第 36 卷 第 1 期 应 用 数 学 学 报 Vol. 36 No. 1
2013 年 1 月 ACTA MATHEMATICAE APPLICATAE SINICA Jan., 2013

求解互补问题的一族非单调光滑牛顿法 *

李向利

(桂林电子科技大学数学与计算科学学院，桂林 541004)

(E-mail: lixiangli213@gmail.com)

刘红卫

(西安电子科技大学应用数学系，西安 710071)

摘 要 基于广义 Fischer-Burmeister 函数，在本文我们提出了求解互补问题的一族非单调光滑牛顿法. 该方法的全局和局部收敛性在理想情况下得到了证明，并且也给出了实验结果.

关键词 互补问题；一族非单调光滑牛顿法；全局收敛性；局部收敛性

MR(2000) 主题分类 90

中图分类 O221

1 引言

设 $F: R^n \to R^n$ 是连续可微的. 互补问题就是寻找一向量 $x \in R^n$ 使得

$$x \geq 0, \qquad F(x) \geq 0, \qquad x^T F(x) = 0. \tag{1.1}$$

互补问题在理论上和实际应用都是非常有用的，并已用于研究和制定各种由经济和工程中产生的均衡问题，如纳什均衡问题，交通均衡问题，接触力学问题等等. 问题 (1.1) 等价于下面的非光滑方程

$$\begin{aligned} &\Phi(x) = 0, \qquad \Phi: R^n \longrightarrow R^n, \\ &\Phi(x) = (\phi(x_1, F_1(x)), \phi(x_2, F_2(x)), \cdots, \phi(x_n, F_n(x)))^T, \end{aligned} \tag{1.2}$$

这里

$$\phi(a, b) = 0 \Longleftrightarrow a, b \geq 0, \qquad ab = 0.$$

本文 2010 年 8 月 24 日收到. 2011 年 9 月 29 日收到修改稿.

* 国家自然科学基金 (61072144, 61179040, 71001015, 71101033) 资助项目.

满足上式的函数较多，例如， Fischer-Burmeister 函数：

$$\phi(a,b)=\sqrt{a^2+b^2}-(a+b). \tag{1.3}$$

有很多方法求解 (1.2), 如非光滑牛顿法 [1], 雅克比光滑法 [2] 和光滑的牛顿法 [3−5]. 在这些方法中，越来越多的注意力集中在光滑牛顿法上 [6,7]. 一类光滑牛顿法就是求解序列问题

$$H(z)=H(\mu,x)=\begin{pmatrix}\mu\\ \Phi(\mu,x)\end{pmatrix}=0,\qquad \mu>0 \tag{1.4}$$

对每一 μ 且逐渐使得 μ 趋于 0. 这里的 $\Phi(\mu,x)$ 是 Φ 的一光滑近似且当 $\mu\to 0$ 时，$\Phi(\mu,x)\to\Phi$. μ 通常称为光滑参数. 显然 $H(z)=0$ 等价于 $\mu=0$ 且 x 是 (1.1) 的一个解.

最近，在 [8] 和 [9] 中，广义 Fischer-Burmeister 函数得到了研究. 特别是，定义 $\phi_p: R^2\to R$ 为

$$\phi_p(a,b)=\sqrt[p]{|a|^p+|b|^p}-(a+b). \tag{1.5}$$

这里 p 是 $(1,+\infty)$ 上任意固定的实数. 由性质 3.1[8] 知， ϕ_p 是一族 NCP- 函数.

在本文，我们给出 $\phi_p(a,b)$ 的光滑近似 $\phi_p(\mu,a,b)$, 并推广 (1.4) 到如下的形式：

$$H_p(z)=H_p(\mu,x)=\begin{pmatrix}k(\mu)\\ \Phi_p(\mu,x)\end{pmatrix}=0,\qquad \mu>0. \tag{1.6}$$

这里 $[\Phi_p(\mu,x)]_i=\phi_p(\mu,x_i,F_i(x))$, 函数 $k(\mu)$ 的性质将在后面进行讨论. 我们提出了一族非单调光滑牛顿法求解 (1.6), 提出的算法的全局和局部收敛性在一定的条件下得到了证明.

本文结构如下：在第 2 部分，提出了一些预备知识. 在第 3 部分，提出了一新的光滑函数并分析了它的性质. 在第 4 部分，给出了新的方法. 提出算法的全局和局部收敛性分别在第 5 部分和第 6 部分进行讨论. 第 7 部分给出数据实验. 第 8 部分进行了小结.

在整个文章中，所有的向量都是列向量. T 表示转置. R^n_+ 和 R^n_{++} 分别表示 R^n 中的非负数和正数. $O(t)$ 表示使得 $\lim\limits_{t\downarrow 0}\sup\frac{|O(t)|}{t}<\infty$ 成立的任一函数. 对可微函数 $F: R^n\to R^m$, 我们记 $F'(x)$ 为 F 在点 x 的雅克比，而 $\nabla F(x)$ 为雅克比的转置. 设 $F: R^n\to R^m$ 是局部 Lipschitz 连续， $\partial F(x)$ 表示 F 在点 x 处的 Clarke 广义雅克比矩阵.

2 预备知识

定义 2.1 (1) 矩阵 $M\in R^n$ 被称为 P_0 矩阵如果它的所有主子式都是非负的. 这类矩阵记为 P_0.

