对称锥规划的邻域跟踪算法

刘长河^{①*}, 刘红卫^②, 尚有林^①

① 河南科技大学数学与统计学院, 洛阳 471023;

② 西安电子科技大学理学院数学系, 西安 710071

E-mail: changheliu@163.com, hwliu@mail.xidian.edu.cn, ylshang@haust.edu.cn

收稿日期: 2011-07-31; 接受日期: 2013-02-21; * 通信作者

国家自然科学基金 (批准号: 61072144 和 61179040) 资助项目

摘要 本文把艾文宝的邻域跟踪算法推广到对称锥规划, 定义中心路径的宽邻域 $\mathcal{N}(\tau, \beta)$, 并证明该邻域的一个重要性质, 该性质在算法的复杂性分析中起到关键作用. 取宽邻域 $\mathcal{N}(\tau, \beta)$ 中一点为初始点并采用 Nesterov-Todd (NT) 搜索方向, 则该算法的迭代复杂界为 $\mathcal{O}(\sqrt{r} \log \varepsilon^{-1})$, 其中, r 是 Euclid Jordan 代数的秩, ε 是允许误差. 这是对称锥规划的宽邻域内点算法最好的复杂界.

关键词 对称锥规划 Euclid Jordan 代数 邻域跟踪算法 宽邻域 内点法 多项式复杂性

MSC (2010) 主题分类 65K05, 90C05, 90C25, 90C51

1 引言

对称锥规划问题, 即带有线性约束和对称锥约束的极小化线性目标函数的问题, 包括了运筹学中十分重要的一些优化问题, 如线性规划、半定规划和二阶锥规划. Nesterov 和 Todd^[1,2] 首先提出了对称锥规划的原 - 对偶内点算法, Faybusovich^[3,4] 利用 Euclid Jordan 代数的工具, 探讨了对称锥规划解的严格可行性、约束非退化性和中心路径的存在唯一性以及最优性系统的可解性等, Yoshise^[5] 研究了对称锥规划的最优性系统衍生出来的对称锥互补问题的中心路径的存在唯一性和有界性等. 这些性质为对称锥上各种内点算法的设计与分析提供了理论基础. Schmieta 和 Alizadeh^[6] 把 Monteiro 和 Zhang^[7] 提出的求解半定规划的内点算法推广到了对称锥规划, 基于搜索方向的“可交换类”, 他们证明了短步、半长步和长步路径跟踪算法的迭代复杂性.

在线性规划的内点算法中, 一个值得关注的成果是艾文宝^[8] 提出的邻域跟踪算法. 艾文宝设法让所有的迭代点都跟踪通常的宽邻域 $\mathcal{N}_{\infty}(1 - \tau)$, 但不一定属于这个宽邻域. 运用这一策略加上引入新的中心向量, 使得该邻域跟踪算法具有 $\mathcal{O}(\sqrt{nL})$ 迭代复杂性和局部二次收敛性, 其中, n 是线性规划问题中变量的个数, L 是输入数据的长度. 这是第一次得到与窄邻域具有相同复杂性的宽邻域内点算法. 基于文献 [8] 的思想, Ai 和 Zhang^[9] 提出了求解单调线性互补问题的具有 $\mathcal{O}(\sqrt{nL})$ 迭代复杂性的宽邻域路径跟踪算法. 近来, Feng 和 Fang^[10] 与 Li 和 Terlaky^[11] 分别把文献 [8] 和 [9] 中的宽邻域内点算法推广到半定规划问题, 并证明了算法的多项式复杂性. 本文把文献 [8] 中的路径跟踪算法从线性规划推广到对称锥规划问题, 基于 NT 方向, 证明了该算法的迭代复杂性是 $\mathcal{O}(\sqrt{r} \log \varepsilon^{-1})$, 其中, r 是 Euclid Jordan 代数的秩, ε 是允许误差. 这与 Schmieta 和 Alizadeh^[6] 提出的窄邻域内点算法的复杂性相同, 这也是对称锥上内点算法最好的复杂性结果.

英文引用格式: Liu C H, Liu H W, Shang Y L. Neighborhood-following algorithms for symmetric cone programming (in Chinese). Sci Sin Math, 2013, 43: 691–702, doi: 10.1360/012011-577

2 Euclid Jordan 代数和对称锥

本节主要介绍 Euclid Jordan 代数和对称锥的基本概念与相关理论, 并给出在内点算法分析中用到的一些重要结果, 详细内容参见文献 [6,12]. 设 $\mathcal{J}$ 是实数域 $\mathbb{R}$ 上的 n 维向量空间, 定义乘积 “ $\circ$ ” 使得双线性映射 $(x, y) \rightarrow x \circ y$ 满足下面性质:

- (1) $x \circ y = y \circ x$;
- (2) $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$, 其中, $x^2 = x \circ x$,

则称 $(\mathcal{J}, \circ)$ 是一个 Jordan 代数, 简记为 $\mathcal{J}$. 若还存在一个对称正定的二次型 Q 使得 $Q(x \circ y, z) = Q(x, y \circ z)$, 则称 Jordan 代数 $\mathcal{J}$ 是 Euclid 的, 简称为 Euclid Jordan 代数. 若元素 $e \in \mathcal{J}$ 满足 $x \circ e = e \circ x = x, \forall x \in \mathcal{J}$, 则称 e 为 Jordan 代数 $\mathcal{J}$ 的单位元. Jordan 代数 $\mathcal{J}$ 的平方锥定义为 $\mathcal{K} := \{x^2 : x \in \mathcal{J}\}$. 下面的定理建立了对称锥与 Euclid Jordan 代数之间的联系, 为利用 Euclid Jordan 代数研究对称锥上的优化问题奠定了基础, 参见文献 [12, 定理 III.2.1 和 III.3.1].

定理 2.1 一个锥是对称锥当且仅当它是某个 Euclid Jordan 代数的平方锥.

由于乘积 “ $\circ$ ” 是双线性的, 则对任意的 $x \in \mathcal{J}$, 存在线性映射 L_x 使得 $x \circ y = L_x y, \forall y \in \mathcal{J}$. 特别地, 有 $L_x e = x$. 对于 $x, y \in \mathcal{J}$, 定义 $Q_{x,y} := L_x L_y + L_y L_x - L_{x \circ y}$, $Q_x := Q_{x,x} = 2L_x^2 - L_{x^2}$. Q_x 称为 x 的二次表示, 它将在对称锥规划的算法分析中起到重要作用, 特别地, 有 $Q_x e = x^2$.

设 $x \in \mathcal{J}$, 使得 $\{e, x, x^2, \dots, x^r\}$ 线性相关的最小整数 r 称为 x 的次数, 记为 $\deg(x)$. 所有 $\deg(x)$ 的最大值称为 x 的秩, 记为 $\text{rank}(\mathcal{J})$. 若 $\mathcal{J}$ 中的非零元 c 满足 $c^2 = c$, 则称它是幂等的; 若一个幂等元不能表示成两个幂等元的和, 则称它是本原的. 若一个幂等元集合 $\{c_1, c_2, \dots, c_k\}$ 满足 $c_i \circ c_j = 0, \forall i \neq j$, 且 $c_1 + c_2 + \dots + c_k = e$, 则称该集合为正交幂等元完备系; 进一步, 若其中的每个幂等元都是本原的, 则称它为 $\mathcal{J}$ 的 Jordan 标架.

定理 2.2 [12, 定理 III.1.2] (谱分解) 设 Euclid Jordan 代数 $\mathcal{J}$ 的秩为 r , 则对于任意的 $x \in \mathcal{J}$, 存在 Jordan 标架 $\{c_1, c_2, \dots, c_r\}$ 与一组实数 $\lambda_1, \lambda_2, \dots, \lambda_r$ 使得 $x = \lambda_1 c_1 + \lambda_2 c_2 + \dots + \lambda_r c_r$, 其中, $\lambda_1, \lambda_2, \dots, \lambda_r$ 是 x 的所有的特征值 (包含重数).

