

解 P_0 非线性互补问题的光滑 Levenberg-Marquardt 方法

陈金雄¹, 刘 宁²

(1. 福建武夷学院数学与计算机系, 福建 武夷山 354300)

(2. 广西玉林市第十二中学数学教研组, 广西 玉林 537000)

摘 要: 本文研究了一个 P_0 非线性互补问题. 利用信赖域技术获得了求解该问题的光滑 Levenberg-Marquardt 算法, 该算法在一定条件下具有全局性. 利用局部误差界还获得了该算法的超线性和二次收敛. 数值结果表明该算法是有效的.

关键词: 非线性互补问题; Levenberg-Marquardt 方法; 全局性; 局部收敛性

MR(2010) 主题分类: 90C30

中图分类号: O221.2

文献标识码: A

文章编号: 0255-7797(2015)04-0905-12

1 引言

考虑下面的非线性互补问题: 求向量 $x \in R^n$, 使得

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

其中连续可微函数 $F: R^n \rightarrow R^n$.

非线性互补问题在许多实际部门有着广泛的应用, 如力学、工程、经济、交通等. 正因为非线性互补问题的重要性, 因此当前已有不少关于求解互补问题的数值算法, 如信赖域法, 内点与非内点法, 连续化方法, 投影收缩法, 优化方法等. 而 Levenberg-Marquardt 是求解互补问题的重要算法之一, 在实际中有着重要应用. 本文将采用 Levenberg-Marquardt 方法求解 P_0 非线性互补问题. 首先利用 FB 函数 $\phi(a, b) = a + b - \sqrt{a^2 + b^2}$ 将非线性互补问题 (1.1) 转化为下面的方程组:

$$\Phi(x) = \begin{pmatrix} \phi(x_1, F_1(x)) \\ \vdots \\ \phi(x_n, F_n(x)) \end{pmatrix}. \quad (1.2)$$

引入参数 τ 对 FB 函数进行光滑逼近:

$$\phi_\tau(a, b) = a + b - \sqrt{a^2 + b^2 + 2\tau^2}. \quad (1.3)$$

从而有如下方程组:

$$\Phi_\tau(x) = \begin{pmatrix} \phi_\tau(x_1, F_1(x)) \\ \vdots \\ \phi_\tau(x_n, F_n(x)) \end{pmatrix}. \quad (1.4)$$

*收稿日期: 2012-10-09

接收日期: 2012-11-18

基金项目: 福建省教育厅科技 A 类项目 (JA13315); 福建省武夷学院青年教师专项科研基金项目 (XQ1201).

作者简介: 陈金雄 (1983-), 男, 广西玉林, 讲师, 主要研究方向: 数值代数和最优化理论与算法研究.

定义价值函数:

$$\theta(x) = \|\Phi(x)\|^2, \quad \theta_\tau(x) = \|\Phi_\tau(x)\|^2. \quad (1.5)$$

目前有大量关于 Levenberg-Marquardt 方法的文章, 如文献 [2] 通过引入两个 NCP 函数 (FB 函数和 minimum 函数) 将广义线性互补问题转化为非线性方程组 Φ , 同时引入一个参数 τ 得到 Φ 的一个光滑逼近 Φ_τ , 提出了一个求解该方程组的光滑算法. 算法具有全局性和在严格互补条件下获得该算法的超线性. 文献 [3] 基于 FB 函数, 引入一个参数将非线性互补问题转化为光滑方程组, 在这个方程组中参数被视为一个独立的变量, 提出了一个解 P_0 非线性互补问题的光滑 Levenberg-Marquardt 方法, 该算法是全局收敛的, 在严格互补条件或者非奇异条件满足情况下利用局部误差界都可得到算法的超线性; 文献 [4] 针对非线性方程组将文献 [5] 中关于信赖域半径选择的技巧与文献 [6] 中关于 LM 方法的参数更新技巧结合起来获得一个结合信赖域技术的修正 Levenberg-Marquardt 方法, 所提算法有全局性, 同时在局部误差界条件下获得该算法的局部收敛性. 本文基文献 [1] 的算法思想提出一个解 P_0 非线性互补问题的光滑 Levenberg-Marquardt 方法, 采用新的 LM 参数 $\mu_k = \alpha_k \|\nabla \Phi_{\tau_k}^T \Phi_{\tau_k}\|^\delta$, 而 α_k 的更新则采用文献 [6] 所给信赖域技术. 所提算法具有全局性, 不需要严格互补条件的假设, 而是采用较弱的局部误差界条件来获得该算法的局部收敛性.

2 算法

本节提出一个解 P_0 -NCP 的光滑 Levenberg-Marquardt 方法. 首先设

$$Q_k(d_k) = \frac{1}{2} \|\Phi_{\tau_k}(x_k) + \nabla \Phi_{\tau_k}^T d_k\|^2, \quad \text{pred} = Q_k(0) - Q_k(d_k),$$

$$\text{ared} = \frac{1}{2} \|\Phi_{\tau_k}(x_k)\|^2 - \frac{1}{2} \|\Phi_{\tau_k}(x_{k+1})\|^2.$$

算法 1 (光滑 Levenberg-Marquardt 方法)

步 0 给定 $x_0 \in R^n$, $b_0 \in (0, \frac{1}{2})$, $a > 0$, $b_0 < b_1 \leq 1$, $\eta < 1$, $\mu < 1$, $\alpha_0 > m > 0$, $\delta \in (0, 2)$, $\delta_2 < 1 \leq \delta_1$, $\beta_0 = \|\Phi(x_0)\|$, $\tau_0 = \frac{\mu}{2\sqrt{2n}}\beta_0$, $k := 0$.

