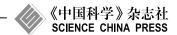
link.springer.com

math.scichina.com



矫正低秩相关系数矩阵的松弛序列凸近似方法

王博①*、张立卫①、初丽①②

- ① 大连理工大学数学科学学院, 大连 116024;
- ② 大连理工大学城市学院, 大连 116600

 $E-mail: bowang@mail.dlut.edu.cn, \\ lwzhang@dlut.edu.cn, \\ sophiatruly@126.com$

收稿日期: 2013-09-14; 接受日期: 2014-12-22; * 通信作者 国家自然科学基金 (批准号: 11071029 和 91130007) 资助项目

摘要 本文主要讨论带有秩约束以及简单上下界约束的相关系数矩阵矫正问题的求解方法. 该问题可以写成一个含有 DC (两个凸函数之差) 约束的优化问题,于是考虑利用求解 DC 优化问题的序列凸近似 (SCA) 方法求解. 然而对本文讨论的问题,经典的序列凸近似方法收敛所需的约束规范不成立,于是,本文提出一种松弛的序列凸近似方法. 本文证明当松弛参数趋于零时,松弛的 DC 问题的稳定点趋于原问题的稳定点.另一方面,可以利用序列凸近似方法求解松弛的 DC 问题. 可以证明,序列凸近似方法生成的一系列凸子问题的解的聚点就是该松弛 DC 问题的稳定点. 数值实验验证了该方法的有效性.

关键词 DC 优化 相关系数矩阵 低秩问题 **MSC (2010) 主题分类** 49M05, 90C22, 90C26

1 引言

一个半正定矩阵, 如果其对角线元素都为数字 1, 则称这个矩阵为一个相关系数矩阵. 相关系数矩阵是统计学的基本概念, 它在诸如数值代数、金融和保险等领域有着重要的应用^[1,2]. 最近的数十年间, 相关系数矩阵的矫正问题逐渐受到关注, 发展出了很多方法 (参见文献 [2–4]). 最新的成果参见文献 [5,6].

最近几年,一个更加有挑战性的问题引起了学者们的注意:如何找到一个满足低秩条件且与已知矩阵"最接近"的协方差矩阵?计量经济学中提出了很多这种类型的问题.文献 [7-9] 讨论了 LIBOR (London inter-bank offered rate) 和交换市场模型下的利率导数定价问题. Lillo 和 Mantegna [10] 讨论了多变量时间序列的因子模型,该模型也基于低秩最优相关系数矩阵.一个低秩的相关系数矩阵在因子数少的模型中更有优势,因为这样能显著减少 Monte Carlo 模拟中生成随机数的代价.另一方面,工程领域也涉及低秩相关系数矩阵,参见文献 [11-13].

设 S^n 和 S^n_+ 分别表示 $n \times n$ 阶对称矩阵和对称半正定矩阵构成的矩阵锥. 符号 $\langle \cdot, \cdot \rangle$ 表示内积, $\| \cdot \|$ 表示由内积诱导的范数. 设 C 是 S^n 中的一个给定的矩阵, r 是一个给定的正整数.

英文引用格式: Wang B, Zhang L W, Chu L. A relaxed SCA approach for calibrating low rank correlation matrix problem (in Chinese). Sci Sin Math, 2015, 45: 411–426, doi: 10.1360/N012013-00136

本文主要讨论如下带有简单上下界约束和低秩约束的最优相关系数矩阵矫正问题:

min
$$\frac{1}{2}||X - C||^2,$$
s.t.
$$X_{ii} = 1, \quad 1, \dots, n,$$

$$X_{ij} = e_{ij}, \quad (i, j) \in \mathcal{B}_e,$$

$$X_{ij} \geqslant l_{ij}, \quad (i, j) \in \mathcal{B}_l,$$

$$X_{ij} \leqslant u_{ij}, \quad (i, j) \in \mathcal{B}_u,$$

$$X \in S_+^n,$$

$$\operatorname{rank}(X) \leqslant r,$$

$$(1.1)$$

其中 \mathcal{B}_e , \mathcal{B}_l 和 \mathcal{B}_u 是指标集, 且满足条件 $\mathcal{B}_e \cap \mathcal{B}_l = \emptyset$, $\mathcal{B}_e \cap \mathcal{B}_u = \emptyset$, $|\mathcal{B}_e| = k_1$, $|\mathcal{B}_l| = k_2$, $|\mathcal{B}_u| = k_3$. 令

$$\mathcal{A}X = \begin{bmatrix} \langle E_{ii}, X \rangle, & i = 1, \dots, n \\ \langle E_{ij}, X \rangle, & (i, j) \in \mathcal{B}_e \\ -\langle E_{ij}, X \rangle, & (i, j) \in \mathcal{B}_l \\ \langle E_{ij}, X \rangle, & (i, j) \in \mathcal{B}_u \end{bmatrix}, \quad b = \begin{bmatrix} 1_n \\ \{e_{ij}\}, (i, j) \in \mathcal{B}_e \\ \{-l_{ij}\}, (i, j) \in \mathcal{B}_l \\ \{u_{ij}\}, (i, j) \in \mathcal{B}_u \end{bmatrix},$$

其中 E_{ij} 是一个 $n \times n$ 阶零矩阵, 只有第 (i,j) 个分量被换成了 1, 且 $\mathbb{Q} = \{0^{n+k_1}\} \times \mathbb{R}^{k_2+k_3}$, 于是可以把问题 (1.1) 重新写成更 "紧凑"的形式:

min
$$\theta(X) := \frac{1}{2} ||X - C||^2,$$

s.t. $\mathcal{A}X \in b + \mathbb{Q},$
 $X \in S^n_+,$
 $\operatorname{rank}(X) \leqslant r.$ (1.2)

类似的问题在许多文献中被讨论过. Simon [14] 综述了 13 种不同的方法. Gao 和 Sun [15] 也给出了详细的综述,本文不再赘述. 详细的内容参见文献 [3,16,17].

我们需要指出, Gao 和 Sun [15] 提出的 MPA (majorized penalty approach) 方法, 可以求解更一般 的目标函数中含有所谓 "H 权重" 范数的模型. 本文提出的方法理论上可以应用到目标函数是任意凸函数的问题, 比 "H 权重"模型更一般. 之所以限制为无权重的问题, 是因为 "H 权重"的凸子问题目前还没有足够高效的求解算法. 而其他文献中的方法也都无法求解带有上下界约束的问题 (1.1).

在接下来的讨论中, 用记号 $N_{\Omega}(x)$, $\delta_{\Omega}(\cdot)$ 和 $\Pi_{\Omega}(\cdot)$ 分别表示集合 Ω 在 x 点处的法锥, 集合 Ω 在 Ω 凸分析意义下的指示函数, 即

$$\delta_{\Omega}(x) = \begin{cases} 0, & \text{if } x \in \Omega, \\ +\infty, & \text{if } M \end{cases}$$

和到凸集合 Ω 上的投影映射.

2 预备知识

为了刻画可行集合的性质, 需要讨论集值映射的极限. 考虑集值映射 $M: H^n \rightarrow H^m$, 其中 H^n

和 H^m 分别是 n 维和 m 维的 Hilbert 空间. 下面的定义来自文献 [18, 第 152 页]. M 在 $\bar{\epsilon}$ 处的外极限定义为

$$\limsup_{\varepsilon\to\bar\varepsilon}M(\varepsilon):=\{x:\exists\,\varepsilon^k\to\bar\varepsilon,\exists\,x^k\in M(\varepsilon^k),x^k\to x\};$$

内极限定义为

$$\liminf_{\bar{\omega}} M(\varepsilon) := \{x: \forall \, \varepsilon^k \to \bar{\varepsilon}, \exists \text{ } \bar{\mathbb{E}} \text{ } \underline{\mathbb{E}} \text{ } \underline{\mathbb{Y}} \text{ } N, \forall \, k > N, \exists \, x^k \in M(\varepsilon^k), x^k \to x \}.$$

如果内外极限相等, 则把它们定义为集值映射 M 在 ε 处的极限, 记为 $\lim_{\varepsilon \to \varepsilon} M(\varepsilon)$. 如果 M 满足

$$\limsup_{\varepsilon \to \bar{\varepsilon}} M(\varepsilon) \subseteq M(\bar{\varepsilon}),$$

则称 M 是外半连续的; 如果

$$\liminf_{\varepsilon \to \bar{\varepsilon}} M(\varepsilon) \supseteq M(\bar{\varepsilon}),$$

则称 M 是内半连续的; 如果 M 既是内半连续又是外半连续的, 则称 M 在 $\bar{\varepsilon}$ 处连续, 即当 $\varepsilon \to \bar{\varepsilon}$ 时, $M(\varepsilon) \to M(\bar{\varepsilon})$.

