Vol. 42 (2022)
No. 5

数 学 杂 志
J. of Math. (PRC)

一个求解广义圆锥互补问题的无导数光滑算法

邵灿燃, 汤京永
(信阳师范学院 数学与统计学院, 河南 信阳 464000)

摘 要: 本文研究了一个求解广义圆锥互补问题的无导数光滑算法. 利用光滑函数将广义圆锥互补问题等价转化成一个光滑方程组, 然后再利用牛顿法求解此方程组. 该算法采用了一种新的非单调无导数线搜索技术, 并且在适当条件下具有全局和局部二次收敛性质. 数值实验结果表明算法是非常有效的.

关键词: 广义圆锥互补问题; 光滑算法; 无导数线搜索; 二次收敛
MR(2010) **主题分类号:** 90C30; 65K05 **中图分类号:** O221.1
文献标识码: A **文章编号:** 0255-7797(2022)05-0461-10

1 引言

近年来, 设计求解圆锥优化问题的数值算法已成为优化领域的一个研究热点. 比如, 文献 [1–3] 研究了求解圆锥规划和圆锥互补问题的内点算法, 文献 [4–6] 研究了求解圆锥规划和圆锥互补问题的光滑算法, 文献 [7] 研究了求解圆锥上绝对值方程的光滑算法. 本文用 L_θ^m 表示 m 维圆锥, 其定义如下:

$$L_\theta^m := \{x = (x_1, \bar{x}) \in R \times R^{m-1} |\ \|x\|\cos\theta \le x_1\},$$

这里 $\theta \in (0, \frac{\pi}{2})$ 是一个旋转角，$\|\cdot\|$ 表示欧几里得范数. 当 $\theta = \frac{\pi}{4}$ 时, 圆锥 L_θ^m 即为如下二阶锥 K^m

$$K^m := \{x = (x_1, \bar{x}) \in R \times R^{m-1} |\ \|\bar{x}\| \le x_1\}.$$

由文献 [8] 的定理 2.1 可知, L_θ^m 的对偶锥为

$$(L_\theta^m)^* = L_{\frac{\pi}{2}-\theta}^m = \{x = (x_1, \bar{x}) \in R \times R^{m-1} |\ \|x\|\sin\theta \le x_1\}.$$

因此, 当 $\theta \ne \frac{\pi}{4}$ 时, 圆锥 L_θ^m 不是自对偶的, 因此它是一种非对称锥.

在本文中, 我们考虑广义圆锥互补问题, 其数学模型为: 求解 $(x, y, t) \in R^n \times R^n \times R^\ell$ 使得

$$x \in L_\theta,\ y \in L_\theta^*,\ \langle x, y\rangle = 0,\ F(x, y, t) = 0, \tag{1.1}$$

其中 $x = (x_1, ..., x_r) \in R^n$, $y = (y_1, ..., y_r) \in R^n$, $x_i, y_i \in R^{n_i}$, $\langle\cdot,\cdot\rangle$ 表示欧几里得内积, $F : R^n \times R^n \times R^\ell \to R^{n+\ell}$ 是一个连续可微函数, $L_\theta \subset R^n$ 和 $L_\theta^* \subset R^n$ 分别是一些圆锥和它们的对偶锥的笛卡尔积, 即

$$L_\theta = L_\theta^{n_1} \times \cdots \times L_\theta^{n_r},\ L_\theta^* = (L_\theta^{n_1})^* \times \cdots \times (L_\theta^{n_r})^*,$$

*收稿日期: 2021-04-15 接收日期: 2021-05-31
基金项目: 河南省自然科学基金资助 (222300420520); 河南省高等学校重点科研项目基金资助 (22A110020).
作者简介: 邵灿燃 (1997–), 女, 河南信阳, 硕士, 主要研究方向: 最优化理论与算法.

其中 $r, n_1, ..., n_r \geq 1$ 且 $n = \sum_{i=1}^{r} n_i$. 广义圆锥互补问题是一类非常值得研究的数学模型. 一方面是因为当 $\theta \neq \frac{\pi}{4}$ 时, 该问题是一类非对称锥优化问题, 而目前对于非对称锥优化问题的研究成果还不是很多. 另一方面，广义圆锥互补问题是一类非常广泛的互补系统, 其包含文献 [1–6] 中研究的圆锥互补问题以及圆锥规划问题的最优性条件. 此外, 当 $\theta = \frac{\pi}{4}$ 时, 广义圆锥互补问题即为文献 [9,10] 中研究的二阶锥互补问题.

