

一类张量特征值互补问题^{*1)}

罗 刚

(南开大学数学科学学院, 天津 300071)

杨庆之

(喀什大学数学与统计学院, 喀什 844006;

南开大学数学科学学院, 天津 300071)

摘 要

矩阵特征值互补问题在力学系统领域有广泛的应用. 在本文中, 我们提出了一类特殊的四阶张量特征值互补问题, 它是矩阵特征值互补问题的推广. 我们对该特征值互补问题解的存在性, 计算复杂度等性质进行了初步的研究. 在一定条件下, 我们建立了该互补问题同一类非线性约束优化问题的等价性联系, 并由此提出了平移投影幂法来求解该特征值互补问题.

关键词: 半正定锥; 特征值互补问题; NP 难; 平移投影幂法

MR (2010) 主题分类: 65K10, 65F15, 15A18, 15A69, 90C22, 90C33

1. 引 言

矩阵特征值互补问题 (eigenvalue complementarity problem: EiCP) 是求解满足如下条件的 $(\lambda, x) \in (\mathbb{R}, \mathbb{R}^n)$ 的问题:

$$\begin{cases} x^T(\lambda B - A)x = 0, \\ (\lambda B - A)x \geq 0, \\ x \geq 0, \end{cases} \quad (1.1)$$

其中 B, A 是 $n \times n$ 实矩阵. 在一些具体应用中, 该问题有额外的约束条件: $\lambda \geq 0$. EiCP 在单侧摩擦力学系统平衡态^[3], 结构力学系统的动力学分析^[15, 16]等问题中有广泛的应用, 此外还有很多关于 EiCP 解的性质及计算方法的研究文章, 如 [10, 18].

在 [19] 中, 作者研究了更一般的特征值互补问题:

$$\begin{cases} \langle x, Ax - \lambda x \rangle = 0, \\ Ax - \lambda x \in K^*, \\ x \in K, \end{cases} \quad (1.2)$$

其中 K 是希尔伯特空间 $(H, \langle \cdot, \cdot \rangle)$ 中的一个闭凸锥, K^* 是它的对偶锥 (定义为 $\{x \in H \mid \langle x, y \rangle \geq 0, \forall y \in K\}$), A 是 H 上的一个连续线性映射: $H \rightarrow H$. 在该文中, 作者特别研究了当 K 是帕累托锥 (Pareto cone), 即 $K = \{x \in \mathbb{R}^n \mid x \geq 0\}$ 时, 特征值互补问题解的性质. 在 [1, 6] 中, 作者研究了 K 为二阶锥 ($\{(x, t) \in \mathbb{R}^{n+1} \mid \|x\| \leq t, t > 0\}$) 时特征值互补问题解的相关性质.

^{*} 2018 年 3 月 28 日收到.

¹⁾ 基金项目: 国家自然科学基金项目 (11671217), 新疆自然科学基金项目 (2017D01A14).

自从高阶张量 (矩阵推广) 特征值概念分别由 Qi^[17] 和 Lim^[13] 在 2005 年提出之后, 张量特征值相关的问题被广泛地研究了. 在 [2, 5, 14] 中, 作者研究了如下形式的张量特征值互补问题:

$$\text{求 } (\lambda, x) \text{ 满足 } \begin{cases} \lambda Bx^m - Ax^m = 0, \\ \lambda Bx^{m-1} - Ax^{m-1} \geq 0, \\ x \geq 0, \end{cases} \quad (1.3)$$

其中 B, A 是 m 阶 n 维张量, $Ax^m = \sum_{i_1, i_2, \dots, i_m=1}^n a_{i_1 i_2 \dots i_m} x_{i_1} x_{i_2} \cdots x_{i_m}$, Ax^{m-1} 是 n 维向量: $(Ax^{m-1})_{i_1} = \sum_{i_2, \dots, i_m=1}^n a_{i_1 i_2 \dots i_m} x_{i_2} \cdots x_{i_m}$, $i_1 = 1, \dots, n$, Bx^m 和 Bx^{m-1} 的定义类似.

在张量 A, B 超对称条件下, [2] 文的作者将问题 (1.3) 转化为一类非线性约束优化问题, 并提出了一个平移投影幂算法 (shifted projected power method) 来求解该优化问题. 在 [5] 中, 作者将该张量特征值互补问题转化为多项式优化问题. 当 A, B 中有一个是严格协正时, 文中利用 Lasserre 层次结构的半定松弛方法来求解这个多项式优化问题. 在 [14] 中, 作者研究了如下一般张量特征值互补问题:

$$\text{求 } (\lambda, x) \text{ 满足 } \begin{cases} \lambda Bx^m - Ax^m = 0, \\ \lambda Bx^{m-1} - Ax^{m-1} \in K^*, \\ x \in K, \end{cases} \quad (1.4)$$

其中 K 是欧式空间 \mathbb{R}^n 中的一个闭凸锥, K^* 是其对偶锥. 当 $K = \{x \in \mathbb{R}^n \mid x \geq 0\}$, 他们证明了当 $Bx^m \neq 0, \forall x \in K \setminus \{0\}$ 成立时, 张量特征值互补问题至少存在一组解. 当张量对 (B, A) 是非奇异时, 他们给出了张量特征值互补问题解个数的一个上界. 在 [8] 中, 作者研究了 K 为二阶锥时张量特征值互补问题解的性质, 并给出了计算算法.

本文中, 我们将矩阵特征值互补问题由帕累托锥推广到半正定锥, 称作半正定特征值互补问题 (SDPEiCP). 这个推广的问题涉及到四阶张量, 因此可以看成一类特殊的张量特征值互补问题, 并且该问题不同于上述文献中研究的张量特征值互补问题. 我们给出了 SDPEiCP 解存在的一个充分条件, 然后在一定假设条件下, 将该互补问题同两个非线性半正定规划问题联系起来. 最后利用它们之间的等价联系, 我们提出了一类迭代幂法来求解半正定特征值互补问题.

本文其余内容安排如下: 在第 2 节, 我们提出半正定特征值互补问题 (SDPEiCP), 研究了它的解存在的一些充分条件. 在第 3 节, 在一定条件下, 我们建立了 SDPEiCP 同一类非线性约束优化问题的等价性联系. 在第 4 节, 我们提出了一个平移投影幂法来求解该类约束优化问题, 同时证明了算法在一定条件下的收敛性. 在第 5 节, 我们给出了一些初步的数值模拟实验结果.

