

交替投影法求解对称随机逆特征值问题

党婵娟, 王湘美*

贵州大学数学与统计学院, 贵州 贵阳
Email: *xmwang2@gzu.edu.cn

收稿日期: 2021年5月21日; 录用日期: 2021年6月9日; 发布日期: 2021年6月25日

摘要

本文主要研究对称随机矩阵的逆特征值问题。通过将该问题转化为求两个集合交点的可行性问题, 提出用交替投影法进行求解。因为其中一个集合不是凸集, 关于凸可行性问题的收敛性结果不能用来分析算法的收敛性。对于算法的收敛性, 本文在已有关于两个黎曼流形的交替投影算法收敛性的研究结果上, 建立了交替投影算法在一定条件下的线性收敛性。最后数值例子也表明了算法的有效性。

关键词

对称随机逆特征值问题, 可行性问题, 交替投影算法, 黎曼流形, 线性收敛

Alternating Method for Symmetric Stochastic Inverse Eigenvalue Problems

Chanjuan Dang, Xiangmei Wang*

School of Mathematics and Statistics, Guizhou University, Guiyang Guizhou
Email: *xmwang2@gzu.edu.cn

Received: May 21st, 2021; accepted: Jun. 9th, 2021; published: Jun. 25th, 2021

Abstract

The symmetric stochastic inverse eigenvalue problem (StIEP) is considered. By reformulating it into a feasibility problem of two sets, the alternating method is proposed for solving it. Since one of the sets is not convex, the convergence analysis technique for the convex cases does not work. For the convergence of the algorithm, this paper establishes the linear convergence of the alternating projection algorithm under certain conditions on the existing results on the convergence of

*通讯作者。

the alternating projection algorithms for two Riemannian manifolds. At last, some numerical experiments are provided to illustrate the efficiency of the method.

Keywords

Symmetric Stochastic Inverse Eigenvalue Problem, Feasibility Problem, Alternating Method, Riemannian Manifold, Linear Convergence

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

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



Open Access

1. 引言

求一个 n 阶非负矩阵(所有元素都是非负数), 使得它的特征值是给定的 n 个常数 $\lambda_i (i=1, \dots, n)$ 的问题称为非负逆特征值问题(简称 *NIEP*)。 *NIEP* 在很多领域有广泛的应用, 例如在博弈论, 数值分析, 数据分析, 群论和经济学等领域中, 详见([1]-[20])及其参考文献。 1937 年 *Kolmogorov* [1] 提出逆特征值问题: 一个复数 z 何时为某个非负矩阵的特征值的问题? 后来, *Suleimanova* ([2]) 在 1949 年提出了更一般的非负逆特征值问题, 并证明了当 $\{\lambda_i\}$ 中只有一个数大于 0 (其他数小于等于 0) 时, *NIEP* 有解的充分必要条件是这些数的和大于 0。 此后, *NIEP* 有解的充分和必要条件一直是学者们研究的热点问题之一。 一些学者研究了 *NIEP* 有解的必要条件, 取得了一些重要的研究成果。 例如, *Loewy* 和 *Johnson* ([3])、 *Johnson* ([4]) 建立了 *NIEP* 有解的四个必要条件。 1998 年, *Laffey* 和 *Meehan* 给出了奇数阶非负矩阵迹为零的必要条件。 此外, 一些学者对低阶的情形进行研究, 1978 年 *Loewy* 和 *Londow* 给出了 $n=3$ 时, *NIEP* 有解的四个充分必要条件。 此外, 1996 年 *Reams* 解决了 $n=4$ 时非负矩阵迹为零的问题, 1999 年 *Laffey* 和 *Meehan* 解决了 $n=5$ 时同样的问题(详见[5])。 其他研究成果参考([6]-[20])及其参考文献。

本文主要关注求解 *NIEP* 的有效算法。 据我们所知, 目前关于 *NIEP* 的算法并不多, 主要有梯度流下降算法, 交替投影算法, 非光滑牛顿法以及共轭梯度法。 在[21]中, *Chu* 和 *Guo* 将 *NIEP* 转化为以下优化问题:

$$\min_{P, R \in \mathbb{R}^{n \times n}} F(P, R) = \frac{1}{2} \|P \Lambda P^{-1} - R \circ R\|^2, \quad (1.1)$$

其中 Λ 是以给定的 n 个常数为主对角元的 n 阶对角矩阵, “ \circ ” 表示 *Hadamard* 积(详见[21])。 他们提出了求解问题(1.1)的基于微分方程的梯度流下降算法, 并通过一些应用说明了算法的有效性。 文献[22]中, *Orsi* 将 *NIEP* 转化为含两个集合的可行性问题:

$$\text{求 } X \in E \cap N \quad (1.2)$$

其中 E 是以给定的 n 个常数为特征值的 n 阶实对称矩阵集合, N 是 n 阶非负矩阵的集合。 在[22]中, *Orsi* 给出了集合 E 上的投影算子的计算式, 并提出了求解 *NIEP* (即问题(1.2))的交替投影算法(交替地在集合 E 和 N 上进行投影)。 因为集合 E 不是凸集, 所以问题(1.2)不是凸可行问题。 关于非凸的可行性问题的交替投影算法的收敛性分析非常困难, 收敛性结果也非常少。 在[22]中, 作者证明了当 $E \cap \text{int } N \neq \emptyset$, 且迭代点充分接近 $x \in E \cap \text{int } N$ 时, 算法再经过一次迭代得到的 *NIEP* 解, 其中 $\text{int } N$ 是集合 N 的内点集。 类似地假设条件也被用在[23]算法的分析中。 此外, 参考文献[24]和[25]研究了已知部分特征对的 *NIEP* 问题的算法, 即已知矩阵的部分特征值及其特征向量, 求一个非负矩阵。 这两篇