在本文中, 我们给出了一个新的求解广义圆锥互补问题的光滑型算法. 与文献 [4–6] 中研究的光滑牛顿法不同, 我们的算法采用了一种新的非单调无导数线搜索技术. 在适当条件下, 我们证明了算法有下面三个收敛性质: (a) 算法产生的迭代序列的任何聚点都是问题 (1.1) 的解. (b) 如果迭代序列有一个孤立的聚点, 则整个迭代序列会收敛到该聚点. (c) 在雅可比非奇异假设条件下, 整个迭代序列局部二次收敛于问题 (1.1) 的一个解. 数值实验结果表明我们的算法是非常有效的.

2 预备知识

首先, 我们简单介绍与二阶锥 $K^m := \{x = (x_1, \bar{x}) \in R \times R^{m-1} |\ \|\bar{x}\| \leq x_1\}$ 相关的欧几里得约当代数. 对于任意的 $x = (x_1, \bar{x}) \in R \times R^{m-1}$, $y = (y_1, \bar{y}) \in R \times R^{m-1}$, 定义约当积为

$$x \circ y = (x^T y,\ x_1 \bar{y} + y_1 \bar{x}),$$

其单位元是 $e_m := (1, 0, ..., 0)^T \in R^m$. 给定 $x = (x_1, \bar{x}) \in R \times R^{m-1}$, 定义对称矩阵

$$L_x := \begin{bmatrix} x_1 & \bar{x}^T \\ \bar{x} & x_1 I_{m-1} \end{bmatrix},$$

这里 $I_{m-1} \in R^{(m-1)\times(m-1)}$ 为单位矩阵. 易知对任意的 $x, y \in R^m$ 都有 $L_x y = x \circ y$.

对于任意 $x = (x_1, \bar{x}) \in R \times R^{m-1}$, 其谱分解为

$$x = \lambda_1(x) c_1(x) + \lambda_2(x) c_2(x),$$

其中

$$\lambda_i(x) = x_1 + (-1)^i \|\bar{x}\|,\ \ c_i(x) = \begin{cases} \frac{1}{2}\left(1, (-1)^i \frac{\bar{x}}{\|\bar{x}\|}\right), & \bar{x} \neq 0, \\ \frac{1}{2}\left(1, (-1)^i \omega\right), & \bar{x} = 0, \end{cases} \quad i = 1, 2,$$

为特征值及其相关的特征向量, 这里 $\omega \in R^{m-1}$ 是满足 $\|\omega\| = 1$ 的任意向量. 利用特征值分解, 我们定义

$$x^2 := \lambda_1(x)^2 c_1(x) + \lambda_2(x)^2 c_2(x).$$

此外, 如果 $x \in K^m$, 那么 $\lambda_2(x) \geq \lambda_1(x) \geq 0$, 故我们定义

$$\sqrt{x} := \sqrt{\lambda_1(x)} c_1(x) + \sqrt{\lambda_2(x)} c_2(x).$$

易知 $x^2 = x \circ x$, $x = \sqrt{x} \circ \sqrt{x}$.

接下来, 我们将广义圆锥互补问题等价转化成一个光滑非线性方程组. 为此, 对于任意 $\theta \in (0, \frac{\pi}{2})$, 我们定义矩阵

$$A_m := \begin{bmatrix} \tan\theta & 0 \\ 0 & I_{m-1} \end{bmatrix} \in R^{m \times m}.$$

显然 A_m 是正定的, 它的逆矩阵是

$$A_m^{-1} := \begin{bmatrix} \mathrm{ctan}\theta & 0 \\ 0 & I_{m-1} \end{bmatrix} \in R^{m\times m}.$$

由文献 [8] 中的定理 2.1 可知

$$L_\theta^m = A_m^{-1}K^m, \quad A_m L_\theta^m = K^m.$$

本文采用如下光滑函数:

$$\psi_{A_m}(\mu,a,b) = A_m a + A_m^{-1}b - \sqrt{(A_m a)^2 + (A_m^{-1}b)^2 + 2\mu^2 e_m},$$

这里 $e_m := (1,0,...,0)^T \in R^m$. 下面引理给出了 ψ_{A_m} 的一些性质, 其具体证明可参见文献 [11] 中的定理 1 和定理 2.

引理 2.1 (i) ψ_{A_m} 在任意的点 $(\mu,a,b) \in R_{++} \times R^m \times R^m$ 处是连续可微的, 并且有

$$(\psi_{A_m})'_\mu = -2\mu L_w^{-1}e_m, \ \ (\psi_{A_m})'_a = (I_m - L_w^{-1}L_{A_m a})A_m, \ \ (\psi_{A_m})'_b = (I_m - L_w^{-1}L_{A_m^{-1}b})A_m^{-1},$$

其中 $w := \sqrt{(A_m a)^2 + (A_m^{-1}b)^2 + 4\mu^2 e_m}$.
(ii) ψ_{A_m} 满足互补性质, 即 $\psi_{A_m}(0,a,b) = 0 \Longleftrightarrow a \in L_\theta^m,\ b \in (L_\theta^m)^*,\ \langle a,b\rangle = 0$.

