文章编号:1000-0887(2010)12-1503-10

有界约束非线性系统的结合 Lanczos 分解技术不精确 Newton 法*

张 勇[1], 朱德通[2]

(1. 上海师范大学 数理学院, 上海 200234;
2. 上海师范大学 商学院, 上海 200234)

摘要: 提出了结合 Lanczos 分解技术不精确 Newton 法求解有界变量约束非线性系统. 通过 Lanczos 分解技术解一个仿射二次模型获得迭代方向. 利用内点回代线搜索技术,沿着这个方向得到一个可接受的步长. 在合理的假设条件下,证明了算法的整体收敛性与局部超线性收敛速率. 此外,数值结果表明了算法的有效性.

关 键 词: 非线性系统; Lanczos 方法; 不精确 Newton 法; 非单调技术

中图分类号: O221.2 **文献标志码:** A

DOI: 10.3879/j.issn.1000-0887.2010.12.011

引 言

信赖域方法广泛用于解非线性系统问题. 如 Coleman 和 Li[1] 提出了一种双信赖域算法求解有界约束非线性最优化问题. 近来, Bellavia 等[2] 进一步延伸了这些算法,进而提出用仿射信赖域方法求解有界约束非线性系统. 信赖域方法解非线性最优化问题时能确保整体收敛性. 然而,为获得成功迭代,需要多次求解信赖域子问题,因此计算量相当大且困难的. Lanczos 分解方法主要是在 Krylov 空间形成一组正交基,并且在解非线性方程组问题时能减少计算量. 如 Jia 和 Zhu[3] 提出了仿射内点 Lanczos 方法求解有界约束的非线性方程组问题. 此外,许多不精确 Newton 方法求解光滑的非线性方程的分析基于对收敛性的分析,如文献[4-5]. 然而,不精确 Newton 法求解光滑系统问题时不能确保得到整体收敛性,即收敛仅仅是局部的.

由上述启发,本文结合 Lanczos 分解技术和不精确 Newton 法,构建出结合 Lanczos 分解技术的不精确 Newton 法求解有界变量约束非线性系统问题.

本文的结构如下:第 1 节,给出非线性系统求解的预备知识;第 2 节,提出结合 Lanczos 分解技术的不精确 Newton 法求解有界变量约束非线性系统问题的算法;第 3 节,证明算法的一些性质;第 4 节,讨论了结合 Lanczos 分解技术不精确 Newton 法的整体收敛性和局部收敛性的

---

* **收稿日期:** 2010-03-14; **修订日期:** 2010-11-01

基金项目: 国家自然科学基金资助项目(10871130);教育部博士点基金资助项目(20093127110005);上海市重点学科资助项目(T0401)

作者简介: 张勇(1981—),男,江苏人,博士(E-mail: yzhyzhang@163.com);
朱德通(1954—),男,博士,博士生导师(联系人. Tel: +86-21-64323361; E-mail: dtzhu@shnu.edu.cn).

阶数. 第 5 节,给出了算法的数值结果.

1 预备知识

本节分析有界变量约束的非线性系统解的问题

$$\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{0}, \qquad \boldsymbol{x} \in \Omega = \{\boldsymbol{x} \mid \boldsymbol{l} \leqslant \boldsymbol{x} \leqslant \boldsymbol{u}\}, \tag{1}$$

这里, $\boldsymbol{F}: \mathcal{X} \to \Re^n$ 是一给定的可微连续映射, $\mathcal{X} \subseteq \Re^n$ 是一包含 n- 维有界约束 Ω 的开集,其中向量 $\boldsymbol{l} \in (\Re \cup \{-\infty\})^n$ 和 $\boldsymbol{u} \in (\Re \cup \{+\infty\})^n$ 是可行域的上下界且 $\mathrm{int}(\Omega) \stackrel{\mathrm{def}}{=} \{\boldsymbol{x} \mid \boldsymbol{l} < \boldsymbol{x} < \boldsymbol{u}\}$ 非空. 引入价值函数

$$f(\boldsymbol{x}) = \frac{1}{2} \| \boldsymbol{F}(\boldsymbol{x}) \|^2. \tag{2}$$

因此,问题(1)等价转化为

$$\min f(\boldsymbol{x}), \qquad \boldsymbol{x} \in \Omega. \tag{3}$$

问题(3)在无约束的情况下, $\boldsymbol{x}_k$ 处的近似二次模型为

$$m_k(\boldsymbol{p}) = \frac{1}{2} \| \boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{p} + \boldsymbol{F}(\boldsymbol{x}_k) \|^2 =$$

$$\frac{1}{2}\boldsymbol{p}^{\mathrm{T}}\boldsymbol{F}'^{\mathrm{T}}(\boldsymbol{x}_k)\boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{p} + \boldsymbol{F}^{\mathrm{T}}(\boldsymbol{x}_k)\boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{p} + \frac{1}{2}\boldsymbol{F}^{\mathrm{T}}(\boldsymbol{x}_k)\boldsymbol{F}(\boldsymbol{x}_k). \tag{4}$$

另外,为解问题(3),我们需引入一个仿射矩阵. 现在定义向量函数 $\boldsymbol{\phi}(\boldsymbol{x})$, 它的分量定义如下:

$$\phi^i(x) \stackrel{\mathrm{def}}{=} \begin{cases} 1, & \text{当 } l^i = -\infty \text{ 或 } u^i = +\infty, \\ \min\{x^i - l^i + \gamma\max\{0, -g^i(x)\}, & \\ \qquad u^i - x^i + \gamma\max\{0, g^i(x)\}\}, & \text{否则}, \end{cases} \tag{5}$$

其中, x^i, l^i 和 u^i 分别是 $\boldsymbol{x}, \boldsymbol{l}$ 和 $\boldsymbol{u}$ 的第 i 个分量. $\boldsymbol{g}(\boldsymbol{x}) \stackrel{\mathrm{def}}{=} \boldsymbol{F}'^{\mathrm{T}}(\boldsymbol{x})\boldsymbol{F}(\boldsymbol{x})$, $g^i(x)$ 是 $\boldsymbol{g}(\boldsymbol{x})$ 的第 i 个分量, $\gamma > 0$(γ 是一个常数).

