arXiv:2304.04043v1 [stat.ME] 8 Apr 2023

统计学 a nd 高秩张量估计中的计算速率

Chanwoo Lee MiaoyanWang 威斯康星大学——麦迪逊分校 威斯康星大学——麦迪逊分校chanwoo.lee@wisc.edu miaoyan.wang@wisc.edu

摘要

高阶张量数据集常见于推荐系统、神经影像学和社交网络。本文提出了用于从含噪观测中估计可能具有高秩的信号张量的概率方法。我们考虑一种生成式潜变量张量模型，它同时涵盖高秩模型和低秩模型，包括但不限于简单超图函数模型、单指标模型、低秩CP模型和低秩Tucker模型。我们针对信号张量估计，全面研究了统计极限和计算极限。我们发现，高维潜变量张量的秩为对数级；这一事实解释了低秩张量在应用中的普遍性。此外，我们提出了一种多项式时间谱算法，可达到计算最优速率。我们证明，统计-计算差距仅在三阶或更高阶的潜变量张量中出现。我们通过数值实验和两个真实数据应用展示了所提方法的实际优势。

关键词：Te 张量估计、潜变量张量模型、统计计算效率

1 引言

高阶张量的分析近年来在统计学、机器学习和数据科学领域受到广泛关注。高阶张量数据集被收集并应用于推荐系统（Baltrunas 等人，2011年；Bi 等人，2018年）、社交网络（Bickel和Chen，2009年）、神经影像学（Zhou等人，2013年）、基因组学（Hore 等人，2016年）以及纵向数据分析（Hoff，2015年）等领域。一个例子是多组织表达数据（Wang 等人，2019年）。该数据集收集了多个个体不同组织的全基因组表达谱，从而形成了基因 × 个体 × 组织的三阶张量。另一个例子是超图网络，其中边可以连接两个以上的顶点。基于超图考察多元交互有助于理解分子系统（Michoel 和 Nachtergaele，2012年）和计算机视觉（Agarwal 等人，2006年）中的复杂网络。张量被自然地用于表示这类超图结构。除了许多重要应用之外，张量方法还在数据分析中展现出了有效性，这是经典的基于向量或基于矩阵的方法无法实现的（Han 等人，2022年；Lee 和 Wang，2021a）。

施加于目标张量的一种常见结构是低秩性。常见的低秩模型包括 CP低秩模型（Kolda 和 Bader，2009年；Sun 等人，2017年）、Tucker 低秩模型（Zhang 和 Xia，2018年）和块模型（Wang 和 Zeng，2019年）。尽管低秩假设广受欢迎，但假设张量维度趋于无穷大时张量的秩保持固定，未免过于局限。特别是，低秩假设对

表1：我们与广义线性模型（Hu 等，2021年；Wang 和 Li，2020年；Zhang 和 Xia，2018年）和简单超图函数模型（Balasubrama-nian，2021年）以往研究结果的比较。我们提出了两种方法，即最小二乘估计量（LSE）和双投影谱估计量（DSE）。我们的结果支持高秩张量估计，并在均方误差（MSE）速率方面实现了统计与计算最优性。 *对于广义线性模型，只有在强信噪比条件下才存在多项式算法。

	广义线性模型	简单超图函数模型	**我们的方法** (LSE)	**我们的方法** (DSE)
m 阶张量的均方误差率	$d^{-(m-1)}$	$d^{-2m/(m+2)}$	$d^{-(m-1)}$	$d^{-m/2}$
（例如，当 $m=3$ 时）	d^{-2}	$(d^{-6/5})$	(d^{-2})	$(d^{-3/2})$
允许高秩性	×	√	√	√
最优性分析	√	×	√	√
多项式算法	×/√*	×	×	√

到逐元素变换，且不足以表示张量的特殊结构（Lee 和 Wang，2021a）。此外，低秩张量处处稀疏，随机矩阵/张量几乎必然是满秩（Udell 和 Townsend，2019）。这促使我们开发一种更灵活的模型，以处理可能具有高秩的张量。

1.1 我们的贡献

我们提出了 一种潜变量张量模型，兼顾低秩和高秩张量。我们的模型涵盖但不限于大多数现有的张量模型，例如 CP模型（Kolda 和 Bader，2009年）、Tucker模型（Zhang 和 Xia，2018年）、广义线性模型（Wang 和 Li，2020年；Hu 等人，2021年）、单指标模型（Ganti 等人，2017）以及简单超图函数模型（Balasubramanian，2021年）。表 1 从统计和计算视角将我们的工作与以往结果进行了比较，下面对此进行总结。

首先，尽管真实数据应用中高秩张量十分常见，我们仍为低秩方法在实践中的成功提供了严谨的理论依据。我们证明，由潜变量张量模型生成的 d 维张量的对数秩为 $\mathcal{O}(\log d)$，当 $d \to \infty$ 时。这一关键谱性质从统计学视角解释了低秩近似的合理性。

其次，我们发现了高阶张量估计中统计与计算最优性之间的差距。我们证明，该问题的 $d^{-(m-1)}$统计最小最大最优速率为。然而，根据超图植入团猜想（HPC），我们证明该速率无法由任何多项式时间算法实现（Luo 和 Zhang，2022年）。随后，我们证明一个更慢的速率 $d^{-m/2}$ 是计算上最优的，并且可以由多项式时间算法实现。基于这两个界，我们揭示了差距区域 $[d^{-(m-1)},\ d^{-m/2}]$，在该区域内，估计在统计上是可能的，但在计算上是不可能的。这一现象不同于具有 $m=2$ 的矩阵问题，在后者中不存在差距区域。图1 展示了高阶张量估计中的这一统计-计算差距。

第三，我们提出了两种具有精度保证的估计方法：最小二乘估计（LSE）和双投影谱估计（DSE）。LSE 达到了信息论下界 $d^{-(m-1)}$，展现了其统计最优性。然而，LSE 的计算可能需要非多项式时间复杂度。随后，我们利用双投影谱方法（Zhang 和 Xia，2018年）的思想以及潜变量张量的对数秩性质，提出了DSE。我们证明，DSE 在多项式时间估计量的子类中达到了最优界 $d^{-m/2}$ 。

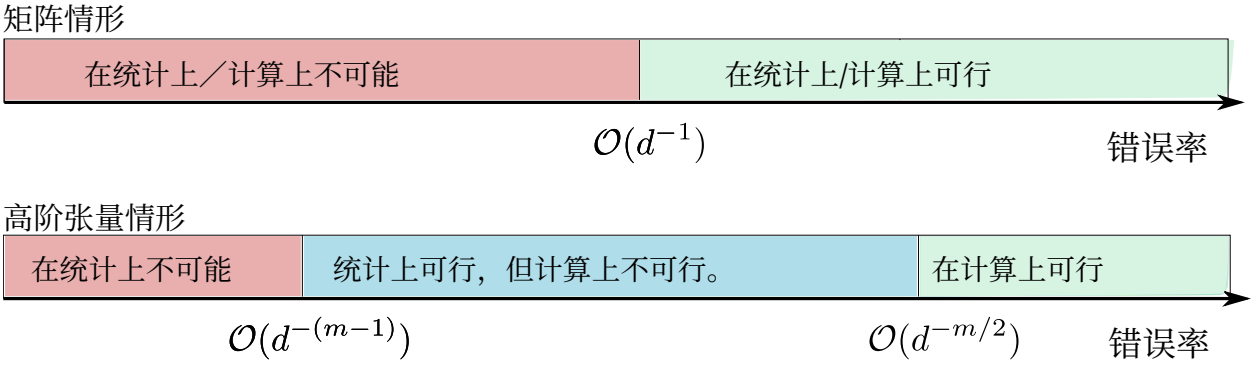


图1：高秩张量估计问题中均方误差的统计边界和计算边界。这里，信号是一个阶数为 m、维度为（d，……，d）的张量。我们发现，统计-计算差距仅出现在满足 $m \geq 3$ 的高阶张量中。

最后，我们通过模拟和数据示例展示了我们方法的有效性。我们将模型应用于功能性磁共振成像脑图像和作物产量数据集。我们的估计量成功地从含噪观测中恢复出真实的脑图像。对作物产量数据集进行聚类分析揭示了各国的区域模式。数值分析证明了所提出方法的实际效用。

1.2 相关工作

这项工作与张量分析领域的广泛文献相关，但也与之有明显区别。我们回顾了几条相关研究脉络，以资比较。

低秩模型下的张量模型。过去几十年来，关于低秩模型下结构化张量估计的研究成果大量涌现，包括 CANDECOMP/PARAFAC（CP）模型（Kolda和Bader，2009年；Sun等，2017年）、Tucker模型（Zhang和Xia，2018年）以及块模型（Wang和Zeng，2019年；Han 等人，2022年）。为克服对信号张量的低秩假设，研究者提出了单指标模型（Ganti等，2017年）和广义线性模型（Wang和Li，2020年；Hu等，2021年；Han 等人，2022年）。然而，这些工作仍然假设底层潜在张量在链接函数意义下具有低秩性。基于低秩的模型属于参数化方法，因为它们使用有限数量的参数，即一组 r 分解因子，对数据张量进行建模。相比之下，我们的方法不假设低秩结构。我们从潜变量张量模型的角度考虑高秩性，并采用非参数方法，允许参数数量为无限多个。在许多统计问题中，非参数方法相较于参数化方法的优势已经得到报道（Pananjady和Samworth，2022年；Gao等，2015年；Bickel和Chen，2009年）。

图函数与超图子函数 关于图函数与超图子函数的研究与我们的工作相关。图函数是一种可测函数，表示一系列可交换随机图的极限（Udell and Townsend，2019年；Xu，2018年；Klopp 等，2017年；Gao 等，2015年；Chan and Airoldi，2014年）。与图函数类似，超图子函数（Zhao，2015年；Lov´asz，2012年）是 m一致超图的极限函数，其中的边可以连接具有 $m \geq 3$ 的 m 个顶点。超图子函数为建模多元交互提供了强大工具，正如图函数能够用于建模成对交互一样。不同于矩阵（其中二元函数足以表示图函数）（Lov´asz and Szegedy，2006年），一般超图子由（$2^m - 2$）元函数表示（Zhao,

2015年）。简单超图子（Balasubramanian，2021年）将 m元函数作为模型灵活性与高效估计之间的折中。简单超图子与我们的潜变量张量模型具有共同点，即二者都考虑 m元函数。然而，我们的模型使用向量值潜变量，而简单超图子使用标量值潜变量。两种方法的比较将在第7节讨论。

高秩张量估计 近期已有一些尝试对高秩信号张量进行分析。例如，Lee 和 Wang (2021b) 提出了一种称为符号秩的新张量秩概念，可以用较低的复杂度对高秩张量进行建模。尽管该工作提供了一种多项式时间算法，但所提出的算法需要求解数量多达张量维度阶数的子优化问题，且极小极大最优性尚不明确。另一个工作中，Lee 和 Wang (2021b) 提出了另一种高秩模型，称为置换平滑张量模型。置换平滑张量模型与简单超图子（Balasubramanian，2021年）相似，因为二者都假设潜变量为标量值潜变量。相比之下，我们的模型允许一般的向量值潜变量。从算法角度来看，Lee 和 Wang(2021b) 要求潜函数满足严格的单调性假设，才能保证算法精度。相反，我们的双投影谱算法无需对模型作任何进一步假设，即可达到计算最优速率。详细比较将在第5节进行。

张量问题中的统计-计算差距 另一个相关主题是张量问题中的统计-计算差距。在张量补全（Barak 和 Moitra，2016年；Wang 和 Li，2020年）、张量主成分分析（Richard 和 Montanari，2014年；Zhang 和 Xia，2018年；Han 等人，2022年）以及多路聚类（Luo 和 Zhang，2022年）中，均发现了内在的统计-计算差距。由于理解矩阵与张量之间的根本区别十分重要，计算极限分析受到了广泛关注。对于矩阵，计算极限是基于平均情形归约方案推导得出的，该方案建立在 Erd¨os-R´enyi 图中植入团问题的常见困难性假设之上（Hazan 和 Krauthgamer，2011年；Ma 和 Wu，2015年；Berthet 和 Baldin，2020年；Wu 和 Xu，2021年）。然而，由于张量具有多路结构，将植入团检测直接应用于张量问题较为复杂（Luo 和 Zhang，2020年）。近年来，人们提出了植入超图团，以解决这一问题（Zhang 和 Xia，2018年；Luo 和 Zhang，2022年）。我们的工作也致力于理解统计-计算差距，但针对的是更具挑战性的高秩张量估计问题。

1.3 记号与组织

我们用 $\mathbb{R}$、 $\mathbb{N}$、 $\mathbb{N}_+$ 分别表示实数、整数和正整数集合。令 $[d] = \{1, \ldots, d\}$ 表示具有 $d \in \mathbb{N}_+$的 d 集合。对于两个正数列 $\{a_d\}, \{b_d\}$，若对某个常数 $c > 0$有 $\lim_{d\to\infty} a_d/b_d \le c$，则记为$a_d \lesssim b_d$；若对某些常数 $c_1, c_2 > 0$有 $c_1 \le \lim_{d\to\infty} a_d/b_d \le c_2$，则记为 $a_d \asymp b_d$。我们用 $\mathcal{O}(\cdot)$表示大 O 记号，用 $\mathcal{O}(\cdot)$表示隐藏对数因子的变体。我们用粗体小写字母（例如 $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{u}$）表示向量，用粗体大写字母（例如 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$）表示矩阵。对于矩阵 $\boldsymbol{A} \in \mathbb{R}^{d_1 \times d_2}$，$\mathrm{SVD}_r(\boldsymbol{A})$表示由 $\boldsymbol{A}$的前 r 个左奇异向量组成的矩阵，$\|\boldsymbol{A}\|_{\mathrm{sp}}$ 表示 $\boldsymbol{A}$的谱范数。

设 $\Theta \in \mathbb{R}^{d_1 \times \cdots \times d_m}$为一个阶数为$m$（$d_1$，……，$d_m$）维的张量。我们用 $\Theta(i_1$，……，$i_m)$ 表示由（i_1，……，i_m）$\in [d_1] \times \cdots \times [d_m]$索引的张量元素。我们定义弗罗贝尼乌斯范数和

∞张量的范数 Θ 如下:

$$\|\Theta\|_F = \sqrt{\sum_{(i_1,\ldots,i_m)\in[d_1]\times\cdots\times[d_m]} \Theta^2(i_1,\ldots,i_m)}, \quad \|\Theta\|_\infty = \max_{(i_1,\ldots,i_m)\in[d_1]\times\cdots\times[d_m]} |\Theta(i_1,\ldots,i_m)|.$$

张量 $\mathcal{C} \in \mathbb{R}^{r_1\times\cdots\times r_m}$与矩阵 $\boldsymbol{U}_{(k)}\in \mathbb{R}^{d_k\times r_k}$, $k \in [m]$的多线性乘法定义为

$$(\mathcal{C} \times_1 \boldsymbol{U}^{(1)} \times \cdots \times_m \boldsymbol{U}^{(m)})(i_1,\ldots,i_m) = \sum_{j_1=1}^{r_1} \cdots \sum_{j_m=1}^{r_m} \mathcal{C}(j_1,\ldots,j_d)\boldsymbol{U}^{(1)}(i_1,j_1)\cdots\boldsymbol{U}^{(m)}(i_m,j_m),$$

这会产生一个阶为m $(d_1,\ldots,d_m)$维的张量。我们使用Unfold$_k(\cdot)$表示展开操作，即沿模态 k 将张量重塑为矩阵，其中 $k \in [m]$。如果张量 $\Theta \in \mathbb{R}^{d_1\times\cdots\times d_m}$ 满足秩(Unfold$_k(\Theta)$) r_k =，则称其具有 Tucker 秩（$r_1,\ldots,r_m$）。在本文中，除非另有说明，我们保留“张量秩”一词，用于表示上述定义的 Tucker 秩。如果 E 事件以高概率发生，则 $\mathbb{P}(E)$随着张量维度$d \to \infty$趋于 1。最后，对于一个 m维向量 $\boldsymbol{d} = (d_1,\ldots,d_m)$，我们使用以下简写表示法，

$$\bar{d} := \max_{k\in[m]} d_k, \quad \underline{d} := \min_{k\in[m]} d_k, \quad d_* := \prod_{k\in[m]} d_k.$$

对于一个 s 维向量 $\boldsymbol{x} = (x_1,\ldots,x_m)$ 以及一个满足$\alpha_i \in \mathbb{N}$的整数分拆 $\alpha = \alpha_1 + \cdots + \alpha_s$，我们使用简写记号

$$\partial \boldsymbol{x}^\alpha = \partial x_1^{\alpha_1}\cdots\partial x_s^{\alpha_s}. \tag{1}$$

本文其余部分安排如下。在第2节，我们提出一种潜变量张量模型。该模型允许高秩张量，并将以往文献中的现有张量模型纳入其中，作为特殊情形。我们发现，良好的高维潜变量张量具有对数秩；这一事实解释了低秩张量在应用中的普遍性。在第3节，我们建立了该问题的极小极大速率及相应的最小二乘估计量。第4 节说明，没有任何多项式时间算法能够达到统计最优性，从而揭示了统计与计算最优性能之间的差距。随后，我们给出一种多项式时间双投影谱算法，可实现计算最优速率。第5节展示合成数据与真实数据分析。主要定理的证明见第6节。最后，我们在第7节讨论并总结全文。所有技术引理和补充结果均置于附录。

2 高阶张量的潜变量模型

2.1 模型表述

假设我们观测到一个由以下模型生成的阶数为 m （d_1、.....、 d_m）维数据张量 $\mathcal{Y}$

$$\mathcal{Y} = \Theta + \mathcal{E}, \tag{2}$$

其中， $\Theta \in \mathbb{R}^{d_1\times\cdots\times d_m}$ 是一个待研究的未知信号张量， $\mathcal{E}$ 是由独立同分布（i.i.d.）标准高斯随机变量组成的噪声张量。

潜变量张量模型 信号张量 Θ由以下模型生成。

- 非参数函数：存在一个潜在（未知）多元函数 f：$\mathbb{R}^{s_1} \times \cdots \times \mathbb{R}^{s_m} \to \mathbb{R}$，使得

$$\Theta(i_1, \ldots, i_m) = f(\boldsymbol{a}_{i_1}^{(1)}, \ldots, \boldsymbol{a}_{i_m}^{(m)}), \text{ for all } (i_1, \ldots, i_m) \in [d_1] \times \cdots \times [d_m], \tag{3}$$

其中，$\boldsymbol{a}_{i_k}^{(k)} \in \mathbb{R}^{s_k}$ 表示每个 $k \in [m]$ 和 $i_k \in [d_k]$对应的潜在 s_k 维向量。

- 正则性假设：

  – 完备潜空间：潜向量的支撑集包含于闭球中，即对所有 $k \in [m]$和 $i_k \in [d_k]$，$\boldsymbol{a}_{i_k}^{(k)} \in B(\mathbb{R}^{s_k}, \|\cdot\|_\infty)$。这里，$B(\mathbb{R}^{s_k}, \|\cdot\|_\infty)$ 表示 $\mathbb{R}^{s_k}$ 中关于无穷范数的单位球。

  – 解析潜函数：我们假设潜函数 f 是 M-解析的，使得

$$\sup_{\boldsymbol{x}_k \in B(\mathbb{R}^{s_k}, \|\cdot\|_\infty), k \in [m]} \left| \frac{\partial^{|\boldsymbol{\alpha}|} f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m)}{\partial (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m)^{\boldsymbol{\alpha}}} \right| \leq M^{|\boldsymbol{\alpha}|} \boldsymbol{\alpha}!, \tag{4}$$

对于某个常数 $M \in \mathbb{R}$ 以及所有多重指标 $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m) \in \mathbb{N}^m$。这里，我们采用简写记号 $|\boldsymbol{\alpha}| := \sum_{k \in [m]} \alpha_k$, $\boldsymbol{\alpha}! := \prod_{k \in [m]} \alpha_k!$, and (1).

