arXiv:2301.04487v1 [stat.ME] 11 Jan 2023

连续函数型数据的可分离性检验

Holger Dette、Gauthier Dierickx、Tim Kutta

2023年1月12日

摘要

分析数据的协方差结构是统计学的一项基础任务。对于低维观测而言，这项任务很简单；但面对多元函数等更复杂的对象时，任务便颇具挑战性。在这种情况下，协方差可能复杂到仅保存一个非参数估计都不切实际，因此需要结构性假设来简化模型。时空数据中一种常见的假设是，协方差可分离为纯空间因子和纯时间因子。

本文针对相依函数型时间序列，提出一种新的可分离性检验。尽管大多数相关研究在由平方可积函数构成的希尔伯特空间中研究函数型数据，我们将观测建模为赋有上确界范数的连续函数空间中的对象。我们认为，这种在数学上颇具挑战的设定能增强用户对结果的可解释性，也更符合实际中的预处理方式。我们的检验统计量衡量估计的协方差核与可分离近似之间的最大偏差。临界值通过一种针对相关数据的非标准乘子自助法获得。我们证明了该方法在统计学上的有效性，并通过模拟研究和一个数据示例展示了其实用性。

关键词：相依乘子自助法；时空数据；可分离性；巴拿赫空间；函数型时间序列。

数学主题分类：62G10；
62R10。

1 引言

过去几十年来，高维时空数据分析已成为地统计学的基石。新技术使人们能够采集温度、磁场或污染物浓度等变量的高频、高分辨率测量数据（例如，Gromenko 等人（2012）；Aue 等人（2018）；King 等人（2018））。分析此类数据的一种方法是对其在空间和时间上进行平滑处理，从而得到时空过程的时间序列。这种重建和分析随机函数的方法遵循函数型数据分析（FDA）的范式，近年来在地统计学界日益受到关注（综述见 Martínez-Hernández 和 Genton（2020），以及 Mateu 和 Giraldo（2022）的专著）。

虽然分析这类过程有望带来深刻的科学见解，但它们的高度复杂性可能使统计方法面临极限。例如，主成分分析和克里金法等常用工具都依赖于

对过程协方差算子的逼近——该算子可能过于庞大，以至于无法存储或求逆来进行预测。关于这些计算挑战的简要概述，请参阅 Masak 等人（2020）的表1。为降低复杂性，许多研究会对协方差施加结构性假设，其中之一便是可分离性。

粗略地说，可分离性意味着时空过程的协方差可以分解为纯时间分量和纯空间分量。由此消除时空交互，可减少模型参数的数量，使协方差再次变得易于处理（参见 Genton（2007））。此外，可分离性使主成分具有乘积结构，从而便于构造估计量和推断方法（参见 Gromenko 等人（2012、2016））。为了严格定义可分离性，考虑随机过程 $\{X(s,t) \mid s \in K_1, t \in K_2\}$，其中一个自变量属于空间域 K_1，另一个属于时间域 K_2。那么，在适当条件下（例如参见 Janson 和 Kaijser（2015）），其协方差算子可以逐点定义为

$$C(s,t,s',t') := \mathbb{E}\left[(X(s,t) - \mathbb{E}X(s,t))(X(s',t') - \mathbb{E}X(s',t'))\right].$$

如果存在两个函数 C_1（空间）和 C_2（时间），使得 C，则称其为可分离的。

$$C(s,t,s',t') = C_1(s,s') \cdot C_2(t,t') \qquad \forall s, s' \in K_1,\ t, t' \in K_2.$$

