arXiv:2602.09762v1 [math.ST] 10 Feb 2026

高维部分含噪数据的高斯核矩阵渐近分析

Kensuke Aishima *

2026年2月11日

摘要

高斯核是最重要的核函数之一，适用于包括科学计算和数据科学在内的许多研究领域。本文针对含噪数据的统计模型，给出了高维情形下高斯核矩阵的渐近分析。主要成果是将 Karoui 的渐近分析与约束低秩矩阵逼近方法巧妙结合。具体而言，Karouli 阐明了高斯核矩阵一种重要的渐近结构，由此得出特征向量具有强相合性，但特征值并不相合。本文聚焦于上述结果，并提出一种利用最小特征值的相合估计量，适用于目标核矩阵在渐近区域趋于低秩的情形。值得注意的是，本文在表示部分噪声的统计模型下进行了渐近分析。尽管朴素估计量并不相合，但通过应用带约束的低秩近似优化方法，我们克服了由不相合性导致的困难，从而在秩亏情形下得到一种具有强相合性的新估计量。

关键词：特征值问题、高斯核、低秩逼近、一致估计$PACS$: 65F15, 65F55, 15A18, 62H12, 60B12

1 引言

核方法是适用于许多研究领域的标准技术，例如科学计算和数据科学。尤其是高斯核，它是最重要的核之一，通常与支持向量机（SVM）、主成分分析（PCA）等结合使用。本研究讨论高斯核的渐近分析。

下面对高斯核进行数学描述。全文中，对于正整数 p 和 q，I_p 为 $p \times p$ 单位矩阵，$0_{p\times q}$ 为 $p \times q$ 零矩阵。此外，$[\cdot]_{pq}$ 表示参数矩阵的（p，q）元素。关于范数，$\|\cdot\|$ 表示 2-范数。给定 n 个向量 $\boldsymbol{x}_i \in \mathbb{R}^d$（$i = 1, \ldots\ldots, n$），定义

$$[K(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)]_{ij} = \exp\left(-\frac{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2}{c_d}\right) \qquad (i, j = 1, \ldots, n), \tag{1}$$

其中，c_d 是一个缩放参数。上述矩阵的元素表示为

$$\begin{bmatrix} 1 & \exp\left(-\frac{\|\boldsymbol{x}_1-\boldsymbol{x}_2\|^2}{c_d}\right) & \cdots & \exp\left(-\frac{\|\boldsymbol{x}_1-\boldsymbol{x}_n\|^2}{c_d}\right) \\ \exp\left(-\frac{\|\boldsymbol{x}_2-\boldsymbol{x}_1\|^2}{c_d}\right) & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \exp\left(-\frac{\|\boldsymbol{x}_{n-1}-\boldsymbol{x}_n\|^2}{c_d}\right) \\ \exp\left(-\frac{\|\boldsymbol{x}_n-\boldsymbol{x}_1\|^2}{c_d}\right) & \cdots & \exp\left(-\frac{\|\boldsymbol{x}_n-\boldsymbol{x}_{n-1}\|^2}{c_d}\right) & 1 \end{bmatrix},$$

其中，所有对角元素均为 1。

由于核技巧，上述矩阵被视为格拉姆矩阵。严格来说，上述矩阵是由作为非线性特征映射之特征空间的再生核希尔伯特空间中的内积所定义的格拉姆矩阵。在此背景下，核方法的渐近理论通常会与

*日本东京 184-8585，法政大学计算机与信息科学学部（aishima@hosei.ac.jp）。

积分核逼近理论 [11, 24]结合起来讨论。此外，在统计模型下，高斯核矩阵的渐近结构通常会结合一些应用进行分析 [34, 35]。对于含噪数据，通常使用反卷积核密度估计量来有效去除随机误差 [13, 28]。在应用研究方面，也有专门研究含有缺失值的数据中高斯核的工作 [6, 33]。在众多研究中，本研究与 Karoui [23]的工作密切相关。该工作具有开创性，重点关注谱分析，具体证明了渐近区域内高斯核矩阵特征向量的一致性。在此基础上，关于谱分布及其在聚类中的应用的研究得到了积极推进 [4, 10, 14]。

本研究旨在将 Karoui 的渐近分析 [23] 与约束低秩矩阵逼近方法相结合。与 Karoui 的表述类似，观测向量 $\boldsymbol{x}_1, \ldots\ldots, \boldsymbol{x}_n \in \mathbb{R}^d$ 受到随机噪声污染。在温和假设下，由于核矩阵具有渐近矩阵结构，其特征向量是相合估计量。基于这一视角，我们为秩亏情形构造了真实核矩阵的相合估计量。本文的主要贡献是推进上述渐近分析，并针对一种更复杂、表示部分噪声的统计模型，给出一个具有强相合性的具体估计量；对于该模型，现有的朴素估计方法无法保持特征向量的相合性。

从另一个角度来看，本文的贡献涉及数值线性代数中的矩阵特征值分析和低秩矩阵近似。通过奇异值分解（SVD） [17, 38], , 可以有效求解带约束的问题 [9, 18], 从而得到矩阵的低秩近似。此类低秩近似的重要结果通常应用于总体最小二乘（TLS） [16]。由于带约束的结构化 TLS 十分重要，研究者已积极开发了许多高效的数值方法 [19, 20, 21, 22, 30], , 并对其可解条件和条件数进行了深入分析 [29, 36, 41, 42]。此外，针对低秩系数矩阵的 TLS 已得到充分研究 [15, 31, 32, 37]。值得注意的是，相关的统计渐近分析也在积极发展中 [1, 2, 8, 26, 27]。此外，核矩阵的低秩近似近年来备受关注 [3, 5, 25, 39, 40]。基于这一背景，对于高斯核矩阵，我们采用上述约束低秩矩阵逼近方法，在秩亏情形下构造了一个具有强相合性的估计量。据作者所知，这是首次将上述带约束的低秩矩阵近似方法直接应用于核方法的渐近分析。

本文结构如下。第2节介绍我们的问题设定，并在其中进行渐近分析。第3节给出高斯核矩阵的渐近分析，并以与 Karoui 的结果相同的方式证明特征向量的强相合性，同时 [23], 重新构造了一个适用于秩亏情形的核矩阵相合估计量。第4节介绍了在表示部分噪声的统计模型下渐近分析的主要结果，并由此得到一个在秩亏情形下具有强相合性的新估计量。最后，第5节给出结论。

