文章编号:1000-0887(2013)11-1216-09

非线性不等式系统带折线步的信赖域方法及其收敛性*

何郁波, 林晓艳

(怀化学院 数学与应用数学系, 湖南 怀化 418008)

摘要: 针对一类非线性不等式系统求解的问题,利用一系列目标函数二次可微的带参数优化问题来逐次逼近非线性不等式系统的解,从而提出了针对参数最优化问题带折线步的信赖域算法.在较弱的条件下,算法的全局收敛性得到了保证.数值试验显示算法有效.

关 键 词: 非线性不等式系统; 信赖域算法; Cauchy 点; 折线步; 全局收敛

中图分类号: O178.2 **文献标志码:** A

DOI:10.3879/j.issn.1000-0887.2013.11.011

1 预备知识

本文求解非线性不等式系统(nonlinear inequalities system, NIS):

$$\boldsymbol{F}(\boldsymbol{x}) \leqslant 0, \qquad \boldsymbol{x} \in R^n, \tag{1}$$

其中 $\boldsymbol{F}(\boldsymbol{x}) = (F_1(\boldsymbol{x}), F_2(\boldsymbol{x}), \cdots, F_m(\boldsymbol{x}))^{\mathrm{T}}$ 且 $F_i: R^n \to \mathbf{R}, i \in I = \{1,2,\cdots,m\}$ 二阶连续可微.

非线性不等式系统频繁出现于集合分离、数据分离、计算机辅助设计和图像处理等领域.特别地,在求解约束最优化问题中,可行方向法是一种重要算法.其典型策略是从可行点出发,沿着下降的可行方向进行搜索,从而求出使目标函数值下降的新的可行点.可行方向法的初始可行点的确定需要通过求解非线性不等式系统而获得.

目前,非线性不等式系统的求解方法并不多见.针对凸函数构成的不等式系统,文献[1]结合 SQP(sequential quadratic programming)算法和扰动技术,提出了一类有限步收敛算法.文献[2]和[3]中借助于投影技术和广义投影法,给出了求解非线性不等式系统的新算法,在该算法中,搜索方向仅由两个显式表达式即可产生,从而使计算量相对减少;文献[4]采用目标函数为复合函数的信赖域算法;文献[5]利用光滑辅助函数将非线性不等式系统转化成最小二乘问题后再采用光滑 Gauss-Newton 算法逐次逼近最小二乘问题的解;文献[6]和[7]针对非线性不等式系统分别采用广义投影技术和 Levenberg-Marquardt 方法建立了新的算法,效果也都比较理想.受以上研究成果的启发,本文利用光滑辅助函数将 NIS 系统的求解转化成一系列目标函数二次可微的最优化问题来逼近原问题的解.在求解最优化问题过程中,采用带折线步的

* **收稿日期:** 2013-04-09; **修订日期:** 2013-06-19

基金项目: 湖南省教育厅资助重点项目(08A503);湖南省普通高校青年骨干教师培养基金资助项目(湘教通[2012]510 号)

作者简介: 何郁波,男,讲师,硕士(通讯作者. E-mail: heyinprc@yahoo.com.cn);
林晓艳,女,教授,博士(E-mail: xiaoyanlin98@hotmail.com).

信赖域算法,降低算法的计算量.

若存在光滑函数 $\phi(\cdot,\cdot)$ 满足

$$\phi(\varepsilon,t)=0\Leftrightarrow t\leqslant-\varepsilon,$$

其中 $\varepsilon>0$ 为光滑因子,则问题(1)的求解等价于解光滑方程组

$$\phi(\varepsilon,F_i(\boldsymbol{x}))=0,\qquad i\in I.\tag{2}$$

关于函数 ϕ 的构造,作者在文献[5]中给出了一系列满足条件的函数,本文取为

$$\phi(\varepsilon,\boldsymbol{x})=\frac{1}{2}(\varepsilon+\boldsymbol{x}+\sqrt{\boldsymbol{x}^2+\varepsilon^2})^2,\tag{3}$$

其中 $\varepsilon\geqslant 0$ 为光滑参数.

光滑函数 $\phi(\varepsilon,\boldsymbol{x})$ 有很好的性质,本身是连续可微的,其梯度函数强半光滑,具体性质见文献[5].

