第26卷第1期 数 学 研 究 与 评 论 Vol.26, No.1
2006年2月 JOURNAL OF MATHEMATICAL RESEARCH AND EXPOSITION Feb., 2006

文章编号: 1000-341X(2006)01-0179-10 **文献标识码**: **A**

求解非线性互补问题的逐次逼近拟牛顿法的收敛性分析

马昌凤 [1,2]
(1. 桂林电子工业学院计算科学与数学系，广西 桂林 541004;
2. 浙江师范大学数理学院，浙江 金华 321004)
(E-mail: macf62@163.com)

摘 要: 提出了求解非线性互补问题的一个逐次逼近拟牛顿算法. 在适当的假设下，证明了该算法的全局收敛性和局部超线性收敛性.

关键词: 非线性互补问题；逐次逼近；拟牛顿法；收敛性分析.
MSC(2000): 90C33
中图分类: O224.2

1 引言和算法

本文考虑如下模型问题的数值方法：求 $x \in R^n$, 使得

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

其中 $F : R^n \to R^n$ 是连续可微的非线性映射. 问题 (1.1) 通常称为非线性互补问题，简记为 NCP(F).

非线性互补问题是变分不等式的重要的类型之一, 主要来源于经济领域中的实际问题 (如平衡问题等) 和物理，力学及工程领域中的有关问题 (一般为无穷维问题的离散化)[1−3]. 由于其应用的广泛性, 对互补问题的数值解法的研究近年来很受重视. 牛顿型方法是求解非线性互补问题的一类重要的数值迭代算法，其局部收敛性质的研究取得了很好的成果 [2]. 20 世纪 90 年代中期以来，此类算法的全局收敛性研究也得到了诸多进展 [4−8]. 然而，对于计算上更为实用的拟牛顿算法的研究目前尚不多见. 本文提出一种求解非线性互补问题的逐次逼近拟牛顿算法, 在适当的线搜索方式下, 此算法是适定的. 而且在一定的条件下, 算法具有全局收敛性和局部超线性收敛速度.

借助于 Fischer-Burmeister 函数 [9]

$$\varphi(a,b) = \sqrt{a^2+b^2} - a - b,$$

问题 (1.1) 等价于下面的非光滑方程组：

$$H(x) = \big(\varphi(x_1,\ F_1(x)),\ \cdots\cdots,\ \varphi(x_n,\ F_n(x))\ \big)^T = 0. \tag{1.2}$$

收稿日期: 2003-08-20
基金项目: 广西自然科学基金 (0448075), 中国博士后科学基金 (2004036133), 浙江师范大学博士科学基金.

此处定义的函数 $H:\ R^n\to R^n$ 是半光滑的 [6]. 由于 H 不是 Frechét 可微的，故而经典的牛顿法用于求解 (1.2) 存在困难. 但我们可对 H 作如下分解:

$$H(x)=f_k(x)+g_k(x), \tag{1.3}$$

其中 f_k 可微而 g_k 的模相对地小. 事实上，对任何 $\mu_k>0$, $f_k(x)$ 和 $g_k(x)$ 的分量可定义如下:

$$[f_k(x)]_i=\begin{cases} H_i(x), & i\in A_k(x),\\ \dfrac{x_i-2\mu_k}{2\mu_k}x_i+\dfrac{F_i(x)-2\mu_k}{2\mu_k}F_i(x)+\dfrac{\mu_k}{2}, & i\in B_k(x),\end{cases} \tag{1.4}$$

$$[g_k(x)]_i=\begin{cases} 0, & i\in A_k(x),\\ \dfrac{(\sqrt{x_i^2+F_i^2(x)}-\mu_k)^2}{2\mu_k}, & i\in B_k(x),\end{cases} \tag{1.5}$$

其中 $H_i(x)=\varphi(x_i,F_i(x))=\sqrt{x_i^2+F_i^2(x)}-x_i-F_i(x)$, $A_k(x)=\{i:\sqrt{x_i^2+F_i^2(x)}\geq\mu_k\}$, $B_k(x)=\{i:\sqrt{x_i^2+F_i^2(x)}<\mu_k\}$. 显然，只要 μ_k 充分小，就可保证 $\|g_k\|$ 足够小，且有如下估计式

$$\|H(x)-f_k(x)\|=\|g_k\|\leq\frac{\sqrt{n}}{2}\mu_k. \tag{1.6}$$

不难证明， $f_k(x)$ 是可微的，其在点 x 处的 Jacobian 矩阵 $\nabla f_k(x)=([\nabla f_k(x)]_1,[\nabla f_k(x)]_2,\cdots,[\nabla f_k(x)]_n)^T$ 元素为:

$$[\nabla f_k(x)]_i=\begin{cases} (1-\dfrac{x_i}{\sqrt{x_i^2+F_i^2(x)}})e_i+(1-\dfrac{F_i(x)}{\sqrt{x_i^2+F_i^2(x)}})\nabla F_i(x), & i\in A_k(x),\\ (1-\dfrac{x_i}{\mu_k})e_i+(1-\dfrac{F_i(x)}{\mu_k})\nabla F_i(x), & i\in B_k(x),\end{cases} \tag{1.7}$$

现在可以建立本文的拟牛顿算法:

算法 1(QNM 算法) 给定 ρ, $\alpha\in(0,1)$, $x_0\geq0$, B_0 非奇异， $\kappa=\frac{\sqrt{n}}{2}$, $\mu_0=\frac{\alpha}{2\kappa}\|H(x_0)\|$, 令 $k:=0$.

