arXiv:2303.13847v1 [math.SP] 24 Mar 2023

鲁棒　正则单纯形张量的特征向量：猜想证明 ⋆

Lei Wang[a, *], Xiurui Geng[b, c, d], Lei Zhang[a]

[a] *the School of Microelectronics and Communication Engineering, Chongqing University, Chongqing 400044, China*[b] *Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China*
[c] *University of the Chinese Academy of Sciences, Beijing 100049, China*
[d] *Key Laboratory of Technology in Geo-Spatial Information Processing and Application System, Chinese Academy of Science, Beijing 100190, China*

摘要

张量特征对这一概念在过去几十年中受到了越来越多的关注。近期研究关注了一类特殊的对称张量，称为正则单纯形张量，它由 $n+1$ 维空间中的等角紧框架构造而成，并研究了特征对的鲁棒性。文献最后提出了一个猜想：正则单纯形张量的鲁棒特征向量恰好是框架中的向量。之后的一项研究从理论上证明了 $n=2$ 的情形成立。本文进一步推进研究，完成了上述猜想的证明。最后讨论了一些值得期待的未来研究方向。

关键词：正则单纯形张量、特征对、鲁棒性分析、局部最优性、约束优化。

美国数学学会主题分类：15A69，90C26

1. 引言

近年来，张量分析及其应用受到广泛研究，其中，张量特征对这一概念已在理论研究中得到广泛应用 [1, 2, 3, 4, 5] ，并在许多学科中催生了大量应用，例如潜变量模态 [6], 、高光谱图像处理 [7, 8], 、信号处理 [9] 等。

然而，已有研究也表明，大多数张量问题都是 NP难 [10], ，包括计算张量的全部特征对。据我们所知，迄今为止，只有两种算法能够求出张量的全部特征对 [4, 5]。因此，先前的大多数工作都致力于求出最大或最小的特征对，并开发了不同的优化算法，例如 [1, 11, 2]。其中一种经典算法称为张量幂法（TPM），

⋆本研究部分获得以下资助：中国国家自然科学基金（62271090）、重庆市自然科学基金（cstc2021jcyj-jqX0023）、国家重点研发计划（2021YFB3100800）、CCF海康威视开放基金（CCF-HIKVISION OF 20210002）、中国人工智能学会-华为MindSpore开放基金，以及北京智源人工智能研究院（BAAI）。

*通讯作者*Email address:* wanglei179@mails.ucas.ac.cn (Lei Wang)

它基于不动点迭代方案。TPM 的不动点（可能不止一个）就是张量的特征向量。关于张量幂方法，一个重要的问题是考察得到的特征向量是否鲁棒（详情见第 2.2 小节）。针对这一问题，先前的工作主要关注一些特殊类型的对称张量，其中一种受到广泛研究的类型称为正交可分解（odeco）张量 [12, 3, 6, 13, 14, 15],，它是正交矩阵分解的自然推广。odeco 张量的一个优良性质是，经典张量幂法可以精确提取生成该 odeco 张量的向量。换言之，可以通过张量幂法找到这类张量的鲁棒特征对。

然而遗憾的是，大多数对称张量都无法正交分解。基于这一点，近年来，一些研究者进一步将正交可分解张量（odeco 张量）的情形推广到更一般的情况：对称张量由某个等角集合（ES）或等角紧框架（ETF）生成，其中包含 r 个位于 n 维空间 [16] 中的向量。例如， $r = n$ 是紧框架的一种特殊情形，它构成标准正交基，并对应于正交可分解张量。当 $r = n + 1$ 时，该框架称为正则单纯形框架，由此生成的张量则称为正则单纯形张量。在 [16], 中，他们讨论了特征向量满足何种条件才具有鲁棒性，并通过一些理论证明研究了由 ES 或 ETF 生成的这些特殊类型的张量。

此外，在文献 [16], 的末尾，作者提出了一个猜想：正则单纯形张量的鲁棒特征向量恰好是框架中的向量（正式表述见第 2.3 小节的猜想 1）。后来，几位研究者进一步关注这一猜想。他们将正则单纯形张量的特征对方程重新表述为代数方程组，研究所有特征对的实特征结构，并试图确定其中哪些具有鲁棒性[17]。然而，他们只证明了 $n = 2$ 情形下的鲁棒性，并为 $n = 3$ 情形提供了实验结果作为依据。对于更大的 n 情形，证明将是一项艰巨的任务。此类证明方法主要面临两项困难：1）：随着 n 增大，所有特征对的数量呈指数增长，确定它们各自的谱半径需要大量计算；2）：所有特征对的特征结构也十分复杂，因此，对于大多数特征对而言，计算雅可比矩阵并进一步确定其谱半径都可能十分困难。本文基于仅有的两项相关工作 [16, 17] ，为避免这些困难，进一步研究并给出这一猜想的完整证明。本文的主要贡献总结如下:

张量的局部最大化特征对是优化理论中的另一个重要概念。我们首次研究了鲁棒特征对与局部最大化特征对之间的联系，并从理论上证明，鲁棒特征对是所有这些局部最大化特征对的子集。这一结论使我们能够简化 [17] 中采用的证明方法。具体而言，无需检查所有特征对的鲁棒性，只需关注这些局部最大化特征对即可。这是我们处理鲁棒性问题的主要发现，同时我们也给出了尚未解决的猜想的证明。