定义下面运算: 当全部特征值 $\lambda_i \neq 0$, 逆运算 $x^{-1} := \lambda_1^{-1} c_1 + \lambda_2^{-1} c_2 + \dots + \lambda_r^{-1} c_r$; 当全部特征值 $\lambda_i \geq 0$, 平方根运算 $x^{1/2} := \lambda_1^{1/2} c_1 + \lambda_2^{1/2} c_2 + \dots + \lambda_r^{1/2} c_r$; 迹 $\text{tr}(x) := \lambda_1 + \lambda_2 + \dots + \lambda_r$; 行列式 $\det(x) := \lambda_1 \lambda_2 \dots \lambda_r$. $x \in \mathcal{J}$ 的最小和最大特征值分别记为 $\lambda_{\min}(x)$ 和 $\lambda_{\max}(x)$. 若 $x \in \mathcal{J}$ 的所有特征值都不等于零, 则称 x 是可逆的. 若 $x \in \mathcal{J}$ 的特征值都是非负的 (正的), 则称 x 是半正定的 (正定的), 并记作 $x \succeq 0$ ($x \succ 0$). 显然, x 是半正定的 (正定的) 当且仅当 x 属于 $\mathcal{J}$ 的平方锥 (平方锥的内部), 即 $x \in \mathcal{K}$ ($x \in \text{int}\mathcal{K}$).

引理 2.1 [13, 引理 2.15] 若 $x \circ y \in \text{int}\mathcal{K}$, 则 $\det(x) \neq 0$.

设 $\mathcal{J}$ 是秩为 r 的 Euclid Jordan 代数, 由于对称二次型 $\text{tr}(x \circ y)$ 满足结合性和正定性 (参见文献 [12, 命题 III.1.5 和 II.4.3]), 则可定义内积 $\langle x, y \rangle := \text{tr}(x \circ y)$. 并且对任意的 $x, y, z \in \mathcal{J}$, 都有 $\langle x \circ y, z \rangle = \langle x, y \circ z \rangle$. 设 $x \in \mathcal{J}$ 的特征值为 $\lambda_i, 1 \leq i \leq r$, 定义 Frobenius 范数和谱范数分别为 $\|x\|_F := \sqrt{\langle x, x \rangle} = \sqrt{\sum \lambda_i^2}$ 和 $\|x\|_2 := \max_i |\lambda_i|$. 显然, 由于单位元 e 有 r 重特征值 1, 则 $\|e\|_F = \sqrt{r}$, $\|e\|_2 = 1$. 由于内积 $\langle \cdot, \cdot \rangle$ 是结合的, 则算子 L_x 和 L_x^{-1} 关于内积对称, 即 $\langle L_x y, z \rangle = \langle y, L_x z \rangle$, $\langle L_x^{-1} y, z \rangle = \langle y, L_x^{-1} z \rangle$. 同时由定义可知, 二次表示 Q_x 也是关于内积对称的.

引理 2.2 [14, 引理 2.9] 设 $x, y \in \mathcal{J}$, 则 $\|x \circ y\|_F \leq \|x\|_F \|y\|_F$.

设 $\mathcal{J}$ 是带有单位元 e 的 Jordan 代数, $\{c_1, c_2, \dots, c_r\}$ 是 $\mathcal{J}$ 的一组 Jordan 标架. 对于 $i, j \in \{1, 2, \dots, r\}$, 定义 Peirce 空间: $\mathcal{J}_{ii} := \{x \in \mathcal{J} : x \circ c_i = x\}$, $\mathcal{J}_{ij} := \{x \in \mathcal{J} : x \circ c_i = \frac{1}{2}x = x \circ c_j\}$, 当 $i \neq j$.

定理 2.3 [12, 定理 IV.2.1] (Peirce 分解) 空间 $\mathcal{J}$ 是 Peirce 空间 $\mathcal{J}_{ij}$ ($i \leq j$) 的直和.

因此, 给定 Jordan 标架 $\{c_1, c_2, \dots, c_r\}$, 元素 $x \in \mathcal{J}$ 可以表示成

$$x = \sum_{i=1}^r x_i c_i + \sum_{i < j} x_{ij}, \quad (2.1)$$

其中, $x_i \in \mathbb{R}$, $x_{ij} \in \mathcal{J}_{ij}$. 该表达式称为 x 关于 Jordan 标架 $\{c_1, c_2, \dots, c_r\}$ 的 Peirce 分解.

下面给出关于二次表示 Q_x 的两个命题.

命题 2.1 [12, 命题 III.2.2] 若 $x, y \in \text{int } \mathcal{K}$, 则 $Q_x y \in \text{int } \mathcal{K}$.

命题 2.2 [6, 命题 21] 设 $x, y, p \in \text{int } \mathcal{K}$, 并令 $\tilde{x} := Q_p x, \tilde{y} := Q_{p^{-1}} y$, 则 $Q_{x^{1/2}} y$ 和 $Q_{y^{1/2}} x$ 有相同的谱; $Q_{x^{1/2}} y$ 和 $Q_{\tilde{x}^{1/2}} \tilde{y}$ 有相同的谱.

设 $x, y \in \mathcal{J}$, 若 $L_x L_y = L_y L_x$, 则称 x 和 y 算子可交换.

引理 2.3 [6, 引理 30] 设 $x, y \in \text{int } \mathcal{K}$, 记 $w := Q_{x^{1/2}} y$, 则有 $\text{tr}(x \circ y) = \text{tr}(w)$, $\lambda_{\min}(x \circ y) \leq \lambda_{\min}(w)$, $\lambda_{\max}(x \circ y) \geq \lambda_{\max}(w)$. 并且, 若 x 和 y 算子可交换, 则 $x \circ y = w$.

若一个 Jordan 代数不能表示成两个 Jordan 代数的直和, 则称它是简单的. 下面给出 Euclid Jordan 代数的分类定理.

定理 2.4 [6, 定理 22] 设 $\mathcal{J}$ 一个简单 Euclid Jordan 代数, 则它与下面某个代数同构, 其中, 矩阵代数的乘积定义为 $X \circ Y = \frac{1}{2}(XY + YX)$, 称为 Jordan 积.

(1) $n \times n$ 实对称矩阵代数 $\text{Sym}(\mathbb{R}^{n \times n})$ [6, 例 1].

(2) 二次型代数 $\mathcal{E}_{n+1}$ [6, 例 2].

(3) $n \times n$ 复 Hermite 矩阵代数 $\text{Herm}(\mathbb{C}^{n \times n})$.

(4) $n \times n$ 四元数 Hermite 矩阵代数 $\text{Herm}(\mathcal{H}^{n \times n})$.

(5) 3×3 八元数 Hermite 矩阵代数 $\text{Herm}(\mathcal{O}^{3 \times 3})$.

下面的内容参见文献 [15].

设 $\mathcal{K}$ 是 $\mathcal{J}$ 中的闭凸锥, 令 $\Pi_{\mathcal{K}}$ 为 $\mathcal{K}$ 上的投影映射, 即 $x^* = \Pi_{\mathcal{K}}(x)$ 当且仅当 $x^* \in \mathcal{K}$, 且 $\|x - x^*\|_F \leq \|x - y\|_F, \forall y \in \mathcal{K}$. 若 $\mathcal{K}$ 是自对偶的, 则对任意的 $x \in \mathcal{J}$, 定义 x 的正部和负部分别为 $x^+ := \Pi_{\mathcal{K}}(x)$ 和 $x^- = x - x^+$. 设 x 有谱分解 $x = \lambda_1 c_1 + \lambda_2 c_2 + \dots + \lambda_r c_r$, 容易验证 $x^+ = \lambda_1^+ c_1 + \lambda_2^+ c_2 + \dots + \lambda_r^+ c_r$, $x^- = \lambda_1^- c_1 + \lambda_2^- c_2 + \dots + \lambda_r^- c_r$, 其中 $\lambda_i^+ = \max(\lambda_i, 0)$, $\lambda_i^- = \min(\lambda_i, 0)$.

定理 2.5 任何一个 Euclid Jordan 代数都可以唯一地表示成一些简单 Euclid Jordan 代数的直和. 并且, 它们的平方锥 (即对称锥) 也满足相应的直和表示关系, 且表示方式是唯一的.