如前所述，可分离性的乘积结构精简了模型参数，使模型在统计和计算上都更易于处理。

然而，对时空数据而言，可分离性并非毫无代价。事实上，错误地假设模型可分离，可能导致估计不一致和推断结果有偏。这一点至关重要，因为可分离性很少是不言自明的，许多作者都注意到了这一点（参见 Scaccia 和 Martin（2005）；Genton（2007）；Aston 等人（2017）等）。为解决这一问题，人们提出了用于检验可分离性的统计检验，例如针对有限维数据的 Matsuda 和 Yajima（2004）；Scaccia 和 Martin（2005）；Fuentes（2006）；Crujeiras 等人（2010）提出的方法。近年来，Aston 等人（2017）；Constantinou 等人（2017、2018）；Bagchi 和 Dette（2020）；Dette 等人（2022）设计了适用于函数型观测的非参数方法。尽管这些后续研究在推断策略上有所不同，但它们采用了相同的数学框架，即在平方可积函数空间中对观测进行建模。这种方法是函数型数据分析中的标准做法（参见 Horváth 和 Kokoszka（2012）；Hsing 和 Eubank（2015）的专著），也是将有限维方法推广到函数空间最直接的途径。然而，选择 L^2-空间通常更多是出于数学上的便利，而非实际应用的考虑。事实上，正如许多本科教材所指出的，L^2-距离难以从几何角度解释，其不直观的收敛概念也常令初学者困惑。就函数型数据分析而言，以 L^2-空间为基础构建理论，通常会忽略函数的结构特征，例如连续性。值得注意的是，函数型数据分析的大部分内容都以连续非参数曲线估计作为预处理。在这类情况下，坚持采用 L^2 框架可能会让用户觉得与直觉脱节，也会使结果难以作启发式解释。

继 Degras（2011）、Cao 等人（2012）、Degras（2017）以及 Dette 等人（2020）的近期研究之后，我们选择在我们看来更自然的连续函数空间上进行函数型数据分析，该空间配备了上确界范数。在这个更复杂的空间中，我们研究函数型时空过程，并提出一种检验协方差可分离性的新方法，该方法基于对最大

协方差算子与由可分过程的协方差算子构成的逼近之间偏差的估计。我们结合随机过程弱收敛的深厚理论（例如，参见 van der Vaart 和 Wellner（1996）或 Giné 和 Nickl（2016），以及众多其他文献），与这些可分性度量的新可微性结果，来研究相应估计量的渐近性质。为改善有限样本性质并避免估计复杂的干扰参数，我们提出一种适用于相关数据的乘子自助法，作为更实用的备选方案。特别是，与以往研究相比，我们的结果在较弱的依赖性和平稳性假设下仍然成立。

本文其余部分安排如下。第2节简要介绍连续函数空间上的随机变量以及可分离性的概念。随后，我们讨论相关文献中的三种方法，用可分离形式逼近协方差核 C （我们将在后文为这三种逼近方法分别建立理论）。第 3 节中，我们给出可分离性假设的检验统计量，并证明其弱收敛。为逼近极限分布，我们讨论一种针对相关数据的非标准乘子自助法。第 4 节通过模拟和一个数据实例，考察我们检验的有限样本性质。最后，所有证明和技术细节均置于附录中。

2 数学概念

本节在连续函数空间中阐述函数型数据分析的数学基础。第 2.1 节首先回顾随机连续函数，以及函数型期望和协方差核等基本概念。随后，我们定义可分离性这一模型假设，它是下文统计分析的重点。第 2.2 节中，我们讨论三种将核 A 逼近为可分离形式的方法，并证明每个逼近映射（$A \mapsto A^{\mathsf{x}}$）都有定义且可微（见定理 2.2）。

2.1 数学预备知识

Let

$$\mathcal{C}(K) := \big\{ f : K \to \mathbb{R} \,|\, f \text{ continuous} \big\},$$

表示定义在非空紧集 $K \subset \mathbb{R}^d$ 上的连续实值函数空间，该空间配备常用的上确界范数（简称“上确界范数”）

$$\|f\| := \sup_{t \in K} |f(t)| \tag{2.1}$$

是一个巴拿赫空间。若不会引起混淆，我们有时会用函数 $f \in \mathcal{C}(K)$ 的取值 $f(t)$ 来指代该函数。此外，我们还用记号 $\|\cdot\|$ 表示矩阵或向量的最大范数（即最大绝对值元素，对应于 (2.1) 中的有限集 K ）。