本文其余部分安排如下。第 2 节介绍与主题相关的一些预备知识，包括张量特征对和鲁棒特征对的定义，以及焦点正则单纯形张量。第 3 节探讨张量的鲁棒特征对与局部最大化特征对之间的关系。随后，第 4 节给出该猜想的证明。第 5 节讨论一些未来的研究方向。

2. 预备知识

本文将介绍一些必要的记号、定义和引理。在本文中，与许多张量相关研究的做法一致，[18, 19, 20], 高阶张量用粗体欧拉花体字母表示，例如 $\mathcal{A}$。矩阵用粗体大写字母表示，例如 A；向量用粗体小写字母表示，例如 a。集合和子集用黑板粗体大写字母表示，例如 $\mathbb{A}$。

一个 d阶张量记作 $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_d}$，其中 d 是 $\mathcal{A}$的阶数，I_j ($j \in \{1, 2, \ldots, d\}$)是 j第 $\mathcal{A}$模的维度。i_1的元素以整数元组 ($i_2, i_d, \ldots, a_{i_1}$) 为索引，记作 ($i_2$,$i_d$,...,$1 \le i_1 \le I_1$)$1 \le i_d \le I_d$,...,[18]。若张量的元素在对指标 T^m进行任意置换后保持不变，则称其为对称张量。令 $\mathbb{R}^n$(m) 表示所有此类实对称张量构成的空间。给定一个 n阶、$\mathcal{S}$ 维的对称张量 n和一个向量 $\in \mathbb{R}^{n \times 1}$u $\mathcal{S}$，有 $\sum_{i_1\ i_2\ \cdots\ \mathrm{u} i_m=1}$,,,i_2,i_m $u_{i_1} \cdots u_{i_m}$,..., $\mathcal{S}$，且 $^{m-1}$u n表示一个n维列向量，其第 j个元素为 ($\mathcal{S}$u^{m-1})$_j$ = $\sum_{i_2\ \cdots\ , , i_m=1}$ s_j ,i_2,...,i_m $u_{i_2} \cdots u_{i_m}$[4]。此外，$n\mathcal{S}$u^{m-2} 是一个矩阵，其第（ $n \times n$,i）个元素为 (ju$\mathcal{S}$)$^{m-2}$,i, , $_j$ = $\sum_{i_3\ \cdots\ , i_m=1}$ s_i ,j,...,i_3

$\mathbf{1}_n$ 表示一个 $n \times 1$ 列向量，I$_n$ 表示一个 $n \times n$ 单位矩阵。
∇ 是梯度算子。null(A) 表示 A 的零空间。P$_{\mathbf{A}}^{\perp}$ 是 A 的正交补算子。

定义 1（谱半径）。矩阵 A $\in \mathbb{R}^{n \times n}$ 的谱半径是 A 的所有特征值绝对值中的最大值，记作 ρ(A) $= max|\sigma_i$(A)|，其中 σ_i(A) i =1, 2, . . . , n 是 A 的 n 特征值。

定义 2（外积）。给定 m 向量 a$^{(i)} \in \mathbb{R}^{I_i \times 1}$ (i = 1, 2, . . ., m)，它们的外积 a$^{(1)} \circ$ a$^{(2)} \circ \cdots \circ$ a$^{(m)}$ 是一个 m阶张量，记作 $\mathcal{A}$，其大小为 $I_1 \times I_2 \times \cdots \times I_m$。其元素是对应向量元素的乘积，即$a_{i_1,\ i_2,\ \ldots\ldots\ i_m}$ = a$_{i_1}^{(1)}$ a$_{i_2}^{(2)}$ . . . a$_{i_m}^{(d)}$.当 a$^{(1)}$ = a$^{(2)}$ = $\cdots$ = a$^{(m)}$ = a ($I_1 = I_2 = \cdots = I_m = I$) 时，为简洁起见，我们使用记号 $\mathcal{A}$ = a$^{\circ m}$，其中 $\mathcal{A}$ 是一个阶数为 m、维度为 I的对称张量。

2.1. 张量特征对的优化理论

本节简要介绍与张量特征对问题相关的优化理论。通过考虑以下约束优化模型，可以理解并推导张量特征对的概念:

$$\begin{cases} \max\limits_{\mathbf{v}} & \mathcal{S}\mathbf{v}^m \\ \text{s.t.} & \mathbf{v}^{\mathrm{T}}\mathbf{v} = 1 \end{cases} . \tag{1}$$

式（1）的拉格朗日函数定义如下:

$$L(\mathbf{v}, \lambda) = \frac{1}{m}\mathcal{S}\mathbf{v}^m + \frac{\lambda}{2}(1 - \mathbf{v}^{\mathrm{T}}\mathbf{v}). \tag{2}$$

当 L(v λ) 对 v 的梯度（也称为 ）为 0 时，可以推导出对称张量的特征对，这一概念由 Lim 和 Qi 于 2005年 独立定义:

定义 3. *[21, 22] [*对称张量的特征对*]* 给定张量 $\mathcal{S} \in T^m(\mathbb{R}^n)$，一对 $(\lambda, \mathbf{v})$ 是 $\mathcal{S}$ 的特征对，当且仅当

$$\mathcal{S}\mathbf{v}^{m-1} = \lambda\mathbf{v}, \tag{3}$$

其中 $\lambda \in \mathbb{C}$ 是特征值，$\mathbf{v} \in \mathbb{C}^{n\times 1}$ 是对应的特征向量，满足 $\mathbf{v}^{\mathrm{T}}\mathbf{v} = 1$.