因为计算的需要, 本文引入 Ky Fan r- 范数的概念. 如果 X 是一个 n 维对称矩阵, 即 $X \in S^n$, 则可以假设它有如下的特征值分解:

$$X = Q\Lambda Q^{\mathrm{T}} = Q\mathrm{Diag}(\lambda_1, \lambda_2, \dots, \lambda_n)Q^{\mathrm{T}},$$

其中 Q 是正交矩阵. 这些特征值的绝对值记为 $\hat{\mu_i} := |\lambda_i|$. 把它们按照降序排列, 重新记为 $\mu_1 \ge \mu_2 \ge \dots \ge \mu_n$. X 处的 Ky Fan r- 范数, 记为 $\|X\|_{(r)}$, 定义为 $\|X\|_{(r)} := \sum_{i=1}^r \mu_i$. 当 r = n 时, 这个范数也称为核范数, 记为 $\|X\|_*$.

下面的讨论只限于 $X \in S^n_+$ 时的情形. 正交矩阵 Q 可以写成分块的形式:

$$Q = [P_1 \ Q_1 \ P_0],$$

其中 P_1 是一个 $n \times r_1$ 阶的矩阵, 对应大于 μ_r 的特征值; Q_1 是 $n \times t_1$ 阶矩阵, 对应等于 μ_r 的特征值; 矩阵 P_0 对应余下的部分.

下面的经典结果刻画了范数 $\|\cdot\|_{(r)}$ 在对称情形下的次微分, 最早可以追溯到文献 [19].

引理 1 Ky Fan r- 范数在 $X \in S^n_+$ 处的次微分是如下的非空紧致凸集合:

$$\partial \| \cdot \|_{(r)}(X) = \{ W \in S^n : 存在 \tilde{U} \in S^{t_1} \notin \mathcal{P}_1 \cup \tilde{U} \leq I,$$

 $\operatorname{trace}(\tilde{U}) = r - r_1, W = P_1 P_1^{\mathrm{T}} + Q_1 \tilde{U} Q_1^{\mathrm{T}} \}.$

注 1 集合 $\partial \|\cdot\|_{(r)}(X)$ 中的一个典型的元素是 $\tilde{P}\tilde{P}'$, 其中 \tilde{P} 是正交矩阵 Q 的前 r 列. 本文数值实验中次微分集合中的一个元素即按此方式取得.

不等式系统的误差估计, 可以通过如下的 Hoffman 引理[20] 完成.

引理 2 对相容的不等式系统 $Ax \leq b$, 其中 $b \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times m}$, 存在一个常数 L, 使得对任意的 $x \in \mathbb{R}^m$, 都存在该不等式系统的解 x_0 , 满足

$$||x - x_0|| \le L ||\Pi_{\mathbb{R}^n} (Ax - b)||.$$

3 一个一般的序列凸近似算法框架

本节讨论一个一般的序列凸近似 (SCA) 算法框架, 用来求解目标是水平有界凸函数, 约束是凸约束加 DC 约束的优化问题. 具体而言, 我们讨论如下模型的 SCA 算法:

min
$$f(x)$$
,
s.t. $x \in \Omega$, (3.1)
 $g_1(x) - g_2(x) \leq 0$,

其中 $\Omega := \{x \in \mathcal{X} \mid \mathcal{A}_e x = b_e, h(x) \leq 0, x \in \mathcal{C}\}, \mathcal{X}$ 是有限维 Hilbert 空间, \mathcal{A}_e 是有界线性算子, $\mathcal{A}_e : \mathcal{X} \to \mathbb{R}^p, b_e \in \mathbb{R}^p$; h, g_1 和 g_2 是从 \mathcal{X} 到 \mathbb{R} 的凸函数; $\mathcal{C} \subseteq \mathcal{X}$ 是一个凸集合.

为了求解上述问题 (3.1), 我们要考虑对应问题 (3.1) 的凸子问题 $CP(y,\xi)$:

min
$$f(x)$$
,
s.t. $x \in \Omega$, (3.2)
 $g_1(x) - g_2(y) - \langle \xi, x - y \rangle \leq 0$,

其中 $y \in \Omega$, $\xi \in \partial g_2(y)$. 问题 (3.1) 和 (3.2) 的可行域分别为记为 \mathbb{F} 和 $\mathbb{F}(y,\xi)$.

研究问题 (3.1) 和 (3.2) 的性质, 需要如下的基本假设和稳定点定义.

假设 1 可行域 \mathbb{F} 非空. 函数 f, h, g_1 和 g_2 是连续的凸函数, 且函数 f 还是水平有界的.

定义 1 设 \bar{x} 是问题 (3.1) 的一个可行点, 即 $\bar{x} \in \mathbb{F}$. 如果条件

$$0 \in \partial f(\bar{x}) + (\partial g_1(\bar{x}) - \partial g_2(\bar{x})) N_{\mathbb{R}_-}(g_1(\bar{x}) - g_2(\bar{x})) + N_{\Omega}(\bar{x})$$

$$(3.3)$$

成立, 则称 \bar{x} 为问题 (3.1) 的一个稳定点.

问题 (3.1) 可以通过序列凸近似 (SCA) 方法求解, 参见文献 [21,22]. 这里给出 SCA 方法的一种形式.

算法 1 序列凸近似 (SCA) 方法.

第 1 步 选定 $\delta > 0$ 和初值 x^0 , 使得 $g_1(x^0) - g_2(x^0) \leq 0$, 令 k = 0;

第 2 步 取 $\xi^k \in \partial g_2(x^k)$. 求解问题 $CP(x^k, \xi^k)$:

$$\label{eq:force_eq} \begin{aligned} & \min \quad f(x), \\ & \text{s.t.} \quad x \in \Omega, \\ & g_1(x) - g_2(x^k) - \langle \xi^k, x - x^k \rangle \leqslant 0, \end{aligned}$$

且把其解记为 x^{k+1} ;

第 3 步 如果 $||x^{k+1} - x^k|| \le \delta$, 则终止, 否则设 k = k + 1, 转第 2 步.

注 2 由后面的定理 5, 此处初值 x^0 应满足如下条件: 存在一个 $\xi^0 \in \partial g(x^0)$, 使得子问题 $\mathrm{CP}(x^0,\xi^0)$ 的可行集 $\mathbb{F}(y,\xi)$ 非空.

算法 1 的收敛性证明依赖于如下引理, 参见 Rockafellar 和 Wets 的专著 [18, 定理 1.17].

引理 3 考虑参数优化问题

$$p(u) := \inf_x f(x,u), \quad P(u) := \arg\min_x f(x,u),$$

其中 $f: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ 是一个正常的下半连续函数, 且满足 f(x,u) 关于变量 x 是水平有界的, 关于变量 u 是局部一致的, 则有

- (a) 函数 $p \in \mathbb{R}^m$ 上的正常下半连续函数, 且对每个 $u \in \text{dom } p$, 集合 P(u) 是非空紧致的, 其中 P(u) 在 $u \notin \text{dom } p$ 时为空集;
- (b) 如果 $x^k \in P(u^k)$, $u^k \to \bar{u} \in \text{dom } p$ 且满足 $p(u^k) \to p(\bar{u})$, 则序列 $\{x^k\}_{k \in \mathbb{N}}$ 是有界的, 且它的所有聚点都在 $P(\bar{u})$ 中;
- (c) 保证 p 在点 \bar{u} 处相对于一个包含 \bar{u} 的集合 U 连续的一个充分条件是, 存在某个 $\bar{x} \in P(\bar{u})$ 使得 $f(\bar{x},u)$ 关于变量 u 在 \bar{u} 处相对于集合 U 连续.

SCA 方法的收敛性依赖于如下的假设.

假设 2 问题 $CP(y,\xi)$ 的 Slater 条件^[23] 成立, 即存在 \bar{z} , 使得

$$\begin{split} &\mathcal{A}_e\bar{z}=b_e,\quad h(\bar{z})<0,\\ &\bar{z}\in\operatorname{int}\mathcal{C},\\ &g_1(\bar{z})-g_2(y)+\langle\xi,\bar{z}-y\rangle<0 \end{split}$$

成立.