令 $(\Omega, \mathcal{A}, \mathbb{P})$ 表示一个完备概率空间。若映射 $X : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathcal{C}(K)$ 关于上确界范数是博雷尔可测的，则称其为随机 $\mathcal{C}(K)$ 值函数。我们指出， X （作为巴拿赫空间值函数）的博雷尔可测性，等价于对所有 $t \in K$，实值边缘函数 $X(t) : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}$ 的可测性。假设 X 的一阶绝对矩存在，即 $\mathbb{E}\|X\| < \infty$，则 X 的期望（在博赫纳意义下）是良定义的，其中 $\mathbb{E}X \in \mathcal{C}(K)$，并且

对任意 $t \in K$，该期望等于逐点期望 $\mathbb{E}[X(t)]$。若更强的矩条件 $\mathbb{E}\|X\|^2 < \infty$ 成立，我们可以定义乘积函数 $\{(X(s)-\mathbb{E}X(s))(X(t)-\mathbb{E}X(t))\}_{s,\ t\in K}$ 在空间 $\mathcal{C}(K^2) := \mathcal{C}(K \times K)$ 上的期望。我们称之为协方差核，并将其逐点定义为

$$C(s,t) := \mathbb{E}[X(s) - \mathbb{E}X(s)][X(t) - \mathbb{E}X(t)].$$

有关如何定义取值于巴拿赫空间的随机变量的矩，更多详情请参阅 Janson 和 Kaijser（2015）。与函数型数据研究中的惯例相同，对于连续核函数$A \in \mathcal{C}(K^2)$，我们有时会考虑相应的积分算子 $\{f(t)\}_{t\in K} \mapsto \{\int_K A(s,t)f(s)ds\}_{t\in K}$。为简化记号，我们通常将积分算子与其对应的核等同起来（注意二者之间是一一对应的），并定义表达式 $A[f] := \{\int_K A(s,t)f(s)ds\}_{t\in K}$。特别地，我们将与 C 相对应的积分算子称为协方差算子。最后，我们注意到，$\mathcal{C}(K)$可视为 $L^2(K)$的子空间；后者是配备典范半范数$\|f\|_{L^2} := \{\int_K f(s)^2 ds\}^{1/2}$的平方可积函数空间。回想一下，$L^2$-拓扑弱于一致拓扑，因为我们考虑的是具有紧支撑的函数。

接下来，我们考察可分离性这一数学性质。设 $K_1 \subset \mathbb{R}^d, K_2 \subset \mathbb{R}^q$ 是紧且非空的集合。若存在核 $A_i \in \mathcal{C}(K_i^2)$，其中 $i = 1,\ 2$，使得……，则称核 $A \in \mathcal{C}((K_1 \times K_2)^2)$ 是可分离的。

$$A(s,t,s',t') = A_1(s,s')\,A_2(t,t') \qquad \forall s,s' \in K_1,\ t,t' \in K_2.$$

之后，我们将研究随机函数 X 的可分离协方差核 C。为了评估可分离性，我们将使用估计量 C_N，并将其与一个可分离近似进行比较。构造此类可分离近似是下一节的主题。

2.2 可分离近似

本节讨论如何用两个函数 $B = A_1 \cdot A_2$ 的（可分离）乘积逼近核函数 $A \in \mathcal{C}((K_1 \times K_2)^2)$，其中 $A_i \in \mathcal{C}(K_i^2)$ 和 $i = 1,\ 2$。这种逼近是后续分析的关键，因为它通过拟合优度度量 $\|A - B\| = 0$ 趋于零来刻画可分离性。那么，我们应当如何构造可分离近似呢？最自然的方法或许是针对范数 $\|\cdot\|$，在所有可分离映射上进行优化，即求得

$$B^{\mathrm{opt}} \in \operatorname{argmin}\{\|A - B\| : B \,\text{separable}\}.$$