步 1 求解下面的方程组

$$(\nabla \Phi_{\tau_k}^T \nabla \Phi_{\tau_k} + \alpha_k \|\nabla \Phi_{\tau_k}^T \Phi_{\tau_k}\|^\delta I) d = -\nabla \Phi_{\tau_k}^T \Phi_{\tau_k} \quad (2.1)$$

获得 d_k 满足

$$\text{pred} \geq a \|\Phi_{\tau_k}\| \|d_k\|.$$

步 2 计算 $r_k = \frac{\text{ared}}{\text{pred}}$. 若 $r_k \geq b_0$, 那么 $x_{k+1} = x_k + d_k$, 否则 $x_{k+1} = x_k$. 更新 α_{k+1} ,

$$\alpha_{k+1} = \begin{cases} \delta_1 \alpha_k, & \text{若 } rk < b_0, \\ \alpha_k, & \text{若 } rk \in [b_0, b_1], \\ \max\{\delta_2 \alpha_k, m\}, & \text{若 } rk > b_1. \end{cases}$$

步 3 若 $\|\nabla \theta(x_{k+1})\| = 0$, 停算.

步 4 若

$$\|\Phi(x_{k+1})\| \leq \max\{\eta \beta_k, \mu^{-1} \|\Phi(x_{k+1}) - \Phi_{\tau_k}(x_{k+1})\|\}.$$

令 $\beta_{k+1} = \|\Phi(x^{k+1})\|$, 并选取 τ_{k+1} 满足

$$0 < \tau_{k+1} < \min \left\{ \frac{\mu}{4\sqrt{n}} \beta_{k+1}, \frac{\tau_k}{2} \right\}.$$

否则置 $\beta_{k+1} := \beta_k$, $\tau_{k+1} := \tau_k$, 转步 5.

步 5 令 $k := k + 1$, 转步 1.

注 1 定义集合 S :

$$S = \{0\} \cup \{k \mid \|\Phi(x_k)\| \leq \max\{\eta\beta_{k-1}, \mu^{-1}\|\Phi(x_k) - \Phi_{\tau_{k-1}}(x^k)\|\}\}.$$

3 全局收敛性分析

本节将对算法 1 的全局收敛性进行分析.

定义 1 设函数 $F: R^n \rightarrow R^n$ 对 $\forall x, y \in R^n, x \neq y$, 有 $\max_{x_i \neq y_i} (x_i - y_i)(F(x_i) - F(y_i)) \geq 0$,

则称 F 为 P_0 函数.

引理 1 对 $\forall x \in R^n$, 以及 $\tau \geq 0$, 都有 $\|\Phi(x_k) - \Phi_\tau(x_k)\| \leq \sqrt{2n\tau}$.

证 由题意有

$$\begin{aligned} \|\Phi(x_k) - \Phi_\tau(x_k)\| &= \left(\sum_{i=1}^n \left[\varphi(x_i, F_i(x)) - \varphi_\tau(x_i, F_i(x)) \right]^2 \right)^{1/2} \\ &= \left(\sum_{i=1}^n \left[x_i + F_i(x) - \sqrt{x_i^2 + (F_i(x))^2} - \left(x_i + F_i(x) - \sqrt{x_i^2 + (F_i(x))^2 + 2\tau^2} \right) \right]^2 \right)^{1/2} \\ &= \left(\sum_{i=1}^n \left[\sqrt{x_i^2 + (F_i(x))^2 + 2\tau^2} - \sqrt{x_i^2 + (F_i(x))^2} \right]^2 \right)^{1/2} \\ &= \left(\sum_{i=1}^n \left[\frac{2\tau^2}{\sqrt{x_i^2 + (F_i(x))^2 + 2\tau^2} + \sqrt{x_i^2 + (F_i(x))^2}} \right]^2 \right)^{1/2} \\ &\leq \left(\sum_{i=1}^n \left(\frac{2\tau^2}{\sqrt{2\tau^2}} \right)^2 \right)^{1/2} = \sqrt{2n\tau}. \end{aligned}$$

引理 2 ^[7] 对 $\forall k \in S$, 都有 $\|\Phi(x_k) - \Phi_{\tau_k}(x_k)\| \leq \mu\|\Phi(x_k)\|$.

注 2 引理 2 意味着 $\forall k \in S$, 有 $\|\Phi_{\tau_k}(x_k)\| \neq 0$.

引理 3 ^[7] 如果 F 为 P_0 函数, 那么对 $\forall k \in S, \tau > 0$, 有 $\nabla \Phi_{\tau_k}(x_k)$ 是非奇异的.

引理 4 记水平集 $L_0 = \{x \in R^n \mid \theta(x_k) \leq (1 + \mu)\theta(x_0)\}$, 序列 $\{x_k\}$ 由算法 1 产生, 若 F 为 P_0 函数, 那么 $\{x_k\} \subseteq L_0$.

证 类似于文献 [1] 命题 4.1 的证明过程, 立即可得引理 4 的结论.

定理 1 如果 F 为 P_0 函数, 那么算法 1 是适定的.

证 由算法步 2 知, 要证明算法 1 的适定性, 只需证明 $Q_k(0) - Q_k(d_k) > 0$.

反证法: 假定对于 $j \in S$, 使得 $Q_j(0) - Q_j(d_j) = 0$, 即

$$-\nabla\theta_{\tau_j}(x_j)^T d_j - \frac{1}{2} d_j^T \nabla\Phi_{\tau_j}^T(x_j) \nabla\Phi_{\tau_j}(x_j) d_j = 0.$$

这就意味着 $d_j = 0$ 是以上方程的一个解, 从而由 (2.1) 式有 $-\nabla\Phi_{\tau_j}^T(x_j)\Phi_{\tau_j}(x_j) = 0$, 但这与由引理 2 的注 2 和由引理 3 所得的 $-\nabla\Phi_{\tau_j}^T(x_j)\Phi_{\tau_j}(x_j) \neq 0$ 相矛盾. 故命题成立.