1. 求解如下线性方程组，确定搜索方向 d_k:

$$H(x_k)+B_kd_k=0. \tag{1.8}$$

2. 定义函数

$$q_k(\lambda)=\frac{(1+2^{-(k+1)})\|f_k(x)\|^2-f_k(x_k)^Tf_k(x_k+\lambda d_k)}{\max\{\|f_k(x_k)-f_k(x_k+\lambda d_k)\|^2,\min\{\frac{3}{4}\|f_k(x_k)\|^2,\|\lambda d_k\|^2\}\}},$$

确定 $\lambda_k=\rho^{i_k}$, 其中 i_k 是满足下式的最小非负整数:

$$q_k(\rho^{i_k})\geq\frac{1}{2}+\varepsilon. \tag{1.9}$$

此处 ε 为某正常数.

3. 令 $x_{k+1}:=x_k+\lambda_kd_k$. 若 $\|(x_{k+1})\|=0$, 停算，否则转步 4.

4. 由 Broyden 类公式修正 B_k 得到 B_{k+1}, 即

$$B_{k+1} = B_k + \varphi_k \frac{(y_k - B_k s_k)s_k^T}{\|s_k\|^2}, \tag{1.10}$$

此处 $s_k = x_{k+1} - x_k = \lambda_k d_k$, $y_k = f_k(x_{k+1}) - f_k(x_k)$, 且选择参数 φ_k 满足 $|\varphi_k - 1| \le \varphi$ (其中常数 $\varphi \in (0,1)$).

5. 若 $\mu_k < \alpha\|H(x_{k+1})\|$, 令 $\mu_{k+1} := \mu_k$; 否则，选择 μ_{k+1} 满足

$$\mu_{k+1} \le \min\{\frac{\alpha}{2\kappa}\|H(x_{k+1})\|, \frac{1}{2}\mu_k\}, \tag{1.11}$$

$k := k+1$ 转步 1.

2. 全局收敛性

本节我们讨论算法 1 的全局收敛性，先证明下面的引理.

引理 1 设序列 $\{x_k\}$ 由算法 1 生成，则

$$\|f_k(x_{k+1})\| \le (1 + 2^{-k})\|f_k(x_k)\|. \tag{2.1}$$

证明 由线搜索 (1.9), 成立 $2q_k(\lambda_k) \ge 2\varepsilon + 1)$, 可得

$$\begin{aligned}
&(2 + 2^{-k})\|f_k(x_k)\|^2 - 2f_k(x_k)^T f_k(x_{k+1}) \\
&\quad \ge (2\varepsilon + 1)\max\{\|f_k(x_k) - f_k(x_{k+1})\|^2, \min\{\frac{3}{4}\|f_k(x_k)\|^2, \|\lambda_k d_k\|^2\}\} \\
&\quad \ge (2\varepsilon + 1)\|f_k(x_k) - f_k(x_{k+1})\|^2 \\
&\quad = (2\varepsilon + 1)(\|f_k(x_k)\|^2 - 2f_k(x_k)^T f_k(x_{k+1}) + \|f_k(x_{k+1})\|^2),
\end{aligned}$$

即

$$\begin{aligned}
(2\varepsilon + 1)\|f_k(x_{k+1})\|^2 + (2\varepsilon - 1 - 2^{-k})\|f_k(x_k)\|^2 &\le 2\varepsilon f_k(x_k)^T f_k(x_{k+1}) \\
&\le 2\varepsilon\|f_k(x_k)\|\|f_k(x_{k+1})\|.
\end{aligned}$$

故有

$$\begin{aligned}
\|f_k(x_{k+1})\| &\le \frac{\varepsilon + \sqrt{\varepsilon^2 - (2\varepsilon+1)(2\varepsilon - 1 - 2^{-k})}}{2\varepsilon + 1}\|f_k(x_k)\| \\
&\le \frac{2\varepsilon + \sqrt{1 + (2\varepsilon+1)2^{-k}}}{2\varepsilon + 1}\|f_k(x_k)\| = (1 + \frac{\sqrt{1 + (2\varepsilon+1)2^{-k}} - 1}{2\varepsilon + 1})\|f_k(x_k)\| \\
&= (1 + \frac{2^{-k}}{1 + \sqrt{1 + (2\varepsilon+1)2^{-k}}})\|f_k(x_k)\| \le (1 + 2^{-k})\|f_k(x_k)\|.
\end{aligned}$$

