

A New Smoothing Newton Method for Second-Order Cone Optimization Based on SOC Function

Xiaojuan Liang, Youfang Zeng

School of Mathematics and Information Science, Guangxi University, Nanning Guangxi
Email: 1533090300@qq.com, zengyf@gxu.edu.cn

Received: Mar. 27th, 2019; accepted: Apr. 11th, 2019; published: Apr. 18th, 2019

Abstract

Based on the Fischer-Burmeister function and the Natural-Residual function, a new smoothing Newton method is proposed for solving the second-order cone programming. This algorithm adopts a Newton equation with disturbance to gain the search direction. Under suitable assumptions, we prove that the proposed method is globally and locally quadratically convergent. Finally, some numerical results are given.

Keywords

Second-Order Cone Programming, Smoothing Newton Method, Global Convergence, Quadratical Convergence

二阶锥规划基于SOC函数的新光滑牛顿算法

梁晓娟, 曾友芳

广西大学数学与信息科学学院, 广西 南宁
Email: 1533090300@qq.com, zengyf@gxu.edu.cn

收稿日期: 2019年3月27日; 录用日期: 2019年4月11日; 发布日期: 2019年4月18日

摘要

本文在向量值Fischer-Burmeister (FB)函数和向量值Natural-Residual (NR)函数的基础上, 提出一种求解二阶锥规划(SOCP)问题的光滑函数。用一个带扰动的牛顿方程组去获得搜索方向, 在适当假设下, 分

析了算法的全局收敛和局部收敛速度, 给出了数值实验结果。

关键词

二阶锥规划, 光滑牛顿法, 全局收敛, 局部二次收敛

Copyright © 2019 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

本文主要考虑以下线性二阶锥规划问题:

$$\begin{aligned} \min & \sum_i^r c_i^T x_i \\ \text{s.t.} & \sum_i^r A_i x_i = b, \\ & x_i \succeq_{K^{n_i}} 0, i \in I = \{1, 2, \dots, r\} \end{aligned} \quad (1.1)$$

其中, $b \in R^m, c_i \in R^{n_i}, A_i \in R^{m \times n_i}$ 是已知量, $x_i \succeq_{K^{n_i}} 0 (i \in I)$ 表示 $x_i \in K^{n_i}$ 是变量。 K^{n_i} 是维数为 n_i 的二阶锥, 即

$$K^{n_i} = \{x_i = (x_{i0}; \bar{x}_i) \in R^{n_i} \mid x_{i0} \geq \|\bar{x}_i\|\},$$

这里 $\|\cdot\|$ 为向量的欧几里得范数。易知 K^{n_i} 是自对偶的, 即 $K^{n_i} = (K^{n_i})^* = \{s_i \in R^{n_i} \mid x_i^T s_i \geq 0, \forall x_i \in K^{n_i}\}$ 其对偶规划为:

$$\begin{aligned} \max & b^T y \\ \text{s.t.} & A_i^T y + s_i = c_i, i \in I, \\ & s_i \succeq_{K^{n_i}} 0, i \in I. \end{aligned} \quad (1.2)$$

其中 $y \in R^m, s_i \in K^{n_i}, I = \{1, 2, \dots, r\}$ 。为便于问题的描述, 使用如下记号:

$$\begin{aligned} n & \triangleq \sum_{i=1}^r n_i, K = K^{n_1} \times K^{n_2} \times \dots \times K^{n_r}, A = (A_1, A_2, \dots, A_r) \in R^{m \times n}, \\ c & = (c_1, c_2, \dots, c_r) \in R^n, x = (x_1; x_2; \dots; x_r) \in K, s = (s_1; s_2; \dots; s_r) \in K. \end{aligned}$$

这里用 R^n 表示 n 维实列向量, 且将 $R^{n_1} \times R^{n_2} \times \dots \times R^{n_r}$ 定义为 $R^{n_1+n_2+\dots+n_r}$, 因此 $(x_1; x_2; \dots; x_r) \in K$ 是 $R^{n_1+n_2+\dots+n_r}$ 中的列向量。则上面(1.1)和(1.2)式可转化为如下简洁形式:

$$\min \{c^T x \mid Ax = b, x \in K\}. \quad (1.3)$$

其对偶问题:

$$\max \{b^T y \mid A^T y + s = c, s \in K, y \in R^m\}. \quad (1.4)$$

二阶锥规划是一类非光滑凸规划问题, 它是在仿射空间与有限个二阶锥的笛卡尔积的交集上极小化或极大化一个线性函数。在现实生活中有着广泛的应用: 许多工程, 力学, 投资组合, 线阵列等都可以转化为二阶锥问题求解。此外, 许多数学问题也可转化为二阶锥求解, 如线性规划, 二次规划, 凸二

次约束二次规划, 范数极小化, 因而研究二阶锥规划具有重要的理论与现实意义。本文在向量值 FB 函数和向量值 NR 函数的基础上进行改进, 得到了一个新的光滑函数, 利用得到的光滑函数, 把二阶锥互补函数转化为一个非线性方程组问题, 并给出求解二阶锥问题的一个光滑化算法, 在无严格互补条件的假设下, 证明算法全局收敛和局部二次收敛。

