Hindawi Publishing Corporation
ISRN Operations Research
Volume 2013, Article ID 495378, 7 pages
http://dx.doi.org/10.1155/2013/495378

Research Article

A Nonmonotone Trust Region Algorithm Based on the Average of the Successive Penalty Function Values for Nonlinear Optimization

Zhensheng Yu[1] and Jinhong Yu[2]

[1] *College of Science, University of Shanghai for Science and Technology, Shanghai 200093, China*
[2] *College of Language, University of Shanghai for Science and Technology, Shanghai 200093, China*

Correspondence should be addressed to Zhensheng Yu; zhenshengyu@usst.edu.cn

Received 9 April 2013; Accepted 30 April 2013

Academic Editors: E. E. Ammar and G. Silva

We present a nonmonotone trust region algorithm for nonlinear equality constrained optimization problems. In our algorithm, we use the average of the successive penalty function values to rectify the ratio of predicted reduction and the actual reduction. Compared with the existing nonmonotone trust region methods, our method is independent of the nonmonotone parameter. We establish the global convergence of the proposed algorithm and give the numerical tests to show the efficiency of the algorithm.

1. Introduction

In this paper, we consider the equality constrained optimization problem as follows:

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

where $f(x) : R^n \to R$, $c(x) = (c_1(x), c_2(x), \ldots, c_m(x))^T$, $c_i(x) : R^n \to R^m$, $(i = 1, 2, \ldots, m)$, and $(m \le n)$ are assumed to be twice continuously differentiable.

Trust region method is one of the well-known methods for solving problem (1). Due to its strong convergence and robustness, trust region methods have been proved to be efficient for both unconstrained and constrained optimization problems [1–9].

Most traditional trust region methods are of descent type methods; namely, they accept only a trial point as the next iterate if its associated merit function value is strictly less than that of the current iterate. However, just as pointed out by Toint [10], the nonmonotone techniques are helpful to overcome the case that the sequence of iterates follows the bottom of curved narrow valleys, a common occurrence in difficult nonlinear problems. Hence many nonmonotone algorithms are proposed to solve the unconstrained and constrained optimization problems [11–20]. Numerical tests show that the performance of the nonmonotone technique is superior to those of the monotone cases.

The nonmonotone technique was originally proposed by Grippo, Lampariello and Lucidi [13] for unconstrained optimization problems based on Newton's method, in which the stepsize α_k satisfies the following condition:

$$f(x_k + \alpha_k d_k) \le \max_{0 \le j \le m_k} f(x_{k-j}) + \beta \alpha_k \nabla f(x_k)^T d_k, \tag{2}$$

where $\beta \in (0, 1)$, $0 \le m_k \le \min\{m_{k-1} + 1, M\}$, and M is a prefixed nonnegative integer.

Although the nonmonotone technique based on (2) works well in many cases, there are some drawbacks. Firstly, a good function value generated in any iteration is essentially discarded due to the maximum in (2). Secondly, in some cases, the numerical performance is heavily dependent on the choice of M (see, e.g., [16, 21]). To overcome these drawback, Zhang and Hager [21] proposed another nonmonotone algorithm, and they used the average of function values to replace the maximum function value in (2). The numerical tests show that their nonmonotone line search algorithm used fewer function and gradient evaluations, on average, than either the monotone or the traditional nonmonotone scheme.

Recently, Mo and Zhang [16] extended Zhang and Hager's nonmonotone technique to unconstrained optimization with trust region global scheme and discussed the global and local convergence of the proposed algorithm.

In this paper, we further extend the nonmonotone technique [16, 21] to equality constrained optimization. To design our algorithm, we first introduce some notations as follows: denote $g(x) = \nabla f(x)$ and $A(x) = (\nabla c_1(x), \nabla c_2(x), \dots, \nabla c_m(x)) \in R^{n\times m}$. Assuming that $A(x)$ has full column rank, we define the projective matrix

$$Z(x) = I - A(x)\left(A(x)^T A(x)\right)^{-1} A(x)^T \in R^{n\times n} \tag{3}$$

and the Lagrange function

$$L(x,\lambda) = f(x) + \lambda^T c(x), \tag{4}$$

where λ is a projective version of the multiplier vector as follows:

$$\lambda(x) = \left(A(x)^T A(x)\right)^{-1} A(x)^T g(x). \tag{5}$$

For convenience, we denote the previous quantities at x_k by f_k, c_k, g_k, A_k, Z_k, and λ_k. At each iteration, we calculate the trust region trial step as follows (see [22]): firstly, we calculate

$$\nu(x_k) = -\alpha_k A(x)\left[A(x_k)^T A(x_k)\right]^{-1} c(x_k), \tag{6}$$

where

$$\alpha_k = \begin{cases} 1, & c_k = 0, \\ \min\left\{1, \dfrac{\Delta_k}{\left\| A_k\left[A(x_k)^T A(x)\right]^{-1} c_k\right\|}\right\}, & \text{otherwise}. \end{cases} \tag{7}$$