引理 1 对 $\forall(\varepsilon,x)\in\mathbf{R}_+\times\mathbf{R}$,函数 $\phi(\varepsilon,x)$ 满足

1) $\phi(\varepsilon,x)\geqslant 0$;

2) $\lim\limits_{\varepsilon\to 0^+}\phi(\varepsilon,x)=\frac{1}{2}x_+^2$;

3) $\phi(\varepsilon,x)=0$, 当且仅当 $\varepsilon=0$ 和 $x\leqslant 0$;

4) 若 $\varepsilon>0$,则 $\phi(\cdot,\cdot)$ 二次连续可微.

根据引理 1 知, $\phi(\varepsilon,F_i(\boldsymbol{x}))\geqslant 0,i\in I$.因此方程组(2)的解可以由最优化问题

$$\min_{\boldsymbol{x}\in R^n,\varepsilon>0}f(\varepsilon,\boldsymbol{x})=\sum_{i=1}^{m}\phi(\varepsilon,F_i(\boldsymbol{x}))\tag{4}$$

求得,接下来本文将针对问题(4)设计带折线步的信赖域算法.

对 $\forall\varepsilon\in\mathbf{R}_+,\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}\in R^n$,记 $\boldsymbol{w}=(\varepsilon,\boldsymbol{x}^{\mathrm{T}})^{\mathrm{T}}\in(\mathbf{R}_+,R^n)$,则问题(4) 目标函数 $f(\boldsymbol{w})$ 在点 $\boldsymbol{w}^k$ 的邻域内二次近似函数为

$$f(\boldsymbol{w})\approx Q(\boldsymbol{w})=f(\boldsymbol{w}^k)+\boldsymbol{g}^{\mathrm{T}}(\boldsymbol{w}^k)(\boldsymbol{w}-\boldsymbol{w}^k)+\frac{1}{2}(\boldsymbol{w}-\boldsymbol{w}^k)^{\mathrm{T}}G(\boldsymbol{w}^k)(\boldsymbol{w}-\boldsymbol{w}^k),\tag{5}$$

其中 $\boldsymbol{g}(\boldsymbol{w})=\nabla f(\boldsymbol{w}),G(\boldsymbol{w})=\nabla\boldsymbol{g}(\boldsymbol{w})$.

引理 2 对 $\forall\boldsymbol{w}=(\varepsilon,\boldsymbol{x})^{\mathrm{T}}\in(\mathbf{R}_+,R^n)$, 有

1) 若 $\|\boldsymbol{g}(\boldsymbol{w})\|=0$, 则 $f(\boldsymbol{w})=0$;

2) 若 $\boldsymbol{x}^*\in\{\boldsymbol{x}\in R^n\mid\|\boldsymbol{g}(\boldsymbol{w})\|=0\}$, 则 $\boldsymbol{x}^*$ 是问题(4)的稳定点.

下一节,将介绍针对问题 (4)的带折线步的信赖域算法.

2 带折线步信赖域算法

设 $\boldsymbol{w}^k$ 为第 k 次迭代点,记 $f_k=f(\boldsymbol{w}^k)$,$\boldsymbol{g}_k=\nabla f(\boldsymbol{w}^k)$,$\boldsymbol{B}_k$ 为 $\nabla\boldsymbol{g}(\boldsymbol{w}^k)$ 的第 k 次近似,则第 k 次迭代步的信赖域子问题有如下的形式:

$$\min_{\boldsymbol{d}\in\mathbf{R}_+\times R^n}m_k(\boldsymbol{d})=f_k+\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}+\frac{1}{2}\boldsymbol{d}^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{d}\qquad \text{s.t. }\|\boldsymbol{d}\|\leqslant\Delta_k,\tag{6}$$

其中 $\Delta_k>0$ 为信赖域半径.

对于子问题(6),其 Cauchy 点为

$$
\boldsymbol{d}_k^{\mathrm{C}}=\begin{cases}-\dfrac{\Delta_k}{\|\boldsymbol{g}_k\|}\boldsymbol{g}_k, & \boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{g}_k\leqslant 0,\\ -\min\left\{\dfrac{\|\boldsymbol{g}_k\|^2}{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{g}_k},\dfrac{\Delta_k}{\|\boldsymbol{g}_k\|}\right\}\boldsymbol{g}_k, & \text{其它.}\end{cases} \tag{7}
$$

为了保证子问题(6)中 $\boldsymbol{B}_k$ 的正定性,本文选取

$$
\boldsymbol{B}_{k+1}=\boldsymbol{B}_k+\tau_k\left[\frac{(\boldsymbol{y}_k-\boldsymbol{B}_k\boldsymbol{d}_k)\boldsymbol{d}_k^{\mathrm{T}}}{\|\boldsymbol{d}_k\|^2}-\frac{\boldsymbol{d}_k^{\mathrm{T}}(\boldsymbol{y}_k-\boldsymbol{B}_k\boldsymbol{d}_k)\boldsymbol{d}_k\boldsymbol{d}_k^{\mathrm{T}}}{\|\boldsymbol{d}_k\|^2}\right], \tag{8}
$$

其中 $\boldsymbol{y}_k=\boldsymbol{g}_{k+1}-\boldsymbol{g}_k$, $|\tau_k-1|\leqslant\tau$, $(\tau\in(0,1))$.