在本文中, 我们采用如下记号: 小写字母 (如 a) 表示标量, 大写字母 (如 A) 表示矩阵, 花体字母 (如 \mathcal{A}) 表示高阶张量. S^n 表示所有 n 阶对称矩阵的集合. 矩阵 $A(\succeq) \succ 0$ 表示 A 是 (半) 正定矩阵, 所有 (半) 正定矩阵的集合记为 $(S_+^n)S_{++}^n$. $\langle \cdot, \cdot \rangle$ 表示矩阵内积: $\langle X, Y \rangle = \sum_{i,j=1}^n x_{ij} y_{ij}$. 矩阵的 Frobenius 范数 $\|X\|_F^2 := \langle X, X \rangle = \sum_{i,j=1}^n x_{ij}^2$. $\text{diag}(a_1, a_2, \dots, a_n)$ 表示对角元素为 a_1, \dots, a_n 的对角矩阵. $\text{trace}(X)$ 表示矩阵 X 的迹: $\text{trace}(X) := \sum_{i=1}^n x_{ii}$. 我们定义矩阵和张量之间特殊的运算: $A \otimes B$ 表示由矩阵 A, B 生成的一个四阶张量, 其中张量的各元素为

$$(A \otimes B)_{ijkl} = A_{ij} B_{kl}, \quad \forall i, j, k, l = 1, \dots, n.$$

$\mathcal{A}X$ 表示四阶张量 \mathcal{A} 与矩阵 X 相乘生成的矩阵, 其元素为

$$(\mathcal{A}X)_{ij} = \sum_{k,l=1}^n a_{ijkl} x_{kl}, \quad \forall i, j = 1, \dots, n. \quad (1.5)$$

我们记 $\mathcal{A}X^2$ 为 $\langle X, \mathcal{A}X \rangle$, 即 $\mathcal{A}X^2 = \sum_{i_1, i_2, i_3, i_4=1}^n a_{i_1 i_2 i_3 i_4} x_{i_1 i_2} x_{i_3 i_4}$. $M(\mathcal{A})$ 表示由四阶张量 \mathcal{A} 展成的矩阵, 其元素为

$$(M(\mathcal{A}))_{pq} = a_{ijkl}, \quad p = (i-1) \times n + j, \quad q = (k-1) \times n + l, \quad i, j = 1, \dots, n.$$

2. 半正定特征值互补问题

我们将矩阵特征值互补问题 (EiCP) 从帕累托锥推广到半正定锥:

$$\text{求 } (\lambda, X) \in (\mathbb{R}, S^{n \times n}) \text{ 满足 } \begin{cases} \langle \lambda B X - \mathcal{A} X, X \rangle = 0, \\ \lambda B X - \mathcal{A} X \succeq 0, \\ X \succeq 0, \end{cases} \quad (2.1)$$

其中 B, \mathcal{A} 是四阶张量. 我们称该问题为半正定特征值互补问题, 记为 SDPEiCP. 从张量矩阵乘法定义 (1.5) 可以看出, 为了保证矩阵 $\lambda B X - \mathcal{A} X$ 对称, 我们要求张量 B, \mathcal{A} 的前两个指标可交换, 又因为 X 是对称的, 所以我们不妨假设 B, \mathcal{A} 后两个指标可交换. 我们假定张量 B, \mathcal{A} 都满足如下部分对称条件: $a_{ijkl} = a_{jikl} = a_{ijlk}, \forall i, j, k, l = 1, \dots, n$. 我们称满足问题 (2.1) 的 X, λ 分别为半正定互补特征向量, 半正定互补特征值. 当 $\lambda > 0$ 时, 我们称其为正互补特征值. $X = 0$ 时, 显然任何 $\lambda \in \mathbb{R}$ 都是问题 (2.1) 的解, 因此本文我们仅考虑非平凡解, 即 $X \neq 0$ 的情况.

2.1. SDPEiCP 解的存在性

对于任意的张量 \mathcal{A}, B , SDPEiCP 不一定存在解, 这里我们给出 SDPEiCP 解存在的一个充分条件. 记 $\Omega = \{X \mid X \succeq 0, \|X\|_F = 1\}$. 定义连续映射 $F: \Omega \rightarrow S^n$

$$F(X) = \left(\frac{\mathcal{A}X^2}{BX^2} B - \mathcal{A} \right) X.$$

这里我们假设

$$BX^2 \neq 0, \quad \forall X \in \Omega. \quad (2.2)$$

定义变分问题 $VI(F, \Omega)$: 求 $\bar{X} \in \Omega$,

$$\langle F(\bar{X}), X - \bar{X} \rangle \geq 0, \quad \forall X \in \Omega.$$

我们可以证明在假设 (2.2) 下 SDPEiCP 同变分问题 $VI(F, \Omega)$ 之间的等价性.

定理 1. 在假设 (2.2) 下, \bar{X} 是 $VI(F, \Omega)$ 的一个解当且仅当 \bar{X} 是 SDPEiCP 的一个解.

证明. ‘ \leftarrow ’: 若 \bar{X} 是 SDPEiCP 的一个解, 即

$$\begin{cases} \bar{\lambda} B \bar{X} - \mathcal{A} \bar{X} \succeq 0, \\ \bar{\lambda} B \bar{X}^2 - \mathcal{A} \bar{X}^2 = 0, \\ \bar{X} \succeq 0, \end{cases} \quad (2.3)$$

其中 $\bar{\lambda}$ 是所对应的互补特征值, 满足 $\bar{\lambda} = \frac{A\bar{X}^2}{B\bar{X}^2}$. 因此

$$\begin{aligned}\langle F(\bar{X}), X - \bar{X} \rangle &= \left\langle \left(\frac{A\bar{X}^2}{B\bar{X}^2} B - A \right) \bar{X}, X - \bar{X} \right\rangle \\ &= -\bar{\lambda} B \bar{X}^2 + A \bar{X}^2 + \langle \bar{\lambda} B \bar{X} - A \bar{X}, X \rangle \\ &= \langle \bar{\lambda} B \bar{X} - A \bar{X}, X \rangle \\ &\geq 0,\end{aligned}\quad (2.4)$$

对所有 $X \in \Omega$ 成立, 所以 \bar{X} 是 $VI(F, \Omega)$ 的一个解.