Then we solve the trust region subproblem

$$\begin{aligned} \min \quad & (Z_k g_k)^T \omega + \left(\frac{1}{2}\right)\omega^T (Z_k B_k Z_k)\,\omega \\ \text{s.t.} \quad & \|\omega\| \le \Delta_k, \end{aligned} \tag{8}$$

where B_k denotes the Hessian matrix of the Lagrange function $L(x_k, \lambda_k)$, $\Delta_k > 0$ is the trust region radius. Let ω_k be the solution of (8) and

$$h_k = Z_k \omega_k. \tag{9}$$

The trust region trial step is taken as

$$d_k = h_k + \nu_k. \tag{10}$$

To test whether the point $x_k + d_k$ can be accepted as the next iteration, we use the Fletcher's exact penalty function as the merit function as follows:

$$\psi(x,\lambda,\sigma) = f(x) + \lambda^T c(x) + \sigma \|c(x)\|^2, \tag{11}$$

where $\sigma > 0$ is the penalty parameter.

To define our nonmonotone algorithm, we define

$$F_k = \begin{cases} \psi(x_k, \lambda_k, \sigma_k), & \text{if } k = 0, \\ \dfrac{(\eta_{k-1} Q_{k-1} F_{k-1} + \psi(x_k, \lambda_k, \sigma_k))}{Q_k}, & \text{if } k \ge 1, \end{cases} \tag{12}$$

where

$$Q_k = \begin{cases} 1, & \text{if } k = 0, \\ \eta_{k-1} Q_{k-1} + 1, & \text{if } k \ge 1, \end{cases} \tag{13}$$

where $\eta_{k-1} \in [\eta_{\min}, \eta_{\max}]$, $\eta_{\min} \in [0,1]$, $\eta_{\max} \in [\eta_{\min}, 1)$, and $\eta_{\min}, \eta_{\max}$ are two chosen parameters.

From (12) and (13), we observe that F_k is a convex combination of the function values

$$\psi(x_0, \lambda_0, \sigma_0), \psi(x_1, \lambda_1, \sigma_1), \dots, \psi(x_k, \lambda_k, \sigma_k), \tag{14}$$

so F_k is regarded as the weighted average of the merit function values.

The paper is organized as follows. We describe our algorithm in Section 2 and analyze the global convergence in Section 3. The numerical tests are given in Section 4, and the conclusion is presented in Section 5.

2. Algorithm

In this section, we give the details of the nonmonotone trust region algorithm. We first recall the definition of a stationary point of problem (1). A point x is called a stationary point of problem (1) if it satisfies

$$\left\| Z(x)^T g(x)\right\| + \|c(x)\| = 0. \tag{15}$$

We define the actual reduction from x_k to $x_k + d_k$ by

$$\text{Ared}_k = \psi(x_k, \lambda_k, \sigma_k) - \psi(x_k + d_k, \lambda_{k+1}, \sigma_k) \tag{16}$$

and the nonmonotone actual reduction by

$$\text{NAred}_k = F_k - \psi(x_k + d_k, \lambda_{k+1}, \sigma_k). \tag{17}$$

The predicted reduction is defined as

$$\begin{aligned} \text{Pred}_k = {} & -g_k^T d_k - \left(\frac{1}{2}\right) d_k^T B_k d_k - \nabla\lambda_k^T \left(c_k + A_k^T d_k\right) - \lambda_k^T A_k^T d_k \\ & + \sigma_k \left(\|c_k\|^2 - \left\|c_k + A_k^T d_k\right\|^2\right). \end{aligned} \tag{18}$$

Furthermore, we define the monotone ratio by

$$r_k = \frac{\text{Ared}_k}{\text{Pred}_k} \tag{19}$$

and the nonmonotone ratio by

$$\text{Nr}_k = \frac{\text{NAred}_k}{\text{Pred}_k}, \tag{20}$$

where F_k is computed by (12) and (13).

The description of the algorithm is given as follows.

Algorithm 1. Step 0. Set $x_0 \in \mathbb{R}^n$, $\Delta_0 > 0$, $\sigma_0 > 0$, $\mu \in (0,1)$, $0 < c_1 < c_2 < 1$, $c_3 > 0$, a symmetric matrix $B_0 \in \mathbb{R}^{n\times n}$, parameters $\eta_{\min} \in [0,1)$ and $\eta_{\max} \in [\eta_{\min}, 1)$, and $k := 0$.

Step 1. If $\|Z_k^T g_k\| + \|c_k\| = 0$, stop; otherwise, go to Step 2.

