A trust region interior point algorithm for solving bound-constrained nonlinear systems

JIA Chun-xia, ZHU De-tong

(Mathematics and Sciences College, Shanghai Normal University, Shanghai 200234, China)

Abstract: We propose a trust region interior algorithm for solving nonlinear equality systems subject to bound constraints on variables. Under some reasonable conditions, the proposed algorithm is globally convergent and has locally fast convergent rate. Results of numerical experiments are reported to show the effectiveness of the proposed algorithm.

Key words: trust region method; interior points; nonlinear equation

CLC number: 0221.2 **Document code:** A **Article ID:** 1000-5137(2005)04-0001-07

1 Introduction

We consider the trust region interior point algorithm for solving the nonlinear systems subjective to bound constraints on variables:

$$F(x) = 0, \Omega = \{ x \mid l \leqslant x \leqslant u \}. \tag{1.1}$$

where $F:\mathscr{X} \to \mathbf{R}^n$ is a given continuously differentiable mapping and $\mathscr{X} \subseteq \mathbf{R}^n$ is an open set containing the n-dimensional box constraint Ω. The vector $l \in (\mathbf{R} \cup \{-\infty\})^n$ and $u \in (\mathbf{R} \cup +\infty)^n$ are respectively specified lower and upper bounds on the variables such that $\operatorname{int}(\Omega) \overset{\text{def}}{=} \{ x \mid l < x < u \}$ is nonempty. Generally, the line search and trust region have been used in order to ensure global convergence towards local minima. Quite a few papers have used these two strategies for solving problem (1.1) in [1] and [4].

In order to describe and design the algorithms for solving (1.1), we first introduce the squared Eudidean norm reformulation of linear model of the systems (1.1) and state the trust region algorithm with backtracking interior point technique for the nonlinear equations in the next section. In Section 3, we prove the global convergence of the proposed algorithm. We discuss some further convergence properties such as strong global convergence and characterize the order of local convergence of the Newton methods in terms of the rates of the relative residuals in Section 4. Finally, the results of our numerical experiments of the proposed algorithm are reported in section 5.

2 Algorithm

In this section we describe a trust region strategy for solving the bound-constrained problem (1.1). The

Received date: 2005-09-01

Biography: JIA Chun-xia (1976 –), Female, Lecturer, Mathematics and Sciences College, Shanghai Normal University; ZHU De-tong (1954 –), Male, Doctor, Professor, Mathematics and Sciences College, Shanghai Normal University.

method involves choosing a matrix D_k and a quadratic model for the objective function. For every interior point x_k, we define the matrix $D_k \overset{\text{def}}{=} \{D_k^{11}, \cdots, D_k^{nn}\}$ at the kth iteration with the component D_k^{ii} defined as follows:

$$D_k^{ii} = \begin{cases} x_k^i - l_i & \text{if } l_i \text{ is finite and } u_i \text{ is not} \\ u_i - x_k^i & \text{if } u_i \text{ is finite and } l_i \text{ is not} \\ \min\{x_k^i - l_i, u_i - x_k^i\}, & \text{if both } l_i \text{ and } u_i \text{ are finite} \\ 1 & \text{if either } l_i \text{ nor } u_i \text{ is finite} \end{cases}.$$

where x_k^i, l_i and u_i are the ith components of the vectors x_k, l and u respectively. Another matrix B_k is a symmetric approximation of the Hessian of the objective function.

Now we state the trust region interior point algorithm for the problem (1.1).

Main steps:

(1) Evaluate $f_k = f(x_k) \overset{\text{def}}{=} \frac{1}{2}\|F(x_k)\|^2$, $g_k = \nabla f(x_k) \overset{\text{def}}{=} (F_k')^T F_k$ and D_k.

(2) If $\|D_k g_k\| = \|D_k(F_k')^T F_k\| \leqslant \epsilon$, stop with the approximate solution x_k.

(3) Solve a step d_k based on the trust region subproblem

$$\min \quad \psi_k(d) \overset{\text{def}}{=} \frac{1}{2}\|F_k' d + F_k\|^2$$

$$(S_k) \text{ s.t.} \quad \|D_k^{-1} d\| \leqslant \Delta_k,\ l < x_k + d < u.$$

(4) Choose $\alpha_k = 1, \omega, \omega^2, \cdots$, until the following inequality is satisfied:

$$f(x_k + \alpha_k d_k) \leqslant f(x_{l(k)}) + \alpha_k \beta g_k^T d_k,$$

where $f(x_{l(k)}) = \max_{0 \leqslant j \leqslant m(k)} \{f(x_{k-j})\}$. Set $x_{k+1} = x_k + \alpha_k d_k$.