2. 基础知识

二阶锥规划建立在与二阶锥相伴的代数基础上, 为了更好地研究二阶锥规划, 下面将简要介绍与二阶锥相伴的欧几里得若当代数的基本概念和与之相关的重要结论, 二阶锥规划的对偶理论和最优性条件等。先介绍一些记号, 其中 U 表示反射矩阵, E_{n-1} 表示 $(n-1)$ 阶单位阵, $Arw(x)$ 表示一个箭形矩阵, 本文用相应的大写字母表示箭形矩阵, $X = Arw(x)$, $S = Arw(s)$, $W = Arw(w)$ 。

$$U \triangleq \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & -E_{n-1} \end{pmatrix} \in R^{n \times n}, Arw(x) \triangleq \begin{pmatrix} x_1 & \bar{x}^T \\ \bar{x} & x_0 E_{n-1} \end{pmatrix}, \mathbf{e} \triangleq (1; \mathbf{0}) \in R^n.$$

对于 $\forall x = (x_0; \bar{x}) \in R \times R^{n-1}, s = (s_0; \bar{s}) \in R \times R^{n-1}$, 定义如下若当积:

$$x \circ s \triangleq \begin{pmatrix} x^T s \\ x_0 \bar{s} + s_0 \bar{x} \end{pmatrix} = Arw(x)s = Arw(x)Arw(s)\mathbf{e} = XSe.$$

记 $x^2 \triangleq x \circ x$, $x + y$ 表示通常的向量的向量加法。

定理 2.1: (谱分解定理) (文[1])

对向量 x 定义与二阶锥有关的谱分解为: $x = \lambda_1 c_1 + \lambda_2 c_2$, 其中 x 的谱值 λ_i 和与之对应的谱向量 c_i 如下:

$$\lambda_i = x_0 + (-1)^{i+1} \|\bar{x}\|, i = 1, 2.$$

$$c_i = \begin{cases} \frac{1}{2} \begin{pmatrix} 1; (-1)^{i+1} \frac{\bar{x}}{\|\bar{x}\|} \end{pmatrix}, \bar{x} \neq \mathbf{0}; \\ \frac{1}{2} \begin{pmatrix} 1; (-1)^{i+1} \omega \end{pmatrix}, \bar{x} = \mathbf{0}; \end{cases} \quad i = 1, 2. \quad \omega \in R^{n-1} \text{ 是满足 } \|\omega\| = 1 \text{ 的任意向量。}$$

关于谱值, 谱向量的相关性质归纳如下:

- 1) $x \in K \Leftrightarrow \lambda_2 \geq \lambda_1 \geq 0, x \in K^0 \Leftrightarrow \lambda_2 \geq \lambda_1 > 0$,
- 2) 对任意的 $x \in R^n$, $x^2 = \lambda_1^2 c_1 + \lambda_2^2 c_2 \in K$,
- 3) 对任意的 $x \in K$, $\sqrt{x} = \sqrt{\lambda_1} c_1 + \sqrt{\lambda_2} c_2 \in K$ 。

在线性规划中, 互补性条件为: 若 $x \geq \mathbf{0}, s \geq \mathbf{0}$, 则 $x^T s = \mathbf{0}$ 当且仅当 $x_i s_i = \mathbf{0}, i = 1, \dots, r$, 二阶锥规划有着类似于线性规划的互补性条件。

引理 2.1: (互补性条件) 设 $x, s \in K$, 即 $x_i, s_i \in K_i, i = 1, \dots, r$, 则 $x^T s = \mathbf{0}$ 当且仅当 $x_i \circ s_i = \mathbf{0}$ 。

在线性规划中, 若原问题与对偶问题任意一个存在最优解, 则另一个也存在最优解, 且原问题与对偶问题最优值相等, 但二阶锥规划不具有这种性质, 若要二阶锥规划的强对偶理论成立, 必须使得原问题与对偶问题同时存在严格可行解。因此二阶锥规划的最优性条件为:

定理 2.2: (最优性条件) (文[2]) 如果(1.3)和(1.4)都有严格可行解, 则 $(x, (y, s))$ 是(1.3)和(1.4)的最优解对当且仅当

$$\begin{cases} Ax = b, x \in K, \\ A^T y + s = c, s \in K, \\ x \circ s = \mathbf{0}. \end{cases} \quad (2.1)$$

对二阶锥规划, 其二阶锥互补函数定义如下:

定义 2.1: (文[3])如果向量值函数 $\phi^{\text{soc}}: R^n \times R^n \rightarrow R^n$ 满足

$$\phi^{\text{soc}}(x, s) = \mathbf{0} \Leftrightarrow x \circ s = \mathbf{0}, x \in K, s \in K. \quad (2.2)$$

则称 ϕ^{soc} 为二阶锥互补函数。

结合问题(1.3)和(1.4)的最优性条件, 定义函数 $H(x, y, s): R^{n+m+n} \rightarrow R^{n+m+n}$ 如下:

$$H(x, y, s) = \begin{pmatrix} b - Ax \\ c - A^T y - s \\ \phi^{\text{soc}}(x - s) \end{pmatrix}, \quad (2.3)$$

其中 $\phi^{\text{soc}}(x - s)$ 是任意的二阶锥互补函数。易见 $H(x, y, s) = \mathbf{0}$ 是与最优性条件等价的方程组, 由此可知, $H(x, y, s) = \mathbf{0}$ 的解 (x^*, y^*, s^*) 满足最优性条件, 是问题(1.3)和(1.4)的最优解对。由上易知, 求解此方程组的关键在于构造合适的二阶锥互补函数。由于常见的二阶锥互补函数在 $(0, 0)$ 点不可微(即非光滑), 例如向量值 FB 函数, 向量值 NR 函数等, 不能直接用牛顿法求解。为了进一步研究, 先给出光滑化的概念。

定义 2.2: (文[4])对于不可微函数 $h: R^n \rightarrow R^m$, 考虑带有参数 $\mu > 0$ 的函数 $h_\mu: R^n \rightarrow R^m$, 如果它具有如下性质:

