中国科学 A 辑: 数学 2009 年 第 39 卷 第 1 期: 1～14
www.scichina.com math.scichina.com
《中国科学》杂志社 SCIENCE IN CHINA PRESS

一个求解对称锥互补问题的具有非单调线搜索的光滑算法的全局收敛性

黄正海 ①*, 胡胜龙 ①, 韩继业 ②

① 天津大学理学院数学系, 天津 300072
② 中国科学院数学与系统科学研究院, 北京 100080
* E-mail: huangzhenghai@tju.edu.cn

收稿日期: 2008-02-19; 接受日期: 2008-06-05
国家自然科学基金 (批准号: 10571134, 10671010) 和天津市自然科学基金 (批准号: 07JCYBJC05200) 资助项目

摘要 提出一个求解单调对称锥互补问题 (简记为 SCCP) 的具有非单调线搜索的光滑算法, 并且证明提出的算法在所求解问题的解集非空的条件下是全局收敛的. 这样的假设比现有的大多数求解对称锥优化问题的算法中所使用的假设都要弱. 最后在适当的条件下, 证明所提算法得到的解是一个极大互补解.

关键词 互补问题 对称锥 欧氏 Jordon 代数 光滑算法 全局收敛

MSC(2000) 主题分类 **90C25, 90C33**

1 引言

光滑算法已经成功地运用到求解很多的优化问题, 包括线性规划、互补和变分不等式问题、具有平衡约束的数学规划、二阶锥规划、半定规划等. 在适当的假设下, 已经证明很多光滑算法具有全局收敛性和局部超线性收敛性. 以前使用的很多假设都蕴含了所考虑问题的解集是非空有界的 [1]. 最近, 求解互补问题的光滑算法的全局收敛性假设被减弱到所考虑问题的解集只是非空的 [2,3]. 此外, 光滑算法具有很好的数值效果. 为得到更好的数值计算结果, 在光滑算法的实际数值计算中, 已经采用了一些非单调线搜索技巧 [4], 然而缺乏相应的理论分析.

最近, 对称锥优化问题已成为优化领域的研究热点之一. 称一个集合 $K \subset V$ 是一个对称锥, 如果对每个 $x, s \in \mathrm{int}(K)$ (K 的内部), 存在一个可逆的线性变换 $\mathcal{A}$, 使得 $\mathcal{A}(x) = s$ 且 $\mathcal{A}(K) = K$, 其中 V 是一个具有内积的有限维的线性空间. 众所周知, 非负卦限 ($\Re^n_+$)、二阶锥 ($\mathcal{L}^n_+$)、对称半正定矩阵锥 ($\mathcal{S}^n_+$) 都是对称锥的特例. 因此, 对称锥优化问题为许多优化问题提供了统一的框架, 包括传统的优化问题 (即 $\Re^n_+$ 上的优化问题)、二阶锥规划、二阶锥互补问题、半定规划和半定互补问题.

除了一些理论的分析 [5−10], 学者们对如何求解这类问题已做了大量的努力和工作, 主要是把内点算法推广到求解某些对称锥优化问题 [11−15]. 最近, 一些光滑算法被提出来求解对称锥线性规划 [16] 和对称锥互补问题 [17, 18], 并且对对称锥互补问题也提出了一些效用函数法 [19, 20]. 容易证明: 以上算法中的全局收敛性假设都蕴含了所考虑问题的解集是非空有界的. 一个很自然的问题是如何求解具有无界解集的对称锥优化问题?

本文考虑求解对称锥互补问题. 设 V 是一个具有内积 $\langle\cdot,\cdot\rangle$ 的 n 维向量空间, $K\subset V$ 是一个对称锥. 如果存在双线性映射 $V\times V\to V$, 记作 "$\circ$", 使得对任意的 $x,y,z\in V$ 有

$$x\circ y=y\circ x;\quad \langle x\circ y,z\rangle=\langle x,y\circ z\rangle;\quad L_xL_{x^2}=L_{x^2}L_x,$$

其中 $x^2:=x\circ x$, $L_x:V\to V$ 是一个线性算子, 定义为 $L_xy:=x\circ y$, $\forall y\in V$, 则 $(V,\circ,\langle\cdot,\cdot\rangle)$ 称为一个欧氏 Jordon 代数. 集合 $\{x^2:x\in V\}$ 称为欧氏 Jordon 代数 $(V,\circ,\langle\cdot,\cdot\rangle)$ 的平方锥. 已经证明: 一个锥是对称锥当且仅当它是某个欧氏 Jordon 代数的平方锥. 因此, 一个对称锥互补问题可以表述成: 求解 $x,s\in V$, 使得

$$x\succeq 0,\quad s=F(x)\succeq 0,\quad x\circ s=0, \tag{1.1}$$

其中序 $\succeq$ 是由 $K\subset V$(K 是某个欧氏 Jordon 代数 $(V,\circ,\langle\cdot,\cdot\rangle)$ 的平方锥) 导入的, 即: $x\succeq 0$ 是指 $x\in K$ (类似地, $x\succ 0$ 是指 $x\in\mathrm{int}(K)$), $F:V\to V$ 是一个非线性函数. 我们记这类问题为 SCCP. 若对所有的 $u,v\in V$ 有 $\langle u-v,F(u)-F(v)\rangle\geqslant 0$ 成立, 则称函数 F 是单调的. 若 F 是单调的, 则称 (1.1) 式是一个单调对称锥互补问题.

本文提出一个求解 SCCP(1.1) 的光滑算法. 在提出的算法中, 嵌入了一个非单调线搜索技术. 我们证明在如下的假设下这个算法是全局收敛的:

假设 1.1 SCCP(1.1) 的解集非空.

我们也证明: 在适当的假设下, 算法得到的解是 (1.1) 式的一个极大互补解.

2 基本理论

以下介绍欧氏 Jordon 代数的基本结论. 更详细的内容可参见文献 [21].

假定 V 有一个单位元, 即: 存在唯一的元素 $e\in V$, 使得对所有的 $x\in V$ 都有 $x\circ e=e\circ x=x$ 成立. 元素 $c\in V$ 称作幂等元, 如果 $c\circ c=c$. 幂等元 c 称作本征的, 如果它是非零的而且不能写成两个非零的幂等元之和. 对任意 $x\in V$, 令 $m(x)$ 是使得 $\{e,x,\ldots,x^{m(x)}\}$ 线性相关的最小正整数. 定义 V 的秩为 $\mathrm{Rank}(V):=\max\{m(x):x\in V\}$. 假定 $\mathrm{Rank}(V)=r$; 记 $\mathcal{I}:=\{1,\ldots,r\}$. 称一个幂等元的集合 $\{c_1,\ldots,c_k\}$ 是完备的正交幂等系, 如果对任意 $i,j\in\{1,\ldots,k\}$ 且 $i\neq j$ 都有 $c_i\circ c_j=0$ 和 $\sum_{i=1}^k c_i=e$ 成立. 称一个完备的正交幂等系 $\{c_1,\ldots,c_r\}$ 是一个 Jordon 系, 如果它的每个元素是本征的.

定理 2.1 (谱分解定理, 类型 I[21, 定理 III.1.1]) 对 $x\in V$, 存在唯一的互不相同的实数 $\lambda_1(x),\lambda_2(x),\ldots,\lambda_k(x)$ 和唯一的完备的正交幂等系 $\{c_1,\ldots,c_k\}$, 使得 $x=\lambda_1(x)c_1+\cdots+\lambda_k(x)c_k$.

定理 2.2 (谱分解定理, 类型 II[21, 定理 III.1.2]) 设 $(V,\circ,\langle\cdot,\cdot\rangle)$ 是一个 $\mathrm{Rank}(V)=r$ 的欧氏 Jordon 代数. 则对任意的 $x\in V$, 存在 Jordon 系 $\{c_1,\ldots,c_r\}$ 和实数 $\lambda_1(x),\ldots,\lambda_r(x)$, 使得 $x=\sum_{i=1}^r\lambda_i(x)c_i$. 实数 $\lambda_1(x)$, …, $\lambda_r(x)$ (包括它们的重数) 由 x 唯一决定.

每个 $\lambda_i(x)(i\in\mathcal{I})$ 称作 x 的一个特征值, 它们是 x 的连续函数 [6]. 定义 $\mathrm{Tr}(x):=$

$\sum_{i=1}^{r}\lambda_i(x)$, 其中 $\mathrm{Tr}(x)$ 表示 x 的迹. 从定理 2.2 可知对于 $x\in V$, x 的特征值是由 x 唯一决定的, 但是 x 可能有不同的 Jordon 系.