进而定义内点仿射矩阵

$$\boldsymbol{D}(\boldsymbol{x}) \stackrel{\mathrm{def}}{=} \mathrm{diag}(\phi^1(x), \cdots, \phi^n(x)), \tag{6}$$

则问题(3)的一阶必要条件等价表示为

$$\boldsymbol{D}(\boldsymbol{x})^{-1} \nabla f(\boldsymbol{x}) = 0. \tag{7}$$

为求解方程(7),令 $\bar{\boldsymbol{p}} = \boldsymbol{D}_k \boldsymbol{p}$,则 $\boldsymbol{p} = \boldsymbol{D}_k^{-1}\bar{\boldsymbol{p}}$,其中 $\boldsymbol{D}_k = \boldsymbol{D}(\boldsymbol{x}_k)$. 此时,在 $\boldsymbol{x}_k$ 处,仿射二次模型定义为

$$\bar{m}_k(\bar{\boldsymbol{p}}) = \frac{1}{2}\bar{\boldsymbol{p}}^{\mathrm{T}}\boldsymbol{D}_k^{-1}\boldsymbol{F}'^{\mathrm{T}}(\boldsymbol{x}_k)\boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{D}_k^{-1}\bar{\boldsymbol{p}} + \boldsymbol{F}^{\mathrm{T}}(\boldsymbol{x}_k)\boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{D}_k^{-1}\bar{\boldsymbol{p}} + \frac{1}{2}\boldsymbol{F}^{\mathrm{T}}(\boldsymbol{x}_k)\boldsymbol{F}(\boldsymbol{x}_k). \tag{8}$$

2 算法

本节给出结合 Lanczos 分解技术的不精确 Newton 法,并结合非单调技术求解有界变量约束问题(8)的算法.

初始步

选定参数 $\mu \in (0,1)$, $\omega \in (0,1)$, $\varepsilon > 0$ 和正整数 M 作为非单调参数. 令 $m(0) = 0$,选取初始严格可行点 $\boldsymbol{x}_0 \in (\boldsymbol{l}, \boldsymbol{u}) \subseteq \Re^n$. 令 $k = 0$, 转主步.

主步

1）计算 $F_k' = F'(x_k)$，D_k，D_k^{-1}，$M_k = D_k^2$，$F_k'^{\mathrm{T}} F_k$，$D_k^{-1} F_k'^{\mathrm{T}} F_k$，并且选取一个控制序列 $\eta_k \in (0,1)$；

2）如果 $\| D_k^{-1} F_k'^{\mathrm{T}} F_k \| \leqslant \varepsilon$，则停止计算，$x_k$ 作为最优解；否则转入下一步；

3）置 $i = 1$，$v_0 = \mathbf{0}$，$r_1 = \nabla m_k(v_1) = F_k'^{\mathrm{T}} F_k$，$\omega_0 = \mathbf{0}$，$d_0 = \mathbf{0}$，可得 $d_1 = - M_k^{-1} F_k'^{\mathrm{T}} F_k$；

4）令 $\omega_i = F_k' d_i$，检验

$$\omega_i^{\mathrm{T}} \omega_i \neq 0, \tag{9}$$

$$D_k^{-1} r_i \neq 0, \tag{10}$$

若式(9)和式(10)都满足，转入下一步，否则转入步 6)；

5）计算

$$y_i = M_k^{-1} r_i,\ \gamma_i = \sqrt{\langle r_i, y_i \rangle},\ \omega_i = \frac{r_i}{\gamma_i},\ q_i = \frac{y_i}{\gamma_i},\ \delta_i = \| F_k' q_i \|^2,$$

$$r_{i+1} = F_k'^{\mathrm{T}} F_k' q_i - \delta_i \omega_i - \gamma_i \omega_{i-1},\ \lambda_i = \frac{\theta_i^2 r_i^{\mathrm{T}} y_i}{\omega_i^{\mathrm{T}} \omega_i} > 0,\ \theta_{i+1} = - \lambda_i \theta_i \gamma_i \qquad (\theta_1 = 1),$$

$$\beta_{i-1} = \frac{\theta_i y_i^{\mathrm{T}} F_k'^{\mathrm{T}} \omega_{i-1}}{\omega_{i-1}^{\mathrm{T}} \omega_{i-1}} > 0,\ d_i = - \theta_i y_i + \beta_{i-1} d_{i-1},\ v_i = v_{i-1} + \lambda_{i-1} d_{i-1};$$

检验

$$\| F(x_k) + F'(x_k) v_i \| \leqslant \eta_k \| F(x_k) \|, \tag{11}$$

若式(11)满足，转入下一步，否则，令 $i \Leftarrow i + 1$ 转步 4)；

6）若 $i = 1$，$p_k = d_1$；$i \geqslant 2$，$p_k = v_i$；

7）选取 $\alpha_k = 1, \omega, \omega^2, \cdots$，直到下式成立：

$$\| F(x_k + \alpha_k p_k) \| \leqslant \| F(x_{l(k)}) \| + \mu(\eta_k - 1) \| F(x_k) \|, \tag{12}$$

且
$$x_k + \alpha_k p_k \in (l, u), \tag{13}$$

其中
$$\| F(x_{l(k)}) \| = \max_{0 \leqslant j \leqslant m(k)} \{ \| F(x_{k-j}) \| \};$$

8）令

$$x_{k+1} = x_k + \alpha_k p_k; \tag{14}$$

9）取 $m(k+1) = \min\{ m(k) + 1, M \}$. 令 $k \leftarrow k + 1$，再转步 1)。

注释 1 步 1) 中 η_k 有一些合适控制序列。本文只选取 $\eta_k = \min\{ 1/(k+2), \| F(x_k) \| \}$.

注释 2[6] 知

$$\nabla m_k(v_{i+1}) = F'(x_k)^{\mathrm{T}} F'(x_k) v_{i+1} + F_k'^{\mathrm{T}} F_k = \theta_{i+1} r_{i+1},$$

其中 $\theta_{i+1} = \langle e_{i+1}, h_{i+1} \rangle$，得到 $\theta_{i+1} = - \lambda_i \theta_i \gamma_i (\theta_1 = 1)$.

注释 3[1] 在式(13)中 α_k 表示沿 p_k 受有界变量约束 $l < x_k + \alpha_k p_k < u$ 的步长，即

$$\alpha_k^* \overset{\text{def}}{=} \min\left\{ \max\left\{ \frac{l^i - x_k^i}{p_k^i}, \frac{u^i - x_k^i}{p_k^i} \right\},\ i = 1, 2, \cdots, n \right\}, \tag{15}$$

令 $\alpha_k = \vartheta_k \alpha_k^*$，其中 $\vartheta_k \in (\vartheta_l, 1]$，$0 < \vartheta_l < 1$ 和 $\vartheta_k - 1 = O(\| p_k \|)$. l^i, u^i, x_k^i 和 p_k^i 分别是 l, u, x_k 和 p_k 的第 i 个分量。若 $p_k^i = 0$，则令

$$\frac{l^i - x_k^i}{p_k^i} = \frac{u^i - x_k^i}{p_k^i} = + \infty.$$

α_k 的选取要确保 $\alpha_k p_k$ 不会超出有界约束的边界。

3　算法的性质

本节给出算法的一些性质.