(2) 函数 $F: R^n \to R^n$ 被称为 P_0 函数如果对任意 $x, y \in R^n$ 且 $x \neq y$, 存在指标 i_0 使得

$$x_{i_0} \neq y_{i_0}, \qquad (x_{i_0} - y_{i_0})[F_{i_0}(x) - F_{i_0}(y)] \geq 0.$$

定义 2.2 函数 $F: R^n \to R^n$ 被称为 R_0 函数如果对满足

$$\|x^k\| \to \infty, \qquad \lim_{k\to\infty} \inf \frac{\min_i x_i^k}{\|x^k\|} \geq 0, \qquad \lim_{k\to\infty} \inf \frac{\min_i F_i(x^k)}{\|x^k\|} \geq 0$$

的任意序列 $\{x^k\}$, 存在指标 j 使得

$$\{x_j^k\} \longrightarrow \infty \qquad \text{和} \qquad \{F_j(x^k)\} \longrightarrow \infty.$$

由定义 2.1 和定理 2.1.24[10], 任一连续可微的 P_0 函数的雅克比必为 P_0 矩阵.

定义 2.3 $F: R^n \to R^m$ 是局部 Lipschitz 连续. 如果对任意的 $d \to 0$ 和 $V \in \partial F(x+d)$,

$$F(x+d) - F(x) - Vd = o(\|d\|),$$

F 在 x 处称为半光滑; 如果对任意的 $d \to 0$ 和 $V \in \partial F(x+d)$,

$$F(x+d) - F(x) - Vd = O(\|d\|^2),$$

F 在 x 处称为强半光滑.

3 新的光滑函数及其性质

在本文, 我们给出 $\phi_p(a,b)$ 的光滑近似函数如下:

$$\phi_p(\mu, a, b) = \sqrt[p]{|a|^p + |b|^p + |\mu|^p} - (a+b). \tag{3.1}$$

这里的 μ 是一光滑参数. 对任意 $\mu \neq 0$, $\phi_p(\mu, a, b)$ 是光滑的且梯度为

$$(\nabla_{(a,b)}\phi_p(\mu,a,b))^T = \left(\frac{|a|^{p-1}\mathrm{sgn}(a)}{(|a|^p+|b|^p+|\mu|^p)^{1-\frac{1}{p}}} - 1, \frac{|b|^{p-1}\mathrm{sgn}(b)}{(|a|^p+|b|^p+|\mu|^p)^{1-\frac{1}{p}}} - 1\right)^T$$

这里 $\mathrm{sgn}\,(\cdot)$ 是 sign 函数.

下面的结论很容易证明.

性质 3.1 对所有的 $(a,b) \in R^2$ 和所有的 μ_1, μ_2, 函数 $\phi_p(\mu, a, b)$ 满足不等式

$$|\phi_p(\mu_1, a, b) - \phi_p(\mu_2, a, b)| \leq |\mu_1 - \mu_2|;$$

特别的, 对所有的 $(a,b) \in R^2$ 和所有的 μ, 我们有

$$|\phi_p(\mu, a, b) - \phi_p(a, b)| \leq |\mu|.$$

设

$$\Phi_p(\mu,x)=\begin{pmatrix}\phi_p(\mu,x_1,F_1(x))\\ \vdots\\ \phi_p(\mu,x_n,F_n(x))\end{pmatrix}.$$

由性质 3.1, 容易得到下面的推论.

推论 3.1 对所有的 $x\in R^n$ 和所有的 μ_1, μ_2, 函数 $\Phi_p(\mu,x)$ 满足不等式

$$\|\Phi_p(\mu_1,x)-\Phi_p(\mu_2,x)\|\le n|\mu_1-\mu_2|;$$

特别的，对所有的 $x\in R^n$ 和所有的 μ, 我们有

$$\|\Phi_p(\mu,x)-\Phi_p(x)\|\le n|\mu|.$$

性质 3.2 $\phi_p(\mu,a,b)$ 定义为 (3.1) 且 p 是 $(1,+\infty)$ 中任意固定的实数，则

i) $\phi_p(\mu,a,b)$ 是凸的.

ii) $\phi_p(\mu,a,b)$ 是 Lipschitz 连续，当 $1<p<2$ 时， $L_1=\sqrt{3}+3^{(1/p-1/2)}$, 当 $p\ge 2$ 时， $L_2=\sqrt{3}+1$.

iii) $\phi_p(\mu,a,b)$ 在 R^3 上是强半光滑的.

证　i) 和 ii) 的证明类似于 [8] 中性质 3.1 的 d) 和 e), 在这我们省略该过程．设 $h=(h_1,h_2,h_3)^T$, 对任意 $\xi\in\partial\phi(h_1,h_2,h_3)$ 和 $h\to 0$, 因为 $\phi_p(\mu,a,b)$ 在 $R^3\setminus(0,0,0)$ 上是光滑的，我们只需要证明

$$\xi^Th-\phi_p'((0,0,0);(h_1,h_2,h_3))=O(\|h\|^2).$$

由 i), 我们知 $\phi_p(\mu,a,b)$ 在 R^3 上是半光滑的．结合 [10] 的性质 5.1.2, 有

$$\phi_p'((0,0,0);(h_1,h_2,h_3))=\phi_p(h_1,h_2,h_3).$$

因为 $h\neq 0$, 所以

$$\xi=\Big(\frac{|h_1|^{p-1}\mathrm{sgn}\,(h_1)}{(|h_2|^p+|h_3|^p+|h_1|^p)^{1-\frac{1}{p}}},\frac{|h_2|^{p-1}\mathrm{sgn}\,(h_2)}{(|h_2|^p+|h_3|^p+|h_1|^p)^{1-\frac{1}{p}}}-1,\frac{|h_3|^{p-1}\mathrm{sgn}\,(h_3)}{(|h_2|^p+|h_3|^p+|h_1|^p)^{1-\frac{1}{p}}}-1\Big).$$

因此，

$$\xi^Th-\phi_p'((0,0,0);(h_1,h_2,h_3))=\phi_p(h_1,h_2,h_3)-\phi_p(h_1,h_2,h_3)=0.$$

下面的引理来自于 [8] 的引理 3.1.

引理 3.1 设 $\phi_p:R^2\to R$ 定义为 (1.5), 这里 $p>1$. 如果 $\{(a^k,b^k)\}\subseteq R^2$ 并有 $(a^k\to\infty$且$b^k\to\infty)$ 或 $(a^k\to-\infty$或$b^k\to-\infty)$, 则有 $|\phi_p(a^k,b^k)|\to\infty$ 对 $k\to\infty$.