我们用 $\mathcal{P}$（$\boldsymbol{d}$、$\boldsymbol{s}$、M）表示满足上述假设的潜变量张量模型，其中$\boldsymbol{d}$ =（d_1，……，d_m）表示（2）中的张量维度，$\boldsymbol{s}$ =（s_1，……，s_m）表示（3）中的潜维度，且 $M \in \mathbb{R}$ 是（4）中的正则性常数。我们将 $\Theta \in \mathcal{P}$（$\boldsymbol{d}$、$\boldsymbol{s}$、M）记为（2）中的信号张量。

接下来我们将说明，我们的潜变量张量模型是一个广泛的模型族，涵盖了许多现有张量模型作为特例。

例 1（CP低秩模型）。CP低秩张量是广受欢迎的张量模型之一（Hitchcock，1927；Kolda 和 Bader，2009）。现在我们说明，CP低秩张量属于我们的潜变量张量模型。设 Θ 是一个具有 CP s-秩的低秩张量，满足

$$\Theta = \sum_{r=1}^{s} \lambda_r \boldsymbol{u}_r^{(1)} \otimes \cdots \otimes \boldsymbol{u}_r^{(m)},$$

其中 $\lambda_r > 0$和 $\boldsymbol{u}_r^{(k)} \in B(\mathbb{R}^{d_k}, \|\cdot\|_\infty)$ 对所有 $r \in [s]$和 $k \in [m]$. 均成立。这里，$\otimes$表示向量的外积。定义向量 $\boldsymbol{\lambda}$ =（λ_1，……，λ_s）$\in \mathbb{R}^s$ 以及因子矩阵 $\boldsymbol{A}^{(k)} \in \mathbb{R}^{d_s \times s}$ 如下：

$$\boldsymbol{A}^{(k)} = \begin{pmatrix} | & | & & | \\ \boldsymbol{u}_1^{(k)} & \boldsymbol{u}_2^{(k)} & \cdots & \boldsymbol{u}_s^{(k)} \\ | & | & & | \end{pmatrix} = \begin{pmatrix} — & (\boldsymbol{a}_1^{(k)})^T & — \\ — & (\boldsymbol{a}_2^{(k)})^T & — \\ & \vdots & \\ — & (\boldsymbol{a}_{d_k}^{(k)})^T & — \end{pmatrix}, \quad \text{for all } k \in [m].$$

那么，CP s-秩张量 Θ属于我们的潜变量张量模型，其潜维度为（$s, \ldots, s$），满足

$$\Theta(i_1, \ldots, i_m) = f(\boldsymbol{a}_{i_1}^{(1)}, \ldots, \boldsymbol{a}_{i_m}^{(m)}) \quad \text{for all } (i_1, \ldots, i_m) \in [d_1] \times \cdots \times [d_m],$$

其中， $\boldsymbol{a}_{i_k}^{(k)} \in \mathbb{R}^s$ 是因子矩阵 $\boldsymbol{A}^{(k)}$ 的第 i_k 行， $f\colon \mathbb{R}^s \times \cdots \times \mathbb{R}^s \to \mathbb{R}$ 是定义为 $f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m) = \langle \boldsymbol{\lambda}, \boldsymbol{x}_1 \circ \cdots \circ \boldsymbol{x}_m \rangle$ 的解析函数。此处， $\circ$ 是逐元素乘积，也称为 Hadamard 乘积。

示例2（Tucker 低秩模型）。Tucker 低秩性广泛应用于张量分析（Tucker、1966；Kolda 和 Bader，2009）。我们展示 Tucker 低秩张量也属于我们的潜变量张量模型。令 Θ 为一个具有 Tucker（s_1，...，s_m）-秩的低秩张量，使得

$$\Theta = \mathcal{C} \times_1 \boldsymbol{A}^{(1)} \times_2 \cdots \times_m \boldsymbol{A}^{(m)},$$

其中， $\mathcal{C} \in \mathbb{R}^{s_1 \times \cdots \times s_m}$ 是一个核心张量， $\boldsymbol{A}^{(k)} \in \mathbb{R}^{d_k \times s_k}$ 是 $k \in [m]$ 的因子矩阵。张量 Θ 可以表示为具有（$s_1, \ldots, s_m$）-潜维度的潜变量张量模型，

$$\begin{aligned}\Theta(i_1, \ldots, i_m) &= \mathcal{C} \times_1 \boldsymbol{a}_{i_1}^{(1)} \times_2 \cdots \times \boldsymbol{a}_{i_m}^{(m)} \\ &= f(\boldsymbol{a}_{i_1}^{(1)}, \ldots, \boldsymbol{a}_{i_m}^{(m)}) \text{ for all } (i_1, \ldots, i_m) \in [d_1] \times_2 \cdots \times [d_m],\end{aligned}$$

其中， $\boldsymbol{a}_{i_k}^{(k)} \in \mathbb{R}^{s_k}$ 是因子矩阵 $\boldsymbol{A}^{(k)}$ 的第 i_k 行， $f\colon \mathbb{R}^{s_1} \times \cdots \times \mathbb{R}^{s_m} \to \mathbb{R}$ 是定义为 $f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m) = \mathcal{C} \times_1 \boldsymbol{x}_1^T \times \cdots \times_m \boldsymbol{x}_m^T$ 的解析函数。

示例 3（广义线性模型）。设 $\mathcal{Y}$ 是来自逻辑模型（Wang和Li，2020）的二值张量，其均值为 Θ（i_1，...，i_m）$= f$（$\mathcal{Z}$（i_1，...，i_m））， 其中 $\mathcal{Z}$ 是一个潜在的低秩张量（关于 CP秩或 Tucker 秩）， 而 f 是定义为 f（x）$= 1/$（$1 + e^{-x}$）的逻辑链接函数。注意，信号张量 Θ本身通常是高秩的。由于解析函数的复合仍然是解析的，根据示例 1和 2，我们可以得出结论：逻辑模型包含在我们的潜变量张量模型中。对于具有一个（已知的）解析链接函数的一般指数族模型，同样的结论也成立（Hong等，2020年）。

示例 4（带解析函数的单指标模型）。单指标模型是一种灵活的半参数模型，由经济学（Robinson，1988）和高维统计学（Balabdaoui等，2019；Ganti等，2017）提出。单指标模型假设存在未知的链接函数 $f\colon \mathbb{R} \to \mathbb{R}$，使得信号张量 $\Theta(i_1, \ldots, i_m) = f(\mathcal{Z}(i_1, \ldots, i_m))$，其中 $\mathcal{Z}$ 是一个潜在的低秩张量。假设链接函数是 M-解析的。我们看到，单指标模型出于与示例 3 中相同的原因，属于潜变量张量模型。

示例 5（简单超图函数模型）。 图函数是一个可测函数，表示一系列可交换随机图（矩阵）的极限（Klopp 等，2017；Gao 等，2015年；Chan 和 Airoldi， 2014年）。类似地，超图子（Zhao， 2015年； Lov´asz， 2012）是 m-一致超图的极限函数，其边可以连接 m-个顶点，并具有 $m \geq 3$。简单超图子模型（Balasubramanian， 2021年）是超图子的一种特殊情形，考虑m-元潜函数。具体而言，简单超图子模型假设信号张量 Θ 具有如下形式

$$\Theta(i_1, \ldots, i_m) = f(a_{i_1}, \ldots, a_{i_m}) \text{ for all } i_1 < \ldots < i_m,$$

其中 $a_i \in [0, 1]$ 是一个标量值潜变量，而 $f\colon [0, 1]^m \to [0, 1]$ 是一个称为简单超图子的函数。假设 f 是解析的，我们可以看到，简单超图子模型属于我们的潜变量张量模型，其潜维度为 $(1, \ldots, 1)$。

2.2 第一个主要结果：潜变量张量具有对数秩 k

本节研究高秩潜变量张量的近似理论。一个值得注意的结果是，我们可以用一个对数秩张量来近似潜变量张量，使得逐项误差很小。

定理 1（与维度无关的一致逼近）。 对于每个张量 $\Theta \in \mathcal{P}(\boldsymbol{d},\ \boldsymbol{s},\ M)$ 和每个整数 $r \in \mathbb{R}_+$，存在一个 $(r,\ \ldots,\ r)$-秩张量 $\mathcal{X}$，使得

$$\|\Theta - \mathcal{X}\|_\infty \leq e^{-c(\boldsymbol{s},M)r^{1/\bar{s}}}, \tag{5}$$

其中 $c(\boldsymbol{s},\ M) > 0$ 是一个与张量维度无关的常数。

定理 1 表明，逐元素错误率相对于逼近秩呈指数级快速衰减。逐元素错误率在现代数据应用中非常有用。例如，人们通常希望在尽可能少地扰动每个元素的情况下逼近一个规模极其庞大的张量；这正是（5）中的无穷范数所刻画的内容。

基于定理 1，我们得到以下命题。

命题 1（潜变量张量具有对数秩）。 固定任意一个 $\varepsilon > 0$，并考虑 $\Theta \in \mathcal{P}(\boldsymbol{d},\ \boldsymbol{s},\ M)$。当 $d \to \infty$ 时，我们有

$$\min\{r \in \mathbb{N}_+ \colon \text{Tucker-rank}(\mathcal{X}) \leq (r, \ldots, r),\ \|\Theta - \mathcal{X}\|_F \leq \varepsilon\} \lesssim \log^{\bar{s}}(\bar{d}),$$

其中，asy 渐近概念 $\lesssim$ 吸收了不依赖于张量维度的常数 .

命题 1 表明，对数秩张量足以将潜变量张量近似到较小误差。该结果解释了尽管高秩张量普遍存在，基于低秩的方法为何仍能取得经验上的成功。尽管现实世界中生成的张量通常具有高秩，但其中许多都可以用对数秩张量很好地近似。

备注1（与先前关于矩阵的研究比较）。 定理 1 是通过非参数工具、张量代数和泛函分析的巧妙结合建立的。（4）中的解析函数是将一元解析函数推广到多元情形的一种形式。以往的研究（Udell 和Townsend， 2019； Xu， 2018年）引入了多元解析函数的另一种定义：如果存在一个常数 f，使得 $M > 0$ 满足

$$\sup_{\boldsymbol{x}_k \in B(\mathbb{R}^{s_1}, \|\cdot\|_\infty), k \in [m]} \left| \frac{\partial^{|\boldsymbol{\alpha}|} f(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m)}{\partial^{\boldsymbol{\alpha}} \boldsymbol{x}_1} \right| \leq M^{|\boldsymbol{\alpha}|} \boldsymbol{\alpha}!, \tag{6}$$

对于所有多重指标 $\boldsymbol{\alpha} \in \mathbb{N}^m$。一个关键区别是，函数（6）仅在第一个模态上是部分解析的，而我们的函数（4）在所有 m 模态上都是联合解析的。以往的工作（Udell 和Townsend，2019；Xu，2018年）利用（6）表明，潜变量矩阵（即我们模型中的特殊情形$m = 2$ ）可以用对数秩矩阵很好地近似。

然而，将以往的技术（Udell and Townsend，2019年；Xu，2018年）直接应用于张量，结果证明是不可能的。事实上，我们证明，（6）中的解析函数只能保证第一个模态上的对数秩，而其他模态上的秩可能是满秩。对于矩阵而言，第一个模态上的对数秩已经足够，因为列秩和行秩相同。然而，对于高阶张量，某一模态上的对数秩并不能保证其他模态也具有对数秩。我们提出联合解析函数（4），以确保所有模态上的对数秩，如命题 1所述。与矩阵相比，这凸显了高阶张量所面临的挑战。

3 通过最小二乘估计实现统计最优速率

在本节中，我们首先提出最小二乘估计量（LSE），并推导其收敛速率。随后，我们给出估计问题相匹配的下界。我们关注 $\boldsymbol{d} \to \infty$ 下的渐近区间，同时将模型配置（$\boldsymbol{s}, M$）视为固定。

我们针对 Θ 提出一种秩约束最小二乘估计量，

$$\hat{\Theta}^{\mathrm{LSE}} = \underset{\text{Tucker-rank}(\Theta)\le(r_1,\ldots,r_k)}{\arg\min} \|\mathcal{Y} - \Theta\|_F. \tag{7}$$

最小二乘估计量 $\hat{\Theta}^{\mathrm{LSE}}$ 取决于约束中的秩（$r_1, \ldots, r_k$）。我们选择命题1 所建议的对数秩。我们的定理 2建立了最小二乘估计量的收敛速率。

定理 2 (统计上界)。考虑一个数据张量 $\mathcal{Y}$，它由（2）生成，其信号张量为 $\Theta \in \mathcal{P}$（$\boldsymbol{d}$, $r = r_1 = \cdots = r_m = c\ \boldsymbol{s}\ M^{-\bar{s}}\ ^{\bar{s}}(d_*/\bar{d})$ $k \in [m]$ $c\ \boldsymbol{s}\ M$ ^{1}, M）。令 $\hat{\Theta}^{\mathrm{LSE}}$ 表示式（7）中的最小二乘估计量，其中（,）对所有 取对数，且（,）是定理 中的常数。那么，以高概率有

$$\|\hat{\Theta}^{\mathrm{LSE}} - \Theta\|_F^2 \lesssim \underbrace{r^m + r\bar{d}}_{\text{low-rank estimation error}} + \underbrace{d_* \exp(-c(\boldsymbol{s}, M) r^{1/s})}_{\text{high-rank approximation error}} = \tilde{O}(\bar{d}). \tag{8}$$

我们讨论结果（8）的含义。LSE 的速率包含两种误差来源：估计误差和近似误差。前者源于观测模型中的噪声，后者则源于信号的高秩结构。这两项体现了非参数问题中的偏差—方差权衡。选择较高的拟合秩会降低近似误差，但由于需要估计更多参数，会增加估计误差。我们的对数秩近似理论实现了这两者之间的权衡。

对于经典的低秩模型，我们的定理 2 通过放宽拟合秩的选择，改进了以往的工作。以往的工作要求拟合秩 $(r_1, \ldots, r_k)$ 必须与真实信号秩 (Zhang 和 Xia，2018年) 完全相同；而我们的结果允许在估计时使用略微过拟合的秩。我们发现，与低秩张量的速率 $\mathcal{O}(d)$ 相比，这种过拟合仅引入可忽略的对数因子（Zhang 和 Xia，2018年）。

我们现在证明，任何估计方法都无法将速率 (8) 提高超过一个对数因子。以下定理给出了该估计问题的下界。

定理 3 (统计下界)。考虑与定理2相同的设定。令 Θ 为基于观测到的数据张量 $\mathcal{Y}$的任意估计量。那么，存在一个绝对常数 $p_0 \in (0, 1)$，使得

$$\inf_{\hat{\Theta}} \sup_{\Theta \in \mathcal{P}(\boldsymbol{d}, \boldsymbol{s}, M)} \mathbb{P}\left(\|\hat{\Theta} - \Theta\|_F^2 \gtrsim \bar{d} \right) \ge p_0. \tag{9}$$

定理 3 保证，不存在收敛速率快于 $\mathcal{O}(d)$ 的估计量。该统计下界通过信息论分析获得。该下界适用于从观测张量 $\mathcal{Y}$ 获得的所有估计量，包括但不限于后续章节中介绍的最小二乘估计量（LSE）和双投影谱估计量（DSE）。(9) 中的下界与 (8) 中的上界相匹配，意味着 (7) 中 LSE 的统计最优性。

4 通过双重投影实现计算最优的谱估计速率

与矩阵不同，对（7）中的秩约束最小二乘问题求解通常是 NP-困难且数值上病态的（Hillar 和 Lim，2013）。一个问题是，能否通过可在多项式时间内计算的估计量达到统计最优速率。本节将回答这个问题。我们首先基于超图植入团问题构造计算下界。然后，我们提出一种多项式时间的双投影谱估计方法，可达到计算最优速率。该结果保证不存在任何多项式时间算法能够达到统计最优速率，从而揭示了阶数为 $m \geq 3$ 的张量存在不可避免的统计-计算差距。

4.1 超图植入团检测

超图植入团检测在构造我们问题的计算下界方面发挥着重要作用。本节将简要介绍超图植入团（HPC）模型和HPC猜想。

考虑一个 m一致超图 $G = (V, E)$，其中 V 是顶点集合， E 是该超图的超边集合。我们定义与超图 G 对应的邻接张量 $\mathcal{A} = \mathcal{A}(G)$如下：

$$\mathcal{A}(i_1, \ldots, i_m) = \begin{cases} 1, & \text{if } (i_1, \ldots, i_m) \in E, \\ 0, & \text{otherwise.} \end{cases}$$

具有 d 个顶点的Erd¨os-R´enyi随机超图记为 $\mathcal{G}_m(d, 1/2)$，它是一个随机的 m一致超图，每条超边被连接的概率均为1/2。团大小为 $\tau > 0$的超图植入团（HPC）记为 $\mathcal{G}_m(d, 1/2, \tau)$，其生成方式如下：首先，从 $G_m(d, 1/2)$生成一个Erd¨os-R´enyi随机超图。然后，以相同的概率从 d 个顶点中独立选取 τ 个顶点。最后，仅保留所有顶点都属于所选τ 个顶点的超边，由此得到一个HPC。

HPC 检测指的是以下假设检验问题，

$$H_0 : G \sim \mathcal{G}_m(d, 1/2) \quad \text{v.s.} \quad H_1 : G \sim \mathcal{G}_m(d, 1/2, \tau). \tag{10}$$

给定一个邻接张量 $\mathcal{A}(G)$以及一个针对 ϕ （10）的检验，检验性能由第一类和第二类错误之和来衡量，即

$$\mathrm{Err}(\phi) = \mathbb{P}_{H_0}(\phi(\mathcal{A}) = 1) + \mathbb{P}_{H_1}(\phi(\mathcal{A}) = 0). \tag{11}$$

HPC检测问题是对葡萄（即 $m = 2$ 在我们的设定中）的经典研究对象——植入团（PC）检测问题的一种推广。对于满足 $m = 2$的图，著名的植入团猜想已被广泛用作统计问题中的计算困难性假设（Berthet 和 Baldin，2020年；Wang 等人，2016年；Cai 和 Wu，2020年）。与 PC 检测的困难性猜想类似，早期工作（Luo 和 Zhang，2022年）提出了 HPC 检测的困难性猜想。下一节中，我们将基于以下困难性猜想构造潜变量张量估计的计算下界。

猜想1（HPC检测猜想（Luo 和 Zhang，2022年））。考虑式（10）中的HPC问题，并假设 $m \geq 2$ 是一个固定整数。如果

$$\limsup_{d\to\infty} \log \tau / \log \sqrt{d} \leq 1-\varepsilon, \text{ for any } \varepsilon > 0.$$

那么，对于任意多项式时间检验序列 $\{\phi\}_d$: $\mathcal{A} \mapsto \{0,1\}$，我们有

$$\liminf_{d\to\infty} \mathrm{Err}(\phi) \geq \frac{1}{2}.$$

4.2 HPC检测猜想下的计算下界

现在我们准备基于猜想来介绍高秩张量估计问题的计算下界1。

定理4 (计算下界)。 假设猜想 1成立。考虑与定理 2 中相同的设定 ˆ。那么，对于每个多项式时间可计算估计量 Θ，当 $d \to \infty$ 时，我们有:

$$\frac{1}{d_*^{1/2-\epsilon}} \sup_{\Theta\in\mathcal{P}(\boldsymbol{d},\boldsymbol{s},M)} \mathbb{E}\|\hat{\Theta}-\Theta\|_F^2 \to \infty, \quad \text{for any } \epsilon > 0.$$

定理 4 给出了计算上可行的估计量的基本极限。该结果表明，不存在收敛速率快于 $\mathcal{O}(d_*^{1/2})$的多项式时间估计量。下一节将展示，我们的双投影谱估计量可以达到这一计算下界，至多相差对数因子。

4.3 通过双投影谱算法得到的计算上界

现在我们提出一种高效的双投影谱算法，可达到定理 4 中的计算下界。因此，我们的双投影谱估计量（DSE）在计算上是最优的。

我们的 DSE 的主要思想是在投影后结合对数秩近似和矩阵谱分解。给定一个近似秩 $(r_1,\ldots,r_k)$，我们对展开张量应用奇异值分解 (SVD)两次。第一次 SVD 作用于展开的观测张量 $\mathcal{Y}$，即，

$$\tilde{\boldsymbol{U}}_k = \mathrm{SVD}_{r_k}\left(\mathrm{Unfold}_k(\mathcal{Y})\right), \quad k \in [m],$$

