

二阶锥线性互补问题的低阶罚函数算法

赵雯宇, 郝自军, 余国林

(北方民族大学数学与信息科学学院, 宁夏 银川 750021)

摘 要: 本文研究了二阶锥线性互补问题的低阶罚函数算法. 利用低阶罚函数算法将二阶锥线性互补问题转化为低阶罚函数方程组, 获得了低阶罚函数方程组的解序列在特定条件下以指数速度收敛于二阶锥线性互补问题解的结果, 推广了二阶锥线性互补问题的幂罚函数算法. 数值实验结果验证了算法的有效性.

关键词: 二阶锥; 线性互补问题; 低阶罚函数算法; 指数收敛速度

MR(2010) 主题分类号: 90C25; 90C30

中图分类号: O224; O221.2

文献标识码: A

文章编号: 0255-7797(2017)02-0427-12

1 引言

考虑二阶锥线性互补问题: 求向量 $x \in R^n$, 使得

$$x \in K, Ax - b \in K, x^T(Ax - b) = 0, \quad (1.1)$$

其中 R^n 是 n 维欧氏空间, $A \in R^{n \times n}$ 是 $n \times n$ 阶实矩阵, $b \in R^n$ 是 n 维实向量, K 是二阶锥的笛卡尔乘积. 也就是说

$$K = K^{n_1} \times K^{n_2} \times \cdots \times K^{n_m}, \quad (1.2)$$

其中 $m, n_1, \cdots, n_m \geq 1, n_1 + \cdots + n_m = n$, 且 $K^{n_i} \subset R^{n_i}$ 是 n_i 维二阶锥. 即

$$K^{n_i} = \{(x_1, x_2) \in R \times R^{n_i-1} \mid x_1 \geq \|x_2\|\}, \quad (1.3)$$

其中 $\|\cdot\|$ 表示欧几里得范数. 为了方便用 (x_1, x_2) 表示 $(x_1, x_2^T)^T$, 用 $u \succeq_{K^{n_i}} v$ (或 $v \preceq_{K^{n_i}} u$) 表示 $u - v \in K^{n_i}$. 显然 $K \subset R^n$ 是一个闭凸且自对偶锥. 若 $n_i = 1$, K^1 为非负实数集 R_+ . 若 $n_1 = \cdots = n_m = 1$, 则 K^n 退化为 R_+^n , 此时问题 (1.1) 退化为一般的线性互补问题.

二阶锥互补问题是近十多年来研究的新问题, 它是二阶锥规划的推广, 在工程设计、金融、管理科学等领域都有广泛应用^[1,2,3]. 在过去的十多年中, 对二阶锥规划和二阶锥线性互补问题已有广泛研究, 并提出了各种算法, 例如光滑牛顿算法^[4,5,6,7]、半光滑牛顿算法^[8,9]、内点算法^[10]、价值函数算法^[11]、矩阵分裂算法^[12]等, 其中有些算法需要 $O(n^3)$ 浮点运算次数, 这在问题维数 n 变得很大时可能不是有效的.

众所周知, 罚函数方法是求解约束优化问题的重要方法, 其主要思路是将约束优化问题转化为一组无约束优化问题来求解. 文献 [13] 提出了具有良好性质的 l_1 罚函数. 之后, 许

*收稿日期: 2016-01-27

接收日期: 2016-05-23

基金项目: 国家自然科学基金 (11361001; 11661002); 宁夏自然科学基金 (NZ16093); 宁夏高等学校科研项目 (NGY2016136).

作者简介: 赵雯宇 (1991-), 女, 山西吕梁, 硕士, 主要研究方向: 最优化理论与方法.

多学者致力于精确罚函数算法的研究^[14,15,16,17], 相比经典的 l_1 精确罚函数, 这些方法只需较弱的条件即可保证精确性. 2008 年, Wang 和 Yang^[18] 提出一种幂罚函数算法求解线性互补问题. 之后, Huang 和 Wang^[19,20] 将幂罚函数算法推广到求解非线性互补问题和混合互补问题. 2015 年, Hao 和 Wan^[21] 等将幂罚函数算法推广到求解二阶锥线性互补问题. 实际上, 幂罚函数算法就是一类低阶罚函数算法. 文献 [18, 21] 中罚函数的幂参数仅是形如 $1/k (k \in \mathbb{Z}^+)$ 的特殊数. 基于此, 本文将文献 [21] 中幂参数范围推广到 $(0, 1]$ 中任意实数, 提出一般低阶罚函数算法. 同时证明在矩阵 A 正定时, 低阶罚函数方程组的解序列在罚参数趋于无穷时以指数速度收敛到二阶锥线性互补问题的解, 并将低阶罚函数算法的数值结果与著名的光滑 Fischer-Burmeister (F-B) 函数算法进行比较, 结果表明提出的算法是有效的.

本文共分 5 节, 第 2 节给出一些预备知识; 第 3 节探讨二阶锥线性互补问题的低阶罚函数算法, 并证明算法的收敛性; 第 4 节给出一些数值实验结果; 第 5 节进行总结. 在本文中, 对任意 $p > 1$, $\|\cdot\|_p$ 表示 l_p 范数. 特别地, 当 $p = 2$ 时, 它代表欧几里得范数 $\|\cdot\|$.