‘ \rightarrow ’: 如果 \bar{X} 是 $VI(F, \Omega)$ 的一个解. 记 $\bar{\lambda} = \frac{A\bar{X}^2}{B\bar{X}^2}$, 即 $-\bar{\lambda} B \bar{X}^2 + A \bar{X}^2 = 0$. 由 (2.4) 可知, $\langle F(\bar{X}), X - \bar{X} \rangle \geq 0, \forall X \in \Omega$ 成立等价于 $\langle \bar{\lambda} B \bar{X} - A \bar{X}, X \rangle \geq 0, \forall X \in \Omega$ 成立, 即 $\bar{\lambda} B \bar{X} - A \bar{X} \succeq 0$. 所以 \bar{X} 是 SDPEiCP 的一个解.

因此 SDPEiCP 解的存在性问题等价于变分问题 $VI(F, \Omega)$ 解的存在性问题. 由 [4, 12] 知, 当 Ω 是紧凸集, F 是连续映射时, 变分不等式 $VI(F, \Omega)$ 存在解, 所以我们得到了 SDPEiCP 解存在的一个充分条件.

定理 2. SDPEiCP 存在解的一个充分条件: $BX^2 \neq 0$, 对所有 $X \succeq 0 (X \neq 0)$ 成立.

在 [10, 18] 中, 作者研究了在斜对称和对称条件下, 矩阵特征值互补问题解的存在性. 在 [19] 中, 作者研究了更一般的情况: Hilbert 空间一般锥互补问题解的存在性等问题. 在 [3, 10] 中, 作者重点研究了正互补特征值相关的性质, 这里我们给出 SDPEiCP 正互补特征值存在的一个充分条件.

性质 1. SDPEiCP 正互补特征值存在的一个充分条件是: $BX^2 > 0, AX^2 > 0, \forall X \succeq 0 (X \neq 0)$.

证明. 通过定理 1 易得结论.

后面一节, 我们给出了在张量 B, A 更强的对称性假设条件下, SDPEiCP 正互补特征值存在的充分必要条件.

3. 转化为约束优化问题

当 B, A 满足对称性: $a_{ijkl} = a_{klij} = a_{jikl}, \forall i, j, k, l = 1, \dots, n$ 时, 我们可以将求解 SDPEiCP 与特定的非线性约束优化问题联系起来. 这个对称性要求比第二节所列出的基本对称要求更强: 张量的前两个和后两个指标可以整体交换. 这样的对称性要求是基于矩阵函数求导法则的需求, 我们可以验证此时 AX^2 对 X 的导数为 $2AX$. 在本节及后面涉及的算法中, 我们均假设这样的对称性条件成立. 在本节我们仍假定 $BX^2 \neq 0, \forall X \succeq 0 (X \neq 0)$. 不妨假设 $BX^2 > 0, \forall X \succeq 0 (X \neq 0)$ 成立. 记

$$\lambda(X) = \frac{AX^2}{BX^2}.$$

$\lambda(X)$ 的梯度 $\nabla \lambda(X) = \frac{2}{BX^2} (AX - \lambda(X)BX)$. 考察非线性约束优化问题

$$\begin{aligned}(\text{P}) \quad & \max \quad \lambda(X) \\ & \text{s.t.} \quad \|X\|_F = 1, \\ & \quad \quad X \succeq 0.\end{aligned}\quad (3.1)$$

问题 (P) 的拉格朗日函数为

$$L(X, \mu, S) = \lambda(X) + \mu(\|X\|_F^2 - 1) + \langle S, X \rangle. \quad (3.2)$$

其 Karush-Kuhn-Tucker(KKT) 条件为

$$\begin{cases} \nabla_X L(X, \mu, S) = \nabla \lambda(X) + 2\mu X + S = 0, \\ \|X\|_F = 1, \\ S \succeq 0, \\ X \succeq 0, \\ \langle S, X \rangle = 0. \end{cases} \quad (3.3)$$

当 (X, μ, S) 为问题 (P) 的 KKT 点时, 我们有 $\langle \nabla \lambda(X) + 2\mu X, X \rangle = 0$. 因为 $\langle \nabla \lambda(X), X \rangle = 0$, 所以 $\mu = 0$. 于是由 KKT 条件的第一个等式推出 $\frac{2}{\|X\|_F^2}(\mathcal{A}X - \lambda(X)\mathcal{B}X) \preceq 0$. 此时若 $\mathcal{B}X^2 > 0$, $\forall X \succeq 0 (X \neq 0)$ 成立, 则 $\lambda(X)\mathcal{B}X - \mathcal{A}X \succeq 0$, 于是 X 是 SDPEiCP 的一个解. 我们有

定理 3. 若 $\mathcal{B}X^2 > 0$, $\forall X \succeq 0 (X \neq 0)$, 则 (P) 的每一个 KKT 点都是 SDPEiCP 的一个解.

我们可以将 (P) 中的约束 $\|X\|_F = 1$ 替换为 $\text{trace}(X) = 1$, 从而得到另一类非线性约束优化问题:

$$\begin{aligned} \max \quad & \lambda(X) \\ \text{(Q)} \quad \text{s.t.} \quad & \text{trace}(X) = 1, \\ & X \succeq 0. \end{aligned} \quad (3.4)$$

可以看出问题 (Q) 的可行域是凸的, 而 (P) 的可行域是非凸的. 类似地, 我们有如下定理

定理 4. 若 $\mathcal{B}X^2 > 0$, $\forall X \succeq 0 (X \neq 0)$, 则 (Q) 的每一个 KKT 点都是 SDPEiCP 的一个解.

证明. (Q) 的拉格朗日函数为

$$L(X, S, \mu) = \lambda(X) + \langle S, X \rangle + \mu(\text{trace}(X) - 1).$$

其 KKT 条件为

$$\begin{cases} \nabla_X L(X, \mu, S) = \nabla \lambda(X) + \mu I + S = 0, \\ \text{trace}(X) = 1, \\ S \succeq 0, \\ X \succeq 0, \\ \langle S, X \rangle = 0. \end{cases} \quad (3.5)$$

$\langle \nabla \lambda(X) + \mu I + S, X \rangle = 0$, $\langle \nabla \lambda(X), X \rangle = 0$ 和 $\langle S, X \rangle = 0$ 可以推出 $\mu = 0$, 于是 $\nabla \lambda(X) \preceq 0$. 由 $\mathcal{B}X^2 > 0$, 得 $\lambda(X)\mathcal{B}X - \mathcal{A}X \succeq 0$, 所以 X 是 SDPEiCP 的一个解.

我们可以得到 SDPEiCP 正互补特征值存在的充要条件.

性质 2. 若 $\mathcal{B}X^2 > 0$, $X \succeq 0 (X \neq 0)$, SDPEiCP 有正互补特征值的充要条件是存在一个 $X \succeq 0$ 使得 $\mathcal{A}X^2 > 0$.