Step 2. Compute the trust region trial step d_k.

Step 3. Set $\sigma_k = \sigma_k$, if $\mathrm{Pred}_k \geq (1/2)\sigma_k(\|c_k\|^2 - \|c_k + A_k^T d_k\|^2)$, and then set

$$\sigma_k = \max\left\{\sigma_k, 2\frac{g_k^T d_k + (1/2)\, d_k^T B_k d_k + \nabla\lambda_k^T\left(c_k + A_k^T d_k\right) + \lambda_k^T A_k^T d_k}{\|_k\|^2 - \|c - k + A_k^T d_k\|^2}\right\}. \tag{21}$$

Step 4. Compute F_k by (12) and (13), and compute the Nr_k by (20).

Step 5. Set

$$x_{k+1} = \begin{cases} x_k + d_k, & \mathrm{Nr}_k \geq \mu, \\ x_k, & \text{otherwise.} \end{cases} \tag{22}$$

Step 6. Update Δ_{k+1} as

$$\Delta_{k+1} : \begin{cases} \in \left[c_1 \|d_k\|, c_2\Delta_k\right], & \text{if } \mathrm{Nr}_k < \mu, \\ = \Delta_k, & \text{if } \mathrm{Nr}_k \geq \mu,\ \|d_k\| < \Delta_k, \\ \in \left[\Delta_k, c_3\Delta_k\right], & \text{if } \mathrm{Nr}_k \geq \mu,\ \|d_k\| = \Delta_k, \end{cases} \tag{23}$$

go to Step 3.

Step 7. Update B_k, and choose $\eta_k \in [\eta_{\min}, \eta_{\max}]$. Set $k := k+1$; go to Step 1.

3. Global Convergence

In this section, we discuss the global convergence of Algorithm 1. The following assumptions are needed in our convergence analysis:

Assumptions

(A1) The sequence $\{x_k\}$ and $\{x_k + d_k\}$ are contained in a compact set Ω.

(A2) There exists a positive constant $M > 0$ such that for all k, $\|B_k\| \leq M$.

(A3) For all $x \in \Omega$, $A(x)$ is of column full rank.

We define two index sets as follows:

$$I = \{k : \mathrm{Nr}_k \geq \mu\}, \qquad J = \{k : \mathrm{Nr}_k \leq \mu\}. \tag{24}$$

The following lemmas (Lemmas 2–5) are helpful to analyze the convergence of the Algorithm 1, and the proofs are similar to [4].

Lemma 2. *Assume that (A1)–(A3) hold, and then there exists a positive constant K_1 such that*

$$\begin{gathered} \|c_k\|^2 - \|c_k + A_k^T d_k\|^2 \geq K_1 \|c_k\|^2 \min\left\{\|c_k\|^2, \Delta_k\right\}, \\ \mathrm{Pred}_k \geq \left(\frac{1}{2}\right)\sigma_k K_1 \|c_k\| \min\left\{\|c_k\|, \Delta_k\right\}. \end{gathered} \tag{25}$$

Lemma 3. *Let $\zeta_k(d) = g_k^T d + (1/2)d^T B_k d$, and assume that (A1)–(A3) hold. Then there exists a positive constant K_2 such that*

$$\begin{aligned} \zeta_k(d_k) \leq{}& \zeta_k(v_k) - K_2 \|Z_k(g_k + B_k v_k)\| \\ &\times \min\left\{\frac{\|Z_k(g_k + B_k v_k)\|}{M+1}, \Delta_k\right\}. \end{aligned} \tag{26}$$

Lemma 4. *Assume that (A1)–(A3) hold. Then there exists a positive constant K_3 such that*

$$\left|\mathrm{Ared}_k - \mathrm{Pred}_k\right| \leq K_3 \sigma_k \|d_k\|^2. \tag{27}$$

Lemma 5. *Assume that (A1)–(A3) holds. Then there exists a positive constant K_4 such that*

$$\begin{aligned} \mathrm{Pred}_k \geq{}& K_2 \|Z_k(g_k + B_k v_k)\| \min\left\{\frac{\|Z_k(g_k + B_k v_k)\|}{M+1}, \Delta_k\right\} \\ &- K_4 \|c_k\| + \sigma_k\left(\|c_k\|^2 - \|c_k + A_k^T d_k\|^2\right). \end{aligned} \tag{28}$$

The following lemma shows the monotonicity property of the function sequence $\{F_k\}$.