2 预备知识

本节给出单块二阶锥 K^n 的一些基础知识. 对任意的 $x = (x_1, x_2) \in R \times R^{n-1}$, $y = (y_1, y_2) \in R \times R^{n-1}$, 它们的约当积 (Jordan product)^[1] 定义为

$$x \circ y = (x^T y, x_1 y_2 + y_1 x_2). \quad (2.1)$$

本文用 x^2 表示 $x \circ x$, 用 $x + y$ 表示普通向量加法, 这样 “ $\circ$ ”, “ $+$ ” 以及 $e = (1, 0, \dots, 0) \in R \times R^{n-1}$ 构成了 K^n 的约当代数. 易知约当积满足交换律, 但一般不满足结合律, 即 $(x \circ y) \circ z$ 与 $x \circ (y \circ z)$ 并不总相等, 这是研究和分析二阶锥优化算法的主要困难之一. 但约当积在幂意义下结合律成立, 即 $(x \circ x) \circ x = x \circ (x \circ x)$ 对任意 $x \in R^n$ 成立. 因此对任意正整数 p , x^p 有意义, 且对任意正整数 m 和 n , 有 $x^{m+n} = x^m \circ x^n$. 关于与二阶锥相关的欧几里得约当代数的其他概念, 如迹、行列式、可逆向量、平方根、绝对值向量等, 详细内容可参考文献 [1, 4, 22, 23].

与矩阵的谱分解类似, 以 K^n 为基础, R^n 中的任意向量都可以在二阶锥意义下进行谱分解^[4,22]. 对任意 $x = (x_1, x_2) \in R \times R^{n-1}$, x 的谱分解为

$$x = \lambda_1 u^{(1)} + \lambda_2 u^{(2)}, \quad (2.2)$$

$$\lambda_i = x_1 + (-1)^i \|x_2\|, \quad u^{(i)} = \begin{cases} \frac{1}{2} \left(1, (-1)^i \frac{x_2}{\|x_2\|} \right), & \text{如果 } x_2 \neq 0, \\ \frac{1}{2} (1, (-1)^i \omega), & \text{如果 } x_2 = 0, \end{cases} \quad i = 1, 2, \quad (2.3)$$

其中 λ_1, λ_2 和 $u^{(1)}, u^{(2)}$ 分别表示 x 的特征值和相应的特征向量, ω 是 R^{n-1} 中满足 $\|\omega\| = 1$ 的任意向量. 易知 $e = u^{(1)} + u^{(2)}$, $\lambda_1 \leq \lambda_2$, 且 $x_2 \neq 0$ 时分解式 (2.2) 唯一.

下面给出谱分解的一些基本性质.

命题 2.1^[4,22] 对任意的 $x = (x_1, x_2) \in R \times R^{n-1}$, λ_1, λ_2 和 $u^{(1)}, u^{(2)}$ 是按 (2.2) 和 (2.3) 式给出的 x 的特征值和相应的特征向量, 则

- (1) $u^{(1)} \circ u^{(2)} = 0$;
- (2) $\|u^{(1)}\| = \|u^{(2)}\| = \frac{1}{\sqrt{2}}$;
- (3) $u^{(i)} \circ u^{(i)} = u^{(i)}, i = 1, 2$;

(4) λ_1 是非负的 (正的) 当且仅当 $x \in K^n$ ($x \in \text{int}K^n$), 其中 $\text{int}K^n$ 表示 K^n 内部;

(5) $\text{tr}(x) = \lambda_1 + \lambda_2 = 2x_1, \det(x) = \lambda_1\lambda_2 = x_1^2 - \|x_2\|^2, 2\|x\| = \lambda_1^2 + \lambda_2^2$.

由命题 2.1, 对任意的 $x, y \in R^n$ 且 $x \in K^n, y \in K^n, \langle x, y \rangle = 0$ 当且仅当 $x \circ y = 0$. 因此互补条件可以由约当积等价描述. 根据 x 的谱分解 (2.2) 和 (2.3), 一个标量函数 $\hat{f}: R \rightarrow R$ 可以被扩展为与 K^n ($n \geq 1$) 相关的二阶锥向量值函数 [4,24]

$$f(x) = \hat{f}(\lambda_1)u^{(1)} + \hat{f}(\lambda_2)u^{(2)}, \quad (2.4)$$

其中 λ_1, λ_2 是 x 的特征值, $u^{(1)}, u^{(2)}$ 是相应的特征向量.

式 (2.4) 给出了定义二阶锥函数的一种形式, 但其含义还需进一步分析. 对任意 $x \in R^n$, x 到 K^n 的投影定义为在 K^n 上离 x 最近的点, 记作 $[x]_+$, 即 $[x]_+ \in K^n$ 且

$$\|x - [x]_+\| = \min\{\|x - y\| \mid y \in K^n\}. \quad (2.5)$$

显然, 当 $n = 1$ 时, 上述投影退化为

$$[t]_+ = \max\{0, t\} (t \in R).$$

下述命题指出, $|x|$ 和 $[x]_+$ 具有 (2.4) 式的形式 (参见文献 [4] 命题 3.3).

命题 2.2 对任意 $x = (x_1, x_2) \in R \times R^{n-1}$, λ_1, λ_2 和 $u^{(1)}, u^{(2)}$ 是按式 (2.2) 和 (2.3) 给出的 x 的特征值和相应的特征向量, 则

$$(1) |x| = (x^2)^{\frac{1}{2}} = |\lambda_1|u^{(1)} + |\lambda_2|u^{(2)},$$

$$(2) [x]_+ = (x + |x|)/2 = [\lambda_1]_+u^{(1)} + [\lambda_2]_+u^{(2)}.$$

类似于在 K^n 上投影的概念, 定义 [6]

$$[x]_- = [\lambda_1]_-u^{(1)} + [\lambda_2]_-u^{(2)}, \quad (2.6)$$

显然 $[x]_-$ 是 x 在 $-K^n$ 上投影的相反向量, 且 $x = [x]_+ - [x]_-$, $[x]_+, [x]_- \in K^n$, $[x]_+ \circ [x]_- = 0$.

3 低阶罚函数算法及其收敛性分析

本节提出二阶锥线性互补问题 (1.1) 的低阶罚函数算法, 并给出相应的收敛性分析. 不失一般性, 假设向量由一个二阶锥块组成, 即 $K = K^n$, 因为收敛性分析很容易推广到二阶锥的一般形式 (1.2).