(5) Calculate $\text{Pred}_k = \psi_k(0) - \psi_k(\alpha_k d_k)$, $\text{Ared}_k = f(x_k) - f(x_k + \alpha_k d_k)$, and $\rho_k = \frac{\text{Ared}_k}{\text{Pred}_k}$. Update the trust region size from Δ_k to Δ_{k+1}

$$\Delta_{k+1} = \begin{cases} [\gamma_1 \Delta_k,\ \gamma_2 \Delta_k] & \text{if } \rho_k \leqslant \eta_1 \\ (\gamma_2 \Delta_k,\ \Delta_k] & \text{if } \eta_1 < \rho_k < \eta_2 \\ (\Delta_k,\ \min\{\gamma_3 \Delta_k,\ \Delta_{\max}\}] & \text{if } \rho_k \geqslant \eta_2 \end{cases}$$

(6) Take the nonmonotone control parameter $m(k+1) = \min\{m(k)+1, M\}$ and update F_k' to obtain F_{k+1}'. Then set $k \leftarrow k+1$ and go to step 2.

Remark A key property of this transformation in trust region subproblem (D_k) is that $D_k^{-1} d_k$ is at least unit distance from all the bounds in the scaled coordinates, i. e., an arbitrary step $D_k^{-1} d_k$ to the point $x_k + d_k$ does not violate any bound if $d_k^T D_k^{-2} d_k < 1$. It is easy to see that the usual monotone algorithm can be viewed as a special case of the proposed algorithm when $M = 0$.

3 Convergence Analysis

Throughout this section we assume that $F: \mathscr{X} \subset \mathbf{R}^n \to \mathbf{R}^n$ is bounded from below and the level set of f is denoted by $\mathscr{L}(x_0) = \{x \in \mathbf{R}^n \mid f(x) \leqslant f(x_0),\ l \leqslant x \leqslant u\}$. The following assumption is commonly used in convergence analysis of most methods for the box constrained systems.

Assumption 1 The sequence $\{x_k\}$ generated by the algorithm is contained in a compact set $\mathscr{L}(x_0)$ on $\mathbf{R}^n$. There exist some positive constants χ_g and χ_D such that

$$\|F'^T(x) F'(x)\| \leqslant \chi_g,\ \|D(x)\| \leqslant \chi_D,\ \forall\ x \in \mathscr{L}(x_0).$$

Based on solving the above trust region subproblem (S_k), imitating the proof of Lemma 3.4 in [5] which is due to Sorensen, the following lemma establishes a necessary and sufficient condition concerning the pair λ_k, d_k when d_k solves the subproblem (S_k).

Lemma 3.1 d_k is a solution to the subproblem (S_k) if and only if d_k is a solution to the following equations

$$[D_k(F_k'^T F_k')D_k + \lambda_k I]D_k^{-1}d_k = -D_k(F_k')^T F_k, \tag{3.1}$$

$$\lambda_k(\|D_k^{-1}d_k\| - \Delta_k) = 0,\ \lambda_k \geq 0, \tag{3.2}$$

and $[D_k(F_k'^T F_k')D_k + \lambda_k I]$ is positive semidefinite.

It is well known from solving the trust region algorithms that in order to assure the global convergence of the proposed algorithm, it is sufficient to show that at the kth iteration the predicted reduction defined by $\mathrm{Pred}(d_k) = \psi_k(0) - \psi_k(d_k)$, which is obtained by the step d_k from trust region subproblem (S_k), satisfies a sufficient descent condition. The lemma is due to Lemma 3.1 in [6].

Lemma 3.2 Let the step d_k be the solution of the trust region subproblem (S_k), and assume that Assumption 1 holds. Then there exists a $\tau \in (0, \frac{1}{2})$ such that d_k satisfies the following sufficient descent condition

$$\mathrm{Pred}(d_k) \geq \tau \|D_k(F_k')^T F_k\| \min\{1, \Delta_k, \frac{\|D_k(F_k')^T F_k\|}{\|D_k(F_k')^T F_k' D_k\|}\} \tag{3.3}$$

for all F_k', F_k, D_k and Δ_k.