Lemma 6. *Suppose that $\{x_k\}$ is generated by Algorithm 1. Then the following inequality holds for all k:*

$$\psi_{k+1} \leq F_{k+1} \leq F_k. \tag{29}$$

Proof. We first prove that (29) holds for all $k \in I$; that is,

$$\psi_{k+1} \leq F_{k+1} \leq F_k, \quad \forall k \in I. \tag{30}$$

For $k \in$ I, according to Lemma 2, Assumptions (A1) and (A2), we obtain

$$\psi_{k+1} \leq F_k - \left(\frac{1}{2}\right)\sigma_k K_1 \|c_k\| \min\left\{\|c_k\|, \Delta_k\right\}. \tag{31}$$

According to (8)–(13), we have the following inequality:

$$\begin{aligned} F_{k+1} &= \frac{\eta_k Q_k F_k + \psi_{k+1}}{Q_{k+1}} \\ &\leq \frac{\eta_k Q_k F_k + F_k - (1/2)\,\sigma_k K_1 \|c_k\| \min\left\{\|c_k\|, \Delta_k\right\}}{Q_{k+1}} \\ &= F_k - \frac{(1/2)\,\sigma_k K_1 \|c_k\| \min\left\{\|c_k\|, \Delta_k\right\}}{\eta_k Q_k}. \end{aligned} \tag{32}$$

By (12) and (13), if $\eta_k = 0$, we have

$$F_{k+1} = \psi_{k+1}. \tag{33}$$

Otherwise, if $\eta_k \neq 0$, we have

$$F_{k+1} - F_k = \frac{\psi_{k+1} - F_{k+1}}{Q_{k+1}}. \tag{34}$$

So, from (32) to (34), we know that (30) holds.

Next, we prove that (29) holds for all $k \in J$. From Step 4 of Algorithm 1, we get $x_{k+1} = x_k$ and $\psi_{k+1} = \psi_k$ for $k \in J$. Firstly, we prove that $\psi_{k+1} \leq F_{k+1}$.

We consider two cases.

Case 1 ($k-1 \in I$). According to (8), we have $\psi_k \leq F_k$. Then it follows from (12) and (13) and $\psi_{k+1} = \psi_k$ that

$$F_{k+1} \geq \frac{\eta_k Q_k \psi_k + \psi_{k+1}}{Q_{k+1}} = \frac{\eta_k Q_k \psi_{k+1} + \psi_{k+1}}{Q_{k+1}} = \psi_{k+1}. \tag{35}$$

Case 2 ($k-1 \in J$). In this situation, let $K = \{i \mid 1 < i \leq k, k-i \in I\}$. If $K = \emptyset$, from Step 4 of Algorithm 1, we have $F_0 = F_{k-j} = F_{k+1}$, $j = 0, 1, \ldots, k-1$. Consequently, it follows from (12) and (13) that

$$F_{k+1} = F_k = \psi_{k+1}. \tag{36}$$

We suppose that $K \neq \emptyset$ and set $m = \min\{i : i \in K\}$, and then we have

$$\psi_{k-j} = \psi_k = \psi_{k+1}, \quad j = 0, 1, \ldots, m-1. \tag{37}$$

By (12), we obtain

$$Q_k F_k = \eta_{k-1} Q_{k-1} F_{k-1} + \psi_k, \quad k \geq 1. \tag{38}$$

According to (38) repeatedly, we can get

$$\begin{aligned} \eta_k Q_k F_k + \psi_{k+1} &= \prod_{i=0}^{m-1} \eta_{k-i} Q_{k-m+1} F_{k-m+1} \\ &\quad + \sum_{j=0}^{m-2} \prod_{i=0}^{j} \eta_{k-i} \psi_{k-j} + \psi_{k+1}. \end{aligned} \tag{39}$$

Using the definition of K and m, we know that $k-m \in I$ and $F_{k-m+1} \geq \psi_{k-m+1}$ through (8).

From (37) and (39), it follows that

$$\begin{aligned} &\eta_k Q_k F_k + \Psi_{k+1} \\ &\quad \geq \prod_{i=0}^{m-1} \eta_{k-i} Q_{k-m+1} \psi_{k-m+1} \\ &\qquad + \sum_{j=0}^{m-2} \prod_{i=0}^{j} \eta_{k-i} \psi_{k-j} + \psi_{k+1} \\ &\quad = \left(\prod_{i=0}^{m-1} \eta_{k-i} Q_{k-m+1} + \sum_{j=0}^{m-2} \prod_{i=0}^{j} \mu_{k-i} + 1 \right) \psi_{k+1} \\ &\quad = Q_{k+1} \psi_{k+1}. \end{aligned} \tag{40}$$

From (12) and (40) we know that

$$F_{k+1} = \frac{\eta_k Q_k F_k + \psi_{k+1}}{Q_{k+1}} \geq \frac{Q_{k+1} \psi_{k+1}}{Q_{k+1}} = \psi_{k+1} \tag{41}$$