2 高维情形下相合性分析的问题设定

基于上一节的研究背景，我们构建一个噪声模型来构造估计量。基本上，我们的表述是对 Karoui [23]所用表述的简化。在 [23], 中讨论了一般的核矩阵，其中核矩阵的 (i, j) 元素由 $f(\|\boldsymbol{x}_i - \boldsymbol{x}_j\|_2{}^2)$ 给出，其中 f 属于某个特定类别 $\mathcal{F}$。在这类渐近分析中，由于通常会考虑目标函数 G 的 $\sup_{f\in\mathcal{F}} G(f)$，因此需要外概率陈述。尽管这对严谨分析而言并非问题，但为提高可读性，我们将问题重新表述为仅限于高斯核矩阵的情形，从而得到简洁的渐近分析。

下面，我们考虑随机的 d 维向量 $\boldsymbol{x}_i \in \mathbb{R}^d (i = 1, \ldots, n)$，其定义如下：

$$\boldsymbol{x}_i = \boldsymbol{s}_i + \boldsymbol{\xi}_i \qquad (i = 1, \ldots, n), \tag{2}$$

其中， $\boldsymbol{s}_i, \boldsymbol{\xi}_i \in \mathbb{R}^d (i = 1, \ldots, n)$分别是重要的信号向量和随机噪声向量。对于信号向量 $\boldsymbol{s}_i \in \mathbb{R}^d (i = 1, \ldots, n)$，我们作如下假设。

假设1。 对于整数 $i, j\,(1 \le i, j \le n)$，假设$\lim_{d\to\infty} d^{-1} \boldsymbol{s}_i{}^\top \boldsymbol{s}_j$ 存在，且

$$\lim_{d\to\infty} \frac{\|\boldsymbol{s}_i\|^2}{d} > 0 \qquad (i = 1, \ldots, n). \tag{3}$$

如果我们考虑向量 $\boldsymbol{s}_1 \ldots, \boldsymbol{s}_n$ 由 L^2 函数的离散近似得到的情形，那么假设1中的上述条件是自然的。目的是利用 $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$，通过某些估计量得到 $K(\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n)$，因此我们将在下文对这些估计量进行渐近分析。

为了对随机向量进行统计渐近分析，我们先解释一些一般的统计术语和概念。一般来说，随机变量序列 $\mathcal{X}^{(0)}$ $\mathcal{X}^{(1)}, \mathcal{X}^{(2)}, \ldots$ 以概率为1收敛到随机变量 $\mathcal{X}$，当且仅当

$$\mathcal{P}(\{\omega \in \Omega \mid \lim_{m\to\infty} \mathcal{X}^{(m)}(\omega) = \mathcal{X}(\omega)\}) = 1,$$

其中，$\mathcal{P}$ 是样本空间 Ω 上的一个概率测度。强收敛指以概率1收敛。为简洁起见，本文省略随机变量自变量 ω。设 $\widehat{\mathcal{X}}^{(m)}$ 表示利用大小为 m 的样本对目标常数值 θ 的估计量。如果 $\widehat{\mathcal{X}}^{(m)}$ 以概率1收敛到 θ，则称 $\widehat{\mathcal{X}}^{(m)}$ 是一个具有强相合性的估计量。

以下渐近分析讨论维数 $d \to \infty$ 趋于极限的情形，其中随机向量的各元素被视为相应的序列。注意，$\mathbb{R}^d$ 中向量的平方二范数是其 d 个元素的平方和。下文将在以下假设下分析高斯核的相合性。

假设2。假设

$$\lim_{d\to\infty} \frac{\|\boldsymbol{\xi}_i\|^2}{d} = \bar{\sigma}^2 \qquad (1 \le i \le n), \tag{4}$$

$$\lim_{d\to\infty} \frac{\boldsymbol{\xi}_i{}^\top \boldsymbol{\xi}_j}{d} = 0 \qquad (1 \le i \ne j \le n), \tag{5}$$

$$\lim_{d\to\infty} \frac{\boldsymbol{s}_i{}^\top \boldsymbol{\xi}_j}{d} = 0 \qquad (1 \le i, j \le n) \tag{6}$$

以概率*1*。

关于上述假设中的强收敛，通常假设所有 n 个均值向量 $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n$ 均为 $\mathbf{0} \in \mathbb{R}^d$。此外，如果所有向量 $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n$ 的所有元素均独立同分布，且方差有界，则根据强大数定律（SLLN），式（4）、（5）和（6）以概率1成立。事实上，对于独立同分布的情形，收敛到假设2中所述结果并不要求分布相同。

命题1。假设 $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n$ 相互独立且同分布（*i.i.d.*），并且所有 n 均值向量均为 $\mathbf{0} \in \mathbb{R}^d$。令 ${\sigma_1}^2, \ldots, {\sigma_d}^2$ 表示相应随机向量各元素的二阶矩，并假设当 $d \to \infty$ 时，${\sigma_1}^2, \ldots, {\sigma_d}^2$ 一致有界。此外，假设 $\lim_{d\to\infty} \sum_{i=1}^d d^{-1}{\sigma_i}^2$ 存在，并令 $\bar{\sigma}^2 = \lim_{d\to\infty} \sum_{i=1}^d d^{-1}{\sigma_i}^2$。此外，当 $d \to \infty$ 时，所有随机向量各元素的四阶矩一致有界。则 (4)、(5) 和 (6) 以概率*1*成立。

证明。若方差一致有界，则可由柯尔莫哥洛夫理论证明强大数定律（SLLN）；详情见 [12, §2] 。因此，(4) 和 (5) 以概率1成立。对于 (6)，我们使用阿贝尔分部求和公式。对于任意 i, j，将 $s_{i,1}, \ldots, s_{i,d}$ 和 $\xi_{j,1}, \ldots, \xi_{j,d}$ 分别定义为 $\boldsymbol{s}_i$ 和 $\boldsymbol{\xi}_j$ 的元素。此外，定义

$$t_{i,0} = 0, \quad t_{i,d'} = \sum_{k=1}^{d'} {s_{i,k}}^2 \quad (d' = 1, \ldots, d),$$

从而导致

$$\lim_{d\to\infty} d^{-1} t_{i,d} < \infty \tag{7}$$

根据 $d^{-1}\|\boldsymbol{s}_i\|^2$ 有界这一假设。此外，注意到

$$\sum_{k=1}^{d} k^{-2} {s_{i,k}}^2 = \sum_{k=1}^{d} k^{-2}(t_{i,k} - t_{i,k-1}) = \sum_{k=1}^{d} k^{-2} t_{i,k} - \sum_{k=0}^{d-1} (k+1)^{-2} t_{i,k} = d^{-2} t_{i,d} + \sum_{k=1}^{d-1} (k^{-2} - (k+1)^{-2}) t_{i,k}.$$