引理 5 [8] 对 $\forall k$, 假设 $d_k \in R^n$ 是 (2.1) 式的解. 那么

$$\text{pred} \geq \frac{1}{2} \|\nabla\theta_{\tau_k}(x_k)\| \min\{\|d_k\|, \frac{\|\nabla\theta_{\tau_k}\|}{\|\nabla\Phi_{\tau_k}^T \nabla\Phi_{\tau_k}\|}\}. \quad (3.1)$$

定理 2 设序列 $\{x_k\}$ 由算法 1 产生, 如果 F 为 P_0 函数并且由算法 1 所产生的序列存在聚点, 那么集合是无限的且 $\lim_{k \rightarrow \infty} \tau_k = 0$, $\lim_{k \rightarrow \infty} \Phi_{\tau_k}(x_k) = 0$, $\lim_{k \rightarrow \infty} \Phi(x_k) = 0$.

证 首先利用反证法证明集合 S 无限的. 假定集合 S 有限的, 设集合 S 中最大的数记为 $\bar{k}$, 那么 $\forall k \geq \bar{k}$, 有 $\tau_k = \tau_{\bar{k}}$ 以及 $\beta_k = \beta_{\bar{k}}$. 记 $\bar{\tau} = \tau_{\bar{k}}$, $\bar{\beta} = \beta_{\bar{k}}$ 以及 $q(x) = \Phi(x) - \Phi_{\bar{\tau}}(x)$, 那么 $\forall k \geq \bar{k}$, 有

$$\|\Phi(x_k)\| > \max\{\eta\bar{\beta}, \mu^{-1}\|q(x_k)\|\} \quad (3.2)$$

和

$$\Phi(x_k) = q(x_k) + \Phi_{\bar{\tau}}(x_k). \quad (3.3)$$

下面用反证法证明: 序列 $\{x_k\}$ 至少存在一个聚点 $\bar{x} \subseteq L_0$ 使得

$$\nabla\theta_{\bar{\tau}}(\bar{x}) = 0. \quad (3.4)$$

若 (3.4) 式不成立, 记 $K := \{k \in S | r_k \geq b_0\}$, 那么由 $\theta(x)$ 是连续可微的, 那么存在 $\varepsilon_1, \varepsilon_2 > 0$, 有

$$\|\nabla\theta_{\bar{\tau}}(x_k)\| > \varepsilon_1, \|\nabla\Phi_{\bar{\tau}}^T \nabla\Phi_{\bar{\tau}}\| < \varepsilon_2, \forall k \in K. \quad (3.5)$$

由 $r_k \geq b_0$, $\forall k \in K$, 有

$$\begin{aligned} \theta_{\bar{\tau}}(x_k) - \theta_{\bar{\tau}}(x_{k+1}) &\geq b_0 \text{pred} \\ &\geq \frac{1}{2} b_0 \|\nabla\theta_{\bar{\tau}}(x_k)\| \min\left\{\|d_k\|, \frac{\|\nabla\theta_{\bar{\tau}}(x_k)\|}{\|\nabla\Phi_{\bar{\tau}}^T(x_k)\Phi_{\bar{\tau}}(x_k)\|}\right\} \\ &\geq \frac{1}{2} b_0 \varepsilon_1 \min\left\{\|d_k\|, \frac{\varepsilon_1}{\varepsilon_2}\right\}. \end{aligned} \quad (3.6)$$

由于 $\{\theta_{\bar{\tau}}(x_k)\}$ 是减的, 以及由引理 4 知其是有界的, 故 $\{\theta_{\bar{\tau}}(x_k)\}$ 是收敛的. 那么由 (3.6) 式有

$$\begin{aligned} \frac{1}{2} b_0 \varepsilon_1 \min\left\{\|d_k\|, \frac{\varepsilon_1}{\varepsilon_2}\right\} &\leq \sum_{k \in K} [\theta_{\bar{\tau}}(x_k) - \theta_{\bar{\tau}}(x_{k+1})] \\ &\leq \sum_{k=1}^{\infty} [\theta_{\bar{\tau}}(x_k) - \theta_{\bar{\tau}}(x_{k+1})] \\ &< \infty. \end{aligned} \quad (3.7)$$

那么由引理 2, 引理 3 和 (3.7) 式知 $\{d\}_{k \in K} \rightarrow 0$, 由于 (2.1) 式有 $d = -\frac{\nabla \Phi_{\bar{\tau}}^T \Phi_{\bar{\tau}}}{\nabla \Phi_{\bar{\tau}}^T \nabla \Phi_{\bar{\tau}} + \alpha_k \|\nabla \Phi_{\bar{\tau}}^T \Phi_{\bar{\tau}}\|^\delta I}$, 从而

$$\{\alpha_k\}_{k \in K} \rightarrow \infty \quad (3.8)$$

另一方面, 由于 $\theta_{\bar{\tau}}(x_k)$ 是连续可微的, 那么

$$\theta_{\bar{\tau}}(x_k + d_k) = \theta_{\bar{\tau}}(x_k) + \nabla \theta_{\bar{\tau}}(\omega_k)^T d_k, (\omega_k = x_k + \lambda_k d_k, \lambda_k \in (0, 1)). \quad (3.9)$$