令 $z := (\mu,x,y,t) \in R\times R^n \times R^n \times R^\ell$. 利用 ψ_{A_m}, 我们定义函数 $H: R^{1+2n+\ell} \to R^{1+2n+\ell}$ 为

$$H(z) := \begin{pmatrix} \mu \\ F(x,y,t) \\ \psi_{A_{n_1}}(\mu,x_1,y_1) \\ \vdots \\ \psi_{A_{n_r}}(\mu,x_r,y_r) \end{pmatrix}.$$

由文献 [11] 中的命题 1 可知

$$x \in\ L_\theta,\ y \in\ L_\theta^*,\ \langle x,y\rangle = 0 \Longleftrightarrow x_i \in\ L_\theta^{n_i},\ y_i \in (\ L_\theta^{n_i})^*,\ \langle x_i,y_i\rangle = 0\ (i=1,...,r),$$

故由引理 2.1 可知 $H(z)$ 在任意点 $z \in R_{++} \times R^n \times R^n \times R^\ell$ 处连续可微, 并且满足

$$H(z) = 0 \Longleftrightarrow \mu = 0 \text{ 且 } (x,y) \text{ 是广义圆锥互补问题的解}.$$

为确保 H 的雅克比矩阵可逆, 本文作如下假设.

假设 2.1 秩 $(F_t'(x,y,t)) = \ell$, 并且对于任意的 $\Lambda \in R^\ell$, $u = (u_1,...,u_r) \in R^{n_1} \times \cdots \times R^{n_r}$, $v = (v_1,...,v_r) \in R^{n_1} \times \cdots \times R^{n_r}$, $(u,v) \neq 0$ 都满足

$$F'(x,y,t)(u,v,\Lambda) = 0 \Longrightarrow \text{ 存在 } i_0 \in \{1,...,r\} \text{ 使得 } (u_{i_0},v_{i_0}) \neq 0 \text{ 且 } \langle u_{i_0},v_{i_0}\rangle \geq 0.$$

假设 2.1 已被广泛应用于分析求解各种优化问题的光滑牛顿法, 比如文献 [9,10]. 在假设 2.1 条件下，我们有如下引理, 其具体证明可参见文献 [11] 中的定理 3.

引理 2.2 当假设 2.1 成立时, $H(z)$ 的雅克比矩阵 $H'(z)$ 在任意点 $z \in R_{++} \times R^n \times R^n \times R^\ell$ 处是非奇异的.

3 算法

本节给出一个求解广义圆锥互补问题的无导数光滑算法.

算法 3.1 (一个无导数光滑算法)

步骤 0 选取参数 $\lambda_1, \lambda_2, \tau, \delta \in (0,1)$. 选取初始点 $z^0 := (\mu_0, x^0, s^0, y^0) \in R_{++} \times R^n \times R^n \times R^\ell$. 选取参数 $\gamma \in (0,1)$ 满足 $\mu_0 \geq \gamma$. 选取一个正序列 $\{\eta_k\}$ 满足 $\sum_{k=0}^{\infty} \eta_k \leq \eta < \infty$, 这里 $\eta > 0$ 是一个给定的常数. 令 $C_0 := \|H(z^0)\|$, $\beta_0 := \gamma \min\{1, \|H(z^0)\|^2\}$. 令 $p := (1,0,0,0) \in R^{1+2n+\ell}$. 令 $k := 0$.

步骤 1 如果 $\|H(z^k)\| = 0$, 则停止迭代.

步骤 2 通过求解如下线性方程组

$$H(z^k) + H'(z^k)\Delta z^k = \beta_k p, \tag{3.1}$$

得到搜索方向 $\Delta z^k := (\Delta\mu_k, \Delta x^k, \Delta y^k, \Delta t^k) \in R \times R^n \times R^n \times R^\ell$.

步骤 3 如果

$$\|H(z^k + \Delta z^k)\| \leq \tau \|H(z^k)\| - \lambda_1 \|\Delta z^k\|^2, \tag{3.2}$$

则令 $\alpha_k := 1$, 转步骤 5.

步骤 4 令 l_k 是满足下列不等式最小的非负整数 l

$$\|H(z^k + \delta^l \Delta z^k)\| \leq (1+\eta_k) C_k - \lambda_2 \|\delta^l \Delta z^k\|^2. \tag{3.3}$$

令 $\alpha_k := \delta^{l_k}$, 转步骤 5.

步骤 5 令 $z^{k+1} := z^k + \alpha_k \Delta z^k$. 令 $\tau_k := \frac{1}{1+\eta_{k+1}}$. 计算

$$C_{k+1} := (1-\tau_k) C_k + \tau_k \|H(z^{k+1})\|, \tag{3.4}$$

$$\beta_{k+1} := \min\{\gamma, \gamma\|H(z^{k+1})\|^2, \beta_k\}. \tag{3.5}$$

令 $k := k+1$, 转步骤 1.