文献分别通过这类问题转化为欧氏空间和黎曼流形上的优化问题进行求解。[24]提出了求解该问题的非光滑牛顿法, 并证明了算法的二阶收敛性; [25]中提出求解该问题的黎曼流形上的共轭梯度法, 并证明了算法的收敛性。

本文主要研究一类特殊的 *NIEP* 问题的数值算法——对称随机逆特征值问题(简称 *StIEP*), 即求解一个对称随机矩阵(行/列和为 1 的非负实对称矩阵), 使得它的特征值是给定的 n 个常数。受[22]启发, 我们将问题转化为含两个集合的可行性问题:

$$\text{求 } X \in E \cap F \quad (1.3)$$

其中集合 E 的定义同问题(1.2), 集合 F 是双随机矩阵集合(即行和、列和均为 1 的非负矩阵), 并提出交替投影算法进行求解。在算法运行中, 我们采用[22]中计算式计算集合 E 上投影 $P_E(\cdot)$, 而集合 F 上的投影 $P_F(\cdot)$, 我们采用同时投影的子迭代计算, 详见算法 3.2。正如我们在上面提到集合 E 不是凸集, 所以问题(1.3)也是非凸可行性问题, 凸可行问题的交替投影算法的收敛性分析方法不能用来分析算法的收敛性。注意到集合 F 没有内点, 不满足[22]中的假设条件: $E \cap \text{int } N \neq \emptyset$ 。所以这一假设条件不能用来分析求解问题(1.3) (即 *StIEP*) 的交替投影算法的收敛性。注意到 Lewis 在([26], Theorem 13)中研究了当两个集合在解的附近是光滑黎曼流形时(不一定是凸集), 交替投影算法的收敛性, 并建立了算法线性收敛性条件。结合集合 E 和 F 的特性以及对算法产生的迭代点列的分析(注意集合 F 并不是黎曼流形), 用[26]中的结论建立了交替投影算法线性收敛的条件。

本文结构如下。第一节是预备知识, 包括一些记号, 概念和一些重要引理。论文的主要结果在第二节, 我们提出求解 *StIEP* 的交替投影算法, 给出了一些集合上投影的计算式, 并建立了算法线性收敛的条件。第三节的数值实验结果表明交替投影算法是有效的算法。最后是对本文的总结以及展望。

2. 预备知识

这一节介绍本文用到的一些记号, 概念和重要引理。设 R^n , $R^{n \times n}$ 分别表示实数域上的 n 维向量, n 阶方阵构成的欧氏空间。设 $E \subset R^{n \times n}$ 是 $R^{n \times n}$ 中子集, 点 $X \in R^{n \times n}$ 到集合 E 的距离和投影分别记为 $d_E(\cdot)$ 和 $P_E(\cdot)$, 即 $d_E(X) = \inf_{Y \in E} \|X - Y\|$,

$$P_E(X) = \text{Arg min}_{Y \in E} \|X - Y\| = \{Y \in E \mid d_E(X) = \|X - Y\|\},$$

其中 $\|\cdot\|$ 表示 $R^{n \times n}$ 中的 F -范数。

下面介绍黎曼流形上一些概念和结论, 关于黎曼流形更多的结果可以参考([27] [28] [29] [30])。设 $f: R^{n \times n} \rightarrow R^d$ 为定义在 $R^{n \times n}$ 上的光滑函数, 其中 d 是正整数。设 $X \in R^{n \times n}$, 函数 f 在点 X 处的 *Jacobi* 矩阵记为 $Jf(X)$ 。

定义 2.1 设 $M \subset R^{n \times n}$ 是 $R^{n \times n}$ 中子集, $X \in M$ 。我们称 M 在点 X 附近是光滑黎曼流形, 如果存在光滑函数 $f: R^{n \times n} \rightarrow R^d$ 和 X 的邻域 U , 使得 Jf 在 U 上是满射且以下式子成立:

$$M \cap U := \{Y \in U : f(Y) = 0\}.$$

设 M 在点 X 附近是光滑黎曼流形, M 在点 $X \in M$ 处的切空间记为 $T_X M$ 。则有

$$T_X M = \ker Jf(X) := \{Z \in R^{n \times n}, \langle Z, Jf(X) \rangle = 0\},$$

其中 $\langle \cdot, \cdot \rangle$ 是 $R^{n \times n}$ 中内积, 即

$$\langle A, B \rangle = \text{trace}(A^T B), \quad \forall A, B \in R^{n \times n}.$$