在给定优化模型的情况下，二阶导数信息对于判断驻点是否为局部最优解起着重要作用。$L(\mathbf{v}, \lambda)$ 对 $\mathbf{v}$ 的二阶导数，也称为式（2）的海森矩阵，记为

$$\mathbf{H}(\mathbf{v}) = (m-1)\mathcal{S}\mathbf{v}^{m-2} - \lambda\mathbf{I}_n, \tag{4}$$

其中 $\mathbf{I}_n$ 是一个 $n \times n$ 单位矩阵。

然后，可以通过检验下列矩阵是否负定，并确定其每个特征值的符号，来识别式（1）的局部最优解:

$$\mathbf{K} = \mathbf{P}_{\mathbf{v}}^{\perp}\mathbf{H}(\mathbf{v})\mathbf{P}_{\mathbf{v}}^{\perp} \tag{5}$$

a式（5）的详细推导说明可参见我们的先前工作 [23].

2.2. 对称张量的鲁棒特征对

一种广泛使用的张量特征对求解算法称为张量幂法，该方法基于以下映射函数

$$\phi(\mathbf{v}) = \frac{\mathcal{S}\mathbf{v}^{m-1}}{\|\mathcal{S}\mathbf{v}^{m-1}\|} \tag{6}$$

并执行以下迭代过程:

$$\mathbf{v}_{k+1} \mapsto \frac{\mathcal{S}\mathbf{v}_k^{m-1}}{\left\|\mathcal{S}\mathbf{v}_k^{m-1}\right\|}. \tag{7}$$

对比式（7）与 KKT 条件（3）可知，张量幂法可理解为利用一阶 KKT 方程的迭代法。

给定一个随机初始化向量 $\mathbf{v}_0$，并迭代执行上述公式（7），直至满足某个终止条件，即可得到一个特征对。于是，$\mathcal{S}$ 的鲁棒特征向量是张量幂法的一个吸引不动点特征向量 $\mathbf{v}$，可通过计算映射函数的雅可比矩阵的谱半径来识别，该矩阵记为并定义为

$$\mathbf{J}(\mathbf{v}) = \frac{m-1}{\lambda}\left(\mathcal{S}\mathbf{v}^{m-2} - \lambda\mathbf{v}\mathbf{v}^{\top}\right), \tag{8}$$

其中 $\lambda \neq 0$，这表明鲁棒性检验应排除这些特征值为 0 的特征对。如果 $\mathbf{J}(\mathbf{v})$ 的谱半径小于 1，则称相应的特征对是鲁棒特征对。详细推导可参见 [16] 中的引理 3.2 和 3.3。

注1。需要指出的是，*[16]* 中引理3.3的证明不够严谨，或者虽然正确，但对式（8）的最终表达式缺少必要的说明。

严格来说，根据 *[16],* 中定理 3.3 证明所给出的推导，只能得到如下雅可比矩阵：

$$\mathbf{J}(\mathbf{v}) = \frac{d-1}{|\lambda|}\left(\mathcal{S}\mathbf{v}^{d-2} - \lambda\mathbf{v}\mathbf{v}^{\top}\right), \tag{9}$$

这在讨论同一问题的后续研究中有所采用 *[17]*。细微差别在于，分子上额外进行了一个取绝对值的操作 λ。请注意，尽管这一差别不会影响谱半径的确定，但在确定谱半径时，我们仍需再计算一次绝对值，如定义 1 所示。从这个意义上说，(8) 与 (9) 之间的差异微不足道。

这里，我们仍希望说明 (8) 也是正确的，因此可以直接使用原式 (8)。理由如下。当 d 为奇数时，(λ,v) 和 (−λ, –v) 都是特征对；不失一般性，我们只考虑特征值非负的特征对；当 d 为偶数时，特征对以具有相同特征值的 $\pm$ 对出现，我们只选择满足 $\sum_i v_i > 0$ 的 v。此外，我们可以推出，对于偶数 m，$\lambda = \mathcal{S}\mathrm{v}^m = \sum_{i=1}^{n+1} (\mathrm{w}_j^{\mathrm{T}}\mathrm{w}_i)^m > 0$ 。

2.3. 正则单纯形标架与张量

本节将介绍一类特殊的对称张量，称为正则单纯形张量。首先介绍广义等角集的定义如下：

定义 4（参考文献 [16] 中的定义 4.1）。等角集合（ES）是由向量 w_1、.....、$\mathrm{w}_r \in \mathbb{R}^n$ 组成的集合，其中 $r \geq n$；若存在 $\alpha \in \mathbb{R}$，使得

$$\alpha = |\langle \mathbf{w}_i, \mathbf{w}_j \rangle|, \forall i \neq j \quad and \quad \|\mathbf{w}_i\| = 1, \forall i.$$

此外，若等角集合（ES）满足以下条件，则称其为等角紧框架（ETF）：

$$\mathbf{W}\mathbf{W}^{\mathrm{T}} = a\mathbf{I}, \quad \mathbf{W} := (\mathbf{w}_1, \cdots, \mathbf{w}_r) \in \mathbb{R}^{(n-1)\times r} \tag{10}$$