若公式(8)中令 $\tau_k=1$, 则是文献[8]中 PSB 公式的情形.Powell 在文献[8]中证明了由式(8)产生的对称矩阵 $\boldsymbol{B}_k$ 满足

$$
\frac{\|(\nabla^2 f(\boldsymbol{w}^k)-\boldsymbol{B}_k)\boldsymbol{d}_k\|}{\|\boldsymbol{d}_k\|^2}\to 0. \tag{9}
$$

因此,由式(8)产生的近似矩阵 $\boldsymbol{B}_k$ 在每步迭代中满足对称正定的条件.

这样,问题(6)的近似解为

$$
\boldsymbol{d}_k(\lambda)=\begin{cases}-\dfrac{\Delta_k}{\|\boldsymbol{g}_k\|}\boldsymbol{g}_k, & \|\boldsymbol{d}_k^{\mathrm{C}}\|\geqslant\Delta_k,\\ \boldsymbol{d}_k^{\mathrm{N}}=-\boldsymbol{B}_k^{-1}\boldsymbol{g}_k, & \|\boldsymbol{d}_k^{\mathrm{C}}\|<\Delta_k\text{和}\ \|\boldsymbol{d}_k^{\mathrm{N}}\|\leqslant\Delta_k,\\ \lambda\boldsymbol{d}_k^{\mathrm{N}}+(1-\lambda)\boldsymbol{d}_k^{\mathrm{C}}, & \text{其它,}\end{cases} \tag{10}
$$

其中 $\lambda\in[0,1]$.

引理 3 若 $\boldsymbol{B}_k$ 正定,则对 $\forall\lambda\in[0,1]$ 有

(i) $m(\boldsymbol{d}_k(\lambda))$ 关于 λ 递减;

(ii) $m_k(\boldsymbol{d}_k^{\mathrm{N}})\leqslant m(\boldsymbol{d}_k^{\mathrm{C}})\leqslant m_k(0)$.

证明 (i),令

$$
h(\lambda)=m(\boldsymbol{d}_k(\lambda))=
$$

$$
f_k+\boldsymbol{g}_k^{\mathrm{T}}[\lambda\boldsymbol{d}_k^{\mathrm{N}}+(1-\lambda)\boldsymbol{d}_k^{\mathrm{C}}]+\frac{1}{2}[\lambda\boldsymbol{d}_k^{\mathrm{N}}+(1-\lambda)\boldsymbol{d}_k^{\mathrm{C}}]^{\mathrm{T}}\boldsymbol{B}_k[\lambda\boldsymbol{d}_k^{\mathrm{N}}+(1-\lambda)\boldsymbol{d}_k^{\mathrm{C}}],
$$

则

$$
h'(\lambda)=\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k^{\mathrm{N}}-\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k^{\mathrm{C}}+\lambda(\boldsymbol{d}_k^{\mathrm{N}})^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{d}_k^{\mathrm{N}}+\frac{1}{2}(1-2\lambda)(\boldsymbol{d}_k^{\mathrm{N}})^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{d}_k^{\mathrm{C}}+
$$

$$
\frac{1}{2}(1-2\lambda)(\boldsymbol{d}_k^{\mathrm{C}})^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{d}_k^{\mathrm{N}}-(1-\lambda)(\boldsymbol{d}_k^{\mathrm{C}})^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{d}_k^{\mathrm{C}}=
$$

$$
-\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k^{-1}\boldsymbol{g}_k-\frac{\|\boldsymbol{g}_k\|^4}{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{g}_k}+\lambda\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k^{-1}\boldsymbol{g}_k-(1-2\lambda)\frac{\|\boldsymbol{g}_k\|^4}{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{g}_k}+(1-\lambda)\frac{\|\boldsymbol{g}_k\|^4}{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{g}_k}=
$$

$$
(\lambda-1)\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k^{-1}\boldsymbol{g}_k+(\lambda-1)\frac{\|\boldsymbol{g}_k\|^4}{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{g}_k}=
$$

$$
(\lambda-1)\left(\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k^{-1}\boldsymbol{g}_k+\frac{\|\boldsymbol{g}_k\|^4}{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{g}_k}\right)\leqslant 0.
$$

因此 $m(\boldsymbol{d}_k(\lambda))$ 关于 λ 递减.