证明. 假设 (X, λ) 是 SDPEiCP 的一个解, 其中 $\lambda > 0$. 我们有 $\lambda BX^2 - AX^2 = 0$, 所以 $AX^2 = \lambda BX^2 > 0$.

另一方面, 若存在 $X \succeq 0$, 使得 $AX^2 > 0$ 成立, 令 $\bar{X} := \frac{X}{\|X\|_F}$, 所以 $\lambda(\bar{X}) = \frac{A\bar{X}^2}{B\bar{X}^2} > 0$. 因为问题 (P) 的可行域是有界闭的, 并且 $\lambda(X)$ 连续, 所以其最大值是正的且是可达的. 该问题的最优值显然满足 KKT 条件, 从定理 3 知, KKT 点所对应的目标函数 $\lambda(X)$ 恰好是 SDPEiCP 的互补特征值, 因此 SDPEiCP 有一个正互补特征值.

类似地, 我们有

性质 3. 若 $BX^2 < 0, \forall X \succeq 0 (X \neq 0)$, SDPEiCP 存在一个正互补特征值的充要条件是存在一个 $X \succeq 0$ 使得 $AX^2 < 0$.

事实上, 当 $BX^2 > 0, \forall X \succeq 0 (X \neq 0)$ 时, 我们有 SDPEiCP 的非平凡解和问题 (P), (Q) 的 KKT 点是一一对应的. 这样 SDPEiCP 的互补特征值 λ 的上, 下界可由问题 (P), (Q) 的目标函数 $\lambda(X)$ 的上下界给出, 并且此时上界是可达的. 即

性质 4. 若 $BX^2 > 0, \forall X \succeq 0 (X \neq 0)$, $\max_{X \succeq 0} \frac{AX^2}{BX^2}, \min_{X \succeq 0} \frac{AX^2}{BX^2}$ 分别是 SDPEiCP 互补特征值 λ 的上下界, 且上界是可达的.

证明. 假设 (X, λ) 是 SDPEiCP 任意一个非平凡解, 我们不妨令 $\|X\|_F = 1$. 由 SDPEiCP 的定义可知, 该解一定满足问题 (P) 的 KKT 条件, 此时 λ 对应目标函数 $\lambda(X)$, 所以我们有 $\min_{X \succeq 0} \frac{AX^2}{BX^2} \leq \lambda \leq \max_{X \succeq 0} \frac{AX^2}{BX^2}$. 因为问题 (P) 的最优解也满足 KKT 条件, 此时目标函数为 $\max_{X \succeq 0} \frac{AX^2}{BX^2}$, 所以该上界是可达的.

当张量 B 为如下特殊张量,

$$B_{ijkl} = \begin{cases} \frac{1}{2}, & i = k \neq j = l \text{ or } i = l \neq j = k, \\ 1, & i = j = k = l, \\ 0, & \text{else.} \end{cases} \quad (3.6)$$

我们可以验证 $BX^2 = \|X\|_F^2$ 对所有对称矩阵 X 成立, 此时 SDPEiCP 互补特征值有一个便于计算的上下界.

性质 5. 当 B 为 (3.6) 定义的张量时, SDPEiCP 的互补特征值 λ 满足 $\lambda_{\min} M(A) \leq \lambda \leq \lambda_{\max} M(A)$, 其中 $\lambda_{\min} M(A), \lambda_{\max} M(A)$ 分别是折叠矩阵 $M(A)$ 的最小, 最大特征值.

证明. 当 B 为 (3.6) 定义的张量时, $BX^2 = \|X\|_F^2$, 满足性质 4 的条件, 此时 SDPEiCP 的上下界分别为 $\max_{X \succeq 0} \frac{AX^2}{\|X\|_F^2}$ 和 $\min_{X \succeq 0} \frac{AX^2}{\|X\|_F^2}$. 由四阶张量的矩阵化定义可知, $AX^2 = \text{vec}(X)^T M(A) \text{vec}(X)$, 其中 $\text{vec}(X)$ 是矩阵 X 的向量化: $(\text{vec}(X))_{(i-1) \times n + j} = x_{ij}, \forall i, j = 1, \dots, n$, 此时我们有 $\|\text{vec}(X)\|^2 = \|X\|_F^2$. 所以性质 4 给出的 SDPEiCP 互补特征值 λ 的上下界分别等价于 $\max_{X \succeq 0} \frac{\text{vec}(X)^T M(A) \text{vec}(X)}{\|\text{vec}(X)\|^2}$ 和 $\min_{X \succeq 0} \frac{\text{vec}(X)^T M(A) \text{vec}(X)}{\|\text{vec}(X)\|^2}$. 显然, 我们有 $\lambda_{\max} M(A) \geq \max_{X \succeq 0} \frac{\text{vec}(X)^T M(A) \text{vec}(X)}{\|\text{vec}(X)\|^2}, \lambda_{\min} M(A) \leq \min_{X \succeq 0} \frac{\text{vec}(X)^T M(A) \text{vec}(X)}{\|\text{vec}(X)\|^2}$.

在这一节的最后, 我们指出 SDPEiCP 解的计算是 NP 难的. 考虑特殊的张量 B, A , 如果此时问题是 NP 难的, 则整个问题也是 NP 难的. 我们假设 B 为 (3.6) 定义的张量, A 是超对称张量: $a_{ijkl} = a_{jikl} = a_{kijl} = a_{ljk i}, i, j, k, l = 1, \dots, n$. 由性质 2 可知, 计算全部 SDPEiCP 的

解后, 我们可以判断是否存在一个非零 $X \succeq 0$, 使得 $\mathcal{A}X^2 > 0$. 然而下面我们将说明判断是否存在一个非零 $X \succeq 0$, 使得 $\mathcal{A}X^2 > 0$ 是 NP 难的.

在 [9](定理 3.3) 中, 作者证明下面二个锥实际上是同一个锥.

$$\begin{aligned} & \text{cone}\{a \circ a \circ a \circ a | a \in \mathbb{R}^n\} \\ & \text{sym cone}\{X \otimes X | X \succeq 0\} \end{aligned} \quad (3.7)$$

其中 sym 是一个对称化算子, 具体定义在 [9] 中. 当 \mathcal{A} 是超对称时, 判断是否存在 $X \succeq 0$, 使得 $\mathcal{A}X^2 > 0$ 等价于判断是否存在非零 $x \in \mathbb{R}^n$ 使得 $\mathcal{A}x^4 > 0$, 而后者在 [7](定理 11.2) 中被证明是 NP 难的, 所以我们有

性质 6. 计算 SDPEiCP 的全部解是 NP 难的.