引理 3.2 设 $\phi_p:R^3\to R$ 定义为 (3.1), 这里 $p>1$. 如果 $\{(a^k,b^k)\}\subseteq R^2$ 并有 $(a^k\to\infty$且$b^k\to\infty)$ 或 $(a^k\to-\infty$或$b^k\to-\infty)$, 则有 $|\phi_p(\mu,a^k,b^k)|\to\infty$ 对 $k\to\infty$.

证　由性质 3.1, $|\phi_p(\mu,a^k,b^k)-\phi_p(a^k,b^k)|\le|\mu|$. 结合引理 3.1, 不难得到结论.

4 算法

设 $z:=(\mu,x)\in R_+\times R^n$ 且

$$H_p(z)=H_p(\mu,x)=\begin{pmatrix} k(\mu) \\ \Phi_p(\mu,x) \end{pmatrix}. \tag{4.1}$$

这里 $[\Phi_p(\mu,x)]_i=\phi_p(\mu,x_i,F_i(x))$. 如果 $k(\mu)$ 满足下面的假设 1, 则 $H_p(z)=0$ 等价于 $\mu=0$ 且 x 是 (1.1) 的一个解.

假设 1 $k:R_+\to R$ 满足下面的条件.

i) k 是连续可微的.

ii) $k(\mu)=0\Leftrightarrow\mu=0$.

iii) $k(\mu)\ge 0$.

iv) 当 $\mu>0$ 时, $k'(\mu)>0$.

v) 当 $\mu\to\infty$ 时, $k(\mu)\to\infty$.

vi) 当 $\mu>0$ 时, $\frac{k(\mu)}{k'(\mu)}\le\mu$.

vii) 存在实数 $a>0$, $b\ge 0$, 使得 $k'(\mu)\le ak(\mu)+b$.

由假设 1, 易得到下面的定理.

定理 4.1 如果 $k_1(\mu)$ 和 $k_2(\mu)$ 满足假设 1, 则 $\omega_1k_1(\mu)+\omega_2k_2(\mu)$ 和 $k_1(k_2(\mu))$ 也满足假设 1, 这里 $\omega_1,\omega_2>0$.

定理 4.2 下面的函数满足假设 1.

1) $k(\mu)=c_n\mu^n+c_{n-1}\mu^{n-1}+c_{n-2}\mu^{n-2}+\cdots+c_2\mu^2+c_1\mu,\ n\ge 1$. 这里 $0<c_1\le b,\ 0<c_2\le\frac{ac_1}{2},\ 0<c_3\le\frac{ac_2}{3},\cdots,0<c_{n-1}\le\frac{ac_{n-2}}{n-1},\ 0<c_n\le\frac{ac_{n-1}}{n}$.

2) $k(\mu)=c^\mu-1,\ c>1,\ a\ge\ln c,\ b\ge a$.

3) $k(\mu)=\mu^n\ln(1+\mu),\ n\ge 1,\ a\ge n+\frac{1}{2\ln 2},\ b\ge n\ln 2+1$.

4) $k(\mu)=(\mu+1)^n\ln(1+\mu),\ n\ge 1,\ a\ge n+\frac{1}{2\ln 2},\ b\ge 2^{n-1}n\ln 2+2^{n-1}$.

5) $k(\mu)=\mu^n(c^\mu-1),\ n\ge 1,\ c>1,\ a\ge n+\frac{c\ln c}{c-1},\ b\ge n(c-1)+c\ln c$.

6) $k(\mu)=(\mu+1)^n(c^\mu-1),\ n\ge 1,\ c>1,\ a\ge n+\frac{c\ln c}{c-1},\ b\ge n2^{n-1}(c-1)+2^nc\ln c$.

证 容易证明 1) 和 2) 满足假设 1, 并且 3) 也满足假设 1 的 i), ii), iii), iv) 和 v). 因此，我们只需证明 3) 满足假设 1 的 vi) 和 vii) 即可.

因为

$$k(\mu)=\mu^n\ln(1+\mu)\le n\mu^n\ln(1+\mu)\le\left[n\mu^{n-1}\ln(1+\mu)+\frac{\mu^n}{1+\mu}\right]\mu=k'(\mu)\mu,$$

$$\frac{k(\mu)}{k'(\mu)}\le\mu,$$

这表明 3) 满足假设 1 的 vi).

为了证明 3) 也满足假设 1 的 vii), 我们考虑两种情况.

a) $\mu \geq 1$. 因为

$$n\mu^{n-1}\ln(1+\mu) \leq n\mu^n \ln(1+\mu) \tag{4.2}$$

且

$$\frac{\mu^n}{1+\mu} \leq \frac{\mu^n}{2} \leq \frac{\mu^n \ln(1+\mu)}{2\ln 2}, \tag{4.3}$$

得

$$\begin{aligned} k'(\mu) =& n\mu^{n-1}\ln(1+\mu) + \frac{\mu^n}{1+\mu} \leq \Big(n + \frac{1}{2\ln 2}\Big)\mu^n \ln(1+\mu) \\ \leq& ak(\mu) \leq ak(\mu) + b. \end{aligned} \tag{4.4}$$

b) $0 \leq \mu < 1$. 易得

$$k'(\mu) = n\mu^{n-1}\ln(1+\mu) + \frac{\mu^n}{1+\mu} < n\ln 2 + 1 \leq b + ak(\mu).$$

由 a) 和 b), 得 $k'(\mu) \leq ak(\mu) + b$. 这蕴含 3) 满足假设 1 的 vii).

4), 5) 和 6) 的证明类似于 3). 证毕.