By (35), (36), and (42), we get

$$\psi_{k+1} \leq F_{k+1}, \quad \forall k \in J. \tag{42}$$

Now we prove that $F_{k+1} \leq F_k$. If $\eta_k \neq 0$, from (34) and (42), the conclusion is obvious. If $\eta_k = 0$, then by (12), (13) and $k \in J$, we have $F_{k+1} = F_k$. Thus (29) holds for all $k \in J$. The proof is completed. □

Theorem 7. *Suppose that the Assumptions (A1)–(A3) hold and the sequence $\{x_k\}$ is generated by Algorithm 1. Then the algorithm is well defined.*

Proof. Since the algorithm does not stop in Step 2, then we have either $\|c_k\| \neq 0$ or $\|Z_k^T g_k\| \neq 0$. We prove the conclusion by contradiction; if the conclusion is not true, by the algorithm, we have $x_{k+1} = x_k$, but

$$Nr_k < \mu, \qquad \lim_{k \to \infty} \Delta_k = 0. \tag{43}$$

Case 1 ($\|c_k\| \neq 0$). Then from Lemmas 2 and 4, we have

$$\begin{aligned} \lim_{k\to\infty} \left| r_k - 1 \right| &= \lim_{k\to\infty} \left| \frac{\mathrm{Ared}_k - \mathrm{Pred}_k}{\mathrm{Pred}_k} \right| \\ &\leq \lim_{k\to\infty} \frac{2K_3 \sigma_k \left\| \Delta_k \right\|^2}{\sigma_k K_1 \left\| c_k \right\| \min\left\{ \left\| c_k \right\|, \Delta_k \right\}} \\ &\leq \lim_{k\to\infty} \frac{2K_3 \sigma_k \left\| \Delta_k \right\|^2}{\sigma_k K_1 \left\| c_k \right\| \Delta_k} = 0, \end{aligned} \tag{44}$$

which means that $r_k > \mu$ for k large enough, according to Lemma 6, and we have that $\mathrm{NAred}_{k,i} = F(x_k + d_k) - \psi(x_k + d_k) > \mathrm{Ared}_k$, so $\mathrm{Nr}_k \geq r_k > \mu$, which contradicts (43).

Case 2 ($\|c_k\| = 0$). In this case, we have $\|v_k\| = 0$ and $\|Z_k^T g_k\| \neq 0$, By Lemma 3, and we can have

$$\begin{aligned} \mathrm{Pred}_k &= -\zeta(d_k) \geq K_2 \left\| Z_k (g_k + B_k v_k) \right\| \\ &\quad \times \min\left\{ \frac{\left\| Z_k (g_k + B_k v_k) \right\|}{M+1}, \Delta_k \right\} \\ &= K_2 \left\| Z_k g_k \right\| \Delta_k. \end{aligned} \tag{45}$$

Combining with Lemma 4, we have

$$\begin{aligned} \lim_{k\to\infty} \left| r_k - 1 \right| &= \lim_{k\to\infty} \left| \frac{\mathrm{Ared}_k - \mathrm{Pred}_k}{\mathrm{Pred}_k} \right| \\ &\leq \lim_{k\to\infty} \frac{K_3 \sigma_k \left\| \Delta_k \right\|^2}{K_2 \left\| Z_k g_k \right\| \Delta_k} = 0. \end{aligned} \tag{46}$$

Then similar to Case 1, we can get a contradiction. Combining Cases 1 and 2, we can get the conclusion. □

Similar to Lemma 7.11 in [4], we get the proposition of the penalty parameter as follows.

Lemma 8. *Under Assumption A1, if* $\|Z_k^T g_k\| + \|c_k\| \neq 0$, *then there exist a integer* k_0 *and a positive constant* $\sigma^\star$ *such that for all* $k \geq k_0$, $\sigma_k = \sigma^\star$.

Without loss of generality, we assume that $\sigma_k = \sigma^\star$ for all k. The following theorem gives the convergence proposition of the constraint sequence $\{\|c_k\|\}$.

Theorem 9. *Under the Assumptions (A1)–(A3), we have*

$$\lim_{k \to \infty} \|c_k\| = 0. \tag{47}$$

Proof. First, we prove that

$$\liminf_{k \to \infty} \|c_k\| = 0. \tag{48}$$

Assume by contradiction that (48) does not hold, then there exists a constant $\varepsilon > 0$ such that $\|c_k\| \geq \varepsilon$ for all k. According to Lemma 6, we have

$$\begin{aligned} F_{k+1} &\leq F_k - \frac{\mathrm{Pred}_k}{Q_{k+1}}, \\ F_k - F_{k+1} &\geq \frac{\mathrm{Pred}_k}{Q_{k+1}}, \\ F_{k-1} - F_k &\geq \frac{\mathrm{Pred}_k}{Q_k}, \\ F_{k-2} - F_{k-1} &\geq \frac{\mathrm{Pred}_k}{Q_{k-1}}, \\ &\vdots \\ F_1 - F_2 &\geq \frac{\mathrm{Pred}_k}{Q_2}. \end{aligned} \tag{49}$$