由于计算 SDPEiCP 全部解的复杂性, 在下一节我们考虑用迭代优化算法来求解该问题.

4. 平移投影幂法

由上一节讨论我们知当 $\mathcal{B}X^2 > 0, \forall X \succeq 0 (X \neq 0)$ 成立时, SDPEiCP 的解与非线性约束优化问题 (P), (Q) 的 KKT 点一一对应, 利用问题 (P) 的特殊结构, 同时结合 [2, 11] 中提出的方法, 我们提出了一个平移投影幂法来求问题 (P) 的 KKT 点, 由此来计算 SDPEiCP 的解. 注意到问题 (Q) 的可行域是一个单纯形, 所以可以应用投影梯度法等优化方法来求其 KKT 点, 这里我们不再讨论. 首先我们给出一个关键定理, 它能保证我们提出的算法单调收敛.

定理 5. 假设 $W \in \Omega$, 函数 $F(X)$ 在 W 的一个开邻域 Σ 上凸且连续可微. 定义 $X := P_{\mathbb{R}_+^n}(\nabla F(W))$ ($P_{\mathbb{R}_+^n}$ 是半正定锥投影算子), 如果 $X \neq 0, V := \frac{X}{\|X\|_F} \in \Sigma, V \neq W$, 则 $F(V) - F(W) > 0$.

证明. 因为 X 是 $\nabla F(W)$ 在半正定锥上的投影, 所以 $\langle X, \nabla F(W) \rangle = \|X\|_F^2$. 对 $W \succeq 0, \langle \nabla F(W), W \rangle \leq \langle X, W \rangle$, 所以如果 $X \neq 0, V := \frac{X}{\|X\|_F} \neq W$,

$$\langle \nabla F(W), \frac{X}{\|X\|_F} - W \rangle \geq \|X\|_F - \langle X, W \rangle > 0, \quad (4.1)$$

对于 $\|W\|_F = 1$ 成立. 从 $F(X)$ 的凸性得, $F(V) - F(W) > 0$.

对于函数 $\lambda(X) = \frac{\mathcal{A}X^2}{\mathcal{B}X^2}$ 而言, 它的凸性一般不能得到保证, 我们可以在 $\lambda(X)$ 上加一个‘常数项’ $\alpha\|X\|_F^2$ 来使得目标函数凸, 也即我们可以将问题 (P) 等价地变换为

$$\begin{aligned} (\bar{P}) \quad & \max \quad \bar{F}(X) := \frac{\mathcal{A}X^2}{\mathcal{B}X^2} + \alpha\|X\|_F^2 \\ & \text{s.t.} \quad \|X\|_F = 1, \\ & \quad \quad X \succeq 0, \end{aligned} \quad (4.2)$$

其中参数 α 是预先给定的正常数. 根据定理 3 的推导过程, 我们可以验证问题 (\bar{P}) 的 KKT 条件等价于问题 (P) 的 KKT 条件, 因此我们可以通过求解问题 (\bar{P}) 的 KKT 点来求 SDPEiCP 的解. 矩阵函数 $\bar{F}(X)$ 的梯度是

$$\nabla \bar{F}(X) = \frac{2}{\mathcal{B}X^2}(\mathcal{A}X - \lambda(X)\mathcal{B}X) + 2\alpha X.$$

矩阵函数 $\bar{F}(X)$ 的 Hessian 阵本质上是一个四阶张量, 这里我们称之为 Hessian 张量, 记为 $\bar{\mathcal{H}}(X)$. 我们有

$$\begin{aligned}\bar{\mathcal{H}}(X) &= \nabla \left(\frac{2}{BX^2} \left(AX - \frac{AX^2}{BX^2} BX \right) \right) + 2\alpha \mathcal{I}. \\ &= 2 \frac{ABX^2 - 2AX \otimes BX - BAX^2 - 2BX \otimes AX}{(BX^2)^2} + \\ &\quad 8 \frac{AX^2 BX \otimes BX}{(BX^2)^3} + 2\alpha \mathcal{I},\end{aligned}\quad (4.3)$$

其中 \mathcal{I} 是 (3.6) 中定义的部分对称张量. 这里我们引入四阶张量的矩阵半正定的定义.

定义 1. 对部分对称四阶张量 \mathcal{A} : $a_{ijkl} = a_{klij} = a_{jikl}$, $\forall i, j, k, l = 1, \dots, n$. \mathcal{A} 是矩阵半正定的, 如果 $\langle AX, X \rangle \geq 0$, $\forall X \in S^n$.

记

$$\mathcal{H}(X) = 2 \frac{ABX^2 - 2AX \otimes BX - BAX^2 - 2BX \otimes AX}{(BX^2)^2} + 8 \frac{AX^2 BX \otimes BX}{(BX^2)^3}. \quad (4.4)$$

为了利用定理 5 中的结论, 我们需要保证 $\bar{F}(X)$ 凸. $\bar{F}(X)$ 的凸性等价于 $\bar{\mathcal{H}}(X)$ 的矩阵半正定性, 即

$$\langle \bar{\mathcal{H}}(X)Y, Y \rangle = \langle \mathcal{H}(X)Y, Y \rangle + 2\alpha \|Y\|_F^2 \geq 0, \forall Y \in S^n.$$

因为 $\langle \mathcal{H}(X)Y, Y \rangle + 2\alpha \|Y\|_F^2 \geq (\lambda_{\min}(M(\mathcal{H}(X))) + 2\alpha) \|Y\|_F^2$, 所以当 $\alpha \geq -\frac{\lambda_{\min}(M(\mathcal{H}(X)))}{2}$ 时, $\bar{\mathcal{H}}(X)$ 是矩阵半正定的. 令

$$\alpha = \max\left\{0, \frac{-\lambda_{\min}(M(\mathcal{H}(X)))}{2} + \tau\right\}, \quad (4.5)$$

其中 τ 是一个非负常量. 此时函数 $\bar{F}(X)$ 在 X 的邻域内凸. 求解问题 (\bar{P}) 的平移幂迭代算法见算法 1.

