Vol. 36 (2016)
No. 4

数 学 杂 志
J. of Math. (PRC)

非线性互补问题的两种数值解法

周光辉 [1], 张从军 [2], 张成虎 [2], 王月虎 [3]

(1. 淮北师范大学数学科学学院; 信息学院, 安徽 淮北 235000)
(2. 南京财经大学应用数学学院, 江苏 南京 210023)
(3. 南京财经大学管理科学与工程学院, 江苏 南京 210023)

摘 要: 本文研究了非线性互补问题的两类数值求解方法. 在经典 LQP 算法及 Levenberg-Marquardt 算法的基础上, 构造了两种新算法, 并证明了这两种新算法的收敛性. 数值实验表明, 新算法对测试问题优于已有算法.

关键词: 非线性互补问题; LQP 算法; Levenberg-Marquardt 算法
MR(2010) **主题分类号**: 65K10; 49J40 **中图分类号**: O224; O221
文献标识码: A **文章编号**: 0255-7797(2016)04-0794-15

1 引言

考虑非线性互补问题 NCP (F): 寻找一个向量 $x \in R^n$, 使

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

其中 $F: R^n \to R^n$ 是一个非线性映射.

一个 NCP (F) 可转化为寻找一个极大单调算子 T 的零点问题. 由 Martinet [1] 提出的邻近点算法, 将极大单调算子的零点问题转化为一个极大包含问题. 将极大单调包含问题中的线性项用非线性项代替, 使上述问题转化为非线性方程组的问题, 此类算法称为 LQP 算法. He [2]、Bnouhachem [3]、Noor 和 Bnouhachem [4]、Xu [5] 引进了基于 "预测校正方法" 的 LQP 算法, 能很好的求解上述非线性方程组问题.

预测校正方法, 在每次迭代过程中包括两个步骤: 预测步与校正步. 预测子 (predictor) 是通过在一个误差准则下近似求解 LQP 方程组获得的; 校正步, 亦称迭代步 (new iterate), 一般是由一个投影算子表示的.

例如 Noor 和 Bnouhachem 在文献 [4] 中, 提出了如下算法:

步骤 1 寻找近似解 $\tilde{x}^k$, 使得

$$0 \approx \beta_k F(\tilde{x}^k) + \tilde{x}^k - x^k + \mu X_k \log \frac{\tilde{x}^k}{x^k} =: \xi^k,$$

其中 $\left\|\xi^k\right\| \leq \eta \left\|x^k - \tilde{x}^k\right\|$, $\eta \in (0,1)$.

*收稿日期: 2015-10-29 接收日期: 2016-01-19
基金项目: 安徽省高校自然科学重点项目 (KJ2016A651) 和安徽省教育厅教育教学重点项目 (2014jyxm161) 资助.
作者简介: 周光辉 (1973–), 男, 安徽定远, 副教授, 主要研究方向: 最优化理论与算法、变分法等.

步骤 2 对 $\alpha>0$, $x^{k+1}(\alpha)$ 是下面方程组的正数解:

$$\left(\frac{1-\mu}{1+\mu}\right)\alpha\beta_k F(\tilde{x}^k)+x-x^k+\mu X_k\log\frac{x}{x^k}=0.$$

Yuan 在文献 [6] 中提出了以下算法:

步骤 1 寻找 $\tilde{x}^k\in R_+^n$, 满足 $0\approx\beta_k F(\tilde{x}^k)+\tilde{x}^k-(1-\mu)x^k-\mu X_k^2\left(\tilde{x}^k\right)^{-1}=:\xi^k$, 其中 ξ^k 满足如下误差准则:

$$\left\|\xi^k\right\|\leq\eta\sqrt{1-\mu^2}\left\|x^k-\tilde{x}^k\right\|.$$

步骤 2 令 $x^{k+1}(\alpha_k)=P_{R_+^n}\left[x^k-\alpha_k\frac{\beta_k}{1+\mu}F\left(\tilde{x}^k\right)\right]$.

受文献 [4, 6] 所提算法的启发, 本文提出了一种基于预测校正方法的 LQP 算法. 在新算法中设置了两个预测步, 并使用 x^k 与投影算子的凸组合构成算法的校正步.

此外, 可借助非线性互补函数, 将 NCP (F) 转化为非线性方程组问题:

$$\Phi(x)=0, \tag{1.2}$$

其中 $\Phi: R^n\to R^m$ 是一个非线性映射.

Levenberg-Marquardt 算法可以看作是一种非线性最小二乘法, 是用于求解上述非线性方程组问题的一种改进的 Gauss-Newton 算法. 由于此算法在优化问题中的广泛应用, 对于该算法的研究成果颇丰, 例如可参见文献 [7–11].

Levenberg-Marquardt 算法在每个迭代点处的搜索方向 d_L^k 通过如下线性方程组获得

$$\left(\Phi'\left(x^k\right)^T\Phi'\left(x^k\right)+\mu_k I\right)d=-\Phi'\left(x^k\right)^T\Phi\left(x^k\right), \tag{1.3}$$

其中 $\Phi'(x)$ 表示 $\Phi(x)$ 的 Jacobian, $\mu_k>0$ 是一个参数. 参数 μ_k 的引入克服了 Gauss-Newton 算法中要求矩阵 $\Phi'(x^*)$ 满秩的困难. Yamashita 和 Fukushima [12]、Dan [13] 等学者提出了该参数的更新准则: $\mu_k=\left\|\Phi\left(x^k\right)\right\|^\delta$, 其中 $\delta\in(0,2]$. 本文将沿用这一更新准则.

值得注意的是, 由于 $\Phi'\left(x^k\right)^T\Phi'\left(x^k\right)$ 的半正定性, 正参数 μ_k 的引入使得搜索方向 (试探步)d_L^k 远离了矩阵广义逆下的迭代步 $d_{MP}^k=-\Phi'\left(x^k\right)^+\Phi\left(x^k\right)$. 所以本文将考虑将下式的解 d_G^k 作为修正的搜索方向:

$$\left(\Phi'\left(x^k\right)^T\Phi'\left(x^k\right)+\mu_k I\right)d_G^k=-\Phi'\left(x^k\right)^T\Phi\left(x^k\right)+\mu_k d_L^k. \tag{1.4}$$

这样 d_G^k 很可能比 d_L^k 更靠近 d_{MP}^k.

另外, 为了保证此类非线性方程组算法的全局收敛性, 通常需要假设 F 为 P_0 函数, 并且在证明其超线性收敛性时, 通常需要一些非奇异性和严格互补条件的假设, 参见文献 [2, 12].

考虑到以上两点, 本文给出了一种改进的 Levenberg-Marquardt 算法. 该算法的优点在于搜索方向经过修正能更接近 Moore-Penrose 步, 并且其全局收敛性的证明不需要 F 为 P_0 函数的假设.

本文接下来的部分, 分别给出一种新的 LQP 算法和 Levenberg-Marquardt 算法, 阐述了算法中参数的选取方法, 给出了收敛性定理, 通过数值实验说明了新算法相对已有算法的优劣, 表明了算法的有效性.

2 非线性互补问题的 LQP 算法

2.1 LQP 算法的构造

算法 2.1

步骤 0 设置参数 $\beta_0 > 0$, $\varepsilon > 0$, $\mu \in (0,1)$, $\delta \in (1,2)$, $\sigma \in (1,2)$, $\eta \in (0,1)$, $\tau \in (0,1]$. 初始点为 $x^0 > 0$.