定理 3.1 如果假设2.1成立, 那么算法3.1可以产生一个无穷序列 $\{z^k = (\mu_k, x^k, s^k, y^k)\}$ 并且对所有的 $k \geq 0$ 都有 $\mu_k > 0$ 和 $\|H(z^k)\| < (1+\eta_k) C_k$.

证 假设对某个 k 满足 $z^k \in R_{++} \times R^n \times R^n \times R^\ell$ 和 $\|H(z^k)\| < (1+\eta_k) C_k$. 由引理 2.2 可知 $H'(z^k)$ 是非奇异的, 所以步骤 2 是可行的. 因为

$$\lim_{l\to\infty} \|H(z^k + \delta^l \Delta z^k)\| = \|H(z^k)\| < (1+\eta_k) C_k = \lim_{l\to\infty} [(1+\eta_k) C_k - \lambda_2 \|\delta^l \Delta z^k\|^2],$$

故至少存在一个非负整数 l 满足 (3.3). 这表明步骤 4 是可行的. 因此, 我们可以在步骤 5 中得到第 $k+1$ 个迭代点 $z^{k+1} = z^k + \alpha_k \Delta z^k$. 现在我们证明 $z^{k+1} \in R_{++} \times R^n \times R^n \times R^\ell$ 和 $\|H(z^{k+1})\| < (1+\eta_{k+1}) C_{k+1}$. 事实上, 根据方程组 (3.1) 中的第一个等式可以得到 $\Delta\mu_k = -\mu_k + \beta_k$, 再结合 $\alpha_k \in (0,1]$ 以及 $\beta_k > 0$ 可知

$$\mu_{k+1} = \mu_k + \alpha_k \Delta\mu_k = (1-\alpha_k)\mu_k + \alpha_k \beta_k > 0. \tag{3.6}$$

这表明 $z^{k+1} \in R_{++} \times R^n \times R^n \times R^\ell$. 此外, 由步骤 5 可知

$$C_{k+1} = \left(1 - \frac{1}{1+\eta_{k+1}}\right)C_k + \frac{1}{1+\eta_{k+1}}\|H(z^{k+1})\|,$$

进而可得

$$(1+\eta_{k+1})C_{k+1} = \eta_{k+1}C_k + \|H(z^{k+1})\| > \|H(z^{k+1})\|.$$

因此我们可以得到如下结论: 如果 $z^k \in R_{++} \times R^n \times R^n \times R^\ell$ 和 $\|H(z^k)\| < (1+\eta_k)C_k$ 对于某个 k 成立, 那么可由算法 3.1 产生 z^{k+1} 并且满足 $z^{k+1} \in R_{++} \times R^n \times R^n \times R^\ell$ 和 $\|H(z^{k+1})\| < (1+\eta_{k+1})C_{k+1}$. 因为 $z^0 \in R_{++} \times R^n \times R^n \times R^\ell$ 和 $\|H(z^0)\| = C_0 < (1+\eta_0)C_0$, 故由数学归纳法可知定理成立.

4 收敛性分析

4.1 全局收敛性

引理 4.1 设 $\{z^k = (\mu_k, x^k, y^k, t^k)\}$ 是算法 3.1 产生的迭代序列, 则对于所有的 $k \geq 0$ 满足 $\mu_k \geq \beta_k$ 和 $\mu_k \geq \mu_{k+1}$.

证明 如果 $\mu_k \geq \beta_k$ 对于某个 k 成立, 则由 (3.6) 可知 $\mu_{k+1} \geq (1-\alpha_k)\beta_k + \alpha_k\beta_k = \beta_k \geq \beta_{k+1}$. 因为 $\mu_0 \geq \gamma \geq \beta_0$, 故由数学归纳法可知对于所有的 $k \geq 0$ 满足 $\mu_k \geq \beta_k$. 利用这个结论, 我们可以由 (3.6) 进一步得到 $\mu_{k+1} \leq (1-\alpha_k)\mu_k + \alpha_k\mu_k = \mu_k$.

引理 4.2 由算法3.1产生的序列 $\{C_k\}$ 收敛.

证明 如果 α_k 由步骤 3 产生, 则根据定理 3.1 可知

$$\|H(z^{k+1})\| \leq \tau\|H(z^k)\| < \|H(z^k)\| < (1+\eta_k)C_k; \tag{4.1}$$

否则 α_k 由步骤 4 产生, 我们也有

$$\|H(z^{k+1})\| \leq (1+\eta_k)C_k. \tag{4.2}$$

从 (4.1) 和 (4.2) 可知对于所有的 $k \geq 0$ 都有 $\|H(z^{k+1})\| \leq (1+\eta_k)C_k$, 进而由 (3.4) 可得

$$\begin{aligned}
C_{k+1} &= (1-\tau_k)C_k + \tau_k\|H(z^{k+1})\| \\
&\leq (1-\tau_k)C_k + \tau_k(1+\eta_k)C_k \\
&= (1+\tau_k\eta_k)C_k \\
&= \left(1 + \frac{\eta_k}{1+\eta_{k+1}}\right)C_k \\
&\leq (1+\eta_k)C_k.
\end{aligned}$$

因为 $\sum_{k=0}^{\infty} \eta_k < \infty$, 故由文献 [12] 中的引理 2.2 可知 $\{C_k\}$ 收敛.