例如，当 $r = n$ 且 $a = 1$时，标准正交基 $\{\mathrm{w}_1$、.....、$\mathrm{w}_n\} \subset \mathbb{R}^{n\times n}$ 构成一个等角紧框架（ETF），其中 $\alpha = 0$。当 $r = n+1$ 且 $a = \frac{n+1}{n}$时，$\{\mathrm{w}_1$、.....、$\mathrm{w}_{n+1}\} \subset \mathbb{R}^{(n)\times(n+1)}$称为正则单纯形标架，$\alpha = -\frac{1}{n}$。$\mathrm{w}_i$、$i = 1$、2.....、$n+1$ 称为框架中的向量。在二维空间中，正则单纯形标架是一个正三角形。

那么，正则单纯形张量是由正则单纯形标架和 t he 以下公式 orm:

$$\mathcal{S} := \sum_{i=1}^{n+1} \mathbf{w}_i^{\circ m} \tag{11}$$

其中，$\circ$ 是定义 2 中定义的外积，$\mathcal{S}$ 是一个阶数为 m 、维度为 n的对称张量。本文主要分析正则单纯形张量的鲁棒特征对，并尝试解决以下猜想：

猜想 1（参考文献 [16] 中的猜想 4.7）。正则单纯形张量的鲁棒特征向量恰好是标架中的向量。

这一猜想最初由 [16] 提出。后来，也有几位研究者在 [17] 中探讨了这个问题。需要指出的是，上述猜想包含以下含义：1）正则单纯形张量的鲁棒特征对可能并不存在，这一点已由

[16], 中的实验观察到，后续分析也将从理论上加以论证；2）如果存在，它们也只能是标架中的这些向量。在 [16], 中，作者仅证明了，对于维度为 n 、阶数为 m ，且 $n \geq 2$、 $m \geq 3$ 和 $n+m \geq 7$ 的张量，标架中的这些向量确实是鲁棒特征向量。（详情请参见 [16] 的定理 4.5。）然而，这些标架中的向量是唯一鲁棒向量这一点尚未得到证明，因此便留下了上述猜想 1。换言之，猜想中的“恰好”一词仍有待证明。为此，必须检查给定正则单纯形张量的所有特征向量的鲁棒性，并从所有候选向量中确定哪些是鲁棒的。后续研究 [17] 采用了这一策略。具体而言，他们首先分析了与正则单纯形张量相关的所有特征对的实特征结构，然后通过检查每个特征对处雅可比矩阵的谱半径，判断各特征对是否具有鲁棒性。然而，可以看出，随着 n 和 m 增大，这项任务的计算开销很大，因为特征对的总数会呈指数增长。因此，他们只给出了最简单情形 $n=2$ 的证明，而更高阶的 n 情形则难以计算和检验。此外，对于大多数特征对，隐式计算雅可比矩阵也很困难，更不用说进一步确定其谱半径了。

据我们所知，迄今为止，这两篇文献是仅有的两项聚焦于识别正则单纯形张量鲁棒特征对的研究。本文以这两项研究为基础，进一步探讨上述猜想，并最终完成了对该猜想的证明。详情如下。

3. 鲁棒特征对与局部最大化特征对之间的关系

请注意，在相关研究中，通常通过第 2.2 小节中定义的雅可比矩阵的谱半径来检验特征对的鲁棒性。然而，如上所述，确定所有特征对的谱半径有些复杂，尤其是在较高的 n 和 m情形下。

本节中，我们将给出一种用于鲁棒性检验的辅助判据。比较 (8) 与 (4) 可知，两者都包含一项 $\mathcal{S}\mathbf{v}^{m-2}$。进一步考察二者之间的关系，可以得出以下引理：

引理 1. 给定张量 $\mathcal{S} \in T^m(\mathbb{R}^n)$ 的一个特征对 $(\lambda, \mathbf{v})$，如果该特征对是鲁棒的，那么它必定是模型 (1) 的局部极大解。反之则不一定成立。

证明. 首先，将 (5) 中的 $\mathbf{K}$ 改写为以下形式：

$$
\begin{aligned}
\mathbf{K} &= \mathbf{P}_{\mathbf{v}}^{\perp}\mathbf{H}(\mathbf{v})\mathbf{P}_{\mathbf{v}}^{\perp} \\
&= (\mathbf{I}_n - \mathbf{v}\mathbf{v}^{\mathrm{T}})[(m-1)\mathcal{S}\mathbf{v}^{m-2} - \lambda\mathbf{I}_n](\mathbf{I}_n - \mathbf{v}\mathbf{v}^{\mathrm{T}}) \\
&= (\mathbf{I}_n - \mathbf{v}\mathbf{v}^{\mathrm{T}})[(m-1)\mathcal{S}\mathbf{v}^{m-2} - \lambda\mathbf{I}_n - (m-2)\mathbf{v}\mathbf{v}^{\mathrm{T}}] \\
&= (m-1)\mathcal{S}\mathbf{v}^{m-2} - \lambda\mathbf{I}_n - (m-2)\mathbf{v}\mathbf{v}^{\mathrm{T}} - (m-1)\mathcal{S}\mathbf{v}^{m-2}\mathbf{v}\mathbf{v}^{\mathrm{T}} + \lambda\mathbf{v}\mathbf{v}^{\mathrm{T}} + (m-2)\mathbf{v}\mathbf{v}^{\mathrm{T}} \\
&= (m-1)\mathcal{S}\mathbf{v}^{m-2} - \lambda\mathbf{I}_n - (m-1)\lambda\mathbf{v}\mathbf{v}^{\mathrm{T}} + \lambda\mathbf{v}\mathbf{v}^{\mathrm{T}},
\end{aligned}
$$