下面考虑低阶罚函数方程组: 求 $x_\eta \in R^n$, 使得

$$Ax_\eta - \eta[-x_\eta]_+^r = b, \quad (3.1)$$

其中 $0 < r \leq 1$ 是幂参数, $\eta \geq 1$ 是罚参数. 当 $-x_\eta$ 的谱分解为 $-x_\eta = \gamma_1 v^{(1)} + \gamma_2 v^{(2)}$ 时, 定义

$$[-x_\eta]_+^r = [\gamma_1]_+^r v^{(1)} + [\gamma_2]_+^r v^{(2)}. \quad (3.2)$$

当 $x \succeq_K 0$ 不成立时, 方程组 (3.1) 中的罚项 $\eta[-x_\eta]_+^r$ 惩罚 x_η 的“负”项. 由于 $\eta[-x_\eta]_+^r \in K$, 故 $Ax_\eta - b \succeq_K 0$ 总成立. 当 $\eta \rightarrow +\infty$ 时, 我们希望低阶罚函数方程组 (3.1) 的解序列收敛于

二阶锥线性互补问题 (1.1) 的解. 因此二阶锥线性互补问题 (1.1) 被转化为低阶罚函数方程组序列. 基于此, 给出如下低阶罚函数算法.

算法 3.1 步骤 1 取任意向量 $b \in R^n$, 矩阵 $A \in R^{n \times n}$, 令 $\tilde{x} = 0 \in R^n$.

步骤 2 若 $b \preceq_K 0$, 停止, 转步骤 7; 否则, 转步骤 3.

步骤 3 若 $A^{-1}b \succeq_K 0$, 令 $\tilde{x} = A^{-1}b$, 停止, 转步骤 7; 否则, 转步骤 4.

步骤 4 设置幂参数 $0 < r \leq 1$, 罚参数 $\eta \geq 1$, 误差限为 eps, 倍数参数 $c > 1$, 选择一个初始点 $(x_{01}, x_{02}) \in R \times R^{n-1}$, 且当 $x_{01} < 0$ 时, $x_{02} \neq 0$.

步骤 5 对罚参数 η 和初始点 x_0 , 解非线性方程组

$$Ax - \eta[-x]_+^r = b, \quad (3.3)$$

假设 x_η 是 (3.3) 式的解, 令 $\text{Tol} = |x_\eta^T(Ax_\eta - b)|$.

步骤 6 如果 $\text{Tol} \leq \text{eps}$, 令 $\tilde{x} = x_\eta$, 停止, 转步骤 7; 否则, 令 $x_0 = x_\eta, \eta = c\eta$, 转步骤 5.

步骤 7 向量 $\tilde{x}$ 是二阶锥线性互补问题 (1.1) 的近似最优解.

下面讨论算法 3.1 的收敛性分析, 为此给出如下假设.

假设 3.1 矩阵 A 正定, 但不一定对称, 即存在正常数 ω , 使得

$$y^T Ay \geq \omega \|y\|^2, \forall y \in R^n. \quad (3.4)$$

众所周知, 二阶锥线性互补问题 (1.1) 与下述变分不等式等价: 求 $x^* \succeq_K 0$, 使得

$$(y - x^*)^T Ax^* \geq (y - x^*)^T b, \forall y \succeq_K 0. \quad (3.5)$$

由假设 3.1 及文献 [3] 中定理 2.3.3 知变分不等式 (3.5) 有唯一解, 因而二阶锥线性互补问题 (1.1) 亦有唯一解. 同时, 由于方程组可看作无约束变分不等式, 故 (3.1) 式也有唯一解.

下述命题 3.1 给出低阶罚函数方程组 (3.1) 解的有界性, 命题 3.2 给出 $\|[-x_\eta]_+\|$ 的一个上界, 其证明只需将文献 [21] 命题 3.1 和 3.2 中的幂参数 $1/k (k \in Z^+)$ 换为式 (3.1) 中的实数 $r \in (0, 1]$ 即可, 这里不再赘述.

命题 3.1 对任意的 $\eta \geq 1, 0 < r \leq 1$, 低阶罚函数方程组 (3.1) 的解有界, 即存在正常数 W , 独立于 x_η, η 和 r , 使得 $\|x_\eta\| \leq W$.

命题 3.2 对任意 $\eta \geq 1, 0 < r \leq 1$, 存在正常数 M , 独立于 x_η 和 η , 使得

$$\|[-x_\eta]_+\| \leq \frac{M}{\eta^{1/r}}. \quad (3.6)$$

利用命题 3.1 和命题 3.2, 可以得到如下重要结果, 它给出了算法 3.1 求解二阶锥线性互补问题 (1.1) 的收敛性分析.

定理 3.1 对任意 $\eta \geq 1, 0 < r \leq 1$, 令 x_η 是低阶罚方程组 (3.1) 的解, 且 x^* 是二阶锥线性互补问题 (1.1) 的解. 则存在独立于 x^*, x_η 和 η 的正常数 M , 使得

$$\|x^* - x_\eta\| \leq \frac{M}{\eta^{1/r}}. \quad (3.7)$$

证 利用 $-x_\eta = [-x_\eta]_+ - [-x_\eta]_-$, 向量 $x^* - x_\eta$ 可以分解为

$$x^* - x_\eta = x^* - [-x_\eta]_- + [-x_\eta]_+ = r_\eta + [-x_\eta]_+, \quad (3.8)$$