引理 3.1[3]　设 $\boldsymbol{q}_i$ 和 $\boldsymbol{d}_i$ 由算法第 5) 步在 k 次迭代中得到, $1 \leqslant i \leqslant l \leqslant m+1$, 则下列性质成立:

$$\boldsymbol{q}_i^{\mathrm{T}} \boldsymbol{D}_k^2 \boldsymbol{q}_j = 0, \qquad 1 \leqslant j < i \leqslant l \leqslant m+1, \tag{16}$$

$$\boldsymbol{Q}_i^{\mathrm{T}} \boldsymbol{F}'^{\mathrm{T}}(\boldsymbol{x}_k) \boldsymbol{F}'(\boldsymbol{x}_k) \boldsymbol{Q}_i = \boldsymbol{T}_i, \qquad i = 1, 2, \cdots, m+1, \tag{17}$$

$$\boldsymbol{r}_i^{\mathrm{T}} \boldsymbol{d}_j = 0, \qquad 1 \leqslant j < i \leqslant l \leqslant m+1, \tag{18}$$

$$\boldsymbol{\omega}_i^{\mathrm{T}} \boldsymbol{\omega}_j = 0, \qquad i \neq j, \tag{19}$$

$$\boldsymbol{d}_i^{\mathrm{T}} \boldsymbol{D}_k^2 \boldsymbol{d}_j > 0, \qquad i \neq j, \tag{20}$$

$$\boldsymbol{d}_j^{\mathrm{T}} \boldsymbol{r}_j = -\theta_j \boldsymbol{r}_j^{\mathrm{T}} \boldsymbol{D}_k^{-2} \boldsymbol{r}_j, \qquad j = 1, 2, \cdots, m+1, \tag{21}$$

其中 $\boldsymbol{Q}_i = [\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_i]$ 和三对角矩阵 $\boldsymbol{T}_i$ 为

$$\boldsymbol{T}_i = \begin{bmatrix} \delta_1 & \gamma_2 & & & \\ \gamma_2 & \delta_2 & \gamma_3 & & \\ & \ddots & \ddots & \ddots & \\ & & \gamma_{i-1} & \delta_{i-1} & \gamma_i \\ & & & \gamma_i & \delta_i \end{bmatrix}.$$

引理 3.2　设 $\boldsymbol{v}_j$ 由算法第 5) 步产生的迭代点列, 则其范数单调增加, 即

$$\| \boldsymbol{D}_k \boldsymbol{v}_j \|_2 < \| \boldsymbol{D}_k \boldsymbol{v}_{j+1} \|_2, \qquad 1 \leqslant j \leqslant m. \tag{22}$$

证明　由 $\boldsymbol{v}_0 = \boldsymbol{0}$ 和算法第 5) 步, 可得

$$\boldsymbol{v}_j = \boldsymbol{v}_0 + \sum_{i=0}^{j-1} \lambda_i \boldsymbol{d}_i = \sum_{i=0}^{j-1} \lambda_i \boldsymbol{d}_i.$$

既然 $\lambda_i > 0$, $\boldsymbol{d}_i^{\mathrm{T}} \boldsymbol{D}_k^2 \boldsymbol{d}_j > 0\ (i \neq j)$, 有

$$\boldsymbol{v}_j^{\mathrm{T}} \boldsymbol{D}_k^2 \boldsymbol{d}_j = \left(\sum_{i=0}^{j-1} \lambda_i \boldsymbol{d}_i \right)^{\mathrm{T}} \boldsymbol{D}_k^2 \boldsymbol{d}_j = \sum_{i=0}^{j-1} \lambda_i \boldsymbol{d}_i^{\mathrm{T}} \boldsymbol{D}_k^2 \boldsymbol{d}_j > 0.$$

进而

$$\begin{aligned} \| \boldsymbol{D}_k \boldsymbol{v}_{j+1} \|_2^2 = \boldsymbol{v}_{j+1}^{\mathrm{T}} \boldsymbol{D}_k^2 \boldsymbol{v}_{j+1} = (\boldsymbol{v}_j + \lambda_j \boldsymbol{d}_j)^{\mathrm{T}} \boldsymbol{D}_k^2 (\boldsymbol{v}_j + \lambda_j \boldsymbol{d}_j) = \\ \boldsymbol{v}_j^{\mathrm{T}} \boldsymbol{D}_k^2 \boldsymbol{v}_j + 2\lambda_j \boldsymbol{v}_j^{\mathrm{T}} \boldsymbol{D}_k^2 \boldsymbol{d}_j + \lambda_j^2 \boldsymbol{d}_j^{\mathrm{T}} \boldsymbol{D}_k^2 \boldsymbol{d}_j > \\ \| \boldsymbol{D}_k \boldsymbol{v}_j \|_2^2. \end{aligned}$$

因此, 不等式 (22) 成立.

引理 3.3　二次函数 $m_k(\boldsymbol{p}) = \| \boldsymbol{F}'(\boldsymbol{x}_k) \boldsymbol{p} + \boldsymbol{F}(\boldsymbol{x}_k) \|^2 / 2$ 是单调下降的, 即

$$\| \boldsymbol{F}'(\boldsymbol{x}_k) \boldsymbol{v}_{j+1} + \boldsymbol{F}(\boldsymbol{x}_k) \| < \| \boldsymbol{F}'(\boldsymbol{x}_k) \boldsymbol{v}_j + \boldsymbol{F}(\boldsymbol{x}_k) \|, \qquad 1 < j \leqslant m. \tag{23}$$