注意

$$[\Phi_p(z)]'_x = [\Phi_p(\mu, x)]'_x = A(z) + B(z)F'(x).$$

这里 $A(z) = \mathrm{diag}\,(A_{ii}(z)),\ B(z) = \mathrm{diag}\,(B_{ii}(z))$,

$$A_{ii}(z) = \frac{|x_i|^{p-1}\mathrm{sgn}\,(x_i)}{(|x_i|^p + |F_i(x)|^p + |\mu|^p)^{1-\frac{1}{p}}} - 1, \qquad B_{ii}(z) = \frac{|F_i(x)|^{p-1}\mathrm{sgn}\,(F_i(x))}{(|x_i|^p + |F_i(x)|^p + |\mu|^p)^{1-\frac{1}{p}}} - 1.$$

显然的，当 $\mu \neq 0$ 时， $A_{ii}(z) < 0,\ B_{ii}(z) < 0$. 如果 $F'(z)$ 是 P_0 矩阵，由 [10] 的引理 5.3.12, 我们知 $[\Phi_p(z)]'_x$ 是非奇异的.

性质 4.1 假设 F 是连续可微的 P_0 函数，则对所有的 $z \in R_{++} \times R^n$，H_p 是连续可微的，对任意 $z \in R_{++} \times R^n$，$H'_p(z)$ 是非奇异的.

对任意 $z = (\mu, x) \in R \times R^n$, 定义 $\theta(z) = \|H_p(z)\|^2 = k(\mu)^2 + \Psi(z)$, 这里 $\Psi(z) = \|\Phi_p(\mu, x)\|^2$.

算法 1

(s.0) 选取 $\delta, \sigma \in (0,1)$, $0 < \tau < 1$ 和 $\overline{\mu} > 0$. 选取 $\gamma \in (0,1)$ 使得 $(a+b)\gamma\overline{\mu} < 1$. 令 $\beta(z^{-1}) = \gamma,\ \mu^0 = \overline{\mu},\ x^0 \in R^n$. 令 $z^0 = (\mu^0, x^0),\ \overline{z} = (\overline{\mu}, 0),\ M \geq 1,\ C_0 = \theta(z^0),\ m(0) = 0$, 且 $k := 0$.

(s.1) 如果 $\|H(z^k)\| = 0$, 停，否则，

$$\beta(z^k) = \min\big\{\gamma, \gamma\theta(z^k), \beta(z^{k-1})\big\}. \tag{4.5}$$

(s.2) 计算牛顿方向 $\Delta z^k := (\Delta\mu^k, \Delta x^k) \in R^{n+1}$ 根据下式

$$H_p(z^k) + H'_p(z^k)\Delta z^k = k'(\mu^k)\beta(z^k)\overline{z}. \tag{4.6}$$

(s.3) 设 λ_k 是 $1,\delta,\delta^2,\cdots$ 中使得下式成立的最大者

$$\theta(z^k+\lambda_k\Delta z^k)\le\{1-2\sigma[1-(a+b)\gamma\overline{\mu}]\lambda_k\}C_k. \tag{4.7}$$

(s.4) 令 $z^{k+1}=z^k+\lambda_k\Delta z^k$.

(s.5) 令 $m(k+1)=\min\{m(k)+1,M\}$, $C_{k+1}=\frac{\tau\cdot m(k+1)C_k+\theta(z^{k+1})}{\tau\cdot m(k+1)+1}$.

(s.6) 令 $k=k+1$. 转 s.1.

算法说明 1 C_k 起着重要的作用，它可以使得 $\theta(z^k)$ 是非单调的． C_k 的单调性可见定理 5.1.

5 收敛性分析

我们现在考虑算法 1 的全局收敛性．在这之前，我们首先要做下面假设，这些假设将在算法的 1 的讨论中用到.

假设 2 $F:R^n\to R^n$ 是连续可微 P_0+R_0 函数.

引理 5.1 若假设 1 和 2 成立．对任意 $z:=(\mu,x)\in R_{++}\times R^n$, 定义水平集

$$\mathcal{L}(z^0):=\{z\in R_{++}\times R^n:\theta(z)\le\theta(z^0)\},$$

这里 z^0 是算法 1 的初始点，则集合 $\mathcal{L}(z^0)$ 是有界的.

证 我们通过反证法来证明这个引理．事实上，如果引理的结论不真，则存在序列 $\{z^k=(\mu^k,x^k)\in R_{++}\times R^n$, 使得

$$\theta(z^k)\le\theta(z^0),\qquad \|z^k\|\longrightarrow\infty.$$

我们考虑两种情况.

i) 如果 $\mu^k\to\infty$ 对每一个 k, 由假设 1, 我们有 $k(\mu^k)\to\infty$, $\theta(z^k)=k(\mu^k)^2+\Psi(z^k)\to\infty$. 这是不可能的.

ii) 基于 i) 的分析，存在 $0<\mu_1\le\mu_2$, 使得 $\mu_1\le\mu^k\le\mu_2$, 因此 $\|x^k\|\to\infty$. 结合引理 3.2 和假设 2, 我们得到的结果和 $\theta(z^k)\le\theta(z^0)$ 不一致.

由 i) 和 ii), 结论成立.

定理 5.1 如果假设 1 和 2 成立．算法 1 是有定义的.