给定 Euclid Jordan 代数 $\mathcal{J}$ 以及相应的对称锥 $\mathcal{K}$, 记 $\mathcal{J} = \mathcal{J}_1 \times \mathcal{J}_2 \times \dots \times \mathcal{J}_m$, $\mathcal{K} = \mathcal{K}_1 \times \mathcal{K}_2 \times \dots \times \mathcal{K}_m$, 其中, 每个 $\mathcal{J}_i$ 都是一个简单 Euclid Jordan 代数, $\mathcal{K}_i$ 为相应的对称锥. 对于 $\mathcal{J}$ 中的两个元素 $x = (x_1, x_2, \dots, x_m)$, $y = (y_1, y_2, \dots, y_m)$, $x_i, y_i \in \mathcal{J}_i$, 有

$$x \circ y = (x_1 \circ y_1, x_2 \circ y_2, \dots, x_m \circ y_m), \quad \|x\|_F^2 = \sum \|x_i\|_F^2. \quad (2.2)$$

由此可得

$$\Pi_{\mathcal{K}} = (\Pi_{\mathcal{K}_1}, \Pi_{\mathcal{K}_2}, \dots, \Pi_{\mathcal{K}_m}), \quad (2.3)$$

其中, $\Pi_{\mathcal{K}_i}$ 是 $\mathcal{K}_i$ 上的投影映射.

3 对称锥规划与宽邻域

设 Euclid Jordan 代数 $\mathcal{J}$ 的维数为 n , 秩为 r , 平方锥为 $\mathcal{K}$, 对称锥规划的原问题和对偶问题分别为

$$(P) \quad \min \langle c, x \rangle, \quad \text{s.t.} \quad Ax = b, \quad x \in \mathcal{K}, \quad (3.1)$$

$$(D) \quad \max \langle b, y \rangle, \quad \text{s.t.} \quad A^*y + s = c, \quad s \in \mathcal{K}, \quad y \in \mathbb{R}^m, \quad (3.2)$$

其中, $c \in \mathcal{J}$, $b \in \mathbb{R}^m$. A 是从 $\mathcal{J}$ 映到 $\mathbb{R}^m$ 的线性算子, A^* 是 A 的共轭算子, 即 $\langle x, A^*y \rangle = \langle Ax, y \rangle$, $\forall x \in \mathcal{J}, y \in \mathbb{R}^m$. 记原 - 对偶问题 (P) 和 (D) 的 (原 - 对偶) 可行集和 (原 - 对偶) 可行内点集分别为

$$\mathcal{F} := \{(x, y, s) \in \mathcal{K} \times \mathbb{R}^m \times \mathcal{K} : Ax = b, A^*y + s = c\},$$

$$\mathcal{F}^0 := \{(x, y, s) \in \text{int } \mathcal{K} \times \mathbb{R}^m \times \text{int } \mathcal{K} : Ax = b, A^*y + s = c\}.$$

若 $(x, y, s) \in \mathcal{F}$, 则称 (x, y, s) 是可行点; 若 $(x, y, s) \in \mathcal{F}^0$, 则称 (x, y, s) 是严格可行点. 本文假设 A 是满射, 且 $\mathcal{F}^0 \neq \emptyset$. 在这个假设下, x^* 和 (y^*, s^*) 是原问题 (P) 和对偶问题 (D) 的最优解当且仅当它们满足下面最优系统^[4, 命题 2.1]:

$$\begin{aligned} Ax &= b, \quad x \in \mathcal{K}, \\ A^*y + s &= c, \quad s \in \mathcal{K}, \quad y \in \mathbb{R}^m, \\ x \circ s &= 0, \end{aligned} \quad (3.3)$$

其中, 最后的等式称为互补条件. 对 $\mu > 0$, 把 (3.3) 中的互补条件松弛为 $x \circ s = \mu e$, 得到松弛系统

$$\begin{aligned} Ax &= b, \quad x \in \mathcal{K}, \\ A^*y + s &= c, \quad s \in \mathcal{K}, \quad y \in \mathbb{R}^m, \\ x \circ s &= \mu e. \end{aligned} \quad (3.4)$$

若 A 是满射, 且 $\mathcal{F}^0 \neq \emptyset$, 则对任意的 $\mu > 0$, 松弛系统 (3.4) 有唯一解, 解的集合构成中心路径. 当 μ 趋近于零时, 中心路径的极限存在且为原 - 对偶问题 (P) 和 (D) 的最优解^[4].

引理 3.1 设 $x, y \in \mathcal{J}$, 则 $\|(x+y)^+\|_F \leq \|x^+ + y^+\|_F \leq \|x^+\|_F + \|y^+\|_F$.

证明 设 $x+y$ 具有谱分解 $x+y = \lambda_1 c_1 + \lambda_2 c_2 + \cdots + \lambda_r c_r$, 其中, $\{c_1, c_2, \dots, c_r\}$ 是 Jordan 标架, 特征值满足 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq 0 \geq \lambda_{k+1} \geq \lambda_{k+2} \geq \cdots \geq \lambda_r$. 由定义可得 $(x+y)^+ = \lambda_1 c_1 + \lambda_2 c_2 + \cdots + \lambda_k c_k$. 因此,

$$\|(x+y)^+\|_F^2 = \lambda_1^2 + \lambda_2^2 + \cdots + \lambda_k^2.$$

另一方面, 设 $a = x^+ + y^+$ 关于 $\{c_1, c_2, \dots, c_r\}$ 的 Peirce 分解为 $a = \sum_{i=1}^r a_i c_i + \sum_{i < j} a_{ij}$, 其中 $a_i \in \mathbb{R}$, $a_{ij} \in \mathcal{J}_{ij}$. 则, 对于 $1 \leq i \leq k$, 有

$$a_i = \langle x^+ + y^+, c_i \rangle \geq \langle x + y, c_i \rangle = \lambda_i \geq 0,$$

其中, 第一个不等式由 $x^+ + y^+ = x + y + (-x^- - y^-) \succeq x + y$ 以及 $\mathcal{K}$ 是自对偶锥可得.

利用 Peirce 空间 $\mathcal{J}_{ij}$ 的正交性可得

$$\|a\|_F^2 = \left\| \sum_{i=1}^r a_i c_i + \sum_{i < j} a_{ij} \right\|_F^2 = \left\| \sum_{i=1}^r a_i c_i \right\|_F^2 + \left\| \sum_{i < j} a_{ij} \right\|_F^2,$$

故

$$\|x^+ + y^+\|_F^2 \geq \left\| \sum_{i=1}^r a_i c_i \right\|_F^2 = \sum_{i=1}^r a_i^2 \geq \sum_{i=1}^k a_i^2 \geq \sum_{i=1}^k \lambda_i^2 = \|(x+y)^+\|_F^2,$$

由此可证第一个不等式, 第二个不等式由范数的三角不等式即证. $\square$

宽邻域内点算法通常把迭代点限制在中心路径的 $-\infty$ 邻域里, 其定义为

$$\mathcal{N}_{\infty}^-(1-\tau) := \{(x, y, s) \in \mathcal{F}^0 : \lambda_{\min}(Q_{x^{1/2}}s) \geq \tau\mu\},$$

其中, $\tau \in (0, 1)$ 为参数, $\mu = \langle x, s \rangle / r$. 宽邻域 $\mathcal{N}_{\infty}^-(1-\tau)$ 可以等价的写成下面形式:

$$\mathcal{N}_{\infty}^-(1-\tau) = \{(x, y, s) \in \mathcal{F}^0 : \|(\tau\mu e - Q_{x^{1/2}}s)^+\|_F = 0\}.$$

因此, 我们定义对称锥上的宽邻域

$$\mathcal{N}(\tau, \beta) := \{(x, y, s) \in \mathcal{F}^0 : \|(\tau\mu e - Q_{x^{1/2}}s)^+\|_F \leq \beta\tau\mu\}, \quad (3.5)$$

其中, $\tau, \beta \in (0, 1)$ 是邻域参数. 容易验证下面包含关系成立, 其表明 $\mathcal{N}(\tau, \beta)$ 是一个宽邻域,

$$\mathcal{N}_{\infty}^-(1-\tau) \subseteq \mathcal{N}(\tau, \beta) \subseteq \mathcal{N}_{\infty}^-(1-(1-\beta)\tau), \quad \forall \tau, \beta \in (0, 1). \quad (3.6)$$

由命题 2.2 可知, $Q_{x^{1/2}}s$ 和 $Q_{s^{1/2}}x$ 有相同的谱, 故邻域 $\mathcal{N}_{\infty}^-(1-\tau)$ 和 $\mathcal{N}(\tau, \beta)$ 关于 x 和 s 对称.