其中，$\mathrm{SVD}_{r_k}(\cdot)$返回矩阵的前-$r_k$ 个左奇异向量。第二次SVD作用于展开张量，该张量已投影到其他 ($m-1$) 个模态的预估子空间上，即，

$$\hat{\boldsymbol{U}}_k = \mathrm{SVD}_{r_k}\left(\mathrm{Unfold}_k\left(\mathcal{Y}\times_1 \tilde{\boldsymbol{U}}_1^T \times_2 \cdots \times_{k-1} \tilde{\boldsymbol{U}}_{k-1}^T \times_{k+1} \tilde{\boldsymbol{U}}_{k+1}^T \times \cdots \times_m \tilde{\boldsymbol{U}}_m^T\right)\right), \quad k \in [m]. \tag{12}$$

最后，我们通过 Θ估计信号张量

$$\hat{\Theta}^{\mathrm{DSE}} = \mathcal{Y}\times_1 (\hat{\boldsymbol{U}}_1\hat{\boldsymbol{U}}_1^T) \times \cdots \times_m (\hat{\boldsymbol{U}}_m\hat{\boldsymbol{U}}_m^T). \tag{13}$$

DSE 的完整流程总结于算法1中。该估计过程仅涉及矩阵运算，因此可在多项式时间内完成。我们的方法不同于张量高阶奇异值分解（HOSVD）（De Lathauwer 等人，2000年）和高阶正交迭代（HOOI）（Zhang 和

算法1 双投影谱算法

输入： 含噪数据张量 $\mathcal{Y} \in \mathbb{R}^{d_1 \times \cdots \times d_m}$ 和近似秩（r_1，…，r_k）。~1: 对 $k = 1$，…，m.，计算 $\boldsymbol{U}_k = \text{SVD}_{r_k}(\text{Unfold}_k(\mathcal{Y}))$2: **对所有** $k = 1$，…，m **执行**3: 计算奇异子空间估计

$$\hat{\boldsymbol{U}}_k = \text{SVD}_{r_k}\left(\text{Unfold}_k\left(\mathcal{Y} \times_1 \tilde{\boldsymbol{U}}_1^T \times \cdots \times_{k-1} \tilde{\boldsymbol{U}}_{k-1}^T \times_{k+1} \tilde{\boldsymbol{U}}_{k+1}^T \times \cdots \times_m \tilde{\boldsymbol{U}}_m^T\right)\right).$$

4: **结束循环**

5: 通过 $\hat{}\Theta^{\text{DSE}} = \mathcal{Y} \times_1 (\hat{}\boldsymbol{U}_1 \hat{}\boldsymbol{U}_1^T) \times \cdots \times_m (\hat{}\boldsymbol{U}_m \hat{}\boldsymbol{U}_m^T)$ 估计信号张量. **输出：** 估计得到的信号张量 $\hat{}\Theta^{\text{DSE}}$。

~Xia，2018年）。高阶奇异值分解仅将观测张量 $\mathcal{Y}$ 投影一次，使用 $\{\boldsymbol{U}_k\}_{k=1}^m$，见（13）。HOOI（Zhang 和 Xia，2018年）在 $\mathcal{O}(\log p)$ 次迭代中反复进行投影。我们发现，我们的双重投影算法通过两次投影减轻展开时的噪声影响（Han 等人，2022年），改进了高阶奇异值分解，并减少了 HOOI 所需的迭代次数。

现在我们建立 DSE 的统计精度。

定理 5（通过 DSE 算法得到的计算上界）。考虑与ˆ定理2相同的设定。令 Θ 为通过算法1 使用输入张量 $\mathcal{Y}$和近似秩 $r_k = c(\boldsymbol{s}, M)^{-\bar{s}} \log^{\bar{s}}(d_*/\bar{d})$ 对所有 $k \in [m]$得到的估计量。则以高概率有

$$\begin{aligned} \|\hat{\Theta}^{\text{DSE}} - \Theta\|_F^2 &\lesssim d_*^{1/2} \log^{\bar{s}}\left(\frac{d_*}{\bar{d}}\right) + \bar{d} \log^{2\bar{s}}\left(\frac{d_*}{\bar{d}}\right), \\ &= \tilde{\mathcal{O}}(d_*^{1/2} \vee \bar{d}). \end{aligned} \tag{14}$$

我们看到，（14）中的上界与定理 4中的计算下界Θ^{DSE} 仅相差对数因子。因此，我们的估计量 ˆ

备注2（统计与计算最优性之间的差距）。定理 4-5 表明，估计信号张量 $\Theta \in \mathcal{P}(\boldsymbol{d}, \boldsymbol{s}, M)$ 的计算最优速率为 $\tilde{\mathcal{O}}(d_*^{1/2})$。~ ¯定理 2-3 表明，在包括计算上难以处理的所有估计方法中，统计最优速率为 $\mathcal{O}(d)$，为便于比较这两种速率，我们将等维考虑为 $d_1 = \cdots = d_m = d$。图 1 展示了相对于均方误差的统计-计算差距

$$\text{MSE}(\hat{\Theta}, \Theta) = \frac{1}{\prod_{k \in [m]} d_k} \|\hat{\Theta} - \Theta\|_F^2.$$

我们看到，统计最优均方误差为 $\tilde{\mathcal{O}}(d^{-(m-1)})$，而计算最优均方误差为 $\tilde{\mathcal{O}}(d^{-m/2})$。两者之间的差距被称为“统计上可能但计算上不可能”区域。我们的两个估计量 LSE 和 DSE分别位于两条边界上。随着张量阶数 m 增加，统计-计算差距变得更加明显，而对于具有 $m = 2$ 的矩阵则不存在差距。这一现象揭示了高阶张量所面临的根本性挑战。

备注3（双重投影程序的必要性）。对于矩阵（即 $m = 2$），第二个˜ ˆ投影（12）变得多余，因为 $\boldsymbol{U}_k$ 和 $\boldsymbol{U}_k$ 对于 $k = 1$而言是相同的，2。将观测矩阵 $\mathcal{Y} \in \mathbb{R}^{d_1 \times d_2}$ 投影到最佳秩-r 列空间上，等于投影到最佳秩-r行空间上。步骤（13）归结为截断-SVD 估计量，其准确性由艾克特-杨定理（Eckart 和 Young，1936）保证。

然而，对于高阶张量，投影（12）到一个模态的子空间上会改变其他模态的子空间。事实上，在所有模态上联合计算最佳秩-r 子空间 $\mathcal{Y}$ 是 NP-困难的（Hillar 和 Lim，2013）。我们必须以更复杂的方式控制子空间估计中的噪声聚合，并且需要将 $\mathcal{Y}$ 额外投影到预估子空间上（Han 等人，2022年）。我们引入双重投影程序来克服这些挑战。将 $\mathcal{Y}$ 额外投影到预估子空间上，可以大幅降低 ˆ在 $\boldsymbol{U}_k$中的噪声影响。我们在第5节验证了双重投影相较于传统高阶奇异值分解的优势。

备注 4（超参数调优）。我们的算法 1 将近似秩作为输入。秩的理论选择见定理 5。在实践中，由于模型配置未知，我们需要搜索近似秩作为调优参数。基于我们的¯模拟，我们发现将 $r_k = r = c\log\, d$设置为对所有 $k \in [m]$，并通过交叉验证选择常数 c就足够了。我们在第5节研究了调优参数在合成数据和真实数据应用中的实际影响。

5 数值分析

我们研究高秩张量估计中的统计与计算速率。 通过模拟和真实数据应用评估我们方法的性能 .

5.1 模拟

我们首先使用合成数据验证理论结果。我们基于潜变量张量模型模拟三阶张量（3）。我们抽取相互独立的潜变量 $\boldsymbol{a}_i \sim$ 均匀分布$[0,1]^s$ 对于$i \in [d]$，并生成信号张量 $\Theta \in \mathbb{R}^{d\times d\times d}$ 基于

$$\Theta(i_1, i_2, i_3) = f(\boldsymbol{a}_{i_1}, \boldsymbol{a}_{i_2}, \boldsymbol{a}_{i_3}) \text{ for all } (i_1, i_2, i_3) \in [d]^3, \tag{15}$$

其中 f: $\mathbb{R}^s \times \mathbb{R}^s \times \mathbb{R}^s \to \mathbb{R}$ 是一个解析函数。我们考虑表 2中列出的三个模拟模型。这些函数涉及多项式、对数、指数和余弦等运算的复合。值得注意的是，由这些解析函数生成的信号张量均为满秩。

表 2：式（15）中的解析函数，用于生成信号张量。

模型 ID	$f(\boldsymbol{x},\ \boldsymbol{y},\ \boldsymbol{z})$
1	$\exp\left(-(\|\boldsymbol{x}-\boldsymbol{y}\|_2^2 + \|\boldsymbol{y}-\boldsymbol{z}\|_2^2 + \|\boldsymbol{z}-\boldsymbol{x}\|_2^2)/3\right)$
2	$\cos\left((\|\boldsymbol{x}-\boldsymbol{y}\|_2^2 + \|\boldsymbol{y}-\boldsymbol{z}\|_2^2 + \|\boldsymbol{z}-\boldsymbol{x}\|_2^2)/3\right)$
3	$\log\left(1 + (\|\boldsymbol{x}-\boldsymbol{y}\|_2^2 + \|\boldsymbol{y}-\boldsymbol{z}\|_2^2 + \|\boldsymbol{z}-\boldsymbol{x}\|_2^2)/3\right)$

第一个实验按照我们的命题 15 中的建议，考察由（15）生成的信号张量的数值对数秩。我们的理论表明，由潜变量生成的张量

张量模型（3）可以用对数秩张量很好地近似。我们将张量 ϵ的数值 Θ 秩定义为

$$\text{rank}_\epsilon(\Theta) = \min\{r \in \mathbb{N}_+ \colon \text{rank}(\mathcal{X}) \leq (r, r, r), \|\Theta - \mathcal{X}\|_F \leq \epsilon\|\Theta\|_F\}.$$

计算精确的 ϵ秩对于 $0 < \epsilon < 1$ 是 NP 困难的。近似秩的一种常见方法是$\epsilon(\Theta)$ 通过 HOOI 算法将其投影到 Θ 上的 r秩张量（Zhang 和 Xia，2018年），并检查近似误差。我们将 ϵ秩设置为使相对近似误差低于 ϵ的最小秩。我们改变潜变量的维度 $s \in \{5, 10, 15\}$ 和张量维度 $d \in \{20, 40, \ldots, 200\}$，其中 $\epsilon = 0.01$.

图 2 绘制了三种模拟模型下数值秩与张量维度的关系。我们发现，数值秩似乎按张量维度的对数阶增长 d。这一趋势验证了我们的高秩张量对数秩近似理论。此外，我们观察到，随着潜维度 s <style id='11}>增加，曲线整体上移。这一现象与命题 1一致，该命题指出 ε-秩的上界为 $\mathcal{O}$(对数^s (d))。直观而言，更高的潜维度意味着更高的模型复杂度，因此需要更高的近似秩。

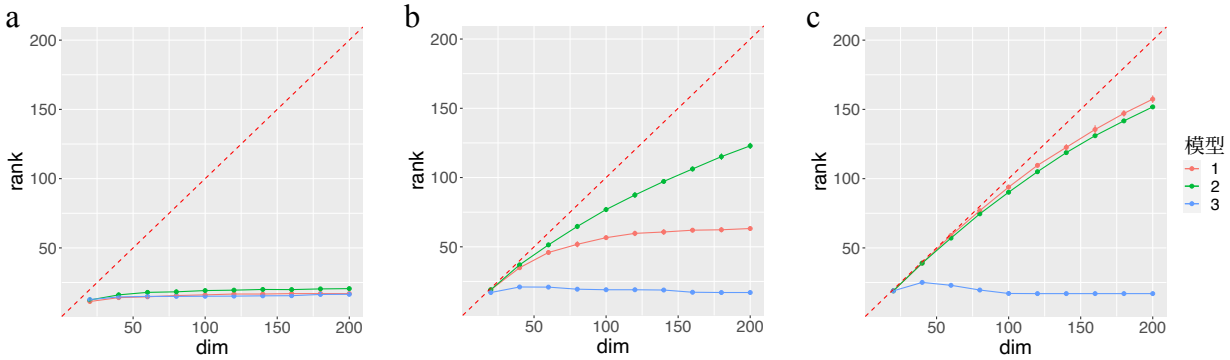


图 2：的数值秩 Θ与张量维度 d 在表 2中的三个模型下的关系。面板 a-c 分别对应潜维度 $s \in \{5, 10, 15\}$。

第二个实验研究了我们的 DSE 算法在信号估计中的性能。我们生成带噪声的张量 $\mathcal{Y} = \Theta + \mathcal{E}$ 来自（15），其中 Θ 是信号张量，而 $\mathcal{E}$ 是噪声张量，其条目独立同分布于 N（0，1）。我们使用输入张量 $\mathcal{Y}$ 和近似秩 r的DSE 算法来估计信号张量。性能使用均方误差进行评估。我们将近似秩设为 $r = c\log^s d$，基于定理 5，其中我们选择 c 以获得最佳结果。

图 3 展示了在潜维度为 $s \in \{1, 2, 3\}$. 时，均方误差随张量维度的变化情况。我们看到，DSE 算法在所有场景中都能成功估计信号张量。均方误差随张量维度 d 呈多项式衰减趋势。这一结果再次与我们的定理 5 一致。此外，我们发现，对于不同的s，衰减趋势相似，这再次表明我们的算法对潜维度具有稳健性。

最后，我们将 DSE 算法与另外两种广受欢迎的张量估计算法进行比较：高阶奇异值分解（De Lathauwer 等人，2000年）和波达计数法（Lee 和 Wang，2021b）。高阶奇异值分解采用单重投影，而我们的 DSE 算法采用双重投影程序。波达计数法假设潜函数满足单调性，并使用排序

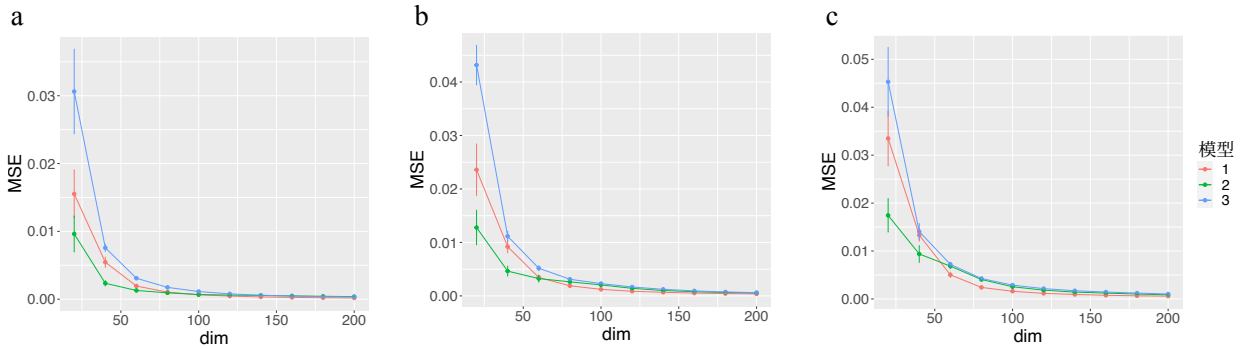


图 3：均方误差与张量维度的关系 d，对应表 2 中的三个模型。对于每种设置，我们绘制 20 次重复实验的平均均方误差及其标准误。面板 a–c 分别对应潜维度 $s \in \{1, 2, 3\}$。

和平滑进行估计。图4 展示了高阶奇异值分解、波达计数法和 DSE 算法的均方误差。我们固定 $s = 2$ 并改变 $d \in \{20, 40, ..., 200\}$。可以看到，在大多数情形下，我们的 DSE 算法表现最佳，而波达计数法往往表现最差。波达计数法表现不佳，可能是因为潜函数不满足单调性。我们还发现，双重投影始终优于单重投影，具体表现为 DSE 的性能始终优于高阶奇异值分解。这一结果验证了我们的直觉：双重投影能够减轻高阶张量中的噪声影响。

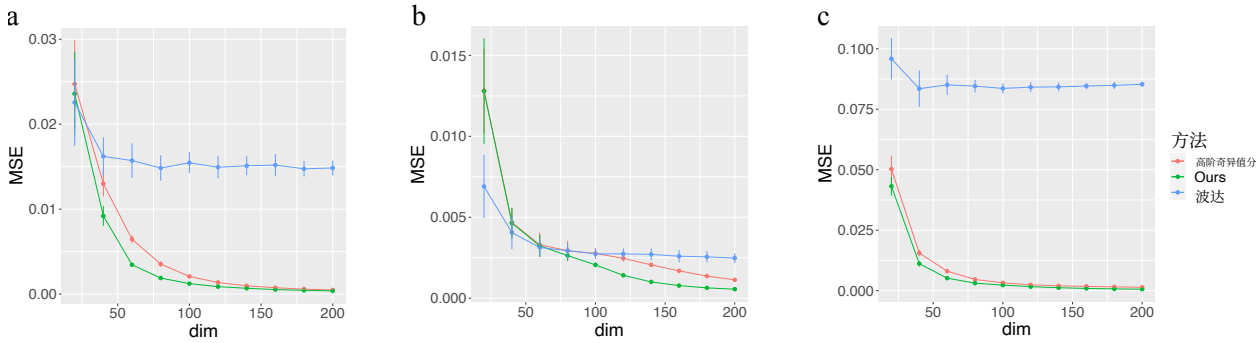


图 4：均方误差与张量维度的关系 d，对应表 2 中的不同模型及估计算法（高阶奇异值分解、波达计数法和我们的方法）。面板 a–c 分别对应表 2中的模拟模型 1–3。

5.2 功能性磁共振成像 3D 脑图像的应用

我们将 DSE 算法应用于视觉运动功能性磁共振成像数据库（B¨uchel 和 Friston，1997）。功能性磁共振成像脑图像由体素组成，相当于沿长度、宽度和高度分布的三维像素。观测到的三维脑图像张量 Θ 大小为 $157 \times 189 \times 68$ 且为满秩。图 5（a）绘出了张量脑图像的一个水平切片，图 6（a）则通过垂直线和水平线展示原始三维图像，以便更好地观察。

在本研究中，我们向原始张量的每个元素添加独立同分布的高斯噪声 $N(0, \sigma_\gamma^2)$，其中

$$\sigma_\gamma = \gamma \left(\frac{\|\Theta\|_F^2}{157 \times 189 \times 68} \right)^{1/2}, \quad \text{with noise level } \gamma \in \{0, 0.2, 0.4, \ldots, 0.8, 1\}.$$

请注意，噪声水平 $\gamma = 0$ 表示不含噪声的原始信号张量。我们通过改变近似秩 $r \in \{3, 6, \ldots, 60\}$，对受污染的图像张量应用 DSE 算法。图 5(b) 展示了不同噪声水平 γ下均方误差随近似秩的变化情况。正如定理 2所预期的，在没有噪声时，设置更高的秩会得到更小的近似误差。然而，在存在噪声时，设置更高的秩会导致次优的偏差—方差权衡。我们发现，最佳性能通常出现在秩为 20–40 时，且噪声越低，对应的秩越高。