由引理 5, (3.5), (3.6), (3.9) 式以及 Cauchy-Schwarz 不等式有

$$\begin{aligned} |r_k - 1| &= \left| \frac{\theta_{\bar{\tau}}(x_k) - \theta_{\bar{\tau}}(x_{k+1}) + Q_k(x_k) - Q_k(0)}{Q_k(0) - Q_k(x_k)} \right| \\ &= \left| \frac{\theta_{\bar{\tau}}(x_k) - \theta_{\bar{\tau}}(x_{k+1}) + \nabla \theta_{\bar{\tau}}(x_k)^T d_k + \frac{1}{2} d_k^T \nabla \Phi_{\bar{\tau}}^T(x_k) \Phi_{\bar{\tau}}(x_k) d_k}{Q_k(0) - Q_k(x_k)} \right| \\ &\leq \left| \frac{\theta_{\bar{\tau}}(x_k) - \theta_{\bar{\tau}}(x_k) - \nabla \theta_{\bar{\tau}}(\omega_k)^T d_k + \nabla \theta_{\bar{\tau}}(x_k)^T d_k + \frac{1}{2} d_k^T \nabla \Phi_{\bar{\tau}}^T(x_k) \Phi_{\bar{\tau}}(x_k) d_k}{\frac{1}{2} \varepsilon_1 \min\{\|d_k\|, \frac{\varepsilon_1}{\varepsilon_2}\}} \right| \\ &\leq \frac{2}{\varepsilon_1 \|d_k\|} \left(\left| \nabla \theta_{\bar{\tau}}(x_k)^T d_k - \nabla \theta_{\bar{\tau}}(\omega_k)^T d_k + \frac{1}{2} d_k^T \nabla \Phi_{\bar{\tau}}^T(x_k) \Phi_{\bar{\tau}}(x_k) d_k \right| \right) \\ &\leq \frac{2}{\varepsilon_1 \|d_k\|} \left(\left| \nabla \theta_{\bar{\tau}}(x_k)^T d_k - \nabla \theta_{\bar{\tau}}(\omega_k)^T d_k \right| + \left| \frac{1}{2} d_k^T \nabla \Phi_{\bar{\tau}}^T(x_k) \Phi_{\bar{\tau}}(x_k) d_k \right| \right) \\ &\leq \frac{2}{\varepsilon_1 \|d_k\|} \left(\|\nabla \theta_{\bar{\tau}}(x_k) - \nabla \theta_{\bar{\tau}}(\omega_k)\| \|d_k\| + \frac{1}{2} \|d_k\|^2 \varepsilon_2 \right) \\ &\leq \frac{2}{\varepsilon_1} \left(\|\nabla \theta_{\bar{\tau}}(x_k) - \nabla \theta_{\bar{\tau}}(\omega_k)\| + \frac{1}{2} \|d_k\| \varepsilon_2 \right). \end{aligned} \quad (3.10)$$

由于 $\{\omega_k\}_{k \in K} \rightarrow \bar{x}$, 当 $k \in K$ 充分大式 (3.10) 式是趋向于零的. 即 $r_k \rightarrow 1 (k \in K)$. 因此由算法 1 步 2 中 α_k 的取值知存在常数 $N > m$, 使对充分大 $k \in K$ 有

$$\alpha_k < N. \quad (3.11)$$

而 (3.8) 与 (3.11) 式矛盾. 从而 (3.4) 式是成立的.

因此由引理 1 和 (3.4) 式有 $\{\theta_{\bar{\tau}}(x_k)\} \rightarrow \theta_{\bar{\tau}}(\bar{x}) = 0$, 那么存在 $\hat{k} \geq \bar{k}$, 有 $k \geq \hat{k}, k \in K$,

$$\|\Phi_{\bar{\tau}}\| \leq (1 - \mu)\eta\bar{\beta}.$$

对于 $k \geq \hat{k}, k \in K$, 由 (3.2), (3.3) 式有

$$\|\Phi_{\bar{\tau}}(x_k)\| \leq (1 - \mu)\|\Phi(x_k)\| \leq (1 - \mu)(\|q(x_k)\| + \|\Phi_{\bar{\tau}}(x_k)\|), \quad (3.12)$$

由 (3.12) 式得到

$$\|\Phi_{\bar{\tau}}(x_k)\| \leq (\mu^{-1} - 1)\|q(x_k)\|, \quad (3.13)$$

而由 (3.3), (3.13) 式有

$$\begin{aligned}\|\Phi(x_k)\| &= \|q(x_k) + \Phi_{\bar{\tau}}(x_k)\| \\ &\leq \|q(x_k)\| + \|\Phi_{\bar{\tau}}(x_k)\| \\ &\leq \|q(x_k)\| + (\mu^{-1} - 1)\|q(x_k)\| \\ &\leq \mu^{-1}\|q(x_k)\|.\end{aligned}\quad (3.14)$$

故 (3.2) 与 (3.14) 式矛盾, 所以集合 S 是无限的.

由算法 1 的步 4 关于 τ_k 的更新准则知 $\{\tau_k\} \rightarrow 0, k \in \infty$. 由引理 4 可推得

$$\|\Phi(x_k)\| \leq p_j(1 + \mu)\|\Phi(x_0)\|, k_j \leq k \leq k_{j+1}, \quad (3.15)$$

其中 $p_j = \max\{\frac{1}{2}, \eta\}$.

由于集合 S 是无限的, 那么由引理 2 和 (3.15) 式有 $\lim_{k \rightarrow \infty} \Phi(x_k) = 0, \lim_{k \rightarrow \infty} \Phi_{\tau_k}(x_k) = 0$.