设 $x \in \mathcal{J}$ 的谱分解为 $x = \lambda_1 c_1 + \lambda_2 c_2 + \cdots + \lambda_r c_r$, 其中, 实数 $\lambda_1, \lambda_2, \dots, \lambda_r$ 为 x 的谱特征值 (为了与四元数矩阵的左、右特征值相区别, 本节称上述特征值为谱特征值, 并记作 $\sigma_{sp}(x) := \{\lambda_1, \lambda_2, \dots, \lambda_r\}$). 本节的主要内容是证明下面的定理, 它在算法的复杂性分析中起到关键作用.

定理 3.1 设 $\mathcal{J}$ 是 Euclid Jordan 代数, $\mathcal{K}$ 是它的平方锥, 则对任意的 $x, y \in \text{int } \mathcal{K}$ 和 $\tau, \mu > 0$, 下面不等式成立:

$$\|(\tau\mu e - w)^+\|_F \leq \|(\tau\mu e - x \circ y)^+\|_F, \quad (3.7)$$

其中, $w = Q_{x^{1/2}}y$.

我们首先考察四元数和四元数矩阵, 分别记作 $\mathcal{H}$ 和 $\mathcal{H}^{m \times n}$. 关于四元数的详细内容可参见文献 [16–18]. 类似于复数矩阵, 对于四元数矩阵 $A = [a_{ij}] \in \mathcal{H}^{m \times n}$, 定义共轭 $\bar{A} = [\bar{a}_{ij}]$, 转置 $A^T = [a_{ji}] \in \mathcal{H}^{n \times m}$, 共轭转置 $A^* = (\bar{A})^T$. 若方阵 $A \in \mathcal{H}^{n \times n}$ 满足 $A^* = A$, 则称为 Hermite 矩阵; 若满足 $A^*A = I$, 则称为酉矩阵; 若存在 $B \in \mathcal{H}^{n \times n}$ 使得 $AB = BA = I$, 则称 A 是可逆的; 若存在可逆的四元数矩阵 P 使得 $P^{-1}AP = B$, 则称 A 和 B 相似.

设 $A, B \in \mathcal{H}^{n \times n}$, 则有下面一些性质 (参见文献 [18]): $(AB)^* = B^*A^*$; 若 A 和 B 都可逆, 则 $(AB)^{-1} = B^{-1}A^{-1}$; 若 A 可逆, 则 $(A^*)^{-1} = (A^{-1})^*$; 若 $AB = I$, 则 $BA = I$. 对于 $x, y \in \mathcal{H}$, 一般情况下, $xy \neq yx$. 若存在非零四元数 u 使得 $u^{-1}xu = y$, 则称 x 和 y 是相似的. 若存在非零的向量 $x \neq 0$ 使得 $Ax = \lambda x$ ($Ax = x\lambda$), 则称 λ 是 A 的左 (右) 特征值. 称集合 $\{\lambda \in \mathcal{H} : Ax = \lambda x, x \neq 0\}$ 为 A 的左谱, 记作 $\sigma_l(A)$; 类似地, 定义 A 的右谱, 记作 $\sigma_r(A)$. 下面给出几个关于四元数矩阵的左、右特征值以及三角化的结果, 它们的证明可参见相应的文献.

引理 3.2 [16, 定理 13] 相似矩阵有相同的右特征值.

引理 3.3 [18, 定理 6.1] 设 $A \in \mathcal{H}^{n \times n}$, 则存在酉矩阵 U 使得 U^*AU 是上三角矩阵.

引理 3.4 [18, 定理 6.2] 若 A 是上三角矩阵, 则 A 的每一个对角元素都是它的右特征值; 相反地, A 的每一个右特征值都相似于它的一个对角元素.

引理 3.5 ^[17, 命题 3.8] 若 A 是 Hermite 矩阵, 则 A 的所有右特征值都是实数; 而且, λ 是 A 的谱特征值当且仅当 λ 是实数, 且 λ 是 A 的左 (右) 特征值.

矩阵理论中的 Schur 受控定理可以推广到简单 Jordan 代数.

引理 3.6 ^[19, 推论 4.6] 设 $\mathcal{J}$ 是简单 Jordan 代数, $\{c_1, c_2, \dots, c_r\}$ 是一个 Jordan 标架, 且 $a \in \mathcal{J}$ 关于 $\{c_1, c_2, \dots, c_r\}$ 的 Peirce 分解为 $a = \sum_{i=1}^r a_i c_i + \sum_{i < j} a_{ij}$. 则有 $\sum_{i=1}^k a_i^\downarrow \leq \sum_{i=1}^k \lambda_i^\downarrow(a)$, $\forall k = 1, 2, \dots, r$, 并且当 $k = r$ 时等式成立, 其中, $\lambda_1^\downarrow(a), \lambda_2^\downarrow(a), \dots, \lambda_r^\downarrow(a)$ 表示按降序排列.

下面分别考虑 5 种基本的简单 Jordan 代数: 实对称矩阵 $\text{Sym}(\mathbb{R}^{n \times n})$ 、复 Hermite 矩阵 $\text{Herm}(\mathbb{C}^{n \times n})$ 、四元数 Hermite 矩阵 $\text{Herm}(\mathcal{H}^{n \times n})$ 、二次型代数 $\mathcal{E}_{n+1}$ 和 3×3 八元数 Hermite 矩阵 $\text{Herm}(\mathcal{O}^{3 \times 3})$.

引理 3.7 设 $\mathcal{J} = \text{Herm}(\mathcal{H}^{n \times n})$, 则对 $X, Y \in \mathcal{J}$, $X, Y \succ 0$, $w = Q_{X^{1/2}} Y = X^{1/2} Y X^{1/2}$, 存在矩阵 $D \in \mathcal{J}$, 使得 D 和 $X \circ Y$ 有相同的谱特征值, 且 D 的对角元素与 w 的谱特征值相同.

证明 由引理 3.3 可知, 存在酉矩阵 $U \in \mathcal{H}^{n \times n}$ 使得 $U^* X Y U = T$, 其中, T 是上三角矩阵, 则

$$T = U^* X Y U = U^* X^{1/2} (X^{1/2} Y X^{1/2}) X^{-1/2} U = (X^{-1/2} U)^{-1} w (X^{-1/2} U),$$

即 T 和 w 相似. 由引理 3.5 可知, w 的每一个右特征值都是实数, 且

$$\lambda \in \sigma_{sp}(w) \Leftrightarrow \lambda \in \sigma_r(w). \quad (3.8)$$

由于相似矩阵有相同的右特征值, 则

$$\lambda \in \sigma_r(w) \Leftrightarrow \lambda \in \sigma_r(T). \quad (3.9)$$

因此, T 的右特征值都是实数. 又由引理 3.4 得

$$\lambda \in \sigma_r(T) \Leftrightarrow \lambda = t_{ii}, \quad (3.10)$$

其中, t_{ii} 是 T 的对角元素.

令 $D = \frac{1}{2}(T + T^*)$, 则 D 是 Hermite 矩阵, 且有

$$X \circ Y = \frac{1}{2}(XY + (XY)^*) = \frac{1}{2}U(T + T^*)U^* = UDU^*,$$

即 $X \circ Y$ 与 D 相似. 又由引理 3.2 和 3.5 可知, 它们具有相同的谱特征值. 同时, 由等价关系 (3.8)–(3.10) 可知, λ 是 w 的谱特征值当且仅当 λ 是 D 的对角元素. $\square$

对于复 Hermite 矩阵 $A \in \mathbb{C}^{n \times n}$, 其左特征值和右特征值都等同于谱特征值, 因此, 在不引起混淆时下面简称为特征值.

引理 3.8 设 $\mathcal{J} = \text{Herm}(\mathbb{C}^{n \times n})$ 或 $\text{Sym}(\mathbb{R}^{n \times n})$, 则对 $X, Y \in \mathcal{J}$, $X, Y \succ 0$,

$$w = Q_{X^{1/2}} Y = X^{1/2} Y X^{1/2},$$

存在矩阵 $D \in \mathcal{J}$, 使得 D 和 $X \circ Y$ 有相同的特征值, 且 D 的对角元素与 w 的特征值相同.