引理 2 设序列 $\{x_k\}$ 由算法 1 生成，则

$$\{x_k\} \subset \Omega = \{x : \|H(x) \le 4\|H(x_0)\|\}.$$

证明 由算法 1 步 5 知， $\{\mu_k\}$ 是单调下降的．设

$$K=\{0\}\cup\{k:\mu_k\geq\alpha\|H(x_{k+1})\|\}=\{k_0,k_1,k_2,\cdots\}.$$

不失一般性, 假定 $0=k_0<k_1<k_2<\cdots$, 对任意的非负整数 k, 设 k_j 是 K 中满足条件 $k_j<k$ 的最大整数，则有 $f_k=f_{k_j}$ 及

$$\begin{aligned}&\|H(x_k)\|\leq\|f_k(x_k)\|+\|g_k(x_k)\|\\&\quad=\|f_{k_j}(x_k)\|+\|H(x_k)-f_{k_j}(x_k)\|\leq(1+2^{-k})\|f_{k_j}(x_k)\|+\kappa\mu_{k_j}\\&\quad\leq(1+2^{-k})\|H(x_{k_j})\|+(1+2^{-k})\|H(x_{k_j})-f_{k_j}(x_{k_j})\|+\kappa\mu_{k_j}\\&\quad\leq(1+2^{-k_j})\|H(x_{k_j})\|+(2+2^{-k_j})\kappa\mu_{k_j}.\end{aligned}$$

若 $j=0$, 则

$$\begin{aligned}&\|H(x_k)\|\leq2\|H(x_0)\|+3\kappa\mu_0\leq2\|H(x_0)\|+3\kappa\frac{\alpha}{2\kappa}\|H(x_0)\|\\&\quad\leq2(1+\alpha)\|H(x_0)\|\leq4\|H(x_0)\|;\end{aligned}$$

若 $j>0$, 则有

$$\begin{aligned}\|H(x_k)\|&\leq(1+2^{-k_j})\frac{1}{\alpha}\mu_{k_j-1}+(2+2^{-k_j})\frac{1}{2}\kappa\mu_{k_j-1}\\&\leq(\frac{1+2^{-k_j}}{\alpha}+\frac{2+2^{-k_j}}{2}\kappa)\mu_{k_{j-1}}\leq\cdots\\&\leq(\frac{1+2^{-k_j}}{\alpha}+\frac{2+2^{-k_j}}{2}\kappa)\frac{1}{2^{j-1}}\mu_{k_{0-1}}\\&\leq(\frac{1+2^{-k_j}}{\alpha}+\frac{2+2^{-k_j}}{2}\kappa)\frac{1}{2^{j-1}}\frac{\alpha}{2\kappa}\|H(x_0)\|\\&\leq(1+2^{-k_j})(1+\alpha)\frac{1}{2^{j-1}}\|H(x_0)\|\\&\leq2(1+\alpha)\frac{1}{2^{j-1}}\|H(x_0)\|.\end{aligned}\tag{2.2}$$

因此， $\{x_k\}\subset\Omega.$ □

下面的论述中需要用到序列 Lipschitz 连续可微的概念，我们先叙述其定义如下.

定义 1 给定函数 $f:R^n\to R^n$ 及序列 $\{x_k\}$, 若存在常数 $L>0$ 使得当 K 充分大时，成立

$$\begin{aligned}&\|f(x_k+ts_k)-f(x_{k-1}+ts_{k-1})\|\leq L\|(x_k+ts_k)-(x_{k-1}+ts_{k-1})\|\\&\qquad\qquad\leq L(t\|s_k\|+(1-t)\|s_{k-1}\|),\quad\forall t\in(0,1),\end{aligned}$$

则称函数 f 关于 $\{x_k\}$ 是序列 Lipschitz 连续的，其中 $s_k=x_{k+1}-x_k$. 如果 ∇f 关于 $\{x_k\}$ 是序列 Lipschitz 连续的，即存在常数 $L>0$, 使得当 k 充分大时，成立

$$\|\nabla f(x_k+ts_k)-\nabla f(x_{k-1}+ts_{k-1})\|\leq L(t\|s_k\|+(1-t)\|s_{k-1}\|),\quad\forall t\in(0,1),\tag{2.3}$$

则称函数 f 关于 $\{x_k\}$ 是序列 Lipschitz 连续可微的，其中 $\|\cdot\|_F$ 是矩阵的 Frobenius 范数.

由上述定义不难知道， f 的 Lipschitz 连续可微性蕴含了 f 关于 $\{x_k\}$ 的序列 Lipschitz 连续可微性.

引理 3 设序列 $\{x_k\}$ 由算法 1 产生，集合 K 是有限集，且 $\lim_{k\to\infty}\|x_{k+1}-x_k\|=0$. 则当函数 F 在 $\bar{\Omega}=\{tx+(1-t)y: t\in[0,1], x,y\in\Omega\}$ 上 Lipschitz 连续可微时，对充分大的 k, f_k 关于 $\{x_k\}$ 是序列 Lipschitz 连续可微的，且 Lipschitz 常数 L 与 k 无关.