引理 4 如果存在 $\bar{\xi}$ 和 \bar{y} , 满足 $\bar{\xi} \in \partial g_2(\bar{y})$ 和 $\bar{y} \in \arg\min(\operatorname{CP}(\bar{y},\bar{\xi}))$, 且假设 2 对问题 $\operatorname{CP}(\bar{y},\bar{\xi})$ 成立, 则 \bar{y} 是问题 (3.1) 的一个稳定点.

证明 注意到 $\mathrm{CP}(\bar{y},\bar{\xi})$ 是一个凸优化问题, 且广义的 Slater 条件成立, 于是, 点 \bar{y} 是问题 $\mathrm{CP}(\bar{y},\bar{\xi})$ 的最优解当且仅当 $\bar{y}\in\mathbb{F}(\bar{y},\bar{\xi})$ 且

$$\begin{cases}
0 \in \partial f(\bar{y}) + \lambda(\partial g_1(\bar{y}) - \bar{\xi}) + N_{\Omega}(\bar{y}), \\
\lambda \in N_{\mathbb{R}_-}(g_1(\bar{y}) - g_2(\bar{y}))
\end{cases}$$
(3.4)

成立. 根据定义 $1, \bar{y}$ 是问题 (3.1) 的稳定点.

有了上述准备工作, 现在可以表述算法 1 的收敛性定理. 类似定理可以在文献 [24] 中找到, 注意这里的条件与文献 [22] 中的条件略有不同, 文献 [24] 和本文采用的条件是目标函数的水平有界性, 而文献 [22] 中要求可行集的有界性. 出于完整性的考虑, 这里同时也给出证明过程.

定理 5 假设问题 $CP(x^0,\xi^0)$ 有可行解. 令 $\{x^k\}$ 是由算法 1 生成的序列, 且 $\delta=0$. 如果对于充分大的 $k,\,x^k\equiv\bar{x}$ 成立, 则 \bar{x} 是问题 (3.1) 的稳定点. 否则, $\{x^k\}$ 是有界的, 且至少有一个聚点. 进一步, 对任意一个聚点 \bar{x} , 一定存在点 $\bar{\xi}$ 、一个子序列 $\{x^{k_j}\}$ 和序列 $\{\xi^{k_j}\}$, 使得 $\xi^{k_j}\in\partial g(x^{k_j})$, $\xi^{k_j}\to\bar{\xi}$, $x^{k_j}\to\bar{x}$ 且 $\bar{\xi}\in\partial g_2(\bar{x})$. 如果假设 2 对问题 $CP(\bar{x},\bar{\xi})$ 成立, 则 \bar{x} 是问题 (3.1) 的一个稳定点.

证明 如果对充分大的 $k, x_k \equiv \bar{x}$ 成立, 则由引理 4 可得定理结论成立. 否则, 不妨设 $\bar{f}(x,y,\xi)$:= $f(x) + \delta_{\mathbb{F}(y,\xi)}(x)$. 由于 x^k 和 x^{k+1} 是问题 $\mathrm{CP}(x^k,\xi^k)$ 的可行解, 由 x^{k+1} 的最优性可知, $\{f(x^k)\}$ 是单调递减的序列. $\{x^k\}$ 无界将与 f 的水平有界性矛盾. 接下来不妨设 $x^k \to \bar{x}$. 由次微分的外半连续性, 存在 $\bar{\xi}$ 、序列 $\{x^k\}$ 的子序列 $\{x^{k_j}\}$ 和 $\{\xi^k\}$ 的子序列 $\{\xi^{k_j}\}$, 满足 $\xi^{k_j} \in \partial g(x^{k_j})$, $\xi^{k_j} \to \bar{\xi}$, $x^{k_j} \to \bar{x}$ 且 $\bar{\xi} \in \partial g_2(\bar{x})$.

直接验证可以得到:

- (i) $\bar{f}(x,y,\xi)$ 关于变量 x 是水平有界的, 且该水平有界性关于变量 (y,ξ) 是局部一致的;
- (ii) \bar{f} 是闭的、下半连续的.

П

由引理 3(b), 把其中的参数 \bar{u} , u^k 和 x^k 对应地用 $(\bar{x},\bar{\xi})$, (x^{k_j},ξ^{k_j}) 和 x^{k_j+1} 代替, 得到子问题最优解集合的外半连续性, 即

$$\limsup_{i \to \infty} \arg \min_{x} \operatorname{CP}(x^{k_j}, \xi^{k_j}) \subseteq \arg \min_{x} \operatorname{CP}(\bar{x}, \bar{\xi}). \tag{3.5}$$

设 $\{x^{k_j+1}\}$ 的一个聚点为 \hat{x} . 由 $\{f(x^k)\}$ 的单调性得 $f(x^{k_j}) \ge f(x^{k_j+1}) \ge f(x^{k_{j+1}})$,于是,由 $f(x^{k_j}) \to f(\bar{x})$ 可得 $f(x^{k_j+1}) \to f(\hat{x}) = f(\bar{x})$. 再由 $\hat{x} \in \arg\min_x \operatorname{CP}(\bar{x}, \bar{\xi})$ 和 \bar{x} 对问题 $\operatorname{CP}(\bar{x}, \bar{\xi})$ 的可行性得 $\bar{x} \in \arg\min_x \operatorname{CP}(\bar{x}, \bar{\xi})$. 最后应用引理 4 可得定理结论成立. 证毕.

4 秩约束二次半定规划问题

正如许多作者都注意到的 (如文献 [15]), 秩约束 $rank(X) \leq r$ 等价于

$$||X||_* - ||X||_{(r)} = 0$$
 $\vec{\boxtimes}$ $||X||_* - ||X||_{(r)} \le 0.$

通过设 $g_1(X) := \|X\|_*, g_2(X) := \|X\|_{(r)}, f(X) := \frac{1}{2}\|X - C\|^2$ 及 $\Omega := \{X \in S^n_+ \mid \mathcal{A}X \in b + \mathbb{Q}\}$,问题 (1.2) 可以写成如下的 DC 优化模型的形式:

min
$$f(X)$$
,
s.t. $X \in \Omega$, (4.1)
 $g_1(X) - g_2(X) \leq 0$.

一种直接的想法是应用算法 1. 然而通过如下简单的验证可知, 上述问题的 Slater 条件不成立, 无法保证收敛性: 由核范数和 k- 范数的定义以及 X 的半正定性, 不存在 X, 使得

$$g_1(X) - g_2(X) < 0.$$

为了克服上述问题, 本文提出了一种 " ε - 松弛" 方法: 每步迭代求解松弛模型

$$\begin{aligned} & \min \quad \frac{1}{2} \|X - C\|^2, \\ & \text{s.t.} \quad \mathcal{A}X \in b + \mathbb{Q}, \\ & \quad X \in S^n_+, \\ & \quad \|X\|_* - \|X\|_{(r)} \leqslant \varepsilon. \end{aligned} \tag{4.2}$$

当 ε 充分小的时候, 可以认为原问题 (1.2) 已经被近似求解了. 给定矩阵 $\bar{X} \in S^n_+$, 设 $w \in \partial \|\bar{X}\|_{(r)}$, 问题 (4.2) 在点 (\bar{X}, w) 处的对应凸子问题是

min
$$\frac{1}{2} ||X - C||^2$$
,
s.t. $\mathcal{A}X \in b + \mathbb{Q}$,
 $X \in S_+^n$,
 $\langle I, X \rangle - g_2(\bar{X}) + \langle w, X - \bar{X} \rangle - \varepsilon \leqslant 0$. (4.3)

这样就避免了 Slater 约束规范不成立的问题.

对 S^n 中的一个固定的元素 ω 和 S^n_+ 中取定的 \bar{X} , 令

$$\Phi_{\varepsilon}(\bar{X}, w) := \{ X \in S_{+}^{n} \mid \mathcal{A}X \in b + \mathbb{Q}, \langle I, X \rangle - g_{2}(\bar{X}) + \langle w, X - \bar{X} \rangle - \varepsilon \leqslant 0 \},
\Upsilon_{\varepsilon}(\bar{X}, w) := \{ X \in S^{n} \mid \mathcal{A}X \in b + \mathbb{Q}, \langle I, X \rangle - g_{2}(\bar{X}) + \langle w, X - \bar{X} \rangle - \varepsilon \leqslant 0 \},$$

则如下的估计成立.