其中 $[-x_\eta]_-$ 如 (2.6) 式的形式所定义, 且

$$r_\eta = x^* - [-x_\eta]_-. \quad (3.9)$$

下面考虑 r_η 的估计值. 如果 $r_\eta = 0$, 由 (3.6) 和 (3.8) 式可知不等式 (3.7) 显然成立, 故下面只讨论 $r_\eta \neq 0$ 的情形. 显然 $[-x_\eta]_- \succeq_K 0$, 由二阶锥线性互补问题与变分不等式 (3.5) 的等价性及 (3.9) 式, 在 (3.5) 式中用 $[-x_\eta]_-$ 代替 y 可得

$$-r_\eta^T A x^* \geq -r_\eta^T b. \quad (3.10)$$

在 (3.1) 式两端同时左乘 r_η , 可得

$$r_\eta^T A x_\eta - \eta r_\eta^T [-x_\eta]_+^r = r_\eta^T b. \quad (3.11)$$

将 (3.10) 和 (3.11) 式相加得

$$r_\eta^T A(x_\eta - x^*) - \eta r_\eta^T [-x_\eta]_+^r \geq 0. \quad (3.12)$$

由 (3.2) 式知 $[-x_\eta]_+^r \succeq_K 0$, 又由于 $x^* \succeq_K 0$, 根据 (3.9) 式可得

$$\begin{aligned} r_\eta^T [-x_\eta]_+^r &= (x^* - [-x_\eta]_-)^T [-x_\eta]_+^r \\ &= (x^*)^T [-x_\eta]_+^r - [-x_\eta]_-^T [-x_\eta]_+^r \\ &= (x^*)^T [-x_\eta]_+^r \geq 0. \end{aligned} \quad (3.13)$$

将不等式 (3.13) 应用于 (3.12) 式得 $r_\eta^T A(x_\eta - x^*) \geq 0$, 即

$$r_\eta^T A(x^* - x_\eta) \leq 0. \quad (3.14)$$

由 (3.8) 式知 $r_\eta = x^* - x_\eta - [-x_\eta]_+$, 代入 (3.14) 式得

$$(x^* - x_\eta)^T A(x^* - x_\eta) \leq [-x_\eta]_+^T A(x^* - x_\eta). \quad (3.15)$$

由假设 3.1, 式 (3.15), Cauchy-Schwarz 不等式及范数相容性知, 存在常数 $b_0 > 0$, 使得

$$\begin{aligned} b_0 \|x^* - x_\eta\|^2 &\leq (x^* - x_\eta)^T A(x^* - x_\eta) \\ &\leq [-x_\eta]_+^T A(x^* - x_\eta) \\ &\leq \|[-x_\eta]_+\| \|A(x^* - x_\eta)\| \\ &\leq \|[-x_\eta]_+\| \|A\| \|x^* - x_\eta\|, \end{aligned} \quad (3.16)$$

其中 $\|A\|$ 表示矩阵 A 的 2-范数. 用 $\|x^* - x_\eta\|$ 同除 (3.16) 式两端得

$$\|x^* - x_\eta\| \leq \frac{\|A\|}{b_0} \|[-x_\eta]_+\|. \quad (3.17)$$

最后由 (3.6) 及 (3.17) 式知结论 (3.7) 成立. 证毕.

根据定理 3.1, 在假设 3.1 下, 当 $\eta \rightarrow +\infty$ 时, 低阶罚函数方程组 (3.1) 的解序列 $\{x_\eta\}$ 以指数速度收敛到二阶锥线性互补问题 (1.1) 的解. 但若假设 3.1 不成立, 又会有怎样的收敛性

结果呢? 下述定理 3.2 证明了此种情形下低阶罚函数方程组 (3.1) 解序列的任意极限点都是二阶锥线性互补问题 (1.1) 的解.

定理 3.2 对任意 η_i , 假设低阶罚函数方程组 (3.1) 有解 x_{η_i} , 对任意幂参数 $0 < r \leq 1$, 令 $\{\eta_i\}$ 是一列满足 $\eta_i \geq 1$ 的正数序列, 且 $\lim_{i \rightarrow \infty} \eta_i = +\infty$, 则序列 $\{x_{\eta_i}\}$ 的任意极限点都是二阶锥线性互补问题 (1.1) 的解.

证 对任意 $\eta_i \geq 1$, 因为 x_{η_i} 是低阶罚函数方程组 (3.1) 关于 η_i 的解, 故

$$Ax_{\eta_i} - b - \eta_i[-x_{\eta_i}]_+^r = 0. \quad (3.18)$$

设 $\bar{x}$ 是序列 $\{x_{\eta_i}\}$ 的任意极限点, 为了方便, 不妨假设当 $i \rightarrow \infty$ 时 $x_{\eta_i} \rightarrow \bar{x}$, 否则用 $\{x_{\eta_i}\}$ 的一个收敛到 $\bar{x}$ 子序列代替 $\{x_{\eta_i}\}$ 即可. 因而存在正数 P , 使得

$$\|x_{\eta_i}\| \leq P, i = 1, 2, \dots \quad (3.19)$$

注意到

$$\lim_{i \rightarrow \infty} x_{\eta_i} = \bar{x}, \quad (3.20)$$

从式 (3.18) 得 $Ax_{\eta_i} - b = \eta_i[-x_{\eta_i}]_+^r \in K$, 在此式中令 $i \rightarrow \infty$, 由 K 的闭性知

$$A\bar{x} - b = \lim_{i \rightarrow \infty} Ax_{\eta_i} - b \in K. \quad (3.21)$$

下面证明

$$\bar{x} \in K. \quad (3.22)$$

反设 $\bar{x} \notin K$, 此时令 $\varepsilon := \|[-\bar{x}]_+^r\|$, 则 $\varepsilon > 0$. 故由 (3.20) 式知当 i 充分大时, $x_{\eta_i} \notin K$ 且 $\|[x_{\eta_i}]_+^r\| \geq \frac{1}{2}\varepsilon > 0$. 同时, 由 (3.18) 式得当 $i \rightarrow \infty$ 时,

$$\|Ax_{\eta_i} - b\| = \|\eta_i[-x_{\eta_i}]_+^r\| \geq \frac{1}{2}\eta_i\varepsilon \rightarrow \infty,$$

这与式 (3.19) 矛盾, 故式 (3.22) 成立.

下面再证明

$$\bar{x}^T(A\bar{x} - b) = 0. \quad (3.23)$$

因为 $\bar{x} \in K$, 所以有以下三种情形.