步骤 1 如果 $\left\|\min\left\{x^k, F\left(x^k\right)\right\}\right\|_\infty \le \varepsilon$, 算法终止. 否则, 转入 步骤 2.

步骤 2 (预测步 1) 寻找 $\tilde{x}^k \in R^n_+$, 满足

$$0 \approx \beta_k F(\tilde{x}^k) + \tilde{x}^k - x^k + \mu X_k \log \frac{\tilde{x}^k}{x^k} =: \xi^k, \tag{2.1}$$

其中 ξ^k 满足如下非精确误差准则:

$$\left\|\xi^k\right\| \le \eta\sqrt{1-\mu^2}\left\|x^k - \tilde{x}^k\right\|. \tag{2.2}$$

计算 $\xi^k := \beta_k\left(F\left(\tilde{x}^k\right) - F\left(x^k\right)\right)$, $r = \frac{\left\|\xi^k\right\|}{\sqrt{1-\mu^2}\left\|x^k-\tilde{x}^k\right\|}$. 如果 $r > \eta$, 令 $\beta_k = \beta_k * \frac{0.8}{r}$, 重复步骤 2; 否则, 转入步骤 3.

步骤 3 (预测步 2) 计算最优参数 α_k. 然后, 寻找如下方程组的正数解 $\widehat{x}^k$:

$$\left(\frac{1-\mu}{1+\mu}\right)\alpha\beta_k F(\tilde{x}^k) + x - x^k + \mu X_k \log \frac{\tilde{x}^k}{x^k} = 0. \tag{2.3}$$

步骤 4 (校正步) 计算最优参数 γ_k. 然后, 生成新的迭代点 x^{k+1}:

$$x^{k+1}(\gamma) = \tau x^k + (1-\tau)P_{R^n_+}\left[x^k - \gamma(x^k - \widehat{x}^k)\right]. \tag{2.4}$$

步骤 5 如果 $r \le 0.5$, 令 $\beta_{k+1} = \frac{\beta_k * 0.7}{r}$; 否则, 令 $\beta_{k+1} = \beta_k$.

步骤 6 令 $k = k+1$, 转入 步骤 1.

注 2.1 算法 2.1 中包括了两个预测步和一个校正步. 校正步由一个凸组合表示, 见式 (2.4) 所示. 此外, 算法中对于参数 α 和 γ 的确定是本节研究的重点.

2.2 参数的确定及收敛性分析

令 $\Pi(\alpha) = \left\|x^k - x^*\right\|^2 - \left\|\widehat{x}^k(\alpha) - x^*\right\|^2$, $\Psi(\gamma) = \left\|x^k - x^*\right\|^2 - \left\|x^{k+1}(\gamma) - x^*\right\|^2$.

下面说明如何确定合适的参数 α 和 γ, 同时给出算法 2.1 的收敛性定理.

首先给出两个常用结论.

引理 2.1 [14] 假设 $P_{R^n_+}(.)$ 表示欧式范数下的投影, 即 $P_{R^n_+}(y) = \min\left\{\|y-x\|, x \in R^n_+\right\}$, 那么对任意的 $u \in R^n_+$ 和 $v \in R^n_+$, 有下面两式成立

$$\left\langle v - P_{R^n_+}(v), P_{R^n_+} - u\right\rangle \ge 0,$$

$$\left\|P_{R^n_+}(v) - u\right\|^2 \le \|v-u\|^2 - \left\|v - P_{R^n_+}(v)\right\|^2.$$

引理 2.2 [4] 对给定的 $x^k>0$, $q\in R^n$, 假定 x 是如下方程组的一个正数解

$$q+x-x^k+\mu X_k\log\frac{x}{x^k}=0,$$

其中 $X_K=\operatorname{diag}(x_1^k,\cdots,x_n^k)$, $\log\frac{x}{x^k}=(\log\frac{x_1}{x_1^k},\cdots,\log\frac{x_n}{x_n^k})^T$, 那么对任意的 $y\geq 0$, 有下式成立

$$\langle x-y,-q\rangle\geq\frac{1+\mu}{2}\left(\|x-y\|^2-\left\|x^k-y\right\|^2\right)+\frac{1-\mu}{2}\left\|x^k-x\right\|^2.$$

下面的定理用来说明如何确定参数 α.

定理 2.1 [4] 假设对任意的 $x^*\in\Omega^*$, $\alpha>0$, $\tilde{x}^k$ 和 $\widehat{x}^k$ 是由算法 2.1 生成的, 则有下式成立:

$$\Pi(\alpha)\geq\frac{1-\mu}{1+\mu}\Phi(\alpha),\tag{2.5}$$

其中

$$\Phi(\alpha)=2\alpha\phi_k-\alpha^2\left\|d_k\right\|^2,\tag{2.6}$$

$$d_k=(x^k-\tilde{x}^k)+\frac{1}{1+\mu}\xi^k,\tag{2.7}$$

$$\phi_k=\frac{1}{1+\mu}\left\|x^k-\tilde{x}^k\right\|^2+\frac{1}{1+\mu}\left\langle x^k-\tilde{x}^k,\xi^k\right\rangle.\tag{2.8}$$

注 2.2 通过选取合适的 α 极大化 $\Phi(\alpha)$, 使每次迭代产生的迭代量 $\Pi(\alpha)$ 极大化. 由式 (2.6) 可知 $\Phi(\alpha)$ 是关于 α 的二次函数, 则极大值点位于

$$\alpha_k^*=\frac{\phi_k}{\left\|d_k\right\|^2},\tag{2.9}$$

此时

$$\Phi(\alpha)=\alpha^*\phi_k.\tag{2.10}$$

引理 2.3 假设 $x^*\in\Omega^*$, $\tilde{x}^k$ 和 $\widehat{x}^k$ 是由算法 2.1 生成的, 那么就有下式成立:

$$\Phi(\alpha_k)\geq\frac{1+\mu^2-\eta^2}{4(1+\mu)}\left\|x^k-\widehat{x}^k\right\|^2.\tag{2.11}$$

证 一方面, 由式 (2.2)、式 (2.8)、柯西 - 施瓦茨不等式及 $\mu\in(0,1)$ 有

$$\begin{aligned}\phi_k&=\frac{1}{1+\mu}\left\|x^k-\tilde{x}^k\right\|^2+\frac{1}{1+\mu}\left\langle x^k-\tilde{x}^k,\xi^k\right\rangle\\&\geq\frac{1}{1+\mu}\left\|x^k-\tilde{x}^k\right\|^2-\frac{1}{1+\mu}\left\|x^k-\tilde{x}^k\right\|\left\|\xi^k\right\|\\&\geq\frac{1}{1+\mu}\left(\left\|x^k-\tilde{x}^k\right\|^2-\frac{1-\mu^2}{2}\left\|x^k-\tilde{x}^k\right\|^2-\frac{1}{2(1-\mu^2)}\left\|\xi^k\right\|^2\right)\\&\geq\frac{1}{1+\mu}\frac{1+\mu^2-\eta^2}{2}\left\|x^k-\tilde{x}^k\right\|^2\\&=\frac{1+\mu^2-\eta^2}{2(1+\mu)}\left\|x^k-\tilde{x}^k\right\|^2.\end{aligned}\tag{2.12}$$

