} \le 0$, then $0 \le \mu_k \le 1$, $f(x^{k+1}) \ge (f^{k+1}) \ge f(x^k)$ is also possible. Therefore, the nonmonotonic aim is achieved. From Eq. 25, we have $f(x^{k+1}) \le \lambda^r |f(x^0)|$ and $\lambda^r \ge 1$, so the aim of $f(x^{k+1}) > f(x^0)$ can be achieved and the purpose of relaxation can be achieved too. Our greatest breakthrough point is to let x^k out of x^0 near the bottom to search better solutions.

Combinatorial search criteria Eq. 22 is easy to be realized and there are a lot of h_k, such as:

$$h_k = \frac{1}{(k+1)^p}$$

constant p>1, then the series:

$$\sum_{k=0}^{\infty} \frac{1}{(k+1)^p}$$

converges in a finite number of γ, then:

$$\lambda^{\sum_{k=0}^{\infty} \frac{1}{(1+k)^p}} = \lambda^{\gamma}$$

ACKNOWLEDGMENTS

This study is supported by the National Natural Science Foundation of China (61074188).

REFERENCES

Grippo, L., F. Lampariello and S. Lucidi, 1986. A nonmonotone line search technique for Newton's method. SIAMM J. Numer. Anal., 23: 707-716.

Grippo, L., F. Lampariello and S. Lucidi, 1989. A truncated Newton method with nonmonotone line search for unconstrained optimization. J. Optim. Theory Appl., 60: 401-419.

Hu, P. and Q. Ni, 2010. A relaxation nonmonotone line search method. J. Numer. Methods Comput. Appl., 3: 191-199.

Hu, P., Z.H. Jia and N. Qi, 2012. A new algorithm of moving asymptotes for solving unconstrained optimization problems. Chin. J. Eng. Math., 3: 366-374.

Sun, W.Y. and Q.Y. Zhou, 2007. An unconstrained optimization method using nonmonotone second order Goldstein's line search. Sci. China Ser. A: Math., 50: 1389-1400.

Sun, W.Y., J.Y. Han and J. Sun, 2002. Global convergence of nonmonotone descent methods for unconstrained optimization problems. J. Comput. Applied Math., 146: 89-98.

Svanberg, K., 1987. The method of moving asymptotes: A new method for structural optimization. Int. J. Numer. Methods. Eng., 24: 359-373.

Wang, H. and Q. Ni, 2008. A new method of moving asymptotes for large-scale unconstrained optimization. Applied Math. Comput., 203: 62-71.

Xao, Y.H., H.J. Sun and Z.G. Wang, 2009. A globally convergent BFGS method with nonmonotone line search for non-convex minimization. J. Comput. Applied Math., 230: 95-106.

Yu, Z.S. and D.G. Pu, 2008. A new nonmonotone line search technique for unconstrained optimization. J. Comput. Applied Math., 219: 134-144.

Yuan, Y.X., 2008. Calculation Method of Nonlinear Optimization. Science Press, Beijing, China.

Zhou, Q.Y. and W.Y. Sun, 2008. Adaptive nonmonotone line search method for unconstrained optimization. Front. Math. China, 3: 133-148.