(ii),由(i)的证明可知

$$h(1)\leqslant h(0),$$

即

$$m_k(\boldsymbol{d}_k^{\mathrm{N}})\leqslant m_k(\boldsymbol{d}_k^{\mathrm{C}}),$$

因此,结论成立.

下面给出带折线步的信赖域算法.

算法 A

(s.0) 给定 $\bar{\Delta}>0$, $\Delta_0\in(0,\bar{\Delta})$, $\boldsymbol{w}_0=(\varepsilon_0,\boldsymbol{x}_0^{\mathrm{T}})^{\mathrm{T}}\in\mathbf{R}_+\times R^n$, $\boldsymbol{B}_0=\boldsymbol{I}$,且 $0<\eta_1<\eta_2<1$, $0<\gamma_1<1<\gamma_2$,$\tau\in(0,1)$,$\varepsilon\geqslant 0$.令 $k=0$.

(s.1) 计算 $\boldsymbol{g}_k=\nabla f(\boldsymbol{w}_k)$,若 $f(\boldsymbol{w}^k)\leqslant\varepsilon$ 或 $\|g(\boldsymbol{w}^k)\|\leqslant\varepsilon$ 成立, 停止.

(s.2) 执行算法 B,得搜索方向 $\boldsymbol{d}_k$.

(s.3) 计算 r_k

$$r_k=\frac{f(\boldsymbol{w}_k)-f(\boldsymbol{w}_k+\boldsymbol{d}_k)}{m_k(0)-m_k(\boldsymbol{d}_k)}. \tag{11}$$

(s.4) 调整信赖域半径 Δ_k

$$\Delta_{k+1}=\begin{cases}\gamma_1\Delta_k, & \text{若 } r_k\leqslant\eta_1,\\ \min\{\gamma_2\Delta_k,\bar{\Delta}\}, & \text{若 } r_k\geqslant\eta_2,\ \|\boldsymbol{d}_k\|=\Delta_k,\\ \Delta_k, & \text{其他.}\end{cases} \tag{12}$$

(s.5) 如果 $r_k>\eta_1$, 令

$$\boldsymbol{w}_{k+1}=\boldsymbol{w}_k+\boldsymbol{d}_k,$$

否则,令

$$\boldsymbol{w}_{k+1}=\boldsymbol{w}_k.$$

利用式(8)更新矩阵 $\boldsymbol{B}_{k+1}$,令 $k=k+1$, 转(s.1).

算法 B

当 $k=0,1,2,\cdots$,计算 $\boldsymbol{d}_k^{\mathrm{C}}=-\min\left\{\frac{\|\boldsymbol{g}_k\|^2}{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{g}_k},\frac{\Delta_k}{\|\boldsymbol{g}_k\|}\right\}\boldsymbol{g}_k$ 和 $\boldsymbol{d}_k^{\mathrm{N}}=-\boldsymbol{B}_k^{-1}\boldsymbol{g}_k$.

若 $\|\boldsymbol{d}_k^{\mathrm{C}}\|\geqslant\Delta_k$,

$$\boldsymbol{d}_k=-\frac{\Delta_k}{\|\boldsymbol{g}_k\|}\boldsymbol{g}_k,$$

再若 $\|\boldsymbol{d}_k^{\mathrm{N}}\|\leqslant\Delta_k$,

$$\boldsymbol{d}_k=-\boldsymbol{B}_k^{-1}\boldsymbol{g}_k,$$

再选择 $\lambda\in[0,1]$ 使得

$$\boldsymbol{d}_k=\lambda\boldsymbol{d}_k^{\mathrm{C}}+(1-\lambda)\boldsymbol{d}_k^{\mathrm{N}}\quad\text{和}\quad\|\boldsymbol{d}_k\|=\Delta_k, \tag{$*$}$$

结束.

注 2.1　算法 B 中,满足($*$)的 λ 存在.

3　收敛性分析

本节将证明算法 A 的收敛性.