证明　若 $\boldsymbol{r}_j \neq \boldsymbol{0}$, 由算法第 5) 步和引理 3.1, 可得

$$\begin{aligned} 2(\boldsymbol{F}_k'^{\mathrm{T}} \boldsymbol{F}_k)^{\mathrm{T}} \boldsymbol{d}_j + \theta_j^2 \boldsymbol{r}_j^{\mathrm{T}} \boldsymbol{y}_j = \boldsymbol{r}_1^{\mathrm{T}} \boldsymbol{d}_j + \boldsymbol{r}_1^{\mathrm{T}} \boldsymbol{d}_j + \theta_j \boldsymbol{r}_j^{\mathrm{T}} \theta_j \boldsymbol{y}_j = \\ \boldsymbol{r}_1^{\mathrm{T}} \boldsymbol{d}_j + \boldsymbol{r}_1^{\mathrm{T}} \boldsymbol{d}_j + \theta_j \boldsymbol{r}_j^{\mathrm{T}} (\beta_{j-1} \boldsymbol{d}_{j-1} - \boldsymbol{d}_j) = \\ \boldsymbol{r}_1^{\mathrm{T}} \boldsymbol{d}_j + \boldsymbol{d}_j^{\mathrm{T}} (\boldsymbol{r}_1 - \theta_j \boldsymbol{r}_j) = \\ \boldsymbol{r}_1^{\mathrm{T}} \boldsymbol{d}_j + \boldsymbol{d}_j^{\mathrm{T}} (\boldsymbol{r}_1 - \boldsymbol{r}_1 - \boldsymbol{F}_k'^{\mathrm{T}} \boldsymbol{F}_k' \boldsymbol{v}_j) = \\ \boldsymbol{r}_1^{\mathrm{T}} \boldsymbol{d}_j + \boldsymbol{d}_j^{\mathrm{T}} \left(\boldsymbol{r}_1 - \boldsymbol{r}_1 - \sum_{i=0}^{j-1} \lambda_i \boldsymbol{F}_k'^{\mathrm{T}} \boldsymbol{F}_k' \boldsymbol{d}_i \right) = \end{aligned}$$

$$\boldsymbol{r}_1^{\mathrm{T}}\boldsymbol{d}_j - \boldsymbol{d}_j^{\mathrm{T}}\sum_{i=0}^{j-1}\lambda_i \boldsymbol{F}_k^{'\mathrm{T}}\boldsymbol{F}_k'\boldsymbol{d}_i =$$
$$\boldsymbol{r}_1^{\mathrm{T}}\boldsymbol{d}_j = -\boldsymbol{d}_1^{\mathrm{T}}\boldsymbol{D}_k^2\boldsymbol{d}_j < 0.$$

由 $\theta_j^2\boldsymbol{r}_j^{\mathrm{T}}\boldsymbol{y}_j = \theta_j^2\boldsymbol{r}_j^{\mathrm{T}}\boldsymbol{M}_k^{-1}\boldsymbol{r}_j > 0$ 和 $\boldsymbol{\omega}_j^{\mathrm{T}}\boldsymbol{\omega}_j > 0$, 得

$$\|\boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{v}_{j+1} + \boldsymbol{F}(\boldsymbol{x}_k)\|^2 - \|\boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{v}_j + \boldsymbol{F}(\boldsymbol{x}_k)\|^2 =$$
$$2(m_k(\boldsymbol{v}_{j+1}) - m_k(\boldsymbol{v}_j)) =$$
$$2\lambda_j(\boldsymbol{F}_k^{'\mathrm{T}}\boldsymbol{F}_k)^{\mathrm{T}}\boldsymbol{d}_j + \left(\sum_{i=0}^{j}\lambda_i\boldsymbol{d}_i\right)^{\mathrm{T}}\boldsymbol{F}_k^{'\mathrm{T}}\boldsymbol{F}_k'\sum_{i=0}^{j}\lambda_i\boldsymbol{d}_i -$$
$$\left(\sum_{i=0}^{j-1}\lambda_i\boldsymbol{d}_i\right)^{\mathrm{T}}\boldsymbol{F}_k^{'\mathrm{T}}\boldsymbol{F}_k'\sum_{i=0}^{j-1}\lambda_i\boldsymbol{d}_i =$$
$$2\lambda_j(\boldsymbol{F}_k^{'\mathrm{T}}\boldsymbol{F}_k)^{\mathrm{T}}\boldsymbol{d}_j + \lambda_j^2\boldsymbol{\omega}_j^{\mathrm{T}}\boldsymbol{\omega}_j =$$
$$\frac{\theta_j^2\boldsymbol{r}_j^{\mathrm{T}}\boldsymbol{y}_j(2(\boldsymbol{F}_k^{'\mathrm{T}}\boldsymbol{F}_k)^{\mathrm{T}}\boldsymbol{d}_j + \theta_j^2\boldsymbol{r}_j^{\mathrm{T}}\boldsymbol{y}_j)}{\boldsymbol{\omega}_j^{\mathrm{T}}\boldsymbol{\omega}_j} < 0.$$

因此,不等式(23)成立.

引理 3.4 假设 $\|\boldsymbol{D}_k^{-1}\boldsymbol{F}_k^{'\mathrm{T}}\boldsymbol{F}_k\| \neq 0$.

(a) $\boldsymbol{v}_j$ 由算法第 5)步得到,则它有限步满足式(11).

(b) 若 $\boldsymbol{p}_k = \boldsymbol{v}_j$ 是满足式(11)的一个下降方向,则可找到一个合适的步长 α_k, 使其满足式(12).

证明(a) 由算法第 5)步,可得

$$\|\boldsymbol{F}(\boldsymbol{x}_k)\|^2 - \|\boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{v}_2 + \boldsymbol{F}(\boldsymbol{x}_k)\|^2 =$$
$$2\left(-\lambda_1(\boldsymbol{F}_k^{'\mathrm{T}}\boldsymbol{F}_k)^{\mathrm{T}}\boldsymbol{d}_1 - \frac{1}{2}\lambda_1^2\boldsymbol{\omega}_1^{\mathrm{T}}\boldsymbol{\omega}_1\right) =$$
$$2\lambda_1\left((\boldsymbol{F}_k^{'\mathrm{T}}\boldsymbol{F}_k)^{\mathrm{T}}\boldsymbol{D}_k^{-2}\boldsymbol{F}_k^{'\mathrm{T}}\boldsymbol{F}_k - \frac{1}{2}(\boldsymbol{F}_k^{'\mathrm{T}}\boldsymbol{F}_k)^{\mathrm{T}}\boldsymbol{D}_k^{-2}\boldsymbol{F}_k^{'\mathrm{T}}\boldsymbol{F}_k\right) =$$
$$\lambda_1\|\boldsymbol{D}_k^{-1}\boldsymbol{F}_k^{'\mathrm{T}}\boldsymbol{F}_k\|^2 > 0. \tag{24}$$

由式(23),知

$$\|\boldsymbol{F}(\boldsymbol{x}_k)\| - \|\boldsymbol{F}(\boldsymbol{x}_k) + \boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{v}_i\| > \|\boldsymbol{F}(\boldsymbol{x}_k)\| - \|\boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{v}_2 + \boldsymbol{F}(\boldsymbol{x}_k)\| > 0,$$

则

$$\|\boldsymbol{F}(\boldsymbol{x}_k)\| > \|\boldsymbol{F}(\boldsymbol{x}_k) + \boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{v}_i\|.$$