情形 1 如果 $\bar{x} = 0$, 显然 $\bar{x}^T(A\bar{x} - b) = 0$.

情形 2 如果 $\bar{x} \in \text{int}K$, 则由 (3.20) 式知当 i 充分大时, $[-x_{\eta_i}]_+^r = 0$. 由 (3.18) 式知

$$Ax_{\eta_i} - b = \eta_i[-x_{\eta_i}]_+^r,$$

在此式中令 $i \rightarrow \infty$, 则 $A\bar{x} - b = 0$, 所以 $\bar{x}^T(A\bar{x} - b) = 0$.

情形 3 如果 $\bar{x} \in \text{bd}K \setminus \{0\}$, 其中 $\text{bd}K$ 表示 K 的边界, 此时令 $\bar{x} = (z_1, z_2) \in R \times R^{n-1}$, 则 $z_1 = \|z_2\| > 0$, 根据向量的谱分解 (2.2)–(2.3), $\bar{x}$ 可以分解为 $\bar{x} = \mu_1 v^{(1)} + \mu_2 v^{(2)}$, 这里 μ_1, μ_2 和 $v^{(1)}, v^{(2)}$ 分别表示 $\bar{x}$ 的特征值和相应的特征向量. 因此

$$\mu_1 = z_1 - \|z_2\| = 0, \mu_2 = z_1 + \|z_2\| > 0, v^{(1)} = \left(\frac{1}{2}, \frac{-z_2}{2\|z_2\|}\right), v^{(2)} = \left(\frac{1}{2}, \frac{z_2}{2\|z_2\|}\right),$$

即 $\bar{x} = \mu_2 v^{(2)}$. 假设 x_{η_i} 的谱分解为

$$x_{\eta_i} = \mu_1(i) v^{(1)}(i) + \mu_2(i) v^{(2)}(i),$$

由于对任意 $y = (y_1, y_2) \in R \times R^{n-1}$, 当 $\|y_2\| > 0$ 时,

$$f_{1,2}(y) = y_1 \pm \|y_2\|, \quad f_{3,4}(y) = \left(\frac{1}{2}, \frac{\pm y_2}{2\|y_2\|}\right)$$

都连续, 故由 (3.20) 式知当 $i \rightarrow \infty$ 时,

$$\mu_1(i) \rightarrow \mu_1 = 0, \quad \mu_2(i) \rightarrow \mu_2 > 0, \quad v^{(1)}(i) \rightarrow v^{(1)}, \quad v^{(2)}(i) \rightarrow v^{(2)}.$$

特别地, 当 i 充分大时有 $\mu_2(i) \geq \frac{1}{2}\mu_2 > 0$, 此时由 (3.18) 式得

$$\begin{aligned} Ax_{\eta_i} - b &= \eta_i [-x_{\eta_i}]_+^r \\ &= \eta_i ([-\mu_1(i)]_+^r v^{(1)}(i) + [-\mu_2(i)]_+^r v^{(2)}(i)) \\ &= \eta_i [-\mu_1(i)]_+^r v^{(1)}(i). \end{aligned} \quad (3.24)$$

在 (3.24) 式两边令 $i \rightarrow \infty$, 并记 $\lim_{i \rightarrow \infty} (\eta_i [-\mu_1(i)]_+^r) = c \geq 0$, 则 $A\bar{x} - b = cv^{(1)}$. 因此

$$\bar{x}^T (A\bar{x} - b) = \mu_2 c (v^{(2)})^T v^{(1)} = 0.$$

由以上 3 种情形知式 (3.23) 成立.

最后, 由 (3.21)–(3.23) 式可得 $\bar{x}$ 是二阶锥线性互补问题 (1.1) 的解. 由极限点 $\bar{x}$ 的任意性, 可知结论成立. 证毕.

4 数值实验

本节用一些数值算例来检验算法 3.1 的有效性, 所有例子均采用离散牛顿法求解其对应于 (3.1) 式的非线性方程组, 所有的数值实验运行环境为 Intel(R) Core(TM) i5-3570K CPU 3.40GHz, 4G of RAM, and windows7 with MATLAB R2009a.

例 4.1 考虑在 K^2 上的二阶锥线性互补问题 (1.1), 其数据如下

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 4 \end{pmatrix}.$$

例 4.1 的精确解为 $x^* = (1, 1)^T$. 下面考虑用算法 3.1 求解, 首先取幂参数 $r = 1$, 初始点 $x_0 = (-1, 1)^T$, 取罚参数 $\eta = 10 \times 2^i, i = 2, \dots, 7$, 其数值实验结果如表 4.1 所示, 其中 NS 代表数值解, Err 代表数值解 NS 与 x^* 的距离. 事实上, 本例对应的低阶罚函数方程组 (3.1) 可以利用分析方法求解, 且 $x_\eta = ((\eta - 2)/(1 + \eta), (\eta + 2)/(1 + \eta))$, 容易得到

$$\|x^* - x_\eta\| = \sqrt{10}/(1 + \eta) \leq \sqrt{10}/\eta.$$

因此当 $\eta \rightarrow +\infty$ 时, x_η 以指数速度 $O(\eta^{-1})$ 收敛到 x^* , 这与式 (3.7) 的结论一致. 由表 4.1 可以看出数值实验结果和分析结果是一致的.