The following lemma shows the relation between the gradient $g_k = (F_k')^T F_k$ of the objective function and the step d_k generated by the proposed algorithm. We can see from the lemma that the direction of the trial step is a sufficiently descent direction since $D_k(F_k'^T F_k')D_k$ is positive semidefinite.

Lemma 3.3 At the iteration, let d_k be generated in the trust region subproblem (S_k). Then

$$\nabla f(x_k)^T d_k \leq -\tau \|D_k(F_k')^T F_k\| \min\{1, \Delta_k, \frac{\|D_k(F_k')^T F_k\|}{\|D_k(F_k')^T F_k' D_k\|}\}, \tag{3.4}$$

where the constant τ is given in (3.3).

Lemma 3.4 Let f be differentiable and assume that its gradient is such that

$$\|\nabla f(x) - \nabla f(y)\|_2 \leq \gamma \|x - y\|_2,\ \forall x, y \in \mathbf{R}^n, \tag{3.5}$$

where γ is the Lipschitz constant. Let $\beta \in (0,1)$ and d_k be proposed by the subproblem (S_k). If $\|D_k g_k\| \neq 0$, then Algorithm 2.1 will produce an iteration $x_{k+1} = x_k + \alpha_k d_k$ in a finite number of backtracking steps. (The proof is omitted since the space is limited.)

We are going to state one of our main results of the proposed algorithm, but it requires the following assumptions.

Assumption 2 $D_k(F_k')^T F_k' D_k$ is bounded, i.e., there exists a constant $\chi > 0$, such that $b_k \overset{\mathrm{def}}{=} \|D_k(F_k')^T F_k' D_k\| \leq \chi,\ \forall\ k$.

Assumption 3 The first order optimality system associated to problem (1.1) has no nonisolated solutions.

Theorem 3.4 Assume that Assumptions 1 ~ 3 hold. Let $\{x_k\} \subset \mathbf{R}^n$ be a sequence generated by the algorithm. Then

$$\liminf_{k\to\infty} \|D_k(F_k')^T F_k\| = 0. \tag{3.6}$$

Proof According to the acceptance of α_k, we have

$$f(x_{l(k)}) - f(x_k + \alpha_k d_k) \geqslant -\alpha_k \beta g_k^T d_k = -\alpha_k \beta [D_k (F_k')^T F_k]^T (D_k^{-1} d_k). \quad (3.7)$$

Notice $m(k+1) \leqslant m(k) + 1$, $f(x_{k+1}) \leqslant f(x_{l(k)})$, we get:

$$f(x_{l(k+1)}) = \max_{0 \leqslant j \leqslant m(k+1)} f(x_{k+1-j}) \leqslant \max_{0 \leqslant j \leqslant m(k)+1} f(x_{k+1-j}) = \max_{0 \leqslant j \leqslant m(k)} f(x_{k-j}) = f(x_{l(k)}).$$

So $\{f(x_{l(k)})\}$ decreases, and $\lim_{k\to\infty} f(x_{l(k)})$ exist. Similar to the theorem of [4], the equality

$$\lim_{k\to\infty} f(x_{l(k)}) = \lim_{k\to\infty} f(x_k) \quad (3.8)$$

holds. By step 4, we have

$$f(x_{l(k)}) - f(x_k + \alpha_k d_k) \geqslant -\alpha_k \beta g_k^T d_k \geqslant \beta \tau \alpha_k \| D_k (F_k')^T F_k \| \min \left\{1, \Delta_k, \frac{\| D_k (F_k')^T F_k \|}{\| D_k (F_k')^T F_k' D_k \|}\right\}.$$

If the conclusion is not true, there exists some $\epsilon > 0$ such that

$$\| D_k (F_k')^T F_k \| \geqslant \epsilon, \ k > K, \quad (3.9)$$

where K is a positive integer. So $f(x_{l(k)}) - f(x_k + \alpha_k d_k) \geqslant \alpha_k \beta \tau \epsilon \min \{1, \Delta_k, \frac{\epsilon}{\chi}\}$. Notice that $\{f(x_{l(k)})\}$ is convergent. We obtain $\lim_{k\to\infty} \alpha_k \Delta_k = 0$, which implies that either

$$\liminf_{k\to\infty} \alpha_k = 0, \quad (3.10)$$

or

$$\lim_{k\to\infty} \Delta_k = 0. \quad (3.11)$$

Now, we can prove that if (3.11) holds, $\alpha_k = 1$ satisfies:

$$f(x_k + \alpha_k d_k) \leqslant f(x_{l(k)}) + \alpha_k \beta g_k^T d_k, \text{ that is: } f(x_k + d_k) \leqslant f(x_{l(k)}) + \beta g_k^T d_k. \quad (3.12)$$