在最后一个等式中，我们利用了 $\mathcal{S}\mathbf{v}^{m-2}\mathbf{v} = \mathcal{S}\mathbf{v}^{m-1} = \lambda\mathbf{v}$。通过与式 (8) 对比，可得

$$
\lambda\mathbf{J} = \mathbf{K} + \lambda(\mathbf{I}_n - \mathbf{v}\mathbf{v}^{\mathrm{T}}). \tag{12}
$$

此外，可以验证

$$\mathbf{J}\mathbf{v} = \frac{m-1}{\lambda}(\mathcal{S}\mathbf{v}^{m-2}\mathbf{v} - \lambda\mathbf{v}\mathbf{v}^{\mathrm{T}}\mathbf{v}) = \frac{m-1}{\lambda}(\lambda\mathbf{v} - \lambda\mathbf{v}) = 0 \cdot \mathbf{v}, \tag{13}$$

其中我们重复使用 $\mathcal{S}\mathrm{v}^{m-2}\mathrm{v} = \mathcal{S}\mathrm{v}^{m-1} = \lambda\mathrm{v}$，而上述结果意味着 v 是 J 的特征向量，对应特征值 0，因此，其他特征向量必定位于 v 的零空间中，记作 $null$(v)。

类似地，向量 v 同时也是 K、$\mathrm{I}_n - \mathrm{v}\mathrm{v}^{\mathrm{T}}$ 的特征向量，其中满足

$$\mathbf{K}\mathbf{v} = 0 \cdot \mathbf{v}, \quad (\mathbf{I}_n - \mathbf{v}\mathbf{v}^{\mathrm{T}})\mathbf{v} = 0 \cdot \mathbf{v}, \tag{14}$$

K、$\mathrm{I}_n - \mathrm{v}\mathrm{v}^{\mathrm{T}}$ 的其他特征向量也必定位于 $null$(v) 中。

这表明，它们的特征分解共享同一个特征向量矩阵，记作 $\mathrm{V} = [\mathrm{v}\ null(\mathrm{v})] \in \mathbb{R}^{n\times n}$。换言之，对于式 (12)，它们可以由同一个正交矩阵 V 对角化，如下所示：

$$\lambda\mathbf{V}^{\mathrm{T}}\mathbf{J}\mathbf{V} = \mathbf{V}^{\mathrm{T}}[\mathbf{K} + \lambda(\mathbf{I}_n - \mathbf{v}\mathbf{v}^{\mathrm{T}})]\mathbf{V}. \tag{15}$$

注意，$\mathrm{I}_n - \mathrm{v}\mathrm{v}^{\mathrm{T}} \in \mathbb{R}^{n\times n}$ 是一个秩为 $n-1$ 的投影矩阵，其特征值为 1, 1, . . . , 1, 0，其中特征值 1 的个数为 $n-1$。

根据上述关系式 (15)，一旦获得 K 的特征值分布，就可以确定 J 的相应特征值及其谱半径。因此，下面将讨论在什么条件下，特征对可能是鲁棒的。

Assu明 J 和 K 的特征值按以下方式排序 为

$$\sigma_1^{\mathbf{J}} \le \sigma_2^{\mathbf{J}} \le \cdots \le \sigma_n^{\mathbf{J}}. \tag{16}$$

$$\sigma_1^{\mathbf{K}} \le \sigma_2^{\mathbf{K}} \le \cdots \le \sigma_n^{\mathbf{K}}. \tag{17}$$

于是，我们可以得到

$$\lambda\sigma_i^{\mathbf{J}} = \begin{cases} \sigma_i^{\mathbf{K}} + \lambda, & \sigma_i^{\mathbf{I}_n - \mathbf{v}\mathbf{v}^{\mathrm{T}}} = 1 \\ \sigma_i^{\mathbf{K}} = 0, & \sigma_i^{\mathbf{I}_n - \mathbf{v}\mathbf{v}^{\mathrm{T}}} = 0 \end{cases}. \tag{18}$$

Acc根据 K 的所有特征值的符号，下面讨论三种不同情况 :

情况1：如果 (λ, v) 是模型 (1) 的局部极小解，则其对应的矩阵 K 是半正定的。结合 (14) 可得 $\sigma_{min}^{\mathbf{K}} = \sigma_1^{\mathbf{K}} = 0$，且 $0 \le \sigma_2^{\mathbf{K}} \le \cdots \le \sigma_{n-1}^{\mathbf{K}}$。结合 (15) 可得

$$\sigma_{min}^{\mathbf{J}} = \sigma_1^{\mathbf{J}} = 0, \quad \lambda\sigma_i^{\mathbf{J}} = \sigma_i^{\mathbf{K}} + \lambda \ge \lambda, \quad i = 2, \ldots, n. \tag{19}$$

这表明 J 的所有特征值均为非负数，并且可以推导出 $\rho(\mathrm{J}) = \sigma_n^{\mathbf{J}} \ge 1$，这表明相应的特征对不可能是鲁棒的。