算法 1 平移投影幂法

Require: 给定张量 \mathcal{B} , \mathcal{A} , 和一个初始点 $X^0 \in \Omega$, τ

for $k = 1, 2, 3, \dots$ **do**

1. 计算 $\frac{\mathcal{H}(X^{k-1})}{2}$ (式 (4.4)) 展成的矩阵 $M(\frac{\mathcal{H}(X^{k-1})}{2})$ 的最小特征值, 由公式 (4.5) 得参数 α_{k-1}

2. 计算 $\lambda(X) + \alpha_{k-1} \|X\|_F^2$ 在点 X^{k-1} 处的梯度, 然后对其特征值分解为 $V_{k-1} D_{k-1} V_{k-1}^\top$, 将 D_{k-1} 中所有负值变为 0, 记为 \bar{D}_{k-1} , 得到投影为 $\bar{d}(X^{k-1}) = V_{k-1} \bar{D}_{k-1} V_{k-1}^\top$

3. 如果 $\bar{d}(X^{k-1}) = 0$ 则停止, 否则 $X^k = \frac{\bar{d}(X^{k-1})}{\|\bar{d}(X^{k-1})\|_F}$

end for

我们能够得到如下的算法收敛性结论.

定理 6. 在算法 1 中, 如果对某 k , $\bar{d}(X^k) = 0$, 或 $X^{k+1} = X^k$, 则 $(X^k, \lambda(X^k))$ 是 SDPEiCP 的一个解.

证明. 如果 $\bar{d}(X^k) = 0$, 即 $\nabla \bar{F}(X^k) = \frac{2}{B(X^k)^2}(\mathcal{A}X^k - \lambda(X^k)BX^k) + 2\alpha_k X^k$ 是负半定的. 但 $\langle \nabla \bar{F}(X^k), X^k \rangle = 2\alpha_k \|X^k\|_F^2 \geq 0$, 所以 $\alpha_k = 0$. 在这种情况下, 可以看出 $(X^k, \lambda(X^k))$ 是 SDPEiCP 的一个解.

若 $\bar{d}(X^k) \neq 0$, $X^{k+1} = X^k$, 则 $\nabla \bar{F}(X^k)$ 在半正定锥上的投影与矩阵 X^k 成正比. 由矩阵相关理论知, 若 X^k 是正定矩阵, 则 $\nabla \bar{F}(X^k)$ 也是正定矩阵且存在 $l > 0$ 使得 $\nabla \bar{F}(X^k) = lX^k$, 即 $\frac{2}{B(X^k)^2}(\mathcal{A}X^k - \lambda(X^k)BX^k) = (l - 2\alpha_k)X^k$. 由 $\langle \mathcal{A}X^k - \lambda(X^k)BX^k, X^k \rangle = 0$, 可知 $\mathcal{A}X^k - \lambda(X^k)BX^k = 0$, 所以此时 $(X^k, \lambda(X^k))$ 是 SDPEiCP 的一个解.

当 $X^k \succeq 0 (X \neq 0)$ 时, 存在单位正交矩阵 U 使得 $U^T X^k U = \text{diag}(\rho_1, \rho_2, \dots, \rho_r, 0, \dots, 0)$, 其中 r 是 X^k 的秩. 由 $\nabla \bar{F}(X^k)$ 的半正定锥的投影与 X^k 正相关和投影相关理论可知,

$$U^T \nabla \bar{F}(X^k) U = \frac{2}{B(X^k)^2} \begin{pmatrix} \beta_1 & & & & 0 \\ & \ddots & & & \\ & & \beta_r & & \\ & & & \beta_{r+1} & \\ & & & & \ddots \\ 0 & & & & & \beta_n \end{pmatrix} + 2\alpha_k \begin{pmatrix} \rho_1 & & & & 0 \\ & \ddots & & & \\ & & \rho_r & & \\ & & & 0 & \\ & & & & \ddots \\ 0 & & & & & 0 \end{pmatrix}, \quad (4.6)$$

其中 $\beta_{r+1}, \dots, \beta_n \leq 0$. 存在 $l > 0$ 使得 $\frac{2}{B(X^k)^2} \beta_i + 2\alpha_k \rho_i = l \rho_i$, $i = 1, \dots, r$. $\langle \mathcal{A}X^k - \lambda(X^k)BX^k, X \rangle = 0$ 等价于 $\sum_{i=1}^r \beta_i \rho_i = 0$. 综上所述我们有 $\beta_i = 0$, $i = 1, \dots, r$, $\beta_i \leq 0$, $i = r+1, \dots, n$, 即 $\mathcal{A}X^k - \lambda(X^k)BX^k$ 半负定, 此时 $(X^k, \lambda(X^k))$ 是 SDPEiCP 的一个解.

如果 $\bar{d}(X^k) = 0$ 在整个计算中没有出现, 在一定条件下, 我们可以得到算法 1 生成的序列 $\{(\lambda(X^k), X^k)\}$ 的单调收敛性.

定理 7. 假设算法 1 中, 对所有的 k , 都有 $\bar{d}(X^k) \neq 0$ 且 $X^{k+1} \neq X^k$, X^{k+1} 属于 $\bar{F}(X)$ 包含 X^k 的凸可行区域, 则 $\{\lambda(X^k)\}$ 是单调增且收敛到 SDPEiCP 的一个解.

证明. 由定理 5 可知, 序列 $\{\lambda(X^k)\}$ 单调增. 区域 Ω 为有界闭集且目标函数 $\lambda(X)$ 是连续函数, 序列 $\{\lambda(X^k)\}$ 有上界, 即存在 λ^* 使得 $\lim_{k \rightarrow \infty} \lambda(X^k) = \lambda^*$. 由 Ω 的有界闭性, $\{X^k\}$ 存在聚点. 我们证明它的每个聚点 X^* 都是 SDPEiCP 的解, 其对应的互补特征值为 λ^* . 由定理 5 中 (4.1) 式, 知 $\langle \nabla \bar{F}(X^k), X^{k+1} - X^k \rangle \geq 0$. 由 $\{\lambda(X^k)\}$ 单调有上界, 知 $\bar{F}(X^{k+1}) - \bar{F}(X^k) \rightarrow 0$. 由 $\bar{F}(X)$ 的凸性, 再结合 $\langle \nabla \bar{F}(X^k), X^{k+1} - X^k \rangle \geq 0$, 我们有

$$\langle \nabla \bar{F}(X^k), X^{k+1} - X^k \rangle \rightarrow 0,$$

即

$$\sqrt{\rho_{(1,k)}^2 + \dots + \rho_{(r_k,k)}^2} - \langle \nabla \bar{F}(X^k), X^k \rangle \rightarrow 0, \quad (4.7)$$

其中 r_k 表示 $\nabla \bar{F}(X^k)$ 非负特征值的个数, $\rho_{(i,k)}$, $i = 1, \dots, r_k$ 是 $\nabla \bar{F}(X^k)$ 的非负特征值. 由 α_k 的选取, 我们知 α_k 是有界的, 因此我们不妨假设存在子列 $\{k_i\}$ 使得 $\{X^{k_i}\}_{i=1}^\infty \rightarrow X^*$, $\{\alpha_{k_i}\}_{i=1}^\infty \rightarrow \alpha^*$. 由 $\nabla \bar{F}(X)$ 的连续性可知,

$$\sqrt{\rho_{(1,*)}^2 + \dots + \rho_{(r_*,*)}^2} - \langle \nabla \bar{F}(X^*), X^* \rangle = 0, \quad (4.8)$$

其中 r_* 表示 $\nabla \bar{F}(X^*)$ 非负特征值的个数, $\rho_{(i,*)}$, $i = 1, \dots, r_*$ 是 $\nabla \bar{F}(X^*)$ 的非负特征值.