另一方面, 由式 (2.2)、式 (2.8) 有

$$\begin{aligned}\phi_k &= \frac{1}{1+\mu}\left\|x^k-\tilde{x}^k\right\|^2+\frac{1}{1+\mu}\left\langle x^k-\tilde{x}^k,\xi^k\right\rangle \\ &= \frac{1}{2}\left[\left\|x^k-\widehat{x}^k\right\|^2+\frac{1-\mu}{1+\mu}\left\|x^k-\widehat{x}^k\right\|^2+\frac{2}{1+\mu}\left\langle x^k-\widehat{x}^k,\xi^k\right\rangle\right] \\ &> \frac{1}{2}\left[\left\|x^k-\widehat{x}^k\right\|^2+\frac{1}{(1+\mu)^2}\left\|\xi^k\right\|^2+\frac{2}{1+\mu}\left\langle x^k-\widehat{x}^k,\xi^k\right\rangle\right].\end{aligned}$$

由于 $\|a+b\|^2=\|a\|^2+\|b\|^2+2\langle a,b\rangle$ 及 $d_k=(x^k-\tilde{x}^k)+\frac{1}{1+\mu}\xi^k$, 则有 $\phi_k\geq\frac{1}{2}\|d_k\|^2$. 再由式 (2.9), 有

$$\alpha_k^*>\frac{1}{2}. \tag{2.13}$$

由式 (2.10)、式 (2.12) 及式 (2.13) 知, 式 (2.11) 成立. 证毕.

注 2.3 为了加快收敛速度, 用松弛因子 $\delta\in(1,2)$ 乘以 α_k^*, 得到 α_k. 即

$$\alpha_k=\delta\alpha_k^*=\delta\frac{\left\|x^k-\tilde{x}^k\right\|^2+\left\langle x^k-\tilde{x}^k,\xi^k\right\rangle}{(1+\mu)\left\|(x^k-\tilde{x}^k)+\frac{1}{1+\mu}\xi^k\right\|^2}.$$

此时

$$\Phi\left(\alpha_k\right)=\Phi\left(\delta\alpha_k^*\right)=\left(2\delta-\delta^2\right)\Phi\left(\alpha_k^*\right)\geq\frac{\delta\left(2-\delta\right)\left(1+\mu^2-\eta^2\right)}{4(1+\mu)}\left\|x^k-\tilde{x}^k\right\|^2. \tag{2.14}$$

由定理 2.1 中式 (2.5), 立有下式成立

$$\begin{aligned}\Pi\left(\alpha_k\right)=\Pi\left(\delta\alpha_k^*\right)&=\left\|x^k-x^*\right\|^2-\left\|\widehat{x}^k\left(\delta\alpha_k^*\right)-x^*\right\|^2 \\ &\geq\frac{1-\mu}{1+\mu}\Phi\left(\delta\alpha_k^*\right)\geq\frac{\delta\left(2-\delta\right)\left(1-\mu\right)\left(1+\mu^2-\eta^2\right)}{4(1+\mu)^2}\left\|x^k-\tilde{x}^k\right\|^2.\end{aligned}$$

采用类似于定理 2.1 的思想, 给出定理 2.2, 分析参数 γ 的选取方法. 为描述简单, 令 $x_*^k\left(\gamma\right):=P_{R_+^n}\left[x^k-\gamma(x^k-\widehat{x}^k)\right]$.

定理 2.2 假设 $x^*\in\Omega^*$, 且 $\tilde{x}^k$ 和 $\widehat{x}^k$ 是由算法 2.1 生成的, 则有 $\Psi\left(\gamma\right)\geq(1-\tau)\Theta(\gamma)$, 其中

$$\Theta(\gamma)=\gamma\left[\left\|x^k-\widehat{x}^k\right\|^2+\left\|x^k-x^*\right\|^2-\left\|\widehat{x}^k-x^*\right\|^2\right]-\gamma^2\left\|x^k-\widehat{x}^k\right\|^2. \tag{2.15}$$

证 由 $\|a+b\|^2=\|a\|^2+\|b\|^2+2\langle a,b\rangle$ 及 $\langle a+b,b\rangle=\langle a,b\rangle+\|b\|^2$, 有下面两式成立

$$\begin{aligned}&2\left\langle x^k-x_*^k\left(\gamma\right),x_*^k\left(\gamma\right)-x^*\right\rangle=\left\|x^k-x^*\right\|^2-\left\|x^k-x_*^k\left(\gamma\right)\right\|^2+\left\|x_*^k\left(\gamma\right)-x^*\right\|^2, \\ &2\left\langle x^*-\widehat{x}^k,x^k-\widehat{x}^k\right\rangle=\left\|x^*-\widehat{x}^k\right\|^2-\left\|x^*-x^k\right\|^2+\left\|x^k-\widehat{x}^k\right\|^2.\end{aligned}$$

由引理 2.1, 有

$$\left\|x_*^k\left(\gamma\right)-x^*\right\|^2\leq\left\|x^k-\gamma\left(x^k-\widehat{x}^k\right)-x^*\right\|^2-\left\|x^k-\gamma\left(x^k-\widehat{x}^k\right)-x_*^k\left(\gamma\right)\right\|^2.$$

注意到 $2\langle a+b,b\rangle=\|a+b\|^2-\|a\|^2+\|b\|^2$ 及 $\tau\in(0,1)$, 有

$$\begin{aligned}\left\|x^{k+1}(\gamma)-x^*\right\|^2=&\left\|\tau\left(x^k-x^*\right)+\tau(1-\tau)\left(x_*^k(\gamma)-x^*\right)\right\|^2\\=&\tau\left\|x^k-x^*\right\|^2+(1-\tau)\left\|\left(x_*^k(\gamma)-x^*\right)\right\|^2-\tau(1-\tau)\left\|x^k-x_*^k(\gamma)\right\|^2\\\le&\tau\left\|x^k-x^*\right\|^2+(1-\tau)\left[\left\|x^k-\gamma\left(x^k-\widehat{x}^k\right)-x^*\right\|^2\right.\\&\left.-\left\|x^k-\gamma\left(x^k-\widehat{x}^k\right)-x_*^k(\gamma)\right\|^2-\tau\left\|x^k-x_*^k(\gamma)\right\|^2\right],\end{aligned}$$

那么

$$\begin{aligned}\left\|x^{k+1}(\gamma)-x^*\right\|^2=&\left\|x^k-x^*\right\|^2-(1-\tau)\left[\left\|x^k-x_*^k(\gamma)-\gamma\left(x^k-\widehat{x}^k\right)\right\|^2\right.\\&\left.+\tau\left\|x^k-x_*^k(\gamma)\right\|^2-\gamma^2\left\|x^k-\widehat{x}^k\right\|^2+2\gamma\left\langle x^k-x^*,x^k-\widehat{x}^k\right\rangle\right]\\\le&\left\|x^k-x^*\right\|^2-(1-\tau)\left[2\gamma\left\langle x^k-x^*,x^k-\widehat{x}^k\right\rangle-\gamma^2\left\|x^k-\widehat{x}^k\right\|^2\right].\end{aligned}\tag{2.16}$$

由 $\Psi(\gamma)$ 的定义、式 (2.16) 及 $\|a-c\|^2=\langle a-b,a-c\rangle+\langle b-c,a-c\rangle$, 有

$$\begin{aligned}\Psi(\gamma)=&\left\|x^k-x^*\right\|^2-\left\|x^{k+1}(\gamma)-x^*\right\|^2\\\ge&(1-\tau)\left[2\gamma\left\langle x^k-x^*,x^k-\widehat{x}^k\right\rangle-\gamma^2\left\|x^k-\widehat{x}^k\right\|^2\right]\\=&(1-\tau)\left[2\gamma\left(\left\|x^k-\widehat{x}^k\right\|^2-\left\langle x^*-\widehat{x}^k,x^k-\widehat{x}^k\right\rangle\right)-\gamma^2\left\|x^k-\widehat{x}^k\right\|^2\right]\\\ge&(1-\tau)\left[\gamma\left(2\left\|x^k-\widehat{x}^k\right\|^2-\left\|x^*-\widehat{x}^k\right\|^2+\left\|x^*-x^k\right\|^2-\left\|x^k-\widehat{x}^k\right\|^2\right)-\gamma^2\left\|x^k-\widehat{x}^k\right\|^2\right]\\=&(1-\tau)\Theta(\gamma).\end{aligned}$$

证毕.