By using (13), we can prove that

$$\begin{aligned} Q_{k+1} &= 1 + \sum_{j=0}^{k} \prod_{i=0}^{j} \eta_{k-i} \\ &\leq 1 + \sum_{j=0}^{k} \eta_{\max}^{j+1} \leq \sum_{j=0}^{\infty} \eta_{\max}^{j} = \frac{1}{1-\eta_{\max}}. \end{aligned} \tag{50}$$

Adding all the previous inequalities and by Lemma 2, we have

$$\begin{aligned} F_1 - F_{k+1} &\geq \sum_{i=1}^{k+1} \frac{\mathrm{Pred}_i}{Q_{i+1}} \\ &\geq \frac{1}{2(1-\eta_{\max})} \sum_{i=1}^{k+1} \sigma_i K_1 \|c_i\| \min\{\|c_i\|, \Delta_i\}. \end{aligned} \tag{51}$$

By Assumption (A1), we know that $F_1 - F_{k+1}$ is bounded, let $k \to \infty$, and we have

$$\begin{aligned} +\infty &> F_1 - F_{k+1} \\ &\geq \frac{1}{2(1-\eta_{\max})} \sum_{k=1}^{\infty} \sigma^* K_1 \|c_k\| \min\{\|c_k\|, \Delta_k\}. \end{aligned} \tag{52}$$

Since $\|c_k\| \geq \varepsilon$ for all k, we have $\lim_{k\to\infty} \Delta_k = 0$. But similar to the proof of Theorem 7, we get $\mathrm{Nr}_k > \mu$, and therefore we have $\Delta_{k+1} > \Delta_k$, which contradicts to $\lim_{k\to\infty}\Delta_k = 0$. This contradiction shows that (48) holds.

Next we prove (47). Assume that (47) does not hold, then there exist a subsequence $\{m_j\}$ and a positive constant ε_1 such that

$$\|c_{mj}\| \geq \varepsilon_1. \tag{53}$$

On the other hand, according to (48) we know that there exists another subsequence $\{l_i\}$ such that for $\varepsilon_2 = \varepsilon_1/2$, we have

$$\begin{gathered} \|c_k\| \geq \varepsilon_2 \qquad m_j \leq k \leq l_j, \\ \|c_{lj}\| \leq \varepsilon_2. \end{gathered} \tag{54}$$

We define $\mathcal{K} = \{k \mid m_j \leq k \leq l_j\}$. According to Lemma 2, we get the following inequality:

$$F_1 - F_{k+1} \geq \frac{1}{2(1-\eta_{\max})} \sum_{k\in\mathcal{K}} \sigma_k K_1 \|c_k\| \min\{\varepsilon_2, \Delta_k\}. \tag{55}$$

By Assumption (A1), F_k is bounded, so we have that $\min\{\varepsilon_2/2, \Delta_j\} = \varepsilon_2/2$ can be true only finite number of times. Thus there exists $k_1 > 0$ such that for $j > k_1$, we have $\min\{\varepsilon_2/2, \Delta_j\} = \Delta_j$. Hence for $j > k_1$, we have

$$\sum_{j\in\mathcal{K}, j=k_1}^{k} \Delta_j \leq \frac{2(1-\eta_{\max})}{K_1 K_2 \varepsilon_2^2} \left[F_1 - \min_{x\in\Omega} F(x)\right] < \infty. \tag{56}$$

Then we know that

$$\sum_{j\in\mathcal{K}, j=k}^{\infty} \Delta_j \longrightarrow 0 \quad (k \longrightarrow \infty). \tag{57}$$

Now, for large j,

$$\|x_{lj} - x_{mj}\| \leq \sum_{k=m_j}^{l_j-1} \|x_{k+1} - x_k\| \leq \sum_{k=m_j}^{l_j-1} \Delta_k < \sum_{k=m_j}^{\infty} \Delta_k \longrightarrow 0. \tag{58}$$

Since $c(x)$ is continuous, thus for j large enough we have $\|c_{mj} - c_{lj}\| < \varepsilon_2$,

$$\|c_{mj}\| \leq \|c_{mj} - c_{lj}\| + \|c_{lj}\| < 2\varepsilon_2, \tag{59}$$

and this contradicts to the assumption $\|c_{mj}\| \geq 2\varepsilon_2$, which means that (47) holds. □

Theorem 10. *If (A1) holds, we have*

$$\liminf_{k\to\infty} \left\| Z_k^T g_k \right\| = 0. \tag{60}$$

Proof. Similar to the proof of Theorem 4 in [18]. □

Based on Theorems 9 and 10, we get the following global convergence result.