令

$$\xi_i = \frac{\|\boldsymbol{F}(\boldsymbol{x}_k) + \boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{v}_i\|}{\|\boldsymbol{F}(\boldsymbol{x}_k)\|},$$

得到

$$\xi_i = \frac{\|\boldsymbol{F}(\boldsymbol{x}_k) + \boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{v}_i\|}{\|\boldsymbol{F}(\boldsymbol{x}_k)\|} \to 0, \qquad i \to \infty.$$

此外,由 $\eta_k \in (0,1)$ 是一个控制序列,则存在 ξ_j,使得 $\xi_j \leqslant \eta_k$, 即

$$\xi_j = \frac{\|\boldsymbol{F}(\boldsymbol{x}_k) + \boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{v}_j\|}{\|\boldsymbol{F}(\boldsymbol{x}_k)\|} \leqslant \eta_k. \tag{25}$$

因此

$$\|\boldsymbol{F}(\boldsymbol{x}_k) + \boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{v}_j\| \leqslant \eta_k\|\boldsymbol{F}(\boldsymbol{x}_k)\|, \tag{26}$$

所以结论(a)成立.

证明(b)　令

$$\eta \equiv \frac{\| \boldsymbol{F}(\boldsymbol{x}_k) + \boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{v}_j \|}{\| \boldsymbol{F}(\boldsymbol{x}_k) \|},\ \epsilon \equiv \frac{(1-\mu)(1-\eta)\| \boldsymbol{F}(\boldsymbol{x}_k) \|}{\| \boldsymbol{v}_j \|},$$

$$\eta_{\min} \equiv \max\left\{ \eta, 1 - \frac{(1-\eta)\delta}{\| \boldsymbol{v}_j \|} \right\},$$

其中 $\delta > 0$ 是一个充分小的量,若 $\| \alpha \boldsymbol{v}_j \| \leqslant \delta$, 可知

$$\| \boldsymbol{F}(\boldsymbol{x}_k + \alpha \boldsymbol{v}_j) - \boldsymbol{F}(\boldsymbol{x}_k) - \boldsymbol{F}'(\boldsymbol{x}_k)\alpha \boldsymbol{v}_j \| \leqslant \epsilon \| \alpha \boldsymbol{v}_j \|. \tag{27}$$

给定一个 $\eta_k \in [\eta_{\min}, 1)$.

令 $$\boldsymbol{p}_k' = \frac{1-\eta_k}{1-\eta}\boldsymbol{v}_j,$$

因为

$$\| \boldsymbol{p}_k' \| = \frac{1-\eta_k}{1-\eta}\| \boldsymbol{v}_j \| \leqslant \frac{1-\eta_{\min}}{1-\eta}\| \boldsymbol{v}_j \| \leqslant \delta, \tag{28}$$

且

$$\begin{aligned} &\| \boldsymbol{F}(\boldsymbol{x}_k) + \boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{p}_k' \| \leqslant \\ &\quad \frac{\eta_k - \eta}{1-\eta}\| \boldsymbol{F}(\boldsymbol{x}_k) \| + \frac{1-\eta_k}{1-\eta}\| \boldsymbol{F}(\boldsymbol{x}_k) + \boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{v}_j \| = \\ &\quad \frac{\eta_k - \eta}{1-\eta}\| \boldsymbol{F}(\boldsymbol{x}_k) \| + \frac{1-\eta_k}{1-\eta}\eta \| \boldsymbol{F}(\boldsymbol{x}_k) \| = \\ &\quad \eta_k \| \boldsymbol{F}(\boldsymbol{x}_k) \|, \end{aligned} \tag{29}$$

从而

$$\begin{aligned} &\| \boldsymbol{F}(\boldsymbol{x}_k + \boldsymbol{p}_k') \| \leqslant \\ &\quad \| \boldsymbol{F}(\boldsymbol{x}_k + \boldsymbol{p}_k') - \boldsymbol{F}(\boldsymbol{x}_k) - \boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{p}_k' \| + \| \boldsymbol{F}(\boldsymbol{x}_k) + \boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{p}_k' \| \leqslant \\ &\quad \epsilon \frac{1-\eta_k}{1-\eta}\| \boldsymbol{v}_j \| + \eta_k \| \boldsymbol{F}(\boldsymbol{x}_k) \| \leqslant \\ &\quad (1-\mu)(1-\eta_k)\| \boldsymbol{F}(\boldsymbol{x}_k) \| + \eta_k \| \boldsymbol{F}(\boldsymbol{x}_k) \| = \\ &\quad [1 + \mu(\eta_k - 1)]\| \boldsymbol{F}(\boldsymbol{x}_k) \|, \end{aligned} \tag{30}$$

因此令 $\alpha_k = (1-\eta_k)/(1-\eta)$,再结合 $\| \boldsymbol{F}(\boldsymbol{x}_k) \| \leqslant \| \boldsymbol{F}(\boldsymbol{x}_{l(k)}) \|$, 结论(b)也成立.

4　整体和局部收敛性分析

在分析整体收敛性之前,需作如下假设:

假设1　$f: \mathfrak{R}^n \to \mathfrak{R}$ 连续可微且有界. 设取 $\boldsymbol{x}_0 \in \mathfrak{R}^n$, f 的水平集 $\mathcal{L}(\boldsymbol{x}_0) = \{ \boldsymbol{x} \in \mathfrak{R}^n \mid f(\boldsymbol{x}) \leqslant f(\boldsymbol{x}_0), \boldsymbol{l} \leqslant \boldsymbol{x} \leqslant \boldsymbol{u} \}$ 是紧的,算法产生的迭代序列 $\{ \boldsymbol{x}_k \} \subseteq \mathcal{L}(\boldsymbol{x}_0)$.

假设2　$\| \boldsymbol{F}_k'^{\mathrm{T}} \boldsymbol{F}_k' \|$ 和 $\| \boldsymbol{D}_k^{-1} \boldsymbol{F}_k'^{\mathrm{T}} \|$ 一致有界,存在 κ_{F} 和 κ_{d} 使得对所有的 k, 均有

$$\| \boldsymbol{F}_k'^{\mathrm{T}} \boldsymbol{F}_k' \| \leqslant \kappa_{\mathrm{F}},\ \| \boldsymbol{D}_k^{-1} \boldsymbol{F}_k'^{\mathrm{T}} \| \leqslant \kappa_{\mathrm{d}}.$$

引理4.1　设 $\boldsymbol{F}_k'$ 可逆且算法没有终止,假设存在一个 η,使得 $1-\eta_k > 1-\eta$,则存在 $\tau > 0$ 使得 $\| \boldsymbol{p}_k \| \leqslant \tau(1-\eta)\| \boldsymbol{F}_k \|$.