情况2：如果（λ, v）是模型 (1) 的鞍点解，则其对应的矩阵 K 不确定，这表明其特征值包括负值、正值和 0。假设第 k 个特征值等于 0，则有：

$$\sigma_1^{\mathbf{K}} \le \cdots \le \sigma_{k-1}^{\mathbf{K}} \le \sigma_k^{\mathbf{K}} = 0 \le \sigma_{k+1}^{\mathbf{K}} \cdots \le \sigma_{n-1}^{\mathbf{K}}. \tag{20}$$

由于 (14) 表明 J 始终有一个特征值为 0，根据 (15)，这表明 $\sigma_k^{\mathbf{J}} = 0$，且 $\sigma^{\mathbf{I}_n - \mathbf{v}\mathbf{v}_k}{}^{\mathrm{T}} = 0$。这进一步意味着，除第 k 个特征值外，$\lambda(\mathbf{I}_{n-1} - \mathbf{v}\mathbf{v}^{\mathrm{T}})$ 的其余特征值均为 λ，于是有

$$\lambda\sigma_i^{\mathbf{J}} = \sigma_i^{\mathbf{K}} + \lambda \leq \lambda, \quad i = 1, \ldots, k-1. \tag{21}$$

$$\lambda\sigma_i^{\mathbf{J}} = \sigma_i^{\mathbf{K}} + \lambda \geq \lambda, \quad i = k+1, \ldots, n-1. \tag{22}$$

可以看出，对于 $i = k+1, \ldots, n-1$，相应的特征值 $\lambda\sigma_i^{\mathbf{J}}$ 确实大于 λ。从这个意义上说，无论 $\lambda\sigma_i^{\mathbf{J}}$ 对于 $i = 1, \ldots, k-1$ 的取值范围如何，始终有 $\rho(\mathbf{J}) > 1$，且相应的特征对不可能是稳健的。

情形 3：如果 $(\lambda, \mathbf{v})$ 是模型 (1) 的局部极大解，则其对应矩阵 K 为负定，这表明 $\sigma_{max}^{\mathbf{K}} = \sigma_n^{\mathbf{K}} = 0$ 和 $\sigma_1^{\mathbf{K}} \leq \cdots \leq \sigma_{n-1}^{\mathbf{K}} \leq 0$。结合 (15)，可以得出

$$\sigma_{max}^{\mathbf{J}} = \sigma_{n-1}^{\mathbf{J}} = 0, \quad \lambda\sigma_i^{\mathbf{J}} = \sigma_i^{\mathbf{K}} + \lambda \leq \lambda, \quad i = 1, \ldots, n-2. \tag{23}$$

在这种情况下，可以看出所有特征值均被限制在小于 λ 的范围内。此外，如果

$$\lambda\sigma_i^{\mathbf{J}} = \sigma_i^{\mathbf{K}} + \lambda > -\lambda, i = 1, \ldots, n-2. \tag{24}$$

且 (23) 不可能满足严格等式，则可以得出所有特征值都位于区间 $(-\lambda, \lambda)$ 内，这表明 $\rho(\mathbf{J}) < 1$,

以上对三种情形的讨论表明，当且仅当 $(\lambda, \mathbf{v})$ 是模型 (1) 的局部极大解，且满足如 (24) 所示的附加条件时，它才可能是鲁棒解。而局部极小解和鞍点解必定不是鲁棒解。

反过来，如果一个特征对是鲁棒特征对，且 $\rho(\mathbf{J}) < 1.$，这等价于表明所有特征值都被限制在 $(-\lambda, \lambda)$ 的范围内。同样，由 (18) 可得 $\sigma_i^{\mathbf{K}} < 0$，这意味着 $(\lambda, \mathbf{v})$ 是模型 (1) 的局部极大解。证明完毕。 □

总而言之，如果一个特征对是鲁棒特征对，那么它必定是相应约束模型 (1) 的局部最大化特征对。反之则不总是成立。因此，所有鲁棒特征对构成的集合是这些局部最大化特征对的子集，其规模始终不大于所有特征对的集合。换言之，只有这些局部最大化特征对才可能成为鲁棒特征对的候选项。

这样，给定一个张量（这里不要求它是正则单纯形张量，因为上述引理适用于任意对称张量），我们无需像先前工作那样分析并确定所有特征对中 J 的谱半径。相反，只需关注该张量的这些局部最大化特征对即可。

从这个意义上说，这样的引理将为猜想 1 中鲁棒特征对的识别带来两项优势：

1）需要检查的特征对数量大幅减少，因为这些局部最大化特征对的数量始终不超过所有特征对的数量，尤其是在n 变大时。

2)：自然，沿着第一个优点，我们可以避免先前工作中存在的困难，即计算雅可比矩阵以及确定大多数特征对的谱半径，因为这里我们只关注这些局部最大化特征对（标架中的这些向量），其计算较为容易，具体可见下文。

因此，接下来的目标是确定哪些是正则单纯形张量的局部极大解，这也可能是另一个重要问题。我们在先前工作[23] 中系统地研究了这一问题，并从理论上证明，正则单纯形张量的这些局部极大解仅为正则单纯形标架中的这些向量。详情请参阅 [23] 中的引理1至5。