上述方程右侧的第二项为 $\sum_{k=1}^{d-1} (k^{-1}(k+1)^{-2}(2k+1)) k^{-1} t_{i,\,k}$。于是，$\sum_{k=1}^{d-1} k^{-1}(k+1)^{-2}(2k+1) < \infty$ 得出 (7) $\sum_{k=1}^{d} k^{-2} {s_{i,\,k}}^2 < \infty$。因此，根据假设 ${\sigma_1}^2, \ldots, {\sigma_d}^2$ 一致有界，我们得到

$$\lim_{d\to\infty} \sum_{k=1}^{d} k^{-2} \mathbb{E}((s_{i,k}\xi_{j,k})^2) = \lim_{d\to\infty} \left(\sup_{1\le k\le d} {\sigma_k}^2 \right) \sum_{k=1}^{d} k^{-2} {s_{i,k}}^2 < \infty,$$

其中 $\mathbb{E}$ 表示期望值。注意到 [7, 式 (7) 中的推论（定理 5.4）1], ，我们得到

$$\lim_{d\to\infty}\frac{\sum_{k=1}^{d}s_{i,k}\xi_{j,k}}{d}=0$$

以概率为1成立，从而以概率为1推出（6）。 □

根据近期关于强大数定律（SLLN）的渐近分析，在更弱的条件下或许可以建立假设2中的强收敛。由于本文聚焦于高斯核的渐近分析，我们将不再进一步考察随机变量序列的强大数定律，而是讨论假设2下的收敛。

3 高维含噪数据的高斯核矩阵的强相合估计量

本节将给出高斯核的渐近分析结果，形式为 $d\to\infty$ 。更具体地说，我们证明了在假设 1 和假设 2 下，高斯核矩阵的特征向量具有相合性。我们的渐近分析表明，核矩阵及其特征值是不相合的。上述渐近分析实质上是对 [23, §2.3.2] 的复现。不过，该分析专门针对高斯核，并在上一节中清晰地重述了相关条件，以便后续分析，同时阐明了后续分析所需的要点。更具体地说，基于这些结果，我们提出了一种修正的核矩阵，它是秩亏情形下真实核矩阵的相合估计量。下一节将把这一相合性分析应用于更一般的噪声模型。因此，这项分析被视为一项重要的基础分析。

对于渐近分析，回顾一下，高斯核矩阵 $K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 的定义如式 (1) 所示。对于缩放参数 c_d，我们作如下假设。

假设 3。 极限$\lim_{d\to\infty}d^{-1}c_d$ 存在，令

$$\lim_{d\to\infty}\frac{c_d}{d}=\gamma>0. \tag{8}$$

注意，由于在假设 1 和假设 2 下，核矩阵 $K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 由式 (1) 针对 n 个样本向量 $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n\in\mathbb{R}^d$ 给出，因此，假设 3 对于归一化 $K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 是合理的。

3.1 渐近分析：不变子空间的强相合性

这里我们证明，当 $d\to\infty$ 时，式 (1) 中 $K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 的特征向量具有强相合性。下一个引理是理解核矩阵 $K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 和 $K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 的收敛性质的关键。

引理 1。 在假设 *1* 和假设 *3* 下，极限$\lim_{d\to\infty}K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 存在。令 i、j 表示满足 $1\le i$、$j\le n$ 和 $i\neq j$ 的整数。那么，在假设 *2* 下，我们有

$$\lim_{d\to\infty}\left[K(\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_n)\right]_{ij}=\exp\left(-2\gamma^{-1}\bar{\sigma}^2\right)\quad \textit{with probability one.} \tag{9}$$

此外，我们有

$$\lim_{d\to\infty}\left[K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)\right]_{ij}=\exp\left(-2\gamma^{-1}\bar{\sigma}^2\right)\lim_{d\to\infty}\left[K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)\right]_{ij}\quad \textit{with probability one.} \tag{10}$$

作为矩阵表示，

$$\lim_{d\to\infty}K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)=\lim_{d\to\infty}K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)\odot\lim_{d\to\infty}K(\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_n)\quad \textit{with probability one,} \tag{11}$$

其中，$\odot$表示两个矩阵的哈达玛积。

证明。 首先，显然，在假设 1 和 3 下，$\lim_{d\to\infty}K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 存在。接下来，对于 (9)，我们看到

$$\left[K(\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_n)\right]_{ij}=\exp\left(-\frac{d}{c_d}\frac{\|\boldsymbol{\xi}_i-\boldsymbol{\xi}_j\|^2}{d}\right)=\exp\left(-\frac{d}{c_d}\frac{\|\boldsymbol{\xi}_i\|^2+\|\boldsymbol{\xi}_j\|^2+\boldsymbol{\xi}_i{}^{\top}\boldsymbol{\xi}_j}{d}\right). \tag{12}$$

因此，在假设 2 和 3 下，由 (4)、(5) 和 (8) 可得 (9)。对于 (10)，我们有

$$\begin{aligned}\exp\left(-\frac{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|^2}{c_d}\right) &= \exp\left(-\frac{\|\boldsymbol{s}_i-\boldsymbol{s}_j+\boldsymbol{\xi}_i-\boldsymbol{\xi}_j\|^2}{c_d}\right)\\ &= \exp\left(-\frac{\|\boldsymbol{s}_i-\boldsymbol{s}_j\|^2+2(\boldsymbol{s}_i-\boldsymbol{s}_j)^\top(\boldsymbol{\xi}_i-\boldsymbol{\xi}_j)+\|\boldsymbol{\xi}_i-\boldsymbol{\xi}_j\|^2}{c_d}\right)\\ &= \exp\left(-\frac{\|\boldsymbol{s}_i-\boldsymbol{s}_j\|^2+2(\boldsymbol{s}_i-\boldsymbol{s}_j)^\top(\boldsymbol{\xi}_i-\boldsymbol{\xi}_j)+\|\boldsymbol{\xi}_i\|^2-2\boldsymbol{\xi}_i{}^\top\boldsymbol{\xi}_j+\|\boldsymbol{\xi}_j\|^2}{c_d}\right).\end{aligned}$$

注意到，在假设1下，$\lim_{d\to\infty} d^{-1}\|\boldsymbol{s}_i-\boldsymbol{s}_j\|^2$ 存在。因此，在假设2和假设3下，我们有

$$\lim_{d\to\infty}\exp\left(-\frac{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|^2}{c_d}\right)=\exp\left(-2\gamma^{-1}\bar{\sigma}^2\right)\lim_{d\to\infty}\exp\left(-\frac{\|\boldsymbol{s}_i-\boldsymbol{s}_j\|^2}{c_d}\right).$$