- 1) 对于任意的 $\mu > 0$, h_μ 是光滑的;
- 2) $\lim_{\mu \rightarrow 0} h_\mu(x) = h(x), \forall x \in R^n$,

则称 h_μ 是 h 的光滑函数。

3. 一个新的光滑函数及其性质

光滑函数在二阶锥规划光滑算法研究中起着重要作用(文[5] [6] [7] [8] [9])。由于向量值 FB 函数和向量值 NR 函数并不处处连续可微, 大大影响了其实际应用。本文通过光滑化对称扰动得到一个新的向量值函数 $\phi(\mu, x, s): R_+ \times R^n \times R^n \rightarrow R^n$:

$$\phi(\mu, x, s) = x + s - \sqrt{\mu(x-s)^2 + (1-\mu)(x^2 + s^2) + 2\mu^2 e} \quad (3.1)$$

为了减少变量的个数, 使证明简单, 令 $z = (\mu, x, y)$, 定义函数 $H(z)$ 如下:

$$H(z) = \begin{pmatrix} e^\mu - 1 \\ b - Ax \\ \phi(\mu, x, c - A^T y) \end{pmatrix}, \quad (3.2)$$

显然, 若 $(0, x^*, y^*)$ 是 $H(z) = 0$ 的解, 则 $(x^*, y^*, c - A^T y^*)$ 是原问题(1.3)及其对偶问题(1.4)的最优解。因此, 可考虑用牛顿法求解 $H(z) = \mathbf{0}$, 但对于所构造的 $H(z)$, 其雅可比矩阵必须是非奇异的。在讨论 $H(z)$ 的性质之前先介绍相关结论。

引理 3.1: (文[3])对于任意的 $a^1, \dots, a^p \in R^n$, 令 $\chi(a^1, \dots, a^p) = \sqrt{\sum_{i=1}^p (a^i)^2}$ 。则 χ 是全局利普希茨连续的; 若 $v_0 \neq \|\bar{v}\|$, 其中 $v = (v_0; \bar{v}) \in R \times R^{n-1}$ 且 $v = \sum_{i=1}^p (a^i)^2$, 则 χ 在 (a^1, \dots, a^p) 的任意邻域是连续可微的。

定理 3.1: 设函数 $H(z)$ 由(3.2)定义, 令 $\omega = \sqrt{\mu(x-s)^2 + (1-\mu)(x^2 + s^2) + 2\mu^2 e}$, $\mu \in (0, 1)$, 则对任意的 $\phi(\mu, x, s) \in R_+ \times R^n \times R^n$, 有以下结论:

1) $\phi(\mu, x, s)$ 是全局利普希茨连续的处处光滑函数, 且在任意 (μ, x, s) 处连续可微, 有

$$\phi'_\mu(\mu, x, s) = -W^{-1}(x \circ s + 2\mu e), \tag{3.3}$$

$$\phi'_x(\mu, x, s) = E_n - W^{-1}(X - \mu S), \tag{3.4}$$

$$\phi'_s(\mu, x, s) = E_n - W^{-1}(S - \mu X). \tag{3.5}$$

2) $\phi(\mu, x, s)$ 在 $\mu = 0$ 处是光滑的。

证明: 1) 由引理 3.1 可知 $\phi(\mu, x, s)$ 是全局利普希茨连续且在任意 (μ, x, s) 处连续可微。由 $\omega'_\mu(\mu, x, s) = W^{-1}(x \circ s + 2\mu e)$, 知

$$\phi'_\mu(\mu, x, s) = -W^{-1}(x \circ s + 2\mu e), \tag{3.6}$$

同理可得:

$$\omega'_x(\mu, x, s) = W^{-1}(X - \mu S),$$

$$\omega'_s = W^{-1}(S - \mu X).$$

由此可得:

$$\phi'_x(\mu, x, s) = E_n - W^{-1}(X - \mu S), \tag{3.7}$$

$$\phi'_s(\mu, x, s) = E_n - W^{-1}(S - \mu X). \tag{3.8}$$

2) 对于任意的 $x = (x_0; \bar{x}), s = (s_0; \bar{s}) \in R \times R^{n-1}$, 由谱分解定理有

$$\omega^2 = \lambda_1(\mu)u_1(\mu) + \lambda_2(\mu)u_2(\mu),$$

因此

$$\phi(\mu, x, s) = x + s - \left(\sqrt{\lambda_1(\mu)}u_1(\mu) + \sqrt{\lambda_2(\mu)}u_2(\mu) \right),$$

其中

$$\lambda_i(\mu) = \mu(x-s)^2 + (1-\mu)(x^2 + s^2) + 2\mu^2 + 2(-1)^{i+1} \|v(\mu)\|, i = 1, 2.$$

$$u_i(\mu) = \begin{cases} \frac{1}{2} \left(1; (-1)^{i+1} \frac{v(\mu)}{\|v(\mu)\|} \right), & v(\mu) \neq \mathbf{0}; \\ \frac{1}{2} \left(1; (-1)^{i+1} w \right), & v(\mu) = \mathbf{0}. \end{cases} \quad i = 1, 2.$$

其中 $v(\mu) = \mu(x_0 - s_0)(\bar{x} - \bar{s}) + (1-\mu)(x_0\bar{x} + s_0\bar{s})$, $\omega \in R^{n-1}$ 是满足 $\|\omega\| = 1$ 的任意向量。同理有 $\phi(0, x, s) = x + s - (\sqrt{\lambda_1}u_1 + \sqrt{\lambda_2}u_2)$, 其中 $\lambda_i = \|x\|^2 + \|s\|^2 + 2(-1)^{i+1} \|v\|, i = 1, 2$ 。

$$u_i = \begin{cases} \frac{1}{2} \left(1; (-1)^{i+1} \frac{v}{\|v\|} \right), & v \neq \mathbf{0}; \\ \frac{1}{2} \left(1; (-1)^{i+1} w \right), & v = \mathbf{0}. \end{cases} \quad i = 1, 2.$$