证明 因 $K=\{0\}\cup\{k:\mu_k\ge\alpha\|H(x_{k+1})\|\}$ 是有限集，故由算法 1 步 5 知，当 k 充分大时，μ_k 是一个常数, 不失一般性, 设 $\mu_k=\mu>0$, 则 $f_k\equiv f, g_k\equiv g, A_k(z)=A(z), B_k(z)=B(z)$. 记 $\tilde{x}_k=x_k+ts_k$, $N=\{1,2,\cdots,n\}$, $(\nabla f(\tilde{x}_k)-\nabla f(\tilde{x}_{k-1}))_i$ 表示 $\nabla f(\tilde{x}_k)-\nabla f(\tilde{x}_{k-1})$ 的第 i 列, 则由定义 1, 只需证明对任何 $t\in(0,1)$, 存在常数 $L>0$, 使得下式成立：

$$\|(\nabla f(\tilde{x}_k)-\nabla f(\tilde{x}_{k-1}))_i\|\le L\|\tilde{x}_k-\tilde{x}_{k-1}\|\le L(t\|s_k\|+(1-t)\|s_{k-1}\|). \tag{2.4}$$

事实上，对于所有的 k, 我们分四种情形讨论：

(i) 若 $i\in A(\tilde{x}_k)\cap A(\tilde{x}_{k-1})$, 此时 f 是 Lipschitz 连续可微的，因而也是关于 $\{\tilde{x}_k\}$ 序列 Lipschitz 连续可微的，故 (2.4) 成立.

(ii) 若 $i\in A(\tilde{x}_k)\cap B(\tilde{x}_{k-1})$, 则

$$\begin{aligned}
&\|(\nabla f(\tilde{x}_k)-\nabla f(\tilde{x}_{k-1}))_i\|\le\|(1-\frac{\tilde{x}_k^i}{\sqrt{(\tilde{x}_k^i)^2+F_i^2(\tilde{x}_k)}})e_i-(1-\frac{\tilde{x}_{k-1}^i}{\mu})e_i\|+\\
&\quad\|(1-\frac{F_i(\tilde{x}_k)}{\sqrt{(\tilde{x}_k^i)^2+F_i^2(\tilde{x}_k)}})\nabla F_i(\tilde{x}_k)-(1-\frac{F_i(\tilde{x}_{k-1})}{\mu})\nabla F_i(\tilde{x}_{k-1})\|\\
&\le\frac{1}{\mu}\|(\tilde{x}_k^i-\tilde{x}_{k-1}^i)e_i\|+\|\nabla F_i(\tilde{x}_k)-\nabla F_i(\tilde{x}_{k-1})\|+\frac{1}{\mu}\|F_i(\tilde{x}_k)\nabla F_i(\tilde{x}_k)-F_i(\tilde{x}_{k-1})\nabla F_i(\tilde{x}_{k-1})\|\\
&\le L_1\|\tilde{x}_k-\tilde{x}_{k-1}\|+L_2\|\tilde{x}_k-\tilde{x}_{k-1}\|+\frac{1}{\mu}\|[F_i(\tilde{x}_{k-1})+\\
&\quad\nabla F_i(\tilde{x}_{k-1}+\theta(\tilde{x}_k-\tilde{x}_{k-1}))(\tilde{x}_k-\tilde{x}_{k-1})]\nabla F_i(\tilde{x}_k)-F_i(\tilde{x}_{k-1})\nabla F_i(\tilde{x}_{k-1})\|\\
&\le(L_1+L_2)\|\tilde{x}_k-\tilde{x}_{k-1}\|+\frac{1}{\mu}\|F_i(\tilde{x}_{k-1})\|\|\nabla F_i(\tilde{x}_k)-\nabla F_i(\tilde{x}_{k-1})\|+\\
&\quad\frac{1}{\mu}\|\nabla F_i(\tilde{x}_{k-1}+\theta(\tilde{x}_k-\tilde{x}_{k-1}))\|\|\nabla F_i(\tilde{x}_k)\|\|\tilde{x}_k-\tilde{x}_{k-1}\|\\
&\le(L_1+L_2+L_3+L_4)\|\tilde{x}_k-\tilde{x}_{k-1}\|=L\|\tilde{x}_k-\tilde{x}_{k-1}\|
\end{aligned}$$

(iii) 若 $i\in A(\tilde{x}_{k-1})\cap B(\tilde{x}_k)$, 与 (ii) 的证明类似，可得 (2.4) 式成立.

(iv) 若 $i\in B(\tilde{x}_k)\cap B(\tilde{x}_{k-1})$, 则

$$\begin{aligned}
&\|(\nabla f(\tilde{x}_k)-\nabla f(\tilde{x}_{k-1}))_i\|\\
&\le\|(1-\frac{\tilde{x}_k^i}{\mu})e_i-(1-\frac{\tilde{x}_{k-1}^i}{\mu})e_i\|+\|(1-\frac{F_i(\tilde{x}_k)}{\mu})\nabla F_i(\tilde{x}_k)-(1-\frac{F_i(\tilde{x}_{k-1})}{\mu})\nabla F_i(\tilde{x}_{k-1})\|\\
&\le\frac{1}{\mu}\|(\tilde{x}_k^i-\tilde{x}_{k-1}^i)e_i\|+\|\nabla F_i(\tilde{x}_k)-\nabla F_i(\tilde{x}_{k-1})\|+\frac{1}{\mu}\|F_i(\tilde{x}_k)\nabla F_i(\tilde{x}_k)-F_i(\tilde{x}_{k-1})\nabla F_i(\tilde{x}_{k-1})\|
\end{aligned}$$