因此，我们得到（10）。

对于（11）中的矩阵表示，注意到高斯核矩阵的对角元素为1，由（9）和（10）可得（11）。

□

下述定理中给出的核矩阵收敛表示十分重要。

定理 1. 在假设 *1* 和 *3* 下，$\lim_{d\to\infty} K(\boldsymbol{s}_1、\ldots、\ \boldsymbol{s}_n)$ 存在。此外，在假设*2*下，我们有

$$\lim_{d\to\infty} K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)=\exp\left(-2\gamma^{-1}\bar{\sigma}^2\right)\left(\lim_{d\to\infty} K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)-I_n\right)+I_n \quad \textit{with probability one.} \tag{13}$$

证明。注意到（10）和 $[K(\boldsymbol{x}_1、\ldots、\ \boldsymbol{x}_n)]_{ii}=1\,(i=1、\ldots、\ n)$，我们得到（13）。 □

上述定理 阐明了核矩阵特征对的重要性质。为说明细节，令

$$\lambda_1\le\lambda_2\le\cdots\le\lambda_n \tag{14}$$

表示 $K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 的特征值。同样，令

$$\mu_1\le\mu_2\le\cdots\le\mu_n \tag{15}$$

表示 $K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 的特征值。那么，定理 1 中的 (13) 给出了以下结论。

推论1. 在与定理 *1* 相同的假设下，我们有

$$\lim_{d\to\infty}\lambda_i=\exp(-2\gamma^{-1}\bar{\sigma}^2)\lim_{d\to\infty}\mu_i-\exp(-2\gamma^{-1}\bar{\sigma}^2)+1\quad(i=1,\ldots,n)\quad \textit{with probability one.} \tag{16}$$

此外，特征向量 *o* *f* $K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 是特征向量 $\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n$ 的相合估计量，其中极限为 $\lim_{d\to\infty} K($ $)$.

下面说明特征向量相合性的含义。特征向量的大小可以任意取值；当特征值重数大于 1 时，它们可以是属于相应子空间的任意向量。由于我们考虑的是对称矩阵，重数为 k 的特征值对应的特征空间维数也为 k。因此，由于特征向量的连续性，在渐近区域中，当 $d\to\infty$ 时，$K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 的任意特征向量都等于 $K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 的某个特征向量。

3.2 秩亏矩阵重构的修正估计量 s

下面我们将超越 Karoui 的结果 [23, §2.3.2]，进入本文全新的核心内容。基于上述渐近分析，我们针对秩亏情形下 $\lim_{d\to\infty} K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 的重构提出一种修正估计量。因此，我们作如下假设。

假设 4. $\lim_{d\to\infty} K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 是一个低秩矩阵。

在进行渐近分析之前，我们先解释一种 $\lim_{d\to\infty} K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 为低秩矩阵的情形。例如，假设

$$\lim_{d\to\infty} \frac{\|\boldsymbol{s}_i-\boldsymbol{s}_j\|^2}{d}=0 \tag{17}$$

对于某些固定的 i 和 j，那么 $\lim_{d\to\infty} K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 中对应的两行和两列相同，这意味着 $\lim_{d\to\infty} K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 是低秩的。具体来说，对于任意满足 $1\le k\le n$ 的整数 k，我们有

$$\lim_{d\to\infty}\frac{\|\boldsymbol{s}_i-\boldsymbol{s}_k\|^2}{d}=\lim_{d\to\infty}\frac{\|\boldsymbol{s}_i-\boldsymbol{s}_k-\boldsymbol{s}_j+\boldsymbol{s}_j\|^2}{d}=\lim_{d\to\infty}\frac{\|\boldsymbol{s}_j-\boldsymbol{s}_k\|^2+\|\boldsymbol{s}_i-\boldsymbol{s}_j\|^2+2(\boldsymbol{s}_j-\boldsymbol{s}_k)^\top(\boldsymbol{s}_i-\boldsymbol{s}_j)}{d}=\lim_{d\to\infty}\frac{\|\boldsymbol{s}_j-\boldsymbol{s}_k\|^2}{d},$$

其中，最后一个等式是由于

$$\lim_{d\to\infty}\left|\frac{(\boldsymbol{s}_j-\boldsymbol{s}_k)^\top(\boldsymbol{s}_i-\boldsymbol{s}_j)}{d}\right|\le\lim_{d\to\infty}\sqrt{\frac{\|\boldsymbol{s}_j-\boldsymbol{s}_k\|^2}{d}}\sqrt{\frac{\|\boldsymbol{s}_i-\boldsymbol{s}_j\|^2}{d}}=0$$

根据柯西–施瓦茨不等式。因此，$\lim_{d\to\infty} K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 中的两行和两列相同。下文将在渐近区域中考虑秩亏情形。

由于 $K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 的特征值定义如式 (15) 所示，最小特征值的极限值为

$$\lim_{d\to\infty}\mu_1=0 \tag{18}$$

在假设 4 下。然后，在满足式 (18) 中假设的条件下，推论1表明

$$\lim_{d\to\infty}\lambda_1=\exp(-2\gamma^{-1}\bar{\sigma}^2)\lim_{d\to\infty}\mu_1-\exp(-2\gamma^{-1}\bar{\sigma}^2)+1=-\exp(-2\gamma^{-1}\bar{\sigma}^2)+1, \tag{19}$$

这是 $\lim_{d\to\infty} K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 的最小特征值。

注意到上述收敛特性，我们构造 $\widetilde{K}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 对于对角元素，$[\widetilde{K}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)]_{ii}=[K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)]_{ii}(=1)$ $(i=1,\ldots,n)$。此外，令

$$\text{for } i\neq j,\quad \left[\widetilde{K}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)\right]_{ij}=(1-\lambda_1)^{-1}\left[K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)\right]_{ij}.$$

换言之，$\widetilde{K}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 定义为

$$\widetilde{K}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)=(1-\lambda_1)^{-1}(K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)-I_n)+I_n. \tag{20}$$

该 p s. 计算所提估计量 $\widetilde{K}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 的过程总结如下

[所提估计量]

- 按照式 (1) 计算高斯核 $K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 的格拉姆矩阵
- 计算 $K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 的最小特征值 λ_1
- 将式 (20) 中的 $K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 作为估计量进行计算

那么，$\widetilde{K}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 具有如下强相合性。

定理 2. 在假设 *1* 和 *3* 下，$\lim_{d\to\infty} K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 存在。此外，在假设 *2* 和 *4* 下，我们有

$$\lim_{d\to\infty}\widetilde{K}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)=\lim_{d\to\infty}K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)\quad \textit{with probability one}, \tag{21}$$

其中 $\widetilde{K}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 的定义同式 (20)。