证 显然，由 (4.5), $\beta(z)\le\gamma\theta(z)^{1/2}$. 因此，有

$$\begin{aligned}k(\mu+\lambda\Delta\mu)^2=&k(\mu)^2+2\lambda k(\mu)k'(\mu)\Delta\mu+o(\lambda)\\=&(1-2\lambda)k(\mu)^2+2\lambda k'(\mu)k(\mu)\beta(z)\overline{\mu}+o(\lambda)\\\le&(1-2\lambda)k(\mu)^2+2\lambda[ak(\mu)+b]k(\mu)\beta(z)\overline{\mu}+o(\lambda)\\\le&(1-2\lambda)k(\mu)^2+2\lambda ak(\mu)^2\beta(z)\overline{\mu}+2\lambda bk(\mu)\beta(z)\overline{\mu}+o(\lambda)\\\le&(1-2\lambda)k(\mu)^2+2\lambda a\gamma\overline{\mu}\theta(z)+2\lambda b\gamma\overline{\mu}\theta(z)+o(\lambda).\end{aligned} \tag{5.1}$$

从 (4.6), $\nabla\Psi(z)^T\Delta z=-2\Psi(z)$, 因此,

$$\Psi(z+\lambda\Delta z)=\Psi(z)+\lambda\nabla\Psi(z)^T\Delta z+o(\lambda)=(1-2\lambda)\Psi(z)+o(\lambda). \tag{5.2}$$

结合 (5.1), 我们有

$$\begin{aligned}&\theta(z+\lambda\Delta z)\\ \le&(1-2\lambda)\theta(z)+2(a+b)\lambda\gamma\overline{\mu}\theta(z)+o(\lambda)\\ =&\big\{1-2[1-(a+b)\gamma\overline{\mu}]\lambda\big\}\theta(z)+o(\lambda)\\ =&\big\{1-2\sigma[1-(a+b)\gamma\overline{\mu}]\lambda\big\}\theta(z)-2(1-\sigma)[1-(a+b)\gamma\overline{\mu}]\lambda\theta(z)+o(\lambda).\end{aligned} \tag{5.3}$$

因为 $\sigma\in(0,1)$ 且 $(a+b)\gamma\overline{\mu}<1$, 则

$$2(1-\sigma)[1-(a+b)\gamma\overline{\mu}]\lambda\theta(z)>0. \tag{5.4}$$

当 $\theta(z)>0$. 由 (5.3) 和 (5.4), 我们可得

$$\theta(z+\lambda\Delta z)\le\{1-2\sigma[1-(a+b)\gamma\overline{\mu}]\lambda\}\theta(z). \tag{5.5}$$

当 $k=0$ 时,

$$\begin{aligned}\theta(z^0+\lambda_0\Delta z^0)\le&\{1-2\sigma[1-(a+b)\gamma\overline{\mu}]\lambda_0\}\theta(z^0)\\ =&\{1-2\sigma[1-(a+b)\gamma\overline{\mu}]\lambda_0\}C_0.\end{aligned} \tag{5.6}$$

现假设

$$\theta(z^k+\lambda_k\Delta z^k)\le\{1-2\sigma[1-(a+b)\gamma\overline{\mu}]\lambda_k\}C_k. \tag{5.7}$$

这蕴含 $\theta(z^{k+1})\le C_k$, 此外,

$$C_{k+1}=\frac{\tau\cdot m(k+1)C_k+\theta(z^{k+1})}{\tau\cdot m(k+1)+1}\ge\theta(z^{k+1}),\qquad C_{k+1}\le C_k.$$

结合 (5.5), 我们有

$$\theta(z^{k+1}+\lambda_{k+1}\Delta z^{k+1})\le\{1-2\sigma[1-(a+b)\gamma\overline{\mu}]\lambda_{k+1}\}C_{k+1}.$$

上面的引理表明 C_k 是单调下降的.

引理 5.2 如果假设 1 和 2 成立且 $\{z^k\}$ 是由算法 1 产生. 如果 $\mu^k>0,\ \mu^k\ge\beta(z^k)\overline{\mu}$, 则 $\mu^{k+1}>0,\ \mu^{k+1}\ge\beta(z^{k+1})\overline{\mu}$.

证 因为 $\mu^k>0,\ \beta(z^k)>0$. 对任意 $\lambda\in(0,1]$, 我们有

$$\mu^k+\lambda\Delta\mu^k=\mu^k+\lambda\Big[\beta(z^k)\overline{\mu}-\frac{k(\mu^k)}{k'(\mu^k)}\Big]\ge(1-\lambda)\mu^k+\lambda\beta(z^k)\overline{\mu}>0. \tag{5.8}$$

因为 $\beta(z^{k+1}) \le \beta(z^k)$, 从而得

$$\begin{aligned}\mu^{k+1} - \beta(z^{k+1})\overline{\mu} =& \mu^k + \lambda_k \Delta\mu^k - \beta(z^{k+1})\overline{\mu} \\ \ge& (1-\lambda_k)\mu^k + \lambda_k \beta(z^k)\overline{\mu} - \beta(z^{k+1})\overline{\mu} \\ \ge& (1-\lambda_k)\beta(z^k)\overline{\mu} + \lambda_k \beta(z^k)\overline{\mu} - \beta(z^{k+1})\overline{\mu} \\ \ge& 0. \end{aligned} \tag{5.9}$$

定理 5.2 如果假设 1 和 2 成立且 $\{z^k\}$ 是由算法 1 产生，则 $\{z^k\}$ 的每一个聚点都是 (1.1) 的解.

证 为了证明这个结论，设 z^* 是 z^k 的聚点. 由算法 1 和引理 5.1, 有 $\{z^k\} \subseteq \mathcal{L}(z^0)$. 因此，存在 $\{z^{k_l}\} \subseteq \{z^k\}$, 使得

$$z^{k_l} \to z^*, \quad \beta(z^{k_l}) \to \beta^*, \quad \mu^{k_l} \to \mu^*, \quad C_{k_l} \to C^*.$$

假设 z^* 不是 (1.1) 的解.