引理 6 存在一个常数 c > 0, 使得如下的包含关系成立:

$$\Phi_{\varepsilon}(\bar{X}, w) \subseteq (\Upsilon_0(\bar{X}, w) + \varepsilon c \mathbb{B}) \cap S_+^n. \tag{4.4}$$

当 $\varepsilon \to 0$ 时, $\Phi_{\varepsilon}(\bar{X}, w) \to \Phi_{0}(\bar{X}, w)$, 其中 B 表示闭单位球.

证明 由定义,

$$\Upsilon_{\varepsilon}(\bar{X}, w) \cap S^n_{\perp} = \Phi_{\varepsilon}(\bar{X}, w),$$

应用 Hoffman 引理 (这里把矩阵 X 看成向量, 算子 A 对应一个矩阵, 可以直接应用 Hoffman 引理), 存在常数 c>0, 使得对任意的 $X\in\Upsilon_{\varepsilon}(\bar{X},w)$, 存在一个 $X_0\in\Upsilon_0(\bar{X},w)$, 满足

$$||X - X_0|| \leqslant c \left| \left| \frac{\Pi_{\mathbb{Q}^{\circ}}(\mathcal{A}X - b)}{\Pi_{\mathbb{R}_+}(\langle I, X \rangle - g_2(\bar{X}) + \langle w, X - \bar{X} \rangle)} \right| \right| \leqslant c\varepsilon,$$

其中 ♥。表示 ♥ 的极锥. 这可以推出

$$\Upsilon_{\varepsilon}(\bar{X}, w) \subseteq \Upsilon_0(\bar{X}, w) + \varepsilon c \mathbb{B}.$$

两边同时与 S_+^n 取交集, 得 (4.4). 对 X, 如果存在趋于 0 的序列 $\{\varepsilon^k\}$, $X^k \in \Phi_{\varepsilon^k}(\bar{X}, w)$, 使得 $X^k \to X$, 则由 S_+^n 的闭性和线性函数在有限维空间中的连续性, 可得 X 满足 $\Phi_0(\bar{X}, w)$ 的定义, 于是有

$$\limsup_{\varepsilon \to 0} \Phi_{\varepsilon}(\bar{X}, w) \subseteq \Phi_{0}(\bar{X}, w).$$

另一方面, 由于 $\Phi_0(\bar{X}, w) \subseteq \Phi_{\varepsilon}(\bar{X}, w)$ 对任意 $\varepsilon > 0$ 成立, 所以, 对任意 $X \in \Phi_0(\bar{X}, w)$, 取 $X^k \equiv X$, 则 对任意趋于 0 的序列 ε^k , 有 $X^k \in \Phi_{\varepsilon^k}(\bar{X}, w)$, $X^k \to X$, 即有 $\liminf_{\varepsilon \to 0} \Phi_{\varepsilon}(\bar{X}, w) \supseteq \Phi_0(\bar{X}, w)$. 综上所述, 当 $\varepsilon \to 0$ 时, $\Phi_{\varepsilon}(\bar{X}, w) \to \Phi_0(\bar{X}, w)$. 证毕.

上述引理刻画了松弛子问题与原问题子问题可行集之间的近似关系. 这也从一个侧面提示我们设计一个有收敛性的 ϵ 松弛算法是可行的.

问题 (4.2) 的可行集记为 \mathbb{F}_{ε} , 即令 $\mathbb{F}_{\varepsilon} := \{X \in S^n_+ \mid AX \in b + \mathbb{Q}, g_1(X) - g_2(X) \leq \varepsilon\}$. 下面的命题 说明, 如果每个 ε 松弛子问题的解都能精确求得, 则当 $\varepsilon \to 0$ 时, 其极限就是原问题的解.

命题 7 设假设 1 成立. 如果令 $\varepsilon \to 0$, 则有

$$\limsup_{\varepsilon \to 0} \arg \min_{X \in \mathbb{F}_{\varepsilon}} f(X) \subseteq \arg \min_{X \in \mathbb{F}} f(X). \tag{4.5}$$

- (i) $\tau(X,\varepsilon)$ 是正常的下半连续的;
- (ii) $\tau(X,\varepsilon)$ 是关于变量 X 水平有界的, 且该水平有界性关于变量 ε 是局部一致的. 应用引理 3, 证毕.

实际计算中, 我们逐步减小 ε 来得到一系列的松弛子问题, 然后用每次得到的子问题的近似解来作为下一次迭代的初始值. 这样的迭代直到 ε 小于一个给定的阈值才终止. 换句话说, 前面的迭代都可以看成是后面问题的预处理. 这个过程可以写成如下的算法.

算法 2 ε - 松弛 SCA 算法.

第 1 步 选择 $\varepsilon^0 > 0$, 初值 X^0 , 参数 $0 < \rho < 1$ 以及一个较小的 $\bar{\varepsilon} > 0$, 令 k = 0;

第 2 步 以 X^k 为初值应用算法 1 解如下松弛问题:

$$\min \quad \frac{1}{2} ||X - C||^2,$$
s.t. $\mathcal{A}X \in b + \mathbb{Q},$

$$X \in S^n_+,$$

$$g_1(X) - g_2(X) - \varepsilon^k \leqslant 0,$$

$$(4.6)$$

且把此步骤的解记为 X^{k+1} ;

第 3 步 如果 $\varepsilon^k < \bar{\varepsilon}$, 则终止, 否则令 $\varepsilon^{k+1} = \rho \varepsilon^k$, k = k+1, 转第 2 步. 这里初值 X^0 可通过求如下凸问题得到,

min
$$\frac{1}{2} ||X - C||^2$$
,
s.t. $\mathcal{A}X \in b + \mathbb{Q}$, (4.7)
 $X \in S^n_{\perp}$,

这样得到的初始点 X^0 一定能满足子问题 (4.3) 的可行性: X^0 本身就是一个可行点. 参数 ε^0 可以选为 $\varepsilon^0 := \|X\|_* - \|X\|_r$, 因为这样可以保证问题 (4.2) 和 (4.3) 的相容性, 即两个问题的可行域同时非空. 一般地, 对一个给定的满足 $AY \in b + \mathbb{Q}$, $Y \in S^1$ 的 Y, 当 ε 充分大时, 问题 (4.3) 是可行的.

令 $\mathbf{S}_{\varepsilon} := \{ \text{问题 (4.2) 的所有稳定点} \}$. 下面的定理说明松弛问题 (4.2) 的稳定点在参数 $\varepsilon \to 0$ 时的任意一个聚点都是原问题的稳定点. 从而说明了算法 2 设计的合理性.

定理 8 如果假设 1 成立, 则有 $\limsup_{\varepsilon \to 0} \mathbf{S}_{\varepsilon} \subseteq \mathbf{S}_{0}$.

证明 对任意的 $\bar{X} \in \limsup_{\varepsilon \to 0} \mathbf{S}_{\varepsilon}$, 存在序列 $\{(X^k, \varepsilon^k)\}$ 满足 $\mathbf{S}_{\varepsilon^k} \ni X^k \to \bar{X}$ 和 $\varepsilon^k \to 0$. 于是有

$$\begin{cases} 0 \in \nabla f(X^k) + \lambda^k (\partial g_1(X^k) - \partial g_2(X^k)) + N_{\Omega}(X^k), \\ \lambda^k \in N_{\mathbb{R}_-}(g_1(X^k) - g_2(X^k) - \varepsilon^k). \end{cases}$$
(4.8)

由次微分和法锥的外半连续性,有

$$0 \in \limsup_{k \to \infty} N_{\mathbb{R}_{-}}(g_1(X^k) - g_2(X^k) - \varepsilon^k)(\partial g_1(X^k) - \partial g_2(X^k)) + \nabla f(X^k) + N_{\Omega}(X^k)$$

$$\subseteq \nabla f(\bar{X}) + N_{\mathbb{R}_{-}}(g_1(\bar{X}) - g_2(\bar{X}))(\partial g_1(\bar{X}) - \partial g_2(\bar{X})) + N_{\Omega}(\bar{X}),$$

由此可以推出 $\bar{X} \in \mathbf{S}_0$.

5 数值实验

本节报告若干数值实验结果. 实验中比较了算法 2 与 Gao 和 Sun 的 MPA 方法, 验证算法 2 的有效性. 数值实验中, 子问题由 Gao 和 Sun [6] 提出的非精确光滑化 Newton 法求解, 其中非精确光滑化 Newton 法的代码经过了少量的修改, 以适应序列凸子问题中包含的一般的线性约束.