根据上式，与 (ii) 类似，可推得 (2.4) 式成立. □

引理 4[11] 设序列 $\{x_k\}$ 由算法 1 产生， B_k 由 (1.10) 修正， K 为有限集，且当 k 充分大

时， f_k 关于 $\{x_k\}$ 是序列 Lipschitz 连续可微的，如果

$$\sum_{k=0}^{\infty}\|x_{k+1}-x_k\|^2<\infty, \tag{2.5}$$

那么

$$\lim_{k\to\infty}\frac{1}{k}\sum_{i=0}^{k-1}\frac{\|y_i-B_is_i\|^2}{\|s_i\|^2}=0. \tag{2.6}$$

定理 1 设水平集 Ω 有界， F 于 Ω 的闭包上 Lipschitz 连续可微，序列 $\{x_k\}$ 由算法 1 产生， ∇f_k 在 $\{x_k\}$ 的任何聚点处非奇异，则

$$\lim_{k\to\infty}\inf\|H(x_k)\|=0. \tag{2.7}$$

证明 如果指标集 K 是是无限的，则由 (2.2) 即可得 (2.7) 成立．现设 K 为有限集，当 k 充分大时，令 $\mu_k\equiv\mu$, $f_k\equiv f$, $g_k\equiv g$, $A_k(z)=A(z)$, $B_k(z)=B(z)$．由引理 1, 对充分大的 k, 有

$$\begin{aligned}\|f(x_{k+1})\|&\le(1+2^{-k})\|f(x_k)\|\le(1+2^{-k})(1+2^{-k+1})\|f(x_{k-1})\|\\&\le\cdots\le\prod_{j=0}^{k}(1+2^{-j})\|f(x_0)\|,\end{aligned}$$

由于

$$\prod_{j=0}^{k}(1+2^{-j})\le\Big(\frac{1}{k+1}\sum_{j=0}^{k}(1+2^{-j})\Big)^{k+1}=\Big(1+\frac{1}{k+1}\sum_{j=0}^{k}2^{-j}\Big)^{k+1}\le(1+\frac{2}{k+1})^{k+1}\le e^2,$$

故有 $\|f(x_{k+1})\|\le e^2\|f(x_0)\|$．根据 [7] 中引理 3.3 可知 $\{f(x_k)\}$ 是收敛的．于是由线搜索 (1.9) 可得

$$\begin{aligned}0&\leftarrow(1+2^{-k})\|f(x_k)\|^2-\|f(x_{k+1})\|^2\\&\ge2\varepsilon\max\{\|f(x_k)-f(x_{k+1})\|^2,\min\{\frac{3}{4}\|f(x_k)\|^2,\|x_{k+1}-x_k\|^2\}\}\\&\ge2\varepsilon\min\{\frac{3}{4}\|f(x_k)\|^2,\|x_{k+1}-x_k\|^2\}\end{aligned} \tag{2.8}$$

如果存在可数个 $k=n_1<n_2<\cdots$ 使得 $\frac{3}{4}\|f(x_{n_i})\|^2<\|x_{n_i+1}-x_{n_i}\|^2\}$ 成立，则由 (2.8), $\{f(x_{n_i})\}\to0$. 又由算法 1 步 5 有

$$\|g(x_k)\|\le\kappa\mu_k\le\alpha\|H(x_k)\|.$$

故由 $0<\alpha<1$ 及 $\|H(x_{n_i})\|\le\|f(x_{n_i})\|+\|g(x_{n_i})\|\le\|f(x_{n_i})\|+\alpha\|H(x_{n_i})\|$ 可得 (2.7)．如果当 k 充分大时

$$\frac{3}{4}\|f(x_k)\|^2<\|x_{k+1}-x_k\|^2,$$

于是 (2.8) 对 k 求和可得 (2.5), 特别有 $\|x_{k+1}-x_k\| \to 0(k\to\infty)$. 利用引理 3 及引理 4 可得,

$$\lim_{k\to\infty}\frac{1}{k}\sum_{j=0}^{k-1}\frac{\|y_j-B_js_j\|^2}{\|s_j\|^2}=0.$$

特别存在子列 $k\in K$, 使得 $\{\frac{\|y_k-B_ks_k\|^2}{\|s_k\|^2}\}_{k\in K}\to 0$. 记 $G_{k+1}=\int_0^1\nabla f(x_k+ts_k)\mathrm{d}t$, 得

$$g_k=f(x_{k+1})-f(x_k)=G_{k+1}s_k.$$

于是由 (1.8) 得

$$\begin{aligned}0=H(x_k)+B_kd_k&=H(x_k)+G_{k+1}d_k+\frac{(B_k-G_{k+1})d_k}{\|d_k\|}\|d_k\|\\&=H(x_k)+G_{k+1}d_k+\frac{(B_k-G_{k+1})s_k}{\|s_k\|}\|d_k\|\\&=H(x_k)+G_{k+1}d_k+\frac{B_ks_k-y_k}{\|s_k\|}\|d_k\|,\end{aligned}\tag{2.9}$$