定理 3^[1] 如果 F 为 P_0 函数, 序列 $\{x_k\}$ 由算法 1 生成的, 那么其全部聚点都为 $NCP(F)$ 的解.

4 局部收敛性分析

求解 $(\nabla \Phi_{\tau_k}^T \nabla \Phi_{\tau_k} + \alpha \|\nabla \Phi_{\tau_k}^T \nabla \Phi_{\tau_k}\|^\delta I)d = -\nabla \Phi_{\tau_k}^T \nabla \Phi_{\tau_k}$ 等价于求解以下最小问题的最优解

$$\min_{d \in R^n} \Psi(d), \quad (4.1)$$

其中

$$\Psi(d) = \frac{1}{2} \|\Phi_{\tau_k}(x_k) + \nabla \Phi_{\tau_k}^T d\|^2 + \frac{1}{2} \alpha_k \|\nabla \Phi_{\tau_k}^T \Phi_{\tau_k}\|^\delta \|d\|^2. \quad (4.2)$$

定义 2^[4] 设 M 是 R^n 上的子集, 并且满足 $M \cap X \neq \emptyset$, 其中 X 是 (1.4) 式的非空解集, 若存在常数 $c > 0$, 使得 $\|\Phi_\tau(x)\| \geq c \text{dist}(x, X), \forall x \in M$, 那么 $\Phi_\tau(x)$ 在 M 内局部误差有界. 首先给出一些有用的假设.

假设 A

- (1) $\nabla \Phi_\tau(x_*)$ 是非奇异的, 如果 x_* 是由算法 1 产生的序列 $\{x_k\}$ 的聚点.
- (2) 问题 (1.4) 的任意解 $x_* \in X$, 存在 $b > 0, c > 0$, 有

$$\|\Phi_\tau(x)\| \geq c \text{dist}(x, X), \forall x \in M(x_*, b) \{x \mid \|x - x_*\| \leq b\}, \quad (4.3)$$

则 $\|\Phi_\tau(x)\|$ 在 $M(x_*, b)$ 内有局部误差界. 下面将给出算法 1 的局部收敛性分析, 首先以 $\bar{x}_k$ 记为 x_k 到 X 距离最短的点, 也就是

$$\|x_k - \bar{x}_k\| = \text{dist}(x_k, X), \bar{x}_k \in X. \quad (4.4)$$

引理 6^[3] $\Phi_\tau(x)$ 是局部 Lipschitz 连续, 且是强半光滑的, 那么存在常数 $L_1 > 0, L_2 > 0, x_1, x_2 \in R^n$

$$\|\Phi_\tau(x_1) - \Phi_\tau(x_2)\| \leq L_1 \|x_1 - x_2\|, \quad (4.5)$$

$$\|\Phi_\tau(x_1) - \Phi_\tau(x_2) - V^T(x_1 - x_2)\| \leq L_2\|x_1 - x_2\|^2, \quad (4.6)$$

其中 $V \in \partial\Phi_\tau(x)$, $\partial\Phi_\tau(x)$ 为 Clarke 通常意义下的 Jacobian 矩阵.

引理 7 若假设 A 成立, 对充分大 k , 存在常数 N , 有

$$\alpha_k < N. \quad (4.7)$$

证 由算法 1 的步 1 有 $\text{pred} \geq a\|\Phi_\tau\|\|d_k\|$, 从而由

$$\begin{aligned} |r_k - 1| &= \left| \frac{\text{ared}}{\text{pred}} - 1 \right| \\ &= \left| \frac{\frac{1}{2}\|\Phi_{\tau_k}(x_k) + \nabla\Phi_{\tau_k}^T(x_k)d_k\|^2 - \frac{1}{2}\|\Phi_{\tau_k}(x_{k+1})\|^2}{\text{pred}} \right| \\ &= \frac{\|\Phi_{\tau_k}(x_k) + \nabla\Phi_{\tau_k}^T d_k\|O(\|d_k\|^2) + O(\|d_k\|^4)}{\text{pred}} \\ &\leq \frac{\|\Phi_{\tau_k}\|O(\|d_k\|^2) + O(\|d_k\|^4)}{O(\|\Phi_{\tau_k}\|\|d_k\|)} \\ &= O(\|d_k\|) \rightarrow 0. \end{aligned}$$

即有 $r_k \rightarrow 1$. 因此对充分大 k , 由算法 1 步 2 中 α_k 的取值知存在常数 $N > m$, 有 $\alpha_k < N$.

引理 8 设假设 A 成立, 如果 $x_k \in M(x_*, b)$, 那么存在常数 $c_1 > 0$ 使得

$$c_1\|\bar{x}_k - x_k\|^\delta \leq \alpha_k\|\nabla\Phi_{\tau_k}^T \Phi_{\tau_k}\|^\delta \leq c_2\|\bar{x}_k - x_k\|^\delta.$$

证 由引理 6, 假设 A 的 (2), (4.7) 式有

$$\begin{aligned} \alpha_k\|\nabla\Phi_{\tau_k}^T(x_k)\Phi_{\tau_k}(x_k)\|^\delta &\leq \alpha_k\|\nabla\Phi_{\tau_k}(x_k)\|^\delta\|\Phi_{\tau_k}(x_k)\|^\delta \\ &\leq \alpha_k\|\nabla\Phi_{\tau_k}(x_k)\|^\delta\|\Phi_{\tau_k}(\bar{x}_k) - \Phi_{\tau_k}(x_k)\|^\delta \\ &\leq \alpha_k\|\nabla\Phi_{\tau_k}(x_k)\|^\delta\|L_1(\bar{x}_k - x_k)\|^\delta \\ &\leq c_2\|\bar{x}_k - x_k\|^\delta, \end{aligned}$$

其中 $c_2 = NL_1^{2\delta}$. 右边的不等式得证.