引理 4.3 设 $\{z^k\}$ 是由算法 3.1 产生的迭代序列, 则有 $\lim_{k\to\infty} \|z^{k+1} - z^k\| = 0$.

证明 如果 α_k 由步骤 3 产生, 那么 $\alpha_k = 1$ 并且

$$\|H(z^{k+1})\| \leq \tau\|H(z^k)\| - \lambda_1\|\Delta z^k\|^2 < (1+\eta_k)C_k - \lambda_1\|\alpha_k\Delta z^k\|^2; \tag{4.3}$$

否则 α_k 由步骤 4 产生, 则由 (3.3) 可得

$$\|H(z^{k+1})\| \leq (1+\eta_k)C_k - \lambda_2\|\alpha_k \Delta z^k\|^2. \tag{4.4}$$

令 $\lambda := \min\{\lambda_1, \lambda_2\}$, 则由 (4.3) 和 (4.4) 可知对于所有的 $k \geq 0$ 都有

$$\|H(z^{k+1})\| \leq (1+\eta_k)C_k - \lambda\|\alpha_k \Delta z^k\|^2. \tag{4.5}$$

因此, 对于所有的 $k \geq 0$, 由 (3.4), (4.5) 以及 $\frac{1}{1+\eta} \leq \tau_k = \frac{1}{1+\eta_{k+1}} < 1$ 可知

$$\begin{aligned}
C_{k+1} &= (1-\tau_k)C_k + \tau_k\|H(z^{k+1})\| \\
&\leq (1-\tau_k)C_k + \tau_k[(1+\eta_k)C_k - \lambda\|\alpha_k \Delta z^k\|^2] \\
&= (1+\tau_k\eta_k)C_k - \tau_k\lambda\|\alpha_k \Delta z^k\|^2 \\
&\leq (1+\eta_k)C_k - \frac{\lambda}{1+\eta}\|\alpha_k \Delta z^k\|^2,
\end{aligned}$$

进而可得

$$\frac{\lambda}{1+\eta}\|\alpha_k \Delta z^k\|^2 \leq (1+\eta_k)C_k - C_{k+1}. \tag{4.6}$$

因为 $\{C_k\}$ 收敛并且 $\lim\limits_{k\to\infty} \eta_k = 0$, 故由 (4.6) 可得 $\lim\limits_{k\to\infty} \|\alpha_k \Delta z^k\| = 0$, 再结合 $z^{k+1} - z^k = \alpha_k \Delta z^k$ 可知引理成立.

定理 4.1 设 $\{z^k = (\mu_k, x^k, y^k, t^k)\}$ 是由算法 3.1 产生的迭代序列, 则 $\{z^k\}$ 的任意聚点 z^* 都是 $H(z) = 0$ 的解.

证明 由引理 4.2 可知存在一个常数 $C^* \geq 0$ 使得 $\lim_{k\to\infty} C_k = C^*$. 由于 $\lim\limits_{k\to\infty} \eta_k = 0$, 故 $\lim\limits_{k\to\infty} \tau_k = 1$, 进而由 (3.4) 可得

$$\lim_{k\to\infty} \|H(z^k)\| = \lim_{k\to\infty} \left[\frac{C_k}{\tau_{k-1}} - (1-\tau_{k-1})\frac{C_{k-1}}{\tau_{k-1}}\right] = C^*. \tag{4.7}$$

下面我们证明 $C^* = 0$. 显然, 如果有无穷多的 k 使得 (3.2) 成立, 即 $\|H(z^{k+1})\| \leq \tau\|H(z^k)\|$ 对于无穷多的 k 都成立, 则有 $C^* \leq \tau C^*$, 再结合 $\tau \in (0,1)$ 可得 $C^* = 0$. 现在我们假定存在指标 $\bar{k} > 0$, 当 $k \geq \bar{k}$ 时 α_k 都由步骤 4 确定. 因为 z^* 是 $\{z^k\}$ 的聚点, 不失一般性, 我们假设 $\lim\limits_{k\to\infty} z^k = z^*$, 则由 $H(z)$ 的连续性可得

$$\lim_{k\to\infty} \|H(z^k)\| = \|H(z^*)\| = C^*. \tag{4.8}$$

因为 $\{\beta_k\}$ 单调递减且有下界, 故存在 $\beta^* \geq 0$ 使得 $\lim_{k\to\infty} \beta_k = \beta^*$. 现在假设 $\beta^* > 0$, 我们将推出矛盾. 由 β_k 的定义可知对于所有的 $k \geq 0$ 有

$$\beta_k \leq \min\{\gamma, \gamma\|H(z^k)\|^2\} \leq \gamma\|H(z^k)\|, \tag{4.9}$$

进而可得

$$\|H(z^*)\| \geq \frac{1}{\gamma}\beta^* > 0. \tag{4.10}$$

现在我们把证明分成两个部分.