首先证明对任意 $c_1\in(0,1]$, 由算法 B 产生的问题(6)的近似解满足如下估计式:

$$m_k(0) - m_k(\boldsymbol{d}_k) \geqslant c_1 \|\boldsymbol{g}_k\| \min\left\{\Delta_k, \frac{\|\boldsymbol{g}_k\|}{\|\boldsymbol{B}_k\|}\right\}. \tag{13}$$

引理 4 搜索方向 $\boldsymbol{d}_k^{\mathrm{C}}$ 满足式(13),其中 $c_1 = 1/2$, 即

$$m_k(0) - m_k(\boldsymbol{d}_k^{\mathrm{C}}) \geqslant \frac{1}{2} \|\boldsymbol{g}_k\| \min\left\{\Delta_k, \frac{\|\boldsymbol{g}_k\|}{\|\boldsymbol{B}_k\|}\right\}. \tag{14}$$

证明 若

$$\min\left\{\frac{\|\boldsymbol{g}_k\|^2}{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{g}_k}, \frac{\boldsymbol{\Delta}_k}{\|\boldsymbol{g}_k\|}\right\} = \frac{\|\boldsymbol{g}_k\|^2}{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{g}_k},$$

则有

$$\boldsymbol{d}_k^{\mathrm{C}} = -\frac{\|\boldsymbol{g}_k\|^2}{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{g}_k}\boldsymbol{g}_k.$$

因此

$$\begin{aligned} m_k(\boldsymbol{d}_k^{\mathrm{C}}) - m_k(0) &= -\frac{\|\boldsymbol{g}_k\|^4}{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{g}_k} + \frac{1}{2}\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{g}_k \frac{\|\boldsymbol{g}_k\|^4}{(\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{g}_k)^2} = \\ &-\frac{1}{2}\frac{\|\boldsymbol{g}_k\|^4}{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{g}_k} \leqslant -\frac{1}{2}\frac{\|\boldsymbol{g}_k\|^2}{\|\boldsymbol{B}_k\|} \leqslant \\ &-\frac{1}{2}\|\boldsymbol{g}_k\| \min\left\{\Delta_k, \frac{\|\boldsymbol{g}_k\|}{\|\boldsymbol{B}_k\|}\right\}, \end{aligned} \tag{15}$$

故式(14)成立.

又若

$$\min\left\{\frac{\|\boldsymbol{g}_k\|^2}{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{g}_k}, \frac{\boldsymbol{\Delta}_k}{\|\boldsymbol{g}_k\|}\right\} = \frac{\boldsymbol{\Delta}_k}{\|\boldsymbol{g}_k\|},$$

则有

$$\boldsymbol{d}_k^{\mathrm{C}} = -\frac{\boldsymbol{\Delta}_k}{\|\boldsymbol{g}_k\|}\boldsymbol{g}_k \text{且 } \|\boldsymbol{g}_k\|^3 > \Delta_k\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{g}_k.$$

因此

$$\begin{aligned} m_k(\boldsymbol{d}_k^{\mathrm{C}}) - m_k(0) &= m_k\left(-\frac{\boldsymbol{\Delta}_k}{\|\boldsymbol{g}_k\|}\boldsymbol{g}_k\right) - f_k = \\ &-\frac{\boldsymbol{\Delta}_k}{\|\boldsymbol{g}_k\|}\|\boldsymbol{g}_k\|^2 + \frac{1}{2}\frac{\Delta_k^2}{\|\boldsymbol{g}_k\|^2}\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{g}_k \leqslant \\ &-\frac{1}{2}\Delta_k\|\boldsymbol{g}_k\| \leqslant -\frac{1}{2}\|\boldsymbol{g}_k\| \min\left\{\Delta_k, \frac{\|\boldsymbol{g}_k\|}{\|\boldsymbol{B}_k\|}\right\}. \end{aligned}$$

故结论成立. □

引理 5 若 $\boldsymbol{d}_k$ 由算法 B 产生,则 $\boldsymbol{d}_k$ 满足式(13),且式中 $c_1 = 1/2$.

证明 因为 $\|\boldsymbol{d}_k\| \leqslant \Delta_k$, 则由引理 4 知

$$m_k(0) - m_k(\boldsymbol{d}_k) \geqslant m_k(0) - m_k(\boldsymbol{d}_k^{\mathrm{C}}) \geqslant \frac{1}{2}\|\boldsymbol{g}_k\| \min\left\{\Delta_k, \frac{\|\boldsymbol{g}_k\|}{\|\boldsymbol{B}_k\|}\right\}, \tag{16}$$

即结论成立. □

下面将证明算法的全局收敛性.全文假设目标函数 Hesse 矩阵的近似 $\boldsymbol{B}_k$ 关于范数一致有

界,且水平集

$$L_0 = \{ \boldsymbol{w} \mid f(\boldsymbol{w}) \leqslant f(\boldsymbol{w}_0) \} \tag{17}$$

有界.