具体数值实现中, 算法 1 当下述条件之一满足时终止:

- (a) $\frac{\theta(X^k) \theta(X^{k-1})}{\theta(X^{k-1})} < 10^{-3} \text{ } \exists \epsilon > 10^{-1};$
- (b) $\frac{\theta(X^k) \theta(X^{k-1})}{\theta(X^{k-1})} < 10^{-2} \text{ } (\|X^k\|_* \|X^k\|_r) < 10^{-10};$
- (c) $\frac{\theta(X^k) \theta(X^{k-1})}{\theta(X^{k-1})} < 10^{-5}$.

算法 2 中的 ε^0 和 ρ 按如下规则选取:

$$\varepsilon^0 = \min\left\{\frac{\sqrt{n}}{10}(\|X^k\|_* - \|X^k\|_r), 5\right\}, \quad \rho = \begin{cases} 0.2, & \text{ $ \pm $\'it is } \frac{\|X^k\|_* - \|X^k\|_{(r)}}{\mathrm{rank}(X^k)} < 0.1, \\ 0.3, & \text{ $ \pm $\'it it it is } \end{cases}$$

为了提高精度, 当 ε 足够小 (ε < 10^{-7}) 时, 数值实现中强制令 ε = 0. 调用的非精确 Newton 法的收敛条件与文献 [6] 中的终止条件相同, 误差精度设为 τ = 10^{-6} .

所有的实验都是在配置了 4 GB 内存、3.3 Ghz 的 Intel(R)Core(TM)2 处理器的台式计算机上进行的,程序都在 MATLAB 2011a 环境下运行. 其中 "times(s)" 代表运行时间,以秒为单位; "rank" 表示秩的大小; "residue" 表示所得结果与目标的差的模,即 $\|X-C\|_F$; "infeas" 度量解的可行性,定义为 $\|AX-b\|$; " \mathcal{E}_R " 度量矩阵秩的误差,定义为 $\|X\|_*-\|X\|_{(r)}$; "iter" 代表算法 2 迭代的次数; " N_{sub} " 代表求解子问题 (4.3) 的次数; " N_{eig} " 代表进行特征值分解的次数; " $\mathcal{E}_{\mathrm{rel}}$ " 代表函数 θ 值的相对改善量,即 $\frac{\theta(X^k)-\theta(X^{k-1})}{\theta(X^{k-1})}$.

为了简洁起见, 表格中将省略算法 2 和 MPA 的标注. 每组数据的第一行表示算法 2 的结果, 第 2 行表示 MPA 求解同一问题的结果.

例 1 矩阵 C 来自文献 [25], 其中维数 n = 11. 这是一个关于英镑 (GBP) 的前向的相关系数矩阵, 定义为

1	0.8415	0.6246	0.6231	0.5330	0.4287	0.3274	0.4463	0.2439	0.3326	0.2625	
0.8415	1	0.7903	0.7844	0.7320	0.6346	0.4521	0.5812	0.3439	0.4533	0.3661	
0.6246	0.7903	1	0.9967	0.8108	0.7239	0.5429	0.6121	0.4426	0.5189	0.4251	
0.6231	0.7844	0.9967	1	0.8149	0.7286	0.5384	0.6169	0.4464	0.5233	0.4299	
0.5330	0.7320	0.8108	0.8149	1	0.9756	0.5676	0.6860	0.4969	0.5734	0.4771	
0.4287	0.6346	0.7239	0.7286	0.9756	1	0.5457	0.6583	0.4921	0.5510	0.4581	
0.3274	0.4521	0.5429	0.5384	0.5676	0.5457	1	0.5942	0.6078	0.6751	0.6017	
0.4463	0.5812	0.6121	0.6169	0.6860	0.6583	0.5942	1	0.4845	0.6452	0.5673	
0.2439	0.3439	0.4426	0.4464	0.4969	0.4921	0.6078	0.4845	1	0.6015	0.5200	
0.3326	0.4533	0.5189	0.5233	0.5734	0.5510	0.6751	0.6452	0.6015	1	0.9889	
0.2625	0.3661	0.4251	0.4299	0.4771	0.4581	0.6017	0.5673	0.5200	0.9889	1	

指标集为 $\mathcal{B}_e = \mathcal{B}_l = \mathcal{B}_u = \emptyset$. 见表 1.

例 2 维数取 n=500. 矩阵 $C=E+\alpha R$, 其中 E 由 MATLAB 代码 gallery ("randcorr", n) 生成, 矩阵 R 是一个 $n\times n$ 的对称随机矩阵, 其每个分量 $R_{ij}\in [-1,1], i,j=1,2,\ldots,n$. 参数 $\alpha=0.1$. 指标集取 $\mathcal{B}_e=\mathcal{B}_l=\mathcal{B}_u=\emptyset$. 参见文献 [2,17]. 见表 2.

表 1	仮しっ	的结果
ᅏ	1411	HILL THE

rank	times(s)	residue	infeas	\mathscr{E}_R	iter	$N_{ m sub}$	$N_{ m eig}$	$\mathscr{E}_{\mathrm{rel}}$
2	5.25E - 01	2.25763E+00	3.070E - 08	1.783E - 15	13	26	165	2.170E-10
	5.27E - 01	2.25871E+00	2.111E-07	0.000E+00	2	7	24	$8.406E{-}06$
3	1.34E-01	1.49970E+00	5.620E - 06	$1.709E{-}15$	13	35	369	1.971E - 09
	1.48E-01	1.50213E+00	3.105E - 08	0.000E+00	3	28	86	9.727E - 06
4	1.38E - 01	$9.92227\mathrm{E}{-01}$	1.874E - 06	$1.139E{-}15$	13	26	282	$3.441E{-09}$
	3.70E - 02	$9.93131E{-01}$	1.423E - 08	0.000E+00	1	6	21	7.052E - 06
5	1.48E - 01	6.78057E - 01	$4.716E{-06}$	$2.289E{-}15$	12	28	340	2.045E-09
	8.00E - 02	6.79908E - 01	$3.361E{-08}$	0.000E+00	3	13	37	7.882E - 06

表 2 例 2 的结果

rank	times(s)	residue	infeas	\mathscr{E}_R	iter	$N_{ m sub}$	$N_{ m eig}$	$\mathscr{E}_{\mathrm{rel}}$
20	5.76470E+01	1.06637E+02	1.876E-06	4.941E-14	8	16	901	4.397E-06
	5.63800E+00	1.05760E + 02	2.775E-07	0.000E+00	4	10	32	7.281E-06
30	2.07320E+01	8.39473E+01	8.779E - 07	4.196E - 14	11	47	1108	9.674E - 08
	4.90600E+00	8.38681E + 01	5.705E - 07	0.000E+00	4	10	32	7.899E-06
40	1.50730E+01	7.04857E+01	1.688E - 06	$3.852E{-}14$	12	26	280	3.358E - 08
	4.59400E+00	7.04942E+01	7.493E-09	0.000E+00	4	10	33	7.464E - 06
50	1.46760E + 01	6.11342E+01	1.717E - 06	3.547E - 14	12	25	218	2.827E - 08
	4.52600E+00	6.11417E+01	3.547E - 011	0.000E+00	4	10	32	7.420E-06
60	1.05200E+01	5.40695E+01	9.517E - 07	$3.485E{-}14$	12	23	148	3.129E - 08
	4.62800E+00	5.40751E+01	7.628E - 09	0.000E+00	4	10	29	6.139E - 06
70	9.82300E+00	4.84670E + 01	2.835E - 06	$3.340E{-}14$	12	23	140	1.038E - 08
	4.97700E+00	4.84724E+01	3.883E - 07	0.000E+00	4	10	32	5.375E - 06
80	1.16110E + 01	4.38630E + 01	2.155E - 06	$3.171E{-14}$	12	23	161	$3.586E{-}08$
	4.83000E+00	4.38683E+01	9.454E - 09	0.000E+00	4	10	34	4.902E - 06
90	1.00720E+01	3.99565E+01	6.548E - 06	3.070E - 14	12	23	141	6.546E - 09
	5.06800E+00	3.99612E+01	$2.311E{-08}$	0.000E+00	4	10	34	$4.322E{-06}$
100	1.05860E + 01	3.65724E+01	1.345E - 06	$2.914E{-14}$	13	25	144	9.812E - 09
	3.90200E+00	3.65768E+01	$4.165E{-08}$	0.000E+00	3	8	27	7.697E - 06

例 3 (参见文献 [16,17]) 维数取 n = 500. "利率" 相关系数矩阵 C 是由如下方法生成:

$$C_{ij} = \exp\left\{-\gamma_1 |t_i - t_j| - \frac{\gamma_2 |t_i - t_j|}{\max(t_i, t_j)^{\gamma_3}} - \gamma_4 |\sqrt{t_i} - \sqrt{t_j}|\right\},\,$$

其中 $\gamma_2, \gamma_3, \gamma_4 > 0$, 参数 t_i 表示率 i 的有效期限, 这里取 $t_i = i, i = 1, 2, \ldots, n$, 参数 γ_i 在表 3 中给出. 指标集取 $\mathcal{B}_e = \mathcal{B}_l = \mathcal{B}_u = \emptyset$. 见表 4.