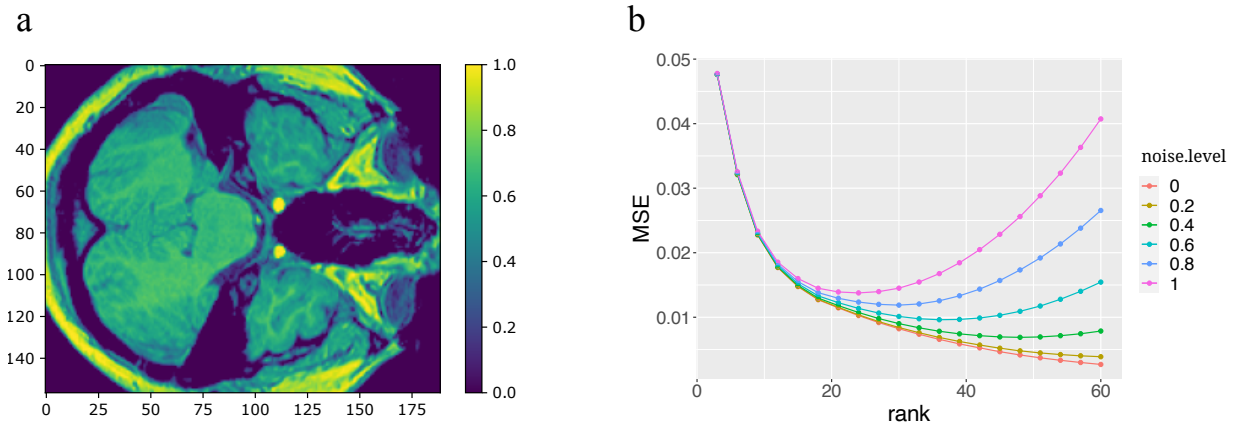
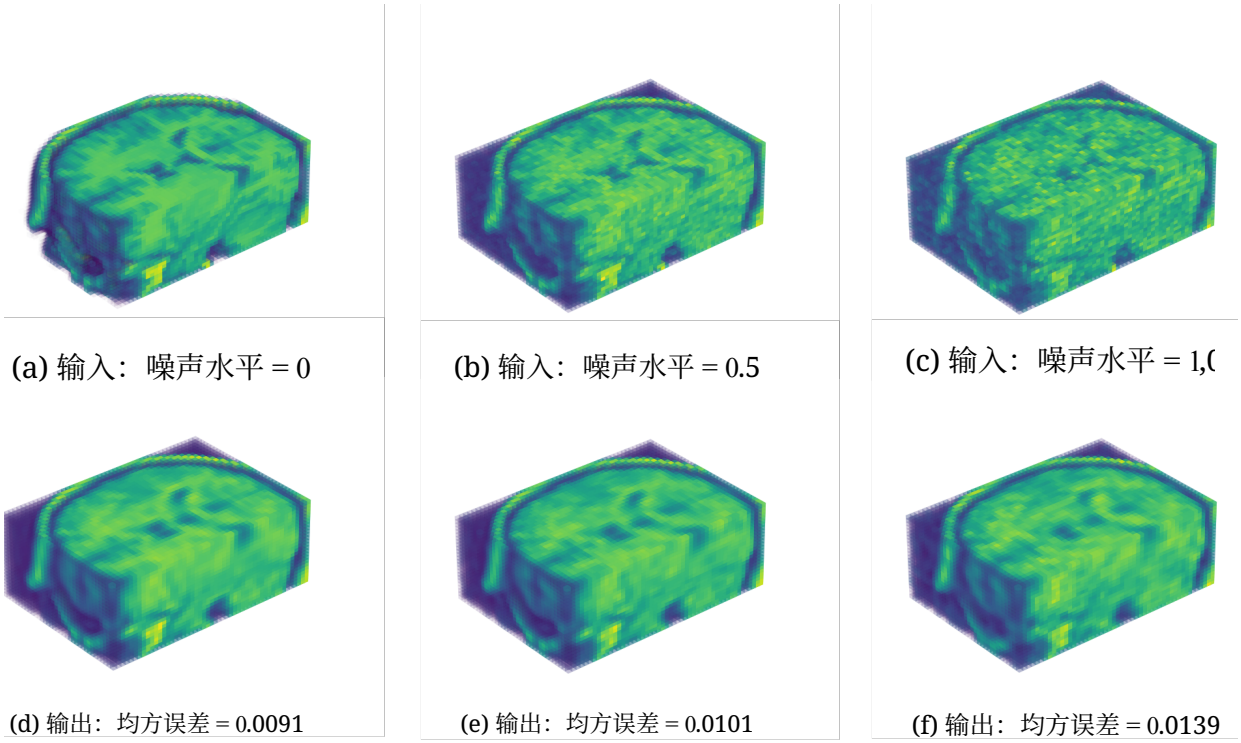


图 5：（a）：张量脑图像 $\{\Theta(i, j, 5) : (i, j) \in [157] \times [189]\}$ 的一个水平切片。（b）：DSE 算法在不同秩近似和噪声水平下的均方误差。

图 6 展示了噪声水平为 $\gamma \in \{0, 0.5, 1\}$ 的输入脑图像数据张量，以及我们算法在近似秩为 $r = 27$ 时得到的相应去噪张量。可以看到，DSE 算法在所有情形下都成功恢复了原始脑图像。

5.3 在作物产量数据中的应用

作为第二个示例，我们将 DSE 算法应用于全球农作物产量数据（数据可从 http://www.fao.org/faostat/en/#data/QC获取）。该数据集包含 1961 年至 2020年 166 个国家的 161 种农作物年度产量。观测数据可整理为一个三阶张量，其元素表示来自 161 种农作物、166 个国家和 60 年的产量对数计数。我们通过交叉验证选择近似秩 $r = 26$，并利用 DSE 算法获得去噪后的张量。我们基于估计的信号张量，在国家和农作物两个模态上使用 K-均值进行聚类分析（具体步骤见附录）。图7 显示，我们的聚类结果成功捕捉到了区域划分。我们发现，簇 1 代表西欧，而簇 2 由地中海周边国家组成。簇 3 主要包含非洲沙漠附近的国家，簇 4 则包括亚洲东部和西部的国家。簇 5 代表东南亚和东非地区，而簇 6 主要由太平洋周边国家组成。值得强调的是，我们的聚类结果与各国的实际地理位置一致，且未将经度或纬度等地理信息作为输入。我们还对农作物模态进行了类似的聚类分析，发现聚类结果能够反映农作物的类型。具体步骤和结果见附录。同一簇内国家和农作物之间呈现出的相似性，说明了我们方法的适用性。



(a) 输入：噪声水平 = 0　(b) 输入：噪声水平 = 0.5　(c) 输入：噪声水平 = 1,(

(d) 输出：均方误差 = 0.0091　(e) 输出：均方误差 = 0.0101　(f) 输出：均方误差 = 0.0139

图 6：面板 (a)-(c) 展示了具有三个噪声水平 $\gamma \in \{0, 0.5, 1\}$ 的输入张量。面板 (d)-(f) 分别展示了对输入张量 (a)-(c) 去噪后得到的张量。

6 证明

6.1 定理1 与命题1的证明

定理*1*的证明。令$\mathcal{B}_{N,\ \ell-1}$ 为引理1中的基张量。我们证明，取

$$N = \lceil M^2 \exp(2m\bar{s}) \rceil, \quad \text{and} \quad \ell = r^{1/\bar{s}} c^{-1/\bar{s}} N^{-m}, \tag{16}$$

其中 $\lceil \cdot \rceil$ 表示向上取整函数。为验证这一点，令 $\ell = r^{1/\bar{s}} c^{-1/\bar{s}} N^{-m}$ ，可得秩的上界，

$$\text{Tucker-rank}(\mathcal{B}_{N,\ell-1}) \le cN^{m\bar{s}}(\ell^{s_1}, \ldots, \ell^{s_k}) \le (r, \ldots, r).$$

接下来，我们考虑近似误差。根据引理 1，近似误差有如下界：

$$\begin{aligned} \|\Theta - \mathcal{B}_{N,\ell-1}\|_\infty &\lesssim \left(\frac{M}{N}\right)^{\ell} \ell^{m\bar{s}} \\ &\overset{(a)}{\le} \exp\left(-\ell(\log N - \log M - m\bar{s})\right) \\ &\overset{(b)}{\le} \exp\left(-\ell(\log M + m\bar{s})\right) \\ &\overset{(c)}{=} \exp(-r^{1/\bar{s}} c'(\log M + m\bar{s})) \end{aligned}$$

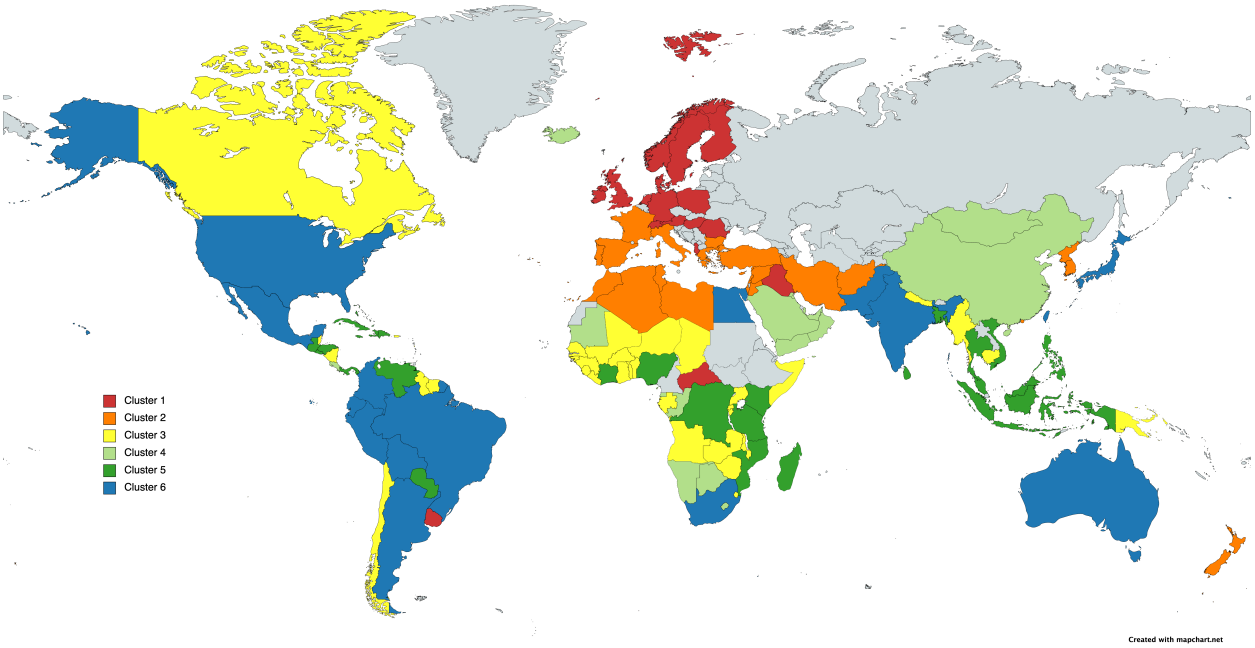


图7：根据我们在作物产量数据分析中使用 DSE 算法得到的估计信号张量，将166个国家分为六个集群。

$$= \exp(-r^{1/\bar{s}} c'')$$

其中，等式 (a) 用到了 log $\ell \le \ell$，等式 (b) 用到了 log $N \ge 2(\log M + m\bar{s})$ 这一事实，等式 (c) 将 $\ell = r^{1/\bar{s}} c'$ 代入 (16)，并使用常数 $c' = c'(\boldsymbol{s},\ M)$，最后一行则将项 $(\log M + m\bar{s})$ 吸收到另一个常数 c''中。

最后，在近似中令 $\mathcal{X} = \mathcal{B}_{N,\ \ell-1}$，即可完成证明。 □

命题*1*的证明。由定理1，可得

$$\|\Theta - \mathcal{X}\|_F^2 \le d_* \|\Theta - \mathcal{X}\|_\infty^2 \le \bar{d}^m e^{-cr^{1/\bar{s}}}.$$

解不等式 $\bar{d}^m e^{-cr^{1/\bar{s}}} \le \varepsilon^2$ 可得 $r \ge c^{-\bar{s}}(m \log^{\bar{s}} \bar{d} - 2\log^{\bar{s}} \varepsilon)$。将所有常数吸收到 $c = c(\varepsilon,\ M,\ \boldsymbol{s})$ 中，我们得到

$$\min\{r \in \mathbb{N}_+ : \text{Tucker-rank}(\mathcal{X}) \le (r, \ldots, r), \|\Theta - \mathcal{X}\|_F \le \varepsilon\} \le c \log^{\bar{s}} \bar{d}.$$

□

引理1（潜变量张量模型的无维度近似）。考虑一个张量$\Theta \in \mathcal{P}(\boldsymbol{d},\ \boldsymbol{s},\ M)$。存在一组基张量 $\{\mathcal{B}_{N,\ell} \in \mathcal{P}(\boldsymbol{d},\ \boldsymbol{s},\ M) : (N,\ \ell) \in \mathbb{N}_+ \times \mathbb{N}_+\}$ 以及一个常数 $c = c(m,\ \boldsymbol{s},\ M)$，使得对于每一对整数 $(N,\ \ell)$，均有

(i) [近似错误率] $\|\Theta - \mathcal{B}_{N,\ \ell-1}\|_\infty \le cM^\ell N^{-\ell} \ell^{m\bar{s}}$；

(ii) [低秩性] Tucker 秩$(\mathcal{B}_{N,\ \ell-1}) \le cN^m \qquad \bar{s}(\ell^{s_1}, \ldots, \ell^{s_k})$.

引理 1。我们首先构造具有较小近似误差的张量$\mathcal{B}_{N,\ \ell-1}$然后证明$\mathcal{B}_{N的\ell-1}$低秩性。为便于表述，我们将在证明中使用以下简写记号。对于一个 m维向量 $\boldsymbol{x}=(x_1,\ \ldots\ldots,\ x_m)$以及一个整数分拆 $\alpha=\alpha_1+\cdots+\alpha_m$满足 $\alpha_i\in\mathbb{N}$，我们记

$$\boldsymbol{x}^{\alpha}:=x_1^{\alpha_1}\cdots x_m^{\alpha_m},\quad \text{and}\quad \partial\boldsymbol{x}^{\alpha}:=\partial x_1^{\alpha_1}\cdots\partial x_m^{\alpha_m},\tag{17}$$

近似误差不失一般性，假设对于所有 $k\in[m]$，潜变量 $\boldsymbol{a}_{i_k}^{(k)}\in[-1,1)^{s_k}$。我们将区间 $[-1,1)$ 划分为 $(2N)$ 个等长区间；即，

$$[-1,1)=\bigcup_{i=-N+1}^{N}\left[\frac{i-1}{N},\frac{i}{N}\right).$$

类似地，对于给定的 $k\in[m]$，我们将 $[-1,1)^{s_k}$ 划分为 s_k 个等长区间的笛卡尔积；即，

$$[-1,1)^{s_k}=\bigcup_{i_{s_k}=-N+1}^{N}\cdots\bigcup_{i_1=-N+1}^{N}\left(\bigtimes_{j=1}^{s_k}\left[\frac{i_j-1}{N},\frac{i_j}{N}\right)\right),$$

其中 $\times$ 表示区间的笛卡尔积。由于潜变量的定义域为$\bigtimes_{k=1}^{m}[-1,1)^{s_k}$，因此划分后总共有 $N_{\mathrm{tot}}:=(2N)^{\sum_{k=1}^{m}s_k}$ 个超立方体。我们将这些超立方体记为 $I_1,\ \ldots\ldots,\ I_{N_{\mathrm{tot}}}$；即，

$$\bigtimes_{k=1}^{m}[-1,1)^{s_k}=\bigcup_{i=1}^{N_{\mathrm{tot}}}I_i,\tag{18}$$

并将这些超立方体的中心记为

$$\underbrace{(\boldsymbol{z}_1^{(1)},\ldots,\boldsymbol{z}_1^{(m)})}_{\in\mathbb{R}^{s_1\times\cdots\times s_m}},\ldots,\underbrace{(\boldsymbol{z}_{N_{\mathrm{tot}}}^{(1)},\ldots,\boldsymbol{z}_{N_{\mathrm{tot}}}^{(m)})}_{\in\mathbb{R}^{s_1\times\cdots\times s_m}}.$$

基于划分（18），我们构造一个次数为 ℓ 的分段多项式，形式如下

$$P_{N,\ell}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_m)=\sum_{i=1}^{N_{\mathrm{tot}}}P_{I_i,\ell}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_m)\mathbb{1}\{(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_m)\in I_i\},\tag{19}$$

其中

$$\begin{aligned}&P_{I_i,\ell}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_m)\\&=\sum_{|\boldsymbol{\alpha}|\le\ell}\frac{1}{\boldsymbol{\alpha}!}\Big((\boldsymbol{x}_1,\ldots,\boldsymbol{x}_m)-(\boldsymbol{z}_i^{(1)},\ldots,\boldsymbol{z}_i^{(m)})\Big)^{\boldsymbol{\alpha}}\frac{\partial^{|\boldsymbol{\alpha}|}f(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_m)}{\partial(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_m)^{\boldsymbol{\alpha}}}\bigg|_{(\boldsymbol{z}_i^{(1)},\ldots,\boldsymbol{z}_i^{(m)})},\end{aligned}$$

并且我们在求和项中使用了简写记号（17）。

根据解析函数的定义（4），对于所有的（$i_1,\ \ldots\ldots,\ i_m$），我们有以下近似误差

$$\sup_{\boldsymbol{a}_{i_k}^{(k)}\in B(k),k\in[m]}|f(\boldsymbol{a}_{i_1}^{(1)},\ldots,\boldsymbol{a}_{i_m}^{(m)})-P_{N,\ell-1}(\boldsymbol{a}_{i_1}^{(1)},\ldots,\boldsymbol{a}_{i_m}^{(m)})|$$

$$
\begin{aligned}
&\le \sup_{1\le i\le N_{\mathrm{tot}}} \sup_{\boldsymbol{a}_{i_k}^{(k)}\in B(k),k\in[m]} |f(\boldsymbol{a}_{i_1}^{(1)},\dots,\boldsymbol{a}_{i_m}^{(m)}) - P_{I_i,\ell-1}(\boldsymbol{a}_{i_1}^{(1)},\dots,\boldsymbol{a}_{i_m}^{(m)})| \\
&\le \sum_{\boldsymbol{\alpha}:\,|\boldsymbol{\alpha}|=\ell} \frac{1}{\boldsymbol{\alpha}!} \sup_{\boldsymbol{x}_k\in B(k),k\in[m]} \left|\frac{\partial^{|\boldsymbol{\alpha}|} f(\boldsymbol{x}_1,\dots,\boldsymbol{x}_m)}{\partial(\boldsymbol{x}_1,\dots,\boldsymbol{x}_m)^{\boldsymbol{\alpha}}}\right| \left(\frac{1}{N}\right)^{\ell} \\
&\le \sum_{\boldsymbol{\alpha}:\,|\boldsymbol{\alpha}|=\ell} \left(\frac{M}{N}\right)^{\ell} \\
&= \left(\frac{M}{N}\right)^{\ell} \binom{\ell+\sum_{k=1}^m s_k}{\ell}.
\end{aligned}
$$

这里为简洁起见，我们将单位球 $B(k) := B(\mathbb{R}^{s_k},\ \|\cdot\|_\infty)$ 记作。

现在，我们定义近似张量 $\mathcal{B}_{N,\ \ell-1} \in \mathcal{P}(\boldsymbol{d},\ \boldsymbol{s},\ M)$，其元素为

$$
\mathcal{B}_{N,\ell-1}(i_1,\dots,i_m) = P_{N,\ell-1}\left(\boldsymbol{a}_{i_1}^{(1)},\dots,\boldsymbol{a}_{i_m}^{(m)}\right),\ \text{for all } (i_1,\dots,i_m)\in[d_1]\times\cdots\times[d_m]. \tag{20}
$$

根据构造，近似误差满足

$$
\|\Theta - \mathcal{B}_{N,\ell-1}\|_\infty \le \left(\frac{M}{N}\right)^{\ell} \binom{\ell+\sum_{k=1}^m s_k}{\ell} \le cM^{\ell}N^{-\ell}\ell^{m\bar{s}},
$$

其中，$c = c(\boldsymbol{s},\ M)$ 表示一个常数。

低秩性 现在我们计算 $\mathcal{B}_{N,\ \ell-1}$ 的 Tucker 秩。只需证明

$$
\operatorname{rank}\left(\operatorname{Unfold}_1(\mathcal{B}_{N,\ell-1})\right) \le (2N)^{\sum_{k=1}^m s_k} \binom{\ell+s_1-2}{s_1-1} \lesssim N^{m\bar{s}}(\ell+s_1)^{s_1}.
$$

给定潜变量 $\boldsymbol{x}_k = (x_{k1},\ \dots,\ x_{ks_k}) \in \mathbb{R}^{s_k}$，我们将由次数至多为 $(\ell-1)$ 的所有单项式构成的基向量定义为

$$
\begin{aligned}
&\phi(\boldsymbol{x}_1,\dots,\boldsymbol{x}_m) \\
&= \{(\boldsymbol{x}_1,\dots,\boldsymbol{x}_m)^{\boldsymbol{\alpha}}\colon |\boldsymbol{\alpha}|\le \ell-1\} \\
&= (1, x_{11},\dots,x_{1s_1}, x_{21},\dots,x_{2s_2},\dots, x_{11}x_{12},\dots, x_{11}^{\ell-2}x_{21},\dots, x_{ms_1}^{\ell-1},\dots,x_{ms_m}^{\ell-1})^T \\
&\in \mathbb{R}^{\binom{\ell+\sum_{k=1}^m s_k-2}{\ell-1}},
\end{aligned}
$$