我们考虑下面的情况.

i) 假设 $\inf \lambda_{k_l} \ge \rho > 0$. 由定理 5.1 的证明，有 $0 \le C_{k_{l+1}} \le C_{k_{l+1}-1} \le C_0$. 结合 (4.7), 有

$$\begin{aligned} & C_{k_{l+1}} \le C_{k_{l+1}-1} \\ & \qquad \vdots \\ \le& C_{k_l+1} \\ =& \frac{\tau \cdot m(k_l+1)C_{k_l} + \theta(z^{k_l+1})}{\tau \cdot m(k_l+1)+1} \\ \le& C_{k_l} - \frac{2\sigma[1-(a+b)\gamma\overline{\mu}]\lambda_{k_l}}{\tau \cdot m(k_l+1)+1} C_{k_l} \\ \le& C_{k_l} - \frac{2\sigma[1-(a+b)\gamma\overline{\mu}]\rho}{\tau \cdot m(k_l+1)+1} C_{k_l}. \end{aligned} \tag{5.10}$$

因为 $\{C_{k_l}\}$ 是有界的，我们可得

$$\sum_{l=0}^{\infty} \frac{2\sigma[1-(a+b)\gamma\overline{\mu}]\rho}{\tau \cdot m(k_l+1)+1} C_{k_l} < \infty.$$

所以

$$\lim_{l\to\infty} \frac{C_{k_l}}{\tau \cdot m(k_l+1)+1} = \frac{C^*}{\tau M+1} = 0.$$

从定理 5.1 的证明，有

$$\theta(z^{k_{l+1}}) \le C_{k_{l+1}-1} \le \cdots \le C_{k_l}. \tag{5.11}$$

对 (5.11) 的两边同时取极限，有

$$\theta(z^*) \le 0, \tag{5.12}$$

这和 $\theta(z^*)>0$ 矛盾.

ii) 假设 $\inf\lambda_k=0$, 则存在 $\{z^{k_p}\}\subseteq\{z^{k_l}\}$, 使得

$$\lim_{k_p\to\infty}\lambda_{k_p}=0.$$

在这种情况下， $\overline{\lambda}_{k_p}=\frac{\lambda_{k_p}}{\delta}$ 不满足线搜索 (4.7) 对充分大的 k_p, i.e.

$$\begin{aligned}\theta(z^{k_p}+\overline{\lambda}_{k_p}\Delta z^{k_p})\geq&\{1-2\sigma[1-(a+b)\gamma\overline{\mu}]\overline{\lambda}_{k_p}\}C_{k_p}\\ \geq&\{1-2\sigma[1-(a+b)\gamma\overline{\mu}]\overline{\lambda}_{k_p}\}\theta(z^{k_p}).\end{aligned}\tag{5.13}$$

因此

$$\frac{\theta(z^{k_p}+\overline{\lambda}_{k_p}\Delta z^{k_p})-\theta(z^{k_p})}{\overline{\lambda}_{k_p}}\geq-2\sigma[1-(a+b)\gamma\overline{\mu}]\theta(z^{k_p}).\tag{5.14}$$

由引理 5.2, 有 $\mu^*\geq\beta^*\overline{\mu}>0$. 这是根据 $\Psi(z)$ 在 z^* 是连续可微的．对 (5.14 的两边同时取极限，有

$$\begin{aligned}-2\sigma[1-(a+b)\gamma\overline{\mu}]\theta(z^*)\leq&2k(\mu^*)k'(\mu^*)\Delta\mu^*-2\Psi(z^*)\\ \leq&2k(\mu^*)[k'(\mu^*)\beta(z^*)\overline{\mu}-k(\mu^*)]-2\Psi(z^*)\\ \leq&-2[1-(a+b)\gamma\overline{\mu}]\theta(z^*).\end{aligned}$$

这表明 $\sigma\geq1$, 这和 $\sigma\in(0,1)$ 矛盾.

接下来我们讨论算法 1 的超线性和二次收敛性．众所周知 H_p 的半光滑和强半光滑是局部收敛性分析中的关键．由性质 3.2, 不难得到下面的性质.

性质 5.1 由 (4.1) 定义的映射 $H_p:R_+\times R^n$ 是半光滑的．此外，如果 F' 是局部 Lipschitz 连续，该映射是强半光滑的.

定理 5.3 如果假设 1 和 2 成立， z^* 是由算法 1 产生的无限序列 $\{z^k\}$ 的聚点，且 $V\in\partial H_p(z^*)$ 是非奇异的，则序列 $\{z^k\}$ 是超线性收敛．此外，如果 F' 在 x^* 的周围是局部 Lipschitz 连续，则收敛速度是二次的.

证 证明类似于定理 6.5.10 (见 [10]).

6 数值试验

在整个数值试验中，我们采用了如下的参数：

$$p=1.5,\ \delta=0.5,\ \sigma=0.00005,\tau=0.3,\ \overline{\mu}=0.1,\ \gamma=\min(0.9,1/(a+b)\overline{\mu}-0.00001),$$

这里 a 和 b 根据 $k(\mu)$ 的选取进行调整.

如果下面的条件有一个满足就终止算法：

$$\|H_p(z^k)\|\leq10^{-12}\qquad\text{或}\qquad k\geq k_{\max}=100.$$

我们把算法 1 和算法 6.5.3 (记为 SMND)[10] 进行了比较, 数值结果可见表 1– 表 5. 这里的 n 表示问题的维数; x_i^0 代表的是初始点 x^0 的第 i 个分量; $\theta_1(x^t)$ 代表的是 $\theta_1(\cdot)$ 在最后的迭代点 x^t 处的函数值, $\theta_1(x)=\|H_2(0,x)\|^2$.