例 4 (参见文献 [3,17]) 维数 n=500. 矩阵 C 是 $n\times n$ 对称随机矩阵, 分量 $C_{ij}\in[-1,1],$ $C_{ii}=1,$ $i,j=1,2,\ldots,n,$ 即 C 可由代码 "C=rand(n);" "C=(C+C')/2;" "C=1-2*C;" "C=C-diag(diag(C))+eye(n)." 生成. 指标集取 $\mathcal{B}_e=\mathcal{B}_l=\mathcal{B}_u=\emptyset.$ 见表 5.

表 3 参数 γ_i , $i=1,\ldots,4$

	γ_1	γ_2	γ_3	γ_4
均值	0.000	0.480	1.511	0.186
标准差	_	0.099	0.289	0.127

表 4 例 3 的结果

rank	times(s)	residue	infeas	\mathscr{E}_R	iter	$N_{ m sub}$	$N_{ m eig}$	$\mathscr{E}_{\mathrm{rel}}$
20	1.97521E+02	1.23333E+02	1.870E-06	4.636E-14	9	26	7254	2.103E-05
	4.20290E+01	7.94708E+01	1.264E - 07	0.000E+00	6	85	319	1.177E - 05
30	1.63127E + 02	9.77278E+01	1.977E - 06	$4.231E{-}14$	9	22	5362	2.382E - 05
	5.09840E+01	5.47351E+01	2.571E - 07	0.000E+00	6	103	389	$1.205 E{-}05$
40	1.44382E+02	$4.93234\mathrm{E}{+01}$	1.123E-06	$3.427E{-}14$	9	18	1839	$1.540E{-}05$
	3.67580E + 01	$4.05271\mathrm{E}{+01}$	3.327E - 08	0.000E+00	6	81	304	$1.141E{-}05$
50	1.08987E + 02	3.79888E+01	1.172E - 06	3.217E - 14	8	17	2148	6.708E - 06
	3.23650E + 01	3.14431E+01	2.112E - 08	0.000E+00	6	78	273	$1.129E{-}05$
60	$4.17610\mathrm{E}{+01}$	$2.55349\mathrm{E}{+01}$	5.033E - 06	$2.832E{-}14$	9	19	1145	$3.501E{-07}$
	3.45360E + 01	$2.52514\mathrm{E}{+01}$	2.030E - 07	0.000E+00	6	86	295	$1.310E{-}05$
70	5.06990E + 01	$2.07394\mathrm{E}{+01}$	5.289E - 06	2.837E - 14	12	56	2513	1.828E - 08
	$2.02760\mathrm{E}{+01}$	$2.07571\mathrm{E}{+01}$	1.578E - 08	0.000E+00	7	48	165	$1.160E{-}05$
80	$2.29790\mathrm{E}{+01}$	1.74176E + 01	$6.061E{-}06$	2.770E - 14	12	46	994	$1.853E{-}08$
	1.43840E + 01	1.74320E+01	$2.250 E{-07}$	0.000E+00	6	36	119	1.093E-05
90	1.96480E + 01	1.48707E + 01	$3.025E{-06}$	$2.621E{-14}$	12	33	411	1.498E - 08
	9.41700E+00	1.48797E+01	2.113E-07	0.000E+00	5	23	74	9.884E - 06
100	1.69080E + 01	1.28653E+01	4.249E - 06	$2.633E{-}14$	13	25	231	4.505E-09
	4.33300E+00	1.28692E + 01	8.726E - 08	0.000E+00	3	9	29	7.923E-06

例 5 矩阵 C 是一个 n 阶方阵, 其每个分量定义为

$$C_{ij} = 0.5 + 0.5e^{-0.05|i-j|}, \quad i, j = 1, \dots, n.$$

该算例约束中都规定秩不大于 100, 维数 n 分别取 $n=1000,900,800,\ldots,500$. 对不同的 n, 约束集合并不相同, 但是生成的规则都是相同的. 其中我们随机固定 10 个非对角元素, 随机的在每一行中限制 20 个非对角元素的上界且限制 20 个非对角元素的下界. 对应的 b 按如下方式选取: 固定元素取 $\alpha_0 \times 1 + (1-\alpha_0) \times \eta$, 其中 η 在 [0,1] 中均匀分布, 参数 $\alpha_0=1$; 所有的下界都定为 -0.1; 上界都定为 0.1. 见表 6. **例 6** 维数 n=500. 矩阵 C 是一个 n 阶方阵, 其每个分量定义为

$$C_{ij} = 0.5 + 0.5e^{-0.05|i-j|}, \quad i, j = 1, \dots, 500.$$

指标集不同. $\mathcal{B}_e \neq \emptyset$, $\mathcal{B}_l \neq \emptyset$, $\mathcal{B}_u \neq \emptyset$. 其中我们随机固定 5 个非对角元素,随机的在每一行中限制 10 个非对角元素的上界且限制 15 个非对角元素的下界. 对应的 b 按如下方式选取: 固定元素取 $\alpha_0 \times 1 + (1 - \alpha_0) \times \eta$, 其中 η 在 [0,1] 中均匀分布,参数 $\alpha_0 = 1$; 所有的下界都定为 -0.1; 上界都定为 0.1. 这个算例来自文献 [15]. 见表 7.

表 5 例 4 的结果

rank	times(s)	residue	infeas	\mathscr{E}_R	iter	$N_{ m sub}$	$N_{ m eig}$	$\mathscr{E}_{\mathrm{rel}}$
20	2.01949E+03	2.87283E+02	1.622E+00	4.539E-14	12	41	42984	9.983E-02
	1.06180E + 01	1.93465E+02	9.511E - 08	0.000E+00	3	17	64	9.374E - 06
30	$1.07480\mathrm{E}{+01}$	1.83757E + 02	$4.955E{-}06$	$4.367E{-}14$	12	23	162	$1.215E{-08}$
	7.20300E+00	1.83764E + 02	4.443E - 07	0.000E+00	3	11	43	$9.925E{-06}$
40	9.92000E+00	1.79238E+02	$4.984E{-}06$	$3.695 E{-14}$	12	23	151	$1.136E{-}08$
	6.25200E+00	1.79244E+02	1.515E - 07	0.000E+00	3	9	38	9.115E - 06
50	$1.07650\mathrm{E}{+01}$	1.76899E + 02	1.607E - 06	$3.553E{-}14$	13	25	165	1.997E - 09
	5.69100E+00	1.76902E+02	$8.114E{-}10$	0.000E+00	3	8	33	4.838E - 06
60	8.89100E+00	$1.75614\mathrm{E}{+02}$	1.099E-06	$3.396E{-}14$	13	25	138	$8.976E{-}10$
	5.38200E+00	1.75616E + 02	$8.592E{-09}$	0.000E+00	3	7	29	$3.730 \mathrm{E}{-06}$
70	1.07760E + 01	1.74896E + 02	5.846E - 06	$3.311E{-14}$	13	25	181	$6.546E{-}10$
	$4.96200\mathrm{E}{+00}$	1.74897E + 02	$5.562E{-09}$	0.000E+00	3	7	29	$6.902 E{-07}$
80	9.14900E+00	1.74503E+02	3.380E - 06	$3.294E{-}14$	12	23	151	$1.737\mathrm{E}{-10}$
	3.69700E+00	1.74503E+02	1.181E - 08	0.000E+00	2	5	20	9.491E - 07
90	7.79500E+00	1.74307E + 02	4.713E - 06	$3.175E{-14}$	12	23	121	$8.073E{-}11$
	9.10700E+00	1.74307E + 02	$3.704 E{-08}$	0.000E+00	6	13	47	8.999E-09
100	8.01300E+00	1.74224E+02	3.629E - 06	$3.121E{-14}$	13	25	126	$3.240 \mathrm{E}{-10}$
	5.28500E+00	1.74224E+02	$3.456E{-}08$	0.000E+00	3	7	25	3.961E-09