遗憾的是，这类逼近在计算上不可行。为说明这一点，我们考虑一个类似的问题：针对 $d \times d$ 矩阵，寻找关于最大范数（也记作 $\|\cdot\|$）的最优可分离近似。这个问题反映了对 A 离散化版本（即实际应用中保存在计算机上的版本）寻找最优可分离近似的过程。在下文的引理中，"$\otimes$" 表示矩阵的众所周知的克罗内克积。

注 2.1。设 $M \in \mathbb{R}^{d\times d}$ 是一个实值矩阵，满足 $d = pq$ 且 $p,\ q \in \mathbb{N}$。那么，寻找以下最小化问题的解

$$\text{“minimize } \|M - M_1 \otimes M_2\| \quad \text{with} \quad M_1 \in \mathbb{R}^{p\times p},\ \ M_2 \in \mathbb{R}^{q\times q}\text{”},$$

是 NP-完备的。

证明。 根据 Gillis 和 Shitov（2017），寻找关于最大范数的最优秩一近似是一个 NP-完备问题。根据 Genton（2007）第2节，寻找可分离逼近的问题等价于寻找秩一近似。

□

上述注 2.1 中的 NP-完全性 概念，指的是一类问题；根据目前的认识，这类问题无法高效求解（详情参见 Garey 和 Johnson（1979））。显然，如果无法为矩阵找到最大范数逼近，那么对于连续核也必然如此（因此，即使在合理大小的网格上，也无法轻易找到连续核的最优可分核逼近）。这一认识表明，应采用其他次优但在计算上可行的近似方法。接下来，我们将讨论三种此前已用于研究 L^2-函数可分离性的方法。

2.2.1 偏迹近似

设 $A \in \mathcal{C}((K_1 \times K_2)^2)$ 为一个核函数。于是，我们可以逐点定义边缘核 $A_1^{\mathrm{tr}} \in \mathcal{C}(K_1^2)$ 和$A_2^{\mathrm{tr}} \in \mathcal{C}(K_2^2)$，如下所示:

$$A_1^{\mathrm{tr}}(s,s') := \int_{K_2} A(s,w,s',w)dw \quad \text{and} \quad A_2^{\mathrm{tr}}(t,t') = \int_{K_1} A(u,t,u,t')du \tag{2.2}$$

并据此将 A 的可分离近似 A^{tr} 定义为:

$$A^{\mathrm{tr}}(s,t,s',t') := \frac{A_1^{\mathrm{tr}}(s,s')A_2^{\mathrm{tr}}(t,t')}{\int_{K_1}\int_{K_2} A(u,w,u,w)du\,dw}. \tag{2.3}$$

注意，如果式 (2.3) 中的分母非零，则 A^{tr} 有明确定义。对于协方差核（对称且半正定），这一条件自动满足，除非该核是平凡的。偏迹近似在量子物理中广为人知（参见 Bhatia（2003）及其中的参考文献），近期也已用于检验 L^2-空间中时空过程的可分离性（参见 Constantinou 等人（2017）、Aston 等人（2017））。Masak 等人（2020）近期的一项研究考察了偏迹的推广形式，并将其应用于“近似可分矩阵”的背景中。

2.2.2 偏乘积近似

偏积与偏迹的概念密切相关。偏乘积近似取决于用户指定的函数 $\psi \in \mathcal{C}(K_2^2)$，Bagchi 和 Dette（2020）讨论了几种典型的选取方式（其中一种是常数 1）。于是，对于 $A \in \mathcal{C}((K_1 \times K_2)^2)$，我们逐点定义边缘核

$$A_1^{\mathrm{pr}}(s,s') := \int_{K_2^2} A(s,w,s',w')\psi(w,w')dw\,dw' \quad \text{and} \quad A_2^{\mathrm{pr}}(t,t') := \int_{K_1^2} A(u,t,u',t')A_1^{\mathrm{pr}}(u,u')du\,du' \tag{2.4}$$