证明　令 $\| \boldsymbol{F}_k'^{-1} \| \leqslant K_1$, $\kappa(\boldsymbol{D}_k) \overset{\mathrm{def}}{=} \| \boldsymbol{D}_k^{-1} \| \| \boldsymbol{D}_k \|$ 表示矩阵 $\boldsymbol{D}_k$ 的条件数,并设 $\kappa(\boldsymbol{D}_k) < K_2$. 由 $\bar{\boldsymbol{p}}_k^N = -\boldsymbol{D}_k \boldsymbol{F}_k'^{-1} \boldsymbol{F}_k$,得到 $\boldsymbol{p}_k^N = -\boldsymbol{F}_k'^{-1} \boldsymbol{F}_k$. 再利用引理 3.2,可得

$$\| \boldsymbol{D}_k \boldsymbol{p}_k \| \leqslant \| -\boldsymbol{D}_k \boldsymbol{F}_k'^{-1} \boldsymbol{F}_k \|, \tag{31}$$

从而

$$\begin{aligned}\|\boldsymbol{p}_k\| \leqslant \|\boldsymbol{D}_k^{-1}\| \|\boldsymbol{D}_k\| \|\boldsymbol{F}_k'^{-1}\| \|\boldsymbol{F}_k\| = \\ \kappa(\boldsymbol{D}_k)\|\boldsymbol{F}_k'^{-1}\| \|\boldsymbol{F}_k\| \leqslant K_1K_2\|\boldsymbol{F}_k\|,\end{aligned} \tag{32}$$

所以当取 $\tau = K_1K_2/(1-\eta)$ 时，$\|\boldsymbol{p}_k\| \leqslant \tau(1-\eta)\|\boldsymbol{F}_k\|$ 成立.

定理 4.1 假设 1、2 成立和 $\eta_k \to 0$，设 $\{\boldsymbol{x}_k\} \subset \Omega$ 为算法产生的迭代序列，则

$$\lim_{k\to\infty}\|\boldsymbol{D}_k^{-1}\boldsymbol{F}_k'^{\mathrm{T}}\boldsymbol{F}_k\| = 0. \tag{33}$$

证明 因为 $m(k+1) \leqslant m(k)+1$ 和 $\|\boldsymbol{F}(\boldsymbol{x}_{k+1})\| \leqslant \|\boldsymbol{F}(\boldsymbol{x}_{l(k)})\|$，所以

$$\begin{aligned}\|\boldsymbol{F}(\boldsymbol{x}_{l(k+1)})\| = \max_{0\leqslant j\leqslant m(k+1)}\|\boldsymbol{F}(\boldsymbol{x}_{k+1-j})\| \leqslant \\ \max_{0\leqslant j\leqslant m(k)+1}\|\boldsymbol{F}(\boldsymbol{x}_{k+1-j})\| = \max_{-1\leqslant j\leqslant m(k)}\|\boldsymbol{F}(\boldsymbol{x}_{k-j})\| = \|\boldsymbol{F}(\boldsymbol{x}_{l(k)})\|.\end{aligned}$$

可知对所有的 k，$\{\|\boldsymbol{F}(\boldsymbol{x}_{l(k)})\|\}$ 是单调递减的，因此 $\|\boldsymbol{F}(\boldsymbol{x}_{l(k)})\|$ 收敛. 由式(12)和引理 4.1，可得

$$\|\boldsymbol{F}(\boldsymbol{x}_k+\alpha_k\boldsymbol{p}_k)\| - \|\boldsymbol{F}(\boldsymbol{x}_{l(k)})\| \leqslant -\mu(1-\eta_k)\|\boldsymbol{F}_k\| \leqslant -\mu\tau^{-1}\|\boldsymbol{p}_k\|. \tag{34}$$

由式(34)，知

$$\begin{aligned}\|\boldsymbol{F}(\boldsymbol{x}_{l(k)})\| = \|\boldsymbol{F}(\boldsymbol{x}_{l(k)-1}+\alpha_{l(k)-1}\boldsymbol{p}_{l(k)-1})\| \leqslant \\ \max_{0\leqslant j\leqslant m(l(k)-1)}\|\boldsymbol{F}(\boldsymbol{x}_{l(l(k)-1-j)})\| - \frac{\mu}{\tau}\|\boldsymbol{p}_{l(k)-1}\| \leqslant \\ \|\boldsymbol{F}(\boldsymbol{x}_{l(l(k)-1)})\| - \frac{\mu}{\tau}\|\boldsymbol{p}_{l(k)-1}\|.\end{aligned}$$

由 $\eta_k \to 0$ 和 $K_1K_2/(1-\eta_k) < K_1K_2/(1-\eta)$，可知 μ/τ 不趋近于0. 因而 $\lim\limits_{k\to\infty}\|\boldsymbol{p}_{l(k)-1}\| = 0$. 类似于文献[7]的证明，有

$$\lim_{k\to\infty}\|\boldsymbol{p}_k\| = 0,\ \lim_{k\to\infty}\|\boldsymbol{F}(\boldsymbol{x}_{l(k)})\| = \lim_{k\to\infty}\|\boldsymbol{F}(\boldsymbol{x}_k)\|. \tag{35}$$

结合式(34)，得到

$$\lim_{k\to\infty}\|\boldsymbol{F}(\boldsymbol{x}_k)\| = 0,\ \|\boldsymbol{D}_k^{-1}\boldsymbol{F}_k'^{\mathrm{T}}\boldsymbol{F}_k\| \leqslant \|\boldsymbol{D}_k^{-1}\boldsymbol{F}_k'^{\mathrm{T}}\| \|\boldsymbol{F}_k\|.$$

再利用假设 2，可得到

$$\lim_{k\to\infty}\|\boldsymbol{D}_k^{-1}\boldsymbol{F}_k'^{\mathrm{T}}\boldsymbol{F}_k\| = 0,$$

因此结论成立.

假设 3 类似文献[8]中的假设：存在 $\omega_k \geqslant \omega_l$ 和 $\sigma_k > 0$，满足

$$\frac{|\phi_k^j|}{|p_k^j|} \geqslant \omega_k\frac{|(\boldsymbol{F}_k'^{\mathrm{T}}\boldsymbol{F}_k)^j| + \sigma_k}{|(\boldsymbol{F}_k'^{\mathrm{T}}\boldsymbol{F}_k)^j + (\boldsymbol{F}_k'^{\mathrm{T}}\boldsymbol{F}_k'\boldsymbol{p}_k)^j|},$$

其中 $\phi_k^j(x)$ 由式(5)定义，$\omega_l > 0$ 是一个常数.