其中 $v = x_0s_0 + \bar{x}\bar{s}$, $\omega \in R^{n-1}$ 是满足 $\|\omega\| = 1$ 的任意向量。

不失一般性, 取 $\omega \in R^{n-1}$ 和 $v(\mu)$ 中的一致, 显然 $\lim_{\mu \rightarrow 0} v(\mu) = v$, 于是 $\lim_{\mu \rightarrow 0} \lambda_i(\mu) = \lambda_i, \lim_{\mu \rightarrow 0} u_i(\mu) = u_i$, 因此 $\lim_{\mu \rightarrow 0} \phi(\mu, x, s) = \phi(0, x, s)$, 结合 $\phi(\mu, x, s)$ 在任意 (μ, x, s) 处连续可微, 知 $\phi(\mu, x, s)$ 是光滑的, 因此由光

滑函数的定义知 $\phi(\mu, x, s)$ 是 $\phi(0, x, s)$ 的一个光滑函数。

4. 算法的描述

记 $z(\mu, x, y): R_+ \times R^n \times R^n, \theta(z) = \|H(z)\|^2, \eta = \sqrt{\theta(z)} + 1$, 令 $\gamma \in (0, 1)$, 当 $H(z) \neq \mathbf{0}$ 时, 定义函数 $\beta, \bar{\beta}: R^{n+m+1} \rightarrow R_+ : \beta(z) = e^\mu \gamma \min\{1, \theta(z)\}, \bar{\beta}(z) = \gamma \min\{1, \theta(z)\}$ 。在矩阵 A 行满秩的假设下, 可证明算法是适定的。

算法 4.1: (二阶锥规划的一个光滑型牛顿法)

步骤 0: 给出常数 $\delta, \sigma, \mu_0 \in (0, 1), \gamma \in (0, 1)$ 满足 $\gamma\mu_0 < \frac{1}{2}, \gamma\eta < \frac{1}{2}$, 任取 $(x^0, y^0) \in R^n \times R^m$, $z^0 = (\mu_0, x^0, y^0), \bar{z} = (\mu_0, 0, 0)$, 令 $k := 0$ 。

步骤 1: 若 $H(z^k) = \mathbf{0}$, 停止, 其中 $\phi(z^k)$ 及 $H(z^k)$ 分别由(2.5)和(2.6)式所定义, 否则令 $\beta_k := \beta(z^k)$ 。

步骤 2: (确定搜索方向)解线性方程组 $H(z^k) + H'(z^k)\Delta z^k = \beta_k \bar{z}$ 得到搜索方向 Δz^k 。

步骤 3: (确定步长)记 α_k 是满足下式的最小非负整数 α :

$$\theta(z^k + \delta^\alpha \Delta z^k) \leq [1 - \sigma(1 - 2\gamma\mu_0\eta)\delta^\alpha] \theta(z^k),$$

令步长 $t_k = \delta^{\alpha_k}$ 。

步骤 4: 令 $z^{k+1} = z^k + t_k \Delta z^k, k := k+1$, 返回步骤 1。

在证明算法的适定性之前, 先介绍相关性质。

引理 4.1: (文[10]) 对任意 $\mu > 0$,

$$-\mu \leq \frac{1 - e^\mu}{e^\mu} \leq -\mu e^{-\mu}. \quad (4.1)$$

引理 4.2: (文[11]) 对于任意的 $x, y \in R$ 和 $\omega \succ_{K^n} \mathbf{0}$, 如果 $\omega^2 \succ_{K^n} x^2 + y^2$, 则

$[Arw(\omega) - Arw(x)][Arw(\omega) - Arw(y)] \succ \mathbf{0}, Arw(\omega) - Arw(x) \succ \mathbf{0}, Arw(\omega) - Arw(y) \succ \mathbf{0}$, 且当 \succ 变为 \succeq 时, 上述结论仍然成立。

定理 4.1: 令 $z = (\mu, x, y) \in R_+ \times R^n \times R^m, H(z): R^{1+n+m} \rightarrow R^{1+m+n}$ 由(3.2)定义, 则有以下结论成立:

1) 在任意点 $z = (\mu, x, y)$ 处, $H(z)$ 全局利普希茨连续, 并且连续可微, 其雅克比矩阵表示为:

$$H'(z) = \begin{pmatrix} e^\mu & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -A & \mathbf{0} \\ \phi'_\mu(z) & \phi'_x(z) & -\phi'_s(z)A^T \end{pmatrix}, \quad (4.2)$$

2) 对任意的 $\mu > 0, H'(z)$ 非奇异。

证明: 由定理 3.1 知(1)成立。下证(2)成立, 对任意给定的 $\mu > 0$, 要证 $H'(z)$ 非奇异, 只需证明方程组 $H'(z)\Delta z = \mathbf{0}$ 只有零解, 即 $\Delta z = (\Delta\mu, \Delta x, \Delta y) = \mathbf{0}$ 。将(4.2)式代入 $H'(z)\Delta z = \mathbf{0}$ 中可知

$$\Delta\mu = 0, A\Delta x = \mathbf{0}, \phi'_x(z)\Delta x - \phi'_s(z)A^T\Delta y = \mathbf{0} \quad (4.3)$$