问题 1[12] 设 $f(x)=Mx+q$, 这里

$$M=\begin{pmatrix}1 & 2 & \cdots & 2 & 2\\ 0 & 1 & \cdots & 2 & 2\\ \vdots & \vdots & \ddots & & \vdots\\ \vdots & \vdots & & 1 & 2\\ 0 & 0 & 0 & \cdots & 0\end{pmatrix},\qquad q=-\begin{pmatrix}1\\1\\ \vdots\\1\\0\end{pmatrix}.$$

表 1 问题 1 的数值结果

			算法 1			SMND		
$k(\mu)$	n	x_i^0	$\theta_1(x^t)$	Iter	CPU	$\theta_1(x^t)$	Iter	CPU
$e^\mu-1$	200	200	3.207e-13	85	11.203	4.930e-03	101	13.813
$\mu^2+2\mu$	150	i	4.206e-13	85	5.281	4.801e-03	101	6.578
$1.5^\mu-1$	100	100	9.590e-17	7	0.172	4.487e-03	101	2.563
	150	150	4.205e-13	85	5.516	5.049e-03	101	6.484
$30^\mu-1$	250	i	4.535e-13	78	17.547	4.997e-03	101	26.500
$\mu\ln(1+\mu)$	50	0	2.023e-19	30	0.188	3.231e-17	70	0.422
	50	1	3.134e-20	13	0.078	5.944e-20	17	0.109
	50	i	1.460e-18	13	0.063	4.955e-03	101	0.609
	50	50	1.138e-18	13	0.078	4.338e-03	101	0.641
	300	300	6.445e-13	77	31.281	4.528e-03	101	42.578
$e^\mu-1+\mu$	400	400	7.647e-13	77	89.563	4.781e-03	101	105.016
$\frac{1}{2}\mu^3+\mu^2+2\mu$	150	150	4.187e-13	85	5.203	5.049e-03	101	6.422

表 2 问题 2 的数值结果

			算法 1			SMND		
$k(\mu)$	n	x_i^0	$\theta_1(x^t)$	Iter	CPU	$\theta_1(x^t)$	Iter	CPU
$2^\mu-1$	500	0	1.395e-15	29	56.750	6.087e-16	49	111.922
	500	1	2.393e-16	18	33.516	2.286e-17	47	100.688
	500	500	1.048e-22	20	36.828	8.339e-19	62	141.063
	500	i	2.635e-16	20	35.516	1.313e-13	41	96.844
$\mu^2+2\mu$	100	i	5.435e-20	14	0.297	3.822e-20	21	0.469
	100	1	2.742e-15	12	0.234	1.022e-14	22	0.469
	100	100	4.853e-18	13	0.266	2.604e-19	13	0.297
$\frac{1}{2}\mu^2+\mu$	50	i	8.819e-16	65	0.281	4.338e-03	101	0.609
	50	50	8.819e-16	65	0.297	4.338e-03	101	0.594
$\mu(2^\mu-1)$	200	0	5.131e-18	18	2.234	1.783e-13	24	3.531
	200	1	2.119e-18	14	1.688	6.898e-16	23	2.719
	200	i	5.479e-17	15	1.766	2.748e-18	34	4.703
	200	200	2.463e-18	15	2.016	1.385e-17	26	3.391
$(\mu+1)\ln(\mu+1)$	50	1	1.175e-13	12	0.047	4.416e-23	15	0.063
	50	i	4.082e-17	11	0.047	4.563e-14	18	0.078

问题 2[12] 随机问题：

$$M=\begin{pmatrix} P+D_1 & P+D_2 \\ -I & 0 \end{pmatrix},\qquad q=-\begin{pmatrix} c-e \\ 0 \end{pmatrix}.$$

这里 $e\in R^{n/2}$, $P=A^TA$, $A\in R^{n/2\times n/2}$ 且 $0<a_{ij}<1$. D_1 和 D_2 都是对角阵且 $0\le (D_1)_{ij},(D_2)_{ij}\le 3$.

问题 3[13] NCP(F) 规模为 n, 这里 $F:R^n\to R^n$ 定义为

$$F_{2i-1}=10(x_{2i}-x_{2i-1}^2),\qquad F_{2i}=x_{2i-1}(1-x_{2i-1}),\qquad 1\le i\le n/2.$$

表 3 问题 3 的数值结果

			算法 1			SMND		
$k(\mu)$	n	x_i^0	$\theta_1(x^t)$	Iter	CPU	$\theta_1(x^t)$	Iter	CPU
$3^\mu-1$	300	i	1.866e-13	23	2.453	2.093e-13	23	2.547
$\mu^2+2\mu$	150	1	1.850e-24	5	0.047	2.177e-13	7	0.063
$30^\mu-1$	250	1	2.175e-23	5	0.250	4.739e-18	6	0.328
$e^\mu-1+\mu$	400	0	2.264e-13	21	5.313	1.982e-13	28	6.906
$\frac{1}{2}\mu^3+\mu^2+2\mu$	150	1	1.873e-24	5	0.047	2.177e-13	5	0.056
	400	i	5.047e-13	23	5.767	1.029e-13	22	5.781
$(\mu+1)\ln(\mu+1)$	300	1	1.717e-22	6	0.609	1.595e-22	8	0.723

问题 4[13] NCP(F) 规模为 n, 这里 $F:R^n\to R^n$ 定义为

$$F_{2i-1}=e^{x_{2i-1}}+x_{2i-1}-2,\qquad F_{2i}=e^{x_{2i}}-2^{x_{2i}}-2x_{2i}^2,\qquad 1\le i\le n/2.$$

表 4 问题 4 的数值结果

			算法 1			SMND		
$k(\mu)$	n	x_i^0	$\theta_1(x^t)$	Iter	CPU	$\theta_1(x^t)$	Iter	CPU
$e^\mu-1$	200	0	1.298e-15	9	0.266	8.022e-15	11	0.328
$1.5^\mu-1$	100	i	1.484e-08	101	0.313	1.054e+01	101	0.453
$30^\mu-1$	250	0	5.441e-16	8	0.438	2.083e-18	12	0.672
$\frac{1}{2}\mu^2+\mu$	150	0	2.562e-13	8	0.078	6.233e-15	11	0.125
$e^\mu-1+\mu$	400	0	6.808e-18	10	2.297	7.106e-18	12	2.922
$3^\mu-1$	300	0	3.237e-16	11	1.109	3.047e-18	12	1.250