表 6 例 5 的结果

\overline{n}	times(s)	vals	infeas	\mathscr{E}_R	iter	$N_{ m sub}$	$N_{ m eig}$	$\mathscr{E}_{\mathrm{rel}}$
1000	2.71140E+02	3.14569E+02	6.771E - 07	1.128E - 13	11	21	404	6.966E-09
	2.73991E+02	3.14576E + 02	3.482E - 08	0.000E+00	3	12	84	8.537E - 06
900	3.21843E+02	3.14616E + 02	$1.305E{-06}$	$1.208E{-}13$	11	21	471	6.968E - 09
	2.84404E+02	3.14622E+02	$1.065E{-}07$	0.000E+00	3	12	93	7.764E-06
800	$1.54101\mathrm{E}{+02}$	$2.57940\mathrm{E}{+02}$	1.307E - 06	$9.088E{-}14$	12	23	367	3.223E-09
	1.44790E + 02	$2.57949\mathrm{E}{+02}$	1.349E - 07	0.000E+00	3	10	72	9.041E - 06
700	$1.28519\mathrm{E}{+02}$	$2.29528\mathrm{E}{+02}$	$1.121E{-06}$	$8.369E{-}14$	13	25	485	1.809E-09
	1.44454E + 02	$2.29531\mathrm{E}{+02}$	2.278E - 07	0.000E+00	6	13	82	3.903E-06
600	8.06040E + 01	$2.00784\mathrm{E}{+02}$	$3.559E{-}06$	$6.678 E{-14}$	13	25	490	$2.741E{-}10$
	$7.21530\mathrm{E}{+01}$	$2.00786\mathrm{E}{+02}$	$9.780 E{-08}$	0.000E+00	4	9	55	$3.001E{-06}$
500	7.79180E+01	1.71537E+02	$1.670 E{-06}$	$4.872E{-}14$	14	27	602	$2.398E{-}10$
	9.08550E+01	1.71537E+02	2.358E - 08	0.000E+00	8	17	110	5.660E-07

例 7 维数 n=500. 矩阵 $C=E+\alpha R$, 其中 E 由 MATLAB 代码 "gallery ('randcorr',n)" 给出, 矩阵 R 是 $n\times n$ 阶对称随机矩阵, 其分量 $R_{ij}\in[-1,1],$ $i,j=1,2,\ldots,n$. 参数 $\alpha=0.1$. 我们随机固定 5 个非对角元素,随机的在每一行中限制 10 个非对角元素的上界且限制 15 个非对角元素的下界. 对应的 b 按如下方式选取: 固定元素取 $\alpha_0\times 1+(1-\alpha_0)\times \eta$, 其中 η 在 [0,1] 中均匀分布,参数 $\alpha_0=1$; 所

有的下界都定为 -0.1; 上界都定为 0.1. 见表 8.

例 8 维数 n=500. 矩阵 C 是一个 n 阶方阵, 其每个分量定义为 $C_{ij}=0.5+0.5\mathrm{e}^{-0.05|i-j|},$ $i,j=1,\ldots,n$. 指标集取 $\mathcal{B}_e=\mathcal{B}_l=\mathcal{B}_u=\emptyset$. 该算例已经在数篇文献中被测试过, 参见文献 [15,17]. 见表 9.

表 7 例 6 的结果

rank	times(s)	residue	infeas	\mathscr{E}_R	iter	$N_{ m sub}$	$N_{ m eig}$	$\mathscr{E}_{\mathrm{rel}}$
30	1.99914E+02	1.68351E+02	2.032E-06	4.914E-14	9	17	1851	1.399E-06
	4.28770E + 02	1.66702E+02	$2.366E\!-\!08$	0.000E+00	4	112	1910	9.933E - 06
40	8.02540E + 01	1.60714E + 02	1.189E - 06	$4.743E{-}14$	10	19	787	2.394E-07
	1.25024E+02	1.60555E+02	6.360E - 09	0.000E+00	4	41	393	9.725E - 06
50	8.89820E + 01	1.57736E+02	1.573E - 06	$5.530E{-}14$	10	19	669	$6.256E{-}08$
	7.04330E+01	1.57701E+02	6.135E-09	0.000E+00	3	24	181	9.330E - 06
60	5.01780E+01	1.56253E+02	1.186E - 06	5.197E - 14	12	23	385	9.518E - 09
	4.98270E + 01	1.56258E+02	1.079E - 08	0.000E+00	3	15	100	$8.485E{-06}$
70	3.66080E + 01	1.55556E+02	1.405E-06	5.173E - 14	12	23	249	5.919E-09
	5.60570E + 01	1.55561E + 02	4.648E - 09	0.000E+00	6	14	83	7.571E - 06
80	5.08230E + 01	1.55225E+02	1.593E - 06	$5.292E{-}14$	15	29	395	$4.913E{-}10$
	2.85270E + 01	1.55226E+02	1.725E - 08	0.000E+00	3	7	39	5.381E - 06
90	7.54740E+01	1.55102E+02	1.565E - 06	5.519E - 14	15	29	722	$3.105E{-}10$
	5.30210E+01	1.55102E+02	6.222E - 08	0.000E+00	6	13	65	2.125E-07
100	2.70830E + 01	1.55072E+02	1.000E-07	$5.554E{-}14$	15	29	255	$9.823E{-}11$
	1.19850E + 01	1.55072E+02	1.789E - 08	0.000E+00	1	3	16	$3.750E{-08}$

表 8 例 7 的结果

rank	times(s)	residue	infeas	\mathscr{E}_R	iter	$N_{ m sub}$	$N_{ m eig}$	$\mathscr{E}_{\mathrm{rel}}$
40	1.51803E+02	7.17460E+01	1.377E-06	3.978E-14	8	17	2771	1.112E-06
	5.99350E+01	7.13194E+01	1.025E - 08	0.000E+00	6	61	275	$1.295 E{-}05$
50	1.36692E + 02	$6.22479\mathrm{E}{+01}$	$1.251E{-06}$	$3.605E{-}14$	9	17	1783	$5.390 E{-07}$
	$4.67140\mathrm{E}{+01}$	6.19127E+01	6.177E - 08	0.000E+00	6	52	220	$1.271E{-}05$
60	9.46640E + 01	5.49929E+01	$9.414E{-07}$	$3.451E{-}14$	10	19	1250	$2.416E{-07}$
	3.84330E + 01	5.48173E + 01	$4.315E{-08}$	0.000E+00	6	45	181	$1.235 E{-}05$
70	1.44479E + 02	5.06406E + 01	$4.604E{-06}$	$3.335E{-}14$	9	17	1661	$3.574 \mathrm{E}{-06}$
	3.27170E + 01	$4.91808\mathrm{E}{+01}$	4.131E - 07	0.000E+00	6	38	155	$1.217E{-05}$
80	1.11506E+03	6.43088E+01	$8.236E{-02}$	$3.451E{-}14$	10	22	16870	$6.210 E{-03}$
	2.86360E + 01	$4.45372\mathrm{E}{+01}$	$1.065E{-}08$	0.000E+00	6	34	133	$1.149E{-}05$
90	$2.78850\mathrm{E}{+01}$	$4.05987\mathrm{E}{+01}$	$4.024E{-06}$	$3.046E{-}14$	13	48	985	9.380E - 09
	2.39940E+01	$4.06155\mathrm{E}{+01}$	8.279E - 08	0.000E+00	6	29	109	$1.154E{-}05$
100	2.00580E + 01	3.72072E+01	$3.356E{-}06$	2.967E - 14	13	34	384	9.486E - 09
	2.14410E+01	3.72290E+01	1.202E - 08	0.000E+00	6	25	95	$1.116E{-}05$