从而

$$\|G_{k+1}d_k\|=\|H(x_k)+\frac{B_ks_k-y_k}{\|s_k\|}\|d_k\|\|\le\|H(x_k)\|+\frac{\|B_ks_k-y_k\|}{\|s_k\|}\|d_k\|.\tag{2.10}$$

令 x^* 为 $\{x_k\}_{k\in K_1}\subset\{x_k\}_{k\in K}$ 的极限点，由 $\nabla f_k(x^*)$ 的非奇异， $\nabla f(x^*)$ 也是非奇异的. 且由 $\|x_{k+1}-x_k\|\to 0$, 当 $k\in K_1$ 充分大时， G_{k+1} 一致非奇异. 由 (2.10) 即知 $\{d_k\}_{k\in K}$ 有界. 由于 $\{\frac{\|B_ks_k-y_k\|}{\|s_k\|}\}_{k\in K_1}\to 0$, 不失一般性，设 $\{d_k\}_{k\in K_1}\to d^*$, 在 (2.9) 中令 $k(\in K_1)\to\infty$, 有

$$H(x^*)+\nabla f(x^*)d^*=0.\tag{2.11}$$

记 $\lambda^*=\inf_{k\in K_1}\lambda_k$, 则 $\lambda^*>0$,

$$\lambda^*d^*=\lim_{k(\in K_1)\to\infty}(x_{k+1}-x_k)=0.$$

若 $\lambda^*>0$, 则 $d^*=0$, 由 (2.11) 即得 $H(x^*)=0$. 如果 ‘$\lambda^*=0$, 不妨设 $\lambda_k\to 0,\ k(\in K_1)\to\infty$. 当 $k\in K_1$ 充分大时，由线搜索条件， $\lambda_k'=\rho^{i_k-1}=\lambda_k/\rho$ 不满足 (1.9), 即

$$\begin{aligned}&f(x_k)^T(f(x_k)-f(x_k+\lambda_k'd_k))\\&\quad\le f(x_k)^T(f(x_k)-f(x_k+\lambda_k'd_k))\le +2^{-k-1}\|f(x_k)\|^2\\&\quad\le(\frac{1}{2}+\varepsilon)\max\{\|f(x_k)-f(x_k+\lambda_k'd_k))\|^2,\|\lambda_k'd_k\|^2\}.\end{aligned}$$

上式两边同除以 λ_k' 并令 $k(\in K_1)\to\infty$ 得

$$-f(^*)^T\nabla f(x^*)d^*\le 0.\tag{2.12}$$

结合 (2.11) 和 (2.12) 得 $H(x^*)^Tf(x^*)\le 0$. 从而

$$\begin{aligned}\|H(x^*)\|^2=H(x^*)^T(f(x^*)+g(x^*))&\le g(x^*)^TH(x^*)\\&\le\|g(x^*)\|\|H(x^*)\|\le\alpha\|H(x^*)\|.\end{aligned}$$

由于 $0<\alpha<1$, 故 $H(x^*)=0$. □

3 超线性收敛性

本节，我们证明算法 1 在一定条件下的超线性收敛性．为此，先给出如下假设：

条件 A (1) $\{x_k\}\to x^*$ 且！$H(x^*)=0$; (2) 指标集 $\{i:x_i^*=F(x^*)=0\}=\emptyset$; (3) $\nabla f_k(x^*)$ 非奇异； (4) $\varepsilon\in(0,\frac{1}{6})$.

条件 A(2) 意味着 $H(x)$ 在 x^* 的一个邻域 $U_1(x^*)$ 内是强 F 可微的．易证存在 x^* 的邻域 $U_2(x^*)$, 使得当 $x\in U_2(x^*)$ 且 k 充分大时， $f_k\equiv f\equiv H$, $B_k(z)=B(z)=\emptyset$. 记 $U=U_1(x^*)\cap U_2(x^*)$, 则当 k 充分大时， $x_k\in U$, 且 $\nabla f_k\equiv\nabla f\equiv\nabla H$ 存在并且关于 $\{x_k\}$ 是序列 Lipschitz 连续的．不难证明下面的引理 [11]:

引理 5 设条件 A 及定理 1 条件成立，则有

$$\sum_{k=0}^{\infty}\|x_{k+1}-x_k\|^2\leq\infty. \tag{3.1}$$

引理 6 设条件 A 及定理 1 条件成立，记

$$G_{k+1}=\int_0^1\nabla H(x_k+ts_k)\mathrm{d}t,\quad \delta_k=\frac{\|y_k-B_ks_k\|}{\|s_k\|}.$$

则存在常数 $\delta>0,a>0,b>1+\sigma$, 使得当 k 充分大时，成立

$$\|H(x_{k+1})\|\leq(a\|H(x_k)\|+b\cdot\frac{\delta_k}{\delta})\|H(x_k)\|. \tag{3.2}$$

而且当 $\delta_k<\delta$, k 充分大时， $\lambda_k\equiv 1$.