表 4.1: 例 4.1 的数值实验结果 ($r = 1$)

η	40	80	160	320	640	1280
NS	$(0.9628, 1)^T$	$(0.9630, 1)^T$	$(0.9814, 1)^T$	$(0.9905, 1)^T$	$(0.9953, 1)^T$	$(0.9977, 1)^T$
Err	0.0771	0.0390	0.0196	0.0099	0.0049	0.0025

表 4.2: 例 4.1 的数值实验结果的误差

η	20	40	80	160	320
$r = \frac{3}{5}$	0.052584	0.016882	0.005350	0.001689	5.3260e-3
$r = \frac{2}{5}$	0.013988	0.002496	4.4280e-4	7.9136e-5	1.4817e-5
$r = \frac{\sqrt{2}}{5}$	2.6530e-3	2.3218e-4	2.2111e-4	2.8069e-6	7.1616e-7

其次令 $\eta = 10 \times 2^i, i = 1, \dots, 5$, $x_0 = (-1, 1)^T$, 且 $r = \frac{3}{5}, \frac{2}{5}, \frac{\sqrt{2}}{5}$, 其数值解的误差结果如表 4.2 所示. 从表 4.2 可以看出对于 $0 < r \leq 1$ 内一个固定的值, 当 η 增大时, 误差 $\|x^* - x_\eta\|$ 趋于 0. 对固定的 $\eta > 1$, 随着 r 的减小, 误差 $\|x^* - x_\eta\|$ 逐渐趋于 0. 这些与式 (3.7) 的结论一致.

下面考虑 Chen 和 Ma^[25,26] 已经测试过的两个例子. 在测试中令 $r = \frac{\sqrt{3}}{4}, c = 10$ 且初始 $\eta = 1000$. 例 4.2 和例 4.3 的终止准则分别设为 $\text{eps} = 10^{-8}$ 和 $\text{eps} = 10^{-7}$.

例 4.2 考虑在 K^5 上的二阶锥线性互补问题 (1.1), 其数据如下

$$A = \begin{pmatrix} 15 & -5 & -1 & 4 & -5 \\ 0 & 5 & 0 & 0 & 1 \\ -1 & -3 & 8 & 2 & -3 \\ 2 & -4 & 2 & 9 & -4 \\ 0 & -5 & 0 & 0 & 10 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

本例的解 $x^* \approx (0.049185, -0.0030997, 0.0096024, 0.0031883, 0.048033)^T$ ^[25], 且矩阵 A 满足假设 3.1. 选取不同的初始点, 测试结果如表 4.3 所示.

表 4.3: 例 4.2 的数值实验结果

初始点	$ \tilde{x}^T(A\tilde{x} - b) $	迭代次数	CPU 时间
$(10^6, \dots, 10^6)$	1.7902e-9	2	2.558
$(10^3, \dots, 10^3)$	1.8076e-9	2	3.276
$(10, \dots, 10)$	1.7953e-9	2	2.652
$(-10, \dots, -10)$	1.7353e-9	2	2.558
$(1, 1, 1, 1, 1)$	1.7679e-9	2	3.010
$(-1, \dots, -1)$	1.8141e-9	2	2.995

例 4.3 考虑在 K^3 上的二阶锥线性互补问题 (1.1), 其数据如下

$$A = \begin{pmatrix} 21 & -9 & 18 \\ -9 & 4 & -7 \\ 18 & -7 & 19 \end{pmatrix}, \quad b = \begin{pmatrix} -3 \\ -7 \\ -1 \end{pmatrix}.$$

本例的解 $x^* \approx (0.183606, -0.154346, -0.099440)^T$ ^[25], A 是对称半正定矩阵. 选取不同初始点, 测试结果如表 4.4 所示.

表 4.4: 例 4.3 的数值实验结果

初始点	$ \tilde{x}^T(A\tilde{x} - b) $	迭代次数	CPU 时间
$(10^6, 10^6, 10^6)$	5.3108e-8	3	2.683
$(10^3, 10^3, 10^3)$	7.1277e-8	3	1.794
$(10, 10, 10)$	5.5847e-8	3	2.059
$(-10, -10, -10)$	7.2508e-8	2	2.293
$(1, 1, 1)$	5.5982e-8	3	2.075
$(-1, -1, -1)$	7.3163e-8	3	2.153

从表 4.3 和 4.4 可以看出, 例 4.2 和例 4.3 将文献 [21] 中形如 $1/k (k \in \mathbb{Z}^+)$ 的幂参数扩展为 $(0, 1]$ 中的一般实数. 而且取幂参数 $r = \frac{\sqrt{3}}{4}$ 时, 其数值精确度比文献 [21] 中 $r = \frac{1}{2}$ 时的情形有明显改进.

众所周知, 当 $K = K^n$ 时, F-B 函数定义为 $\phi_{FB}(x, y) = x + y - \sqrt{x^2 + y^2}$, 满足

$$\phi_{FB}(x, y) = 0 \Leftrightarrow x \in K, y \in K, \langle x, y \rangle = 0.$$

二阶锥互补问题的光滑 F-B 函数 $\phi(\mu, x, y) : R \times R^n \times R^n \rightarrow R^n$ 为^[4]

$$\phi(\mu, x, y) = x + y - \sqrt{x^2 + y^2 + 2\mu^2 e}. \quad (4.1)$$

上述平方和平方根都是指约当积意义下的平方和平方根. 将 $y = Ax - b$ 代入 (4.1) 式, 则二阶锥线性互补问题转化为光滑 F-B 方程, 即求 $x \in R^n$, 使得

$$x + (Ax - b) - \sqrt{x^2 + (Ax - b)^2 + 2\mu^2 e} = 0, \quad (4.2)$$

其中 $\mu > 0$ 是光滑参数.

众所周知, 当 $\mu \downarrow 0$ 时, 方程组 (4.2) 的解序列趋向于二阶锥线性互补问题 (1.1) 的解^[4]. 下面给出相应的光滑 F-B 算法.

光滑 F-B 算法

步骤 1 取任意向量 $b \in R^n$, 矩阵 $A \in R^{n \times n}$, 令 $\tilde{x} = 0$ 且 $\tilde{x} \in R^n$.

步骤 2 如果 $b \preceq_K 0$, 停止, 转步骤 7; 否则, 转步骤 3.

步骤 3 如果 $A^{-1}b \succeq_K 0$, 令 $\tilde{x} = A^{-1}b$, 停止, 转步骤 7; 否则, 转步骤 4.

步骤 4 取光滑参数 $\mu > 0$, 误差限为 eps, 倍数参数 $0 < d < 1$, 初始点 x_0 .