其中第二行采用约定 (17)。这里，ϕ 的维数是所有次数至多为 $(\{a_{ki_k}\}_{k\in[m],\ i_k\in[s_k]})$ 的$\ell-1$的可能组合数。于是，分段多项式 (19) 可表示为

$$
P_{N,\ell-1}(\boldsymbol{x}_1,\dots,\boldsymbol{x}_m) = \sum_{i=1}^{N_{\mathrm{tot}}} \langle\phi(\boldsymbol{x}_1,\dots,\boldsymbol{x}_m),\boldsymbol{\beta}_i\rangle \mathbb{1}\{(\boldsymbol{x}_1,\dots,\boldsymbol{x}_m)\in I_i\},
$$

其中，$\boldsymbol{\beta}_i \in \mathbb{R}^{\binom{\ell+\sum_{k=1}^m s_k-2}{\ell-1}}$ 是区域 I_i. 中的系数向量

现在我们来求 $\operatorname{Unfold}_1(\mathcal{B}_{N,\ \ell-1})$ 的表达式。我们写作 $I_i = I_{i,\ 1}\times I_{i,\ -1}$，其中 $I_{i,\ 1}\subset\mathbb{R}^{s_1}$ 和 $I_{i,\ -1}\subset\mathbb{R}^{s_2\times\cdots\times s_m}$。于是

$$
\mathbb{1}\left\{(\boldsymbol{x}_1,\dots,\boldsymbol{x}_m)\in I_i\right\} = \mathbb{1}\{\boldsymbol{x}\in I_{i,1}\}\mathbb{1}\{(\boldsymbol{x}_2,\dots,\boldsymbol{x}_m)\in I_{i,-1}\}
$$

根据式（$\mathcal{B}_{N,\ \ell-1}$ 中的定义（20）以及张量与其展开之间的关系，我们有

$$
\begin{aligned}
&\text{展开}_1(\mathcal{B}_{N\,\ell-1}) \\
&= \sum_{1\le i\le N_{\text{tot}}} \underbrace{\begin{pmatrix} \left\langle \phi(\boldsymbol{a}_1^{(1)}, \boldsymbol{a}_1^{(2)} \ldots, \boldsymbol{a}_1^{(m)}), \boldsymbol{\beta}_i \right\rangle, & \ldots, & \left\langle \phi(\boldsymbol{a}_1^{(1)}, \boldsymbol{a}_{d_2}^{(2)} \ldots, \boldsymbol{a}_{d_m}^{(m)}), \boldsymbol{\beta}_i \right\rangle \\ \vdots & & \vdots \\ \left\langle \phi(\boldsymbol{a}_{d_1}^{(1)}, \boldsymbol{a}_1^{(2)} \ldots, \boldsymbol{a}_1^{(m)}), \boldsymbol{\beta}_i \right\rangle, & \ldots, & \left\langle \phi(\boldsymbol{a}_{d_1}^{(1)}, \boldsymbol{a}_{d_2}^{(2)} \ldots, \boldsymbol{a}_{d_m}^{(m)}), \boldsymbol{\beta}_i \right\rangle \end{pmatrix}}_{=:\boldsymbol{M}_i} \circ \qquad (21) \\
&\underbrace{\begin{pmatrix} \mathbb{1}\{\boldsymbol{a}_1^{(1)} \in I_{i,1}\}\mathbb{1}\{(\boldsymbol{a}_1^{(2)}, \ldots, \boldsymbol{a}_1^{(m)}) \in I_{i,-1}\}, & \ldots, & \mathbb{1}\{\boldsymbol{a}_1^{(1)} \in I_{i,1}\}\mathbb{1}\{(\boldsymbol{a}_{d_2}^{(2)}, \ldots, \boldsymbol{a}_{d_m}^{(m)}) \in I_{i,-1}\} \\ \vdots & & \vdots \\ \mathbb{1}\{\boldsymbol{a}_{d_1}^{(1)} \in I_{i,1}\}\mathbb{1}\{(\boldsymbol{a}_1^{(2)}, \ldots, \boldsymbol{a}_1^{(m)}) \in I_{i,-1}\}, & \ldots, & \mathbb{1}\{\boldsymbol{a}_{d_1}^{(1)} \in I_{i,1}\}\mathbb{1}\{(\boldsymbol{a}_{d_2}^{(2)}, \ldots, \boldsymbol{a}_{d_m}^{(m)}) \in I_{i,-1}\} \end{pmatrix}}_{=:\mathbb{1}_i},
\end{aligned}
$$

其中， $\boldsymbol{M}_i$ 和 $\mathbb{1}_i$ 是大小为 $d_1 \times \prod_{k=2}^m d_k$的矩阵， $\circ$ 表示逐元素矩阵乘积。我们声称

$$
\operatorname{rank}(\mathbb{1}_i) = 1 \quad \text{and} \quad \operatorname{rank}(\boldsymbol{M}_i) \le \binom{\ell + s_1 - 2}{s_1 - 1} \text{ for all } i \in [N_{\text{total}}]. \qquad (22)
$$

性质（22）将在证明最后给出。结合（21）和（22）即可得出所需结论，

$$
\operatorname{rank}(\text{Unfold}_1(\mathcal{B}_{N,\ell-1})) \le N_{\text{tot}} \operatorname{rank}(\boldsymbol{M}_i) \le (2N)^{\sum_{k=1}^m s_k} \binom{\ell + s_1 - 2}{s_1 - 1} \lesssim N^{m\bar{s}} (\ell + s_1)^{s_1}.
$$

最后 ， 只需验证（22）。$\operatorname{rank}(\mathbb{1}_i) = 1$ 这一事实来自因式分解

$$
\mathbb{1}_i = \begin{pmatrix} \mathbb{1}\{\boldsymbol{a}_1^{(1)} \in I_{i,1}\} \\ \vdots \\ \mathbb{1}\{\boldsymbol{a}_{d_1}^{(1)} \in I_{i,1}\} \end{pmatrix} \left(\mathbb{1}\{(\boldsymbol{a}_1^{(2)}, \ldots, \boldsymbol{a}_1^{(m)}) \in I_{i,-1}\}, \quad \ldots, \quad \mathbb{1}\{(\boldsymbol{a}_{d_2}^{(2)}, \ldots, \boldsymbol{a}_{d_m}^{(m)}) \in I_{i,-1}\} \right).
$$

现在我们证明 $\operatorname{rank}(\boldsymbol{M}_i) \le \binom{\ell+s_1-2}{s_1-1}$。注意到

$$
\boldsymbol{M}_i = \left(\begin{pmatrix} \phi(\boldsymbol{a}_1^{(1)}, \boldsymbol{a}_1^{(2)} \ldots, \boldsymbol{a}_1^{(m)})^T \\ \vdots \\ \phi(\boldsymbol{a}_{d_1}^{(1)}, \boldsymbol{a}_1^{(2)} \ldots, \boldsymbol{a}_1^{(m)})^T \end{pmatrix} \boldsymbol{\beta}_i, \quad \cdots, \quad \begin{pmatrix} \phi(\boldsymbol{a}_1^{(1)}, \boldsymbol{a}_{d_2}^{(2)} \ldots, \boldsymbol{a}_{d_m}^{(m)})^T \\ \vdots \\ \phi(\boldsymbol{a}_{d_1}^{(1)}, \boldsymbol{a}_{d_2}^{(2)} \ldots, \boldsymbol{a}_{d_m}^{(m)})^T \end{pmatrix} \boldsymbol{\beta}_i \right).
$$

对于任意给定的（i_2, ..., i_m） $\in [d_2] \times \cdots \times [d_m]$， $\boldsymbol{M}$的列满足

$$
\begin{aligned}
&\begin{pmatrix} \phi(\boldsymbol{a}_1^{(1)}, \boldsymbol{a}_{i_2}^{(2)} \ldots, \boldsymbol{a}_{i_m}^{(m)})^T \\ \phi(\boldsymbol{a}_2^{(1)}, \boldsymbol{a}_{i_2}^{(2)} \ldots, \boldsymbol{a}_{i_m}^{(m)})^T \\ \vdots \\ \phi(\boldsymbol{a}_{d_1}^{(1)}, \boldsymbol{a}_{i_2}^{(2)} \ldots, \boldsymbol{a}_{i_m}^{(m)})^T \end{pmatrix} \boldsymbol{\beta}_i \\
&\in \text{Column space of } \underbrace{\left\{ \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} a_{11}^{(1)} \\ a_{21}^{(1)} \\ \vdots \\ a_{d_1 1}^{(1)} \end{pmatrix}, \begin{pmatrix} a_{12}^{(1)} \\ a_{22}^{(1)} \\ \vdots \\ a_{d_1 2}^{(1)} \end{pmatrix}, \ldots, \begin{pmatrix} a_{1s_1}^{(1)} \\ a_{2s_1}^{(1)} \\ \vdots \\ a_{d_1 s_1}^{(1)} \end{pmatrix}, \ldots, \begin{pmatrix} (a_{1s_1}^{(1)})^{(\ell-1)} \\ (a_{2s_1}^{(1)})^{(\ell-1)} \\ \vdots \\ (a_{d_1 s_1}^{(1)})^{(\ell-1)} \end{pmatrix} \right\}}_{=:\text{Basis}},
\end{aligned}
$$

其中，基矩阵是一个矩阵，其第 j 行由集合中的所有单项式组成

$$\{(a_{j1}^{(1)}, a_{j2}^{(1)}, \ldots, a_{js_1}^{(1)})^{\boldsymbol{\alpha}} \colon |\boldsymbol{\alpha}| \leq \ell - 1\}.$$

通过计算由 s_1 个元素构成的集合中次数不超过$(\ell - 1)$的单项式数量，我们有

$$\operatorname{rank}(\boldsymbol{M}_i) \leq \operatorname{rank}(\text{Basis}) \leq \text{number of rows(Basis)} \leq \binom{\ell + s_1 - 2}{s_1 - 1}.$$

因此，性质（22）得到验证。 □

6.2 定理 2

证明。 根据定理 1，对于任意 $\Theta \in \mathcal{P}(\boldsymbol{d},\ \boldsymbol{s},\ M)$，我们总能找到一个张量 $\mathcal{X}_r$，其 Tucker 秩为 $(r,\ \ldots,\ r)$，满足

$$\|\Theta - \mathcal{X}_r\|_F^2 \leq d_* e^{-c(\boldsymbol{s},M) r^{1/\bar{s}}}. \tag{23}$$

由三角不等式可得，

$$\|\hat{\Theta}^{\mathrm{LSE}} - \Theta\|_F \leq \|\hat{\Theta}^{\mathrm{LSE}} - \mathcal{X}_r\|_F + \|\mathcal{X}_r - \Theta\|_F.$$

由于上述方程中的第二项可由（23）很好地界定，因此只需界定第一项。由于$\hat{\Theta}^{\mathrm{LSE}}$ 是平方损失的全局最优解，根据泰勒展开，我们有

$$\begin{aligned}
\|\hat{\Theta}^{\mathrm{LSE}} - \mathcal{X}_r\|_F &\leq \left\langle \frac{\hat{\Theta}^{\mathrm{LSE}} - \mathcal{X}_r}{\|\hat{\Theta}^{\mathrm{LSE}} - \mathcal{X}_r\|_F}, \mathcal{E} + (\Theta - \mathcal{X}_r) \right\rangle \\
&\leq \sup_{\text{Tucker-rank}(\Theta_1), \text{Tucker-rank}(\Theta_2) \leq (r,\ldots,r)} \left\langle \frac{\Theta_1 - \Theta_2}{\|\Theta_1 - \Theta_2\|_F}, \mathcal{E} + (\Theta - \mathcal{X}_r) \right\rangle \\
&\leq \sup_{\text{Tucker-rank}(\Theta_3) \leq 2(r,\ldots,r), \|\Theta_3\|_F \leq 1} \langle \Theta_3, \mathcal{E} \rangle + \|\Theta - \mathcal{X}_r\|_F,
\end{aligned} \tag{24}$$

其中，最后一个不等式是由以下事实得到的：Tucker 秩$(\Theta_1 - \Theta_2) \leq$ Tucker 秩(Θ_1)+Tucker 秩$(\Theta_2) \leq 2(r, \ldots, r)$。因此，将引理 3和（23）应用于（24）可得

$$\|\hat{\Theta}^{\mathrm{LSE}} - \mathcal{X}_r\|_F^2 \lesssim r^m + \bar{d}r + d_* e^{-c(\boldsymbol{s},M) r^{1/\bar{s}}},$$

以高概率 $1 - c\exp(-rd)$。令 $r = c(\boldsymbol{s}, M)^{-\bar{s}} \log^{\bar{s}} \left(\frac{d_*}{d} \right)$，证明即告完成。 □

6.3 定理 3 的证明

我们先引入引理2，然后证明定理3。

引理 2 （Wang 和 Li （2020年）中的引理 8）。 将 Tucker 低秩张量类定义为

$$\mathcal{T}(\boldsymbol{d}, \boldsymbol{r}, M) = \{\Theta \in \mathbb{R}^{d_1 \times \cdots \times d_m} \colon \text{Tucker-rank}(\Theta) \leq (r_1, \ldots, r_m), \|\Theta\|_\infty \leq M\}. \tag{25}$$

对于任意给定的常数 $0 \leq \gamma \leq 1$，存在一个由有限个张量组成的集合 $\mathcal{X} = \{\Theta_i \colon i = 1, \ldots\} \subset \mathcal{T}(\boldsymbol{d}, \boldsymbol{r}, M)$，满足以下四个性质：

(i) Card$(\mathcal{X}) \geq 2^{\bar{r}\bar{d}/8} + 1$，其中 Card(·) 表示该集合的基数。

(ii) $\mathcal{X}$ 包含零张量 $\mathbf{0} \in \mathbb{R}^{d_1 \times \cdots \times d_m}$;
(iii) $\|\Theta\|_\infty \le \gamma \left(M \wedge \sqrt{\frac{\bar{r}\bar{d}}{\underline{d}_*}} \right)$ 对所有元素 $\Theta \in \mathcal{X}$;

(iv) $\|\Theta_i - \Theta_j\|_F \ge \frac{\gamma}{4} \left(M\sqrt{\underline{d}_*} \wedge \sqrt{\bar{r}\bar{d}} \right)$ 对于任意两个不同的元素 $\Theta_i \ne \Theta_j \in \mathcal{X}$.

注 5 （低秩张量估计的极小极大下界）。 考虑高斯观测模型

$$\mathcal{Y} = \Theta + \mathcal{E},$$

其中 $\Theta \in \mathcal{T}(\boldsymbol{d}, \boldsymbol{r}, M)$ 是我们关注的低秩信号张量，而 $\mathcal{E}$ 是由独立同分布的标准高斯随机变量组成的噪声张量。将 **?**、定理 2.5 和引理2直接应用于上述情形可得

$$\inf_{\hat{\Theta}} \sup_{\Theta \in \mathcal{T}(\boldsymbol{d},\boldsymbol{r},M)} \mathbb{P}\left(\|\hat{\Theta} - \Theta\|_F \gtrsim \sqrt{\bar{r}\bar{d}} \right) \ge \inf_{\hat{\Theta}} \sup_{\Theta \in \mathcal{X}} \mathbb{P}\left(\|\hat{\Theta} - \Theta\|_F \gtrsim \sqrt{\bar{r}\bar{d}} \right) \ge p_0,$$

对于某个通用常数 $p_0 > 0$.

定理 3的证明。 对于任意给定的核心张量 $\mathcal{C} \in \mathbb{R}^{s_1 \times \cdots \times s_m}$，我们定义一个解析函数$f_{\mathcal{C}}$：$\mathbb{R}^{s_1} \times \cdots \times \mathbb{R}^{s_m} \to \mathbb{R}$ 如下

$$f_{\mathcal{C}}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m) = \mathcal{C} \times_1 \boldsymbol{a}_1^T \times_2 \cdots \times_m \boldsymbol{a}_m^T \quad \text{for any } \boldsymbol{a}_k \in \mathbb{R}^{s_k}, k \in [m].$$

基于 $f_{\mathcal{C}}$ 的概念，我们可以将式（25）中的低秩张量类改写为

$$\mathcal{T}(\boldsymbol{d}, \boldsymbol{s}, M) = \Big\{ \Theta \in \mathbb{R}^{d_1 \times \cdots \times d_m} : \Theta(i_1, \ldots, i_m) = f_{\mathcal{C}}(\boldsymbol{a}_{i_1}^{(1)}, \ldots, \boldsymbol{a}_{i_m}^{(m)}) \text{ for some } \|\mathcal{C}\|_\infty \le M,$$
$$\boldsymbol{a}_{i_k}^{(k)} \in B(\mathbb{R}^{s_k}, \|\cdot\|_\infty), \text{ and all } i_k \in [d_k], k \in [m] \Big\}.$$

注意，根据定义， $\mathcal{T}(\boldsymbol{d}, \boldsymbol{s}, M) \subset \mathcal{P}(\boldsymbol{d}, \boldsymbol{s}, M)$；参见正文中的例 2。直接应用备注 5即可完成证明，因为

$$\inf_{\hat{\Theta}} \sup_{\Theta \in \mathcal{P}(\boldsymbol{d},\boldsymbol{s},M)} \mathbb{P}\left(\|\hat{\Theta} - \Theta\|_F \gtrsim \sqrt{\bar{s}\bar{d}} \right) \ge \inf_{\hat{\Theta}} \sup_{\Theta \in \mathcal{T}(\boldsymbol{d},\boldsymbol{s},M)} \mathbb{P}\left(\|\hat{\Theta} - \Theta\|_F \gtrsim \sqrt{\bar{s}\bar{d}} \right) \ge p_0,$$

对于某个通用常数 $p_0 > 0$. □

6.4 定理 4的证明

证明。 定理4的证明利用了在高阶张量中检测常数植入结构的结果（Luo 和 Zhang，2022年）。假设真实信号 $\Theta \in \mathbb{R}^{d_1 \times \cdots \times d_m}$ 具有给定 $\boldsymbol{r} = (r_1, \ldots, r_m)$ 的常数高阶聚类（CHC）结构，使得

$$\Theta \in \Theta_{\mathrm{CHC}}(\boldsymbol{d}, \boldsymbol{r}, \lambda) := \{\lambda \mathbb{1}_{I_1} \circ \cdots \circ \mathbb{1}_{I_m} : \mathrm{Card}(I_k) = r_k \text{ for all } k \in [m]\},$$

其中， $I_i \subset [d_i]$ 是指标子集，Card(·)表示集合的基数，$\mathbb{1}_{I_i}$ 是一个 d_i维指示向量，使得当 $j \in I_i$ 时，$(\mathbb{1}_{I_i})_j = 1$，否则为 0。注意， Θ_{CHC} 由 CP 秩为 1 的张量组成，且 $\|\Theta\|_\infty = \lambda$。根据正文中秩为 1 的例 1，对于所有 $\boldsymbol{s}$ 和 λ，我们有 $\Theta_{\mathrm{CHC}}(\boldsymbol{d}, \boldsymbol{r}, \lambda) \subset \mathcal{P}(\boldsymbol{d}, \boldsymbol{s}, \lambda)$。

我们考虑基于观测张量 $\mathcal{Y}$检测 CHC 的假设检验，

$$H_0 : \Theta = 0 \quad \text{v.s.} \quad H_1 : \Theta \in \Theta_{\mathrm{CHC}}(\boldsymbol{d}, \boldsymbol{r}, \lambda). \tag{26}$$

以下命题给出了在猜想 ϕ 下使用计算可行检验对 CHC 进行不可能检测的渐近区间。

命题 2 （(Luo 和 Zhang，2022年) 中的定理15）。 考虑高斯观测模型 (26) 下的 CHC 检测问题 (2)，其中渐近区间 $d \to \infty$ 满足

$$d = d_1 = \cdots = d_m, \quad r = r_1 = \cdots = r_m = d^{\alpha}, \quad \lambda = d^{-\beta},$$

其中 $0 \le \alpha \le 1$ 且 $\beta > (m\alpha - m/2) \vee 0$. 那么，在 HPC检测猜想 1下，对于所有多项式时间检验序列 $\{\phi\}_d$: $\mathcal{Y} \mapsto \{0, 1\}$，有

$$\liminf_{d \to \infty} \mathrm{Err}(\phi) \ge \frac{1}{2},$$

其中 Err(ϕ) 定义见 (11)。

为便于技术处理，我们考虑等维情形

$$d_1 = \cdots = d_m = d, \quad \text{and} \quad r_1 = \ldots = r_m = c_1 d^{\alpha} \quad \text{for a constant } c_1 > 0.$$

根据主论文中猜想 1 和命题 2 中的条件，令 $\alpha = 1/2$ 和$\lambda = c_2 d^{-\epsilon/2}$，其中常数 $c_2 > 0$ 以及任意 $\epsilon > 0$，使得所有多项式时间检验 ϕ 满足 $\liminf_{d \to \infty} \mathrm{Err}(\phi) \ge 1/2$.