注 2.4 由于 $\Psi(\gamma)$ 衡量了算法在第 $k+1$ 步相对于第 k 步的改进, 通过选取合适的 γ 极大化 $\Theta(\gamma)$, 使每次迭代产生的迭代量 $\Psi(\gamma)$ 极大化. 由式 (2.15) 可知 $\Theta(\gamma)$ 是关于 γ 的二次函数, 由 $\Phi(\alpha)$ 的定义式及定理 2.1, 则极大值点位于

$$\gamma_k^*=\frac{\left\|x^k-\widehat{x}^k\right\|^2+\Phi(\alpha_k)}{2\left\|x^k-\widehat{x}^k\right\|^2}.$$

注 2.5 为了加快收敛速度, 用松弛因子 $\sigma\in(0,1)$ 乘以 γ_k^*, 得到 γ_k, 即

$$\gamma_k=\sigma\gamma_k^*=\sigma\frac{\left\|x^k-\widehat{x}^k\right\|^2+\Phi(\alpha_k)}{2\left\|x^k-\widehat{x}^k\right\|^2}.\tag{2.17}$$

下面的定理 2.3 是算法 2.1 收敛性分析过程中一个重要的结论.

定理 2.3 假设 $x^*\in\Omega^*$, $\sigma\in(0,1)$, $\tilde{x}^k$, $\widehat{x}^k$ 及 $x^{k+1}(\gamma_k)$ 是由算法 2.1 生成的, 那么就有下式成立:

$$\Psi(\gamma_k)\geq\frac{\sigma(2-\sigma)(1-\tau)\delta^2(2-\delta)^2(1+\mu^2-\eta^2)^2}{64(1+\mu)^2}\left\|x^k-\tilde{x}^k\right\|^2.$$

证 由式 (2.13), (2.14), (2.17) 及定理 2.2, 有

$$\begin{aligned}\Psi(\gamma)&\geq(1-\tau)\Theta(\gamma)\\&=(1-\tau)\left\{\gamma_k\left[\left\|x^k-\widehat{x}^k\right\|^2+\left\|x^k-x^*\right\|^2-\left\|\widehat{x}^k-x^*\right\|^2\right]-\gamma_k^2\left\|x^k-\widehat{x}^k\right\|^2\right\}\\&\geq(1-\tau)\left\{\gamma_k\left[\left\|x^k-\widehat{x}^k\right\|^2+\Phi(\alpha_k)\right]-\gamma_k^2\left\|x^k-\widehat{x}^k\right\|^2\right\}\\&=\sigma(1-\tau)\left\{\gamma_k^*\left[\left\|x^k-\widehat{x}^k\right\|^2+\Phi(\alpha_k)\right]-\sigma(\gamma_k^*)^2\left\|x^k-\widehat{x}^k\right\|^2\right\}\\&=\sigma(2-\sigma)(1-\tau)\gamma_k^*\frac{\left\|x^k-\widehat{x}^k\right\|^2+\Phi(\alpha_k)}{2}\\&\geq\frac{\sigma(2-\sigma)(1-\tau)\delta^2(2-\delta)^2(1+\mu^2-\eta^2)^2}{64(1+\mu)^2}\left\|x^k-\tilde{x}^k\right\|^2.\end{aligned}$$

证毕.

定理 2.3 表明, 算法 2.1 生成的序列 $\{x^k\}$ 关于式 (2.1) 的解点是 Féjer 单调的. 故有如下推论成立.

推论 2.1 令 $x^*\in\Omega^*$, 且 $\{\tilde{x}^k\}$、$\{x^k\}$ 是由算法 2.1 生成, 那么

(1) 序列 $\{x^k\}$ 有界;

(2) 序列 $\{\left\|x^k-x^*\right\|\}$ 是非增的;

(3) $\lim\limits_{k\to\infty}\left\|x^k-\tilde{x}^k\right\|=0$; (4) 序列 $\{\tilde{x}^k\}$ 有界.

下面给出算法 2.1 的收敛性证明过程中的一个引理.

引理 2.4 给定 $x^k\in R^n_{++}$, $\beta_k>0$, 且 $\tilde{x}^k$ 和 ξ^k 满足式 (2.1) 和式 (2.2), 那么对任意的 $x\in R^n_+$ 有下式成立:

$$\left\langle\tilde{x}^k-x,\beta_kF\left(\tilde{x}^k\right)-\xi^k\right\rangle\geq\left\langle x^k-\tilde{x}^k,(1+\mu)x-\mu x^k-\tilde{x}^k\right\rangle. \tag{2.18}$$

证 在引理 2.2 中, 假定 $q=\beta_kF\left(\tilde{x}^k\right)-\xi^k$, $y=x$, 立有式 (2.18) 成立. 证毕.

最后给出算法 2.1 的全局收敛性定理.

定理 2.4 如果 $\inf_{k=0}^{\infty}\beta_k=\beta>0$, 那么由算法 2.1 生成的序列 $\{x^k\}$ 收敛到 $NCP(F)$ 的一个解 x^∞.

利用参考文献 [5] 类似的证明方法, 由定理 2.3, 推论 2.1 及引理 2.4, 此定理可被证明, 不再赘述.

2.3 数值实验

本节将算法 2.1 与文献 [4] 中的算法作比较. 考虑如下 $NCP\,(F)$:

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

其中 $F\,(x) = D\,(x) + Mx + q$. 式中的 $D\,(x)$ 和 $Mx + q$ 分别为 $F\,(x)$ 的非线性项和线性项. 特别地, 令矩阵 $M = A^T A + B$, 其中 A 是元素在区间 $(-5, +5)$ 上随机生成的 $n \times n$ 阶矩阵, B 是元素在区间 $(-5, +5)$ 上随机生成的 $n \times n$ 阶反对称矩阵. 向量 $q \in R^n$ 的元素在区间 $(-200, +300)$ 上服从一致分布. $D\,(x)$ 的分量为 $D_j\,(x) = d_j * \arctan\,(x_j)$, 其中 $d_j \in (0, 1)$. 文献 [4–6, 14] 测试了类似的问题.