下面证明左边的不等式成立. 由

$$\begin{aligned} \|\Phi_{\tau_k}(x_k)\|^2 &= \Phi_{\tau_k}(x_k)^T \Phi_{\tau_k}(x_k) \\ &= \Phi_{\tau_k}(x_k)^T (\Phi_{\tau_k}(\bar{x}_k) + \nabla\Phi_{\tau_k}(x_k)(x_k - \bar{x}_k)) \\ &\quad + \Phi_{\tau_k}(x_k)^T (\Phi_{\tau_k}(x_k) - \Phi_{\tau_k}(\bar{x}_k) - \nabla\Phi_{\tau_k}(x_k)(x_k - \bar{x}_k)), \end{aligned}$$

有

$$\begin{aligned} &\Phi_{\tau_k}(x_k)^T \nabla\Phi_{\tau_k}(x_k)(x_k - \bar{x}_k) \\ &= \|\Phi_{\tau_k}(x_k)\|^2 - \Phi_{\tau_k}(x_k)^T (\Phi_{\tau_k}(x_k) - \Phi_{\tau_k}(\bar{x}_k) - \nabla\Phi_{\tau_k}(x_k)(x_k - \bar{x}_k)). \end{aligned}$$

由假设 A 和引理 6, 有

$$\|\nabla\Phi_{\tau_k}(x_k)^T \Phi_{\tau_k}(x_k)\|\|x_k - \bar{x}_k\| \geq c^2\|x_k - \bar{x}_k\|^2 - L_1L_2\|x_k - \bar{x}_k\|^3,$$

从而有

$$\|\nabla\Phi_{\tau_k}(x_k)^T\Phi_{\tau_k}(x_k)\|^\delta \geq (c^2\|x_k - \bar{x}_k\| - L_1L_2\|x_k - \bar{x}_k\|^2)^\delta.$$

由算法 1 步 2 中 α_k 的取值知 $\alpha_k \geq m$, 于是

$$\alpha_k\|\nabla\Phi_{\tau_k}(x_k)^T\Phi_{\tau_k}(x_k)\|^\delta \geq m(c^2\|x_k - \bar{x}_k\| - L_1L_2\|x_k - \bar{x}_k\|^2)^\delta,$$

即有

$$\alpha_k\|\nabla\Phi_{\tau_k}(x_k)^T\Phi_{\tau_k}(x_k)\|^\delta \geq m\|x_k - \bar{x}_k\|^\delta(c^2 - L_1L_2\|x_k - \bar{x}_k\|)^\delta.$$

因此若 $\|x_k - \bar{x}_k\|$ 充分小, 那么存在常数 $c_1 > 0$, 使得

$$\alpha_k\|\nabla\Phi_{\tau_k}(x_k)^T\Phi_{\tau_k}(x_k)\|^\delta \geq c_1\|x_k - \bar{x}_k\|^\delta.$$

引理 9 若假设 A 成立, 设 $x_{k+1}, x_k \in M(x_*, \frac{1}{2}b)$, 那么有

$$\text{dist}(x_{k+1}, X) \leq c_3\text{dist}(x_k, X)^{\frac{2+\delta}{2}}. \quad (4.8)$$

证 要证 (4.8) 式成立, 首先要证明 $\forall x_k \in M(x_*, \frac{1}{2}b)$, 存在 $c_4 > 0, c_5 > 0$, 使以下不等式成立

$$\|d_k\| \leq c_4\text{dist}(x_k, X) \quad (4.9)$$

及

$$\|\Phi_{\tau_k}(x_k)d_k + \nabla\Phi_{\tau_k}(x_k)^Td_k\| \leq c_5\text{dist}(x_k, X)^{\frac{2+\delta}{2}}. \quad (4.10)$$

首先证明 (4.9) 式成立.

$$\|\bar{x}_k - x_*\| \leq \|\bar{x}_k - x_k\| + \|x_* - x_k\| \leq \|x_* - x_k\| + \|x_* - x_k\| \leq b. \quad (4.11)$$

由 (4.11) 式以及 $x_k \in M(x_*, \frac{1}{2}b)$ 得 $\bar{x}_k \in M(x_*, b)$. 由 (4.2) 式的凸性质知 d_k 是 (4.1) 式全局最优解, 于是有

$$\Psi(d_k) \leq \Psi(\bar{x}_k - x_k). \quad (4.12)$$

那么由假设 A 的 (2), 引理 6, 引理 8, (4.2), (4.9) 式有

$$\begin{aligned} \|d_k\|^2 &\leq \frac{1}{\alpha_k\|\nabla\Phi_{\tau_k}^T\Phi_{\tau_k}\|^\delta}\Psi(d_k) \\ &\leq \frac{1}{\alpha_k\|\nabla\Phi_{\tau_k}^T\Phi_{\tau_k}\|^\delta}\Psi(\bar{x}_k - x_k) \\ &= \frac{1}{\alpha_k\|\nabla\Phi_{\tau_k}^T\Phi_{\tau_k}\|^\delta}\left[\frac{1}{2}\|\Phi_{\tau_k}(x_k) + \nabla\Phi_{\tau_k}^T(\bar{x}_k - x_k)\|^2 + \frac{1}{2}\alpha_k\|\nabla\Phi_{\tau_k}^T\Phi_{\tau_k}\|^\delta\|\bar{x}_k - x_k\|^2\right] \\ &\leq \frac{L_2^2}{2c_1\|\bar{x}_k - x_k\|^\delta}\|\bar{x}_k - x_k\|^4 + \frac{1}{2}\|\bar{x}_k - x_k\|^2 \\ &\leq \frac{L_2^2}{2c_1}\|\bar{x}_k - x_k\|^{4-\delta} + \frac{1}{2}\|\bar{x}_k - x_k\|^2 \\ &\leq \left(\frac{L_2^2}{2c_1} + \frac{1}{2}\right)\|\bar{x}_k - x_k\|^2, \end{aligned} \quad (4.13)$$

从而令

$$c_4 = \sqrt{\frac{L_2^2}{2c_1} + \frac{1}{2}},$$

即得 (4.9) 式.