若 $x^2=y$ 且 $x\succeq 0$, 则 x 可以写成 $y^{1/2}$. 称元素 $x\in V$ 是可逆的, 如果对某个 $k<\infty$ 和实数 γ_i 存在 $y=\sum_{i=0}^{k}\gamma_i x^i$, 使得 $y\circ x=e$, 记 x 的逆为 x^{-1}. 给定 $x\in V$ 具有 $x=\sum_{i=1}^{r}\lambda_i(x)c_i$, 其中 $\{c_1,\ldots,c_r\}$ 是一个 Jordon 系, $\lambda_1(x),\ldots,\lambda_r(x)$ 是 x 的特征值, 则 $x^2=\sum_{i=1}^{r}\lambda_i^2(x)c_i$. 若对所有的 $i\in\mathcal{I}$ 有 $\lambda_i(x)\geqslant 0$ 成立, 则 $x^{1/2}=\sum_{i=1}^{r}\lambda_i^{1/2}(x)c_i$; 而且如果对所有的 $i\in\mathcal{I}$ 都有 $\lambda_i(x)>0$ 成立, 那么 x 是可逆的且 $x^{-1}=\sum_{i=1}^{r}\lambda_i^{-1}(x)c_i$. 一般地, 可将实值函数 g 通过 Jordon 代数中元素的特征值来定义, 把它推广到 Jordon 代数上的映射 G, 即: $G(x):=\sum_{i=1}^{r}g(\lambda_i(x))c_i$, 其中 $x\in V$ 具有谱分解 $x=\sum_{i=1}^{r}\lambda_i(x)c_i$.

对任意的 $x,y\in V$, 定义 x 和 y 的内积为 $\langle x,y\rangle:=\mathrm{Tr}(x\circ y)$, 并且 V 上 x 的由内积导出的范数为 $\|x\|:=\sqrt{\langle x,x\rangle}=\sqrt{\mathrm{Tr}(x^2)}=\sqrt{\sum_{i=1}^{r}\lambda_i^2(x)}$. 下面的结论很容易证明, 我们省去它们的证明.

引理 2.3 (i) 对任意的 $u,v\in V$, 若 $u\succeq 0, v\succeq 0$, 则 $\langle u,v\rangle\geqslant 0$.

(ii) 假定序列 $\{u^k\}\subset V$, 且当 $k\to\infty$ 时 $\|u^k\|\to\infty$. 设 $\hat{u}$ 是序列 $\{u^k/\|u^k\|\}$ 的一个聚点, 则 $\|\hat{u}\|=1$. 进一步, 如果对所有 k 有 $u^k\succeq 0$ 成立, 那么 $\hat{u}\succeq 0$.

(iii) 对 $u,v\in V$, 若 $u\succ 0, v\succeq 0$, 且 $\langle u,v\rangle=0$, 则 $v=0$.

在求解 SCCP 的数值方法的理论分析中, 一个常用的假设是下面的严格可行性.

假设 2.4 $\mathcal{F}^0:=\{(x,s)\in V\times V: x\succ 0,\ s=F(x)\succ 0\}\neq\emptyset$.

下面这个命题表明假设 2.4 比本文所用的假设 1.1 要强.

命题 2.5 假定 F 是一个连续单调函数且假设 2.4 成立, 则 (1.1) 式的解集非空有界.

证明 由于 (1.1) 式的解的存在性可由文献 [22, 定理 4.1] 得到, 所以只需要证明 (1.1) 式的解集 $\mathcal{S}$ 是有界的. 假设 $\mathcal{S}$ 是无界的, 则存在序列 $\{(x^k,s^k)\}\subset\mathcal{S}$, 且当 $k\to\infty$ 时 $\|(x^k,s^k)\|\to\infty$, 我们将导出一个矛盾. 显然对所有的 k 有

$$x^k\succeq 0,\quad s^k=F(x^k)\succeq 0,\quad x^k\circ F(x^k)=0. \tag{2.1}$$

取 $(x^0,s^0)\in\mathcal{F}^0$, 则

$$x^0\succ 0,\quad s^0=F(x^0)\succ 0. \tag{2.2}$$

由于 F 是单调的, 所以对任意的 k 都有 $\langle x^0-x^k,F(x^0)-F(x^k)\rangle\geqslant 0$ 成立. 这与 (2.1) 和 (2.2) 式结合可得: 对所有的 k 有

$$\langle x^0,s^k\rangle+\langle x^k,s^0\rangle\leqslant\langle x^0,s^0\rangle. \tag{2.3}$$

由引理 2.3(i) 我们知道

$$\langle x^0,s^k\rangle\geqslant 0,\quad \langle x^k,s^0\rangle\geqslant 0,\quad \langle x^0,s^0\rangle\geqslant 0. \tag{2.4}$$

由于当 $k\to\infty$ 时 $\|(x^k,s^k)\|\to\infty$, 所以 $\{\|x^k\|\}$ 和 $\{\|s^k\|\}$ 中至少有一个序列是无界的. 不妨假定 $\{\|x^k\|\}$ 是无界的. 如果有必要选取子序列, 我们总可以假定 $\lim_{k\to\infty}\frac{x^k}{\|x^k\|}=\bar{x}$. 因此, 由引理 2.3(ii) 可知 $\bar{x}\succeq 0$ 和 $\|\bar{x}\|=1$. 另一方面, 由 (2.3) 和 (2.4) 式可以得到对所有的 k 有

$$\left\langle\frac{x^k}{\|x^k\|},s^0\right\rangle=\frac{\langle x^k,s^0\rangle}{\|x^k\|}\leqslant\frac{\langle x^0,s^0\rangle}{\|x^k\|},$$

即 $\langle\bar{x},s^0\rangle\leqslant 0$, 所以有 $\langle\bar{x},s^0\rangle=0$. 由引理 2.3(iii) 进一步得到 $\bar{x}=0$, 这和结论 $\|\bar{x}\|=1$ 矛盾, 从而命题得证.

3 算法设计

本文使用如下的函数:

$$\phi(\mu, x, s) = x + s - \sqrt{(x-s)^2 + 4\mu e}, \quad \forall(\mu, x, s) \in \Re_+ \times V \times V. \tag{3.1}$$

下面的这个引理表明 ϕ 在 $\Re_{++} \times V \times V$ 上是连续可微的.

引理 3.1. 设 $x, s, u, v \in V, \mu > 0, h \in \Re$, 函数 ϕ 由 (3.1) 式定义.

(i) 假定 $(x-s)^2$ 的谱分解为

$$(x-s)^2 = \sum_{i=1}^{r} \lambda_i c_i, \tag{3.2}$$

其中 $\{c_1, \ldots, c_r\}$ 是一个 Jordon 系, 实数 $\lambda_1, \ldots, \lambda_r$ (包括其重数) 由 $(x-s)^2$ 唯一确定, 则

$$D_\mu \phi(\mu, x, s)h = -\sum_{i=1}^{r} \frac{2h}{\sqrt{\lambda_i + 4\mu}} c_i$$

且 $D_\mu \phi(\cdot, \cdot, \cdot)$ 在 $\Re_{++} \times V \times V$ 上连续.

(ii) 记 $c := \sqrt{(x-s)^2 + 4\mu e}$, 则

$$D_x \phi(\mu, x, s)u = u - L_c^{-1}[(x-s) \circ u]; \quad D_s \phi(\mu, x, s)v = v - L_c^{-1}[(x-s) \circ (-v)],$$

并且 $D_x \phi(\cdot, \cdot, \cdot)$ 和 $D_s \phi(\cdot, \cdot, \cdot)$ 在 $\Re_{++} \times V \times V$ 上连续.