当 M 和 N 都是 $R^{n \times n}$ 中光滑子流形时, Lewis 在[26]中证明了以下求这两个集合交点的交替投影算法的线性收敛性, 它在我们后面算法收敛性分析中非常重要。需要指出的是, 以下引理将([26], Theorem 13)中的假设条件:

$$T_M(\bar{X}) + T_N(\bar{X}) = R^{n \times n}. \quad (2.1)$$

减弱为

$$T_{M \cap N}(\bar{X}) = T_M(\bar{X}) \cap T_N(\bar{X}). \quad (2.2)$$

(注意(2.1) \Rightarrow (2.2), 见[30]), 这是仔细检查([26], Theorem 13)的证明得到的(也见[30])。同时, 由([26], Theorem 13)的证明也有下面(2.3)成立。

引理 2.1 ([26], Theorem 13) 设 $M, N \in R^{n \times n}$, 点 $\bar{X} \in M \cap N$ 。设 M 和 N 在点 \bar{X} 附近都是光滑黎曼流形, 且满足(2.2)。设 $\{X_k\}$ 是由以下交替投影法产生的点列, 即

$$X_{k+1} := P_N P_M(X_k), \quad \forall k = 0, 1, \dots$$

则存在 $\delta > \delta' > 0$, 使得当初始点 $X_0 \in U(\bar{X}, \delta')$ 时, 点列 $\{X_k\}$ 满足

$$\{X_k\} \subset U(\bar{X}, \delta), \quad (2.3)$$

且线性收敛到 $M \cap N$ 中一点 X^* , 即存在 $\alpha \in R$, $c \in (0, 1)$, 使得 $\|X_k - X^*\| \leq \alpha c^k$, $k = 1, 2, \dots$ 。

3. 交替投影算法及收敛性分

3.1. StIEP 及投影的计算

设 $\lambda := \{\lambda_1, \lambda_2, \dots, \lambda_n\}$ 是给定的 n 个实数。考虑对称随机逆特征值问题(StIEP), 即求一个 n 阶对称随机矩阵 $X \in R^{n \times n}$, 使得它的特征值是 λ 。不失一般性, 设 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$, 并对角阵 Λ 为

$\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, 即

$$\text{diag}(\lambda) = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}.$$

在 $R^{n \times n}$ 中定义以下集合 E , F 和 F_1 如下:

$$E := \left\{ A \in R^{n \times n} \mid A = V \Lambda V^T, V \in R^{n \times n} \text{ 是正交矩阵} \right\},$$

$$F := \left\{ A \in R^{n \times n} \mid A^T = A, \sum_{j=1}^n A_{ij} = 1, A_{ij} \geq 0, i, j = 1, 2, \dots, n \right\}$$

$$F_1 := \left\{ A \in R^{n \times n} \mid \sum_{j=1}^n A_{ij} = 1, A_{ij} \geq 0, i, j = 1, 2, \dots, n \right\},$$

其中 B^T 表示矩阵 $B \in R^{n \times n}$ 的转置, A_{ij} 表示矩阵 A 的第 i 行第 j 列元素。下面我们先讨论集合 E 和 F_1 上投影算子 $P_E(\cdot)$ 和 $P_{F_1}(\cdot)$ 的计算。首先, 我们有以下关于 $P_E(\cdot)$ 的计算式。

引理 3.1 ([22], Theorem 3.2) 设实对称矩阵 $Y \in R^{n \times n}$ 的正交分解为 $Y = V \text{diag}(\mu_1, \dots, \mu_n) V^T$, 其中 $\mu_1 \geq \dots \geq \mu_n$ 。则 $P_E(Y) = V \Lambda V^T$ 。

下面讨论 $P_{f_1}(\cdot)$, 为此设向量 $a = (a_1, \dots, a_n)^T$ 。考虑以下严格凸优化问题:

$$\begin{aligned} \min_{x \in R^n} & \frac{1}{2} \|x - a\|_2^2 \\ \text{s.t.} & \sum_{i=1}^n x_i = 1, \\ & -x_i \leq 0, \quad i = 1, \dots, n, \end{aligned} \quad (3.1)$$

其中 $\|\cdot\|_2$ 表示 R^n 中的 2-范数。设 $x_a^* = ((x_a^*)_i) \in R^n$ 为问题(3.1)的唯一全局最优解, 同时也是它的唯一 KKT 点。所以存在 $\lambda^* = (\lambda_i^*) \in R^n, \mu^* \in R$, 满足以下 KKT 条件:

$$x_a^* - a - \mu^* e - \sum_{i=1}^n \lambda_i^* e_i = 0, \quad (3.2)$$

$$\sum_{i=1}^n (x_a^*)_i = 1, \quad (3.3)$$

$$(x_a^*)_i \geq 0, \lambda_i^* \geq 0, (x_a^*)_i \lambda_i^* = 0, \quad i = 1, \dots, n, \quad (3.4)$$

其中 $e = (1, \dots, 1)$, $e_i = (0, \dots, 0, 1, 0, \dots, 0) \in R^n$ (第 i 个分量为 1, 其他为 0)。为了方便运算, 我们设向量的 a 分量满足 $a_1 \leq a_2 \leq \dots \leq a_n$ 。

引理 3.2 设向量 a 的分量满足 $a_1 \leq a_2 \leq \dots \leq a_n$, 则问题(3.1)的唯一最优解为:

$$x_a^* = \left(0, \dots, 0, b_{l+1} + \frac{1}{n-l} \sum_{j=1}^l b_j, \dots, b_n + \frac{1}{n-l} \sum_{j=1}^l b_j \right), \quad (3.5)$$