步骤 5 对光滑参数 μ 和初始点 x_0 , 解非线性方程组 (4.2), 假设 x_μ 是 (4.2) 式的解, 令 $\text{Tol} = |x_\mu^T(Ax_\mu - b)|$.

步骤 6 如果 $\text{Tol} \leq \text{eps}$, 令 $\tilde{x} = x_\mu$, 停止, 转步骤 7; 否则, 令 $x_0 = x_\mu, \mu = d\mu$, 转步骤 5.

步骤 7 向量 $\tilde{x}$ 是二阶锥线性互补问题 (1.1) 的近似最优解.

例 4.4 用随机产生的二阶锥线性互补问题来比较算法 3.1 和光滑 F-B 算法的数值效果, 为了方便, 例 4.4 中随机产生的矩阵取对称正定矩阵.

例 4.4 考虑 K 为 100 个块的二阶锥线性互补问题, 即 $K = K^{n_1} \times K^{n_2} \times \cdots \times K^{n_{100}}$, 且 $A = \text{diag}(A_1, \cdots, A_{100})$, $b = (b_1, \cdots, b_{100})$. 每一个块矩阵 A_i ($i = 1, \cdots, 100$) 是随机产生的对称正定矩阵. 每一个向量块 b_i ($i = 1, \cdots, 100$) 满足 $q_i^T(Aq_i - b_i) = 0$, 其中 $q_i \in K^{n_i}$ 是随机产生的.

本例中 $A = \text{diag}(A_1, \cdots, A_{100})$ 是块对角矩阵, $q = (q_1, \cdots, q_{100})$ 是构造的精确最优解. 首先固定 $r = \frac{\sqrt{2}}{5}$, 取 $n_i = 2, 3, 4, 5$ ($i = 1, \cdots, 100$), 在算法 3.1 中令 $c = 10$, 初始 $\eta = 10^3$, 而光滑 F-B 算法中令 $d = 0.1$, 初始 $\mu = 0.001$, 其所能达到的数值精度如表 4.5 所示, 其中 $m\text{-Val}$ 代表 $\max_i \{|\tilde{x}_i^T(A\tilde{x}_i - b_i)|\}$, $a\text{-Val}$ 代表 $(\sum_{i=1}^{100} |\tilde{x}_i^T(A\tilde{x}_i - b_i)|)/100$, $m\text{-Err}$ 代表 $\max_i \{\|\tilde{x}_i - q_i\|\}$, $a\text{-Err}$ 代表 $(\sum_{i=1}^{100} \|\tilde{x}_i - q_i\|)/100$, $\tilde{x} = (\tilde{x}_1, \cdots, \tilde{x}_{100})$ 表示数值最优解.

表 4.5: 例 4.4 中不同规模情形的数值结果 ($k = \frac{\sqrt{2}}{5}$)

问题	算法 3.1					光滑 F-B 算法				
	$m\text{-Val}$	$a\text{-Val}$	$m\text{-Err}$	$a\text{-Err}$	CPU	$M\text{-Val}$	$a\text{-Val}$	$m\text{-Err}$	$a\text{-Err}$	CPU
200 维	5.56e-8	9.33e-9	9.89e-8	1.48e-8	23.46	1.00e-6	4.06e-7	8.08e-6	4.96e-7	20.54
300 维	2.80e-7	2.33e-8	5.67e-7	5.12e-8	122.91	1.00e-6	4.16e-7	2.15e-6	4.24e-7	53.26
400 维	7.68e-8	2.13e-8	2.86e-7	5.11e-8	130.19	1.00e-6	3.76e-7	1.22e-6	3.40e-7	163.40
500 维	8.72e-8	1.90e-8	2.70e-6	4.44e-8	195.75	1.00e-6	4.16e-7	1.73e-6	4.03e-7	552.13

其次固定 $n_i = 8$, 且 $r = \frac{\sqrt{3}}{2}, \frac{\sqrt{2}}{3}, \frac{3}{10}$, 在算法 3.1 中令 $c = 10$, 初始 $\eta = 10^3$, 而光滑 F-B 方法中令 $d = 0.1$, 初始 $\mu = 0.001$. 取共同的终止准则为 $\text{eps} = 10^{-6}$, 则数值结果如表 4.6 所示, 其中 $m\text{-Iter}$ 和 $a\text{-Iter}$ 分别代表最大迭代次数和平均迭代次数.

表 4.6: 例 4.4 的数值结果 ($n_i = 8$)

r	$m\text{-Val}$	$a\text{-Val}$	$m\text{-Err}$	$a\text{-Err}$	$m\text{-Iter}$	$a\text{-Iter}$	CPU
$\frac{\sqrt{3}}{2}$	8.8456e-7	8.8174e-7	3.4238e-6	1.8684e-6	3	3	946.13
$\frac{\sqrt{2}}{3}$	4.3339e-7	4.1474e-7	1.2375e-6	8.6969e-7	1	1	413.21
$\frac{3}{10}$	1.9078e-7	2.9165e-8	2.7764e-7	5.5423e-8	1	1	324.79
光滑 F-B 算法	1.0000e-6	9.5400e-7	1.3578e-6	8.8257e-7	2	1.51	4623.68

由表 4.5 可知两种算法对例 4.4 都是可行的, 但算法 3.1 的精度明显优于光滑 F-B 算法; 而表 4.6 说明, 算法 3.1 与光滑 F-B 算法的结果相当, 但算法 3.1 所需时间明显少于光滑 F-B 算法. 由表 4.6 还可看出, 算法 3.1 对不同参数 r 所得结果也不同, 但一般来说, 参数较小时数值结果更优, 这与结论 (3.7) 一致.