将(3.7), (3.8)式中的 $\phi'_x(z), \phi'_s(z)$ 代入(4.3)式并将等式左右的两端同时左乘以 W , 得:

$$[W - (X - \mu S)\Delta x] - [W - (S - \mu X)]A^T\Delta y = 0. \quad (4.4)$$

因 $w^2 - \mu(x-s)^2 - (1-\mu)(x^2 + s^2) = 2\mu^2 e \in K^0$, 由引理 4.2 可知: 矩阵 $W - (S - \mu X)$ 正定, 从而 $W - (S - \mu X)$ 可逆。将方程(4.4)左乘

$\Delta x^T [W - (S - \mu X)]^{-1}$, 结合 $A\Delta x = \mathbf{0}$, 得:

$$\Delta x^T [W - (S - \mu X)]^{-1} [W - (X - \mu S)] \Delta x - \Delta x^T [W - (S - \mu X)]^{-1} [W - (S - \mu X)] A^T \Delta y = \mathbf{0}. \quad \text{令}$$

$\Delta \tilde{x} [W - (S - \mu X)]^{-1} \Delta x$, 则

$$\Delta \tilde{x}^T [W - (X - \mu S)] [W - (S - \mu X)] \Delta \tilde{x} = \mathbf{0}. \quad (4.5)$$

由引理 4.2 可知矩阵 $[W - (X - \mu S)] [W - (S - \mu X)]$ 正定。因此 $\Delta \tilde{x} = \mathbf{0}$, 进而 $\Delta x = \mathbf{0}$ 。结合矩阵 A 行满秩的假设和(4.4)式, 知 $\Delta y = \mathbf{0}$ 。这就说明 $H'(z)\Delta z = \mathbf{0}$ 只有零解, 故 $H'(z)$ 非奇异。

定理 4.2: 设矩阵 A 行满秩, 如果 $\mu_k > 0$, 则算法 4.1 是适定的。

证明: 因为矩阵 A 行满秩且 $\mu_k > 0$, 由定理 4.1 知 $H'(z)$ 非奇异, 所以算法 4.1 的步骤 2 是适定的。令 $\Delta z^k = (\Delta \mu_k, \Delta x^k, \Delta y^k) \in R \times R^n \times R^m$ 是步骤 2 中方程组的解, 则对任意的 $t \in (0, 1]$, 有

$$\mu_{k+1} = \mu_k + t\Delta \mu_k = \mu_k + t \left(\frac{1 - e^{\mu_k}}{e^{\mu_k}} + \frac{\beta_k \mu_0}{e^{\mu_k}} \right) \geq (1-t)\mu_k + t\mu_0 \frac{\beta_k}{e^{\mu_k}} > 0. \quad (4.6)$$

由 $\beta(z)$ 的定义有 $\beta(z) \leq e^\mu \gamma \sqrt{\theta(z)}$,

$$\begin{aligned} e^{\mu_k + t\Delta \mu_k} - 1 &= e^{\mu_k} e^{t\Delta \mu_k} - 1 \\ &= e^{\mu_k} (1 + t\Delta \mu_k + O(t^2)) - 1 \\ &= e^{\mu_k} - 1 + t(1 - e^{\mu_k}) + t\beta_k \mu_0 + O(t^2) \\ &= (1-t)(e^{\mu_k} - 1) + t\beta_k \mu_0 + O(t^2) \end{aligned}$$

从而

$$\begin{aligned} (e^{\mu_k + t\Delta \mu_k} - 1)^2 &= (1-t)^2 (e^{\mu_k} - 1)^2 + 2t(1-t)\beta_k (e^{\mu_k} - 1)\mu_0 + t^2\beta_k \mu_0 + O(t^2) \\ &\leq (1-2t)(e^{\mu_k} - 1)^2 + 2te^{\mu_k} \gamma \sqrt{\theta(z^k)} (e^{\mu_k} - 1)\mu_0 + O(t^2) \\ &\leq (1-2t)(e^{\mu_k} - 1)^2 + 2t\gamma \mu_0 \eta \theta(z^k). \end{aligned}$$

在上述最后一个不等式利用 $e^{\mu_k} - 1 \leq \sqrt{\theta(z^k)}$, $e^{\mu_k} \leq \eta$ 可得。

$$\text{令 } \varphi(z^k) = \begin{pmatrix} b - Ax^k \\ \phi(z^k) \end{pmatrix}, \text{ 并记 } \Phi(z^k) = \|\phi(z^k)\|^2, \text{ 可得 } (\Phi'(z^k))^T \Delta z^k = -2\|\phi(z^k)\|^2 = -2\Phi(z^k),$$

从而

$$\begin{aligned} \Phi(z^k + t\Delta z^k) &= \Phi(z^k) + t(\Phi'(z^k))^T \Delta z^k + o(t) \\ &= \Phi(z^k) - 2t\Phi(z^k) + o(t) \\ &= (1-2t)\Phi(z^k) + o(t). \end{aligned}$$

因此

$$\begin{aligned} \theta(z^k + t\Delta z^k) &= \|H(z^k + t\Delta z^k)\|^2 \\ &\leq (1-2t)(e^{\mu_k} - 1)^2 + 2t\gamma \mu_0 \eta \theta(z^k) + (1-2t)\Phi(z^k) + o(t) \\ &\leq (1-2t)\theta(z^k) + 2t\gamma \mu_0 \eta \theta(z^k) + o(t) \\ &\leq [1 - (1 - 2\gamma \mu_0 \eta)t] \theta(z^k) + o(t). \end{aligned}$$