证明。 由定理 1 中的式 (13)，式 (21) 可由式 (19) 和 (20) 直接推得。 □

注 1. 由于$\lim_{d\to\infty} d^{-1}\|\boldsymbol{s}_i-\boldsymbol{s}_j\|=0$ 在(17) 中的情形极为特殊，因此可能有人会质疑讨论秩亏情形的必要性。考虑降秩情形的一个原因是，近年来使用低秩逼近核的方法日益受到关注。因此，如果存在一种适用于精确降秩情形的一般估计方法，就有可能将其用于低秩近似的误差分析。另一个原因是，当 $n\to\infty$ 时，核矩阵可能近似低秩，其最小奇异值趋于 *0*，不过本文不讨论这类渐近情形。本文讨论固定 n 下的估计量。下一节将给出一个适用于表示部分噪声的统计模型的相合估计量。

4 基于渐近分析的部分噪声模型拟议的相合估计量

In t本节讨论下面描述的扩展噪声模型。对于 $1 \le \ell \le n-1$，le t

$$\boldsymbol{x}_i = \boldsymbol{s}_i \quad (i = 1, \dots, \ell), \qquad \boldsymbol{x}_i = \boldsymbol{s}_i + \boldsymbol{\xi}_i \quad (i = \ell+1, \dots, n). \tag{22}$$

以下是假设2的更新版本，反映了上述噪声模型。

假设5。假设

$$\lim_{d\to\infty} \frac{\|\boldsymbol{\xi}_i\|^2}{d} = \bar{\sigma}^2 \qquad (\ell+1 \le i \le n), \tag{23}$$

$$\lim_{d\to\infty} \frac{\boldsymbol{\xi}_i{}^\top \boldsymbol{\xi}_j}{d} = 0 \qquad (\ell+1 \le i \ne j \le n), \tag{24}$$

$$\lim_{d\to\infty} \frac{\boldsymbol{s}_i{}^\top \boldsymbol{\xi}_j}{d} = 0 \qquad (1 \le i \le n, \ell+1 \le j \le n) \tag{25}$$

概率为*1*。

目的是在 $d \to \infty$的渐近区域内，使用 $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$估计 $K(\boldsymbol{s}_1, \dots, \boldsymbol{s}_n)$，其中$\lim_{d\to\infty} K(\boldsymbol{s}_1, \dots, \boldsymbol{s}_n)$是一个低秩矩阵。为此，我们分析 [9]中的低秩近似优化方法。尽管已有大量研究遵循 [9], ，本文旨在阐明其基本原理。因此，我们聚焦于最基础的优化方法，更为详细的研究将留待未来开展。

4.1 基于高斯消元的低秩近似现有优化方法

为构造相合估计量，我们采用一种用于低秩矩阵近似的优化方法。为此，将 $K(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ 划分为

$$K(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) = \begin{bmatrix} K_{11}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) & K_{12}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \\ K_{21}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) & K_{22}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \end{bmatrix}, \tag{26}$$

$$K_{11}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \in \mathbb{R}^{\ell\times\ell}, \quad K_{12}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \in \mathbb{R}^{\ell\times m}, \quad K_{21}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \in \mathbb{R}^{m\times\ell}, \quad K_{22}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \in \mathbb{R}^{m\times m}, \tag{27}$$

其中 $\ell + m = n$。注意

$$K(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) = K(\boldsymbol{s}_1, \dots, \boldsymbol{s}_\ell, \boldsymbol{x}_{\ell+1}, \dots, \boldsymbol{x}_n)$$

根据（22）中的噪声模型，且 $K_{11}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \in \mathbb{R}^{\ell\times\ell}$ 是由 $\boldsymbol{s}_1, \dots, \boldsymbol{s}_\ell$ 给出的无噪声矩阵。注意到该矩阵结构，我们考虑以下优化问题：

$$\min_{\Delta\in\mathbb{R}^{m\times m}} \|\Delta\|_{\mathrm{F}} \quad \text{such that} \quad \operatorname{rank}\left(\begin{bmatrix} K_{11}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) & K_{12}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \\ K_{21}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) & K_{22}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) + \Delta \end{bmatrix} \right) \le n-1, \tag{28}$$

其中 $\|\cdot\|_{\mathrm{F}}$ 表示 Frobenius范数。为求解上述优化问题，假设 $K_{11}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ 非奇异。然后，与高斯消元类似，

$$\begin{aligned} &\begin{bmatrix} K_{11}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) & K_{12}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \\ K_{21}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) & K_{22}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \end{bmatrix} - \begin{bmatrix} K_{11}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \\ K_{21}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \end{bmatrix} \begin{bmatrix} I_\ell & K_{11}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)^{-1} K_{12}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \end{bmatrix} \\ &= \begin{bmatrix} 0_{\ell\times\ell} & 0_{\ell\times m} \\ 0_{m\times\ell} & K_{22}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) - K_{21}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) K_{11}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)^{-1} K_{12}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \end{bmatrix}. \end{aligned}$$

因此，（28）中的条件等价于

$$\operatorname{rank}\Big(K_{22}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) - K_{21}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) K_{11}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)^{-1} K_{12}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) + \Delta \Big) \le m-1. \tag{29}$$

上述低秩近似问题可通过基于 Eckart-Young-Mirsky定理的奇异值分解（SVD）求解。由于上述矩阵是对称矩阵，绝对值最小的特征值就是最优的 $\|\Delta\|_{\mathrm{F}}$。基于这一点，我们进行渐近分析，然后构造一个相合估计量。

4.2 优化方法中矩阵的渐近分析

对于渐近分析，与 $K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 类似，将 $K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 划分为

$$K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)=\begin{bmatrix}K_{11}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n) & K_{12}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)\\ K_{21}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n) & K_{22}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)\end{bmatrix}, \tag{30}$$

$$K_{11}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)\in\mathbb{R}^{\ell\times\ell},\quad K_{12}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)\in\mathbb{R}^{\ell\times m},\quad K_{21}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)\in\mathbb{R}^{m\times\ell},\quad K_{22}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)\in\mathbb{R}^{m\times m}. \tag{31}$$

根据式（22）中的噪声模型，我们有

$$K_{11}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)=K_{11}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n).$$

在假设1下，$\lim_{d\to\infty}K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 存在，记为

$$\lim_{d\to\infty}K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)=K^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)=\begin{bmatrix}K_{11}^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n) & K_{12}^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)\\ K_{21}^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n) & K_{22}^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)\end{bmatrix}. \tag{32}$$

下一引理中的核矩阵 $K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 的渐近结构对于后续的渐近分析至关重要且具有基础性作用。