证明 由 ϕ 的定义有

$$\phi(\mu + th, x, s) - \phi(\mu, x, s) = \sqrt{(x-s)^2 + 4\mu e} - \sqrt{(x-s)^2 + 4(\mu + th)e}.$$

由 (3.2) 式和 $\sum_{i=1}^{r} c_i = e$, 可以得到

$$(x-s)^2 + 4\mu e = \sum_{i=1}^{r} (\lambda_i + 4\mu) c_i;$$

$$(x-s)^2 + 4(\mu + th)e = \sum_{i=1}^{r} (\lambda_i + 4\mu + 4th) c_i.$$

由于 $(x-s)^2 \succeq 0$, 所以对所有的 $i \in \mathcal{I}$ 都有 $\lambda_i \geqslant 0$ 成立, 从而,

$$\sqrt{(x-s)^2 + 4\mu e} = \sum_{i=1}^{r} \sqrt{\lambda_i + 4\mu} c_i;$$

$$\sqrt{(x-s)^2 + 4(\mu + th)e} = \sum_{i=1}^{r} \sqrt{\lambda_i + 4\mu + 4th} c_i.$$

因此,

$$\begin{aligned}
\lim_{t \to 0} \frac{\phi(\mu + th, x, s) - \phi(\mu, x, s)}{t} &= \lim_{t \to 0} \frac{\sqrt{(x-s)^2 + 4\mu e} - \sqrt{(x-s)^2 + 4(\mu + th)e}}{t} \\
&= \lim_{t \to 0} \frac{\sum_{i=1}^{r} [\sqrt{\lambda_i + 4\mu} - \sqrt{\lambda_i + 4\mu + 4th}] c_i}{t} \\
&= \sum_{i=1}^{r} \lim_{t \to 0} \frac{(\lambda_i + 4\mu) - (\lambda_i + 4\mu + 4th)}{t[\sqrt{\lambda_i + 4\mu} + \sqrt{\lambda_i + 4\mu + 4th}]} c_i \\
&= -\sum_{i=1}^{r} \frac{2h}{\sqrt{\lambda_i + 4\mu}} c_i.
\end{aligned}$$

这就说明了 $\phi(\cdot, x, s)$ 对变量 μ 是可微的且 $D_\mu \phi(\mu, x, s)h = -\sum_{i=1}^{r} \frac{2h}{\sqrt{\lambda_i + 4\mu}} c_i$. 从上面的等式, 容易得到 $D_\mu \phi(\cdot, \cdot, \cdot)$ 是连续的. 这样就证明了结论 (i). 另一方面, 结论 (ii) 可以类似于文献 [16, 引理 1 和 2] 进行证明, 我们省去具体细节. 从而证明了这个引理.

下面的引理建立了 SCCP (1.1) 和函数 ϕ 的联系 (参见文献 [6, 命题 6 和 7]).

引理 3.2 设 $x, s \in V, \mu \in \Re_{++}$, 函数 ϕ 由 (3.1) 式定义. 则,

(i) $x \succeq 0, s \succeq 0, x \circ s = 0$ 当且仅当 $x \succeq 0, s \succeq 0, \langle x, s\rangle = 0$;

(ii) $\phi(0, x, s) = 0$ 当且仅当 $x \succeq 0, s \succeq 0, x \circ s = 0$;

(iii) $\phi(\mu, x, s) = 0$ 当且仅当 $x \succ 0, s \succ 0, x \circ s = \mu e$.

特别地, 在上面的每一种情形下, x 和 s 都是算子可交换的.

定义函数 $H : \Re \times V \times V \to \Re \times V \times V$ 如下:

$$H(\mu, x, s) := \begin{pmatrix} \mu \\ s - F(x) \\ \phi(\mu, x, s) + \mu x \end{pmatrix}. \tag{3.3}$$

显然, 从引理 3.2 可以得到: $H(\mu, x, s) = 0$ 当且仅当 $\mu = 0$ 并且 (x, s) 是 (1.1) 式的解. 从引理 3.1 可以容易得到: 对任意的 $(\mu, x, s) \in \Re_{++} \times V \times V$, 若 F 是连续可微的, 则函数 H 也是连续可微的. 因此, 可以运用某些 Newton 型的方法在每一步迭代中求解 $H(\mu, x, s) = 0$, 并且使得 $H(\mu, x, s) \to 0$, 从而得到 (1.1) 式的一个解.

下面提出一个求解 (1.1) 式的光滑算法, 并且说明这个算法是适定的. 我们用 E 表示单位算子, $\mathcal{J}$ 表示所有非负整数的集合, 即: $\mathcal{J} := \{0, 1, 2, \ldots\}$. 对任意 $(\mu, x, s), (\mu_k, x^k, s^k) \in \Re_+ \times V \times V$, 我们总是使用记号 $z := (\mu, x, s)$ 和 $z^k := (\mu_k, x^k, s^k)$.

算法 3.3 (一个光滑算法) 步骤 0. 选取 $\delta, \sigma \in (0, 1), \mu_0 > 0, x^0, s^0 \in V$. 置 $z^0 := (\mu_0, x^0, s^0)$. 选取 $\beta > 1$, 使得 $\|H(z^0)\| \leqslant \beta\mu_0$. 取定 $p := (1, 0, 0) \in \Re \times V \times V$, $C_0 := \|H(z^0)\|$ 和 $Q_0 := 1$. 选取 $\eta_{\min}$ 和 $\eta_{\max}$ 使得 $0 \leqslant \eta_{\min} \leqslant \eta_{\max} < 1$. 置 $k := 0$.

步骤 1. 如果 $\|H(z^k)\| = 0$, 则终止.

步骤 2. 由下面的式子计算 $\triangle z^k := (\triangle\mu_k, \triangle x^k, \triangle s^k) \in \Re \times V \times V$,

$$DH(z^k) \triangle z^k = -H(z^k) + (1/\beta)C_k p, \tag{3.4}$$

其中

$$DH(z^k) \triangle z^k = \begin{pmatrix} \triangle\mu_k \\ \triangle s^k - DF(x^k) \triangle x^k \\ (D_\mu\phi(z^k) + x^k) \triangle \mu_k + (D_x\phi(z^k) + \mu_k E) \triangle x^k + D_s\phi(z^k) \triangle s^k \end{pmatrix}. \tag{3.5}$$

步骤 3. 令 α_k 是 $1, \delta, \delta^2, \ldots$ 中使得下式成立的最大值:

$$\|H(z^k + \alpha_k \triangle z^k)\| \leqslant [1 - \sigma(1 - 1/\beta)\alpha_k]C_k. \tag{3.6}$$

步骤 4. 置 $z^{k+1} := z^k + \alpha_k \triangle z^k$. 取 $\eta_k \in [\eta_{\min}, \eta_{\max}]$. 置 $Q_{k+1} := \eta_k Q_k + 1$, $C_{k+1} := (\eta_k Q_k C_k + \|H(z^{k+1})\|)/Q_{k+1}$ 和 $k := k + 1$. 返回步骤 1.

在算法 3.3 中, 嵌入了由 Zhang 和 Hager[23] 提出的非单调线搜索. 容易看到: C_{k+1} 是 C_k 和 $\|H(z^{k+1})\|$ 的一个凸组合. 由于 $C_0 = \|H(z^0)\|$, 所以 C_k 是 $\|H(z^0)\|, \|H(z^1)\|, \ldots, \|H(z^k)\|$ 的一个凸组合. η_k 的选取控制了线搜索的非单调程度. 特别地, 如果 $\eta_k = 0$ 对所有的 k 都成立, 那么算法中的线搜索就是通常的 Armijo 单调线搜索. 我们将算法 3.3 涉及到的一些基本结论总结在如下的注中.

注 3.4 设 $\{C_k\}$ 和 $\{z^k = (\mu_k, x^k, s^k)\}$ 是由算法 3.3 产生的序列. 下面的结论成立:

(i) 序列 $\{C_k\}$ 是单调下降的.

事实上, 从 (3.6) 式和 C_{k+1} 与 Q_{k+1} 的定义可以得到: 对所有的 $k \in \mathcal{J}$ 有

$$C_{k+1} \leqslant \frac{\eta_k Q_k C_k + C_k - \sigma(1-\frac{1}{\beta})\alpha_k C_k}{Q_{k+1}} = C_k - \frac{\sigma(1-\frac{1}{\beta})\alpha_k}{Q_{k+1}} C_k \leqslant C_k, \tag{3.7}$$

这就证明了提出的结论.

(ii) 对所有的 $k \in \mathcal{J}$ 都有 $\|H(z^k)\| \leqslant C_k$ 成立.

事实上, 当 $k=0$ 时结论是显然成立的; 假定结论对 $k=l$ 时成立, 则只需要证明结论对 $k=l+1$ 也成立. 注意到

$$\begin{aligned}\|H(z^{l+1})\| &= Q_{l+1}C_{l+1} - \eta_l Q_l C_l \leqslant Q_{l+1}C_{l+1} - \eta_l Q_l C_{l+1} \\ &= (Q_{l+1} - \eta_l Q_l)C_{l+1} = C_{l+1},\end{aligned}$$

其中第 1 个等式是由 C_{l+1} 的定义得到的; 不等式是由上面的结论 (i) 得到的; 最后一个等式是由 Q_{l+1} 的定义得到的, 从而证明了提出的结论.

(iii) 对所有的 $k \in \mathcal{J}$ 都有 $C_k \leqslant \beta\mu_k$ 成立.