因为 $\gamma\mu_0\eta < 1$, 所以存在一个常数 $\bar{t} \in (0, 1]$, 使得对任意的 $t \in (0, \bar{t}]$, 有

$$\theta(z^k + t\Delta z^k) \leq [1 - \sigma(1 - 2\gamma\mu_0\eta)t] \theta(z^k), \quad (4.7)$$

即步骤 3 可在有限步终止, 故步骤 3 是适定的。综上可知, 算法 4.1 是适定的。

5. 收敛性分析

定理 5.1: 1) 设矩阵 A 行满秩且 $\{z^k := (\mu_k, x^k, y^k)\}$ 是算法 4.1 生成的无穷迭代点列, 则对 $\forall k \geq 0$, 若 $\mu_k > 0, z^k \in \Lambda, \theta(z^k) \leq \theta(z^0)$, 则 $z^{k+1} \in \Lambda$, 其中

$$\Lambda := \{z := (\mu, x, y) \in R_+ \times R^n \times R^m \mid \mu \geq \beta(z)\mu_0\}, \quad (5.1)$$

2) (全局收敛性) $\{z^k\}$ 的任意聚点 $z^* = (\mu^*, x^*, y^*)$ 都是 $H(z^*) = \mathbf{0}$ 的解。

证明: 1) 根据 $\beta(z)$ 的定义分两种情况讨论:

i) 若 $\theta(z^k) \leq 1$, 则 $\beta_k = \gamma e^{\mu_k} \|\theta(z^k)\|$ 。对任意的 $t \in (0, 1]$, 从(4.7)式知

$\theta(z^k + t\Delta z^k) \leq [1 - \sigma(1 - 2\gamma\mu_0\eta)t] \theta(z^k)$, 故 $\bar{\beta}(z^k + t\Delta z^k) = \gamma \theta(z^k + t\Delta z^k)$ 。由于 $z^k \in \Lambda$, 可得

$$\begin{aligned} & \mu_k + t\Delta\mu_k - \bar{\beta}(z^k + t\Delta z^k)\mu_0 \\ & \geq (1-t)\mu_k + t\gamma\mu_0\theta(z^k) - \bar{\beta}(z^k + t\Delta z^k)\mu_0 \\ & \geq (1-t)\bar{\beta}_k\mu_0 + t\gamma\mu_0\theta(z^k) - \gamma\theta(z^k + t\Delta z^k)\mu_0 \\ & \geq (1-t)\gamma\theta(z^k)\mu_0 + t\gamma\mu_0\theta(z^k) - \gamma[1 - \sigma(1 - 2\gamma\mu_0\eta)t]\theta(z^k)\mu_0 \\ & = 2\gamma\sigma(1 - \gamma\mu_0\eta)t\theta(z^k)\mu_0 \geq 0. \end{aligned}$$

ii) 若 $\theta(z^k) > 1$, 则 $\beta_k = \gamma e^{\mu_k}, \bar{\beta}_k := \bar{\beta}(z^k) = \gamma$ 。对任意的 $z \in R^{2n+1}$, 有 $\bar{\beta}(z) \leq \gamma$, 从而对任意的 $t \in (0, 1]$, 有

$$\begin{aligned} & \mu_k + t\Delta\mu_k - \bar{\beta}(z^k + t\Delta z^k)\mu_0 \\ & = \mu_k + t \left(\frac{1 - e^{\mu_k}}{e^{\mu_k}} + \frac{\beta_k \mu_0}{e^{\mu_k}} \right) - \bar{\beta}(z^k + t\Delta z^k) \mu_0 \\ & \geq (1-t)\mu_k + t\gamma\mu_0 - \gamma\mu_0 \geq (1-t)\bar{\beta}_k\mu_0 + t\gamma\mu_0 - \gamma\mu_0 \\ & = (1-t)\gamma\mu_0 + t\gamma\mu_0 - \gamma\mu_0 = 0. \end{aligned}$$

从而 $z^{k+1} \in \Lambda$ 。

2) 不失一般性, 设 z^* 是点列 $\{z^k\}$ 的任意聚点, 要证 $H(z^*) = \mathbf{0}$, 用反证法。假设 $\|H(z^*)\| > \mathbf{0}$, 由算法 4.1 的步骤 3 知 $\{\|H(z^k)\|\}$ 单调下降且有界, 结合 $H(z^k)$ 的连续性, 有 $\lim_{k \rightarrow +\infty} \|H(z^k)\| = \|H(z^*)\|$ 。由 $\beta(\cdot)$ 的定义知 $\beta(z^k)$ 单调下降且趋于 $\beta(z^*)$: $\beta_* = \beta(z^*) = e^{\mu_*} \gamma \min\{1, \theta(z^*)\} > 0$ 。另外, 由(5.1)和(4.6)式得到:

$$\begin{aligned} 0 & < \mu_{k+1} = \mu_k + t\Delta\mu_k \\ & = \mu_k + t \left(\frac{1 - e^{\mu_k}}{e^{\mu_k}} + \frac{\beta_k \mu_0}{e^{\mu_k}} \right) \\ & \leq \mu_k + t \left(-\mu_k e^{\mu_k} + \frac{e^{\mu_k} \bar{\beta}_k \mu_0}{e^{\mu_k}} \right) \\ & \leq \mu_k + t(-\mu_k e^{\mu_k} + \mu_k) \leq \mu_k. \end{aligned}$$