证明 首先, 由 $X, Y \succ 0$ 和 $XY = X^{1/2} X^{1/2} Y X^{1/2} X^{-1/2} = X^{1/2} w X^{-1/2}$ 可知, XY 的所有特征值都是实数. 由 Schur 酉三角化定理 ^[20, 定理 2.3.1], 存在酉矩阵 $U \in \mathbb{C}^{n \times n}$ 使得 $U^* X Y U = T = [t_{ij}]$ 是上三角矩阵, 并且其对角元素满足 $t_{ii}^\downarrow = \lambda_i^\downarrow(XY) = \lambda_i^\downarrow(w)$, $i = 1, 2, \dots, n$. 此外, 若 $X, Y \in \text{Sym}(\mathbb{R}^{n \times n})$, 则 U 可以取成实的正交矩阵. 令 $D = \frac{1}{2}(T + T^*)$, 则 $d_{ii} = t_{ii}$, $i = 1, 2, \dots, n$, 且

$$X \circ Y = \frac{1}{2}(XY + (XY)^*) = \frac{1}{2}U(T + T^*)U^* = UDU^*,$$

故 $\lambda_i^\downarrow(D) = \lambda_i^\downarrow(X \circ Y)$, $i = 1, 2, \dots, n$. □

引理 3.9 设 $\mathcal{J} = \text{Sym}(\mathbb{R}^{n \times n})$, $\text{Herm}(\mathbb{C}^{n \times n})$, $\text{Herm}(\mathcal{H}^{n \times n})$, $\text{Herm}(\mathcal{O}^{3 \times 3})$ 或 $\mathcal{E}_{n+1}$, 则下式成立:

$$\sum_{i=1}^k \lambda_i^\downarrow(w) \leq \sum_{i=1}^k \lambda_i^\downarrow(X \circ Y), \quad \forall k = 1, 2, \dots, r,$$

并且当 $k = r$ 时等式成立, 其中, r 是 $\mathcal{J}$ 的秩.

证明 对于 $\mathcal{J} = \text{Sym}(\mathbb{R}^{n \times n})$, $\text{Herm}(\mathbb{C}^{n \times n})$ 或 $\text{Herm}(\mathcal{H}^{n \times n})$, 有 $r = n$ [12, 第 97 页]. 由引理 3.7 和 3.8 可知, 存在矩阵 $D = [d_{ij}] \in \mathcal{J}$ 使得

$$d_{ii}^\downarrow = \lambda_i^\downarrow(w), \quad \lambda_i^\downarrow(D) = \lambda_i^\downarrow(X \circ Y), \quad i = 1, 2, \dots, n.$$

令 $E_i = \text{Diag}(e_i)$, $i = 1, \dots, n$, 其中, e_i 表示第 i 个分量为 1 其他分量全为 0 的列向量, 则 $\{E_1, \dots, E_n\}$ 是 Jordan 标架. 设 D 的 Peirce 分解为 $D = \sum_{i=1}^n d_{ii}E_i + \sum_{i < j} D_{ij}$. 利用引理 3.6 可得

$$\sum_{i=1}^k d_{ii}^\downarrow \leq \sum_{i=1}^k \lambda_i^\downarrow(D), \quad \forall k = 1, 2, \dots, n,$$

且当 $k = n$ 时等式成立. 因此, $\sum_{i=1}^k \lambda_i^\downarrow(w) \leq \sum_{i=1}^k \lambda_i^\downarrow(X \circ Y)$, $\forall k = 1, 2, \dots, n$, 且当 $k = n$ 时等式成立.

对于 $\mathcal{J} = \text{Herm}(\mathcal{O}^{3 \times 3})$, 有 $r = 3$ [12, 推论 V.2.3], 则由引理 2.3 可得

$$\begin{aligned} \lambda_1^\downarrow(w) &\leq \lambda_1^\downarrow(X \circ Y), \quad \lambda_3^\downarrow(w) \geq \lambda_3^\downarrow(X \circ Y), \\ \lambda_1^\downarrow(w) + \lambda_2^\downarrow(w) + \lambda_3^\downarrow(w) &= \lambda_1^\downarrow(X \circ Y) + \lambda_2^\downarrow(X \circ Y) + \lambda_3^\downarrow(X \circ Y). \end{aligned}$$

进而可得

$$\lambda_1^\downarrow(w) + \lambda_2^\downarrow(w) \leq \lambda_1^\downarrow(X \circ Y) + \lambda_2^\downarrow(X \circ Y).$$

对于 $\mathcal{J} = \mathcal{E}_{n+1}$, 有 $r = 2$, 则同样由引理 2.3 可证. □

引理 3.10 设 $\mathcal{J} = \text{Sym}(\mathbb{R}^{n \times n})$, $\text{Herm}(\mathbb{C}^{n \times n})$, $\text{Herm}(\mathcal{H}^{n \times n})$, $\text{Herm}(\mathcal{O}^{3 \times 3})$ 或 $\mathcal{E}_{n+1}$, 则对任意的 $x, y \succ 0$ 和 $\tau, \mu > 0$, 都有 $\|(\tau\mu e - w)^+\|_F \leq \|(\tau\mu e - x \circ y)^+\|_F$, 其中, $w = Q_{x^{1/2}y}$.

证明 设 w 和 $x \circ y$ 的谱特征值满足降序排列, 即 $\lambda_1(w) \geq \dots \geq \lambda_r(w)$, $\lambda_1(x \circ y) \geq \dots \geq \lambda_r(x \circ y)$. 令 $\lambda'_i = \tau\mu - \lambda_i(w)$, $\sigma'_i = \tau\mu - \lambda_i(x \circ y)$, $i = 1, 2, \dots, r$, 则 $\lambda'_1 \leq \dots \leq \lambda'_r$, $\sigma'_1 \leq \dots \leq \sigma'_r$.

由引理 3.9 可得 $\sum_{i=1}^k \lambda'_i \geq \sum_{i=1}^k \sigma'_i$, $\forall k = 1, 2, \dots, r$, 且当 $k = r$ 时等式成立. 由 Schur 受控定理的逆定理 [20, 定理 4.3.32], 存在实对称矩阵 $A = [a_{ij}] \in \mathbb{R}^{r \times r}$ 满足

$$a_{ii} = \lambda'_i, \quad \lambda_i(A) = \sigma'_i, \quad i = 1, 2, \dots, r.$$

由 A^+ 和 A^- 的定义可得 $(A^+)_{ii} = a_{ii} - (A^-)_{ii} \geq a_{ii} = \lambda'_i$, 故 $(A^+)_{ii} \geq (\lambda'_i)^+$, $i = 1, 2, \dots, r$. 因此,

$$\|A^+\|_F^2 \geq \sum_{i=1}^r (A^+)_{ii}^2 \geq \sum_{i=1}^r ((\lambda'_i)^+)^2 = \sum_{\lambda'_i > 0} (\lambda'_i)^2 = \|(\tau\mu e - w)^+\|_F^2.$$

另一方面,

$$\|A^+\|_F^2 = \sum_{i=1}^r (\lambda_i^+(A))^2 = \sum_{i=1}^r ((\sigma'_i)^+)^2 = \sum_{\sigma'_i > 0} (\sigma'_i)^2 = \|(\tau\mu e - x \circ y)^+\|_F^2.$$

证毕. $\square$

定理 3.1 的证明 由定理 2.5 可知, 一个 Euclid Jordan 代数可以唯一的表示成一些简单 Euclid Jordan 代数的直和, 即 $\mathcal{J} = \mathcal{J}_1 \times \mathcal{J}_2 \times \cdots \times \mathcal{J}_m$, 其中, $\mathcal{J}_i$ 是简单 Euclid Jordan 代数. 设 $\mathcal{J}$ 的单位元 $e = (e_1, e_2, \dots, e_m)$, e_i 是 $\mathcal{J}_i$ 的单位元. 则有 $\tau\mu e - x \circ y = (\tau\mu e_1 - x_1 \circ y_1, \dots, \tau\mu e_m - x_m \circ y_m)$, $\tau\mu e - w = (\tau\mu e_1 - w_1, \dots, \tau\mu e_m - w_m)$, 其中, $w_i = Q_{x_i^{1/2}} y_i$. 利用 (2.2) 和 (2.3) 可得

$$\begin{aligned} \|(\tau\mu e - x \circ y)^+\|_F^2 &= \|((\tau\mu e_1 - x_1 \circ y_1)^+, \dots, (\tau\mu e_m - x_m \circ y_m)^+)\|_F^2 = \sum \|(\tau\mu e_i - x_i \circ y_i)^+\|_F^2, \\ \|(\tau\mu e - w)^+\|_F^2 &= \|((\tau\mu e_1 - w_1)^+, \dots, (\tau\mu e_m - w_m)^+)\|_F^2 = \sum \|(\tau\mu e_i - w_i)^+\|_F^2. \end{aligned}$$

由引理 3.10 和定理 2.4 可知, 对任意的 $i = 1, 2, \dots, m$ 都有

$$\|(\tau\mu e_i - w_i)^+\|_F^2 \leq \|(\tau\mu e_i - x_i \circ y_i)^+\|_F^2.$$

因此,

$$\|(\tau\mu e - w)^+\|_F^2 \leq \|(\tau\mu e - x \circ y)^+\|_F^2.$$

由此即得 (3.7). $\square$

4 路径跟踪算法及其复杂性分析

本节首先给出对称锥规划的路径跟踪算法, 然后分析算法的迭代复杂性. 下面的引理给出了互补松弛条件 $x \circ s = \mu e$ 的一个等价条件.