引理 4.2[9] 对任意的 $\boldsymbol{x}$ 和 $\epsilon > 0$，存在 $\delta > 0$ 使得

$$\|\boldsymbol{F}(\boldsymbol{z}) - \boldsymbol{F}(\boldsymbol{y}) - \boldsymbol{F}'(\boldsymbol{y})(\boldsymbol{z}-\boldsymbol{y})\| \leqslant \epsilon\|\boldsymbol{z}-\boldsymbol{y}\|, \tag{36}$$

其中 $\boldsymbol{y},\boldsymbol{z} \in \aleph_\delta(\boldsymbol{x}) = \{\boldsymbol{u} \mid \|\boldsymbol{u}-\boldsymbol{x}\| \leqslant \delta\}$.

定理 4.2 假设 $\boldsymbol{F}_k'^{\mathrm{T}}\boldsymbol{F}_k'$ 正定和 $\boldsymbol{F}_k'$ 可逆，假设 1 ~ 3 成立. $\omega_k \leqslant 1$，$\omega_k - 1 = O(\|\boldsymbol{p}_k\|)$. $\{\boldsymbol{x}_k\} \subset \Omega$ 由算法产生的序列，并且 $\{\boldsymbol{x}_k\}$ 收敛到 $\boldsymbol{x}_*$ 和 $\eta_k \to 0$，则 $\{\boldsymbol{x}_k\}$ 超线性收敛于 $\boldsymbol{x}_*$，即

$$\lim_{k\to\infty}\frac{\|\boldsymbol{x}_{k+1}-\boldsymbol{x}_*\|}{\|\boldsymbol{x}_k-\boldsymbol{x}_*\|} = 0. \tag{37}$$

证明 由文献[3]中引理3.1 的证明知，当 k 充分大时，算法得到 Newton 步. 下面证明当 k 充分大时，Newton 步

$$\bar{\boldsymbol{p}}_k = -(\boldsymbol{D}_k^{-1}\boldsymbol{F}_k'^{\mathrm{T}}\boldsymbol{F}_k'\boldsymbol{D}_k^{-1})^{-1}(\boldsymbol{D}_k^{-1}\boldsymbol{F}_k'^{\mathrm{T}}\boldsymbol{F}_k)$$

满足式(11)~(13).

首先证明 k 充分大时,$\bar{\boldsymbol{p}}_k$ 满足式(11).当 $\eta_k \to 0$, 可知

$$\begin{aligned}&\|\boldsymbol{F}(\boldsymbol{x}_k)+\boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{p}_k\| = \|\boldsymbol{F}(\boldsymbol{x}_k)+\boldsymbol{F}'(x_k)\boldsymbol{D}_k^{-1}\bar{\boldsymbol{p}}_k\| = \\ &\quad \|\boldsymbol{F}(\boldsymbol{x}_k)-\boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{D}_k^{-1}\boldsymbol{D}_k\boldsymbol{F}_k'^{-1}\boldsymbol{F}_k\| = \\ &\quad 0 \leqslant \eta_k\|\boldsymbol{F}(\boldsymbol{x}_k)\|.\end{aligned} \tag{38}$$

因此 $\bar{\boldsymbol{p}}_k$ 满足式(11).

其次证明 $\bar{\boldsymbol{p}}_k$ 满足式(12).令

$$\epsilon_k \equiv \frac{(1-\mu)(1-\eta_k)\|\boldsymbol{F}(\boldsymbol{x}_k)\|}{\|\boldsymbol{p}_k^N\|}, \tag{39}$$

由式(35),得到

$$\lim_{k\to\infty}\|\bar{\boldsymbol{p}}_k\| = \lim_{k\to\infty}\|\boldsymbol{p}_k\| = 0. \tag{40}$$

再由引理 4.2 和式(39),知

$$\begin{aligned}&\|\boldsymbol{F}(\boldsymbol{x}_k+\boldsymbol{D}_k^{-1}\bar{\boldsymbol{p}}_k)\| = \\ &\quad \|\boldsymbol{F}(\boldsymbol{x}_k+\boldsymbol{p}_k^N)\| \leqslant \\ &\quad \|\boldsymbol{F}(\boldsymbol{x}_k+\boldsymbol{p}_k^N)-\boldsymbol{F}(\boldsymbol{x}_k)-\boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{p}_k^N\| + \|\boldsymbol{F}(\boldsymbol{x}_k)+\boldsymbol{F}'(\boldsymbol{x}_k)\boldsymbol{p}_k^N\| \leqslant \\ &\quad \epsilon_k\|\boldsymbol{p}_k^N\| + \eta_k\|\boldsymbol{F}(\boldsymbol{x}_k)\| = \\ &\quad (1-\mu)(1-\eta_k)\|\boldsymbol{F}(\boldsymbol{x}_k)\| + \eta_k\|\boldsymbol{F}(\boldsymbol{x}_k)\| = \\ &\quad [1+\mu(\eta_k-1)]\|\boldsymbol{F}(\boldsymbol{x}_k)\|,\end{aligned} \tag{41}$$

结合 $\|\boldsymbol{F}(\boldsymbol{x}_k)\| \leqslant \|\boldsymbol{F}(\boldsymbol{x}_{l(k)})\|$,当 $\alpha \equiv 1$ 和k 充分大时,$\bar{\boldsymbol{p}}_k$ 满足式(12).

最后,证明 $\bar{\boldsymbol{p}}_k$ 也满足式(13).

由式(40),有

$$\lim_{k\to\infty} p_k^j = 0. \tag{42}$$

此外,存在一个上标 j 使得

$$\alpha_k^* = \min\left\{\max\left\{\frac{l^i-x_*^i}{p_*^i},\frac{u^i-x_*^i}{p_*^i}\right\}, i=1,\cdots,n\right\} = \max\left\{\frac{l^j-x_*^j}{p_*^j},\frac{u^j-x_*^j}{p_*^j}\right\}. \tag{43}$$

当 $l^j < x_*^j < u^j$,再结合式(42),可知$\lim\limits_{k\to\infty}\alpha_k^* = \infty$.否则,不失一般性,假设 $x_*^j = u^j$.考虑下面两种情况:

1) 如果 $p_k^j < 0$,

$$\max\left\{\frac{l^j-x_k^j}{p_k^j},\frac{u^j-x_k^j}{p_k^j}\right\} = \frac{l^j-x_k^j}{p_k^j}.$$

由式(42),可得 $\lim\limits_{k\to\infty}\alpha_k^* = \infty$.