即 μ_k 是单调下降的, 则 $\lim_{k \rightarrow +\infty} \mu_k = \mu_*$. 对 $\mu_k \geq \beta_k \mu_0$ 两边取极限, 有 $\mu_* \geq \beta_* \mu_0 > 0$. 由定理 4.1 知 $H'(z^*)$ 存在且非奇异. 由引理 4.1 知: 存在 z^* 的一个闭邻域 $N(z^*)$ 和正数 $\bar{t} \in (0, 1]$, 使得对于任意的 $z = (\mu, x, y) \in N(z^*)$ 和所有的 $t \in [0, \bar{t}]$, 都有 $\mu > 0$, $H'(z)$ 非奇异且

$$\theta(z^k + t\Delta z^k) \leq [1 - \sigma(1 - 2\gamma\mu_0\eta)t]\theta(z^k).$$

可以找到一个非负整数 τ , 使得 $\delta^\tau \in [0, \bar{t}]$, 当 k 充分大时, 有

$$\theta(z^{k+1}) \leq [1 - \sigma(1 - 2\gamma\mu_0\eta)t]\theta(z^k).$$

由步骤 3 知, 对于充分大的 k , 步长 $t_k = \delta^{\alpha_k} \geq \delta^\tau$, 因此有

$$\theta(z^{k+1}) \leq [1 - \sigma(1 - 2\gamma\mu_0\eta)t_k]\theta(z^k) \leq [1 - \sigma(1 - 2\gamma\mu_0\eta)\delta^\tau]\theta(z^k).$$

这与序列 $\{\theta(z^k)\}$ 的极限 $\theta(z^*) > 0$ 矛盾, 故 $\theta(z^*) = 0$, 即 $H(z^*) = \mathbf{0}$, 因此算法 4.1 是全局收敛的.

接下来证明算法的一个局部收敛性, 在证明之前先介绍相关引理.

引理 5.1 (文[12]): 如果 $V \in \partial F(z^*)$ 都是非奇异的, 那么存在 z^* 的一个邻域 $N(z^*)$ 和一个常数 $C > 0$, 使得对于任意的 $z^k \in Nz^*$ 和任意的 $V \in \partial F(z^k)$. V 非奇异且 $\|V^{-1}\| \leq C$.

定理 5.2 (局部二次收敛): 设矩阵 A 行满秩且 z^* 是算法 4.1 产生的迭代点列 $\{z^k\}$ 的任意聚点, 则当 k 充分大时, $\{z^k\}$ 二阶收敛到 z^* , 即 $\|z^{k+1} - z^*\| = O(\|z^k - z^*\|^2)$, 且 $\mu_{k+1} = O(\mu_k^2)$.

证明: 由定理 5.1 的(2)知 $H(z^*) = 0$, 又由定理 4.1 知 $H'(z^*)$ 存在且非奇异. 故由引理 5.1 知, 存在一个常数 $C > 0$, 对于充分接近 z^* 的所有点 z^k , 有 $\|(H'(z^k))^{-1}\| \leq C$. 由定理 3.1 和定理 4.1 知 $H(z)$ 在 z^k 处是全局利普希茨连续的强半光滑函数, 因为 $H(x + \Delta x) - H(x) - V(\Delta x) = O(\|\Delta x\|^{1+p})$, 当 z^k 充分接近 z^* 时, 由半光滑函数的定义知, 当 z^k 充分接近点 z^* 时, 有

$$\|H(z^k)\| = \|H(z^k) - H(z^*)\| = O(\|z^k - z^*\|)$$

和

$$H(z^*) - H(z^k) - H'(z^k)(z^* - z^k) = O(\|z^k - z^*\|^2).$$

又 $\beta_k = e^{\mu_k} \gamma \min\{1, \|H(z^k)\|\} \leq \|H(z^k)\|^2$, 即 $\beta_k = O(\|H(z^k)\|^2) = O(\|z^k - z^*\|^2)$. 因此,

$$\begin{aligned} \|z^k + \Delta z^k - z^*\| &= \|z^k + (H'(z^k))^{-1}(\beta_k \bar{z} - H(z^k)) - z^*\| \\ &= \|(H'(z^k))^{-1}[H'(z^k)(z^k - z^*) + \beta_k \bar{z} - H(z^k)]\| \\ &\leq \|(H'(z^k))^{-1}\| \|H(z^*) - H(z^k) - H'(z^k)(z^* - z^k) + \beta_k \bar{z}\| \\ &\leq C [\|H(z^*) - H(z^k) - H'(z^k)(z^* - z^k)\| + \beta_k \mu_0] \\ &= O(\|z^k - z^*\|^2). \end{aligned} \tag{5.2}$$

进一步有

$$\|H(z^k + \Delta z^k)\| = O(\|z^k + \Delta z^k - z^*\|) = O(\|z^k - z^*\|^2) = O(\|H(z^k)\|^2) \quad (5.3)$$

即当 z^k 充分接近 z^* 时, 有 $z^{k+1} = z^k + \Delta z^k$ 。由(4.5)知

$$\|z^{k+1} - z^*\| = O(\|z^k - z^*\|^2) \quad (5.4)$$

即算法 4.1 是二次收敛的, 其次, 当 k 充分大时, $H(z^k)$ 单调减少趋向于 $H(z^*)$, 即 $H(z^k) \rightarrow \mathbf{0}$, 所以 $\beta_k = e^{\mu_k} \gamma \|H(z^k)\|^2$, 又因为 $z^{k+1} = z^k + \Delta z^k$, 所以当 k 充分大时有

$$\mu_{k+1} = \mu_k + \Delta \mu_k = \frac{\beta_k \mu_0}{e^{\mu_k}} = \gamma \|H(z^k)\|^2 \mu_0. \quad (5.5)$$