测试上述 NCP (F) 的维数 n 从 100 到 1000. 为了与文献 [4] 中的算法作对比, 选取相同的参数: $\beta_0 = 1$, $\eta = 0.91$, $\mu = 0.01$, $\delta = \sigma = 1.99$, $\tau = 0.01$; 终止准则为

$$\left\|\min\left\{x^k, F\left(x^k\right)\right\}\right\|_\infty \leq 10^{-7},$$

初始点为 $x^0 = (0, 0, \cdots, 0)^T$ 或 $x^0 = (1, 1, \cdots, 1)^T$.

利用 MATLAB2013a 编程, 对比结果见下表 2.1 及 2.2 所示, 其中 No.iter. 表示迭代次数; CPU(s) 表示计算时间, 单位为秒.

表 2.1: 初始点为 $x^0 = (0, 0, \cdots, 0)^T$ 时的数值实验结果

维数 (n)	文献 [4] 中算法		算法 2.1	
	No.iter.	CPU(s)	No.iter.	CPU(s)
100	209	0.04	198	0.03
200	243	0.36	236	0.31
400	269	1.91	278	2.16
600	253	6.02	259	5.83
800	263	13.9	262	12.71

表 2.2: 初始点为 $x^0 = (1, 1, \cdots, 1)^T$ 时的数值实验结果

维数 (n)	文献 [4] 中算法		算法 2.1	
	No.iter.	CPU(s)	No.iter.	CPU(s)
100	229	0.06	218	0.10
200	272	0.37	271	0.32
400	283	2.01	279	2.23
600	296	6.9	303	5.97
800	311	16.63	309	16.32

从表中两种算法的迭代次数和计算时间来看, 虽然数值实验的部分结果较文献 [13] 中提出的算法表现差, 但整体上, 算法 2.1 比文献 [4] 中提出的算法更加有效. 例如, 从表 2.1 中可看出当维数为 400 时, 算法 2.1 的 No.iter. 和 CPU(s) 分别为 278 次和 2.16 秒差于文献 [4] 中所提算法的 269 次的迭代次数和 1.91 秒的计算时间. 但是, 从表 2.2 中可以看出当维数为 800 时, 本文所改进的算法的 No.iter. 和 CPU(s) 分别为 309 次和 16.32 秒, 显然要优于文献 [4] 中所提算法的 311 次的迭代次数和 16.63 秒的计算时间.

3 非线性互补问题的 Levenberg-Marquardt 算法

3.1 Levenberg-Marquardt 算法的构造及适定性分析

算法 3.1

步骤 0 给定初始点 $x^0 \in R^n$. 选取参数 $\eta \in (0,1)$, $\lambda \in (0,1)$, $\alpha \in (0,0.8)$, $\sigma \in \left(0, \frac{1}{2}(1-\alpha)\right)$, $\delta \in (0,2]$. 令 $\mu_0 = \|\Phi(x^0)\|^\delta$, $\kappa := \sqrt{2n}$, $\beta_0 := \|\Phi(x^0)\|$, $\tau_0 := \left(\frac{\alpha}{2\kappa}\beta_0\right)^2$, $k := 0$.

步骤 1 如果 x^k 满足 $\left\|\nabla\Psi\left(x^k\right)\right\| \le \varepsilon$, 则算法终止.

步骤 2 计算如下 LM 线性方程组, 得到搜索方向 $d_L^k \in R^n$,

$$\left(\Phi'_{\tau_k}\left(x^k\right)^T \Phi'_{\tau_k}\left(x^k\right) + \mu_k I\right) d_L^k = -\Phi'_{\tau_k}\left(x^k\right)^T \Phi_{\tau_k}\left(x^k\right), \tag{3.1}$$

其中 $\mu_k = \left\|\Phi\left(x^k\right)\right\|^\delta$.

步骤 3 计算如下方程组, 得到修正的搜索方向 $d_G^k \in R^n$,

$$\left(\Phi'_{\tau_k}\left(x^k\right)^T \Phi'_{\tau_k}\left(x^k\right) + \mu_k I\right) d_G^k = -\Phi'_{\tau_k}\left(x^k\right)^T \Phi_{\tau_k}\left(x^k\right) + \mu_k d_L^k. \tag{3.2}$$

步骤 4 令 $\sigma_k = \min\left\{\sigma, \frac{1}{4}\mu_k\right\}$, 并计算使下式成立的最小的非负整数 $s_k \in \{0,1,\cdots\}$:

$$\Psi_{\tau_k}\left(x^k + \lambda^{s_k} d_G^k\right) \le \Psi_{\tau_k}\left(x^k\right) - \sigma_k \lambda^{s_k} \left\|d_G^k\right\|^2. \tag{3.3}$$

令 $t_k := \lambda^{s_k}$, $x^{k+1} = x^k + t_k d_G^k$.

步骤 5 若

$$\left\|\Phi\left(x^{k+1}\right)\right\| \le \max\left\{\eta\beta_k, \frac{1}{\alpha}\left\|\Phi\left(x^{k+1}\right) - \Phi_{\tau_k}\left(x^{k+1}\right)\right\|\right\}, \tag{3.4}$$

则令 $\beta_{k+1} := \left\|\Phi\left(x^{k+1}\right)\right\|$, 并选取合适的 τ_{k+1}, 满足

$$0 < \tau_{k+1} \le \min\left\{\left(\frac{\alpha}{2\kappa}\beta_{k+1}\right)^2, \frac{1}{4}\tau_k, \bar\tau\left(x^{k+1}, \gamma\beta_{k+1}\right)\right\}, \tag{3.5}$$

其中 $\bar\tau(.)$ 的定义见引理 3.2. 否则, 令 $\beta_{k+1} := \beta_k$, 并选取合适的 τ_{k+1}, 满足

$$0 < \tau_{k+1} \le \min\left\{\left(\frac{\alpha}{2\kappa}\left\|\Phi\left(x^{k+1}\right)\right\|\right)^2, \frac{1}{4}\tau_k\right\}. \tag{3.6}$$

步骤 6 令 $k := k+1$, 转到 步骤 1.

注 3.1 在算法 3.1 中, 对搜索方向 d_L^k 进行了修正, 使得新的搜索方向更接近 Moore-Penrose 步. 另外, 为使算法更加精确有效, 针对文献 [15] 中的算法中所涉及到的一些参数及不等式关系做了适当的修正.

为了讨论简便, 定义指标集

$$K := \{0\} \cup \left\{k \in N \,\middle|\, \left\|\Phi\left(x^k\right)\right\| \le \max\left\{\eta\beta_{k-1}, \frac{1}{\alpha}\left\|\Phi\left(x^k\right) - \Phi_{\tau_k-1}\left(x^k\right)\right\|\right\}\right\}. \tag{3.7}$$

下面先给出连续函数的广义 Jacobian 的定义.

定义 3.1 假定 $G: R^n \to R^n$ 是局部 Lipschitz 连续函数. 那么, G 在点 $x \in R^n$ 处的 Clarke [16] 意义下的广义 Jacobian 定义为

$$\partial G(x) = \mathrm{conv}\left\{ H \in R^{n\times n} \,\middle|\, H = \lim_{x^k \to x} G'\left(x^k\right), x^k \in D_G \right\},$$

其中 D_G 是由 G 的可微点组成的集合, $\mathrm{conv}\,(A)$ 是集合 A 的凸包.

Qi [17] 称 $\partial_C G(x)^T$ 为 G 在点 $x \in R^n$ 处的 C - 次微分, 其中 $\partial_C G(x)^T := \partial G_1(x) \times \cdots\cdots \times \partial G_n(x)$.