Theorem 11. *Under Assumptions (A1)–(A3), we have*

$$\liminf_{k\to\infty} \left\| Z_k^T g_k \right\| + \left\| c_k \right\| = 0. \tag{61}$$

4. Numerical Tests

In this section, we test our algorithm for some typical problems. The program code was written in MATLAB and run in MATLAB 7.1 environment. The parameters in our algorithm are taken as follows: $\Delta_0 = 0.1$, $\sigma_0 = 1$, $\mu = 0.1$, $c_1 = 0.2$, $c_2 = 0.8$, $c_3 = 1.2$, $\eta_k \equiv 0.75$, and $B_0 = I$, and B_k is updated by BFGS formulas as follows:

$$B_{k+1} = \begin{cases} B_k, & \text{if } \delta_k^T y_k \le 0, \\ B_k + \dfrac{y_k y_k^T}{y_k^T \delta_k} - \dfrac{B_k \delta_k \delta_k^T B_k}{\delta_k^T B_k \delta_k}, & \text{if } \delta_k^T y_k > 0, \end{cases} \tag{62}$$

where $\delta_k = x_{k+1} - x_k$, $y_k = (\nabla f(x_{k+1}) - A(x_{k+1})\lambda(x_{k+1}) - (\nabla f(x_k) - A(x_k)\lambda(x_k))$. For deciding when to stop the execution of the algorithm declaring convergence we used the criterion $\|Z_k^T g_k\| + \|c_k\| \le 10^{-5}$. We also stop the execution when 500 iterations were completed without achieving convergence and denoted by fail. Our test problems are chosen from [23], and the problems are numbered in the same way as in [23]. For example, HS28 is the problem 28 in [23]. To test the efficiency of our algorithm, we compare our algorithm with the algorithms in [15, 18], where we choose the nonmonotone parameter $M = 5$.

The test results are given in Table 1: here we use No. to denote the number of the test problems, I_g and I_f denote the number of gradient estimation and the function value estimation, and Time denotes the CPU time when the algorithm is terminated.

From Table 1, we see that our algorithm spend more CPU time than algorithms [15, 18], but we use less function value estimation and gradient value estimation for most of the test problem. These numerical tests show that our algorithm works quiet well.

5. Conclusion

In this paper, we presented a nonmonotone trust region method based on the weighted average of the successive penalty values for equality constrained optimization. Compared with the existing nonmonotone trust region methods for constrained optimization, our method is independent on the nonmonotone parameter M. The numerical comparison with some nonmonotone trust region methods shows the efficiency of our proposed method. How to obtain the local fast convergence of our method deserves further study, and we leave it as the future work.

Table 1: Test results for our method and the methods in [15, 18].

No.	Our method		The method in [18]		The method in [15]	
	I_t/I_g	Time	I_t/I_g	Time	I_t/I_g	Time
H28	11/13	0.3438	13/24	0.2652	13/24	0.1404
H39	59/61	1.4688	24/37	0.3432	64/126	0.2625
H42	45/73	0.8281	133/195	0.2652	fail	
H47	17/21	0.5625	63/121	0.5460	60/118	0.2964
H48	7/10	0.3125	14/26	0.2340	14/26	0.7488
H49	100/197	2.3750	118/234	0.4524	118/234	0.3144
H50	23/27	0.7344	63/124	0.9360	63/124	0.4212
H51	143/223	2.6094	57/88	0.6084	57/88	1.0764
H52	426/658	7.4844	50/100	0.8112	149/188	1.9812
H63	18/20	0.5469	15/27	0.5928	15/27	0.1404
H77	11/15	0.4063	25/48	1.2324	109/132	3.4788

Acknowledgments

This work is supported by National Natural Science Foundation of China (no. 11171221) and National Project and Liberal Base Cultivate Fund of USST (no. 12GXM).

References

[1] M. El-Alem, "A global convergence theory for Dennis, El-Alem, and Maciel's class of trust-region algorithms for constrained optimization without assuming regularity," *SIAM Journal on Optimization*, vol. 9, no. 4, pp. 965–990, 1999.

[2] R. H. Byrd, R. B. Schnab, and G. A. Schultz, "A trust region algorithm for nonlinearly constrained optimization," *SIAM Journal on Numerical Analysis*, vol. 24, no. 5, pp. 1152–1169, 1987.

[3] A. R. Conn, N. Gould, and L. Ph. Toint, *Trust-Region Methods*, SIAM, Philadelphia, Pa, USA, 2000.

[4] J. E. Dennis and L. N. Vicente, "On the convergence theory of trust-region-based algorithms for equality-constrained optimization," *SIAM Journal on Optimization*, vol. 7, no. 4, pp. 927–950, 1997.