In fact, if the above formula is not true, we will get $f(x_k + d_k) > f(x_{l(k)}) + \beta g_k^T d_k \geqslant f(x_k) + \beta g_k^T d_k$. So

$$0 < f(x_k + d_k) - f(x_k) - \beta g_k^T d_k = [\nabla f(x_k + \xi_k d_k)^T - \nabla f(x_k)^T] d_k + (1-\beta) g_k^T d_k \leqslant$$

$$(1-\beta) g_k^T d_k + \gamma \xi_k \| d_k \|^2 \leqslant -\tau (1-\beta) \epsilon \min \{1, \Delta_k, \frac{\epsilon}{\chi}\} + \gamma \chi_D^2 \Delta_k^2,$$

where $\xi_k \in (0,1)$, which contradicts the fact that $\lim_{\Delta_k \to 0} \frac{-\tau (1-\beta)\epsilon \min \{1, \Delta_k, \frac{\epsilon}{\chi}\} + \gamma \chi_D^2 \Delta_k^2}{\Delta_k} = -\tau (1-\beta)\epsilon$. So, we can conclude that when Δ_k is small enough, (3.12) holds. Notice

$$| \mathrm{Ared}(d_k) - \mathrm{Pred}(d_k) | = \| f(x_k + d_k) - f(x_k) - [\varphi_k(d_k) - \varphi_k(0)] \| =$$

$$\| [\frac{1}{2} \| F_k \|^2 - \frac{1}{2} \| F(x_k + d_k) \|^2] - [\frac{1}{2} \| F_k \|^2 - \frac{1}{2} \| F_k' d_k + F_k \|^2] \| \leqslant$$

$$\| F_k + F_k' d_k \| \cdot \| \omega(x_k, d_k) \| + \frac{1}{2} \| \omega(x_k, d_k) \|^2 \leqslant$$

$$[\| F_k \| + \frac{1}{2} \| \omega(x_k, d_k) \|] \cdot \| \omega(x_k, d_k) \|,$$

where $\omega(x_k, d_k) = \int_0^1 [F'(x_k + \xi d_k) - F'(x_k)] d_k \, d\xi$, and $\omega(x_k, d_k) \leqslant \frac{1}{2} \gamma \| d_k \|^2$. We obtain

$$| f_k - f(x_k + d_k) - [\psi_k(0) - \psi_k(d_k)] | \leqslant (\| F_k \| + \frac{1}{2} \gamma \| d_k \|^2) \cdot \frac{1}{2} \gamma \| d_k \|^2.$$

So

$$| \rho_k - 1 | \leqslant \frac{(\| F_k \| + \frac{1}{2} \gamma \| d_k \|^2) \cdot \frac{1}{2} \gamma \| d_k \|^2}{\tau \epsilon \min \{1, \Delta_k, \frac{\epsilon}{\chi}\}} \to 0,$$

which contradicts (3.11). So (3.10) holds, i.e. $\liminf_{k\to\infty} \alpha_k = 0$. Furthermore, the acceptance rule means that

$$f(x_k + \frac{\alpha_k}{\omega}d_k) - f_k \geqslant f(x_k + \frac{\alpha_k}{\omega}d_k) - f_{l(k)} \geqslant \beta \frac{\alpha_k}{\omega} g_k^T d_k.$$

Since $f(x_k + \frac{\alpha_k}{\omega} d_k) - f_k = \frac{\alpha}{\omega} g_k^T d_k + o(\frac{\alpha_k}{\omega} \| d_k \|)$, we have $(1-\beta) \frac{\alpha_k}{\omega} g_k^T d_k + o(\frac{\alpha_k}{\omega} \| d_k \|) \geqslant 0$. And hence $\lim_{k\to\infty} \frac{g_k^T d_k}{\| d_k \|} = 0$, so

$$0 = \lim_{k\to\infty} \frac{g_k^T d_k}{\| d_k \|} \leqslant -\tau \epsilon \min \{ \lim_{k\to\infty} \frac{1}{\| d_k \|}, \lim_{k\to\infty} \frac{\Delta_k}{\| d_k \|}, \lim_{k\to\infty} \frac{\epsilon}{\chi \| d_k \|} \} \leqslant 0.$$