其中对任意 $i = 1, \dots, n$, $b_i = a_i + \frac{1}{n} - \frac{1}{n} \sum_{i=1}^n a_i$, $c_i = b_i + \frac{1}{n-i+1} \sum_{j=1}^{i-1} b_j$, $l \in \{0, 1, \dots, n-1\}$ 使得 $c_i \leq 0$ ($\forall i \leq l$) 且 $c_{l+1} > 0$ 。

证明 由 $\{b_i\}$ 和 $\{c_i\}$ 的定义, 容易证明

$$b_1 \leq \dots \leq b_n, \quad (3.6)$$

且存在 $l \in \{0, 1, \dots, n-1\}$, 使得 $c_i \leq 0$ ($\forall i \leq l$), 且 $c_{l+1} > 0$, 令

$$(x_a^*)_i = \begin{cases} 0, & i \leq l, \\ b_i + \frac{1}{n-l} \sum_{j=1}^l b_j, & i \geq l+1. \end{cases} \quad (3.7)$$

$$\lambda_i^* = \begin{cases} -b_i - \frac{1}{n-l} \sum_{j=1}^l b_j, & i \leq l, \\ 0, & i \geq l+1. \end{cases} \quad (3.8)$$

下面验证 x_a^*, λ^* 满足 KKT 条件(3.2)~(3.4)。首先, 我们有

$$\lambda_i^* \cdot (x_a^*)_i = 0, \quad \forall i = 1, \dots, n.$$

另外, 由(3.6), 显然有 $\lambda_1^* \geq \lambda_2^* \geq \dots \geq \lambda_l^*$ 。要证 $\lambda_i^* \geq 0$ ($i = 1, 2, \dots, n$), 只需证明 $\lambda_l \geq 0$ 。根据 l 的定义, 有 $c_i \leq 0$ 。于是由(3.8)有:

$$\begin{aligned} \lambda_l^* &= -b_l - \frac{1}{n-l} \sum_{j=1}^l b_j \\ &= -b_l - \frac{1}{n-l+1} \sum_{j=1}^{l-1} b_j + \frac{1}{n-l+1} \sum_{j=1}^{l-1} b_j - \frac{1}{n-l} \sum_{j=1}^l b_j \\ &= -c_l - \frac{1}{n-l} c_l \geq 0. \end{aligned}$$

下面证明 $(x_a^*)_i \geq 0 (i=1, 2, \dots, n)$ 。类似地, 由(3.6)和(3.7), 显然有 $(x_a^*)_{l+1} \leq (x_a^*)_{l+2} \leq \dots \leq (x_a^*)_n$, 所以只需证明 $(x_a^*)_{l+1} \geq 0$ 。根据 l 的定义, 有 $c_{l+1} > 0$ 。又 $\{b_j\}$ 单调不减, 我们有

$$(x_a^*)_{l+1} = b_{l+1} + \frac{1}{n-l} \sum_{j=1}^l b_j = c_{l+1} > 0.$$

所以 x_a^*, λ_a^* 满足 KKT 条件(3.4)。又

$$\sum_{i=1}^n (x_a^*)_i = \sum_{i=1}^l (x_a^*)_i + \sum_{i=l+1}^n (x_a^*)_i = \sum_{i=l+1}^n b_i = 1.$$

所以(3.3)成立。令

$$\mu^* = \begin{cases} -a_i - \sum_{i=1}^n \left(-b_i - \frac{1}{n-l} \sum_{j=1}^l b_j \right), & i \leq l, \\ b_i + \frac{1}{n-l} \sum_{j=1}^l b_j - a_i, & i \geq l+1. \end{cases}$$

则有

$$(x_a^*)_i - a_i - \sum_{i=1}^n \lambda_i^* - \mu^* = \begin{cases} b_i + \frac{1}{n-l} \sum_{j=1}^l b_j - a_i - \left(b_i + \frac{1}{n-l} \sum_{j=1}^l b_j - a_i \right) = 0, & i \geq l+1, \\ 0 - a_i + a_i + \sum_{i=1}^n \left(-b_i - \frac{1}{n-l} \sum_{j=1}^l b_j \right) - \sum_{i=1}^n \left(-b_i - \frac{1}{n-l} \sum_{j=1}^l b_j \right) = 0, & i \leq l+1. \end{cases}$$

即

$$x_a^* - a - \mu^* e - \sum_{i=1}^n \lambda_i^* e_i = 0.$$

所以 x_a^*, λ_a^* 满足 KKT 条件(3.2)。综上, x_a^* 是问题(3.1)的 KKT 点, 从而是问题(3.1)最优解。

注 3.1 如果向量 a 的分量不满足 $a_1 \leq a_2 \leq \dots \leq a_n$, 可以先将 a 的分量按从小到大顺序排列, 记为向量 \bar{a} 。设 $x_{\bar{a}}^*$ 是由(3.5)定义的向量, x_a^* 是 $x_{\bar{a}}^*$ 按照分量所在位置还原后得到的向量。则 x_a^* 是问题(3.1)的唯一最优解。

结合引理 3.2 和注 3.1, 将矩阵 $X \in R^{n \times n}$ 按行分块, 即 $X = \begin{pmatrix} X_1^T \\ X_2^T \\ \vdots \\ X_n^T \end{pmatrix}$, 则 X 在集合 F_1 上的投影 $P_{F_1}(\cdot)$ 有以