事实上, 从初始点的选择知道 $C_0 = \|H(z^0)\| \leqslant \beta\mu_0$. 假定对某个 $l \in \mathcal{J}$ 有 $C_l \leqslant \beta\mu_l$ 成立, 则

$$\begin{aligned}\mu_{l+1} - (1/\beta)C_{l+1} &= (1-\alpha_l)\mu_l + \alpha_l(1/\beta)C_l - (1/\beta)C_{l+1} \\ &\geqslant (1/\beta)(C_l - C_{l+1}),\end{aligned}$$

其中的等式由 (3.4) 式的第 1 个方程和 $\mu_{l+1} = \mu_l + \alpha_l \triangle \mu_l$ 得到, 不等式由 $\mu_l \geqslant (1/\beta)C_l$ 得到. 因此, 利用上面的结论 (i), 容易得到提出的结论.

(iv) 对所有的 $k \in \mathcal{J}$ 都有 $\mu_k > 0$ 成立, 并且序列 $\{\mu_k\}$ 是单调下降的.

事实上, 由 (3.4) 式的第 1 个方程容易得到第 1 个结论, 我们证明第 2 个结论. 利用 (3.4) 式的第 1 个方程和以上结论 (iii) 可以得到: 对所有的 $k \in \mathcal{J}$ 有

$$\mu_{k+1} = (1-\alpha_k)\mu_k + \alpha_k(1/\beta)C_k \leqslant (1-\alpha_k)\mu_k + \alpha_k\mu_k = \mu_k,$$

这就说明了序列 $\{\mu_k\}$ 是单调下降的.

(v) 线搜索 (3.6) 式是适定的.

事实上, 如果记 $R(z^k) := H(z^k + \alpha \triangle z^k) - H(z^k) - \alpha DH(z^k) \triangle z^k$, 那么由 (3.4) 式有

$$\begin{aligned}\|H(z^k + \alpha \triangle z^k)\| &= \|R(z^k) + H(z^k) + \alpha DH(z^k) \triangle z^k\| \\ &\leqslant \|R(z^k)\| + [1 - \alpha(1-1/\beta)]\|H(z^k)\| \\ &\leqslant \|R(z^k)\| + [1 - \alpha(1-1/\beta)]C_k,\end{aligned} \tag{3.8}$$

其中最后一个不等式是由于以上结论 (ii). 由于对任意的 $z \in \Re \times V \times V$ 和 $\mu > 0$, 函数 H 是连续可微的, 所以由函数 R 的定义和以上结论 (iv) 可以得到: 对所有的 $k \in \mathcal{J}$ 都有 $\|R(z^k)\| = o(\alpha)$ 成立. 这与 (3.8) 式结合可得到提出的结论.

为了证明 Newton 方程组是可解的, 我们需要下面的结论 (参见文献 [10, 引理 2.6(vi)]).

引理 3.5 对任意的 $a, b, u, v \in V$, 如果 $a \succ 0, b \succ 0, a \circ b \succ 0, \langle u, v\rangle \geqslant 0$, 且 $a \circ u + b \circ v = 0$ 成立, 则 $u = v = 0$.

现在证明 Newton 方程 (3.4) 是可解的.

定理 3.6 假设函数 H 由 (3.3) 定义. 则对任意的 $(\mu, x, s) \in \Re_{++} \times V \times V$, 由 (3.5) 所给出函数 H 的 Jacobi 是可逆的.

证明 假设

$$DH(\mu, x, s)(\triangle\mu, \triangle x, \triangle s) = 0, \tag{3.9}$$

则需要证明 $(\triangle\mu, \triangle x, \triangle s) = 0$. 从 (3.5) 和 (3.9) 式可以得到

$$\begin{aligned} &\triangle\,\mu = 0, \\ &-DF(x)\triangle\,x + \triangle s = 0, \\ &D_\mu\phi(z)\triangle\,\mu + x\triangle\,\mu + D_x\phi(z)\triangle\,x + \mu\triangle\,x + D_s\phi(z)\triangle\,s = 0. \end{aligned} \tag{3.10}$$

由于函数 F 是单调的, 所以对所有的 $u \in V$ 都有 $\langle u, DF(x)u\rangle \geqslant 0$ 成立, 因此, 从 (3.10) 式的第 2 个方程得到

$$\langle\triangle x, \triangle s\rangle = \langle\triangle x, DF(x)\triangle\,x\rangle \geqslant 0. \tag{3.11}$$

利用引理 3.1, 由 (3.10) 式的第 1 和第 3 个方程得到 $D_\mu\phi(z)\triangle\mu + x\triangle\mu = 0$ 和 $(1+\mu)\triangle x + \triangle s - L_c^{-1}((x-s)\circ(\triangle x - \triangle s)) = 0$, 其中 $c := \sqrt{(x-s)^2 + 4\mu e}$. 因此, $L_c((1+\mu)\triangle\,x + \triangle s) - (x-s)\circ(\triangle x - \triangle s) = 0$, 从而

$$((1+\mu)c - (x-s))\circ\triangle x + (c + (x-s))\circ\triangle s = 0. \tag{3.12}$$

从 c 的定义容易得到

$$(1+\mu)c - (x-s) \succ 0, \quad c + (x-s) \succ 0, \tag{3.13}$$

和

$$\begin{aligned} ((1+\mu)c - (x-s))\circ(c + (x-s)) &= \mu c\circ(c + (x-s)) + (c - (x-s))\circ(c + (x-s)) \\ &= \mu c\circ(c + (x-s)) + 4\mu e. \end{aligned} \tag{3.14}$$

令 $x - s = \sum_{i=1}^r \lambda_i c_i$ 是 $x - s$ 的谱分解, 则

$$c = \sum_{i=1}^r \sqrt{\lambda_i^2 + 4\mu}c_i, \quad c + (x-s) = \sum_{i=1}^r \Big(\sqrt{\lambda_i^2 + 4\mu} + \lambda_i\Big)c_i,$$

因此,

$$c\circ(c + (x-s)) = \sum_{i=1}^r \Big[\sqrt{\lambda_i^2 + 4\mu}\Big(\sqrt{\lambda_i^2 + 4\mu} + \lambda_i\Big)\Big]c_i \succeq 0.$$

从而, (3.14) 式变为

$$((1+\mu)c - (x-s))\circ(c + (x-s)) = \mu c\circ(c + (x-s)) + 4\mu e \succeq 4\mu e \succ 0. \tag{3.15}$$

结合 (3.11)–(3.13) 和 (3.15) 式以及引理 3.5, 可得到 $\triangle x = 0$ 和 $\triangle s = 0$, 从而, 定理得证.

结合注 3.4(iv) 和 (v) 及定理 3.6, 我们得到算法 3.3 是适定的.

4 全局收敛性

本节讨论算法 3.3 的全局收敛性.

引理 4.1 对任意的 $(\mu, a, b, c) \in \Re \times V \times V \times V$, 如果 $\mu \in \Re_{++}$, 那么

$$\phi(\mu, a, b) = c \iff a - \frac{1}{2}c \succ 0, \quad b - \frac{1}{2}c \succ 0, \quad \left(a - \frac{1}{2}c\right)\circ\left(b - \frac{1}{2}c\right) = \mu e.$$

证明 由 $\phi(\mu,a,b)=c$ 当且仅当

$$\left(a-\frac{1}{2}c\right)+\left(b-\frac{1}{2}c\right)-\sqrt{\left[\left(a-\frac{1}{2}c\right)-\left(b-\frac{1}{2}c\right)\right]^2+4\mu e}=0,$$

即 $\phi(\mu,a-\frac{1}{2}c,b-\frac{1}{2}c)=0$. 因此, 由引理 3.2(iii) 可以得到 $\phi(\mu,a-\frac{1}{2}c,b-\frac{1}{2}c)=0$ 当且仅当 $a-\frac{1}{2}c\succ 0$, $b-\frac{1}{2}c\succ 0$ 和 $(a-\frac{1}{2}c)\circ(b-\frac{1}{2}c)=\mu e$, 从而结论成立.

引理 4.2 假定 F 是一个连续可微的单调函数, 序列 $\{z^k\}$ 是由算法 3.3 产生的. 如果假设 1.1 成立, 则序列 $\{z^k\}$ 是有界的.

证明 由于假设 1.1 成立, 所以存在 SCCP 的一个解, 记作 $(x^*,s^*)\in V\times V$, 使得

$$x^*\succeq 0,\quad s^*=F(x^*)\succeq 0,\quad \langle x^*,s^*\rangle=0. \tag{4.1}$$

我们将证明分成以下 3 部分:

第 1 部分. 证明序列 $\{\mu_k\}$ 是有界的. 这可以很容易的从序列 $\{\mu_k\}$ 是一个正的单调下降的序列来得到 (参见注 3.4(iv)).