证明 对于 $k\geq\tilde{k}$, 有

$$\begin{aligned}
&\|H(x_k+d_k)\|\\
&\quad=H(x_k+d_k)-H(x_k)-\nabla H(x_k)d_k-B_kd_k+G_{k+1}d_k+\nabla H(x_k)d_k-G_{k+1}d_k\\
&\quad\leq\|H(x_k+d_k)-H(x_k)-\nabla H(x_k)d_k\|+\|(\nabla H(x_k)-G_{k+1})d_k\|+\|(B_k-G_{k+1})d_k\|\\
&\quad\leq[\sup_{0\leq t\leq 1}\|\nabla H(x_k+td_k)-\nabla H(x_k)\|+\|\nabla H(x_k)-G_{k+1}\|+\delta_k]\|d_k\|\\
&\quad\leq(2L\|d_k\|+\delta_k)\|d_k\|,
\end{aligned} \tag{3.3}$$

其中 $L>0$ 是 ∇H 的序列 Lipschitz 常数．由 (2.3), 存在 $\bar{\delta}$ 使得当 $\delta_k<\delta$ 且 k 充分大时, $\|d_k\|/\|H(x_k)\|$ 有界．于是由 (3.3) 及 $H(x_k)\to 0$, 存在常数 δ', 使得当 $\delta_k<\min\{\bar{\delta},\delta'\}$ 时，有

$$H(x_k+d_k)\|\leq\frac{1}{2}\|H(x_k)\|. \tag{3.4}$$

又因

$$q_k(1)-\frac{2}{3}=\frac{(1+2^{-k-1})\|f_k(x_k)\|^2-f_k(x_k)^Tf_k(x_k+d_k)}{\max\{\|f_k(x_k)-f_k(x_k+d_k)\|^2,\min\{\frac{3}{4}\|f_k(x_k)\|^2,\|d_k\|^2\}\}}-\frac{2}{3},$$

则当 k 充分大时，注意到此时 $f_k \equiv f \equiv H$, 有

$$\begin{aligned}
&(1+2^{-k-1})\|H(x_k)\|^2 - H(x_k)^T H(x_k+d_k) - \frac{2}{3}\|H(x_k)-H(x_k+d_k)\|^2 \\
&= (\frac{1}{3}+2^{-k-1})\|H(x_k)\|^2 - \frac{1}{3}H(x_k)^T H(x_k+d_k) - \frac{2}{3}\|H(x_k+d_k)\|^2 \\
&\geq \frac{1}{3}\|H(x_k)\|^2 - \frac{1}{3}\|H(x_k)\|\|H(x_k+d_k)\| - \frac{2}{3}\|H(x_k+d_k)\|^2 \\
&\geq \frac{1}{3}\|H(x_k)\|^2 - \frac{1}{3}\cdot\frac{1}{2}\|H(x_k)\|^2 - \frac{2}{3}\cdot\frac{1}{4}\|H(x_k)\|^2 = 0,
\end{aligned}$$

及

$$\begin{aligned}
&(1+2^{-k-1})\|H(x_k)\|^2 - H(x_k)^T H(x_k+d_k) - \frac{2}{3}\cdot\frac{3}{4}\|H(x_k)\|^2 \\
&\geq \frac{1}{2}\|H(x_k)\|^2 - \|H(x_k)\|\|H(x_k+d_k)\| \geq 0,
\end{aligned}$$

故 $q_k(1) \geq \frac{3}{2} > \frac{1}{2}+\varepsilon$, 从而 $\lambda_k \equiv 1$.

同样，利用 $\|d_k\|/\|H(x_k)\|$ 的有界性及 (3.3), 当 $\delta_k \leq \delta = \min\{1, \delta', \bar{\delta}\}$ 且 k 充分大时，存在 $a>0$ 和 $b>1+\sigma$ 使得

$$\|H(x_{k+1})\| \leq (a\|H(x_k)+b\delta_k)\|H(x_k)\| \leq (a\|H(x_k)\|+b\frac{\delta_k}{\delta})\|H(x_k)\|. \tag{3.5}$$

当 $\delta_k \geq \delta$ 且 k 充分大时，不难证明上面的不等式也成立. □

类似于 [11] 的定理 3.1, 我们有

引理 7　设 $\{x_k\}$ 由算法 1 产生， $f: R^n \to R^n$ 连续可微，其 Jacobian 矩阵 ∇f 关于 $\{x_k\}$ 是序列 lipschitz 连续的，则当 $\sum_{k=0}^{\infty}\|x_{k+1}-x_k\|^2 \leq \infty$ 时，

$$\sum_{k=0}^{\infty}\delta_k^2 = \sum_{k=0}^{\infty}\frac{\|y_k - B_k s_k\|}{\|s_k\|} \leq \infty. \tag{3.6}$$

特别地，有 $\delta_k^2 = \frac{\|y_k-B_k s_k\|^2}{\|s_k\|} \to 0\ (k\to\infty)$.

现在我们可以证明算法 1 的局部超线性收敛性.