定理 1　在算法 A 中令 $\varepsilon = 0$,对任意 $\boldsymbol{w}^0 \in \mathbf{R}_+ \times R^n$,设 f 在水平集 L_0 上连续可微.若存在 $\beta > 0$,使得对任意 k 有 $\| \boldsymbol{B}_k \| \leqslant 2\beta$, 问题(6)的近似解满足式(16).则下列 3 种情形必有一种发生:

1) 存在常数 k 使得 $\boldsymbol{g}_k = \boldsymbol{0}$;

2) $f(\boldsymbol{w}_k) \to -\infty$;

3) $\lim\limits_{k\to\infty} \inf \| \boldsymbol{g}_k \| = 0$.

证明　由式(11),有

$$| r_k - 1 | = \left| \frac{f(\boldsymbol{w}_k) - f(\boldsymbol{w}_k + \boldsymbol{d}_k) - m_k(0) + m_k(\boldsymbol{d}_k)}{m_k(0) - m_k(\boldsymbol{d}_k)} \right| = \left| \frac{f(\boldsymbol{w}_k + \boldsymbol{d}_k) - m_k(\boldsymbol{d}_k)}{m_k(0) - m_k(\boldsymbol{d}_k)} \right|. \tag{18}$$

根据 Taylor 公式得

$$f(\boldsymbol{w}_k + \boldsymbol{d}_k) = f(\boldsymbol{w}_k) + \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{d}_k + \int_0^1 [\boldsymbol{g}(\boldsymbol{w}_k + \tau \boldsymbol{d}_k) - \boldsymbol{g}_k]^{\mathrm{T}} \boldsymbol{d}_k \mathrm{d}\tau, \tag{19}$$

且当 Δ_k 充分小时,有

$$| m_k(\boldsymbol{d}_k) - f(\boldsymbol{w}_k + \boldsymbol{d}_k) | = \left| \frac{1}{2} \boldsymbol{d}_k^{\mathrm{T}} \boldsymbol{B}_k \boldsymbol{d}_k - \int_0^1 \boldsymbol{d}_k^{\mathrm{T}} [\boldsymbol{g}(\boldsymbol{w}_k + \tau \boldsymbol{d}_k) - \boldsymbol{g}_k] \mathrm{d}\tau \right| \leqslant \beta \| \boldsymbol{d}_k \|^2 + o(\| \boldsymbol{d}_k \|). \tag{20}$$

若情形 1) 和 2) 都不成立,用反证法证明情形 3) 一定成立.反设存在 $\varepsilon_0 > 0, K > 0$ 使得当 $k > K$ 时

$$\| \boldsymbol{g}_k \| \geqslant \varepsilon_0, \qquad 对\ \forall k > K. \tag{21}$$

由式(16),当 $k \geqslant K$ 时

$$m_k(0) - m_k(\boldsymbol{d}_k) \geqslant \frac{1}{2} \| \boldsymbol{g}_k \| \min \left\{ \Delta_k, \frac{\| \boldsymbol{g}_k \|}{\| \boldsymbol{B}_k \|} \right\} \geqslant \frac{1}{2} \varepsilon_0 \min \left\{ \Delta_k, \frac{\varepsilon_0}{\beta} \right\}. \tag{22}$$

联合式(18)~(22)得

$$| r_k - 1 | \leqslant \frac{\beta \Delta_k^2 + o(\| \boldsymbol{d}_k \|)}{\varepsilon_0 \min \{ \Delta_k, \varepsilon_0 / \beta \}}. \tag{23}$$

由式(23),存在充分小的 $\tilde{\Delta}$,当 $\Delta_k \leqslant \tilde{\Delta}$ 时

$$| r_k - 1 | \leqslant 1 - \eta_2, \tag{24}$$

即

$$r_k > \eta_2.$$

因此,由算法 A 知

$$\Delta_{k+1} \geqslant \Delta_k,$$

所以存在 $\kappa > 0$,对任意 $\Delta_k \leqslant \tilde{\Delta}$ 和 $k \geqslant K$ 有

$$\Delta_k \geqslant \kappa \tilde{\Delta}. \tag{25}$$

现假设存在无穷序列 $N_0 \in \mathbf{N}$,使得对 $\forall k \in N_0$,有 $r_k \geqslant \eta$.则对 $\forall k \in N_0$,当 $k \geqslant K$, 由式(22)有

$$f_k - f_{k+1} \geqslant \eta_1[m_k(0) - m_k(\boldsymbol{d}_k)] \geqslant \frac{\eta_1}{2}\varepsilon_0 \min\left\{\Delta_k, \frac{\varepsilon_0}{\beta}\right\}. \tag{26}$$

由假设情形 2)不成立,因此 f 下有界,从而由式(26)知

$$\lim_{k \in N_0, k \to \infty} \Delta_k = 0,$$

这与式(25)矛盾.