引理2. 在假设 *1* 和 *3* 下，$\lim_{d\to\infty}K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 存在。此外，在假设*5*下，利用式(32)中的定义，我们有

$$\lim_{d\to\infty}K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)=\begin{bmatrix}K_{11}^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n) & \exp\left(-\gamma^{-1}\bar{\sigma}^2\right)K_{12}^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)\\ \exp\left(-\gamma^{-1}\bar{\sigma}^2\right)K_{21}^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n) & \exp\left(-2\gamma^{-1}\bar{\sigma}^2\right)(K_{22}^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)-I_m)+I_m\end{bmatrix} \tag{33}$$

概率为*1*。

证明。 显然，由 $\boldsymbol{x}_i=\boldsymbol{s}_i(i=1,\ldots,\ell)$可得 $\lim_{d\to\infty}K_{11}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)=K_{11}^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$。此外，由定理 1 可知

$$\lim_{d\to\infty}K_{22}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)=\exp\left(-2\gamma^{-1}\bar{\sigma}^2\right)(K_{22}^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)-I_m)+I_m\quad\text{with probability one.}$$

类似地，{v1}的 $K_{12}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$的强收敛证明如下。对于 $1\le i\le\ell,\ell+1\le j\le n$ ，

$$\exp\left(-\frac{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|^2}{c_d}\right)=\exp\left(-\frac{\|\boldsymbol{s}_i-\boldsymbol{s}_j-\boldsymbol{\xi}_j\|^2}{c_d}\right)=\exp\left(-\frac{\|\boldsymbol{s}_i-\boldsymbol{s}_j\|^2-2(\boldsymbol{s}_i-\boldsymbol{s}_j)^\top\boldsymbol{\xi}_j+\|\boldsymbol{\xi}_j\|^2}{c_d}\right).$$

注意，在假设1下，$\lim_{d\to\infty}d^{-1}\|\boldsymbol{s}_i-\boldsymbol{s}_j\|^2$ 存在。因此，在假设2和假设3下，我们有

$$\begin{aligned}\lim_{d\to\infty}\exp\left(-\frac{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|^2}{c_d}\right)&=\lim_{d\to\infty}\exp\left(-\frac{d}{c_d}\frac{\|\boldsymbol{s}_i-\boldsymbol{s}_j\|^2-2(\boldsymbol{s}_i-\boldsymbol{s}_j)^\top\boldsymbol{\xi}_j+\|\boldsymbol{\xi}_j\|^2}{d}\right)\\&=\exp\left(-\gamma^{-1}\bar{\sigma}^2\right)\lim_{d\to\infty}\exp\left(-\frac{\|\boldsymbol{s}_i-\boldsymbol{s}_j\|^2}{d}\right).\end{aligned}$$

因此，$\lim_{d\to\infty}K_{12}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)=\exp\left(-\gamma^{-1}\bar{\sigma}^2\right)K_{12}^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$，且该结论以概率为1成立。由格拉姆矩阵的对称性，$\lim_{d\to\infty}K_{21}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)=\exp\left(-\gamma^{-1}\bar{\sigma}^2\right)K_{21}^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$，且该结论以概率为1成立。证毕。

□

为应用式 (29) 中的优化问题，我们作如下假设。

假设6。 $K_{11}^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 在(32) 中是非奇异的。

假设6意味着， $K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$的无噪声子矩阵（由 $\boldsymbol{s}_1,\ldots,\boldsymbol{s}_\ell$给出）是非奇异的。下一个定理对于构造相合估计量至关重要，其证明可由引理2中的（33）通过简单计算得到。

定理 3。 在假设 *1* 和 *3* 下，$\lim_{d\to\infty} K$（$\boldsymbol{s}_1$，.....， $\boldsymbol{s}_n$）存在。此外，在假设 *5* 和 *6* 下，利用（32）中的定义，我们有

$$\lim_{d\to\infty} \left(K_{22}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n) - K_{21}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)K_{11}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)^{-1}K_{12}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)\right)$$
$$= \exp\left(-2\gamma^{-1}\bar{\sigma}^2\right)\left(K_{22}^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n) - K_{21}^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)K_{11}^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)^{-1}K_{12}^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n) - I_m\right) + I_m \quad (34)$$

概率为*1*。

重要的是，上述定理对应于定理 1。基于上述定理，令

$$\tau_1 \le \cdots \le \tau_m \quad (35)$$

表示 $K_{22}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n) - K_{21}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)K_{11}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)^{-1}K_{12}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 的特征值。类似地，令

$$\nu_1 \le \cdots \le \nu_m \quad (36)$$

表示 $K_{22}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n) - K_{21}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)K_{11}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)^{-1}K_{12}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 的特征值。

作为定理 3 的推论，我们得到以下结果，对应于推论1 .

推论 2. 在与定理 *3* 相同的假设下，我们有

$$\lim_{d\to\infty} \tau_i = \exp(-2\gamma^{-1}\bar{\sigma}^2) \lim_{d\to\infty} \nu_i - \exp(-2\gamma^{-1}\bar{\sigma}^2) + 1 \quad (i = 1,\ldots,n) \quad \textit{with probability one.} \quad (37)$$

K_{22}（$\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n$） $- K_{21}$（$\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n$） K_{11}（$\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n$） $^{-1}K_{12}$（$\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n$） 是 $K_{22}^{(\infty)}$（$\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n$） $- K_{21}^{(\infty)}$（$\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n$） $K_{11}^{(\infty)}$（$\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n$） $^{-1}K_{12}^{(\infty)}$（$\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n$） 中特征向量的相合估计量。

4.3 秩亏情形的拟议相合估计量

最后，我们根据上述讨论构造一个新的估计量。更具体地说，与上一节一样，在 $K^{(\infty)}$($\boldsymbol{s}_1$，.. . ， $\boldsymbol{s}_n$)的秩亏情形下，我们提出一个修正矩阵，该矩阵在渐近区域 $d \to \infty$中对式（22）中的噪声模型具有强相合性。

为此，假设 $K^{(\infty)}$($\boldsymbol{s}_1$，...， $\boldsymbol{s}_n$) 是一个低秩矩阵，如假设 4 所述。那么，在假设6下， K_{22}($\boldsymbol{s}_1$，...， $\boldsymbol{s}_n$) $- K_{21}$($\boldsymbol{s}_1$，...， $\boldsymbol{s}_n$)K_{11}($\boldsymbol{s}_1$，...， $\boldsymbol{s}_n$)$^{-1}K_{12}$($\boldsymbol{s}_1$，...， $\boldsymbol{s}_n$) 的最小特征值的极限值为

$$\lim_{d\to\infty} \nu_1 = 0,$$

其中，特征值 按照 (36) 中的定义按升序排列。然后，根据 (35) 中的定义，我们有

$$\lim_{d\to\infty} \tau_1 = -\exp(-2\gamma^{-1}\bar{\sigma}^2) + 1 \quad \text{with probability one} \quad (38)$$