2) 如果 $p_k^j > 0$,类似于文献[8]的证明,当 $x_*^j = u^j$,知 $g_*^j < 0$.从而对于 k 充分大时,由 g_k^j 得连续性,得到 $g_k^j < g_*^j < 0$.由 $\phi^j(x)$ 的定义,取 $\phi^j(x) = u^j - x^j$, 再利用假设 3,得到

$$\max\left\{\frac{l^j-x_k^j}{p_k^j},\frac{u^j-x_k^j}{p_k^j}\right\} = \frac{u^j-x_k^j}{p_k^j} = \frac{|\phi_k^j|}{|p_k^j|} \geqslant \omega_k\frac{|(\boldsymbol{F}_k'^{\mathrm{T}}\boldsymbol{F}_k)^j|+\sigma_k}{|(\boldsymbol{F}_k'^{\mathrm{T}}\boldsymbol{F}_k)^j+(\boldsymbol{F}_k'^{\mathrm{T}}\boldsymbol{F}'_k\boldsymbol{p}_k)^j|},$$

又 $\omega_k - 1 = O(\|\boldsymbol{p}_k\|)$, 所以

$$\lim_{k\to\infty}\max\left\{\frac{l^j - x_k^j}{p_k^j},\frac{u^j - x_k^j}{p_k^j}\right\} \geqslant 1. \tag{44}$$

由 $\vartheta_k - 1 = O(\|\boldsymbol{p}_k\|)$ 和式(42),得到$\lim\limits_{k\to\infty}\vartheta_k = 1$. 因此$\lim\limits_{k\to\infty}\alpha_k = 1$ 且 $\bar{\boldsymbol{p}}_k$ 满足式(13).

由上面的讨论可知,假如 $\boldsymbol{F}_k'^{\mathrm{T}}\boldsymbol{F}_k'$ 正定,新的迭代步 $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{p}_k$, 其中

$$\boldsymbol{p}_k = \boldsymbol{D}_k^{-1}\bar{\boldsymbol{p}}_k = -(\boldsymbol{F}_k'^{\mathrm{T}}\boldsymbol{F}_k')^{-1}\boldsymbol{F}_k'^{\mathrm{T}}\boldsymbol{F}_k$$

是 Newton 步,因此式(37)成立.

定理4.2 说明算法的局部收敛速率依赖于目标函数在 $\boldsymbol{x}_*$ 的 Hess 矩阵和步长的局部收敛速率. 如果 $\boldsymbol{p}_k$ 变为 Newton 步,算法产生的序列 $\{\boldsymbol{x}_k\}$ 二次收敛于 $\boldsymbol{x}_*$.

5 数值结果

本节给出一些数值结果(见表1). 为了检验算法的有效性,选取参数如下: $\epsilon = 10^{-6}$, $\mu = 0.5$, $\omega = 0.5$, $\gamma = 1$. 给出的7道题目选自文献[10]和文献[11]. N_{F} 和 N_{G} 分别表示目标函数次数和梯度函数次数, M 代表非单调参数. 数值结果表明了算法的有效性.

表1 数值结果

题　目	$M = 0$		$M = 5$	
	N_{G}	N_{F}	N_{G}	N_{F}
Ferraris-Tronconi	6	6	6	6
Himmelblau 函数	27	32	26	31
Reklaitis Ragsdell 问题	109	109	109	109
SC205	14	15	13	13
SC212	234	234	186	187
SC311	46	52	11	16
SC312	80	93	16	27

参考文献:

[1] Coleman T F, Li Y. An interior trust-region approach for nonlinear minimization subject to bounds[J]. *SIAM J Optim*, 1996, **6**(2): 418-445.

[2] Bellavia S, Macconi M, Morini B. An affine scaling trust-region approach to bound-constrained nonlinear systems[J]. *Appl Numer Math*, 2003, **44**(3): 257-280.

[3] Jia C A, Zhu D T. An affine scaling interior algorithm via Lanczos path for solving bound-constrained nonlinear systems[J]. *Applied Mathematics and Computation*, 2008, **195**(2): 558-575.

[4] Dembo R S, Eisenstat S C, Steihaug T. Inexact Newton methods[J]. *SIAM J Numer Anal*, 1982, **19**(2): 400-408.

[5] Shen W P, Li C. Kantorovich-type convergence criterion for inexact Newton methods[J]. *Appl Numer Math*, 2009, **59**(7): 1599-1611.

[6] Gould Ni I M, Lucidi S, Roma M, Toint P H L. Solving the trust-region subproblem using the Lanczos method[J]. *SIAM Journal on Optimization*, 1999, **9**(2): 504-525.

[7] Gripp R S, Lampariello F, Lucidi S. A nonmonotone line search technique for Newton's methods[J]. *SIAM J Numer Anal*, 1986, **23**(4): 707-716.

[8] Guo P H, Zhu D T. A nonmonotonic reduced projected Hessian method via an affine scaling

interior modified gradient path for bounded-constrained optimization[J]. *Journal of Systems Science and Complexity*, 2008, **21**(1): 85-113.

[9] Ortega J M, Rheinboldt W C. *Iterative Solution of Nonlinear Equations in Several Variables* [M]. New York: Academic Press, 1970.

[10] Floudas C A, Pardalos P M. *Handbook of Test Problems in Local and Global Optimization* [M]. Dordrecht: KluwTer Academic, 1999.

[11] Schittkowski K. *More Test Examples for Nonlinear Programming Codes, Lecture Notes in Economics and Mathematical Systems*[M]. Heidelberg, Berlin: Springer-Verlag, 1981.

Inexact Newton Method via Lanczos Decomposed Technique for Solving Box-Constrained Nonlinear Systems

ZHANG Yong[1], ZHU De-tong[2]

(1. *Mathematics and Science College*, *Shanghai Normal University*, *Shanghai* 200234, *P. R. China*;

2. *Business College*, *Shanghai Normal University*, *Shanghai* 200234, *P. R. China*)

Abstract: An inexact Newton method via Lanczos decomposed technique was proposed for solving the box-constrained nonlinear systems. The iterative direction was obtained by solving an affine scaling quadratic model with Lanczos decomposed technique. By using the interior backtracking line search technique, the acceptable trial step length along this direction will be found. The global convergence and fast local convergence rate of the proposed algorithm were established under some reasonable conditions. Furthermore, the results of the numerical experiments are reported to show the effectiveness of the proposed algorithm.

Key words: nonlinear systems; Lanczos decomposed technique; inexact Newton method; nonmonotonic technique