结合(5.3)式知, 当 z^k 充分接近 z^* 时, 有

$$\frac{\mu_{k+1}}{\mu_k} = \frac{\|H(z^k)\|^2}{\|H(z^{k-1})\|^2} = \frac{O(\|H(z^{k-1})\|^4)}{\|H(z^{k-1})\|^2} = O(\|H(z^{k-1})\|^2) = O(\mu_k),$$

即 $\mu_{k+1} = O(\mu_k^2)$ 。

6. 数值试验

为了检验算法 4.1 的有效性, 本文用 Matlab 2014a 编程, 在 Inter (R) Pentium (R) 4CPU, 3.40 GHz, 2 GB 内存, Windows 7 操作系统的电脑上做数值试验。测试问题是随机生成的规模 $n = 2m$ 从 20 到 800 维不等且 $r = 1$ 的二阶锥规划问题。具体步骤为: 首先生成随机的行满秩矩阵 $A \in R^{m \times n}$ 和随机向量 $x, s \in K_0$, 然后令 $b = Ax, c = A^T y + s$, 则得到的二阶锥规划的原问题和对偶问题都存在最优解且最优值相等。在测试问题过程中, 初始点选取 $x^0 = e \in R^n, y^0 = \mathbf{0} \in R^m$ 。参数取值为: $\mu_0 = 0.01, \delta = 0.65, \sigma = 0.35, \gamma = 0.90$ 。终止准则为 $\|H(z^k)\| \leq 10^{-6}$ 。计算结果列于表 1, 其中 IT 和 CPU(s) 表示每个测试问题进行 10 次所求得最优解时所用的平均迭代次数和平均时间。

Table 1. The numerical experiment results of the algorithm

表 1. 算法的数值实验结果

算法 4.1 文[13]					
n	m	IT	CPU (s)	IT	CPU(s)
20	10	9	0.05	-	-
50	25	11	0.07	-	-
100	50	11	0.12	9	0.11
200	100	11	0.16	9	0.53
300	150	12	0.80	10	1.15
400	200	14	1.60	10	2.55
500	250	14	3.08	10	6.13
600	300	14	5.34	11	9.74
700	350	14	7.31	11	18.93
800	400	18	12.86	11	22.74

由表 1 可以看出, 算法是有效的, 并且可以求解大规模的二阶锥规划问题, 它只需较少的 CPU 时间就可以得到满足终止条件的解。并且本文中的算法 4.1 求解问题所用时间大部分比文献[13]所用时间少。

基金项目

国家自然科学基金(11561005)。

参考文献

- [1] Faraut, J. and Koranyi, A. (1994) Analysis on Symmetric Cone. Clarendon Press, Oxford.
- [2] Alizadeh, F. and Goldfarb, D. (2003) Second-Order Cone Programming. *Mathematical Programming: Series B*, **95**, 3-51.
- [3] Sun, D.F. (2005) Strong Semismoothness of the Fischer-Burmeister SDC and SOC Complementarity Functions. *Mathematical Programming*, **103**, 575-581. <https://doi.org/10.1007/s10107-005-0577-4>
- [4] Hayashi, S., Yamashita, N. and Fukushima, M. (2005) A Combined Smoothing and Regularization Method for Monotone Second-Order Cone Complementarity Problems. *SIAM Journal on Optimization*, **15**, 593-615. <https://doi.org/10.1137/S1052623403421516>
- [5] Chi, X.N. and Liu, S.Y. (2009) A One-Step Smoothing Newton Method for Second-Order Cone Programming. *Journal of Computational and Applied Mathematics*, **223**, 114-123. <https://doi.org/10.1016/j.cam.2007.12.023>
- [6] Chi, X.N. and Liu, S.Y. (2009) A Non-Interior Continuation Method for Second Cone Optimization. *Optimization*, **58**, 965-979. <https://doi.org/10.1080/02331930701763421>
- [7] Tang, J.Y., Dong, L., Fang, L. and Zhou, J.C. (2014) Smoothing Newton Algorithm for the Second-Order Cone Programming with a Nonmonotone Line Search. *Optimization Letters*, **8**, 1753-1771. <https://doi.org/10.1007/s11590-013-0699-1>
- [8] Tang, J.Y., He, G.P., Dong, L., Fang, L. and Zhou, J.C. (2015) A Globally and Quadratically Convergent Smoothing Newton Method for Solving Second-Order Cone Optimization. *Applied Mathematical Modelling*, **39**, 2180-2193. <https://doi.org/10.1016/j.apm.2014.10.027>
- [9] Fang, L. and Feng, Z.Z. (2011) A Smoothing Newton-Type Method for Second-Order Cone Programming Problems Based on a New Smoothing Fischer-Burmeister Function. *Computational & Applied Mathematics*, **30**, 569-588. <https://doi.org/10.1590/S1807-03022011000300005>
- [10] Jiang, H. (1997) Smoothed Fischer-Burmeister Equation Methods for the Complementarity Problem. Technical Report. Department of Mathematics, The University of Melbourne, Parville, Victoria, Australia.
- [11] Fukushima, M., Luo, Z.Q. and Tseng, P. (2001) Smoothing Functions for Second-Order Cone Complementarity Problems. *SIAM Journal on Optimization*, **12**, 436-460. <https://doi.org/10.1137/S1052623400380365>
- [12] Qi, L.Q. and Sun, J. (1993) A Nonsmooth Version of Newton's Method. *Mathematical Programming*, **58**, 353-367. <https://doi.org/10.1007/BF01581275>
- [13] 汤京永, 贺国平. 二阶锥规划的一步光滑牛顿法[J]. 数学物理学报, 2012, 32(4): 768-778.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2324-7991, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: aam@hanspub.org