下表达式, 其中 $x_{X_i^T}^*$ 是按注 3.1 中计算, 即先将 X_i^T 的分量按从小到大排序得到 $\overline{X_i^T}$, 再令 $a := \overline{X_i^T}$ 按(3.5)计算后, 排回原来的位置得到。

引理 3.3 设 $X \in R^{n \times n}$, 且 X 按行分块为 $X = \begin{pmatrix} X_1^T \\ X_2^T \\ \vdots \\ X_n^T \end{pmatrix}$, 则 X 在集合 F_1 上的投影为 $P_{F_1}(X) = \begin{pmatrix} (x_{X_1^T}^*)^T \\ (x_{X_2^T}^*)^T \\ \vdots \\ (x_{X_n^T}^*)^T \end{pmatrix}$.

3.2. 交替投影算法及收敛性分析

显然 *StIEP* 等价于可行性问题(1.3), 即求 $X \in E \cap F$. 我们提出以下求解 *StIEP* 的交替投影算法。

交替投影算法 3.1

- Step 0 给定 $\varepsilon > 0$, 任取实对称矩阵 $X_0 \in R^{n \times n}$, 令 $k := 0$.
- Step 1 令 $X_{k+1} := P_E P_F(X_k)$.
- Step 2 若 $\|X_k - X_{k+1}\| < \varepsilon_1$, 算法终止; 否则, 令 $k := k + 1$ 转 Step 1.

交替投影算法 1 要计算两个集合 E 和 F 上的投影, 其中 $P_E(\cdot)$ 按引理 3.1 计算, 而 $P_F(\cdot)$ 用同时投影算法计算, 具体见以下交替投影算法 3.2 和注 3.2.

交替投影算法 3.2

- Step 0 给定 $\varepsilon_1, \varepsilon_2 > 0$, 任取实对称矩阵 $X_0 \in R^{n \times n}$, 令 $k := 0, l := 0$.
- Step 1 $Y_{k,l} := X_k$.
- Step 2 令 $Z_l := P_{F_1}(Y_{k,l}), Y_{k,l+1} := \frac{Z_l + (Z_l)^T}{2}$.
- Step 3 若 $\|Y_{k,l} - Y_{k,l+1}\| < \varepsilon_2$, 令 $W_k := Y_{k,l+1}, l := 0$, 转下一步; 否则, 令 $l := l + 1$, 转 Step 2.
- Step 4 计算矩阵 W_k 的正交分解 $W_k = V \text{diag}(\mu_1, \dots, \mu_n) V^T$, 满足 $\mu_1 \geq \dots \geq \mu_n$. 令 $X_{k+1} := V \text{diag}(\lambda_1, \dots, \lambda_n) V^T$.
- Step 5 若 $\|X_k - X_{k+1}\| < \varepsilon_1$, 算法终止; 否则, 令 $k := k + 1$, 转 Step 1.

注 3.2 交替投影算法 3.2 给出了投影的计算式。由引理 3.1, Step 4 计算了 $P_E(W_k)$, 即 $P_E(W_k) := X_{k+1}$. 而算法的 Steps 2-3 用同时投影算法计算 $P_F(X_k)$ 。事实上, 记

$$F_2 := \left\{ A \in R^{n \times n} \mid \sum_{i=1}^n A_{ij} = 1, A_{ij} \geq 0, i, j = 1, 2, \dots, n \right\}.$$

则 $P_F(X) = P_{F_1 \cap F_2}(X)$, 且容易验证 $P_{F_2}(X) = (P_{F_1}(X))^T$. 所以交替投影算法 3.2 的 Steps 3-4 实际上是子迭代:

$$Y_{k,l+1} := \frac{P_{F_1}(Y_{k,l}) + P_{F_2}(Y_{k,l})}{2} \quad \forall l \in 0, 1, \dots, \quad (3.9)$$

直到满足精度要求 $\|Y_{k,l} - Y_{k,l+1}\| < \varepsilon_2$. 因为集合 F_1 和 F_2 都是多面体, 求解凸可行问题: 求 $X \in F_1 \cap F_2$ 的同

时投影算法(3.9)是线性收敛的(见[30], Corollary 3.14), 所以交替投影算法 2 的 Steps 2-3 迭代的次数不多。

下面我们分析交替投影算法 3.1 的收敛性。因为集合 E 不是凸集, 所以凸可行性问题的研究方法不能用来分析交替投影算法 3.1 的收敛性。我们将用[22]中的求两个黎曼子流形交点的交替投影算法的收敛性研究结果进行研究。因为集合 F 不是 $R^{n \times n}$ 的子流形。为此, 我们定义集合 F° 如下:

$$F^\circ := \left\{ A \in R^{n \times n} \mid A = A^T, \sum_{j=1}^n A_{ij} = 1, A_{ij} > 0, \quad i, j = 1, 2, \dots, n \right\}.$$

由集合 F° 的定义, 对任意 $X \in F^\circ$, 存在点 X 的邻域 $U(X)$ 使得

$$F^\circ \cap U(X) := \left\{ A \in U(X) \mid A_{ij} = A_{ji}, \sum_{j=1}^n A_{ij} = 1, \quad i, j = 1, 2, \dots, n \right\}.$$