引理 4.1 [6, 引理 28] 设 $x, s \in \text{int } \mathcal{K}$, 且 p 可逆, 则 $x \circ s = \mu e$ 当且仅当 $Q_p x \circ Q_{p^{-1}} s = \mu e$.

因此, 给定 $p \in \text{int } \mathcal{K}$, 中心系统 (3.4) 等价于

$$\begin{aligned} \tilde{A}\tilde{x} &= b, \quad \tilde{x} \in \mathcal{K}, \\ \tilde{A}^*y + \tilde{s} &= \tilde{c}, \quad \tilde{s} \in \mathcal{K}, \quad y \in \mathbb{R}^m, \\ \tilde{x} \circ \tilde{s} &= \mu e, \end{aligned} \tag{4.1}$$

其中, $\tilde{A} = A Q_{p^{-1}}$, $\tilde{c} = Q_{p^{-1}} c$, $\tilde{x} = Q_p x$, $\tilde{s} = Q_{p^{-1}} s$. 由 (4.1) 得到的方向称为 Monteiro-Zhang 族. 记

$$\mathcal{C}(x, s) := \{p : p \in \text{int } \mathcal{K} \text{ 使得 } Q_p x \text{ 和 } Q_{p^{-1}} s \text{ 算子可交换}\},$$

$\mathcal{C}(x, s)$ 称为“可交换类”. 当取 $p = s^{1/2}$ 或 $x^{-1/2}$ 时得到 xs 和 sx 方向; 取

$$p = [Q_{x^{1/2}}(Q_{x^{1/2}} s)^{-1/2}]^{-1/2} = [Q_{s^{-1/2}}(Q_{s^{1/2}} x)^{1/2}]^{-1/2} \tag{4.2}$$

时得到 Nesterov-Todd (NT) 方向. 对于 NT 尺度 p , 容易验证 $Q_p x = Q_{p^{-1}} s$, 即 $\tilde{x} = \tilde{s}$.

令 $\tilde{w} = Q_{\tilde{x}^{1/2}} \tilde{s}$. 由于 $Q_{x^{1/2}}$ 关于内积对称, 则 $\langle x, s \rangle = \langle Q_{x^{1/2}} e, s \rangle = \langle e, Q_{x^{1/2}} s \rangle$. 又由命题 2.2 可知, w 和 $\tilde{w}$ 有相同的谱, 故 $\mu = \langle x, s \rangle / r = \langle \tilde{x}, \tilde{s} \rangle / r = \tilde{\mu}$. 由此可得邻域的尺度不变性, 即 $(x, y, s) \in \mathcal{N}(\tau, \beta)$ 当且仅当 $(\tilde{x}, y, \tilde{s})$ 包含于相应的变尺度的邻域. 我们把变尺度的邻域仍然记为 $\mathcal{N}(\tau, \beta)$.

设当前迭代点 $(\tilde{x}, y, \tilde{s}) \in \mathcal{N}(\tau, \beta)$, 定义一个与当前迭代点有关的量

$$v := (\tau\mu e - \tilde{x} \circ \tilde{s})^+ + \eta(\tau\mu e - \tilde{x} \circ \tilde{s})^- + \tilde{x} \circ \tilde{s}, \tag{4.3}$$

其中, $\tau \in (0, 1)$, $\eta = -\text{tr}((\tau\mu e - \tilde{x} \circ \tilde{s})^+)/\text{tr}((\tau\mu e - \tilde{x} \circ \tilde{s})^-)$. 直接计算可知, $\text{tr}(v)/r = \text{tr}(\tilde{x} \circ \tilde{s})/r = \mu$, 且

$$0 \leq \eta = \frac{\text{tr}(\tau\mu e - \tilde{x} \circ \tilde{s}) - \text{tr}((\tau\mu e - \tilde{x} \circ \tilde{s})^-)}{-\text{tr}((\tau\mu e - \tilde{x} \circ \tilde{s})^-)} = \frac{-(1-\tau)r\mu}{-\text{tr}((\tau\mu e - \tilde{x} \circ \tilde{s})^-)} + 1 < 1.$$

因此, $v - \tau\mu e = (\tau\mu e - \tilde{x} \circ \tilde{s})^+ + \eta(\tau\mu e - \tilde{x} \circ \tilde{s})^- + \tilde{x} \circ \tilde{s} - \tau\mu e = -(1-\eta)(\tau\mu e - \tilde{x} \circ \tilde{s})^- \succeq 0$, 故

$$v \succeq \tau\mu e. \quad (4.4)$$

相应于变尺度系统 (4.1), 算法的搜索方向 $(\Delta\tilde{x}, \Delta y, \Delta\tilde{s})$ 由下面线性系统解得,

$$\begin{aligned} \tilde{A}\Delta\tilde{x} &= 0, \\ \tilde{A}^*\Delta y + \Delta\tilde{s} &= 0, \\ \Delta\tilde{x} \circ \tilde{s} + \tilde{x} \circ \Delta\tilde{s} &= tv - \tilde{x} \circ \tilde{s}, \end{aligned} \quad (4.5)$$

其中, $t \in [0, 1]$ 是待定参数. 则新的迭代为

$$(\tilde{x}(t), y(t), \tilde{s}(t)) := (\tilde{x}, y, \tilde{s}) + (\Delta\tilde{x}, \Delta y, \Delta\tilde{s}). \quad (4.6)$$

算法 1 (邻域跟踪算法) 输入参数 $\varepsilon > 0$, $\tau, \beta \in (0, 1/3]$, 初始点 $(x^0, y^0, s^0) \in \mathcal{N}(\tau, \beta)$. 计算 $\mu_0 = \langle x^0, s^0 \rangle / r$. 置 $k := 0$.

第 1 步 若 $\mu_k \leq \varepsilon\mu_0$, 算法终止;

第 2 步 计算 NT 尺度 p ;

第 3 步 由 (4.5) 计算方向 $(\Delta\tilde{x}^k, \Delta y^k, \Delta\tilde{s}^k)$, 并求满足 $(\tilde{x}^k(t), y^k(t), \tilde{s}^k(t)) \in \mathcal{N}(\tau, \beta)$, $\forall t \in [\bar{t}_k, 1]$ 的最小值 $\bar{t}_k \geq 0$;

第 4 步 令 $(x^{k+1}, y^{k+1}, s^{k+1}) = (Q_{p^{-1}}\tilde{x}^k(\bar{t}_k), y^k(\bar{t}_k), Q_p\tilde{s}^k(\bar{t}_k))$, $\mu_{k+1} = \bar{t}_k\mu_k$. 置 $k := k + 1$, 转第 1 步.

下面先证明一些技术引理, 然后给出算法的多项式收敛性定理. 为简化起见, 在不引起混淆时省略迭代指标 k . 首先, 通过简单计算可得下面的两个引理.

引理 4.2 下面结论成立:

- (1) $\langle \Delta\tilde{x}, \Delta\tilde{s} \rangle = 0$;
- (2) $\tilde{x}(t) \circ \tilde{s}(t) = tv + \Delta\tilde{x} \circ \Delta\tilde{s}$;
- (3) $\mu(t) := \langle \tilde{x}(t), \tilde{s}(t) \rangle / r = t\mu$.

引理 4.3 设 $(\tilde{x}, y, \tilde{s}) \in \mathcal{N}(\tau, \beta)$, 则 $\text{tr}((\tau\mu e - \tilde{x} \circ \tilde{s})^-) \leq -(1-\tau)r\mu$, $\text{tr}((\tau\mu e - \tilde{x} \circ \tilde{s})^+) \leq \sqrt{r}\beta\tau\mu$.