问题 5[10] 设 $f(x)=Mx+q$, 这里

$$\begin{aligned} &M_{ii}=4(i-1)+1,\qquad i=1,\cdots,n,\\ &M_{ij}=M_{ii}+1,\qquad i=1,\cdots,n-1,\quad j=i+1,\cdots,n,\\ &M_{ij}=M_{jj}+1,\qquad j=1,\cdots,n-1,\quad i=j+1,\cdots,n,\\ &q=(-1,-1,-1,\cdots,-1)^T. \end{aligned}$$

表 5 问题 5 的数值结果

$k(\mu)$	n	x_i^0	算法 1 $\theta_1(x^t)$	算法 1 Iter	算法 1 CPU	SMND $\theta_1(x^t)$	SMND Iter	SMND CPU
$3^\mu - 1$	300	1	1.946e-16	9	3.563	1.264e-22	18	7.234
$e^\mu - 1$	200	1	2.500e-16	9	1.219	1.037e-14	17	2.359
$30^\mu - 1$	250	1	1.420e-13	9	2.156	1.292e-24	18	4.391
$1.5^\mu - 1$	100	i	1.112e-16	10	0.219	2.460e-16	10	0.234
$\mu^2 + 2\mu$	150	1	2.225e-16	9	0.578	8.914e-20	16	1.063
$\frac{1}{2}\mu^2 + \mu$	50	1	5.820e-16	9	0.047	1.967e-17	14	0.078
$e^\mu - 1 + \mu$	400	1	3.106e-16	9	10.453	9.182e-20	19	21.984
$\frac{1}{2}\mu^3 + \mu^2 + 2\mu$	150	1	2.102e-16	9	0.563	8.914e-20	16	1.141
	150	1	2.986e-16	9	0.594	8.914e-20	16	1.094
	150	150	4.668e-15	10	0.594	1.124e-24	11	0.625
	100	1	9.584e-16	9	0.219	3.029e-13	15	0.359
	100	100	1.405e-16	10	0.250	2.460e-15	10	0.203
	200	200	9.833e-13	10	1.266	3.411e-24	11	1.266
	150	150	5.190e-15	10	0.609	1.124e-24	11	0.641
$(\mu+1)\ln(\mu+1)$	300	1	2.676e-16	9	3.875	1.264e-22	18	8.016

从上面的数值结果，我们发现，如果对不同的问题选取适当的 $k(\mu)$, 算法 1 大大的减少了运行时间和迭代次数.

7 结论

在本文，基于广义 Fischer-Burmeister 函数，我们提出了一新的光滑函数，并且推广光滑牛顿法到一族非单调光滑牛顿法. 通过分析，在 $P_0 + R_0$ 性质下，我们可以证明全局和局部收敛性. 数值结果表明我们的方法是有效的.

参 考 文 献

[1] Chen Bintong, Chen Xiaojun, Christian Kanzow. A Penalized Fischer-Burmeister NCP-function. *Mathematical Programming*, 2000, 88: 211–216

[2] Chen Xiaojun, Ye Yinyu. On Homotopy-smoothing Methods for Variational Inequalities. *SIAM Journal Control Optimization*, 1999, 37: 589–616

[3] James V, Burke, Xu Song. The Global Linear Convergence of a Noninterior Path-following Algorithm for Linear Complementarity Problems. *Mathematics of Operations Research*, 1998, 23: 719–734

[4] Chen Bintong, Patrick T Harker. Smooth Approximations to Nonlinear Complementarity Problems. *SIAM Journal on Optimization*, 1997, 7: 403–420

[5] Fang Liang. A New One-step Smoothing Newton Method for Nonlinear Complementarity Problem with p0-function. *Applied Mathematics and Computation*, 2010, 216: 1087–1095

[6] Chen Xiaojun, Ye Yinyu. On Smoothing Methods for the p0 Matrix Linear Complementarity Problem. *SIAM Journal on Optimization*, 2000, 11: 341–363

[7] Ma Changfeng, Chen Xiaohong. The Convergence of a One-step Smoothing Newton Method for p0-NCP Based on a New Smoothing NCP-function. *Journal of Computational and Applied Mathematics*, 2008, 216: 1–13

[8] Chen J, Pan S. A Family of NCP Functions and a Descent Method for the Nonlinear Complementarity Problem. *Computational Optimization and Applications*, 2008, 40: 389–404

[9] Chen Jein-Shan. The Semismooth-related Properties of a Merit Function and a Descent Method for the Nonlinear Complementarity Problem. *Journal of Global Optimization*, 2006, 36

[10] Han Jiye, Xiu Naihua, Qi Houduo. Theories and Algorithms for Nonlinear Complementarity Problems. Shanghai: Shanghai Science and Technology Press, 2006

[11] Patrick T Harker, Pang Jong-Shi. A Damped Newton Method for the Linear Complementarity Problem. In: Computational Solution of Nonlinear Systems of Equations, 1990

[12] Qi Houduo, Qi Liqun, Sun Defeng. Solving KKT Systems via the Trust Region and the Conjugate Gradient Methods. Technical Report, Applied Mathematics Research Report AMR99/19, School of Mathematics, University of New South, 1999

[13] Wu Caiying, Chen Guoqing. A Smoothing Conjugate Gradient Algorithm for Nonlinear Complementarity Problems. *Journal of Systems Science and Systems Engineering*, 2008, 17

A Family of Nonmonotone Smoothing Newton Method for Complementarity Problems

LI Xiangli

(*College of Mathematics and Computing Science, Guilin University of Electronic Technology, Guilin* 541004)

(*E-mail: lixiangli213@gmail.com*)

LIU Hongwei

(*Department of Applied Mathematics, Xidian University, Xi'an* 710071)

Abstract In this paper, based on the generalized Fischer-Burmeister function, we propose non-monotone smoothing Newton method for solving the complementarity problem. Global and local convergence of our method are established under reasonable conditions. The preliminary numerical results are also reported.

Key words complementarity problem; a family of non-monotone smoothing Newton method; global convergence; local convergence

MR(2000) Subject Classification 90

Chinese Library Classification O221