由定义 2.1 容易验证以下结论。

引理 3.4 F° 在任意点 $X \in F^\circ$ 附近是光滑黎曼流形。

关于集合 E , 由定义知对任意 $X \in E$, $\lambda := \{\lambda_1, \lambda_2, \dots, \lambda_n\}$ 是它的 n 个特征值, 即 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是关于 λ 的 n 次方程 $|\lambda I - X| = 0$ 的 n 个根, 其中 I 是 n 阶单位阵。由根与系数的关系以及集合 E 的定义有

$$E := \left\{ A \in R^{n \times n} \mid A_{ij} = A_{ji}, f(A) = 0, \quad i, j = 1, 2, \dots, n \right\},$$

其中 $f: R^{n \times n} \rightarrow R^n$ 是定义在 $R^{n \times n}$ 上的多项式函数。设 $\bar{X} \in E \cap F^\circ$, 下面的讨论需要用到以下假设:

E 在点 \bar{X} 附近是光滑黎曼流形, 且满足

$$T_{E \cap F^\circ}(\bar{X}) = T_E(\bar{X}) \cap T_{F^\circ}(\bar{X}). \quad (3.10)$$

下面仅对当 $n=2$ 时, 说明 E 在任意点 $\bar{X} \in E \cap F^\circ$ 附近是光滑黎曼流形, 而当 $n>2$ 时, 类似地可以知道很多情形下相同的结论也成立。事实上容易验证, $n=2$ 时 $StIEP$ 有解(即 $E \cap F \neq \emptyset$)的充分必要条件是

$\lambda = \{1, t\}$, 其中 $t \in [-1, 1]$, 且解为 $X^* := \begin{bmatrix} \frac{1+t}{2} & \frac{1-t}{2} \\ \frac{1-t}{2} & \frac{1+t}{2} \end{bmatrix}$ 。而当 $t \in (-1, 1)$ 时, $E \cap F^\circ \neq \emptyset$ 。设 $t \in (-1, 1)$,

$\bar{X} = \begin{bmatrix} \frac{1+t}{2} & \frac{1-t}{2} \\ \frac{1-t}{2} & \frac{1+t}{2} \end{bmatrix} \in E \cap F^\circ$, 并记

$$E := \left\{ A \in R^{n \times n} \mid A_{12} = A_{21}, A_{11} + A_{22} = 1+t, A_{11}A_{22} - A_{12}A_{21} = t \right\} = \left\{ A \in R^{n \times n} \mid g(A) = 0 \right\}.$$

直接计算知 $Jg(\bar{X})$ 满秩, 所以由定义 2.1, E 在点 \bar{X} 附近是光滑黎曼流形。

考虑以下可行性问题:

$$\text{求 } X \in E \cap F^\circ. \quad (3.11)$$

由引理 2.1 和引理 3.4, 有以下结论(其中 $M := F^\circ$, $N := E$)。

引理 3.5 设 $\bar{x} \in E \cap F^\circ$ 满足假设(3.10)。设 $\{X_k\}$ 是由以下交替投影法产生的点列, 即

$$X_{k+1} := P_E P_{F^\circ}(X_k) \quad \forall k = 0, 1, \dots \quad (3.12)$$

则存在 $\delta > \delta' > 0$, 使得当初始点 $X_0 \in U(\bar{X}, \delta')$ 时, 点列 $\{X_k\}$ 满足 $\{X_k\} \subset U(\bar{X}, \delta)$, 且线性收敛到 $E \cap F^\circ$

中一点 X^* 。

关于交替投影算法 3.1, 我们有以下线性收敛的结论。

定理 3.1 设 $\bar{X} \in E \cap F^\circ$ 满足假设(3.10)。则当初始点 X_0 充分接近 \bar{X} 时, 交替投影算法 3.1 产生的点列 $\{X_k\}$ 线性收敛到 $E \cap F$ 中一点。

证明 设 $\{\tilde{X}_k\}$ 是以 X_0 为初始点(即 $\tilde{X}_0 = X_0$)，用交替投影算法(3.12)产生的点列。由引理 3.5, 存在 $\delta > \delta' > 0$, 使得当初始点 $X_0 \in U(\bar{X}, \delta')$ 时, 点列 $\{\tilde{X}_k\}$ 满足 $\{\tilde{X}_k\} \subset U(\bar{X}, \delta)$, 且线性收敛到 $E \cap F^\circ (\subset E \cap F)$ 中一点。所以, 要证明定理的结论, 只需证明当初始点 X_0 充分接近 \bar{X} 时, 算法 3.1 和算法(3.12)产生相同的点列, 即 $\{X_k\}$ 和 $\{\tilde{X}_k\}$ 相同。设 \bar{X} 到集合 F 的边界 $\partial F := F \setminus F^\circ$ 的距离为 d_0 , 则由 F° 的定义知 $d_0 > 0$ 。不妨设 $\delta < \frac{d_0}{2}$ (否则取更小的 δ 即可)。因为 $\bar{X} \in E \cap F^\circ \subset E \cap F$, 所以对任意 $X \in U(\bar{X}, \delta)$, 我们有 $P_F(X) = P_{F^\circ}(X)$ 。注意到 $\tilde{X}_0 = X_0$ 和 $\{\tilde{X}_k\} \subset U(\bar{X}, \delta)$, 容易得到算法 3.1 和算法(3.12)产生相同的点列。证毕。