引理 4.4 设 $(\tilde{x}, y, \tilde{s}) \in \mathcal{N}(\tau, \beta)$, p 为 NT 尺度. 若 $\tau, \beta \in (0, 1/3]$, 则

$$\|(L_{\tilde{x}}L_{\tilde{s}})^{-1/2}(\tau\mu e - \tilde{x} \circ \tilde{s})^+\|_F^2 \leq \beta\tau\mu/2, \quad \|\eta(L_{\tilde{x}}L_{\tilde{s}})^{-1/2}(\tau\mu e - \tilde{x} \circ \tilde{s})^-\|_F^2 \leq \beta\tau\mu/4.$$

证明 设 $\{c_1, c_2, \dots, c_r\}$ 是 Jordan 标架, 且有谱分解 $\tilde{x} = \tilde{s} = \sum_{i=1}^r \lambda_i c_i$, 则 $\tilde{x} \circ \tilde{s} = \sum_{i=1}^r \lambda_i^2 c_i$. 因此,

$$\begin{aligned} \|(L_{\tilde{x}}L_{\tilde{s}})^{-1/2}(\tau\mu e - \tilde{x} \circ \tilde{s})^+\|_F^2 &= \left\| L_{\tilde{x}}^{-1} \sum (\tau\mu - \lambda_i^2)^+ c_i \right\|_F^2 \\ &= \left\langle L_{\tilde{x}}^{-1} \sum (\tau\mu - \lambda_i^2)^+ c_i, L_{\tilde{x}}^{-1} \sum (\tau\mu - \lambda_i^2)^+ c_i \right\rangle \end{aligned}$$

$$\begin{aligned}
 &= \left\langle \sum (\tau\mu - \lambda_i^2)^+ c_i / \lambda_i, \sum (\tau\mu - \lambda_i^2)^+ c_i / \lambda_i \right\rangle \\
 &= \sum [(\tau\mu - \lambda_i^2)^+]^2 / \lambda_i^2 \\
 &\leq (\beta\tau\mu)^2 / ((1 - \beta)\tau\mu) \leq \beta\tau\mu/2,
 \end{aligned} \tag{4.7}$$

其中, 第三个等式由 $L_{\tilde{x}}^{-1}c_i = c_i/\lambda_i$ 得到, 第一个不等式由 $(\tilde{x}, y, \tilde{s}) \in \mathcal{N}(\tau, \beta)$ 得到, 最后的不等式用到 $\beta \leq 1/3$. 同理可得

$$\|(L_{\tilde{x}}L_{\tilde{s}})^{-1/2}(\tau\mu e - \tilde{x} \circ \tilde{s})^-\|_F^2 = \sum [(\tau\mu - \lambda_i^2)^-]^2 / \lambda_i^2 = \sum [(\lambda_i - \tau\mu/\lambda_i)^+]^2 \leq \sum \lambda_i^2 = r\mu.$$

利用引理 4.3 和 $\tau, \beta \in (0, 1/3]$ 可得

$$\eta = \frac{\text{tr}((\tau\mu e - \tilde{x} \circ \tilde{s})^+)}{-\text{tr}((\tau\mu e - \tilde{x} \circ \tilde{s})^-)} \leq \frac{\sqrt{r}\beta\tau\mu}{(1 - \tau)r\mu} = \frac{\beta\tau}{(1 - \tau)\sqrt{r}} \leq \frac{\sqrt{\beta\tau}}{2\sqrt{r}}.$$

故

$$\|\eta(L_{\tilde{x}}L_{\tilde{s}})^{-1/2}(\tau\mu e - \tilde{x} \circ \tilde{s})^-\|_F^2 = \eta^2 \|(L_{\tilde{x}}L_{\tilde{s}})^{-1/2}(\tau\mu e - \tilde{x} \circ \tilde{s})^-\|_F^2 \leq \beta\tau\mu/4.$$

引理 4.4 证毕. □

下面的引理给出了二次项 $\Delta\tilde{x} \circ \Delta\tilde{s}$ 的一个界.

引理 4.5 设 $(\tilde{x}, y, \tilde{s}) \in \mathcal{N}(\tau, \beta)$, p 为 NT 尺度, 且 $\tau, \beta \in (0, 1/3]$. 若 $1 - \sqrt{\beta\tau/r} \leq t \leq 1$, 则

$$\|\Delta\tilde{x} \circ \Delta\tilde{s}\|_F \leq t\beta\tau\mu.$$

证明 由 $\tilde{x} = \tilde{s}$ 可得 $L_{\tilde{x}} = L_{\tilde{s}}$. 则方程组 (4.5) 中最后一个方程可化为 $L_{\tilde{x}}(\Delta\tilde{x} + \Delta\tilde{s}) = tv - \tilde{x}^2$. 两边同时乘以 $L_{\tilde{x}}^{-1}$, 并取范数然后平方可得

$$\begin{aligned}
 \|\Delta\tilde{x}\|_F^2 + \|\Delta\tilde{s}\|_F^2 + 2\langle \Delta\tilde{x}, \Delta\tilde{s} \rangle &= \|L_{\tilde{x}}^{-1}(tv - \tilde{x}^2)\|_F^2 \\
 &= \|tL_{\tilde{x}}^{-1}(v - \tilde{x}^2) - (1 - t)\tilde{x}\|_F^2 \\
 &= t^2\|L_{\tilde{x}}^{-1}(v - \tilde{x}^2)\|_F^2 + (1 - t)^2\|\tilde{x}\|_F^2,
 \end{aligned} \tag{4.8}$$

其中, 最后的等式用到正交性 $\langle L_{\tilde{x}}^{-1}(v - \tilde{x}^2), \tilde{x} \rangle = \langle v - \tilde{x}^2, e \rangle = \text{tr}(v - \tilde{x}^2) = 0$. 由引理 4.4 可得

$$\begin{aligned}
 \|L_{\tilde{x}}^{-1}(v - \tilde{x}^2)\|_F^2 &= \|L_{\tilde{x}}^{-1}(\tau\mu e - \tilde{x}^2)^+\|_F^2 + \|\eta L_{\tilde{x}}^{-1}(\tau\mu e - \tilde{x}^2)^-\|_F^2 \\
 &\quad + 2\eta\langle L_{\tilde{x}}^{-1}(\tau\mu e - \tilde{x}^2)^+, L_{\tilde{x}}^{-1}(\tau\mu e - \tilde{x}^2)^- \rangle \\
 &\leq \beta\tau\mu/2 + \beta\tau\mu/4 = 3\beta\tau\mu/4,
 \end{aligned} \tag{4.9}$$

其中, 由 (4.7) 的证明过程可知, 第一个等式右端的最后一项等于零.

综合 (4.8) 和 (4.9) 以及正交性 $\langle \Delta\tilde{x}, \Delta\tilde{s} \rangle = 0$ 可得

$$\begin{aligned}
 \|\Delta\tilde{x}\|_F^2 + \|\Delta\tilde{s}\|_F^2 &\leq 3t^2\beta\tau\mu/4 + (1 - t)^2r\mu \\
 &\leq 3t^2\beta\tau\mu/4 + (\sqrt{\beta\tau/r})^2r\mu \\
 &= t(3t/4 + 1/t)\beta\tau\mu \leq 2t\beta\tau\mu,
 \end{aligned}$$

其中, 第二个不等式由 $1 - \sqrt{\beta\tau/r} \leq t \leq 1$ 得到, 最后的不等式用到 $t \geq 2/3$, 以及 $3t/4 + 1/t$ 在区间 $(0, 1]$ 上单调下降. 因此, 由引理 2.2 即得 $\|\Delta\tilde{x} \circ \Delta\tilde{s}\|_F \leq \|\Delta\tilde{x}\|_F \|\Delta\tilde{s}\|_F \leq \frac{1}{2}(\|\Delta\tilde{x}\|_F^2 + \|\Delta\tilde{s}\|_F^2) \leq t\beta\tau\mu$. □

下面的引理给出了迭代点属于邻域 $\mathcal{N}(\tau, \beta)$ 的一个充分条件.