下面的引理是分析算法收敛性时必要的结论.

引理 3.1

(1) [8] 对任意的 $x \in R^n$, $\tau_1, \tau_2 \geq 0$, 成立 $\|\Phi_{\tau_1} - \Phi_{\tau_2}\| \leq \kappa \left|\sqrt{\tau_1} - \sqrt{\tau_2}\right|$, 其中 $\kappa = \sqrt{2n}$. 特别地, 对任意的 $x \in R^n$, $\tau \geq 0$, 有 $\|\Phi_\tau - \Phi\| \leq \kappa\sqrt{\tau}$.

(2) [15] 对所有的 $k \in N$, 成立

$$\left\|\Phi_{\tau_k}\left(x^{k+1}\right)\right\| < \left\|\Phi_{\tau_k}\left(x^k\right)\right\|$$

和

$$\left\|\Phi_{\tau_k}\left(x^{k+1}\right)\right\| + \kappa\sqrt{\tau_{k+1}} \leq \left\|\Phi_{\tau_k}\left(x^k\right)\right\| + \kappa\sqrt{\tau_k}.$$

本章所给出算法中, 涉及到 $\bar{\tau}$ 的表达式, 因此给出下面一个结论.

引理 3.2 [9] 假定 $x \in R^n$ 是任意给定的向量, x 不是 NCP$\,(F)$ 的解. 定义如下常量

$$\gamma(x) := \max_{i \notin \beta(x)} \left\{\left\|x_i e_i + F_i(x) \nabla F_i(x)\right\|\right\} \geq 0$$

和

$$\alpha(x) := \max_{i \notin \beta(x)} \left\{x_i^2 + F_i(x)^2\right\} > 0,$$

其中 $\beta(x) := \{i \,|x_i = F_i(x) = 0\}$. 假定 δ 为一给定的常数, 定义

$$\bar{\tau}(x, \delta) := \begin{cases} 1, & \dfrac{n\gamma(x)^2}{\delta^2} - \alpha(x) \leq 0, \\ \dfrac{\alpha(x)^2}{2} \dfrac{\delta^2}{n\gamma(x)^2 - \delta^2\alpha(x)}, & \dfrac{n\gamma(x)^2}{\delta^2} - \alpha(x) > 0, \end{cases}$$

则对任意的 $\tau \in (0, \bar{\tau}(x, \delta)]$, 有 $\mathrm{dist}\left(\Phi'_\tau(x), \partial_C \Phi(x)\right) \leq \delta$.

下面证明算法 3.1 的适定性.

定理 3.1 算法 3.1 是适定的.

证 只要证明当 λ 足够小时, Armijo 线性搜索准则满足即可. 由式 (3.1), (3.2), $0 < \sigma_k \leq \mu_k$ 及 $\nabla\Psi_\tau(x) = \left(\Psi'_\tau(x)\right)^T = \left(\Phi'_\tau(x)\right)^T \Phi_\tau(x)$, 对足够小的 λ 有

$$\begin{aligned} &\Psi_{\tau_k}\left(x^k + \lambda d_G^k\right) - \Psi_{\tau_k}\left(x^k\right) = \lambda\nabla\Psi_{\tau_k}\left(x^k\right)^T d_G^k + o(\lambda) \\ \leq& -\lambda\left(\Phi_{\tau_k}\left(x^k\right)^T \Phi'_{\tau_k}\left(x^k\right) - \mu_k\left(d_L^k\right)^T\right) d_G^k + o(\lambda) \\ =& -\lambda\left(d_G^k\right)^T\left(\nabla\Phi_{\tau_k}\left(x^k\right)\Phi'_{\tau_k}\left(x^k\right) + \mu_k I\right) d_G^k + o(\lambda) \\ \leq& -\lambda\mu_k\left\|d_G^k\right\|^2 + o(\lambda) \leq -\lambda\sigma_k\left\|d_G^k\right\|^2 + o(\lambda). \end{aligned}$$

故算法 3.1 是适定的. 证毕.

3.2 收敛性分析

这一部分分析算法 3.1 的全局收敛性. 先介绍一个引理.

引理 3.3 [8] 令 $\left\{x^k\right\}$ 是由算法 3.1 生成的序列. 假设该序列至少有一个聚点 x^*, 则有

(1) 如果 x^* 是 $\mathrm{NCP}(F)$ 的一个解, 则指标集 K 是无限集, 且 $\{\tau_k\} \to 0$.

(2) 指标集 K 是无限集, 当且仅当 $\left\{x^k\right\}$ 的每个聚点都是 $\mathrm{NCP}(F)$ 的一个解.

下面利用奇异值分解定理简化 d_L^k, d_G^k. 由式 (3.1), (3.2) 知

$$
\begin{aligned}
d_L^k &= -\left[\nabla\Phi_{\tau_k}\left(x^k\right)\Phi_{\tau_k}'\left(x^k\right)+\mu_k I\right]^{-1}\nabla\Phi_{\tau_k}\left(x^k\right)\Phi_{\tau_k}\left(x^k\right),\\
d_G^k &= -\left[\nabla\Phi_{\tau_k}\left(x^k\right)\Phi_{\tau_k}'\left(x^k\right)+\mu_k I\right]^{-1}\left[\nabla\Phi_{\tau_k}\left(x^k\right)\Phi_{\tau_k}\left(x^k\right)-\mu_k d_L^k\right].
\end{aligned}
$$

由奇异值分解定理, $\Phi_{\tau_k}'\left(x^k\right)$ 被分解为

$$
\Phi_{\tau_k}'\left(x^k\right)=U_k\Sigma_k V_k^T=U_k\begin{bmatrix}\sigma_{k,1} & & & \\ & \sigma_{k,2} & & \\ & & \ddots & \\ & & & \sigma_{k,n}\end{bmatrix}V_k^T,
$$

其中 U_k, $V_k \in R^{n\times n}$ 是正交矩阵, 且 $\sigma_{k,1}\geq\sigma_{k,2}\geq\cdots\geq\sigma_{k,n}\geq 0$. 那么有

$$
\begin{aligned}
d_L^k &= -V_k\Lambda_k V_k^T\nabla\Psi_{\tau_k}\left(x^k\right),\\
d_G^k &= -V_k\Lambda_k V_k^T\left[\nabla\Psi_{\tau_k}\left(x^k\right)-\mu_k d_L^k\right]\\
&= -V_k\Lambda_k V_k^T\left[I+\mu_k V_k\Lambda_k V_k^T\right]\nabla\Psi_{\tau_k}\left(x^k\right)\\
&= -V_k\left(\Lambda_k\Theta_k\right)V_k^T\nabla\Psi_{\tau_k}\left(x^k\right)\\
&= -V_k X_k V_k^T\nabla\Psi_{\tau_k}\left(x^k\right),
\end{aligned}
\tag{3.8}
$$

其中

$$
\Lambda_k=\begin{bmatrix}\frac{1}{\sigma_{k,1}^2+\mu_k} & & \\ & \ddots & \\ & & \frac{1}{\sigma_{k,n}^2+\mu_k}\end{bmatrix},\Theta_k=\begin{bmatrix}1+\frac{\mu_k}{\sigma_{k,1}^2+\mu_k} & & \\ & \ddots & \\ & & \frac{\mu_k}{\sigma_{k,n}^2+\mu_k}\end{bmatrix},
$$
$$
X_k=\begin{bmatrix}\frac{1}{\sigma_{k,1}^2+\mu_k}\left(1+\frac{\mu_k}{\sigma_{k,1}^2+\mu_k}\right) & & \\ & \ddots & \\ & & \frac{1}{\sigma_{k,n}^2+\mu_k}\left(1+\frac{\mu_k}{\sigma_{k,n}^2+\mu_k}\right)\end{bmatrix}.
$$

那么

$$
\left\|d_G^k\right\|^2=\nabla\Psi_{\tau_k}\left(x^k\right)^T V_k X_k^2 V_k^T\nabla\Psi_{\tau_k}\left(x^k\right). \tag{3.9}
$$

最后, 采用文献 [14] 中收敛性分析相似的思想, 证明算法 3.1 的全局收敛性.