定理 2　设条件 A 及定理 1 条件成立，则

$$\sum_{k=0}^{\infty}\|x_{k+1}-x_k\|^2 \leq \infty. \tag{3.7}$$

进一步地，序列 $\{x_k\}$ 超线性收敛.

证明　由引理 6，存在自然数 $\bar{k}$, 当 $k \geq \bar{k}$ 时,

$$\|H(x_{k+1})\|^2 \leq (a\|H(x_k)\|+b\frac{\delta_k}{\delta})^2\|H(x_k)\|^2 \leq 2(a^2\|H(x_k)\|^2+b^2\frac{\delta_k^2}{\delta^2})\|H(x_k)\|^2.$$

上式从 $\bar{k}$ 到 $k>\bar{k}$ 相乘得

$$\begin{aligned}
\|H(x_{k+1})\|^2 &\leq \|H(x_{\bar{k}})\|^2 \prod_{i=\bar{k}}^{k} 2(a^2\|H(x_i)\|^2 + b^2\frac{\delta_i^2}{\delta^2}) \\
&\leq \|H(x_{\bar{k}})\|^2 \Big(\frac{2a^2}{k-\bar{k}+1}\sum_{i=\bar{k}}^{k}\|H(x_i)\|^2 + \frac{2b^2}{\delta^2}\frac{1}{k-\bar{k}+1}\sum_{i=\bar{k}}^{k}\delta_i^2\Big)^{k-\bar{k}+1}.
\end{aligned}$$

上式结合 $\frac{1}{k+1}\sum_{i=0}^{k}\frac{\|y_i-B_s s_i\|^2}{s_i\|^2}\to 0$, 即 $\frac{1}{k+1}\sum_{i=0}^{k}\delta_i^2\to 0$, 可得

$$\|H(x_{k+1})\|^{\frac{1}{k+1}}\to 0. \tag{3.8}$$

由于当 k 充分大时, $\nabla H(x^*)$ 非奇异且 ∇H 连续, $\|x_{k+1}-x_k\|^{\frac{1}{k+1}}\to 0$, 故有

$$\sum_{k=0}^{\infty}\|x_{k+1}-x_k\|^2\leq\infty.$$

由引理 7 有 $\delta_k\to 0$. 因此, 当 k 充分大时, $\lambda_k\equiv 1$. 根据 [10] 的推论 2.3 即可证得 $\{x_k\}$ 的超线性收敛性. □

参考文献:

[1] COTTLE R W, GIANNEESSI F, LIONS T L. *Variational Inequalities and Complementarity Problems, Theory and Applications* [M]. Wiley, New York, 1980.
[2] HARKER P T, PANG J S. *Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications* [J]. Math. Programming, Ser.B, 1990, **48**(2): 161–220.
[3] ISAC G. *Complementarity Problems* [M]. Springer-Verlay, Berlin Heidelberg, 1992.
[4] PANG J S, CHEN D. *Iterative methods for variational and complementarity problems* [J]. Math. Programming, 1982, **24**(3): 284–313.
[5] PANG J S, GABRIEL S A, *NE/SQP: a robust algorithm for the nonlinear complementarity problem* [J]. Math. Programming, Ser.A, 1993, **60**(3): 295–337.
[6] JIANG Hou-yuan, QI Li-qun. *A new nonsmooth equations approach to nonlinear complementarity problems* [J]. SIAM J. Control Optim., 1997, **35**(1): 178–193.
[7] CHEN X, QI L, SUN D. *Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities* [J]. Math. Comp., 1998, **67**(222): 519–540.
[8] Qi Li-qun, CHEN Xiao-jun. *A globally convergent successive approximation method for severely nonsmooth equations* [J]. SIAM J. Control Optim., 1995, **33**(2): 402–418.
[9] FISCHER A. *A special Newton-type optimization method* [J]. Optimization, 1992, **24**(3-4): 269–284.
[10] DENNIES J E, MORÉ J J. *A characterization of superliner convergence and its application to quasi-Newton methods* [J]. Math. Computation, 1974, **28**: 549–560.
[11] 李董辉. 求解无约束最优化问题的非奇异 Broyden 算法的全局收敛性 [J]. 计算数学, 1995, **17**: 321–330.
LI Dong-hui. *Global convergence of nonsingular Broyden's method for solving unconstrained optimizations* [J]. Math. Numer. Sinica, 1995, **17**: 321–330. (in Chinese)

Convergence Analysis of a Successive Approximation Quasi-Newton Method for Solving Nonlinear Complementarity Problems

MA Chang-feng[1,2]
(1. Dept. of Comp. Sci. & Math., Guilin Inst. of Electro. Tech., Guangxi 541004, China;
2. College of Math. & Physics, Zhejiang Normal University, Jinhua 321004, China;)

Abstract: A successive approximation quasi-Newton method for solving nonlinear complementarity problems is proposed. By presenting a suitable line search, the algorithm is well defined. And under certain conditions, the global convergence and locally superlinear convergence of the method are established.

Key words: nonlinear complementarity problem; successive approximation; quasi-Newton method; convergence analysis.