第 1 部分. 假定对所有的 $k \geq \bar{k}$ 都有 $\alpha_k \geq c > 0$, 这里 c 是一个常数. 那么由 (3.3) 可得

$$\lambda_2 c \|\Delta z^k\|^2 \leq \lambda_2 \|\alpha_k \Delta z^k\|^2 \leq (1+\eta_k) C_k - \|H(z^{k+1})\|.$$

这和 (4.8) 以及 $\lim\limits_{k\to\infty} \eta_k = 0$ 表明 $\lim\limits_{k\to\infty} \Delta z^k = 0$. 因此, 令 (3.1) 的两边 $k \to \infty$ 可得 $H(z^*) = \beta^* p$. 再结合 (4.10) 可得 $\|H(z^*)\| = \beta^* \leq \gamma \|H(z^*)\|$. 由于 $\gamma \in (0,1)$, 我们有 $\|H(z^*)\| = 0$, 这和 (4.10) 相矛盾.

第 2 部分. 假定 $\lim\limits_{k\to\infty} \alpha_k = 0$. 令 $\hat{\alpha}_k := \alpha_k/\delta$, 那么 $\lim\limits_{k\to\infty} \hat{\alpha}_k = 0$, 并且当 $k \geq \bar{k}$ 时都有

$$\begin{aligned} \|H(z^k + \hat{\alpha}_k \Delta z^k)\| &> (1+\eta_k) C_k - \lambda_2 \|\hat{\alpha}_k \Delta z^k\|^2 \\ &> \|H(z^k)\| - \lambda_2 \|\hat{\alpha}_k \Delta z^k\|^2. \end{aligned}$$

因此

$$\frac{\|H(z^k + \hat{\alpha}_k \Delta z^k)\| - \|H(z^k)\|}{\hat{\alpha}_k} > -\lambda_2 \hat{\alpha}_k \|\Delta z^k\|^2. \tag{4.11}$$

由引理 4.1 可知 $\mu^* = \lim\limits_{k\to\infty} \mu_k \geq \lim\limits_{k\to\infty} \beta_k = \beta^* > 0$. 这表明 $H(z)$ 在 z^* 处是连续可微的, 故令 (4.11) 中的两边 $k \to \infty$ 可得

$$H(z^*)^T H'(z^*) \Delta z^* \geq 0. \tag{4.12}$$

另一方面, 从 (3.1) 可得

$$H(z^*)^T H'(z^*) \Delta z^* = -\|H(z^*)\|^2 + \mu^* \beta^* \leq -(1-\gamma) \|H(z^*)\|^2, \tag{4.13}$$

这里不等式成立是因为 $\mu^* \leq \|H(z^*)\|$ 以及 $\beta^* \leq \gamma \|H(z^*)\|$. 由 (4.12) 和 (4.13) 可知 $(1-\gamma)\|H(z^*)\|^2 \leq 0$. 由于 $\gamma \in (0,1)$, 故可得 $\|H(z^*)\| = 0$, 这和 (4.10) 矛盾. 因此, 我们可得 $\beta^* = 0$. 由 β_k 的定义, 必存在一个序列 $\{z^{k_n}\}$ 使得 $\lim\limits_{k_n\to\infty} \|H(z^{k_n})\| = 0$, 这和 (4.7) 表明 $C^* = 0$, 即 $\lim\limits_{k\to\infty} \|H(z^k)\| = 0$, 进而再由 $H(z)$ 的连续性可得 $H(z^*) = 0$.

定理 4.2 设 $\{z^k\}$ 是由算法 3.1 产生的迭代序列. 如果 $\{z^k\}$ 有一个孤立的聚点 z^*, 则 $\lim\limits_{k\to\infty} z^k = z^*$.

证 由引理 4.3 和文献 [13] 中的命题 8.3.10 可知结论成立.