ˆ我们通过反证法证明。现在，假设存在一个来自多项式时间算法的假设估计量 Θ，其达到估计误差率 $d_*^{1/2-\epsilon} = d^{m/2-\epsilon}$（为简化记号，令 m 吸收到 ϵ 中）。具体而言，存在常数 $b > 0$ 和 $\varepsilon > 0$, 使得

$$\limsup_{d \to \infty} \frac{1}{d^{m/2-\epsilon}} \sup_{\Theta \in \mathcal{P}(\boldsymbol{d}, \boldsymbol{s}, \lambda)} \mathbb{E} \|\hat{\Theta} - \Theta\|_F^2 \le b. \tag{27}$$

根据马尔可夫不等式，命题（27）表明，当 d 足够大时，对于所有$\Theta \in \mathcal{P}$（$\boldsymbol{d}$， $\boldsymbol{s}$， λ）以及所有 $u \ge 0$,

$$\|\hat{\Theta} - \Theta\|_F \le u d^{\frac{m}{4} - \frac{\epsilon}{2}}, \tag{28}$$

概率至少为 $1 - b/u$。特别地，陈述（28）对所有 $\Theta \in \Theta_{\mathrm{CHC}}$（$\boldsymbol{d}$， $\boldsymbol{r}$， λ） $\subset \mathcal{P}$（$\boldsymbol{d}$， $\boldsymbol{s}$， λ）成立。

考虑 t （26）中的假设检验见正文。然后我们采用以下检验 t

$$\phi(\mathcal{Y}) = \mathbb{1}(\|\hat{\Theta}\|_F \ge u d^{\frac{m}{4} - \frac{\epsilon}{2}}).$$

检验 ϕ 的第一类错误率由以下因素控制。

$$\mathbb{P}_{H_0}(\|\hat{\Theta}\|_F \ge u d^{\frac{m}{4} - \frac{\epsilon}{2}}) = \mathbb{P}_{H_0}(\|\hat{\Theta} - \Theta\|_F \ge u d^{\frac{m}{4} - \frac{\epsilon}{2}}) \le b/u.$$

对于第二类错误率，我们得到

$$\sup_{\Theta \in \Theta_{\mathrm{CHC}}(\boldsymbol{d}, \boldsymbol{r}, \lambda)} \mathbb{P}_{\Theta}(\phi(\mathcal{Y}) = 0) = \sup_{\Theta \in \Theta_{\mathrm{CHC}}(\boldsymbol{d}, \boldsymbol{r}, \lambda)} \mathbb{P}_{\Theta}(\|\hat{\Theta}\|_F < u d^{\frac{m}{4} - \frac{\epsilon}{2}})$$

$$
\begin{aligned}
&\leq \sup_{\Theta\in\Theta_{\mathrm{CHC}}(\boldsymbol{d},\boldsymbol{r},\lambda)} \mathbb{P}_\Theta(\|\hat{\Theta}-\Theta\|_F^2 > \|\Theta\|_F^2 - u^2 d^{\frac{m}{2}-\epsilon}) \\
&\overset{(*)}{\leq} \sup_{\Theta\in\Theta_{\mathrm{CHC}}(\boldsymbol{d},\boldsymbol{r},\lambda)} \mathbb{P}_\Theta(\|\hat{\Theta}-\Theta\|_F^2 > u^2 d^{\frac{m}{2}-\epsilon}) \\
&\overset{(**)}{\leq} b/u.
\end{aligned}
$$

不等式（$*$）成立，因为

$$
\|\Theta\|_F^2 \geq \lambda^2 r^m = c_1^m c_2^2 d^{\frac{m}{2}-\epsilon} \geq 2u^2 d^{\frac{m}{2}-\epsilon},
$$

其中，最后一个不等式在约束 $c_1^m c_2 > 2u^2$下成立。给定某个值 u，我们总能选取这样的常数 c_1 和 c_2。不等式（$**$）成立是因为命题（28）。综合第一类和第二类错误，我们得到

$$
\begin{aligned}
\mathrm{Err}(\phi) &\leq \mathbb{P}_{\mathcal{H}_0}(\phi(\mathcal{Y})=1) + \sup_{\Theta\in\Theta_{\mathrm{CHC}}(\boldsymbol{d},\boldsymbol{r},\lambda)} \mathbb{P}_\Theta(\phi(\mathcal{Y})=0) \\
&\leq 2b/u < 1/2,
\end{aligned} \tag{29}
$$

对于 $u > 4b$，命题（29）与正文中的命题 2 相矛盾。因此，不存在满足（ Θ27）的 ˆ 多项式时间算法。

□

6.5 定理 5 的证明

证明。我们设 $\varepsilon = \sqrt{\bar{d}}$ 在命题 1 的证明中。设 $\mathcal{X}$ 为满足以下条件的近似张量

$$
\|\Theta - \mathcal{X}\|_F \lesssim \sqrt{\bar{d}}, \tag{30}
$$

其 Tucker 秩为（$\mathcal{X}$） = （r_1、.....、r_k），其中 $r_k = c$（$\boldsymbol{s}$， M）$^{-\bar{s}}$对所有 $^{\bar{s}}\left(d_*/\bar{d}\right)$ 取对数 $k \in [m]$. 注意到$\mathcal{X}$ 也满足（30）中的以下不等式

$$
\|\Theta - \mathcal{X}\|_{\mathrm{sp}} := \max_{k\in[m]} \|\mathrm{Unfold}_k(\Theta) - \mathrm{Unfold}_k(\mathcal{X})\|_{\mathrm{sp}} \lesssim \sqrt{\underline{d}}. \tag{31}
$$

定义 $\mathcal{X}_\perp := \Theta - \mathcal{X}$。对于每个 $k = 1, \ldots, m$，我们记

$$
X_k = \mathrm{Unfold}_k(\mathcal{X}), \quad X_{k,\perp} = \mathrm{Unfold}_k(\mathcal{X}_\perp), \quad E_k = \mathrm{Unfold}_k(\mathcal{E}), \quad Y_k = \mathrm{Unfold}_k(\mathcal{Y}),
$$

并定义 $Z_k = X_{k,\perp} + E_k$. 现在我们回顾主文中双投影谱算法的记号。

$$
\begin{aligned}
\tilde{U}_k &= \mathrm{SVD}_{r_k}(Y_k) \\
\hat{U}_k &= \mathrm{SVD}_{r_k}\left(\mathrm{Unfold}_k\left(\mathcal{Y} \times_1 \tilde{U}_1^T \times \cdots \times_{k-1} \tilde{U}_{k-1}^T \times_k \tilde{U}_{k+1}^T \times \cdots \times_m \tilde{U}_m^T\right)\right) \\
\hat{\Theta} &= \mathcal{Y} \times_1 (\hat{U}_1\hat{U}_1^T) \times \cdots \times_m (\hat{U}_m\hat{U}_m^t).
\end{aligned}
$$

记 $U_k = \mathrm{SVD}_{r_k}(X_k)$。对于某个稍后将予以说明的常数 $C_0 > 0$，定义 e

$$
r_k' = \max\{r' \in \{0, \ldots, d_k\} : \sigma_{r'}(X_k) \geq C_0(d_*^{1/4} \vee \bar{d}^{1/2})\}.
$$

若 $\sigma_1(X_k) < C_0(d_*^{1/4} \vee \bar{d}^{1/2})$，则令 $r_k' = 0$。我们用 U_k' 表示U_k的前 r_k' 个奇异向量，并用 V_k表示其余的 $r_k - r_k'$个奇异向量，因此 U_k可写为 $[U_k', V_k]$。接下来定义

$$
X_k' = X_k\left(\mathbb{P}_{U_{k+1}'} \otimes \cdots \otimes \mathbb{P}_{U_m'} \otimes \cdots \otimes \mathbb{P}_{U_{k-1}'}\right),
$$

其中，对于任意正交归一矩阵 $U \in \mathbb{R}^{d\times r}$， $\mathbb{P}_U = UU^T$。我们还记

$$\begin{aligned}
\bar{X}_k &= X_k \left(\tilde{U}_{k+1} \otimes \cdots \otimes \tilde{U}_m \otimes \cdots \otimes \tilde{U}_{k-1}\right), \\
\bar{X}_{k,\perp} &= X_{k,\perp} \left(\tilde{U}_{k+1} \otimes \cdots \otimes \tilde{U}_m \otimes \cdots \otimes \tilde{U}_{k-1}\right), \\
\bar{Y}_k &= Y_k \left(\tilde{U}_{k+1} \otimes \cdots \otimes \tilde{U}_m \otimes \cdots \otimes \tilde{U}_{k-1}\right), \\
\bar{E}_k &= E_k \left(\tilde{U}_{k+1} \otimes \cdots \otimes \tilde{U}_m \otimes \cdots \otimes \tilde{U}_{k-1}\right) \\
\bar{Z}_k &= \bar{X}_{k,\perp} + \bar{E}_k.
\end{aligned}$$

现在我们来界定

$$\|\hat{\Theta} - \Theta\|_F \le \underbrace{\|\Theta \times_1 (\hat{U}_1\hat{U}_1^T) \times \cdots \times_m (\hat{U}_m\hat{U}_m^T) - \Theta\|_F}_{(*)} + \underbrace{\|\mathcal{E} \times_1 \hat{U}_1^T \times \cdots \times_m \hat{U}_m^T\|_F}_{(**)}. \tag{32}$$

为了界定（$*$），我们有

$$\begin{aligned}
(*) &\le \sum_{k\in[m]} \|(I - \hat{U}_k\hat{U}_k^T)\Theta_k\|_F \\
&\le \sum_{k\in[m]} \left(\|(I - \hat{U}_k\hat{U}_k^T)X_k\|_F + \|(I - \hat{U}_k\hat{U}_k^T)X_{k,\perp}\|_F\right) \\
&\le \sum_{k\in[m]} \left(\|\hat{U}_{k,\perp}^T X_k\|_F + \|X_{k,\perp}\|_F\right) \\
&\le \sum_{k\in[m]} \left(\|\hat{U}_{k,\perp}^T X_k'\|_F + \|X_k - X_k'\|_F + \|X_{k,\perp}\|_F\right).
\end{aligned} \tag{33}$$

因此，只需界定 $\|X_k - X_k'\|_F$和 $\|\hat{U}_{k,\perp}^T X_k'\|_F$因为$\|X_{k,\perp}\|_F \le \sqrt{d}$由（30）。

1. $\|X_k - X_k'\|_F$的界：为简化记号，我们重点讨论 $k = 1$，而其他模态的分析也可类似进行。我们有

$$\begin{aligned}
\|X_1 - X_1'\|_F &\le \|X_1 \left((\mathbb{P}_{U_2'} + \mathbb{P}_{V_2'}) \otimes \cdots \otimes (\mathbb{P}_{U_m'} + \mathbb{P}_{U_m'}) - \mathbb{P}_{U_2'} \otimes \cdots \otimes \mathbb{P}_{U_m'}\right)\|_F \\
&\le \sum_{k=2}^m \|V_k'^T X_k\|_F \\
&\le \sum_{k=2}^m \sqrt{r_k - r_k'}\sigma_{r_k'+1}(X_k) \\
&\le \sum_{k=2}^m C_0\sqrt{r_k}(d_*^{1/4} + \bar{d}^{1/2}).
\end{aligned} \tag{34}$$

2. $\|\hat{U}_{k,\perp}^T X_k'\|_F$ 的界：我们有以下两个不等式

$$\|\hat{U}_{k,\perp}^T X_k'(\tilde{U}_{k+1} \otimes \cdots \otimes \tilde{U}_m \otimes \tilde{U}_1 \otimes \cdots \otimes \tilde{U}_{k-1})\|_F \le \|\hat{U}_{k,\perp}^T \bar{X}_k\|_F + \|X_k - X_k'\|_F \tag{35}$$

$$\|\hat{U}_{k,\perp}^T X_k'(\tilde{U}_{k+1} \otimes \cdots \otimes \tilde{U}_m \otimes \tilde{U}_1 \otimes \cdots \otimes \tilde{U}_{k-1})\|_F$$

$$
\begin{aligned}
&= \|\hat{U}_{k,\perp}^T X_k' (\mathbb{P}_{U_{k+1}'} \tilde{U}_{k+1} \otimes \cdots \otimes \mathbb{P}_{U_m'} \tilde{U}_m \otimes \mathbb{P}_{U_1'} \tilde{U}_1 \otimes \cdots \otimes \mathbb{P}_{U_{k-1}'} \tilde{U}_{k-1})\|_F \\
&\geq \|\hat{U}_{k,\perp}^T X_k'\|_F \prod_{\ell \neq k} \sigma_{r_\ell'}(U_\ell^T \tilde{U}_\ell) \\
&= \|\hat{U}_{k,\perp}^T X_k'\|_F \prod_{\ell \neq k} \sqrt{1 - \|\tilde{U}_{\ell,\perp}^T U_\ell'\|_{\mathrm{sp}}^2}. \qquad (36)
\end{aligned}
$$

结合（35）和（36）可得

$$
\|\hat{U}_{k,\perp}^T X_k'\|_F \prod_{\ell \neq k} \sqrt{1 - \|\tilde{U}_{\ell,\perp}^T U_\ell'\|_{\mathrm{sp}}^2} \leq \|\hat{U}_{k,\perp}^T \bar{X}_k\|_F + \|X_k - X_k'\|_F. \qquad (37)
$$

现在，我们对 $\|\hat{U}_{k,\perp}^T \bar{X}_k\|_F$ 和 $\|\tilde{U}_{\ell,\perp}^T U_\ell'\|_{\mathrm{sp}}$ 进行界定，从而得到 $\|\hat{U}_{k,\perp}^T X_k'\|_F$ 的上界。

的上界 $\|\hat{U}_{k,\perp}^T \bar{X}_k\|_F$ 由引理得出 5 以及以下事实 $\bar{Y}_k = \bar{X}_k + \bar{Z}_k$ 当

$$
\begin{aligned}
\|\hat{U}_{k,\perp}^T \bar{X}_k\|_F &\leq 2\sqrt{r_k}\|\bar{Z}\|_{\mathrm{sp}} \\
&\leq 2\sqrt{r_k}\left(\|\bar{X}_{k,\perp}\|_{\mathrm{sp}} + \|\bar{E}_k\|_{\mathrm{sp}}\right) \\
&\leq 2\sqrt{r_k}\left(\|X_{k,\perp}\|_{\mathrm{sp}} + \|\bar{E}_k\|_{\mathrm{sp}}\right) \\
&\lesssim \sqrt{r_k \bar{d}} + \sqrt{r_*} + \sum_{\ell \in [m]} \sqrt{r_\ell \bar{r} d_\ell}, \qquad (38)
\end{aligned}
$$

其中 最后一行使用 $\|\mathcal{X}_\perp\|_{\mathrm{sp}} \leq \sqrt{\bar{d}}$ 根据（31），定义 $\bar{E}_k$ 和引理 4。

根据引理 7，我们对 $\|\tilde{U}_{k,\perp}^T U_k'\|_{\mathrm{sp}}$ 进行界定，其成立的概率至少为 $1 - C\exp(-\underline{c}d)$，对每个 $k \in [m]$,

$$
\begin{aligned}
\|\tilde{U}_{k,\perp}^T U_k'\|_{\mathrm{sp}} &\leq C\left(\frac{\sqrt{d_k} + \|X_{k,\perp}\|_{\mathrm{sp}}}{\sigma_{r'}(X)} + \frac{\sqrt{d_*} + \sqrt{d_*/d_k}\|X_{k,\perp}\|_{\mathrm{sp}} + \|X_{k,\perp}\|_{\mathrm{sp}}^2}{\sigma_{r'}^2(X)}\right) \\
&\leq \frac{C}{C_0}\left(\frac{\sqrt{d_k} + \sqrt{\bar{d}}}{\sqrt{\bar{d}}} + \frac{\sqrt{d_*} + \bar{d}}{\sqrt{d_*} \vee \bar{d}}\right) \\
&\leq \frac{1}{\sqrt{2}} \qquad (39)
\end{aligned}
$$

for 足够大 $C_0 \geq 15$ 其中 15 的设置满足引理 7。

最后，将（34）、（38）和（39）代入（37）可得

$$
\begin{aligned}
\|\hat{U}_{k,\perp}^T X_k'\|_F &\leq 2^{\frac{m-1}{2}}\left(\|\hat{U}_{k,\perp}^T \bar{X}_k\|_F + \|X_k - X_k'\|_F\right) \qquad (40) \\
&\lesssim r_*^{1/2} + \bar{r}\bar{d}^{1/2} + \bar{r}^{1/2} d_*^{1/4}.
\end{aligned}
$$

将（34）和（40）应用于（33）可得

$$
(*) \lesssim r_*^{1/2} + \bar{r}\bar{d}^{1/2} + \bar{r}^{1/2} d_*^{1/4}.
$$

注意到（$**$）项在（32）中以 C（$\sqrt{r_*} + \sum_{\ell \in [m]} \sqrt{d_\ell r_\ell}$）为界，根据引理4 以至少 $1 - \exp(-cd)$的概率。结合（$*$）和（$**$）的上界，我们最终得到

$$
\|\hat{\Theta} - \Theta\|_F \lesssim r_*^{1/2} + \bar{r}\bar{d}^{1/2} + \bar{r}^{1/2} d_*^{1/4}. \qquad (41)
$$

代入 $r_k = c(\boldsymbol{s}, M)^{-\bar{s}}$ 对数$^{\bar{s}}$ $(d_*/\bar{d})$ 对所有 $k \in [m]$ 代入 (41)即可完成证明

f. □

7 讨论

在本文中，我们针对高秩张量估计问题提出了潜变量张量模型。尽管真实数据应用中高秩张量普遍存在，潜变量张量模型仍为低秩方法的经验成功提供了严格的理论依据。我们提出了两种估计方法：统计最优但计算上不可行的最小二乘估计，以及计算最优的DSE估计。我们发现了统计最优速率与计算最优速率之间的内在差距。数值分析证明了我们方法的有效性和适用性。

我们的工作有几个可能的扩展方向。我们将讨论其中的挑战与局限性。**噪声模型的扩展**。我们当前的理论假设噪声张量由来自高斯分布的独立同分布条目组成。事实上，除了定理 4之外，所有定理都可以扩展到独立同分布的次高斯噪声张量。具体而言，我们可以将结果扩展到以下次高斯噪声模型：

- 居中： $\mathbb{E}(\mathcal{E}(i_1,\ldots,i_m)) = 0$ 对于所有 $(i_1,\ldots,i_m) \in [d_1]\times\cdots\times[d_m]$。
- 次高斯：存在一个有界的 $\sigma > 0$，使得对于所有 $(i_1,\ldots,i_m) \in [d_1]\times\cdots\times[d_m]$,

$$\mathbb{E}\exp(\lambda\mathcal{E}(i_1,\ldots,i_m)) \le \exp^{\lambda^2\sigma^2/2}, \quad \text{for all } \lambda\in\mathbb{R}.$$

- 独立同分布（i.i.d.）：对于所有的（$i_1,\ldots,i_m$） $\in[d_1]\times\cdots\times[d_m]$， $\mathcal{E}$（$i_1,\ldots,i_m$）均为独立同分布。

在第 6 节中，我们对定理 2、3 和 5 的证明仅使用次高斯性的性质，因此推广起来很直接。然而，将定理 4 推广到次高斯噪声模型仍然具有挑战性。难点在于如何将命题 2 推广到次高斯情形。命题 2 的主要证明思路是构造随机多项式时间算法 φ，其依据是平均技巧（Ma 和 Wu，2015年）或拒绝核技术（Brennan 等人，2018年），并满足

$$\mathrm{TV}\left(\varphi(\mathrm{HPC}(d,1/x,\tau)),\mathrm{CHC}(\boldsymbol{d},\boldsymbol{r},\lambda)\right)\to 0, \tag{42}$$