从这个意义上说，我们只检验了标架中这些向量的鲁棒性。我们将在下一节完成猜想 1 的证明。

4. 猜想 1 的证明

猜想 1（参考文献 [16] 中的猜想 4.7）。正则单纯形张量的鲁棒特征向量恰好是标架中的向量。

证明。对于正则单纯形标架 $\mathrm{W} = [\mathrm{w}_1、\ldots、\mathrm{w}_n] \in \mathbb{R}^{(n-1)\times n}$ 中的向量，有 $\mathrm{w}_i^{\mathrm{T}}\mathrm{w}_i = 1$、$\mathrm{w}_i^{\mathrm{T}}\mathrm{w}_j = \alpha = -\frac{1}{n}$for $i \neq j$。$\mathrm{WW}^{\mathrm{T}} = \sum_{i=1}^{n} \mathrm{w}_i\mathrm{w}_i^{\mathrm{T}} = \frac{n+1}{n}\mathrm{I}$。可以计算得，当 $\mathrm{v} = \mathrm{w}_j$ 时，有

$$\begin{aligned}
\mathcal{S}\mathbf{v}^{m-2} = \mathcal{S}\mathbf{w}_j^{m-2} &= (\sum_{i=1}^{n+1} \mathbf{w}_i^{\circ m})\mathbf{w}_j^{m-2} = \sum_{i=1}^{n+1} (\mathbf{w}_j^{\mathrm{T}}\mathbf{w}_i)^{m-2}\mathbf{w}_i\mathbf{w}_i^{\mathrm{T}} \\
&= \mathbf{w}_j\mathbf{w}_j^{\mathrm{T}} + \frac{1}{(-n)^{m-2}} \sum_{i=1, i\neq j}^{n+1} \mathbf{w}_i\mathbf{w}_i^{\mathrm{T}} = \mathbf{w}_j\mathbf{w}_j^{\mathrm{T}} + \frac{1}{(-n)^{m-2}}(\frac{n+1}{n}\mathbf{I} - \mathbf{w}_j\mathbf{w}_j^{\mathrm{T}}) \\
&= (1 - \frac{1}{(-n)^{m-2}})\mathbf{w}_j\mathbf{w}_j^{\mathrm{T}} + \frac{n+1}{(-n)^{m-2}n}\mathbf{I}. \qquad (25)
\end{aligned}$$

此外，相应的特征值可计算为

$$\lambda = \mathcal{S}\mathbf{w}_j^m = \sum_{i=1}^{n+1} (\mathbf{w}_j^{\mathrm{T}}\mathbf{w}_i)^m = 1 + \frac{n}{(-n)^m}. \qquad (26)$$

于是，根据 (8)，特征向量 w_j 处的雅可比矩阵满足

$$\mathbf{J}(\mathbf{w}_j)\mathbf{w}_j = \frac{m-1}{\lambda}(\mathcal{S}\mathbf{w}_j^{m-2} - \lambda\mathbf{w}_j\mathbf{w}_j^{\mathrm{T}})\mathbf{w}_j = 0 \cdot \mathbf{w}_j \qquad (27)$$

这意味着 w_j 本身是 J(w_j) 的特征向量，其特征值为 0。自然地，其他特征向量必定位于 w_j 的零空间中，记作 $null(\mathrm{w}_j)$。进一步利用 (25) 和 (26)，可以轻松验证，关于 J(w_j) 只有两个不同的特征值，即

$$0, \underbrace{\frac{(n+1)(m-1)}{1+(-n)^{m-2}n}, \ldots, \frac{(n+1)(m-1)}{1+(-n)^{m-2}n}}_{n-1}. \qquad (28)$$

将 m 的奇数和偶数情形分别考虑后，可得 J(w_j) 的谱半径为

$$\rho(\mathbf{J}(\mathbf{w}_j)) = \begin{cases} \frac{(n+1)(m-1)}{n^{m-1}-1}, & m = 2k+1, k \in \mathbb{N} \\ \frac{(n+1)(m-1)}{1+n^{m-1}}, , & m = 2k, k \in \mathbb{N} \end{cases}, \qquad (29)$$

对于奇数 m 情形，由于 $1+(-n)^{m-2}n = 1-n^{m-1} < 0$，分子应改为正数。

注意到 $n \geq 2, m \geq 3$，可以验证 $n+m \geq 7$，有 $\rho(\mathrm{J}(\mathrm{w}_j)) < 1$，并且对应正则单纯形张量的标架中的向量是稳健的。

□

注 2. 从 (25) 和 (26) 可以看出，当 $\mathrm{v} = \mathrm{w}_j$ 时，我们可以显式地得到 $\mathcal{S}\mathrm{v}^{m-2}$ 的结果以及相应的特征值 λ，然后轻松确定 $\mathrm{J}(\mathrm{w}_j)$ 的特征值分布。然而，对于其他特征对 v，很难确定 (25) 中 $\mathrm{v}^{\mathrm{T}}\mathrm{w}_i$ 的结果以及 λ 的结果，从而推导出关于 $\mathrm{J}(\mathrm{w}_j)$ 的显式形式。这是先前工作采用逐一检查所有特征对这一策略的主要难点。因此，先前工作只检查了 $n = 2$ 的情况，这可能是最简单的情形。当 n 变大时，这将是一项复杂的任务。

注意到，式 (29) 也已在 *[16]* 的定理 4.5 中推导得出。详情请参见该定理证明中的第一个公式。因此，推导出式 (29) 后的后续分析与 *[16],* 中定理 4.5 的证明相同，故省略上述细节。然而，两者的区别在于，定理 4.5 给出的结果仅为上界，而这里通过分析特征值分布，精确确定了谱半径。换言之，我们进一步证明，谱半径可以达到定理 4.5 中的上界。