Noticing $\frac{\Delta_k}{\| d_k \|} \geqslant \frac{\Delta_k}{\| \chi_D \| \, \| D_k^{-1} d_k \|} \geqslant \frac{1}{\chi_D} > 0$, we obtain $\| d_k \| \to +\infty$, which contradicts $\| d_k \| = \| D_k D_k^{-1} d_k \| \leqslant \chi_D \Delta_{\max}$. Hence the conclusion of the theorem is true. □

4 Properties of the Local convergence

Theorem 3.4 indicates that at least one limit point of $\{x_k\}$ is a stationary point. In this section we shall first extend this theorem to a stronger result and the local convergent rate.

Theorem 4.1 Assume that Assumptions 1 - 3 hold. Let $\{x_k\}$ be a sequence generated by the proposed algorithm. Then

$$\lim_{k\to\infty} \| D_k (F_k')^T F_k \| = 0. \tag{4.1}$$

Proof Assume that there is an $\epsilon_1 \in (0,1)$ and a subsequenee $\{ D_{m_i} (F_{m_i}')^T F_{m_i} \}$ such that for all m_i, $i = 1, 2, \cdots s$, $\| D_{m_i} (F_{m_i}')^T_{m_i} \| \geqslant \epsilon_1$. Theorem 3.4 guarantees the existence of another subsequence $\{ D_{l_i} (F_{l_i}')^T F_{l_i} \}$ suchthat $\| D_k (F_k')^T F_k \| \geqslant \epsilon_2$, for $m_i \leqslant k < l_i$ and $\| D_{l_i} (F_{l_i}')^T F_{l_i} \| \leqslant \epsilon_2$ for an $\epsilon_2 \in (0, \epsilon_1)$.

Work similar to the proof of Theorem 3.4 leads to

$$\lim_{k\to\infty} f(x_{l(k)}) = \lim_{k\to\infty} f(x_k). \tag{4.2}$$

Similarly, we also get $\lim_{k\to\infty, m_i \leqslant k \leqslant l_i} \alpha_k \Delta_k = 0$. By the accepting rule of α_k, we get

$$f(x_k + \frac{\alpha_k}{\omega}d_k) - f(x_k) \geqslant f(x_k + \frac{\alpha_k}{\omega}d_k) - f(x_{l(k)}) \geqslant \beta \frac{\alpha_k}{\omega} \nabla f(x_k)^T d_k. \tag{4.3}$$

Using the mean value theorem we have the following equality:

$$f(x_k + \frac{\alpha_k}{\omega}d_k) - f(x_k) = \frac{\alpha_k}{\omega} [\nabla f(x_k + \xi_k \frac{\alpha_k}{\omega} d_k) - \nabla f(x_k)]^T d_k + \frac{\alpha_k}{\omega} \nabla f(x_k)^T d_k \tag{4.4}$$

with $\xi_k \in (0,1)$. Since $f(x)$ is Lipschitz continuously differentiable with constant γ, we have

$$\| [\nabla f(x_k + \xi_k \frac{\alpha_k}{\omega} d_k) - \nabla f(x_k)]^T d_k \| \leqslant \gamma \xi_k \frac{\alpha_k}{\omega} \| d_k \|^2 \leqslant \gamma \frac{\alpha_k}{\omega} \| d_k \|^2. \tag{4.5}$$

From (4.3) ~ (4.5), we get $-(1-\beta) \frac{\alpha_k}{\omega} \nabla f(x_k)^T d_k \leqslant \gamma (\frac{\alpha_k}{\omega})^2 \| d_k \|^2$. So

$$\gamma \alpha_k \| d_k \|^2 \geqslant -\omega(1-\beta) \nabla f(x_k)^T d_k \geqslant \omega(1-\beta)\tau \| D_k (F_k')^T F_k \| \min \{ 1, \Delta_k, \frac{\| D_k (F_k')^T F_k \|}{\| D_k (F_k')^T F_k' D_k \|} \} \geqslant$$

$$\omega (1-\beta) \tau \epsilon_2 \min \{ 1, \Delta_k, \frac{\epsilon_2}{\chi} \} \geqslant 0,$$

which implies $\lim_{k\to\infty, m_i \leqslant k < l_i} \Delta_k = 0$. Hence

$$0 \leqslant \lim_{k\to\infty,\, m_i\leqslant k<l_i} \| d_k \| \leqslant \lim_{k\to\infty,\, m_i\leqslant k<l_i} \| D_k \| \, \| D_k^{-1} d_k \| \leqslant \lim_{k\to\infty,\, m_i\leqslant k<l_i} \chi_D \Delta_k = 0. \tag{4.6}$$