+ ~	/Til	_	的结果
- √ u	1411	×	HVIZ = H

rank	times(s)	residue	infeas	\mathscr{E}_R	iter	$N_{ m sub}$	$N_{ m eig}$	$\mathscr{E}_{\mathrm{rel}}$
20	8.21100E+00	1.57069E+01	5.443E-06	8.405E-14	13	25	145	3.603E-10
	3.20300E+00	1.57080E+01	$1.046E{-}07$	0.000E+00	2	6	22	$5.521E{-06}$
30	7.73300E+00	8.79545E+00	1.987E - 06	$8.532E{-}14$	13	25	140	$8.860E{-}10$
	$2.72400\mathrm{E}{+00}$	8.79626E+00	$1.336E{-}08$	0.000E+00	2	5	19	$4.742E{-06}$
40	$1.17190\mathrm{E}{+01}$	5.76440E+00	1.793E - 06	$1.014E{-}13$	14	27	207	$9.846E{-}10$
	$2.25600\mathrm{E}{+00}$	5.76462E+00	2.626E - 07	0.000E+00	2	5	18	1.890E - 06
50	1.12630E + 01	4.13918E+00	1.557E - 06	$1.001E{-13}$	14	27	191	$6.555E{-}10$
	$1.26700\mathrm{E}{+00}$	4.13946E+00	$6.153E{-}07$	0.000E+00	1	3	10	4.732E - 06
60	9.13900E+00	3.15349E+00	1.386E - 06	$1.743E{-}13$	14	27	150	$4.419E{-}10$
	1.24300E+00	3.15375E+00	$2.046E{-}07$	0.000E+00	1	3	10	$3.856E{-}06$
70	1.00490E + 01	$2.50402\mathrm{E}{+00}$	$1.951E{-06}$	$1.142E{-}13$	14	27	163	$5.704E{-}10$
	1.34000E+00	$2.50427\mathrm{E}{+00}$	$7.003E{-}10$	0.000E+00	1	3	11	$3.264 E{-06}$
80	1.15920E + 01	2.04987E+00	1.757E - 06	$1.174E{-13}$	14	27	182	$4.687\mathrm{E}{-10}$
	1.26600E+00	$2.05011\mathrm{E}{+00}$	$4.605 E{-07}$	0.000E+00	1	3	10	2.771E-06
90	1.28680E + 01	1.71776E+00	$6.529E{-}06$	$1.149E{-}13$	15	29	225	$3.554E{-}10$
	1.31500E+00	1.71799E+00	$3.615 E{-07}$	0.000E+00	1	3	10	$2.369E{-}06$
100	1.39110E+01	1.46628E+00	1.557E - 06	$1.145E{-}13$	15	29	215	$2.769E{-}10$
	1.22200E+00	1.46650E+00	$2.557\mathrm{E}{-07}$	0.000E+00	1	3	9	2.037E - 06

6 结论

带有秩约束和简单上下界约束的相关系数矩阵矫正问题可以转化为 DC 优化问题. 我们通过松弛的序列凸近似 (SCA) 方法, 把转化得来的 DC 优化问题进一步转化为一系列可以有效求解的凸子问题, 这些凸子问题的解的极限就是松弛的 DC 问题的稳定点, 而在松弛参数趋于 0 时, 进一步可以得到原问题的稳定点. 数值实验表明, 该方法可以有效求解带有秩约束和简单上下界约束的相关系数矩阵矫正问题.

数值实验的结果显示, 算法 2 的秩约束的误差很小 ($< 10^{-12}$), 一般可以忽略. 没有不等式的例子中 (例 1–4 和 8), 算法 2 明显速度逊于 MPA, 计算的精度略逊一筹, 但在目标函数值方面略有优势; 而在含有不等式的例子中 (例 5–7), 算法 2 与 MPA 的速度相差不大 (多数在同一数量级). 例 7 中 rank = 40,70 的计算结果和例 5 中 n=500,700,1000 的计算结果显示, 这些情形下算法 2 在计算时间上和目标函数值上略有优势, 而在计算的误差方面稍逊一筹. 这表明在含有不等式的情形下两种算法不相上下.

致谢 作者感谢审稿人给出的宝贵意见.

参考文献 -

¹ Davies P I, Higham N J. Numerically stable generation of correlation matrices and their factors. BIT, 2000, 40:640-651

- 2 Higham N J. Computing the nearest correlation matrix—A problem from finance. IMA J Numer Anal, 2002, 22: 329–343
- 3 Malick J. A dual approach to semidefinite least-squares problems. SIAM J Matrix Anal Appl, 2005, 26: 272-284
- 4 Boyd S, Xiao L. Least-squares covariance matrix adjustment. SIAM J Matrix Anal Appl, 2005, 27: 532–546
- 5 Qi H, Sun D. A quadratically convergent Newton method for computing the nearest correlation matrix. SIAM J Matrix Anal Appl, 2006, 28: 360–385
- 6 Gao Y, Sun D. Calibrating least squares semidefinite programming with equality and inequality constraints. SIAM J Matrix Anal Appl, 2009, 31: 1432–1457
- 7 Wu L X. Fast at-the-money calibration of the LIBOR market model using Lagrange multipliers. J Comput Finance, 2003, 6: 39–77
- 8 Zhang Z, Wu L. Optimal low-rank approximation to a correlation matrix. Linear Algebra Appl, 2003, 364: 161–187
- 9 Brigo D, Mercurio F. Interest Rate Models: Theory and Practice. Berlin: Springer-Verlag, 2001
- 10 Lillo F, Mantegna R N. Spectral density of the correlation matrix of factor models: A random matrix theory approach. Phys Rev E, 2005, 72: 016219
- 11 Burg J P, Luenberger D G, Wenger D L. Estimation of structured covariance matrices. Proc IEEE, 1982, 70: 963–974
- 12 Cybenko G. Moment problems and low rank toeplitz approximations. Circuits Systems Signal Process, 1982, 1: 345–366
- 13 Hoge W S. A subspace identification extension to the phase correlation method. IEEE Trans Med Imaging, 2003, 22: 277–280
- 14 Simon D, Abell J. A majorization algorithm for constrained correlation matrix approximation. Linear Algebra Appl, 2010, 432: 1152–1164
- 15 Gao Y, Sun D. A majorized penalty approach for calibrating rank constrained correlation matrix problems. Technical Report. Singapore: National University of Singapore, 2010
- 16 Pietersz R, Groenen P J F. Rank reduction of correlation matrices by majorization. Quant Finance, 2004, 4: 649–662
- 17 Li Q, Qi H. A sequential semismooth newton method for the nearest low-rank correlation matrix problem. SIAM J Optim, 2011, 21: 1641–1666
- 18 Rockafellar R T, Wets R J-B. Variational Analysis. New York: Springer-Verlag, 1998
- 19 Overton M L, Womersley R S. Optimality conditions and duality theory for minimizing sums of the largest eigenvalues of symmetric matrices. Math Program, 1993, 62: 321–357
- 20 Hoffman A. On approximate solutions of systems of linear inequalities. J Res Natl Bur Stand, 1952, 49: 263–265
- 21 Hong L J, Yang Y, Zhang L. Sequential convex approximations to joint chance constrained programs: A monte carlo approach. Oper Res, 2011, 59: 617–630
- 22 Xiao X, Gu J, Zhang L, et al. A sequential convex program method to DC program with joint chance constraints. J Ind Manag Optim, 2012, 8: 733–747
- 23 Boyd S, Vandenberghe L. Convex Optimization. Cambridge: Cambridge University Press, 2004
- 24 Wang B, Zhang L. Calibrating low-rank correlation matrix problem: An SCA-based approach. Optim Methods Softw, 2014, 29: 561–582
- 25 Brace A, Gatarek D, Musiela M. The market model of interest rate dynamics. Math Finance, 1997, 7: 127–155

A relaxed SCA approach for calibrating low rank correlation matrix problem

WANG Bo, ZHANG LiWei & CHU Li

Abstract In this paper, we consider a class of rank constrained correlation matrix calibrating problems with simple upper and lower bounds. This problem can be reformulated into a DC (difference of convex) constrained problem, thus sequential convex approximation (SCA) type approaches can be considered. However, classical SCA approach cannot be directly applied since the constraint qualification needed in convergence theorem does not hold for this problem. This motivates us to develop a relaxed SCA approach. We prove that for relaxed DC problems, the sets of their stationary points converge to the set of stationary points of the original problem as the relaxation parameter approaching to zero. For each relaxed DC problem, we apply the SCA approach to generate

a sequence of convex subproblems. We show that all cluster points of optimal solutions of these subproblems are stationary points of the relaxed DC problem. Numerical results verify the efficiency of our approach.

Keywords DC programing, correlation matrix, low rank

MSC(2010) 49M05, 90C22, 90C26

doi: 10.1360/N012013-00136