4. 数值实验

本节我们用交替投影算法 3.2 求解 *SHEP* 问题。在数值实验中, 初始值 $X_0 \in R^{n \times n}$ 为随机产生的 n 阶矩阵, *iter* 和 *time(s)* 分别表示算法的迭代次数和 CPU 运行时间(单位为秒)。在每次实验时, 我们先随机生成 n 阶对称随机矩阵 A , 并求出矩阵 A 的 n 个特征值作为 λ 。然后用交替投影算法 3.2 进行求解。算法采用 Matlab R2018b 实现, 实验结果如下。

例 给定 5×5 阶矩阵的特征值为

$$\lambda = \{1.0000, 0.0021, 0.0051, 0.0954, 0.1004\}.$$

用算法 3.2 求一个 5×5 阶对称双随机矩阵 X , 使其特征值为 λ 。算法中的 ε_1 和 ε_2 分别取作 10^{-14} 和 10^{-5} , 实验结果为:

$$X = \begin{bmatrix} 0.9435 & 0.0066 & 0.0058 & 0.0236 & 0.0205 \\ 0.0066 & 0.9501 & 0.0165 & 0.0123 & 0.0145 \\ 0.0058 & 0.0165 & 0.9524 & 0.0035 & 0.0218 \\ 0.0236 & 0.0123 & 0.0035 & 0.9510 & 0.0096 \\ 0.0205 & 0.0145 & 0.0218 & 0.0096 & 0.9336 \end{bmatrix}.$$

更高阶的数值实验结果见表 1, 其中 n 是矩阵的阶数, *iter* 和 *time(s)* 是实验 10 次的平均迭代次数和时间。

Table 1. Experiments for different n
表 1. 不同阶数的实验结果

n	$\varepsilon_1, \varepsilon_2$	time(s)	iter
10	$\varepsilon_1 = 10^{-8}, \varepsilon_2 = 10^{-5}$	0.0214	11
	$\varepsilon_1 = 10^{-9}, \varepsilon_2 = 10^{-5}$	0.0215	12
100	$\varepsilon_1 = 10^{-10}, \varepsilon_2 = 10^{-5}$	0.1240	18
	$\varepsilon_1 = 10^{-11}, \varepsilon_2 = 10^{-5}$	0.1320	21
300	$\varepsilon_1 = 10^{-12}, \varepsilon_2 = 10^{-5}$	1.0340	25
	$\varepsilon_1 = 10^{-13}, \varepsilon_2 = 10^{-5}$	1.1630	28
500	$\varepsilon_1 = 10^{-14}, \varepsilon_2 = 10^{-5}$	4.1470	31

5. 结论

对称随机逆特征值问题(*StIEP*)是非负逆特征值问题(*NIEP*)的一种特殊情形, 通过将问题转化为可行性问题, 本文提出的交替投影算法。因为该问题不是凸可行性问题, 其收敛性分析非常困难。我们运用黎曼流形上可行性问题的线性收敛分析结果, 在一定条件下建立算法的线性收敛性。未来我们将继续研究交替投影法求解其他 *NIEP* 问题, 同时进一步研究非凸可行性问题的交替投影算法的收敛性条件。