4.2 局部二次收敛性

定理 4.3 设 z^* 是由算法 3.1 产生的迭代序列 $\{z^k\}$ 的任意聚点, 并且 F' 在 z^* 点处局部 Lipschitz 连续. 如果所有的 $V \in \partial H(z^*)$ 都是非奇异的, 那么 $\{z^k\}$ 收敛到 z^*, 并且有

$$\|z^{k+1} - z^*\| = O(\|z^k - z^*\|^2), \ \ \|H(z^{k+1})\| = O(\|H(z^k)\|^2).$$

证 由文献 [11] 中的引理 6 可知 $H(z)$ 在 z^* 点是强半光滑的, 故类似于文献 [14] 定理 8 的证明, 我们可以证得对所有充分接近于 z^* 的 z^k 都有

$$\|z^k + \Delta z^k - z^*\| = O(\|z^k - z^*\|^2), \tag{4.14}$$

$$\|H(z^k+\Delta z^k)\| = O(\|H(z^k)\|^2). \tag{4.15}$$

由于所有的 $V\in\partial H(z^*)$ 都是非奇异的, 故由文献 [15] 中的命题 3.1 可知存在一个常数 $C>0$ 使得对所有充分接近于 z^* 的 z^k 都有

$$\|H'(z^k)^{-1}\| \le C. \tag{4.16}$$

因此, 由 (3.1), (4.9) 和 (4.16) 可知对所有充分接近于 z^* 的 z^k 都有

$$\|\Delta z^k\| \le \|H'(z^k)^{-1}\|\|\beta_k p - H(z^k)\| \le C(\gamma+1)\|H(z^k\|.$$

这表明对所有充分接近于 z^* 的 z^k 都有

$$\|\Delta z^k\|^2 = O(\|H(z^k)\|^2). \tag{4.17}$$

由 (4.15) 和 (4.17) 可知对所有充分接近于 z^* 的 z^k 都有

$$\|H(z^k+\Delta z^k)\| + \lambda_1\|\Delta z^k\|^2 \le \tau\|H(z^k)\|.$$

因此, 对所有充分接近于 z^* 的 z^k, $\alpha_k=1$ 都满足 (3.2), 即 $z^{k+1}=z^k+\Delta z^k$. 再结合 (4.14) 和 (4.15) 可知定理成立.

5 数值实验

本节我们对算法 3.1 进行数值实验. 参数取值为

$$\lambda_1=0.01,\ \lambda_2=0.01,\ \tau=0.5,\ \delta=0.8,\ \mu_0=10^{-3},\ \gamma=10^{-4},\ \eta_k=0.95^k.$$

终止准则为 $\|H(z^k)\|\le 10^{-6}$.

考虑如下凸二次圆锥规划:

$$\min f(x)=\frac{1}{2}x^TQx+c^{\mathrm{T}}x \quad \text{s.t.}\ \ Ax=b,\ \ x\in\ L_\theta,$$

这里 $Q\in\ R^{n\times n}$ 是一个半正定矩阵, $c\in\ R^n$, $A\in\ R^{\ell\times n}$ 且 $b\in\ R^\ell$. 该问题的 KKT 最优性条件为如下广义圆锥互补问题:

$$x\in L_\theta,\ y\in L_\theta^*,\ \langle x,y\rangle=0,\ F(x,y,t)=0, \tag{5.1}$$

其中

$$F(x,y,t)=\begin{pmatrix} Qx-A^Tt-y+c \\ Ax-b \end{pmatrix}. \tag{5.2}$$

在数值试验中, 我们产生规模为 $n(=2\ell)$, $L_\theta = L_\theta^{n_1}\times L_\theta^{n_2}\times L_\theta^{n_3}\times L_\theta^{n_4}$ 和 $n_i=\frac{n}{4}$ $(i=1,...,4)$ 的测试问题. 具体而言, 我们首先产生一个行满秩矩阵 $A\in\ R^{\ell\times n}$, 然后产生向量 $\bar x=(\bar x_1,\bar x_2,\bar x_3,\bar x_4)\in \mathrm{int}L_\theta$, 其中 $\bar x_i=((\|a_i\|+1)\mathrm{ctan}\theta,a_i)\in\mathrm{int}L_\theta^{n_i}$ 和 $a_i=\mathrm{rand}(n_i-1,1)$, 令 $b:=A\bar x$. 此外, 我们利用前面产生 $\bar x$ 的方法产生 $c\in\mathrm{int}L_\theta$, 并选取 $Q=nBB^T/\|BB^T\|$, 这里 $B=\mathrm{rand}(n,\ell)$. 我们选取 $x^0=y^0=(1,0,...,0)^T$ 和 $t^0=(0,...,0)^T$ 作为初始点. 每个测试问题我们产生 10 个算例, 数值结果列于表 1, 其中**AIT**和**ACPU**分别表示算法所需迭代次数和 CPU 时间的平均值.