[5] M. J. D. Powell, "On the global convergence of trust region algorithms for unconstrained optimization," *Mathematical Programming*, vol. 29, no. 3, pp. 297–303, 1984.

[6] M. J. D. Powell and Y. X. Yuan, "A trust region algorithms for equality constrained optimization," *Mathematical Programming B*, vol. 49, no. 1, pp. 189–211, 1991.

[7] A. Vardi, "A trust region algorithm for equality constrained minimization: convergence properties and implemention," *SIAM Journal on Numerical Analysis*, vol. 22, no. 3, pp. 575–591, 1985.

[8] Y. X. Yuan, "Convergence of trust region methods," *Chinese Journal of Numerical Mathematics and Applications*, vol. 16, no. 4, pp. 92–106, 1994.

[9] J. L. Zhang, *Trust region algorithm for nonlinear optimization [Ph.D. thesis]*, Institute of Applied Mathematics, Chinese Academy of Science, 2001.

[10] P. L. Toint, "A non-monotone trust-region algorithms for nonlinear optimization subject to convex constraints," *Mathematical Programming B*, vol. 77, no. 1, pp. 69–94, 1997.

[11] Z. W. Chen and X. S. Zhang, "A nonmonotone trust-region algorithm with nonmonotone penalty parameters for constrained optimization," *Journal of Computational and Applied Mathematics*, vol. 172, no. 1, pp. 7–39, 2004.

[12] N. Y. Deng, Y. Xiao, and F. Zhou, "Nonmonotone trust region algorithm," *Journal of Optimization Theory and Applications*, vol. 76, no. 2, pp. 259–285, 1993.

[13] L. Grippo, F. Lampariello, and S. Lucidi, "A nonmonotone line search techniques for Netown's method," *SIAM Journal on Numerical Analysis*, vol. 23, no. 4, pp. 707–716, 1986.

[14] Z. F. Li and N. Y. Deng, "A class of new nonmonotone trust region algorithm and its convergence," *Acta Mathematicae Applicatae Sinica*, vol. 22, no. 3, pp. 139–145, 1999.

[15] X. W. Ke and J. Y. Han, "A nonmonotone trust region algorithm for equality constrained optimization," *Science in China A*, vol. 38, no. 6, pp. 683–695, 1995.

[16] J. T. Mo and K. C. Zhang, "A nonmonotone trust region method for unconstrained optimization," *Applied Mathematics and Computation*, vol. 171, no. 1, pp. 371–384, 2005.

[17] J. T. Mo, C. Y. Liu, and S. C. Yan, "A nonmonotone trust region method based on nonincreasing technique of weighted average of the successive function values," *Journal of Computational and Applied Mathematics*, vol. 209, no. 1, pp. 97–108, 2007.

[18] Z. S. Yu, C. X. He, and Y. Tian, "Global and local convergence of a nonmonotone trust region algorithm for equality constrained optimization," *Applied Mathematical Modelling*, vol. 34, no. 5, pp. 1194–1202, 2010.

[19] H. C. Zhang, "A nonmonotone trust region algorithm for nonlinear optimization subject to general constraints," *Journal of Computational Mathematics*, vol. 21, no. 2, pp. 237–246, 2003.

[20] D. T. Zhu, "A nonmonotonic trust region technique for nonlinear constrained optimization," *Journal of Computational Mathematics*, vol. 13, no. 1, pp. 20–31, 1995.

[21] H. Zhang and W. W. Hager, "A nonmonotone line search technique and its application to unconstrained optimization," *SIAM Journal on Optimization*, vol. 14, no. 4, pp. 1043–1056, 2004.

[22] J. Zhang and D. Zhu, "A projective quasi-Newton method for nonlinear optimization," *Journal of Computational and Applied Mathematics*, vol. 53, no. 3, pp. 291–307, 1994.

[23] W. Hock and K. Schittkowski, *Test Examples For Nonlinear Programming Codes*, Springer, Berlin, Germany, 1981.


Advances in
Operations Research
Advances in
Decision Sciences
Journal of
Applied Mathematics
Algebra
Journal of
Probability and Statistics
The Scientific
World Journal
International Journal of
Differential Equations
Hindawi
Submit your manuscripts at
http://www.hindawi.com
International Journal of
Combinatorics
Advances in
Mathematical Physics
Journal of
Complex Analysis
Journal of
Mathematics
Mathematical Problems
in Engineering
Abstract and
Applied Analysis
Discrete Dynamics in
Nature and Society
International
Journal of
Mathematics and
Mathematical
Sciences
Journal of
Discrete Mathematics
Journal of
Function Spaces
International Journal of
Stochastic Analysis
Journal of
Optimization