Because $f(x)$ is Lipschitz continuously differentiable, we have

$$f(x_k + d_k) = f(x_k) + \nabla f(x_k + \xi_k d_k)^T d_k =$$
$$f(x_k) + \beta \nabla f(x_k)^T d_k + (1-\beta) \nabla f(x_k)^T d_k + [\nabla f(x_k + \xi_k d_k) - \nabla f(x_k)]^T d_k \leqslant$$
$$f(x_{l(k)}) + \beta \nabla f(x_k)^T d_k + \{(1-\beta) \nabla f(x_k)^T d_k + \gamma \| d_k \|^2\},$$

where $\xi_k \in [0,1]$ and γ is Lipschitz constant. By Lemma 3.3 and (4.6), we know that if k is large enough, we will get $f(x_k + d_k) \leqslant f(x_{l(k)}) + \beta \nabla f(x_k)^T d_k$. Similar to Theorem 3.4, the following result is true:

$$\| f(x_k) - f(x_k + d_k) - [\psi_k(0) - \psi_k(d_k)] \| \leqslant (\| F_k \| + \frac{1}{2}\gamma \| d_k \|^2) \gamma \| d_k \|^2.$$

From (3.3), for large enough i, $m_i \leqslant k < l_i$,

$$\mathrm{Pred}(d_k) \geqslant \tau \| D_k (F_k')^T F_k \| \min\{ 1, \Delta_k, \frac{\| D_k (F_k')^T F_k \|}{\chi} \} \geqslant \tau \epsilon_2 \min\{1, \Delta_k, \frac{\epsilon_2}{\chi}\}.$$

We obtain that

$$\rho_k \geqslant 1 - \frac{(\| F_k \| + \frac{1}{2}\gamma \| d_k \|^2) \| d_k \|^2}{\tau \epsilon_2 \min\{\Delta_k, \frac{\epsilon_2}{\chi}\}} \geqslant \eta_2.$$

i. e., when i is large enough, $m_i \leqslant k < l_i$, $f_k - f(x_k + d_k) \geqslant \eta_2 \mathrm{Pred}(d_k) \geqslant \eta_2 \tau \epsilon_2 \min\{1, \Delta_k, \frac{\epsilon_2}{\chi}\}$. It follows that for sufficiently large i, $m_i \leqslant k < l_i$, when $\Delta_k \leqslant \frac{\epsilon_2}{\chi}$, $f_k - f(x_k + d_k) \geqslant \sigma \Delta_k$, where $\sigma = \eta_2 \tau \epsilon_2$. From $\| x_{k+1} - x_k \| = \| d_k \| = \| D_k D_k^{-1} d_k \| \leqslant \| D_k \| \, \| D_k^{-1} d_k \| \leqslant \chi_D \Delta_k$, we deduce

$$\| x_{m_i} - x_{l_i} \| \leqslant \sum_{k=m_i}^{l_i} \| x_k - x_{k+1} \| \leqslant \sum_{k=m_i}^{l_i} \chi_D \Delta_k \leqslant \frac{\chi_D}{\sigma} \sum_{k=m_i}^{l_i} [f_k - f(x_k + d_k)] = \frac{\chi_D}{\sigma}(f_{m_i} - f_{l_i}).$$

Noticing (4.2), we can get $\lim_{i\to\infty} \| x_{m_i} - x_{l_i} \| = 0$, so when i is large enough,

$$\| x_{m_i} - x_{l_i} \| \leqslant \epsilon_2, \qquad \| F_{m_i}' F_{m_i} - F_{l_i}' F_{l_i} \| = \| \nabla f(x_{m_i}) - \nabla f(x_{l_i}) \| \leqslant \gamma \| x_{m_i} - x_{l_i} \| \leqslant \gamma \epsilon_2.$$

So

$$\epsilon_1 \leqslant \| D_{m_i} (F_{m_i}')^T F_{m_i} \| \leqslant$$
$$\| D_{m_i} \| \, \| F_{m_i}' F_{m_i} - F_{l_i}' F_{l_i} \| + \| D_{m_i} - D_{l_i} \| \, \| (F_{l_i}')^T F_{l_i} \| + \| D_{l_i} (F_{l_i}')^T F_{l_i} \| \leqslant$$
$$(\chi_D \gamma + \chi_F + 1) \epsilon_2,$$

which contradicts that $\epsilon_2 \in (0, \epsilon_1)$. □

We now discuss the convergence rate of the proposed algorithm.