其中，我们考虑高斯噪声模型下问题 HPC(d， $1/x$， τ) 与 CHC($\boldsymbol{d}$， $\boldsymbol{r}$， λ) 之间的全变差距离（TV）；详见 Luo 和 Zhang （2022年）。遗憾的是，式（42）的证明大量使用了标准正态分布的显式公式。由于全变差距离的性质，要界定任意次高斯分布下 HPC 与 CHC 问题之间的全变差距离颇具挑战性。

此外，我们的独立同分布噪声假设排除了二元张量，因为伯努利条目通常并非同分布。定理 1-3 对伯努利张量仍然成立，但定理 5可能失效。我们的定理 5 通过使用奇异空间的扰动界（Han 等人，2022年；Cai 和 Zhang，2018年）给出了 DSE 算法的错误率。不幸的是，尽管基于独立同分布次高斯噪声的扰动界已得到充分研究（Cai 和 Zhang，2018年；Fan 等人，2018年），但独立而异方差次高斯噪声下的扰动界尚不明确。在异方差噪声下能否获得类似结果，是一个有趣的扩展方向。

随机与固定的潜在向量。我们目前的模型假设信号张量中的潜在变量未知但固定 Θ。一种可能的扩展是在生成模型（3）中考虑随机潜在变量。我们可以将信号张量建模为 $\Theta(i_1,\ldots,i_m) = f(\boldsymbol{a}_{i_1}^{(1)},\ldots,\boldsymbol{a}_{i_m}^{(m)})$，其中

$\{\boldsymbol{a}_{i_k}^{(k)}: \ i_k \in [d_k]\}$ 是来自某些分布的独立同分布随机向量。针对图子图函数和超图子图函数，已有类似的随机设计模型（Chan 和 Airoldi，2014年；Gao 等，2015年；Klopp 等，2017年；Balasubramanian，2021年）。现在我们证明，在额外满足有界信号条件的情况下，我们的主要定理允许采用这种随机设计： $\|\Theta\|_\infty \leq \alpha$ 对某个常数 $\alpha > 0$成立。具体而言，对于所有 $i_k \in [d_k]$，我们写作 $\boldsymbol{a}_{i_k}^{(k)} \sim \mathcal{A}$ 独立同分布，其中 $\mathcal{A}$ 表示一个分布。我们通过对所有随机性取期望的均方误差来衡量估计性能，定义为 $\mathbb{E}_{\mathcal{Y},\mathcal{A}}(d_*^{-1}\|\hat{}\Theta - \Theta\|_F^2)$。在定理 2 和定理 5 的证明中，我们给出了均方误差的上界：

$$\begin{aligned}
&\mathbb{E}_{\mathcal{Y}}\left(d_*^{-1}\|\hat{\Theta}-\Theta\|_F^2\right) \\
&= \mathbb{E}_{\mathcal{Y}}\left(d_*^{-1}\|\hat{\Theta}-\Theta\|_F^2 \mathbb{1}_{\{\|\hat{\Theta}-\Theta\|_F^2 \geq \epsilon^2\}}\right) + \mathbb{E}_{\mathcal{Y}}\left(d_*^{-1}\|\hat{\Theta}-\Theta\|_F^2 \mathbb{1}_{\{\|\hat{\Theta}-\Theta\|_F^2 < \epsilon^2\}}\right) \\
&\leq 4\alpha^2 \underbrace{\mathbb{P}_{\mathcal{Y}}(\|\hat{\Theta}-\Theta\|_F^2 \geq \epsilon^2)}_{(*)} + \underbrace{d_*^{-1}\epsilon^2}_{(**)},
\end{aligned} \tag{43}$$

其中最后一个不等式是由以下事实得出的： $\|\hat{}\Theta - \Theta\|_\infty \leq 2\alpha$。通过在定理 $\epsilon^2 = \tilde{\mathcal{O}}(\bar{d})$ 中取 2(或在定理 $\varepsilon^2 = \tilde{\mathcal{O}}(d_*^{1/2} + \bar{d})$ 中取 5)，我们使 $(*)$ 指数级递减，从而得到期望均方误差的上界，如 $(**)$ 所示。现在，在随机设计 $\boldsymbol{a}_{i_k}^{(k)} \sim \mathcal{A}$对所有 $i_k \in [d_k]$均独立同分布的条件下，(43) 中的均方误差变为给定$\mathcal{A}$中潜变量的条件期望。由于 (43) 中的上界对 $\mathcal{A}$是一致的，利用，我们得到关于所有随机性的期望均方误差上界

$$\mathbb{E}_{\mathcal{Y},\mathcal{A}}(d_*^{-1}\|\hat{\Theta}-\Theta\|_F^2) = \mathbb{E}_{\mathcal{A}}\left(\mathbb{E}_{\mathcal{Y}}(d_*^{-1}\|\hat{\Theta}-\Theta\|_F^2 \mid \mathcal{A})\right) \lesssim d_*^{-1}\epsilon^2.$$

因此，我们得到了与固定设计情形相同的结论。

界中的对数因子。最后，进一步收紧下界是未来值得研究的问题。我们已证明，统计下界和上界分别为 $\mathcal{O}(d^{-(m-1)})$ 和$\tilde{\mathcal{O}}(d^{-(m-1)})$。类似地，计算下界和上界分别为 $\mathcal{O}(d^{m/2})$ 和$\tilde{\mathcal{O}}(d^{m/2})$。下界与上界之间存在对数级差距。我们猜想，下界分析可以改进至仅相差对数因子。我们将可能的改进留待未来研究。

致谢

本研究部分由 NSF CAREER DMS-2141865、DMS-1915978、DMS-2023239、EF-2133740 项目以及威斯康星校友研究基金会资助。

参考文献

Agarwal、S.、K. Branson 和 S. J. Belongie（2006）。基于图的高阶学习。第*23*届国际机器学习会议论文集。

Balabdaoui、F.、C. Durot 和 H. Jankowski（2019）。单调单指标模型中的最小二乘估计。*Bernoulli 25*(4B)，3276–3310。

Balasubramanian, K.（2021年）。通过超图函数对高阶交互进行非参数建模。机器学习研究期刊 *22*，1–25。

Baltrunas, L., M. Kaminskas, B. Ludwig, O. Moling, F. Ricci, A. Aydin, K.-H. L¨uke 和 R. Schwaiger（2011年）。InCarMusic：面向汽车的情境感知音乐推荐。载于电子商务与网络技术国际会议，第89–100页。Springer。Barak, B. 和 A. Moitra（2016年）。通过平方和层级进行含噪张量补全。载于学习理论会议，第417–445页。PMLR。Berthet, Q. 和 N. Baldin（2020年）。图逻辑回归中的统计与计算速率。载于人工智能与统计国际会议，第2719–2730页。PMLR。Bi, X., A. Qu 和 X. Shen（2018年）。多层张量因式分解及其在推荐系统中的应用。统计学年刊 *46*(6B)，3308–3333。Bickel, P. J. 和 A. Chen（2009年）。网络模型及 Newman–Girvan 模块度和其他模块度的非参数视角。美国国家科学院院刊 *106*(50)，21068–21073。Brennan, M., G. Bresler 和 W. Huleihel（2018年，7月6–9日）。具有植入稀疏结构问题的可约性与计算下界。载于 S. Bubeck、V. Perchet 和 P. Rigollet（编），第*31*届学习理论会议论文集，第75卷，机器学习研究论文集，第48–166页。PMLR。B¨uchel, C. 和 K. J. Friston（1997年）。注意力对视觉通路连接性的调节：利用结构方程建模和功能性磁共振成像评估皮层交互。大脑皮层（纽约，纽约州：*1991*）*7*(8)，768–778。Cai, T. T. 和 Y. Wu（2020年）。稀疏矩阵检测的统计与计算极限。统计学年刊 *48*(3)，1593–1614。Cai, T. T. 和 A. Zhang（2018年）。奇异子空间的速率最优扰动界及其在高维统计学中的应用。统计学年刊 *46*(1)，60–89。Chan, S. 和 E. Airoldi（2014年）。可交换图模型的一致直方图估计量。载于国际机器学习会议，第208–216页。De Lathauwer, L., B. De Moor 和 J. Vandewalle（2000年）。多线性奇异值分解。*SIAM* 矩阵分析与应用期刊 *21*(4)，1253–1278。Eckart, C. 和 G. Young（1936年）。用一个低秩矩阵逼近另一个矩阵。*Psychometrika 1*(3)，211–218。Fan, J., W. Wang 和 Y. Zhong（2018年）。特征向量扰动界及其在稳健协方差估计中的应用。机器学习研究期刊 *18*(207)，1–42。Ganti, R., N. Rao, L. Balzano, R. Willett 和 R. Nowak（2017年）。关于高维结构化单指标模型的学习。载于第三十一届 *AAAI* 人工智能会议论文集，第1898–1904页。Gao, C., Y. Lu 和 H. H. Zhou（2015年）。速率最优的图函数估计。统计学年刊 *43*(6)，2624–2652。Han, R., Y. Luo, M. Wang 和 A. R. Zhang（2022年）。张量块模型中的精确聚类：统计最优性与计算极限。英国皇家统计学会会刊 *B*辑：统计方法学 *84*(5)，1666–1698。

Han, R., R. Willett, 和 A. R. Zhang（2022年）。广义张量估计的最优统计与计算框架。《统计学年鉴》*50*(1), 1–29. Hazan, E. 和 R. Krauthgamer（2011年）。近似求解纳什均衡最优解有多难?《*SIAM*计算杂志》*40*(1), 79–91. Hillar, C. J. 和 L.-H. Lim（2013年）。大多数张量问题都是NP-困难。《*ACM*杂志（*JACM*）》*60*(6), 45. Hitchcock, F. L.（1927年）。张量或多元组表示为乘积之和。《数学与物理杂志》*6*(1-4), 164–189. Hoff, P. D.（2015年）。纵向关系数据的多线性张量回归。《应用统计学年鉴》*9*(3), 1169. Hong, D., T. G. Kolda, 和 J. A. Duersch（2020年）。广义典范多元组张量分解。《*SIAM*评论》*62*(1), 133–163. Hore, V., A. Viˆnuela, A. Buil, J. Knight, M. I. McCarthy, K. Small, 和 J. Marchini（2016年）。多组织基因表达实验的张量分解。《自然遗传学》*48*(9), 1094. Hu, J., C. Lee, 和 M. Wang（2021年）。多个模态带有特征的广义张量分解。《计算与图形统计学杂志》, 1–15. Klopp, O., A. B. Tsybakov, 和 N. Verzelen（2017年）。网络模型与稀疏图子估计的预言机不等式。《统计学年鉴》*45*(1), 316–354. Kolda, T. G. 和 B. W. Bader（2009年）。张量分解及其应用。《*SIAM*评论》*51*(3), 455–500. Lee, C. 和 M. Wang（2021a）。超越符号：通过符号级数进行非参数张量补全。《神经信息处理系统进展》*34*。Lee, C. 和 M. Wang（2021b）。具有未知置换的平滑张量估计。*arXiv*预印本 *arXiv:2111.04681*。Lov´asz, L.（2012年）。大型网络与图极限，第60卷。美国数学学会。Lov´asz, L. 和 B. Szegedy（2006年）。稠密图序列的极限。《组合理论杂志*B*辑》*96*(6), 933–957. Luo, Y. 和 A. R. Zhang（2020年）。开放问题：超图植入团检测的平均情形困难性。载于学习理论会议，第3852–3856页。PMLR。Luo, Y. 和 A. R. Zhang（2022年）。带有隐藏结构的张量聚类：统计最优性与计算极限。《统计学年鉴》*50*(1), 584–613. Ma, Z. 和 Y. Wu（2015年）。极小极大子矩阵检测中的计算障碍。《统计学年鉴》*43*(3), 1089–1116. Michoel, T. 和 B. Nachtergaele（2012年）。通过基于超图的谱聚类对复杂网络进行对齐与整合。《物理评论*E*》*86*(5), 056111. Pananjady, A. 和 R. J. Samworth（2022年）。未知置换下的保序回归：统计、计算与适应。《统计学年鉴》*50*(1), 324–350.

Richard, E. 和 A. Montanari（2014年）。张量主成分分析的统计模型。《神经信息处理系统进展》*27*。Robinson, P. M.（1988年）。根号n一致的半参数回归。《计量经济学：计量经济学会杂志》*56*(4), 931–954. Sun, W. W., J. Lu, H. Liu, 和 G. Cheng（2017年）。可证明的稀疏张量分解。《英国皇家统计学会杂志：*B*辑（统计方法）》*79*(3), 899–916. Tucker, L. R.（1966年）。关于三模态因子分析的一些数学札记。《心理计量学》*31*(3), 279–311. Udell, M. 和 A. Townsend（2019年）。为什么大数据矩阵近似为低秩？《*SIAM*数据科学数学杂志》*1*(1), 144–160. Wang, M., J. Fischer, 和 Y. S. Song（2019年）。使用半非负张量分解对多组织、多个人的基因表达数据进行三路聚类。《应用统计学年鉴》*13*(2), 1103–1127. Wang, M. 和 L. Li（2020年）。从二元多路数据中学习：概率张量分解及其统计最优性。《机器学习研究杂志》*21*(154), 1–38. Wang, M. 和 Y. Zeng（2019年）。通过张量块模型进行多路聚类。《神经信息处理系统进展》*32*。Wang, T., Q. Berthet, 和 R. J. Samworth（2016年）。稀疏主成分估计中的统计与计算权衡。《统计学年鉴》*44*(5), 1896–1930. Wu, Y. 和 J. Xu（2021年）。带有隐藏结构的统计问题：信息理论与计算极限。《数据科学中的信息理论方法》，383。Xu, J.（2018年）。图函数估计中谱方法的收敛速率。载于国际机器学习会议，第5433–5442页。Zhang, A. 和 D. Xia（2018年）。张量SVD：统计与计算极限。《*IEEE* 信息论汇刊》*64*(11), 7311 – 7338. Zhao, Y.（2015年）。超图极限：一种正则性方法。《随机结构与算法》*47*(2), 205–226. Zhou, H., L. Li, 和 H. Zhu（2013年）。张量回归及其在神经影像数据分析中的应用。《美国统计学会杂志》*108*(502), 540–552.

附录

附录包含技术引理和额外的模拟结果。

一个技术引理

引理 3 (引理 E.5，见 Han 等人 (2022年))。 假设 $\mathcal{E} \in \mathbb{R}^{d_1 \times \cdots \times d_m}$ 的所有元素相互独立、均值为零且服从次高斯分布，其代理方差为 σ。那么存在一些通用常数 C， c使得

$$\sup_{\substack{\mathcal{T} \in \mathbb{R}^{d_1 \times \cdots \times d_m}, \|\mathcal{T}\|_F \leq 1 \\ \operatorname{rank}(\mathcal{T}) \leq (r_1, \ldots, r_m)}} \langle \mathcal{T}, \mathcal{E} \rangle \leq C\sigma \left(r_* + \sum_{k=1}^{m} d_m r_k \right)^{1/2},$$

以至少 $1 - \exp(-c \sum_{k=1}^{m} d_k r_k)$ 的概率。

引理 4 (引理 8，见 Han 等人 (2022年))。 设 $E \in \mathbb{R}^{d_1 \times \cdots \times d_m}$ 为噪声张量，其每个元素均相互独立并服从均值为零的次高斯分布，且具有 $\sigma = 1$ ，不失一般性。固定$U_k^* \in \mathbb{O}_{d_k,\ r_k}$。那么，以至少 $1 - \exp(-cd)$ 的概率，以下结论成立：

$$\begin{aligned}
&\|E_k \left(U_{k+1}^* \otimes \cdots \otimes U_m^* \otimes U_1^* \otimes \cdots \otimes U_{k-1}^*\right)\|_{\mathrm{sp}} \leq C(\sqrt{d_k} + \sqrt{r_{-k}}), \\
&\|E_k \left(U_{k+1}^* \otimes \cdots \otimes U_m^* \otimes U_1^* \otimes \cdots \otimes U_{k-1}^*\right)\|_F \leq C\sqrt{d_k r_{-k}}, \\
\sup_{\substack{U_\ell \in \mathbb{O}_{d_\ell, r_\ell} \\ \ell \neq [m]}} &\|E_k \left(U_{k+1} \otimes \cdots \otimes U_m \otimes U_1 \otimes \cdots \otimes U_{k-1}\right)\|_{\mathrm{sp}} \leq C(\sqrt{d_k} + \sqrt{r_{-k}} + \sum_{\ell \neq k} \sqrt{d_\ell r_\ell}), \\
\sup_{\substack{U_\ell \in \mathbb{O}_{d_\ell, r_\ell} \\ \ell \neq [m]}} &\|E_k \left(U_{k+1} \otimes \cdots \otimes U_m \otimes U_1 \otimes \cdots \otimes U_{k-1}\right)\|_F \leq C(\sqrt{d_k r_{-k}} + \sum_{\ell \neq k} \sqrt{d_\ell r_\ell}), \\
\sup_{\substack{U_\ell \in \mathbb{O}_{d_\ell, r_\ell} \\ \ell \neq [m]}} &\|\mathcal{E} \times_1 U_1^T \times \cdots \times_m U_m^T\|_F \leq C(\sqrt{r_*} + \sum_{\ell \in [m]} \sqrt{d_\ell r_\ell}).
\end{aligned}$$

引理 5 (扰动的投影界)。 假设 X， $E \in \mathbb{R}^{m \times n}$ 且秩$(X) = r$ 。令$U \in \mathbb{O}_{m,\ r}$ 为 $Y = X + E$ 的前 r 个奇异向量。那么，

$$\begin{aligned}
&\|(I - UU^T)X\|_{\mathrm{sp}} \leq 2\|E\|_{\mathrm{sp}}, \\
&\|(I - UU^T)X\|_F \leq \min\left(2\sqrt{r}\|E\|_{\mathrm{sp}}, 2\|E\|_F\right).
\end{aligned}$$

引理 5 的证明。 对于矩阵范数界，我们有

$$\begin{aligned}
\|(I - UU^T)X\|_{\mathrm{sp}} &\leq \|(I - UU^T)Y\|_{\mathrm{sp}} + \|E\|_{\mathrm{sp}} \\
&\leq \sigma_{r+1}(Y) + \|E\|_{\mathrm{sp}} \\
&\leq \min_{Z \in \mathbb{R}^{m \times n}:\ \operatorname{rank}(Z) \leq r} \|Y - Z\|_{\mathrm{sp}} + \|E\|_{\mathrm{sp}} \\
&\leq \|Y - X\|_{\mathrm{sp}} + \|E\|_{\mathrm{sp}} \\
&\leq 2\|E\|_{\mathrm{sp}}.
\end{aligned} \tag{44}$$

同样，我们对弗罗贝尼乌斯范数进行界定

$$\|(I - UU^T)X\|_F \leq \|(I - UU^T)Y\|_F + \|E\|_F$$

$$
\begin{aligned}
&\leq \sqrt{\sum_{i=r+1}^{m\wedge n} \sigma_i^2(Y)} + \|E\|_F \\
&\leq \min_{Z\in\mathbb{R}^{m\times n}:\,\mathrm{rank}(Z)\leq r} \|Y-Z\|_F + \|E\|_F \\
&\leq \|Y-X\|_F + \|E\|_F \\
&\leq 2\|E\|_F.
\end{aligned}
$$

此外，直接应用（44）可得

$$
\|(I-UU^T)X\|_F \leq 2\sqrt{r}\|E\|_{\mathrm{sp}}.
$$

□

引理6 (引理2，见 Han 等人 (2022年))。 假设前 r 和其余 d_1-r 奇异向量 of $Y\in\mathbb{R}^{d_1\times d_2}$ 为 $\tilde{U}\in\mathbb{O}_{d_1,\,r}$ 和 $\tilde{U}_\perp\in\mathbb{O}_{d_1,\,d_1-r}$，分别。对于某个 $1\leq r'\leq r$，令 $W\in\mathbb{O}_{d_1,\,r'}$为任意正交归一矩阵，并令 $W_\perp\in\mathbb{O}_{d_1,\,r'}$ 为 W的正交补。给定$\sigma_{r'}(W^TY) > \sigma_{r+1}(Y)$，我们有

$$
\|\tilde{U}_{r\perp}^T W\|_{\mathrm{sp}} \leq \frac{\sigma_{r'}(W^TY)\|W_\perp^T Y\mathbb{P}_{Y^TW}\|_{\mathrm{sp}}}{\sigma_{r'}^2(W^TY)-\sigma_{r+1}^2(Y)}.
$$