定理 3.2 假设 $\left\{x^k\right\}$ 是由算法 3.1 生成的序列, $\left\{x^k\right\}_L$ 是其收敛于 x^* 的一个子列, 则 x^* 是函数 Ψ 的稳定点.

证 假设指标集 K 是无限集, 由引理 3.3 知, 结论成立.

假设指标集 K 是有限集. 令 $K\cap L=\emptyset$, $\widehat{k}$ 是 K 中最大的数. 那么, 由算法 3.1 中 步骤 5 知对所有的 $k>\widehat{k}$, 有

$$\tau_k\le\tau_{\widehat{k}},\beta_k=\beta_{\widehat{k}}=\left\|\Phi\left(x^{\widehat{k}}\right)\right\|,\left\|\Phi\left(x^k\right)\right\|>\eta\beta_{k-1}=\eta\left\|\Phi\left(x^{\widehat{k}}\right)\right\|>0.$$

那么对所有的 $k>\widehat{k}$,

$$\Psi\left(x^k\right)=\frac{1}{2}\left\|\Phi\left(x^k\right)\right\|^2>\frac{1}{2}\eta^2\left\|\Phi\left(x^{\widehat{k}}\right)\right\|^2=\eta^2\Psi\left(x^{\widehat{k}}\right)>0.$$

下面证明 $\nabla\Psi\left(x^*\right)=0$. 如若不然, 有 $\nabla\Psi\left(x^*\right)\neq 0$.

首先证明 $\liminf\limits_{k\in L} t_k=0$. 假设 $\liminf\limits_{k\in L} t_k=t^*>0$. 由 Armijo 线搜索准则 (3.3), 对所有的 $k\in L$,

$$\Psi_{\tau_k}\left(x^{k+1}\right)-\Psi_{\tau_k}\left(x^k\right)\le-\sigma_k t_k\left\|d_G^k\right\|^2. \tag{3.10}$$

由于 $\{x\}_L\to x^*$, $\{\tau_k\}\to 0$, 则存在某个 $m>0$ 使得对所有的 $k\in L$, $i=1,\cdots,n$, 有 $\left|\sigma_{k,i}^2\right|<m$. 由假设 $\nabla\Psi\left(x^*\right)\neq 0$, $\{\mu_k\}_L\to\mu^*=\left\|\Phi\left(x^*\right)\right\|^\delta>0$, 故存在向量 $M>0$, 对充分大的 $k\in L$, 有

$$\left\|d_G^k\right\|^2\ge\left[\frac{1}{m+\mu_k}\left(1+\frac{\mu_k}{m+\mu_k}\right)\right]^2\left\|\nabla\Phi_{\tau_k}\left(x^k\right)\right\|^2\ge M\left\|\nabla\Phi_{\tau_k}\left(x^k\right)\right\|^2. \tag{3.11}$$

由式 (3.10), (3.11), 对充分大的 $k\in L$, 有

$$\Psi_{\tau_k}\left(x^{k+1}\right)-\Psi_{\tau_k}\left(x^k\right)\le-\sigma_k t_k M\left\|\nabla\Phi_{\tau_k}\left(x^k\right)\right\|^2. \tag{3.12}$$

由于 K 是有限集, $\{\sigma_k\}_L\to\sigma^*=\min\left\{\sigma,\frac{1}{4}\left\|\Phi\left(x^*\right)\right\|^\delta\right\}>0$ 及 $\{\tau_k\}\to 0$, 假设 $c:=\sigma^*t^*M\left\|\nabla\Phi_{\tau_k}\left(x^k\right)\right\|^2$, 那么由式 (3.12), 对充分大的 $k\in L$, 有

$$\Psi_{\tau_k}\left(x^{k+1}\right)-\Psi_{\tau_k}\left(x^k\right)\le-\frac{c}{2}. \tag{3.13}$$

假定 $L=\{l_0,l_1,l_2\cdots.\}$. 由引理 3.1 有 $\left\|\Phi\left(x^{l_{j+1}}\right)\right\|\le\left\|\Phi\left(x^{l_j+1}\right)\right\|+2\kappa\sqrt{\tau_{l_j+1}}$. 那么对充分大的 l_j, 有

$$\begin{aligned}\Psi\left(x^{l_{j+1}}\right)&=\frac{1}{2}\left\|\Phi\left(x^{l_{j+1}}\right)\right\|^2\le\left(\left\|\Phi\left(x^{l_j+1}\right)\right\|+2\kappa\sqrt{\tau_{l_j+1}}\right)^2\\&=\Psi\left(x^{l_j+1}\right)+2\kappa\sqrt{\tau_{l_j+1}}\left\|\Phi\left(x^{l_j+1}\right)\right\|+2\kappa^2\tau_{l_j+1}=\Psi\left(x^{l_j+1}\right)+\frac{c}{4}.\end{aligned} \tag{3.14}$$

由式 (3.13)、式 (3.14), 对充分大的 l_j, 有

$$\Psi\left(x^{l_{j+1}}\right)-\Psi\left(x^{l_j}\right)=\left(\Psi\left(x^{l_{j+1}}\right)-\Psi\left(x^{l_j+1}\right)\right)+\left(\Psi\left(x^{l_j+1}\right)-\Psi\left(x^{l_j}\right)\right)\le-\frac{c}{4},$$

这与 $\Psi(x) \geq 0$ 矛盾, 故 $\liminf\limits_{k\in L} t_k = 0$. 因此可以令 $\lim\limits_{k\in L} t_k = 0$. 那么对于足够大的 $k \in L$, λ^{s_k-1} 总不满足式 (3.3), 即对充分大的 $k \in L$, 有

$$\frac{\Psi_{\tau_k}\left(x^k + \lambda^{s_k-1} d_G^k\right) - \Psi_{\tau_k}\left(x^k\right)}{\lambda^{s_k-1}} > -\sigma \left\|d_G^k\right\|^2. \tag{3.15}$$

假定 $\left\{d^k\right\}_L \to d^*$, $\{\sigma_k\}_L \to \sigma^*$, 由式 (3.15) 有

$$\nabla\Psi\left(x^*\right)^T d^* \geq -\sigma^* \left\|d^*\right\|^2. \tag{3.16}$$

假定 $\{V_k\}_L \to V_*$, $\{X_k\}_L \to X_*$, $\mu^* = \left\|\Phi\left(x^*\right)\right\|^\delta > 0$, 其中

$$X_* = \operatorname{diag}\left(\frac{1}{\sigma_1^2 + \mu^*}\left(1 + \frac{\mu^*}{\sigma_1^2 + \mu^*}\right), \cdots, \frac{1}{\sigma_n^2 + \mu^*}\left(1 + \frac{\mu^*}{\sigma_n^2 + \mu^*}\right)\right).$$

由式 (3.8), (3.9) 及式 (3.16), 则有

$$-\nabla\Psi\left(x^*\right)^T V_* X_* V_*^T \nabla\Psi\left(x^*\right) \geq -\sigma^* \nabla\Psi\left(x^*\right)^T V_* X_*^2 V_*^T \nabla\Psi\left(x^*\right). \tag{3.17}$$

又 $\sigma^* \leq \frac{1}{4}\mu^*$, 则由式 (3.17) 有 $\nabla\Psi\left(x^*\right) = 0$. 证毕.