Theorem 4.2 Assume that Assumptions 1 ~ 3 hold, $F(x^*) = 0$ and $F'(x^*)$ is nonsingular at every limit point x^* of $\{x_k\}$. For sufficiently large k, then the step $\alpha_k \equiv 1$, and the trust region constraints is inactive, that is, there exists a $\hat{\Delta} > 0$ such that

$$\Delta_k \geqslant \hat{\Delta}, \ \forall \ k \geqslant K'.$$

where K' is a large enough index.

Theorem 4.2 means that the local convergence rate of the proposed algorithm depends on the Hessian of objective function at x^* and the local convergence rate of the step. If d_k becomes the Newton step, then the sequence $\{x_k\}$ generated by the algorithm converges x^* quadratical.

5 Numerical Experiments

In this section we present some numerical results. In order to check the effectiveness of the method, we select the parameters as follows: $\epsilon = 1\times e^{-6}$, $\eta_1 = 0.1$, $\eta_2 = 0.75$, $\gamma_1 = 0.2$, $\gamma_2 = 0.5$, $\gamma_3 = 2$, $\Delta_{\max} = 0.96$, and $\Delta_0 = 1$.

Tabel 1　Experimental results

Problem results	the optimal solution and the optimal value		M = 0	
	reference results	results of the proposed algorithm	NG	NF
14.1.1	$x^* = (-2.8051, 3.1313)^T$ $f^* = 0$	$x^* = (-2.8051, 3.1313)^T$ $f^* = 1.9725\times 10^{-6}$	6	6
14.1.2	$x^* = (0.003431, 31.325636, 0.068352, 0.859530, 0.036963)^T$	$x^* = (0.0034, 31.3265, 0.0684, 0.8595, 0.0370)^T$ $f^* = 7.3739\times 10^{-19}$	15	19
14.1.4	$x^* = (0.5, 3.14159)^T$	$x^* = (0.5000, 3.1416)^T$ $f^* = 2.2801\times 10^{-15}$	9	10
14.1.5	$x^* = (1,1,1,1,1)^T$	$x^* = (1,1,1,1,1)^T$ $f^* = 9.2564\times 10^{-18}$	8	8
14.1.8	$x^* = (0.007848, 0.010593)^T$	$x^* = (0.0078, 0.0106)^T$ $f^* = 1.1028\times 10^{-15}$	6	7

The experiments are carried out for 5 test problems quoted from [3]. NF and NG stand for the numbers of function evaluations and gradient evaluations, respectively.

References:

[1] BONNANS J F, POLA C. A trust region interior point algorithm for linear constrained optimization[J]. SIAM J Optimization, 1997, 7(3): 717 - 731.

[2] DENG N Y, XIAO Y, ZHOU F J. A nonmonotonic trust region algorithm[J]. Journal of Optimization Theory and Applications, 1993, 76: 259 - 285.

[3] FLOUDAS C A, et al. Handbook of Test Problems In Local Optimization[M]. Dordrecht: Kluwer Academic, 1999.

[4] GRIPP L, LAMPARIELLO F, LUCIDI. A nonmonotone line search technique for Newton's methods[J]. SIAM J Numer Anal, 1986, 23: 707 - 716.

[5] SORENSEN D C. Newton's method with a model trust region modification[J]. SIAM J Numer Anal, 1982, 19: 409 - 426.

[6] ZHU D. Curvilinear paths and trust region methods with nonmonotonic backtracking technique for unconstrained optimization[J]. J of Computational Mathematics, 2001, 19: 241 - 258.

有界变量约束非线性方程组的信赖域内点算法

贾春霞，朱德通
（上海师范大学 数理信息学院，上海 200234）

摘要：提出一种有界变量约束非线性方程组的信赖域内点算法，在合理的条件下所提供的算法不仅能整体收敛于方程组的解而且保持局部收敛速率．数值计算结果说明算法的有效性．

关键词：信赖域；约束优化；内点；非线性方程组