上述证明还证实，对于 $(m, n) = (2, 3)$ 和 $(m, n) = (3, 3)$ 这两种情况，正则单纯形张量不存在鲁棒特征对。这也与参考文献 *[16]* 的表 I 中给出的实验结果一致。

5. 结论与未来工作

本文主要研究关于正则单纯形张量的鲁棒特征对的一项猜想，并给出了证明。与先前工作不同，我们的证明方法首先考察了鲁棒特征对与局部最大化特征对之间的关系；后者是优化领域中的另一个重要概念。随后，建立这一关系有助于我们通过仅考虑所有特征对中的一个较小子集，大幅简化目标猜想的证明过程。借助这一做法，我们最终证明了该猜想中的结论。

原始参考文献 [16] 中提出的其他猜想也值得在下一阶段研究。详细内容可参见猜想 4.8 和 5.2，以及问题 5.1。

参考文献

[1] Tamara G. Kolda 和 Jackson R. Mayo. 用移位幂法计算张量特征对. SIAM 矩阵分析与应用期刊, 32(4):1095–1124, 2011年.[2] Ariel Jaffe、Roi Weiss 和 Boaz Nadler. 用牛顿修正方法计算对称张量的实特征对. SIAM 矩阵分析与应用期刊, 39(3):1071–1094, 2018年.[3] Tamara G Kolda. 正交张量分解. SIAM 矩阵分析与应用期刊, 23(1):243–255, 2001年.

[4] Chun Feng Cui、Yu Hong Dai 和 Jiawang Nie. 对称张量的所有实特征值. SIAM 矩阵分析与应用期刊, 35(4), 2014年.[5] Liping Chen、Lixing Han 和 Liangmin Zhou. 通过同伦方法计算张量特征值. SIAM 矩阵分析与应用期刊, 37(1):290–319, 2016年.[6] Animashree Anandkumar、Rong Ge、Daniel Hsu、Sham M Kakade 和 Matus Telgarsky. 用于学习潜变量模型的张量分解. 机器学习研究杂志, 15(1):2773–2832, 2014年.[7] X. Geng 和 L. Wang. Npsa：非正交主偏度分析. IEEE图像处理汇刊, 29:6396–6408, 2020年.[8] Lei Wang 和 Xiurui Geng. 对称张量的实特征对及其在独立成分分析中的应用. IEEE控制论汇刊, 页码 1–14, 2021年.[9] Liqun Qi 和 Kok Lay Teo. 多元多项式最小化及其在信号处理中的应用. 全局优化杂志, 26(4):419–433, 2003年.[10] Christopher J Hillar 和 Lek-Heng Lim. 大多数张量问题都是 NP困难的. ACM杂志（JACM）, 60(6):45, 2013年.[11] Tamara G. Kolda 和 Jackson R. Mayo. 用于计算广义张量特征对的自适应移位幂法. SIAM 矩阵分析与应用期刊, 35(4):1095–1124, 2014年.[12] Elina Robeva. 对称张量的正交分解. SIAM 矩阵分析与应用期刊, 37(1):86–102, 2016年.[13] Yiju Wang 和 Liqun Qi. 关于高阶超对称张量的逐次超对称秩一分解. 数值线性代数及其应用, 14(6):503–519, 2010年.[14] Cun Mu、Daniel J. Hsu 和 Donald Goldfarb. 近似正交可分解对称张量的逐次秩一近似. SIAM 矩阵分析与应用期刊, 36(4):1638–1659, 2015年.[15] Jianze Li、Konstantin Usevich 和 Pierre Comon. 用于同步正交对称张量对角化的全局收敛Jacobi型算法. SIAM 矩阵分析与应用期刊, 39(1):1–22, 2018年.[16] Tommi Muller、Elina Robeva 和 Konstantin Usevich. 对称张量的鲁棒特征向量. SIAM 矩阵分析与应用期刊, 43(4):1784–1805, 2022年.[17] Adam Czaplinski、Thorsten Raasch 和 Jonathan Steinberg. 正则单纯形张量的实特征结构. CoRR, abs/2203.01865, 2022年.[18] Tamara G Kolda 和 Brett W Bader. 张量分解及其应用. SIAM评论, 51(3):455–500, 2009年.

[19] C. Lu、J. Feng、Y. Chen、W. Liu、Z. Lin 和 S. Yan。《采用新型张量核范数的张量鲁棒主成分分析》。IEEE模式分析与机器智能汇刊，42(4):925–938，2020年。[20] Piotr Koniusz、Lei Wang 和 Anoop Cherian。《用于动作识别的张量表示》。IEEE模式分析与机器智能汇刊，44(2):648–665，2022年。[21] Liqun Qi。《实对称张量的特征值》。符号计算杂志，40(6):1302–1324，2005年。[22] Lek-Heng Lim。《张量的奇异值和特征值：一种变分方法》。载于《计算多传感器自适应处理进展》，2005年首届 IEEE国际研讨会，页码 129–132。IEEE，2005年。[23] Lei Zhang、Lei Wang、Xiurui Geng。《正则单纯形张量的局部最优特征向量》。CoRR，abs/2303.00274，2023年。