由于推论 2 中的收敛性质 (37) 结合引理2，我们提出一种修正后的估计量 r

$$\widetilde{K}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n) = \begin{bmatrix} K_{11}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n) & (1-\tau_1)^{-1/2}K_{12}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n) \\ (1-\tau_1)^{-1/2}K_{21}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n) & (1-\tau_1)^{-1}(K_{22}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n) - I_m) + I_m \end{bmatrix}. \quad (39)$$

~所提估计量 K($\boldsymbol{x}_1$，...， $\boldsymbol{x}_n$) 的计算步骤总结如下。

[所提估计量]

- 按照式 (1) 计算高斯核的格拉姆矩阵 K($\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n$)
- 使用按照式 (26) 和 (27) 定义的子矩阵，计算 K_{22}($\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n$) $-K_{21}$($\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n$)$K_{11}$($\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n$)$^{-1}K_{12}$($\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n$)~的最小特征值 τ_1
- 计算式 (39) 中的 K($\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n$)，并将其作为估计量

上述渐近分析直接导出下一个定理，该定理陈述了所提估计量 ~proposed estimator K($\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n$) 的强相合性。

定理 4. 在假设 *1* 和 *3* 下，$\lim_{d\to\infty} K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 存在。此外，在假设 *4*、*5* 和 *6* 下，使用 (32) 中的定义，我们有

$$\lim_{d\to\infty} \widetilde{K}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n) = \lim_{d\to\infty} K(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n) \quad \textit{with probability one}, \tag{40}$$

其中 $\widetilde{K}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 的定义与(39)中相同。

证明。 注意到引理2中的 (33)，由 (38) 和 (39) 可得 (40)。 □

注 2. 如注 *1* 所述，秩亏情形很重要。可以注意到，如果 $K_{22}^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$是低秩的，那么$\exp(-2\gamma^{-1}\bar{\sigma}^2)$ 可以像上一节那样，通过 $K_{22}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$的最小特征值进行估计。本节提出的估计量适用于 $K_{22}^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$的满秩情形。对于有限的 n，为了保持 $K_{22}^{(\infty)}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$的满秩情形，当 $1\le i\le \ell$ 和 $\ell+1\le j\le n$ 时，秩亏情形发生于$\lim_{d\to\infty} d^{-1}\|\boldsymbol{s}_i-\boldsymbol{s}_j\|^2=0$。此时， $K(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 的 i^{th} 和 j^{th} 列向量在 $d\to\infty$ 时趋于平行，因此$\exp(-2\gamma^{-1}\bar{\sigma}^2)$ 可以通过它们模长之比进行估计。然而，本节提出了一种通用估计方法，以便今后扩展到更一般的情形，具体讨论见注 *1*。

注 3. 有些人可能会想，为什么我们要将上述优化问题(28)应用于所考虑的噪声模型(22) ，下面我们解释这一点。

这一思路源自线性核的渐近分析。令 $K_{\text{linear}}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 表示由线性核给出的格拉姆矩阵，使得

$$[K_{\text{linear}}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)]_{ij} = \boldsymbol{x}_i{}^{\top}\boldsymbol{x}_j \qquad (1\le i,j\le n). \tag{41}$$

对于归一化，我们考虑$\lim_{d\to\infty} d^{-1}K_{\text{linear}}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$，它在假设*1*和假设*5*下具有有界元素。此外，注意到$\lim_{d\to\infty} d^{-1}K_{\text{linear}}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 存在。因此，将 $K_{\text{linear}}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$ 分为

$$K_{\text{linear}}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n) = \begin{bmatrix} K_{\text{linear},11}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n) & K_{\text{linear},12}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n) \\ K_{\text{linear},21}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n) & K_{\text{linear},22}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n) \end{bmatrix},$$

$K_{\text{linear},11}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)\in\mathbb{R}^{\ell\times\ell}$, $K_{\text{linear},12}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)\in\mathbb{R}^{\ell\times m}$, $K_{\text{linear},21}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)\in\mathbb{R}^{m\times\ell}$, $K_{\text{linear},22}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)\in\mathbb{R}^{m\times m}$.

那么，在假设*1*和假设*5*下，我们有

$$\lim_{d\to\infty} d^{-1}K_{\text{linear}}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n) = \begin{bmatrix} \lim_{d\to\infty} d^{-1}K_{\text{linear},11}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n) & \lim_{d\to\infty} d^{-1}K_{\text{linear},12}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n) \\ \lim_{d\to\infty} d^{-1}K_{\text{linear},21}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n) & \lim_{d\to\infty} d^{-1}K_{\text{linear},22}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)+\bar{\sigma}^2 I_m \end{bmatrix}$$

以概率*1*成立。对于线性核函数，这一特征对应于优化问题(28) 。

乍看之下，如(28) 所示，将高斯核矩阵直接进行朴素表述似乎毫无意义，因为引理*2*指出， $K_{12}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 与 $K_{21}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ 彼此不相容，与$d^{-1}K_{\text{linear},12}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$和$d^{-1}K_{\text{linear},21}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$不同。尽管如此，这一思路通过引理*2*中的偏差结构，导向了一种经修正的相合估计量，其重点在于定理*3*。这一技术发现可视为本文的贡献之一。

5 结论

本文对高斯核矩阵进行了渐近分析，结果如 $d\to\infty$所示。具体而言，Karoui [23] 给出的高斯核矩阵渐近结构与带约束的低秩矩阵近似研究相结合，由此得到了一个新估计量；该估计量在第4.3节所述的秩亏情形下具有强相合性。从技术角度看，克服这一困难的创新之处仅在第4节中阐述；该节将分析限定于统计模型下的高斯核，其中部分噪声由式(22)表示。然而，据作者所知，这是首项直接应用上述受约束低秩近似方法的研究。值得注意的是，该方法通过关注偏差的矩阵结构并明确构造具有强相合性的估计量，成功消除了偏差。