若 $r_* = 0$ 时, 则 $\bar{d}(X^*) = 0$; 若 $r_* > 0$, 由 $X^* \in \Omega$, 此时必有

$$X^* = \frac{1}{\sqrt{\rho_{(1,*)}^2 + \cdots + \rho_{(r_*,*)}^2}} U \text{diag}(\rho_{(1,*)}, \dots, \rho_{(r_*,*)}, 0, \dots, 0) U^T,$$

其中 U 是 $\nabla \bar{F}(X^*)$ 特征向量组成的单位正交矩阵. 因此 $\frac{P_{R_+^n}(\nabla \bar{F}(X^*))}{\|P_{R_+^n}(\nabla \bar{F}(X^*))\|_F} = X^*$. 根据定理 6, $(X^*, \lambda(X^*))$ 是 SDPEiCP 的一个解.

由 Hessian 张量 $\bar{\mathcal{H}}(X)$ 的定义可知, 我们可以选择充分大的 α 使得 $\bar{\mathcal{H}}(X)$ 矩阵半正定对任意 $X \in \Omega$ 均成立, 此时定理 7 的凸可行区域条件会自动满足, 这样的选择与 [11] 中平移系数选择相同, 然而增大 α 会使得算法收敛变慢: 较大的 αX^k 会导致算法 1 中较小的 $\|X^{k-1} - X^k\|_F$ 值. 尽管定理 7 的条件看起来很严苛, 我们的数值实验经验表明: 即使 τ 取得很小 (通常意味着 α 很小), 所有的数值模拟例子均满足单调收敛性, 此时算法能够快速收敛.

5. 数值模拟

在这节, 我们给出了一些初步的数值模拟例子来证明提出的算法的有效性. 实验结果表明平移幂法能够有效地求解满足条件的 SDPEiCP. 我们在 MATLAB 2016b, Inter(R) Core(TM) i5-8250U CPU @1.60GHz 1.80GHz, 8.00GB RAM 环境下进行数值模拟.

为了方便实验, 我们让 \mathcal{B} 为张量 \mathcal{I} (定义于 (3.6)), 张量 \mathcal{A} 是 10 维的部分对称四阶张量, 其元素由高斯分布 $N(0, 1)$ 随机生成. 在计算中, 我们采用如下停止准则

$$|\lambda(X^k) - \lambda(X^{k-1})| \leq \text{tol},$$

其中 tol 为预先给定的精度. 我们进行了一些简单的数值实验来检验算法 1 的效果.

首先, 我们生成了一个随机的 \mathcal{A} , 因为元素的个数较大, 这里没有列出. 因为算法的全局收敛性没有得到保证, 仅收敛到问题 (P) 得一个 KKT 点, 即 SDPEiCP 一个解, 因此我们通过随机选取初始点 X^0 大量重复实验来寻找 SDPEiCP 所有的解. 具体随机生成初始点的方法为: 首先随机生成一个对称矩阵, 然后投影到半正定锥上, 最后行进单位化. 针对这个例子, 我们找到了下面两个解, 但是否有其它的解仍无法确定.

$$\lambda_1 = 14.9188, \quad \lambda_2 = 16.8800.$$

因为 α 选取过程会受预设非负常量 τ 的影响, 于是我们观察 τ 的数值变化对 λ 收敛速率的影响. 图 1 显示了对于某个随机生成的 \mathcal{A} 和相同起始点 X^0 , 不同的 τ 对 λ 收敛速度的影响, 在大量的模拟计算中, 我们观察到了相同的现象: τ 越小 λ 收敛越快, τ 越大, λ 收敛越慢. 尽管定理 7 的条件无法进行验证, 在我们所有的实验中, 算法产生的序列均满足单调收敛性.

此外, 我们展示了得到一个稳定点所需迭代次数分布的直方图 2. 它表明了不同停止精度下, 算法所需迭代次数的分布情况. 固定初始点 X^0 为 $\frac{I}{\sqrt{10}}$, I 是单位矩阵. 根据第一个实验所得, τ 设置为 0 效果最好. 我们进行 200 次模拟实验, 每次随机生成张量 \mathcal{A} 来进行实验. 在 2 中, 横轴表示迭代次数, 竖轴表示对应迭代次数的实验频数.

从图 2 可以看出, 不同的停止精度对收敛迭代次数影响不大, 这个现象可由图 1 来解释: 当 $\tau = 0$ 时, λ 值会在一定范围内快速收敛, 从而导致不同精度对迭代次数的影响较小.

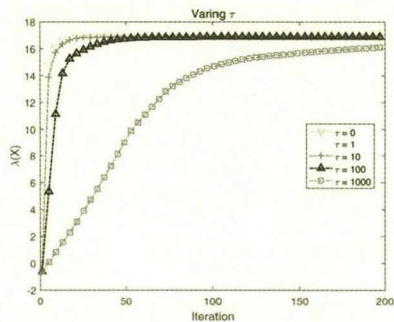
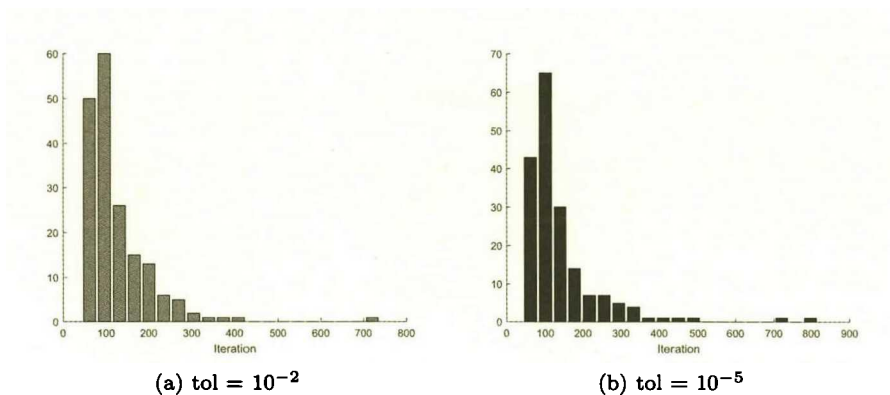
图 1: 不同 τ 对应的 λ 收敛情况