下面再证明 (4.10) 式成立. 事实上, 由 (4.2) 式, 引理 6, 引理 8, 有

$$\begin{aligned} \|\Phi_{\tau_k}(x_k) + \nabla \Phi_{\tau_k}^T d_k\|^2 &\leq \Psi(d_k) \leq \Psi(\bar{x}_k - x_k) \\ &\leq \frac{1}{2} \|\Phi_{\tau_k}(x_k) + \nabla \Phi_{\tau_k}^T(\bar{x}_k - x_k)\|^2 + \frac{1}{2} \alpha_k \|\nabla \Phi_{\tau_k}^T \Phi_{\tau_k}\|^\delta \|\bar{x}_k - x_k\|^2 \\ &\leq \frac{1}{2} L_2^2 \|\bar{x}_k - x_k\|^4 + \frac{1}{2} c_2 \|\bar{x}_k - x_k\|^{2+\delta} \\ &\leq \left(\frac{1}{2} L_2^2 + \frac{1}{2} c_2\right) \|\bar{x}_k - x_k\|^{2+\delta}, \end{aligned} \quad (4.14)$$

在上式中只需令

$$c_5 = \sqrt{\frac{1}{2} L_2^2 + \frac{1}{2} c_2},$$

即得 (4.10) 式成立.

以下证明本引理的结论. 由假设 A 的 (2), 引理 6, (4.9), (4.10) 式有

$$\begin{aligned} \text{dist}(x_{k+1}, X) &\leq \frac{1}{c} \|\Phi_\tau(x_{k+1})\| = \frac{1}{c} \|\Phi_\tau(x_k + d_k)\| \\ &\leq \frac{1}{c} \|\Phi_{\tau_k}(x_k + d_k) - \Phi_{\tau_k}(x_k) - \nabla \Phi_{\tau_k}^T d_k\| + \frac{1}{c} \|\Phi_{\tau_k}(x_k) + \nabla \Phi_{\tau_k}^T d_k\| \\ &\leq \frac{L_2}{c} \|d_k\|^2 + \frac{c_5}{c} \text{dist}(x_k, X)^{\frac{2+\delta}{2}} \\ &\leq \frac{L_2 c_4^2}{c} \text{dist}(x_k, X)^2 + \frac{c_5}{c} \text{dist}(x_k, X)^{\frac{2+\delta}{2}} \\ &\leq \frac{L_2 c_4^2 (\frac{b}{2})^{\frac{2-\delta}{2}}}{c} \text{dist}(x_k, X)^{\frac{2+\delta}{2}} + \frac{c_5}{c} \text{dist}(x_k, X)^{\frac{2+\delta}{2}} \\ &\leq \frac{L_2 c_4^2 (\frac{b}{2})^{\frac{2-\delta}{2}} + c_5}{c} \text{dist}(x_k, X)^{\frac{2+\delta}{2}} \\ &:= c_3 \text{dist}(x_k, X)^{\frac{2+\delta}{2}}, \end{aligned}$$

其中

$$c_3 = \frac{L_2 c_4^2 (\frac{b}{2})^{\frac{2-\delta}{2}} + c_5}{c}.$$

定理 4 在假设 A 成立的条件下,

(a) 由算法 1 产生的序列 $\{x_k\}$ 收敛于 x_* .

(b) 当 $\delta \in (0, 2)$ 时, $\text{dist}(x_k, X)$ 收敛到 0 是超线性的, 当 $\delta = 2$, $\text{dist}(x_k, X)$ 收敛到 0 是二次的.

证 (a) 由定理 3 和假设 A 的 (1) 可知

$$\{x_k\} \rightarrow x_*, x_* \in X. \quad (4.15)$$

(b) 不等式

$$\text{dist}(x_k, X) \leq \text{dist}(x_k + d_k, X) + \|d_k\|. \quad (4.16)$$

由 (4.8) 和 (4.16) 式有

$$\text{dist}(x_k, X) \leq 2\|d_k\|. \quad (4.17)$$

由 (4.8), (4.9), (4.17) 式有

$$\|d_{k+1}\| = O(\|d_k\|^{\frac{2+\delta}{2}}),$$

那么

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|^{\frac{2+\delta}{2}}} = \lim_{k \rightarrow \infty} \frac{\left\| \sum_{l=k+1}^{\infty} d_l \right\|}{\left\| \sum_{l=k}^{\infty} d_l \right\|^{\frac{2+\delta}{2}}} = \lim_{k \rightarrow \infty} \frac{\|d_{k+1}\|}{\|d_k\|^{\frac{2+\delta}{2}}} \leq p$$

($p \geq 0$ 是一个常数).

当 $\delta \in (0, 2)$ 时, $\text{dist}(x_k, X)$ 收敛到 0 是超线性的. 当 $\delta = 2$ 时, $\text{dist}(x_k, X)$ 收敛到 0 是二次的.