第 2 部分. 证明序列 $\{x^k\}$ 是有界的. 假设序列 $\{x^k\}$ 是无界的, 我们将导出矛盾. 通过如下方式构造序列 $\{(u^k,v^k)\}$:

$$u^k:=\frac{1}{\mu_k}(s^k-F(x^k)),\quad v^k:=\frac{1}{\mu_k}(\phi(z^k)+\mu_k x^k),\quad \forall\, k\in\mathcal{J}. \tag{4.2}$$

从注 3.4(ii) 和 (iii) 可以得到: 对所有的 $k\in\mathcal{J}$ 都有 $\|H(z^k)\|\leqslant C_k\leqslant\beta\mu_k$ 成立, 这可推出 $\max\{\|u^k\|,\|v^k\|\}\leqslant\sqrt{\beta^2-1}$, 因此, 序列 $\{u^k\}$ 和 $\{v^k\}$ 是一致有界的.

现在通过以下方式构造另一个序列 $\{(\hat{x}^k,\hat{s}^k)\}$:

$$\hat{x}^k:=x^k+\frac{1}{2}\mu_k x^k-\frac{1}{2}\mu_k v^k,\quad \hat{s}^k:=s^k+\frac{1}{2}\mu_k x^k-\frac{1}{2}\mu_k v^k,\quad \forall\, k\in\mathcal{J}. \tag{4.3}$$

则由引理 4.1 和 (4.2) 式中 v^k 的定义有

$$\hat{x}^k\succ 0,\quad \hat{s}^k\succ 0,\quad \hat{x}^k\circ\hat{s}^k=\mu_k e,\quad \forall\, k\in\mathcal{J}. \tag{4.4}$$

因此, 从 (4.1) 和 (4.4) 式可以得到

$$r\mu_k=\mathrm{Tr}(\mu_k e)=\langle\hat{x}^k,\hat{s}^k\rangle\geqslant\langle\hat{x}^k,\hat{s}^k\rangle-\langle x^*,\hat{s}^k\rangle-\langle\hat{x}^k,s^*\rangle=\langle\hat{x}^k-x^*,\hat{s}^k-s^*\rangle,$$

从而,

$$\begin{aligned}
r\mu_k\geqslant{}&\langle x^k-x^*,s^k-s^*\rangle+\left\langle x^k-x^*,\frac{1}{2}\mu_k x^k-\frac{1}{2}\mu_k v^k\right\rangle\\
&+\left\langle\frac{1}{2}\mu_k x^k-\frac{1}{2}\mu_k v^k,s^k-s^*\right\rangle+\left\langle\frac{1}{2}\mu_k x^k-\frac{1}{2}\mu_k v^k,\frac{1}{2}\mu_k x^k-\frac{1}{2}\mu_k v^k\right\rangle\\
&(\text{由 (4.3) 中 }\hat{x}^k,\hat{s}^k\text{的定义})\\
={}&\langle x^k-x^*,F(x^k)-F(x^*)+\mu_k u^k\rangle+\left\langle x^k-x^*,\frac{1}{2}\mu_k x^k-\frac{1}{2}\mu_k v^k\right\rangle\\
&+\left\langle\frac{1}{2}\mu_k x^k-\frac{1}{2}\mu_k v^k,F(x^k)-F(x^*)+\mu_k u^k\right\rangle+\left\langle\frac{1}{2}\mu_k x^k-\frac{1}{2}\mu_k v^k,\frac{1}{2}\mu_k x^k-\frac{1}{2}\mu_k v^k\right\rangle\\
&(\text{由 (4.2) 中}u^k\text{ 的定义})\\
={}&\langle x^k-x^*,F(x^k)-F(x^*)\rangle+\langle x^k-x^*,\mu_k u^k\rangle+\left\langle x^k-x^*,\frac{1}{2}\mu_k x^k-\frac{1}{2}\mu_k v^k\right\rangle\\
&+\frac{1}{2}\mu_k\langle x^k-v^k,F(x^k)-F(v^k)\rangle+\left\langle\frac{1}{2}\mu_k x^k-\frac{1}{2}\mu_k v^k,F(v^k)-F(x^*)+\mu_k u^k\right\rangle
\end{aligned}$$

$$
\begin{aligned}
&+\left\langle \frac{1}{2}\mu_k x^k-\frac{1}{2}\mu_k v^k, \frac{1}{2}\mu_k x^k-\frac{1}{2}\mu_k v^k\right\rangle\\
\geqslant\ &\mu_k\langle x^k-x^*, u^k\rangle+\mu_k\left\langle x^k-x^*, \frac{1}{2}x^k-\frac{1}{2}v^k\right\rangle\\
&+\mu_k\left\langle \frac{1}{2}x^k-\frac{1}{2}v^k, F(v^k)-F(x^*)+\mu_k u^k\right\rangle+\mu_k^2\left\langle \frac{1}{2}x^k-\frac{1}{2}v^k, \frac{1}{2}x^k-\frac{1}{2}v^k\right\rangle\\
&(\text{由 F 的单调性 和 } \mu_k>0 \text{ 对任意的 } k\in\mathcal{J}\text{都成立})\\
=\ &\mu_k\left\{\left(\frac{1}{2}+\frac{1}{4}\mu_k\right)\|x^k\|^2+g(\mu_k, x^k, u^k, v^k)\right\},
\end{aligned}
\tag{4.5}
$$

其中函数 g 定义如下:

$$
\begin{aligned}
g(\mu_k, x^k, u^k, v^k) :=&\langle x^k-x^*, u^k\rangle+\left\langle x^k, -\frac{1}{2}v^k\right\rangle+\left\langle -x^*, \frac{1}{2}x^k\right\rangle+\left\langle x^*, \frac{1}{2}v^k\right\rangle\\
&+\left\langle \frac{1}{2}x^k-\frac{1}{2}v^k, F(v^k)-F(x^*)+\mu_k u^k\right\rangle\\
&+\mu_k\left\langle \frac{1}{2}x^k, -\frac{1}{2}v^k\right\rangle+\mu_k\left\langle -\frac{1}{2}v^k, \frac{1}{2}x^k\right\rangle+\mu_k\left\langle \frac{1}{2}v^k, \frac{1}{2}v^k\right\rangle.
\end{aligned}
\tag{4.6}
$$

进一步, (4.5) 式变为

$$
\left(\frac{1}{2}+\frac{1}{4}\mu_k\right)\|x^k\|^2+g(\mu_k, x^k, u^k, v^k)\leqslant r. \tag{4.7}
$$

由于序列 $\{(\mu_k, u^k, v^k)\}$ 是有界的, 而且 $\mu_k>0$ 对所有的 $k\in\mathcal{J}$ 成立, 由 (4.6) 式和当 $k\to\infty$ 时 $\|x^k\|\to\infty$, 可以得到序列 $\{\frac{g(\mu_k, x^k, u^k, v^k)}{\|x^k\|}\}$ 是有界的. 从而当 $\|x^k\|\to\infty$ 时, (4.7) 式的左边的式子趋于 ∞, 这就和 (4.7) 式矛盾. 因此, 序列 $\{x^k\}$ 是有界的.

第 3 部分. 证明序列 $\{s^k\}$ 是有界的. 由于序列 $\{x^k\}$ 是有界的, 所以由函数 F 的连续性可得序列 $\{F(x^k)\}$ 是有界的. 这些连同 $\{(\mu_k, u^k)\}$ 的有界性以及 $s^k=F(x^k)+\mu_k u^k$ 可以推出序列 $\{s^k\}$ 是有界的.

结合第 1–3 部分, 该引理得证.

现在我们给出算法 3.3 的收敛性.

定理 4.3 假定 F 是一个连续可微的单调函数, 序列 $\{z^k\}$ 是由算法 3.3 产生的. 如果假设 1.1 成立, 那么序列 $\{z^k\}$ 是有界的并且序列 $\{(x^k, s^k)\}$ 的任意聚点是 SCCP 的一个解.

证明 由引理 4.2 可以得到无穷序列 $\{z^k\}$ 是有界的, 因此只需要证明定理的第 2 个结论. 为此, 只需要证明 $\lim_{k\to\infty}\|H(z^k)\|=0$ (如果有必要可以取子序列), 因为如果这个结论成立, 那么通过一个很简单的连续性讨论就可以得到定理的结论.