图 2: 不同停止精度对收敛速率的影响

6. 总 结

在本文中, 我们提出了一类特殊的张量特征值互补问题, 这类问题是矩阵特征值互补问题的推广, 同时也与已有的一类张量特征值互补问题有所区别. 在一定条件下, 通过将该互补特征值问题转化为变分问题, 我们给出了互补特征值存在的一个充分条件. 在一定对称性假设下, 我们建立了互补特征值问题同一类非线性约束优化问题的联系, 证明了该问题的 NP 难, 然后给出了一个平移投影幂法来计算这类问题. 在最后一节, 我们通过一些数值例子初步显示了算法的效果.

当特征矩阵要求是秩一时, SDPEiCP 等价于如下问题:

$$\text{求解 } (\lambda, x) \in (\mathbb{R}, \mathbb{R}^n) \text{ 满足 } \begin{cases} \lambda \mathcal{B}x^4 - \mathcal{A}x^4 = 0, \\ \lambda \mathcal{B}x^2 - \mathcal{A}x^2 \succeq 0. \end{cases} \quad (6.1)$$

这与已有的张量特征值互补问题不同, 上述问题何时解, 以及与 SDPEiCP, 其它特征值互补问题的关系将是我们下一步研究方向.

致谢. 感谢审稿人对本文给予了许多宝贵的建议与帮助.

参 考 文 献

- [1] Adly S, Rammal H. A new method for solving second-order cone eigenvalue complementarity problems[J]. *Journal of Optimization Theory and Applications*, 2015, 165(2): 563–585.
- [2] Chen Z M, Yang Q Z, Ye L. Generalized eigenvalue complementarity problem for tensors[J]. *Pacific Journal of Optimization*, 2017, 13(3):527–545.
- [3] Da Costa A P, Figueiredo I N, Júdice J J, et al. A complementarity eigenproblem in the stability analysis of finite dimensional elastic systems with frictional contact[M]//*Complementarity: applications, algorithms and extensions*. Springer, Boston, MA, 2001: 67–83.
- [4] Facchinei F, Pang J S. Finite-dimensional variational inequalities and complementarity problems[M]. Springer, 2003.
- [5] Fan J Y, Nie J W, Zhou A W. Tensor eigenvalue complementarity problems[J]. *Mathematical Programming*, 2018,170(2):507–539.
- [6] Fernandes L M, Fukushima M, Júdice J J, et al. The second-order cone eigenvalue complementarity problem[J]. *Optimization Methods and Software*, 2016, 31(1): 24–52.
- [7] Hillar C J, Lim L H. Most tensor problems are NP-hard[J]. *Journal of the ACM (JACM)*, 2013, 60(6): 45.
- [8] Hou J J, Ling C, He H J. A class of second-order cone eigenvalue complementarity problems for higher-order tensors[J]. *Journal of the Operations Research Society of China*, 2017, 5(1): 45–64.
- [9] Jiang B, Li Z, Zhang S. On cones of nonnegative quartic forms[J]. *Foundations of Computational Mathematics*, 2017, 17(1): 161–197.
- [10] Judice J J, Sherali H D, Ribeiro I M, et al. On the asymmetric eigenvalue complementarity problem[J]. *Optimization Methods & Software*, 2009, 24(4-5): 549–568.
- [11] Kolda T G, Mayo J R. Shifted power method for computing tensor eigenpairs[J]. *SIAM Journal on Matrix Analysis and Applications*, 2011, 32(4): 1095–1124.
- [12] Lavilledieu P, Seeger A. Existence de valeurs propres pour les systèmes multivoques: résultats anciens et nouveaux[J]. *Ann. Sci. Math. Québec*, 2001, 25: 47–70.
- [13] Lim L H. Singular values and eigenvalues of tensors: a variational approach[C]//1st IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, 2005. IEEE, 2005: 129–132.
- [14] Ling C, He H, Qi L. On the cone eigenvalue complementarity problem for higher-order tensors[J]. *Computational optimization and applications*, 2016, 63(1): 143–168.
- [15] Martins J A C, da Costa A P. Stability of finite-dimensional nonlinear elastic systems with unilateral contact and friction[J]. *International journal of solids and structures*, 2000, 37(18): 2519–2564.
- [16] Martins J A C, Barbarin S, Raous M, et al. Dynamic stability of finite dimensional linearly elastic systems with unilateral contact and Coulomb friction[J]. *Computer Methods in Applied Mechanics and Engineering*, 1999, 177(3-4): 289–328.
- [17] Qi L. Eigenvalues of a real supersymmetric tensor[J]. *Journal of Symbolic Computation*, 2005, 40(6): 1302–1324.
- [18] Queiroz M, Judice J, Humes Jr C. The symmetric eigenvalue complementarity problem[J]. *Mathematics of Computation*, 2004, 73(248): 1849–1863.
- [19] Seeger A. Eigenvalue analysis of equilibrium processes defined by linear complementarity conditions[J]. *Linear Algebra and its Applications*, 1999, 292(1-3): 1–14.

A CLASS OF TENSOR EIGENVALUE COMPLEMENTARITY PROBLEM

Luo Gang

(School of Mathematical Sciences, Nankai University, Tianjin 300071, China)

Yang Qingzhi

*(School of Mathematics and Statistics, Kashi University, Kashi 844006, China;
School of Mathematical Sciences, Nankai University, Tianjin 300071, China)*

Abstract

In this paper, we generalize the matrix eigenvalue complementarity problem which has wide application in mechanical systems. A positive semidefinite eigenvalue complementarity problem(SDPEiCP) is established using fourth-order tensor form. Some properties, like the existence of the solution, computational complexity, are studied. We show the relation between SDPEiCP and a nonlinear constrained optimization problem. A shifted power method is proposed to compute the solution of SDPEiCP at last.

Keywords: Positive semidefinite cone; eigenvalue complementarity problem; N-P hardness, shifted power method

2010 Mathematics Subject Classification: 65K10, 65F15, 15A18, 15A69, 90C22, 90C33