致谢
本研究获得JSPS KAKENHI资助项目编号JP22K03422的支持。

参考文献

[1] K. Aishima，利用正交投影对含变量误差的线性回归模型进行一致估计，《线性代数及其应用》，684（2024），101–126。[2] K. Aishima，含共线性的变量误差回归模型中基于截断奇异值分解的估计量的强相合性，《线性代数及其应用》，721（2025），520–541。[3] J. M. Altschuler 和 P. A. Parrilo，代数簇上的核近似，《SIAM应用代数与几何杂志》，7（2023），1–28。[4] A. A. Amini 和 Z. S. Razaee，核矩阵的集中性及其在核谱聚类中的应用，《统计年刊》，49（2021），531—556。[5] Y. Chen、E. N. Epperly、J. A. Tropp 和 R. J. Webber，随机主元乔列斯基分解：通过少量条目评估实现核矩阵的实用近似，《纯粹与应用数学通讯》，78（2025），995–1041。[6] J. Choi、Y. Son 和 M. K. Jeong，不完全数据中相关变量的高斯核，《运筹学年刊》，341（2024），223—244。[7] K. L. Chung，《概率论教程》，第3版，学术出版社，圣迭戈，加利福尼亚州，2001。[8] O. O. Dashkov 和 A. G. Kukush，含变量误差的隐式线性模型中正交回归估计量的一致性，《理论概率与数理统计》，97（2018），45–55。[9] J. W. Demmel，降低秩的子矩阵最小扰动与约束总体最小二乘问题，《SIAM数值分析杂志》，24（1987），199–206。[10] X. Ding 和 R. Ma，从高维含噪数据中学习低维非线性结构：一种积分算子方法，《统计年刊》，51（2023），1744—1769。[11] P. Dommel 和 A. Pichler，论核函数的近似，《机器学习研究杂志》，26（2025），第41篇，1–30。[12] R. Durrett，《概率：理论与实例》，第5版，剑桥大学出版社，剑桥，2019。[13] J. Fan，非参数反卷积问题的最优收敛速率，《统计年刊》，19（1991），1257—1272。[14] Z. Fan 和 A. Montanari，随机内积核矩阵的谱范数，《概率论及相关领域》，173（2019），27—85。[15] R. D. Fierro 和 J. R. Bunch，共线性与总体最小二乘，《SIAM矩阵分析与应用杂志》，15（1994），1167—1181。[16] G. H. Golub 和 C. F. Van Loan，总体最小二乘问题分析，《SIAM数值分析杂志》，17（1980），883–893。[17] G. H. Golub 和 C. F. Van Loan，《矩阵计算》，第4版，约翰斯·霍普金斯大学出版社，巴尔的摩，马里兰州，2013。[18] G. H. Golub、A. Hoffman 和 G. W. Stewart，Eckart-Young-Mirsky矩阵近似定理的推广，《线性代数及其应用》，88/89（1987），317–327。[19] I. Hnˇetynkov´a、M. Pleˇsinger、D. M. Sima、Z. Strakoˇs 和 S. Van Huffel，总体最小二乘问题$AX \approx B$：一种与经典著作相关的新分类，《SIAM矩阵分析与应用杂志》，32（2011），748–770。[20] S. Van Huffel 和 J. Vandewalle，在含变量误差的共线性问题中使用总体最小二乘方法进行子集选择，《线性代数及其应用》，88/89（1987），695–714。[21] S. Van Huffel 和 J. Vandewalle，广义总体最小二乘问题的分析与性质： $AX \approx B$当 A 中的部分或全部列存在误差时，《SIAM矩阵分析与应用杂志》，10（1989），294—315。

[22] S. Van Huffel 和 H. Zha，《限制总体最小二乘问题：表述、算法与性质》，《SIAM矩阵分析与应用杂志》，12（1991），292–309。[23] N. E. Karoui，关于信息加噪声核随机矩阵，《统计年刊》，38（2010），3191–3216。[24] T. Karvonen 和 Y. Suzuki，高斯核及相关解析核在希尔伯特空间中的逼近，IMA数值分析杂志，draf050，发表于：2025年7月15日。[25] A. Khan 和 A. K. Saibaba，使用张量列车分解的参数化核低秩逼近，《SIAM矩阵分析与应用杂志》，46（2025），1006–1036。[26] A. Kukush 和 S. Van Huffel，多元变量误差模型中逐元加权总体最小二乘估计量的一致性 $AX = B$，Metrika，59（2004），75—97。[27] A. Kukush、I. Markovsky 和 S. Van Huffel，多元变量误差模型中结构化总体最小二乘估计量的一致性，J. Statist. Plann. Inference，133（2005）315–358。[28] T. Li 和 Q. Vuong，使用多重指标对测量误差模型进行非参数估计，J. Multivariate Anal.，65（1998），139–165。[29] Q. Liu 和 Z. Jia，带线性等式约束的总体最小二乘问题的条件数，Numer. Algorithms，90（2022），363–385[30] Q. Liu、Z. Jia 和 Y. Wei，带线性等式约束的多维总体最小二乘问题，《SIAM矩阵分析与应用杂志》，43（2022），124–150[31] L. Meng、B. Zheng 和 Y. Wei，具有多个解的多维总体最小二乘问题的条件数，Numer. Algorithms，84（2020），887–908[32] Q. Meng、H. Diao 和 Z. Bai，截断总体最小二乘问题的条件数及其估计，Numer. Linear Algebra Appl.，28（2021），论文编号 e2369，共24页。[33] D. P. Mesquita、J. P. Gomes、F. Corona、A. H. S. Junior 和 J. S. Nobre，不完全数据的高斯核，Appl. Soft Comput.，77（2019），356—365。[34] Y. Nakayama、K. Yata 和 M. Aoshima，高维小样本设置下采用高斯核的偏差校正支持向量机，Ann. Inst. Statist. Math.，72（2020），1257–1286。[35] Y. Nakayama、K. Yata 和 M. Aoshima，高维小样本设置下采用高斯核的主成分分析聚类，J. Multivariate Anal.，185（2021），论文编号 104779，1–15。[36] C. C. Paige 和 M. Wei，广义总体最小二乘问题的分析 $AX \approx B$：当 A 的某些列无误差时，Numer. Math.，65（1993），177–202[37] S. Park 和 D. P. O'Leary，隐式加权总体最小二乘，《线性代数及其应用》，435（2011），560–577。[38] Y. Saad，《大型特征值问题的数值方法》，SIAM，费城，宾夕法尼亚州，2011。[39] R. Wang、Y. Li 和 E. Darve，高维径向基函数核的数值秩，《SIAM矩阵分析与应用杂志》，39（2018），1810—1835。[40] R. Wang、Y. Li、M. W. Mahoney 和 E. Darve，用于可扩展核函数求值的块基分解，《SIAM矩阵分析与应用杂志》，40（2019），1497–1526。[41] M. Wei，Eckart-Young-Mirsky定理与约束总体最小二乘问题的扰动理论，《线性代数及其应用》，280（1998），267–287。[42] B. Zheng、L. Meng 和 Y. Wei，多维总体最小二乘问题的条件数。《SIAM矩阵分析与应用杂志》，38（2017），924–948