基金项目

本研究受国家自然科学基金项目(11661019)资助, 在此表示感谢。

参考文献

- [1] Kolmogorov, A.N. (1937) Markov Chains with Countably Many Possible States. *Bull. University. Moscow*, **3**, 1-16.
- [2] Suleimanova, K.R. (1949) Stochastic Matrices with Real Eigenvalues. *Soviet Mathematics—Doklady*, **66**, 343-345.
- [3] Loewy, R. and London, D.A. (1978) Note on an Inverse Problem for Nonnegative Matrices. *Linear and Multilinear Algebra*, **6**, 83-90. <https://doi.org/10.1080/03081087808817226>
- [4] Johnson, C. (1981) Row Stochastic Matrices Similar to Doubly Stochastic Matrices. *Linear and Multilinear Algebra*, **66**, 113-130. <https://doi.org/10.1080/03081088108817402>
- [5] Laffey, T.J. and Meehan, E. (1999) A Characterization of Trace Zero Nonnegative 5×5 Matrices. *Linear Algebra and Its Applications*, **302**, 295-302. [https://doi.org/10.1016/S0024-3795\(99\)00099-3](https://doi.org/10.1016/S0024-3795(99)00099-3)
- [6] Chu, M.T. and Golub, G.H. (2002) Structured Inverse Eigenvalue Problems. *Acta Numerica*, **11**, 1-71. <https://doi.org/10.1017/S0962492902000016>
- [7] Barcilon, V. (1990) A Two-Dimensional Inverse Eigenvalue Problem. *Journal of Inverse and Ill-Posed Problems*, **6**, 11-20. <https://doi.org/10.1088/0266-5611/6/1/004>
- [8] Egleston, P.D., Lenker, T.D. and Narayan, S.K. (2004) The Nonnegative Inverse Eigenvalue Problem. *Linear Algebra and Its Applications*, **379**, 475-490. <https://doi.org/10.1016/j.laa.2003.10.019>
- [9] Soules, G.W. (2008) Constructing Symmetric Nonnegative Matrices. *Linear and Multilinear Algebra*, **13**, 241-251. <https://doi.org/10.1080/03081088308817523>
- [10] Robert, R. (1996) An Inequality for Nonnegative Matrices and the Inverse Eigenvalue Problem. *Linear and Multilinear Algebra*, **41**, 367-375. <https://doi.org/10.1080/03081089608818485>
- [11] Lin, M.M. (2015) An Algorithm for Constructing Nonnegative Matrices with Prescribed Real Eigenvalues. *Applied Mathematics and Computation*, **31**, 582-590. <https://doi.org/10.1016/j.amc.2015.01.033>
- [12] Combettes, P.L. and Trussell, H.J. (1990) Method of Successive Projections for Finding a Common Point of Sets in Metric Spaces. *Journal of Optimization Theory and Applications*, **67**, 487-507. <https://doi.org/10.1007/BF00939646>
- [13] Morel, P. (1976) Des algorithmes pour le problème inverse des valeurs propres. *Linear Algebra and Its Applications*, **13**, 251-273. [https://doi.org/10.1016/0024-3795\(76\)90100-2](https://doi.org/10.1016/0024-3795(76)90100-2)
- [14] Chu, M.T. and Chen, X.Z. (1996) On the Least Squares Solution of Inverse Eigenvalue Problems. *Siam Journal on Numerical Analysis*, **33**, 2417-2430. <https://doi.org/10.1137/S0036142994264742>
- [15] Perfect, H. (1955) Methods of Constructing Certain Stochastic Matrices. *Duke Mathematical Journal*, **22**, 305-311. <https://doi.org/10.1215/S0012-7094-55-02232-8>
- [16] Chu, M.T. (1998) Inverse Eigenvalue Problems. *Siam Review*, **40**, 1-39. <https://doi.org/10.1137/S0036144596303984>
- [17] Xu, M. (2010) On Computing Minimal Realizable Spectral Radii of Non-Negative Matrices. *Numerical Linear Algebra with Applications*, **12**, 77-86. <https://doi.org/10.1002/nla.395>
- [18] Chu, M.T. and Driessel, K.R. (1991) Constructing Symmetric Nonnegative Matrices with Prescribed Eigenvalues by Differential Equations. *SIAM Journal on Numerical Analysis*, **22**, 1372-1387. <https://doi.org/10.1137/0522088>
- [19] Joseph, K.T. (2012) Inverse Eigenvalue Problem in Structural Design. *AIAA Journal*, **30**, 2890-2896. <https://doi.org/10.2514/3.11634>
- [20] Dawson, C.B. and Cha, P.D. (2019) A Sensitivity-Based Approach to Solving the Inverse Eigenvalue Problem for Linear Structures Carrying Lumped Attachments. *Materials Science and Engineering R-Reports*, **120**, 537-566. <https://doi.org/10.1002/nme.6147>

- [21] Chu, M.T. and Guo, Q.Y. (1998) A Numerical Method for the Inverse Stochastic Spectrum Problem. *SIAM Journal on Matrix Analysis and Applications*, **19**, 1027-1039. <https://doi.org/10.1137/S0895479896292418>
- [22] Orsi, R. (2006) Numerical Methods for Solving Inverse Eigenvalue Problems for Nonnegative Matrices. *SIAM Journal on Matrix Analysis and Applications*, **28**, 190-212. <https://doi.org/10.1137/050634529>
- [23] Wu, S.J. and Lin, M.M. (2014) Numerical Methods for Solving Nonnegative Inverse Singular Value Problems with Prescribed Structure. *Inverse Problems*, **30**, 55-69. <https://doi.org/10.1088/0266-5611/30/5/055008>
- [24] Bai, Z.J., Serra-Capizzano, S. and Zhi, Z. (2012) Nonnegative Inverse Eigenvalue Problems with Partial Eigendata. *Bit Numerical Mathematics*, **120**, 387-431. <https://doi.org/10.1007/s00211-011-0415-y>
- [25] Yao, T.T., Bai, Z.J. and Zhao, Z. (2018) A Riemannian Variant of the Fletcher-Reeves Conjugate Gradient Method for Stochastic Inverse Eigenvalue Problems with Partial Eigendata. *Numerical Linear Algebra With Applications*, **26**, 1-19. <https://doi.org/10.1002/nla.2221>
- [26] Lewis, A.S. (2008) Alternating Projections on Manifolds. *Inform Journal on Computing*, **33**, 216-234. <https://doi.org/10.1287/moor.1070.0291>
- [27] 陈维桓, 李兴校. 黎曼几何引论[M]. 北京: 北京大学出版社, 2002.
- [28] Carmo, M.P. (1992) Riemannian Geometry. Birkhauser, Boston. <https://doi.org/10.1007/978-1-4757-2201-7>
- [29] Sakai, T. (1996) Riemannian Geometry. American Mathematical Society, Providence. <https://doi.org/10.1090/mmono/149>
- [30] Bauschke, H.H. and Borwein, J.M. (1993) On the Convergence of von Neumann's Alternating Projection Algorithm for Two Sets. *Set-Valued Analysis*, **1**, 185-212. <https://doi.org/10.1007/BF01027691>