引理 4.6 设 $(\tilde{x}, y, \tilde{s}) \in \mathcal{N}(\tau, \beta)$, p 为 NT 尺度, 且 $\tau, \beta \in (0, 1/3]$. 若 $1 - \sqrt{\beta\tau/r} \leq t \leq 1$, 则

$$(\tilde{x}(t), y(t), \tilde{s}(t)) \in \mathcal{N}(\tau, \beta).$$

证明 由引理 4.5, 4.2(2) 和 4.2(3) 可得

$$\begin{aligned} \|(\tau\mu(t)e - \tilde{x}(t) \circ \tilde{s}(t))^+\|_F &= \|(t\tau\mu e - tv - \Delta\tilde{x} \circ \Delta\tilde{s})^+\|_F \\ &\leq \|t(\tau\mu e - v)^+\|_F + \|(-\Delta\tilde{x} \circ \Delta\tilde{s})^+\|_F \\ &\leq \|\Delta\tilde{x} \circ \Delta\tilde{s}\|_F \leq t\beta\tau\mu = \beta\tau\mu(t), \end{aligned}$$

其中, 第一个不等式由引理 3.1 得到, 第二个不等式由 (4.4) 得到. 由邻域的包含关系 (3.6) 可得

$$\lambda(\tilde{x}(t) \circ \tilde{s}(t)) \geq (1 - \beta)\tau\mu(t) > 0.$$

故由引理 2.1 可知, $\det(\tilde{x}(t)) \neq 0$ 且 $\det(\tilde{s}(t)) \neq 0$. 因此, 利用 $\tilde{x}, \tilde{s} \succ 0$ 和连续性可得 $\tilde{x}(t) \succ 0$, $\tilde{s}(t) \succ 0$. 进一步, 由定理 3.1 可得

$$\|(\tau\mu(t)e - \tilde{w}(t))^+\|_F \leq \|(\tau\mu(t)e - \tilde{x}(t) \circ \tilde{s}(t))^+\|_F \leq \beta\tau\mu(t),$$

其中, $\tilde{w}(t) = Q_{\tilde{x}(t)^{1/2}}\tilde{s}(t)$. 综上可得 $(\tilde{x}(t), y(t), \tilde{s}(t)) \in \mathcal{N}(\tau, \beta)$. □

在实际计算中, 算法 1 的第 3 步的 $\bar{t}$ 并不需要精确求解. 由引理 4.6, $\bar{t}$ 的计算可采用下面方式: 令 $\hat{t} = 1 - \sqrt{\beta\tau/r}$, 在区间 $[0, \hat{t}]$ 中用二分法, 求满足 $(\tilde{x}(\bar{t}), y(\bar{t}), \tilde{s}(\bar{t})) \in \mathcal{N}(\tau, \beta)$ 的最小值.

下面的定理给出了算法 1 终止时的一个迭代界.

定理 4.1 算法 1 至多在迭代 $\mathcal{O}(\sqrt{r} \log \varepsilon^{-1})$ 次后停止.

证明 由引理 4.6 可知, 在每次迭代中都有 $\bar{t}_k \leq 1 - \sqrt{\beta\tau/r}$. 因此,

$$\mu_k = \bar{t}_k \mu_{k-1} \leq (1 - \sqrt{\beta\tau/r})^k \mu_0.$$

由此即证. □

需要指出的是, 若 t 是 $\bar{t}$ 的近似值, 且满足 $t \leq 1 - \sqrt{\beta\tau/r}$, 则算法的迭代复杂性仍为 $\mathcal{O}(\sqrt{r} \log \varepsilon^{-1})$.

5 结论

基于 Nesterov-Todd (NT) 方向, 本文把线性规划的邻域跟踪算法推广到了对称锥规划. 给定邻域中一个初始可行点, 算法的迭代复杂性为 $\mathcal{O}(\sqrt{r} \log \varepsilon^{-1})$, 这是对称锥上内点算法最好的复杂性结果. 然而, 对称锥规划的邻域跟踪算法中仍然有一些问题有待研究. 本文算法基于 NT 方向, 对于其他方向, 如搜索方向的“可交换类”, 邻域跟踪算法是否具有多项式收敛性值得进一步研究. 另一方面, 如何设计邻域跟踪算法的不可行算法, 从而改进不可行内点算法的复杂性也是我们下一步研究内容.

致谢 衷心感谢评阅人给予的评价和指导意义.

参考文献

- 1 Nesterov Y E, Todd M J. Self-scaled barriers and interior-point methods for convex programming. *Math Oper Res*, 1997, 22: 1–42
- 2 Nesterov Y E, Todd M J. Primal-dual interior-point methods for self-scaled cones. *SIAM J Optim*, 1998, 8: 324–364
- 3 Faybusovich L. Euclidean Jordan algebras and interior-point algorithms. *Positivity*, 1997, 1: 331–357
- 4 Faybusovich L. Linear systems in Jordan algebras and primal-dual interior-point algorithms. *J Comput Appl Math*, 1997, 86: 149–175
- 5 Yoshise A. Interior point trajectories and a homogeneous model for nonlinear complementarity problems over symmetric cones. *SIAM J Optim*, 2006, 17: 1129–1153
- 6 Schmieta S H, Alizadeh F. Extension of primal-dual interior-point algorithm to symmetric cones. *Math Program*, 2003, 96: 409–438
- 7 Monteiro R D C, Zhang Y. A unified analysis for a class of long-step primal-dual path-following interior-point algorithms for semidefinite programming. *Math Program*, 1998, 81: 281–299
- 8 艾文宝. 线性规划的邻域跟踪算法. *中国科学 A 辑*, 2004, 34: 40–47
- 9 Ai W, Zhang S. An $O(\sqrt{n}L)$ iteration primal-dual path-following method, based on wide neighborhoods and large updates, for monotone LCP. *SIAM J Optim*, 2005, 16: 400–417
- 10 Feng Z, Fang L. A wide neighbourhood interior-point method with $O(\sqrt{n}L)$ iteration-complexity bound for semidefinite programming. *Optimization*, 2010, 59: 1235–1246
- 11 Li Y, Terlaky T. A new class of large neighborhood path-following interior point algorithms for semidefinite optimization with $O(\sqrt{n} \log \frac{\text{Tr}(X^0 S^0)}{\varepsilon})$ iteration complexity. *SIAM J Optim*, 2010, 20: 2853–2875
- 12 Faraut J, Korányi A. *Analysis on Symmetric Cones*. Oxford: Oxford University Press, 1994
- 13 Gu G, Zangjabad M, Roos C. Full Nesterov-Todd step infeasible interior-point method for symmetric optimization. *European J Oper Res*, 2011, 214: 473–484
- 14 Rangarajan B K. Polynomial convergence of infeasible-interior-point methods over symmetric cones. *SIAM J Optim*, 2006, 16: 1211–1229
- 15 Gowda M S, Sznajder R, Tao J. Some P -properties for linear transformations on Euclidean Jordan algebras. *Linear Algebra Appl*, 2004, 393: 203–232
- 16 Brenner J L. Matrices of quaternions. *Pacific J Math*, 1951, 1: 329–335
- 17 Farenick D F, Pidkoich B A F. The spectral theorem in quaternions. *Linear Algebra Appl*, 2003, 371: 75–102
- 18 Zhang F. Quaternions and matrices of quaternions. *Linear Algebra Appl*, 1997, 251: 21–57
- 19 Gowda M S, Tao J. The Cauchy interlacing theorem in simple Euclidean Jordan algebras and some consequences. *Linear Multilinear Algebra*, 2011, 59: 65–86
- 20 Horn R A, Johnson C R. *Matrix Analysis*. Cambridge: Cambridge University Press, 1990

Neighborhood-following algorithms for symmetric cone programming

LIU ChangHe, LIU HongWei & SHANG YouLin

Abstract The neighborhood-following algorithm of linear programming, developed by Ai, is extended to symmetric cones. We define a new wide neighborhood $\mathcal{N}(\tau, \beta)$ and prove a key property which plays a crucial role in the complexity analysis. Starting with an initial point in $\mathcal{N}(\tau, \beta)$, the complexity bound is $\mathcal{O}(\sqrt{r} \log \varepsilon^{-1})$ for the Nesterov-Todd (NT) direction, where r is the rank of the associated Euclidean Jordan algebra and $\varepsilon > 0$ is the required precision. Then, we get the best complexity bound of wide neighborhood interior-point algorithm for symmetric cone programming.

Keywords symmetric cone programming, Euclidean Jordan algebra, neighborhood-following algorithm, wide neighborhood, interior-point methods, polynomial complexity

MSC(2010) 65K05, 90C05, 90C25, 90C51

doi: 10.1360/012011-577