表 1. 算法 3.1 的数值实验结果

	$\theta=\frac{\pi}{3}$		$\theta=\frac{\pi}{4}$		$\theta=\frac{\pi}{5}$	
n	**AIT**	**ACPU**	**AIT**	**ACPU**	**AIT**	**ACPU**
100	6.8	0.08	6.6	0.09	7.7	0.08
200	6.6	0.19	6.3	0.25	7.4	0.27
300	7.0	0.47	6.5	0.45	7.6	0.59
400	6.9	0.91	6.2	0.79	7.1	1.08
500	6.9	1.41	6.3	1.39	7.3	1.65
600	7.0	2.24	6.4	2.17	7.4	2.34
700	6.8	3.27	6.2	3.17	7.2	3.47
800	7.0	4.35	6.5	4.40	7.3	4.83
900	6.9	5.92	6.3	5.64	7.0	7.34
1000	7.0	8.44	6.4	7.84	7.2	10.77

由表 1 可以看出, 算法 3.1 是非常有效的, 它只需很少的迭代次数和 CPU 时间就可以得到满足终止条件的解. 此外, 我们还发现算法 3.1 求解问题所需的迭代次数几乎不受问题规模的影响, 这说明算法有很好的稳定性.

参 考 文 献

[1] Bai Y Q, Gao X R, Wang G Q. Primal-dual interior-point algorithms for convex quadratic circular cone optimization [J]. Numerical Algebra, Control and Optimization, 2015, 5(2): 211–231.

[2] Bai Y Q, Ma P F, Zhang J. A polynomial-time interior-point method for circular cone programming based on kernel functions [J].Journal of Industrial and Management Optimization, 2016, 12(2): 739–756.

[3] Pirhaji M, Zangiabadi M, Mansouri H. A path following interior-point method for linear complementarity problems over circular cones [J]. Japan Journal of Industrial and Applied Mathematics, 2018, 35 (3): 1103–1121.

[4] Chi X N, Wan, Z P, Zhu Z B, Yuan L Y. A nonmonotone smoothing Newton method for circular cone programming [J]. Optimization, 2016, 65(12): 2227–2250.

[5] Chi X N, Wei H J, Wan Z P, Zhu Z B. Smoothing Newton algorithm for the circular cone programming with a nonmonotone line search [J]. Acta Mathematics Scientia, 2017, 37B(5): 1262–1280.

[6] Chi X N, Wei H J, Wan Z P, Zhu Z B. A nonmonotone smoothing Newton algorithm for circular cone complementarity problems [J]. Journal of Computational Analysis and Applications, 2019, 26(1): 146–162.

[7] Miao X H, Yang J T, Hu S L. A generalized Newton method for absolute value equations associated with circular cones [J]. Applied Mathematics and Computation, 2015, 269(C): 155–168.

[8] Zhou J C, Chen J S. Properties of circular cone and spectral factorization associated with circular cone [J]. Journal of Nonlinear and Convex Analysis, 2013, 14(4): 807–816.

[9] Dong L, Tang J Y, Song X Y. Numerical study of a smoothing algorithm for the complementarity system over the second-order cone [J]. Computational and Applied Mathematics, 2018, 37(3): 2845–2861.

[10] Narushima Y, Sagara N, Ogasawara H. A smoothing Newton method with Fischer-Burmeister function for second-order cone complementarity problems [J]. Journal of Optimization and Theory Applications, 2011, 149(1): 79–101.

[11] Tang J Y, Zhou J C. Smoothing inexact Newtonmethod based on a new derivative-free nonmonotone line search for the NCP over circular cones [J]. Annals of Operations Research, 2020, 295(2): 787–808.

[12] Li D H, Fukushima M. A derivative-free line search and global convergence of Broyden-like method for nonlinear equations [J]. Optimization Methods and Software, 2000, 13(3): 181–201.

[13] Facchinei F, Pang J S. Finite-dimensional variational inequalities and complementarity problems (I&II) [M]. New York: Springer, 2003.

[14] Qi L, Sun D, Zhou G. A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequality problems [J]. Mathematical Programming, 2000, 87(1): 1–35.

[15] Qi L, Sun J. A nonsmooth version of Newton's method [J]. Mathematical Programming, 1993, 58(1–3): 353–367.

A DERIVATIVE-FREE SMOOTHING ALGORITHM FOR SOLVING GENERAL CIRCULAR CONE COMPLEMENTARITY PROBLEMS

SHAO Can-ran, TANG Jing-yong

(*College of Mathematics and Statistics, Xinyang Normal University, Henan 464000, China*)

Abstract: In this paper we study a derivative-free smoothing algorithm for solving the general circular cone complementarity problem. By using a smoothing function, we reformulate the general circular cone complementarity problem as a system of smooth equations and solve it by Newton method. The algorithm adopts a new nonmonotone derivative-free line search and it has global and local quadratic convergence under some suitable conditions. Numerical results show that the algorithm is very effective.

Keywords: general circular cone complementarity problem; smoothing algorithm; derivative-free line search; quadratic convergence

2010 MR Subject Classification: 90C30; 65K05