从注 3.4 (i)–(iv), 可以得到

$$
0\leqslant\|H(z^{k+1})\|\leqslant C_{k+1}\leqslant C_k\leqslant\beta\mu_k\leqslant\beta\mu_0<\infty, \tag{4.8}
$$

因此, 序列 $\{C_k\}$ 是收敛的而且 $\{\|H(z^k)\|\}$ 是有界的. 由于 $\{z^k\}$ 是有界的, 所以序列 $\{z^k\}$ 有一个收敛的子列, 不失一般性, 我们仍然记作 $\{z^k\}$. 记这个子序列的指标集为 $\bar{\mathcal{J}}$. 令 $z^*:=(\mu_*, x^*, s^*)$ 是这个序列的极限, 则 $\lim_{\bar{\mathcal{J}}\ni k\to\infty}\|H(z^k)\|=\|H(z^*)\|$. 记 $C^*:=\lim_{\bar{\mathcal{J}}\ni k\to\infty}C_k$. 如果 $C^*=0$, 那么易知 $\|H(z^*)\|=0$, 从而结论得到证明. 下面假定 $C^*>0$ 且 $\|H(z^*)\|>0$, 我们将导出矛盾. 我们把证明分成如下的两部分:

第 1 部分. 假定对所有的 $k\in\bar{\mathcal{J}}$ 都有 $\alpha_k\geqslant c>0$ 成立, 其中 c 是一个常数. 一方面, 由

(3.7) 式得到对任意的 $k \in \bar{\mathcal{J}}$ (对应的指标 $i_k \in \mathcal{J}$) 有

$$C_{k+1} = C_{i_{k+1}} \leqslant C_{i_k+1} \leqslant C_{i_k} - \frac{\sigma(1-\frac{1}{\beta})c}{Q_{i_k+1}} C_{i_k} = C_k - \frac{\sigma(1-\frac{1}{\beta})c}{Q_{i_k+1}} C_k,$$

则由序列 $\{C_k\}$ 的有界性可以得到

$$\sum_{k=0}^{\infty} \frac{\sigma(1-\frac{1}{\beta})c}{Q_{i_k+1}} C_k < \infty. \tag{4.9}$$

另一方面, 由 Q_k 的定义和从算法知道 $\eta_{\max} \in [0,1)$, 可以得到

$$Q_{k+1} = 1 + \sum_{i=0}^{k} \prod_{j=0}^{i} \eta_{k-j} \leqslant 1 + \sum_{i=0}^{k} \eta_{\max}^{i+1} \leqslant \sum_{i=0}^{\infty} \eta_{\max}^{i} = \frac{1}{1-\eta_{\max}}$$

对所有的 $k \in \mathcal{J}$ 都成立. 再由 (4.9) 式可得 $\lim_{\bar{\mathcal{J}} \ni k \to \infty} C_k = 0$, 这和 $C^* > 0$ 矛盾.

第 2 部分. 假设存在 $\bar{\mathcal{J}}$ 的一个子集, 不失一般性, 仍然记作 $\bar{\mathcal{J}}$, 使得 $\lim_{\bar{\mathcal{J}} \ni k \to \infty} \alpha_k = 0$, 则对充分大的 $k \in \bar{\mathcal{J}}$, 步长 $\hat{\alpha}_k := \alpha_k/\delta$ 不满足线搜索 (3.6), 即

$$\|H(z^k + \hat{\alpha}_k \triangle z^k)\| > [1 - \sigma(1-1/\beta)\hat{\alpha}_k] C_k \tag{4.10}$$

对充分大的 $k \in \bar{\mathcal{J}}$ 成立. 很容易验证以下结论成立:

(a) $\mu_* > 0$. 这个由 (4.8) 式容易得到.

(b) $\{\triangle z^k\}_{k \in \bar{\mathcal{J}}}$ 是收敛的. 事实上, 由于 $\mu_* > 0$, 所以对所有的 $k \in \bar{\mathcal{J}}$, $DH(z^k)$ 是一个连续可逆的线性算子, 因此, 通过 (3.4) 式可以得到提出的结论. 记 $\triangle z^* := \lim_{\bar{\mathcal{J}} \ni k \to \infty} \triangle z^k$.

(c) $C^* = \|H(z^*)\|$. 事实上, 从 (4.8) 式得到 $C^* \geqslant \|H(z^*)\|$; 而从 (4.10) 式得到 $C^* \leqslant \|H(z^*)\|$.

由 (4.10) 式可以得到 $\|H(z^k + \hat{\alpha}_k \triangle z^k)\| > [1 - \sigma(1-\frac{1}{\beta})\hat{\alpha}_k] C_k$. 由于 $C_k \geqslant \|H(z^k)\|$, 所以上面的不等式变成 $(\|H(z^k + \hat{\alpha}_k \triangle z^k)\| - \|H(z^k)\|)/\hat{\alpha}_k > -\sigma(1-1/\beta)C_k$. 由上面的结论 (a) 得到 $\|H(\cdot)\|$ 在 z^* 处是连续可微的, 因此, 令 $k \to \infty$, 由上面的不等式和结论 (b) 有

$$\frac{1}{\|H(z^*)\|} \langle H(z^*), DH(z^*) \triangle z^* \rangle \geqslant -\sigma\left(1 - \frac{1}{\beta}\right) C^*. \tag{4.11}$$

另外, 使用 (3.4) 式和上面的结论 (b) 和 (c) 有

$$\begin{aligned}
\frac{1}{\|H(z^*)\|} \langle H(z^*), DH(z^*) \triangle z^* \rangle &= -\|H(z^*)\| + \frac{C^*}{\beta \|H(z^*)\|} \langle H(z^*), p \rangle \\
&\leqslant -\|H(z^*)\| + \frac{C^* \|H(z^*)\|}{\beta \|H(z^*)\|} \\
&= -\|H(z^*)\| + \frac{C^*}{\beta} \\
&= \left(-1 + \frac{1}{\beta}\right) C^*.
\end{aligned} \tag{4.12}$$

因此, 由 (4.11) 和 (4.12) 式及 $C^* > 0$ 可得 $-1 + \frac{1}{\beta} \geqslant -\sigma(1-\frac{1}{\beta})$, 这和 $\sigma \in (0,1)$ 以及 $\beta > 1$ 矛盾.

这样我们证明了这个定理.

5 算法 3.3 的收敛行为

在上一节, 我们证明了算法 3.3 可以得到 SCCP 的一个解. 那么, 算法 3.3 得到的解是什么类型的呢? 本节对这个问题给出部分回答. 我们将证明在某些假设下, 算法 3.3 得到的解是 SCCP 的一个极大互补解.

假设 $(\bar{x},\bar{s})$ 是 SCCP 的任意一个解, 即: $\bar{x}\succeq 0, \bar{s}=F(\bar{x})\succeq 0, \bar{x}\circ\bar{s}=0$, 则 $\bar{x}$ 和 $\bar{s}$ 算子可交换, 因此, 它们具有一个相同的 Jordon 系 $\{c_1,\ldots,c_r\}$ 且 $\bar{x}=\sum_{i=1}^r\lambda_i(\bar{x})c_i$, $\bar{s}=\sum_{i=1}^r\lambda_i(\bar{s})c_i$. 容易得到

$$\lambda_i(\bar{x})\geqslant 0,\quad \lambda_i(\bar{s})\geqslant 0,\quad \lambda_i(\bar{x})\lambda_i(\bar{s})=0,\qquad i=1,\ldots,r.$$

称 $(\lambda(\bar{x}),\lambda(\bar{s})):=((\lambda_1(\bar{x}),\ldots,\lambda_r(\bar{x}))^{\mathrm{T}},(\lambda_1(\bar{s}),\ldots,\lambda_r(\bar{s}))^{\mathrm{T}})$ 是 $(\bar{x},\bar{s})$ 的一个谱向量 (注意 $\bar{x}$ 和 $\bar{s}$ 的特征值是按它们在这个共同的 Jordon 系下谱分解中特征值的顺序排列的).

称 SCCP 的一个解 $(\bar{x},\bar{s})$ 是一个极大互补解, 如果 $(\bar{x},\bar{s})$ 的谱向量中正特征值的个数是 SCCP 的所有解当中最大的 (因为任意一个元素的特征值是由它唯一确定的, 这样的定义是合理的). 显然, 如果 SCCP 有解, 则它一定有一个极大互补解.