引理 7 （不同维度子空间上的扰动界）。 考虑信号加噪声模型，

$$
Y = X + X_\perp + E \in \mathbb{R}^{d_1\times d_2},
$$

其中， X 是一个信号矩阵，满足秩$(X)=r$； $X_\perp$ 是一个扰动矩阵； E 是一个含有独立同分布标准次高斯元素的噪声矩阵。定义

$$
r' := \max\{r'\in\{0,1,\ldots,r\}\colon \sigma_{r'}(X)\geq \max(\sqrt{3}(d_1+\sqrt{d_1d_2}), 16\|X_\perp\|_{\mathrm{sp}})\}. \tag{45}
$$

我们记

$$
\tilde{U}_r = \mathrm{SVD}_r(Y), \quad U_{r'} = \mathrm{SVD}_{r'}(X).
$$

则以至少 $1-\exp(-cd_1\wedge d_2)$ 的概率，

$$
\|\tilde{U}_{r\perp}^T U_{r'}\|_{\mathrm{sp}} \leq C\left(\frac{\sqrt{d_1}+\|X_\perp\|_{\mathrm{sp}}}{\sigma_{r'}(X)} + \frac{\sqrt{d_1d_2}+\sqrt{d_1\vee d_2}\|X_\perp\|_{\mathrm{sp}}+\|X_\perp\|_{\mathrm{sp}}^2}{\sigma_{r'}^2(X)}\right),
$$

其中 $\tilde{U}_{r\perp}\in\mathbb{R}^{d_1\times d_1-r}$ 是 $\tilde{U}_r$ 的正交补矩阵。

引理 7。 应用引理 6 结合 $W=U_{r'}$，我们有

$$
\|\tilde{U}_{r\perp}^T U_{r'}\|_{\mathrm{sp}} \leq \frac{\sigma_{r'}(U_{r'}^TY)\|U_{r'\perp}^T Y\mathbb{P}_{Y^TU_{r'}}\|_{\mathrm{sp}}}{\sigma_{r'}^2(U_{r'}^TY)-\sigma_{r+1}^2(Y)}, \tag{46}
$$

因此，只需给出 $\sigma_{r'}^2(U_{r'}Y)-\sigma_{r+1}^2(Y)$、 $\|U_{r'\perp}^TY\mathbb{P}_{Y^TU_{r'}}\|_{\mathrm{sp}}$ 和 $\sigma_{r'}(U_{r'}^TY)$ 的概率界。

首先，我们记

$$
\begin{aligned}
&X_1 = \mathbb{P}_{U_{r'}} X, \quad X_2 = X - X_1, \\
&Y = X + X_\perp + E, \quad Y' = X + E, \\
&Y_1 = X_1 + X_\perp + E, \quad Y_1' = X_1 + E.
\end{aligned}
$$

现在给出 $\sigma_{r'}^2(U_{r'}Y) - \sigma_{r+1}^2(Y)$ 和 $\|U_{r'\perp}^T Y \mathbb{P}_{Y^T U_{r'}}\|_{\mathrm{sp}}$，以及 $\sigma_{r'}(U_{r'}^T Y)$ 的界。

该项的界 $\sigma_{r'}^2(U_{r'}Y) - \sigma_{r+1}^2(Y)$ 根据 Cai 和 Zhang，2018年，引理 4），对于所有 $x > 0$，我们有

$$
\begin{aligned}
&\mathbb{P}\left(\sigma_{r'}^2(U_{r'}^T Y_1') \le (\sigma_{r'}^2(X_1) + d_2)(1-x)\right) \le C\exp(Cr - c(\sigma_{r'}(X_1) + d_2)x^2 \wedge x), \\
&\mathbb{P}\left(\sigma_{r+1}^2(Y') \ge d_2(1+x)\right) \le C\exp(Cd_1 - cd_2 x^2 \wedge x), \\
&\mathbb{P}\left(\|U_{r'\perp}^T Y_1 \mathbb{P}_{Y_1^T U_{r'}}\|_{\mathrm{sp}} \ge x\right) \le C\exp\left(C\exp(Cd_1 - cx^2 \wedge x\sqrt{\sigma_{r'}^2(X + X_\perp) + d_2}\right).
\end{aligned}
$$

通过将 x 设为 $\sigma_{r'}^2(X)/3(\sigma_{r'}^2(X)+d_2)$，$\sigma_{r'}^2(X)/3d_2$，以及 $C\left(\sqrt{d_1} + d_1/\sqrt{\sigma_{r'}^2(X+X_\perp)+d_2}\right)$ 分别设定，我们得到

$$
\begin{aligned}
&\sigma_{r'}(U_{r'}^T Y_1') \ge \sqrt{\frac{2\sigma_{r'}^2(X)}{3} + d_2}, \quad \sigma_{r+1}(Y') \le \sqrt{\frac{\sigma_{r'}^2(X)}{3} + d_2}, \\
&\text{and } \|U_{r'\perp}^T Y_1 \mathbb{P}_{Y_1^T U_{r'}}\|_{\mathrm{sp}} \le C\left(\sqrt{d_1} + \frac{d_1}{\sqrt{\sigma_{r'}^2(X + X_\perp) + d_2}}\right),
\end{aligned}
$$

以至少 $1\text{-}C\exp(-cd_1 \wedge d_2)$ 的概率。由于 $Y = Y' + X_\perp$ 和 $Y_1 = Y_1' + X_\perp$，应用 Weyl 不等式可得

$$
\begin{aligned}
&\sigma_{r'}(U_{r'}^T Y) = \sigma_{r'}(U_{r'}^T Y_1) \ge \sqrt{\frac{2\sigma_{r'}^2(X)}{3} + d_2} - \|X_\perp\|_{\mathrm{sp}}, \\
&\sigma_{r+1}(Y) \le \sqrt{\frac{\sigma_{r'}^2(X)}{3} + d_2} + \|X_\perp\|_{\mathrm{sp}}.
\end{aligned} \tag{47}
$$

因此，由（47）可得以下不等式：

$$
\begin{aligned}
\sigma_{r'}^2(U_{r'}^T Y) - \sigma_{r+1}^2(Y) &\ge \frac{\sigma_{r'}^2(X)}{3} - 4\|X_\perp\|_{\mathrm{sp}}\sqrt{\frac{2\sigma_{r'}^2(X)}{3} + d_2} \\
&\ge \frac{\sigma_{r'}^2(X)}{3} - \frac{1}{4}\sigma_{r'}(X)\sqrt{\frac{2\sigma_{r'}^2(X)}{3} + d_2} \\
&\ge \frac{\sigma_{r'}^2(X)}{12},
\end{aligned} \tag{48}
$$

其中第二个不等式使用了 $\sigma_{r'}(X) \ge 16\|X_\perp\|_{\mathrm{sp}}$ 而最后一个不等式使用了 $\sigma_{r'}(X) \ge \sqrt{3d_2}$根据 r' 在（45）中的定义。

界限为 $\|U_{r'\perp}^T Y \mathbb{P}_{Y^T U_{r'}}\|_{\mathrm{sp}}$ 由于 $\sigma_{r'}(X + X_\perp) \geq \sigma_{r'}(X) - \|X_\perp\|_{\mathrm{sp}} \geq \frac{15}{16}\sigma_{r'}(X)$，我们有

$$\|U_{r'\perp}^T Y_1 \mathbb{P}_{Y_1^T U_{r'}}\|_{\mathrm{sp}} \leq C\left(\sqrt{d_1} + \frac{d_1}{\sqrt{\sigma_{r'}^2(X + X_\perp) + d_2}}\right) \leq C\sqrt{d_1},$$

其中我们使用了 r' 的定义，而 C 吸收了所有常数因子。注意，

$$\begin{aligned}
\|U_{r'\perp}^T Y \mathbb{P}_{Y^T U_{r'}}\|_{\mathrm{sp}} &= \|U_{r'\perp}^T Y \mathbb{P}_{Y_1^T U_{r'}}\|_{\mathrm{sp}} \\
&= \|U_{r'\perp}^T (Y_1 + X_2) \mathbb{P}_{Y_1^T U_{r'}}\|_{\mathrm{sp}} \\
&\leq \|U_{r'\perp}^T Y_1 \mathbb{P}_{Y_1^T U_{r'}}\|_{\mathrm{sp}} + \|U_{r'\perp}^T X_2 \mathbb{P}_{Y_1^T U_{r'}}\|_{\mathrm{sp}} \\
&\leq C\sqrt{d_1} + \|U_{r'\perp}^T X_2 \mathbb{P}_{Y_1^T U_{r'}}\|_{\mathrm{sp}}.
\end{aligned}$$

现在我们将重点放在界定 $\|U_{r'\perp}^T X_2 \mathbb{P}_{Y_1^T U_{r'}}\|_{\mathrm{sp}}$ 上。我们有

$$\begin{aligned}
\|U_{r'\perp}^T X_2 \mathbb{P}_{Y_1^T U_{r'}}\|_{\mathrm{sp}} &\leq \frac{\|U_{r'\perp}^T X_2 (Y_1^T U_{r'})\|_{\mathrm{sp}}}{\sigma_{r'}(Y_1^T U_{r'})} \\
&\leq \frac{\|U_{r'\perp}^T X_2 X_\perp^T U_{r'}\|_{\mathrm{sp}} + \|U_{r'\perp}^T X_2 E^T U_{r'}\|_{\mathrm{sp}}}{\sigma_{r'}(Y_1^T U_{r'})} \\
&\leq C\frac{\|U_{r'\perp}^T X_2 X_\perp^T U_{r'}\|_{\mathrm{sp}} + \|U_{r'\perp}^T X_2 E^T U_{r'}\|_{\mathrm{sp}}}{\sigma_{r'}(X)},
\end{aligned} \tag{49}$$

其中我们使用了 $\sigma_{r'}(Y_1^T U_{r'}) \geq \frac{15}{16}\sigma_{r'}(X)$，这是结合（47）以及（ r' （45）中的定义得到的。此外，根据引理 4，以概率 $1 - C\exp(-cd_1)$ 成立以下结果，

$$\|U_{r'\perp}^T X_2 E^T U_{r'}\|_{\mathrm{sp}} \leq C\|X_2\|_{\mathrm{sp}}\sqrt{d_1}.$$

因此，根据（ $\|U_{r'\perp}^T X_2 \mathbb{P}_{Y_1^T U_{r'}}\|_{\mathrm{sp}}$ （49）可得

$$\|U_{r'\perp}^T X_2 \mathbb{P}_{Y_1^T U_{r'}}\|_{\mathrm{sp}} \leq C\frac{\|X_2\|_{\mathrm{sp}}(\|X_\perp\|_{\mathrm{sp}} + \sqrt{d_1})}{\sigma_{r'}(X)} \leq C(\|X_\perp\|_{\mathrm{sp}} + \sqrt{d_1}),$$

其中最后一个不等式使用了 $\|X_2\|_{\mathrm{sp}} \leq \sigma_{r'}$（$X$）。

最后，我们得到 $\|U_{r'\perp}^T Y \mathbb{P}_{Y^T U}$ 的界 $_{r'}\|_{\mathrm{sp}}$ 作为

$$\|U_{r'\perp}^T Y \mathbb{P}_{Y^T U_{r'}}\|_{\mathrm{sp}} \leq C(\|X_\perp\|_{\mathrm{sp}} + \sqrt{d_1}) \tag{50}$$

的界 $\sigma_{r'}(U_{r'}^T Y)$ 我们利用 Weyl 不等式得到上界，

$$\begin{aligned}
\sigma_{r'}(U_{r'}^T Y) &\leq \sigma_{r'}(X) + \|U_{r'}^T X_\perp\|_{\mathrm{sp}} + \|U_{r'}^T E\|_{\mathrm{sp}} \\
&\leq \sigma_{r'}(X) + \|X_\perp\|_{\mathrm{sp}} + \sqrt{d_2},
\end{aligned} \tag{51}$$

其中，最后一个不等式以至少 $1 - C\exp(-cd_2)$ 的概率成立，这由引理 4得出。

最后，将不等式（48）、（50）和（51）代入（46）即可得到，

$$\|\tilde{U}_{r\perp}^T U_{r'}\|_{\mathrm{sp}} \leq C\left(\frac{\sqrt{d_1} + \|X_\perp\|_{\mathrm{sp}}}{\sigma_{r'}(X)} + \frac{\sqrt{d_1 d_2} + \sqrt{d_1 \vee d_2}\|X_\perp\|_{\mathrm{sp}} + \|X_\perp\|_{\mathrm{sp}}^2}{\sigma_{r'}^2(X)}\right).$$

□

B 作物生产分析的补充说明

我们基于估计得到的信号张量 $\hat{\Theta}$ 的 Tucker 表示进行聚类分析。该方法的灵感来自主成分分析（PCA）的高阶扩展。回顾一下，在矩阵情形中，我们基于以下步骤对矩阵 $\boldsymbol{X} \in \mathbb{R}^{m \times n}$ 进行聚类。首先，我们将 X 分解为

$$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T,$$

其中 $\boldsymbol{\Sigma}$ 是一个对角矩阵，$\boldsymbol{U}$，$\boldsymbol{V}$ 是列正交的因子矩阵。其次，我们将 $\boldsymbol{V}$ 的每一列作为一个主轴，并将 $\boldsymbol{U}\boldsymbol{\Sigma}$ 的每一行作为主成分。随后，对$\boldsymbol{U}\boldsymbol{\Sigma}$的 m 行应用多变量聚类方法（例如 K-均值）。

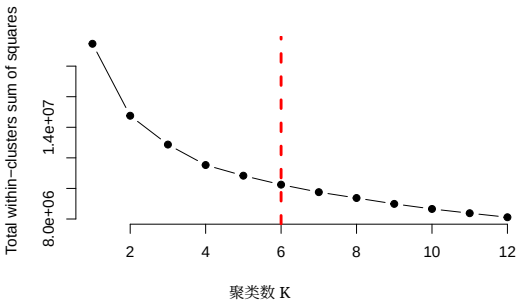
^^ 我们对估计的信号张量 Θ 应用类似的聚类过程。我们基于 Tucker 分解对 Θ 进行因式分解。

$$\hat{\Theta} = \hat{\mathcal{C}} \times_1 \hat{\boldsymbol{U}}_1 \times_2 \cdots \times_m \hat{\boldsymbol{U}}_m, \tag{52}$$

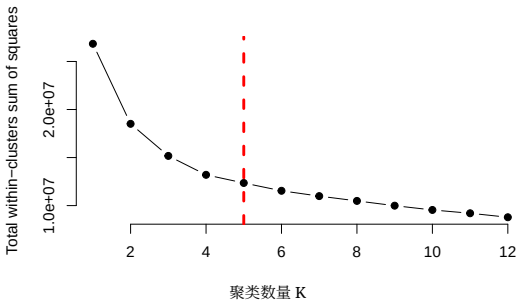
其中，$\hat{\mathcal{C}} \in \mathbb{R}^{r_1 \times \cdots \times r_m}$ 是估计的核心张量，$\hat{\boldsymbol{U}}_k \in \mathbb{R}^{d_k \times r_k}$ 是列正交的估计因子矩阵。模态-k 展开（52）得

$$\text{Unfold}_k(\hat{\Theta}) = \hat{\boldsymbol{U}}_k \text{Unfold}_k(\hat{\mathcal{C}}) \left(\hat{\boldsymbol{U}}_m \otimes \cdots \otimes \hat{\boldsymbol{U}}_1\right).$$

我们对该模态-k 展开后的信号张量进行聚类。我们将 $\left(\hat{\boldsymbol{U}}_m \otimes \cdots \otimes \hat{\boldsymbol{U}}_1\right)$^ ^ 中的列作为主轴，并将 $\boldsymbol{U}_k$展开$_k(\mathcal{C})$ 中的行作为主成分。最后，我们对 K-均值 ^ ^ 聚类方法应用于 d_k 矩阵 $\boldsymbol{U}_k$展开$_k$ $(\mathcal{C})$ 的行。我们根据肘部法则选择聚类数量。图S1 表明，国家和作物分别适合分为六类和五类。



(a) 国家聚类的肘部图



(b) 作物聚类的肘部图

图 S1：用于确定 K-均值聚类数量的肘部图。

正文中研究了国家聚类。这里，我们提供作物的聚类结果。表格 S1 汇总了作物项目的五个聚类。我们发现，五个聚类对应着相似类型的作物。例如，聚类 3 代表浆果和叶菜类植物，而聚类 4 则由主要产自亚洲地区的作物组成。

表S1：基于作物生产数据应用中估计的信号张量对161种作物进行的五个聚类

组	产品项目
1	"Maize" "Nuts nes" "Sunflower seed" "Oats" "Groundnuts, with shell" "Papayas" "Coconuts" "Plantains and others" "Mushrooms and truffles" "Manila fibre (abaca)" "Pyrethrum, dried"
2	"Grapes" "Treenuts, Total" "Tobacco, unmanufactured" "Tomatoes" "Vetches" "Peas, dry" "Bananas" "Flax fibre and tow" "Rapeseed" "Cauliflowers and broccoli" "Raspberries" "Ginger" "Taro (cocoyam)" "Quinoa" "Rubber, natural" "Cherries, sour" "Cow peas, dry" "Okra" "Hazelnuts, with shell" "Chestnut" "Vanilla" "Chicory roots"
3	"Berries nes" "Figs" "Melons, other (inc.cantaloupes)" "Olives" "Plums and sloes" "Pulses nes" "Sesame seed" "Watermelons" "Vegetables Primary" "Beans, green" "Carrots and turnips" "Dates" "Bastfibres, other" "Cocoa, beans" "Coffee, green" "Yams" "Canary seed" "Lupins" "Tung nuts" "Safflower seed" "Spinach" "Strawberries" "Grain, mixed" "Jute" "Fonio" "Persimmons" "Blueberries" "Cranberries" "Hempseed" "Yautia (cocoyam)" "Walnuts, with shell" "Mat\xe9" "Brazil nuts, with shell"
4	"Cereals, Total" "Lentils" "Chillies and peppers, green" "Maize, green" "Sugar crops nes" "Roots and tubers nes" "Vegetables, leguminous nes" "Leeks, other alliaceous vegetables" "Anise, badian, fennel, coriander" "Kapok fruit" "Pistachios"
5	"Apples" "Apricots" "Barley" "Fruit, citrus nes" "Fruit, fresh nes" "Fruit, stone nes" "Linseed" "Millet" "Oranges" "Peaches and nectarines" "Pears" "Potatoes" "Rice, paddy" "Seed cotton" "Sugar beet" "Sugar cane" "Vegetables, fresh nes" "Wheat" "Citrus Fruit, Total" "Fruit Primary" "Pulses, Total" "Roots and Tubers, Total" "Sugar Crops Primary" "Beans, dry" "Rye" "Almonds, with shell" "Artichokes" "Broad beans, horse beans, dry" "Carobs" "Chick peas" "Grapefruit (inc. pomelos)" "Lemons and limes" "Onions, dry" "Cashew nuts, with shell" "Cassava" "Castor oil seed" "Oil palm fruit" "Onions, shallots, green" "Pineapples" "Sisal" "Sweet potatoes" "Cabbages and other brassicas" "Cucumbers and gherkins" "Eggplants (aubergines)" "Pumpkins, squash and gourds" "Asparagus" "Cereals nes" "Chillies and peppers, dry" "Fibre crops nes" "Garlic" "Hops" "Peas, green" "Sorghum" "Soybeans" "String beans" "Tea" "Cherries" "Currants" "Lettuce and chicory" "Mangoes, mangosteens, guavas" "Oilseeds nes" "Pigeon peas" "Areca nuts" "Fruit, tropical fresh nes" "Spices nes" "Karite nuts (sheanuts)" "Pepper (piper spp.)" "Buckwheat" "Mustard seed" "Nutmeg, mace and cardamoms" "Avocados" "Cashewapple" "Quinces" "Bambara beans" "Melonseed" "Hemp tow waste" "Ramie" "Agave fibres nes" "Cinnamon (cannella)" "Cloves" "Gooseberries" "Kola nuts" "Poppy seed"