另一方面,对充分大的 k,若设 $r_k < \eta_1$,则 Δ_k 以 γ_1 的比例压缩,从而

$$\lim_{k \to \infty} \Delta_k = 0,$$

同样与式(25)矛盾.因此,假设不真,从而情形 3)成立. □

4 数值试验

为了检测本文算法的有效性,针对来自文献[9]中的部分问题,将本文算法和文献[3]中的算法 2.2 进行比较.算法 A 中参数选取如下:$\eta_1 = 0.01, \eta_2 = 0.75, \gamma_1 = 0.5, \gamma_2 = 2.0, \Delta_0 = 1, \tau = 0.5, \varepsilon_0 = 0.1$, 所得结果见表 1.

表 1 文献[9]中的算例运行结果

Table 1 Numerical results of examples in ref.[9]

problem	N_f	N_d	algorithm 2.2 in ref.[3]			algorithm A			funcount
			min	max	aiter	min	max	aiter	
P35	4	3	2	6	4.4	3	6	5.5	12
P43	3	4	1	5	2.8	3	8	5.2	40
P76	7	4	4	12	8.0	6	21	8.2	43
P86	15	5	10	38	19.6	8 *	30 *	18.6 *	58
P113	8	10	7	51	21.4	9	26 *	16.9 *	96

注 N_f:不等式系统中不等式的个数;N_d:不等式系统中变量个数;min:迭代次数的最小值;max:迭代次数的最大值;aiter:平均迭代次数;funcount:目标函数计算次数;*:表明本文算法优于文献[3]中的算法.

下面针对较简单问题,运用文中算法 A 进行检测.

算例 1 $\boldsymbol{F}(\boldsymbol{x}) = [F_1(\boldsymbol{x}), F_2(\boldsymbol{x})]^{\mathrm{T}}$, 其中

$$\begin{cases} F_1(\boldsymbol{x}) = x_1^2 - 1, \\ F_2(\boldsymbol{x}) = -x_1 + x_2^2. \end{cases}$$

算例 2 $\boldsymbol{F}(\boldsymbol{x}) = [F_1(\boldsymbol{x}), F_2(\boldsymbol{x})]^{\mathrm{T}}$, 其中

$$\begin{cases} F_1(\boldsymbol{x}) = x_1 - 0.7\sin x_1 - 0.2\cos x_2, \\ F_2(\boldsymbol{x}) = x_2 - 0.7\cos x_1 + 0.2\sin x_2. \end{cases}$$

运行文中算法 A,所得结果见表 2 和表 3.

表 2 算例 1 实验结果

Table 2 Numerical results of example 1

start point	iter	solution (x_1, x_2)	function value (F_1, F_2)
$x^0 = [-1, 1]$	8	(1.000 057 164 786, 0.489 583 333 333)	(0.000 114 332 839, −0.760 365 324 508)
$x^0 = [0.5, 0.5]$	3	(0.500 000 000 000, 0.500 000 000 000)	(−0.75, −0.25)
$x^0 = [1, -1]$	3	(1, −1)	(0, 0)
$x^0 = [1, -0.5]$	2	(0.968 750 000 000, −0.500 000 000 000)	(−0.061 523 437 500, −0.718 750 000 000)
$x^0 = [2, 0.5]$	9	(1.000 060 749 089, 0.500 000 000 000)	(0.000 121 501 868, −0.750 060 740 000)

表 3　算例 2 实验结果

Table 3　Numerical results of example 2

start point	iter	solution (x_1, x_2)	function value (F_1, F_2)
$x^0 = [-1, -1]$	10	(−1.000 000 000 000, −1.000 000 000 000)	(−0.519 030 771 80, −1.546 505 811 06)
$x^0 = [1, 0]$	15	(1.000 000 000 000, 0)	(−0.619 070 311 063, −0.378 211 614 100)
$x^0 = [0, 0]$	7	(−0.2, −0.7)	(−0.061 523 4, −0.718 750 0)
$x^0 = [1, 0]$	27	(0.588 757 905 31, 0)	(0.000 027 964 410, −0.582 141 766 370)

以上算例的运行结果中,关于迭代次数的平均值,采取利用选取 5 组不同参数所得到的结果进行平均而得到.