下面假定 $(\bar{x},\bar{s})$ 是 SCCP 的一个极大互补解, 则易知 $\bar{x}$ 和 $\bar{s}$ 算子可交换, 从而它们有一个相同的 Jordon 系, 记为 $\{\bar{c}_1,\ldots,\bar{c}_r\}$, 即

$$(\bar{x},\bar{s})=\left(\sum_{i=1}^r\lambda_i(\bar{x})\bar{c}_i,\sum_{i=1}^r\lambda_i(\bar{s})\bar{c}_i\right).\tag{5.1}$$

由定理 4.3, 如果有必要可选取子序列, 可以假定 $\lim_{k\to\infty}(x^k,s^k)=(x^*,s^*)$, 其中 (x^*,s^*) 是 SCCP 的一个解. 对所有的 $k\in\mathcal{J}$, 令 u^k 和 v^k 与 $\hat{x}^k$ 和 $\hat{s}^k$ 分别按 (4.2) 和 (4.3) 式所定义, 则容易得到

$$\lim_{k\to\infty}(\hat{x}^k,\hat{s}^k)=\lim_{k\to\infty}(x^k,s^k)=(x^*,s^*).\tag{5.2}$$

从 (4.4) 式知道 $\hat{x}^k$ 和 $\hat{s}^k$ 有一个相同的 Jordon 系, 因此, 可以假定

$$(\hat{x}^k,\hat{s}^k)=\left(\sum_{i=1}^r\lambda_i(\hat{x}^k)c_i^k,\sum_{i=1}^r\lambda_i(\hat{s}^k)c_i^k\right).\tag{5.3}$$

由于 V 中的任意元素的特征值是连续依赖于这个元素的, 故从 (5.2) 式易得: 对任意 $i\in\mathcal{I}$ 有

$$\lim_{k\to\infty}(\lambda_i(\hat{x}^k),\lambda_i(\hat{s}^k))=(\lambda_i(x^*),\lambda_i(s^*)),\tag{5.4}$$

其中 $\lambda_i(x^*)$ 和 $\lambda_i(s^*)$ 分别是 x^* 和 s^* 的特征值. 令 $\sigma(x^*)$ 和 $\sigma(s^*)$ 分别表示 x^* 和 s^* 的所有不相同的特征值; $\bar{r}(x^*)$ 和 $\bar{r}(s^*)$ 分别表示 $\sigma(x^*)$ 和 $\sigma(s^*)$ 中元素的个数; $\bar{\mathcal{I}}(x^*)=\{1,\ldots,\bar{r}(x^*)\}$; $\bar{\mathcal{I}}(s^*)=\{1,\ldots,\bar{r}(s^*)\}$. 记

$$\begin{aligned}
&\sigma(x^*)=\{\gamma_1(x^*),\ldots,\gamma_{\bar{r}(x^*)}(x^*)\};\\
&\sigma(s^*)=\{\gamma_1(s^*),\ldots,\gamma_{\bar{r}(s^*)}(s^*)\};\\
&N_j(x^*)=\{i\in\mathcal{I}:\gamma_j(x^*)=\lambda_i(x^*)\},\quad \forall\, j\in\bar{\mathcal{I}}(x^*);\\
&N_j(s^*)=\{i\in\mathcal{I}:\gamma_j(s^*)=\lambda_i(s^*)\},\quad \forall\, j\in\bar{\mathcal{I}}(s^*).
\end{aligned}$$

则由类型 I 的谱分解可以得到

$$(x^*,s^*)=\left(\sum_{j=1}^{\bar{r}(x^*)}\gamma_j(x^*)b_j^*(x^*),\sum_{j=1}^{\bar{r}(s^*)}\gamma_j(s^*)b_j^*(s^*)\right),\tag{5.5}$$

其中 $\{b_1^*(x^*),\ldots,b_{\bar{r}(x^*)}^*(x^*)\}$ 是 x^* 的唯一的完备正交幂等系, $\{b_1^*(s^*),\ldots,b_{\bar{r}(s^*)}^*(s^*)\}$ 是 s^* 的唯一的完备正交幂等系, 则从 (5.4) 和 (5.5) 式及定理 2.1 容易得到: 对任意的 $j\in\bar{\mathcal{I}}(x^*)$ 有

$$\lim_{k\to\infty}\lambda_i(\hat{x}^k)=\gamma_j(x^*),\quad \forall\, i\in N_j(x^*);\qquad \lim_{k\to\infty}\left(\sum_{i\in N_j(x^*)}c_i^k\right)=b_j^*(x^*),\tag{5.6}$$

以及对任意的 $j \in \bar{\mathcal{I}}(s^*)$ 有

$$\lim_{k\to\infty} \lambda_i(\hat{s}^k) = \gamma_j(s^*), \quad \forall\, i \in N_j(s^*); \qquad \lim_{k\to\infty} \left(\sum_{i\in N_j(s^*)} c_i^k \right) = b_j^*(s^*).$$

现在我们给出在某些假设下, 算法 3.3 得到的解是 SCCP 的一个极大互补解.

定理 5.1 假定 F 是一个连续可微的单调函数, 序列 $\{z^k\}$ 是由算法 3.3 产生的, 且假设 1.1 成立. 假定 $(\bar{x}, \bar{s})$ 是 SCCP 的一个极大互补解, 并且具有谱分解 (5.1). 假设对任意的 $k \in \mathcal{J}$, $(\hat{x}^k, \hat{s}^k)$ 具有谱分解 (5.3), (x^*, s^*) 具有谱分解 (5.5). 如果

(A) 对任意的 $i_0 \in \{i \in \mathcal{I} : \lambda_i(\bar{x}) > 0\}$ 满足 $\lim_{k\to\infty} \lambda_{i_0}(\hat{x}^k) = \gamma_{i^*}(x^*)$, 都有 $\langle b_{i^*}^*(x^*), \bar{x}\rangle \neq 0$; 以及对任意的 $j_0 \in \{i \in \mathcal{I} : \lambda_i(\bar{s}) > 0\}$ 满足 $\lim_{k\to\infty} \lambda_{j_0}(\hat{s}^k) = \gamma_{j^*}(s^*)$, 都有 $\langle b_{j^*}^*(s^*), \bar{s}\rangle \neq 0$,

那么由算法 3.3 得到的解 (x^*, s^*) 是 SCCP 的一个极大互补解.

证明 假定函数 g 是由 (4.6) 式定义但其中用 $\bar{x}$ 代替 x^*. 类似于 (4.5) 式的证明, 可以得到: 对所有的 $k \in \mathcal{J}$ 都有 $\langle \hat{x}^k - \bar{x}, \hat{s}^k - \bar{s}\rangle \geqslant \mu_k g(\mu_k, x^k, u^k, v^k)$ 成立, 从而可以得到 $\langle \hat{x}^k, \bar{s}\rangle + \langle \bar{x}, \hat{s}^k\rangle = -\langle \hat{x}^k - \bar{x}, \hat{s}^k - \bar{s}\rangle + r\mu_k \leqslant \mu_k(|g(\mu_k, x^k, u^k, v^k)| + r)$. 由 (5.1) 和 (5.3) 式可得

$$\sum_{i,j=1}^{r} \lambda_i(\hat{x}^k)\lambda_j(\bar{s})\langle c_i^k, \bar{c}_j\rangle + \sum_{i,j=1}^{r} \lambda_i(\bar{x})\lambda_j(\hat{s}^k)\langle \bar{c}_i, c_j^k\rangle \leqslant \mu_k(|g(\mu_k, x^k, u^k, v^k)| + r). \tag{5.7}$$

因为 $\hat{x}^k, \hat{s}^k, \bar{x}, \bar{s} \succeq 0$, 所以它们的每个特征值都是非负的. 这些连同 $\langle c_i^k, \bar{c}_j\rangle \geqslant 0$ 对所有的 $i, j \in \mathcal{I}$ 都成立, 可推出 (5.7) 式左边式子中的每一项都是非负的.

因此, 对任意的 $i_0 \in \{i \in \mathcal{I} : \lambda_i(\bar{x}) > 0\}$ 满足 $\lim_{k\to\infty} \lambda_{i_0}(\hat{x}^k) = \gamma_{i^*}(x^*)$ 可以得到对所有的 $i \in N_{i^*}(x^*)$ 有

$$\begin{aligned} \lambda_i(\hat{s}^k)\langle c_i^k, \bar{x}\rangle &= \lambda_i(\hat{s}^k)\left\langle c_i^k, \sum_{j=1}^{r} \lambda_j(\bar{x})\bar{c}_j \right\rangle = \lambda_i(\hat{s}^k) \sum_{j=1}^{r} \lambda_j(\bar{x})\langle c_i^k, \bar{c}_j\rangle \\ &\leqslant \mu_k(|g(\mu_k, x^k, u^k, v^k)| + r), \quad \forall\, k \in \mathcal{J}, \end{aligned} \tag{5.8}$$

其中的不等式用到了 (5.7) 式. 利用 (4.4) 式中的等式易得 $\lambda_i(\hat{x}^k)\lambda_i(\hat{s}^k) = \mu_k$, 这个连同 (5.8) 式可得

$$\lambda_i(\hat{x}^k) \geqslant \frac{\langle c_i^k, \bar{x}\rangle}{|g(\mu_k, x^k, u^k, v^k)| + r}, \quad \forall\, k \in \mathcal{J}, \quad \forall\, i \in N_{i^*}(x^*),$$