3.3 数值实验

将算法 3.1 与文献 [15] 中的算法做比较. 给出两个算例, 说明算法 3.1 的优越性及有效性.

参数设定如下: $\varepsilon = 1.0e-7$, $\eta = 0.8$, $\lambda = 0.5$, $\alpha = 0.7$, $\sigma = 0.05$, $\delta = 2$, $\kappa := \sqrt{2n}$.

例 1 假定 $F(x) = Mx + q$, 其中

$$M = \begin{bmatrix} 4 & 1 & 0 & \cdots & 0 \\ -2 & 4 & 1 & \cdots & 0 \\ 0 & -2 & 4 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 4 \end{bmatrix}, \quad q = (-1, -1, \cdots, -1)^T.$$

例 2 (Kojima-Shindo Problem) 假定 $F(x) = \left(f_1(x), f_2(x), f_3(x), f_4(x)\right)^T$, 其中

$$\begin{aligned} f_1(x) &= 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6, \\ f_2(x) &= 2x_1^2 + x_1 + x_2^2 + 10x_3 + 2x_4 - 2, \\ f_3(x) &= 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 9x_4 - 9, \\ f_4(x) &= x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3. \end{aligned}$$

该问题有两个解: $x^1 = \left(\frac{\sqrt{6}}{2}, 0, 0, 0.5\right)^T$ 和 $x^2 = (1, 0, 3, 0)^T$.

利用 MATLAB2013a 编程, 实验结果见下表 3.1 所示, 其中 No.iter. 表示迭代次数. NF. 表示计算函数 F 值的次数.

表 3.1: 数值实验结果

问题	维数	初始点	文献 [15] 中的算法			算法 3.1		
			No.iter.	NF.	$\|\nabla\Psi(x)\|$	No.iter.	NF.	$\|\nabla\Psi(x)\|$
问题 1	4	(1,1,1,1)	6	6	3.2e-08	5	5	6.14e-08
	4	(1,2,3,4)	11	11	4.3e-13	4	4	4.56e-10
问题 2	4	(1,1,1,2)	7	7	2.3e-07	3	3	9.55e-12
	4	(2,2,2,2)	9	9	7.2e-08	3	3	6.38e-12
	4	(1,2,3,4)	11	11	4.3e-13	2	2	1.54e-08

从表中两种算法的迭代次数和函数 F 值的计算次数来看, 算法 3.1 比文献 [15] 中提出的算法更加有效. 例如, 表 3.1 中, 问题 1 和问题 2 经过两种算法的测试后, 算法 3.1 的迭代次数及函数 F 值的计算次数明显小于文献 [15] 中的算法.

参 考 文 献

[1] Martinet B. Determination approachée d'un point fixe d'une application pseudo-contractante [J]. Comptes Rendus de l'Académie des Sci. Ser. I Math., 1972, 274(2): 163–165.

[2] He B S, Liao L Z, Yuan X M. A LQP based interior prediction-correction method for nonlinear complementarity problems [J]. J. Comput. Math. Intern. Edi., 2006, 24(1): 33–44.

[3] Bnouhachem A. An LQP Method for pseudomonotone variational inequalities [J]. J. Glob. Optim., 2006, 36(3):351–363.

[4] Noor M A, Bnouhachem A. Modified proximal point method for nonlinear complementarity problems [J]. J. Comput. Appl. Math., 2006, 197(2): 395–405.

[5] Xu Y, He B S, Yuan X. A hybrid inexact logarithmic-quadratic proximal method for nonlinear complementarity problems [J]. J. Math. Anal. Appl., 2006, 322(1): 276–287.

[6] Yuan X M. The prediction-correction approach to nonlinear complementary problems [J]. Euro. J. Oper. Res., 2007, 176(3): 1357–1370.

[7] 陈金雄, 刘宁. 解 P0 非线性互补问题的光滑 Levenberg-Marquardt 方法 [J]. 数学杂志, 2015, 35(4): 905–916.

[8] Kanzow C, Pieper H. Jacobian smoothing methods for nonlinear complementarity problems [J]. SIAM J. Optim., 1999, 9(2): 342–373.

[9] Yang Y F, QiL. Smoothing trust region methods for nonlinear complementarity problems with P0-functions [J]. Ann. Oper. Res., 2005, 133(1-4): 99–117.

[10] Kanzow C. Some noninterior continuation methods for linear complementarity problems [J]. SIAM J. Matrix Anal. Appl., 1996, 17(4): 851–868.

[11] Qi H, Liao L. A smoothing Newton method for general nonlinear complementarity problems [J]. Comput. Optim. Appl., 2000, 17(2-3): 231–253.

[12] Yamashita N, Fukushima M. On the rate of convergence of the Levenberg–Marquardt method [J]. Comput., 2001,15:239–249.

[13] Dan H, Yamashita N, Fukushima M. Convergence properties of the inexact Levenberg–Marquardt method under local error bound conditions [J]. Optim. Meth. Soft., 2002, 17(4): 605–626.

[14] Bnouhachem A, Noor M A. A new logarithmic-quadratic proximal method for nonlinear complementarity problems [J]. Appl. Math. Comput., 2009, 215(2): 695–706.

[15] Yu H, Pu D. Smooting Levenberg-Marquardt method for general nonlinear complementarity problems under local error bound [J]. Appl. Math. Model., 2011, 35(3): 1337–1348.

[16] Clark F H. Optimization and nonsmooth analysis [M]. New York: John Wiley, 1983.

[17] Qi L. C-differentiability, C-differential operators and generalized Newton methods [J]. Sydney: Appl. Math. Report AMR96/5, Univ. New South Wales, 1996.

TWO CLASSES OF NUMERICAL METHODS FOR THE NONLINEAR COMPLEMENTARITY PROBLEM

ZHOU Guang-hui[1], ZHANG Cong-jun[2], ZHANG Cheng-hu[2], WANG Yue-hu[3]

(*1. School of Mathematical Sciences; Information College, Huaibei Normal University, Huaibei 235000, China*)

(*2.School of Applied Mathematics, Nanjing University of Finance and Economics, Nanjing 210023, China*)

(*3.School of Management Science and Engineering, Nanjing University of Finance and Economics, Nanjing 210023, China*)

Abstract: In this paper, two numerical methods are proposed for the nonlinear complementarity problems. Based on the classical LQP algorithm and the Levenberg-Marquardt algorithm, the new algorithms are designed; the convergence properties of our methods are obtained; the numerical experiments show that our new methods are better than traditional ones for some problems.

Keywords: nonlinear complementarity problem; LQP algorithm; Levenberg-Marquardt algorithm

2010 MR Subject Classification: 65K10; 49J40