5 结论

本文将文献 [21] 中二阶锥线性互补问题的幂罚函数算法推广到一般的低阶罚函数算法. 当矩阵 A 正定时, 低阶罚函数方程组的解序列以指数速度收敛到二阶锥线性互补问题的解. 当矩阵 A 非正定时, 低阶罚函数方程组解序列的任意极限点是二阶锥线性互补问题的解, 但收敛速度未知. 数值实验结果表明, 在矩阵正定条件下, 低阶罚函数算法的数值精度和计算时间优于光滑 F-B 算法, 而且在低阶罚函数算法中, 选取较小的幂参数 r , 数值结果更优.

参 考 文 献

- [1] Alizadeh F, Goldfarb D. Second-order cone programming[J]. *Math. Prog.*, 2003, 95(1): 3–51.
- [2] Lobo M S, Vandenberghe L, Boyd S, et al. Applications of second-order cone programming[J]. *Linear Alg. Appl.*, 1998, 284(1): 193–228.
- [3] Facchinei F, Pang J S. Finite-dimensional variational inequalities and complementarity problems[M]. New York: Springer-Verlag, 2003.
- [4] Fukushima M, Luo Z Q, Tseng P. Smoothing functions for second-order-cone complementarity problems[J]. *SIAM J. Optim.*, 2002, 12(2): 436–460.
- [5] Huang Z H, Ni T. Smoothing algorithms for complementarity problems over symmetric cones[J]. *Comput. Optim. Appl.*, 2010, 45(3): 557–579.
- [6] Chen X D, Sun D, Sun J. Complementarity functions and numerical experiments on some smoothing Newton methods for second-order-cone complementarity problems[J]. *Comput. Optim. Appl.*, 2003, 25(1-3): 39–56.
- [7] 董丽, 王洪芹, 潘虹. 一个求解二阶锥规划的光滑牛顿算法 [J]. *数学杂志*, 2015, 35(6): 1453–1460.
- [8] Kanzow C, Ferenczi I, Fukushima M. On the local convergence of semismooth Newton methods for linear and nonlinear second-order cone programs without strict complementarity[J]. *SIAM J. Optim.*, 2009, 20(1): 297–320.
- [9] Pan S, Chen J S. A damped Gauss-Newton method for the second-order cone complementarity problem[J]. *Appl. Math. Optim.*, 2009, 59(3): 293–318.
- [10] Monteiro R D C, Tsuchiya T. Polynomial convergence of primal-dual algorithms for the second-order cone program based on the MZ-family of directions[J]. *Math. Prog.*, 2000, 88(1): 61–83.
- [11] Chen J S. Two classes of merit functions for the second-order cone complementarity problem[J]. *Math. Meth. Oper. Res.*, 2006, 64(3): 495–519.
- [12] Hayashi S, Yamaguchi T, Yamashita N, et al. A matrix-splitting method for symmetric affine second-order cone complementarity problems[J]. *J. Comput. Appl. Math.*, 2005, 175(2): 335–353.
- [13] Zangwill W I. Non-linear programming via penalty functions[J]. *Manag. Sci.*, 1967, 13(5): 344–358.
- [14] Luo Z Q, Pang J S, Ralph D. Mathematical programs with equilibrium constraints[M]. Cambridge: Cambridge Univ. Press, 1996.
- [15] Han S P, Mangasarian O L. Exact penalty functions in nonlinear programming[J]. *Math. Prog.*, 1979, 17(1): 251–269.
- [16] Di Pillo G, Grippo L. An exact penalty function method with global convergence properties for nonlinear programming problems[J]. *Math. Prog.*, 1986, 36(1): 1–18.
- [17] Bertsekas D P, Nedić A, Ozdaglar A E. Convex analysis and optimization[M]. Massachusetts: Athena Sci. Belmont, 2003.
- [18] Wang S, Yang X. A power penalty method for linear complementarity problems[J]. *Oper. Res. Lett.*, 2008, 36(2): 211–214.
- [19] Huang C, Wang S. A power penalty approach to a nonlinear complementarity problem[J]. *Oper. Res. Lett.*, 2010, 38(1): 72–76.
- [20] Huang C, Wang S. A penalty method for a mixed nonlinear complementarity problem[J]. *Nonl. Anal.: The. Meth. Appl.*, 2012, 75(2): 588–597.
- [21] Hao Z, Wan Z, Chi X. A power penalty method for second-order cone linear complementarity problems[J]. *Oper. Res. Lett.*, 2015, 43(2), 137–142.
- [22] Chen J S. The convex and monotone functions associated with second-order cone[J]. *Optim.*, 2006, 55(4): 363–385.

- [23] Faraut J, Korányi A. Analysis on symmetric cones[M]. Oxford: Oxford Univ. Press, 1994.
- [24] Chen J S, Chen X, Tseng P. Analysis of nonsmooth vector-valued functions associated with second-order cones[J]. Math. Prog., 2004, 101(1): 95–117.
- [25] Ma C. A regularized smoothing Newton method for solving the symmetric cone complementarity problem[J]. Math. Comp. Model., 2011, 54(9): 2515–2527.
- [26] Chen L, Ma C. A modified smoothing and regularized Newton method for monotone second-order cone complementarity problems[J]. Comp. Math. Appl., 2011, 61(5): 1407–1418.

A LOWER ORDER PENALTY METHOD FOR SECOND-ORDER CONE LINEAR COMPLEMENTARITY PROBLEMS

ZHAO Wen-yu, HAO Zi-jun, YU Guo-lin

*(School of Mathematics and Information Science, Beifang University of Nationalities,
Yinchuan 750021, China)*

Abstract: In this paper, a lower order penalty method for solving the second-order cone linear complementarity problems is proposed. By this method, the second-order cone linear complementarity problem is transformed into lower order penalty equations. We prove that the solution sequence of the lower order penalty equations converges to the solution of the second-order cone linear complementarity problems at an exponential rate under a mild assumption, which extend the power penalty method for solving this problem. Numerical results demonstrate that our method is efficient.

Keywords: second-order cone; linear complementarity problem; low order penalty method; exponential convergence rate

2010 MR Subject Classification: 90C25; 90C30