5 数值实验

为了验证该算法的有效性, 本节将选取一些测试问题进行实验, 并与一般 Levenberg-Marquart 方法进行比较. 这里的一般 Levenberg-Marquart 方法由文献 [9] 的算法 3.6.4 改进获得. 记为 ULM. 程序的设计采用的是 Matlab 7.0 进行编译, 其中参数的选取为: $b_0 = 0.05, b_1 = 0.75, \alpha = 0.01, m = 10^{-3}, \delta = 1, \delta_1 = 4, \delta_2 = 0.25, \eta = 0.3, \mu = 0.99$.

例 1 (Kojima-Shindo 问题) 测试函数为

$$F(x) = \begin{cases} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6, \\ 2x_1^2 + x_1 + x_2^2 + 3x_3 + 2x_4 - 2, \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 3x_4 - 1, \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3, \end{cases}$$

该问题有唯一解 $(\frac{\sqrt{6}}{2}, 0, 0, 0.5)^T$. 选取不同的初始点, 数值结果如下表.

表 1: 数值结果

初始点 x_0	算法 1 迭代次数	ULM 迭代次数
(0, 0, 0, 0)	14	32
(1, 1, 1, 1)	9	33
(3, 3, 3, 3)	9	31
(5, 5, 5, 5)	10	30
(15, 15, 10, 10)	10	32
(10, 10, 10, 10)	12	88
(10 ² , 10 ² , 10 ² , 10 ²)	18	90
(10 ³ , 10 ³ , 10 ³ , 10 ³)	31	92
(10 ⁴ , 10 ⁴ , 10 ⁴ , 10 ⁴)	38	94

例 2

$$F(x) = \begin{cases} x_1 - 8, \\ x_2^3 + x_2 - x_3 - 3, \\ x_2 + 2x_3^3 + x_3 - 3, \end{cases}$$

选取不同的初始点, 数值结果如下表.

表 2: 数值结果

初始点 x_0	算法 1 迭代次数	ULM 迭代次数
(0, 0, 0)	7	22
(1, 1, 1)	7	22
(2, 2, 2)	7	22
(8, 1, 1)	3	9
(8, 8, 8)	4	20
(10, 10, 10)	5	21
(10 ² , 10 ² , 10 ² , 10 ²)	13	27
(10 ³ , 10 ³ , 10 ³ , 10 ³)	11	32
(10 ⁴ , 10 ⁴ , 10 ⁴ , 10 ⁴)	15	36

上述数值结果表明, 采用新的 LM 参数并结合信赖域技术的光滑 Levenberg-Marquardt 方法比一般的 Levenberg-Marquardt 方法的计算效率更高.

参 考 文 献

- [1] Yang Yufei, Qi Liqun. Smoothing trust region methods for nonlinear complementarity problems with P_0 functions[J]. Ann. Oper. Res, 2005, 133: 99–117.
- [2] Yu Zhensheng, Su Ku, Lin Ji. A smoothing Levenberg-Marquardt methods for the extended linear complementarity problem[J]. Applied Mathematical Modelling, 2009, 33: 3409–3420.
- [3] Zhang Juliang, Zhang Xiangsun. A smoothing Levenberg-Marquardt method for NCP[J]. Applied Mathematics and Computation, 2006, 178: 212–228.
- [4] 张华仁, 李国维. 一个结合信赖域技术的修正 Levenberg-Marquardt 方法 [J]. 数值计算与计算机应用, 2009, 3(30): 186–194.
- [5] Fan Jinyan. Convergence rate of the trust region method for nonlinear equations under local error bound condition[J]. Computational Optimization and Applications, 2005, 34: 215–227.
- [6] Ma Changfeng, Jiang Lihua. Some research on Levenberg-Marquardt method for nonlinear equations[J]. Applied Mathematics and Computation, 2007, 184: 1032–1040.
- [7] 俞昊东. 解非线性互补问题的非单调信赖域方法 [D]. 上海: 同济大学, 2008.
- [8] Powell M J D. Convergence Properties of a Class of Minimization Algorithms[A]. Mangasarian O L, Meyer R P, Robinson S M, et al. Nonlinear Programming2[C]. New York: Academic Press, 1975: 1–27.
- [9] 袁亚湘, 孙文瑜. 最优化理论与方法 [M]. 北京: 科学出版社, 2007.
- [10] Levenberg K. A method for the solution of certain nonlinear problems in least squares[J]. Quart. Appl. Math., 1944, 2: 164–166.

- [11] Marquardt J E. An algorithm for least-squares estimate of nonlinear inequalities[J]. SIAM J. Annl. Math., 1963, 11: 431–441.
- [12] Jiang Lihua, Ma Changfeng. Damped Gauss-Newton method for solving nonlinear inequalities[J]. Journal of Mathematics, 2009, 29(4): 473–478.

A SMOOTHING LEVENBERG-MARQUARDT METHOD FOR P_0 NONLINERA COMPLEMENTARITY PROBLEM

CHEN Jin-xiong¹, LIU Ning²

(1. Dpt. of Mathematics and Computer Science, Wuyi University, Wuyishan 354300, China)

(2. Mathematics Teaching and Research Group, Yulin No.12 Middle School, Yulin 537000, China)

Abstract: In this paper, we study a P_0 nonlinear complementarity problem. By using the techniques of trust-region, we acquire a smoothing Levenberg-Marquardt algorithm for nonlinear complementarity problem. Under suitable condition the gobal convergence properties of this algorithm are proved. With the local error condition, the local superlinear convergence of this algorithm is also obtained. This algorithm is efficient by the numerical experiments.

Keywords: nonlinear complementarity problem; the method of Levenberg-Marquardt; global convergence; local convergence

2010 MR Subject Classification: 90C30