因此,

$$\sum_{i\in N_{i^*}(x^*)} \lambda_i(\hat{x}^k) \geqslant \frac{\left\langle \sum_{i\in N_{i^*}(x^*)} c_i^k, \bar{x}\right\rangle}{|g(\mu_k, x^k, u^k, v^k)| + r}, \quad \forall\, k \in \mathcal{J}. \tag{5.9}$$

由函数 g 的定义, 可以看到序列 $\{g(\mu_k, x^k, u^k, v^k)\}$ 具有极限点, 记作 $g(0, x^*, u^*, v^*)$, 则 $|g(0, x^*, u^*, v^*)| + r > 0$. 记 $N_{i^*}(x^*)$ 中的元素个数为 $r_{i^*}(x^*)$, 则令 $k \to \infty$, 由 (5.6) 和 (5.9) 式可得 $\lambda_{i_0}(x^*) \geqslant \frac{\langle b_{i^*}^*(x^*), \bar{x}\rangle}{r_{i^*}(x^*)(|g(0,x^*,u^*,v^*)|+r)}$. 因为 $b_{i^*}^*(x^*)$ 是幂等元而且 $\bar{x} \succeq 0$, 由引理 2.3(i) 可以得到 $\langle b_{i^*}^*(x^*), \bar{x}\rangle \geqslant 0$, 这个连同假设 $\langle b_{i^*}^*(x^*), \bar{x}\rangle \neq 0$, 可得 $\langle b_{i^*}^*(x^*), \bar{x}\rangle > 0$, 因此 $\lambda_{i_0}(x^*) > 0$.

由上我们得到: 对所有的 $i_0 \in \{i \in \mathcal{I} : \lambda_i(\bar{x}) > 0\}$ 都有 $\lambda_{i_0}(x^*) > 0$ 成立. 类似的可以得到: 对所有的 $j_0 \in \{i \in \mathcal{I} : \lambda_j(\bar{s}) > 0\}$ 都有 $\lambda_{j_0}(s^*) > 0$ 成立. 这些连同 (x^*, s^*) 是 SCCP 的解以及 $(\bar{x}, \bar{s})$ 是 SCCP 的极大互补解, 得到 (x^*, s^*) 是 SCCP 的一个极大互补解. 因此, 我们证明了这个定理.

推论 5.2 假设除条件 (A) 外定理 5.1 的其余条件都满足. 如果 x^* 和 $\bar{x}$ 以及 s^* 和 $\bar{s}$ 分别算子可交换, 则由算法 3.3 得到的解 (x^*, s^*) 是 SCCP 的一个极大互补解.

证明 事实上, 如果 x^* 和 $\bar{x}$ 算子可交换, 则它们具有一个相同的 Jordon 系 $\{c_1^*, \ldots, c_r^*\}$, 那么对任意的 $i_0 \in \{i \in \mathcal{I} : \lambda_i(\bar{x}) > 0\}$ 满足 $\lim_{k\to\infty} \lambda_{i_0}(\hat{x}^k) = \gamma_{i^*}(x^*)$, 可以得到 $\langle c_{i_0}^*, \bar{x}\rangle = \lambda_{i_0}(\bar{x})\langle c_{i_0}^*, c_{i_0}^*\rangle = \lambda_{i_0}(\bar{x})\|c_{i_0}^*\|^2 > 0$, 从而 $\langle b_{i^*}^*(x^*), \bar{x}\rangle > 0$, 即: 定理 5.1 中假设 (A) 的第 1 个条件满足. 类似地可以得到定理 5.1 中假设 (A) 的第 2 个条件满足. 因此, 推论的结论可直接从定理 5.1 得到.

6 总结

本文提出了一个求解单调 SCCP 的具有非单调线搜索的光滑算法 (即: 算法 3.3). 我们证明了在 SCCP 的解集是非空的假设下, 算法 3.3 是全局收敛的. 这样的假设甚至在 $K = \Re^n_+$ 的情形下都比很多的光滑算法所要求的假设要弱. 此外我们证明了在某些假设下, 算法 3.3 得到的解是 SCCP 的一个极大互补解. 作为进一步的研究, 值得考虑是否可以设计一个算法, 它同时具有本文得到的好的全局收敛性和局部快速收敛性? 这样的问题甚至在 $K = \Re^n_+$ 的情形下也没有解决. 另一个值得考虑的问题是本文中所考虑的算法的全局收敛的假设是否可以应用到非单调的问题上. 在 $K = \Re^n_+$ 的情形下, 文献 [24] 给出了部分回答.

致谢 作者衷心感谢两位评审专家对本文提出的宝贵建议.

参考文献

1 Huang Z H. Sufficient conditions on nonemptiness and boundedness of the solution set of the P_0 function nonlinear complementarity problem. *Oper Res Lett*, **30**: 202–210 (2002)

2 Huang Z H. Locating a maximally complementary solution of the monotone NCP by using non-interior-point smoothing algorithms. *Math Methods Oper Res*, **61**: 41–55 (2005)

3 Zhao Y B, Li D. A globally and locally superlinearly convergent non-interior-point algorithm for P_0 LCPs. *SIAM J Optim*, **13**: 1196–1221 (2003)

4 Zhou G L, Sun D F, Qi L Q. Numerical experiments for a class of squared smoothing Newton Methods for box constrained variational inequality problems. In: Fukushima M, Qi L, eds. Reformulation-Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods. Boston: Kluwer Academic Publishers, 1999, 421–441

5 Gowda M S, Sznajder R. Automorphism invariance of and P- and GUS-properties of linear transformations in Euclidean Jordan algebras. *Math Oper Res*, **31**: 109–123 (2006)

6 Gowda M S, Sznajder R, Tao J. Some P-properties for linear transformations on Euclidean Jordan algebras. *Linear Algebra Appl*, **393**: 203–232 (2004)

7 Kong L C, Xiu N H. New smooth C-functions for symmetric cone complementarity problems. *Optim Lett*, **1**(4): 391–400 (2007)

8 Liu Y J, Zhang L W, Wang Y H. Some properties of a class of merit functions for symmetric cone complementarity problems. *Asia-Pac J Oper Res*, **23**: 473–495 (2006)

9 Sun D F, Sun J. Löwner's operator and spectral functions in Euclidean Jordan algebras. *Math Oper Res*, **33**: 421–445 (2008)

10 Yoshise A. Interior point trajectories and a homogeneous model for nonlinear complementarity problems over symmetric cones. *SIAM J Optim*, **17**: 1129–1153 (2006)

11 Faybusovich L. Euclidean Jordan algebra and interior-point algorithm. *Positivity*, **1**: 331–357 (1997)

12 Faybusovich L. Linear systems in Jordan algebras and primal-dual interior-point algorithm. *J Comput Appl Math*, **86**: 149–175 (1997)

13 Schmieta V, Alizadeh F. Associative and Jordan algebras, and polynonial time interior-point algorithms for symmetrc cones. *Math Oper Res*, **26**: 543–564 (2001)

14 Schmieta S, Alizadeh F. Extension of primal-dual interior-point algorithms to symmetric cones. *Math Program*, **96**: 409–438 (2003)

15 Lu Y, Yuan Y X. An interior-point trust region algorithm for general cone progrmming. *SIAM J Optim*, **18**: 65–86 (2007)

16 Liu Y J, Zhang L W, Wang Y H. Analysis of smoothing method for symmetric conic linear programming. *J Appl Math Comput*, **22**: 133–148 (2006)

17 Kong L C, Sun J, Xiu N H. A regularized smoothing Newton method for symmetric cone complementarity problems. *SIAM J Optim*, **19**(3): 1028–1047 (2008)

18 Huang Z H, Ni T. Smoothing algorithms for complementarity problems over symmetric cones. DOI 10.1007/s10589-008-9180-y. *Comput Optim Appl*, to appear, 2008

19 Kong L C, Tunçel L, Xiu N H. Vector-valued implicit Lagrangian for symmetric cone complementarity problems. *Asia-Pac J Oper Res*, to appear, 2008

20 Pan S H, Chen J S. A one-parametric class of merit functions for the symmetric cone complementarity problem. Department of Mathematics, Taiwan Normal University, preprint, 2008

21 Faraut U, Korányi A. Analysis on Symmetric Cones. In: Oxford Mathematical Monographs. New York: Oxford University Press, 1994

22 Karamardian S. Complementarity problems over cones with monotone and pseudomonotone maps. *J Optim Theory Appl*, **18**: 445–454 (1976)

23 Zhang H C, Hager W W. A nonmonotone line search technique and its application to unconstrained optimization. *SIAM J Optim*, **14**: 1043–1056 (2004)

24 Huang Z H, Xu S W. Convergence properties of a non-interior-point smoothing algorithm for the P_* NCP. *J Ind Manag Optim*, **3**: 569–584 (2007)