5　结　　论

表 1 中,本文选定了 5 组参数分别对所列出的 5 个算例做了实验,得到了每组参数求对应问题的迭代次数及平均迭代次数.表 1 的结果表明本文算法有效,其中标注 * 的结果表明本文算法优于文献[3]中的算法,但是目标函数计算次数稍多.造成目标函数计算次数多的原因是由算法 B 中寻找合适的 λ 而引起的,将 λ 的值进行固定的时候,相应地迭代次数会适当增加.因此如何确定每步迭代算法 B 中 λ 的最优值,将作进一步地研究.

表 2、表 3 分别是针对算例 1 和算例 2 进行的数值试验结果.当算例中函数的非线性程度越高,导致迭代次数会显著增加.如何设计针对非线性程度高的非线性不等式系统的光滑算法,将作进一步的考虑.

综上所述,当采用本文所提的带折线步的信赖域算法求解非线性不等式组时,参数的选定非常重要.在合适的参数选定条件下,针对不同的初始迭代点 x_0 都能有效地求得非线性不等式系统的一个或多个可行点,这对于采用可行方向法求解最优化问题是非常有意义的.

致谢　感谢审稿专家和编辑部老师对本文提出的宝贵意见;感谢怀化学院创新性试验点基金(20121106)的资助;另外非常感谢福建师范大学马昌凤教授的悉心指导.

参考文献(References):

[1] Fukushima M. A finitely convergent algorithm for convex inequalities[J]. *IEEE Trans Autom Contr*, 1982, **27**(5): 1126-1127.

[2] Jian J B, Liang Y M. Finitely convergent algorithm of generalized gradient projection for systems of nonlinear inequalities[J]. *Neural Parallel and Scientific Computing*, 2004, **12**(2): 207-218.

[3] JIAN Jin-bao, CHENG Wei-xin, KE Xiao-yan. Finitely convergent ε-generalized projection algorithm for nonlinear systems[J]. *J Math Anal Appl*, 2007, **332**(2): 1446-1459.

[4] 何郁波, 马昌凤. 非线性不等式组的信赖域算法[J]. 工程数学学报, 2008, **25**(2): 224-230.(HE Yu-bo, MA Chang-feng. Trust region method for nonlinear inequalities[J]. *Journal of Engineering Mathematics*, 2008, **25**(2): 224-230.(in Chinese))

[5] 何郁波, 林晓艳, 董晓亮. 非线性不等式组的光滑近似方法及其收敛性[J]. 应用数学学报, 2011, **34**(4): 723-733.(HE Yu-bo, LIN Xiao-yan, DONG Xiao-liang. On the convergence of smoothing approximate method for nonlinear inequalities[J]. *Acta Mathematicae Applicatae Sinica*, 2011, **34**(4): 723-733.(in Chinese))

[6] Ma C F. A globally convergent Levenberg-Marquardt method for the least l2-norm solution of nonlinear inequalities[J]. *Applied Mathematics and Computation*, 2008, **206**(1): 133-140.

[7] 程维新，陈永强. 求解非线性不等式组的有限步终止算法[J]. 河南师范大学学报(自然科学版)，2010，**38**(2)：25-27.(CHENG Wei-xin，CHEN Yong-qiang. A finitely terminating algorithm for systems of nonlinear inequalities[J]. *Journal of Henan Normal University*(*Natural Science*)，2010，**38**(2)：25-27.(in Chinese))

[8] Powell M J D. Convergence properties of a class of minimization algorithms[C]//Mangasarian O L，Meyer P R，Robinson S M，ed. *Nonlinear Programming* 2. New York：Academic Press，1975.

[9] Hock W，Schittkowski K. *Test Examples for Nonlinear Programming Codes*[M]. Berlin：Springer-Verlag Press，1981.

On the Convergence of Trust Region Method With Dogleg Step for Nonlinear Inequalities Systems

HE Yu-bo，　LIN Xiao-yan

(*Department of Mathematics*，*Huaihua University*，*Huaihua*，*Hunan* 418008，*P.R.China*)

Abstract：The solutions of a class of nonlinear inequalities were studied.The nonlinear inequalities were approximated by a family of parameterized optimization problems with twice continuously differentiable objective functions，then a smoothing trust region method with dogleg steps was applied to solve the parameterized optimization problems.The global convergence of the proposed method was established under some weak conditions. Numerical results show that the method performs well.

Key words：nonlinear inequalities；trust region method；Cauchy point；dogleg step；global convergence