以及相应的可分离近似 A^{pr} 为

$$A^{\mathrm{pr}}(s,t,s',t') := \frac{A_1^{\mathrm{pr}}(s,s')A_2^{\mathrm{pr}}(t,t')}{\int_{K_1^2}\left(A_1^{\mathrm{pr}}(u,u')\right)^2 du\,du'}. \tag{2.5}$$

与偏迹类似，当分子非零时，该近似是良定义的，即，对于 r

$$A_1^{\mathrm{pr}} \neq 0. \tag{2.6}$$

如果对适当选择的 ψ， $A \neq 0$，则满足此条件。最近，偏积被用于 Dette 等人（2022）中，以量化 L^2 空间中的可分离性；更近期，Masak 等人（2020）利用偏积高效推导出关于 L^2 范数的最优可分近似。接下来将讨论后者。

2.2.3 SPCA 近似

SPCA（可分主成分分析）提供了我们希望研究的最后一种逼近方法。对于 L^2-范数，SPCA 近似是最优的；而对于上确界范数，它们并不具有这一特殊地位（见注 2.1）。SPCA 近似在有限维情形中已有数十年的研究历史（Van Loan 和 Pitsianis（1993）；Genton（2007）），最近 Dette 等人（2022）又将其推广到了无限维希尔伯特空间。“SPCA”这一名称由 Masak 等人（2020）提出，他们利用部分乘积提出了一种高效算法，用于计算这些逼近（见该文献中的第 2.2.2 节）。考虑核

$$\tilde{A}_1^{\mathrm{PCA}}(s,s',\bar{s},\bar{s}') := \int_{K_2^2} A(s,w,s',w')A(\bar{s},w',\bar{s}',w)dw\,dw'$$

and

$$\tilde{A}_2^{\mathrm{PCA}}(t,t',\bar{t},\bar{t}') := \int_{K_1^2} A(u,t,u',t')A(u',\bar{t},u,\bar{t}')du\,du',$$

这两个核均为连续、对称（相对于第一对和第二对分量而言）且正定的。根据 Mercer 定理（Steinwart 和 Christmann（2008）中的定理 4.49），我们可以将它们写成

$$\tilde{A}_1^{\mathrm{PCA}}(s,s',\bar{s},\bar{s}') = \sum_{i\geq 1}\lambda_i v_i(s,s')v_i(\bar{s},\bar{s}'), \qquad \tilde{A}_2^{\mathrm{PCA}}(t,t',\bar{t},\bar{t}') = \sum_{i\geq 1}\lambda_i u_i(t,t')u_i(\bar{t},\bar{t}').$$

其中，$\{\lambda_i, v_i\}_{i\in\mathbb{N}}$ 和 $\{\lambda_i, u_i\}_{i\in\mathbb{N}}$ 分别是相应积分算子的特征系统，并假设特征值按降序排列。不难证明，如果严格不等式

$$\lambda_1 > \lambda_2 \tag{2.7}$$

成立，则特征函数 v_1、u_1 是定义良好的（除符号外唯一）且连续的（这两个结论均见附录 A.3）。假设这一点成立，我们定义边缘函数

$$A_1^{\mathrm{PCA}}(s,s') = \sqrt{\lambda_1}v_1(s,s'), \qquad A_2^{\mathrm{PCA}}(t,t') = \sqrt{\lambda_1}u_1(t,t')$$

并由此得到 SPCA 近似

$$A^{\mathrm{PCA}}(s,t,s',t') = A_1^{\mathrm{PCA}}(s,s')A_2^{\mathrm{PCA}}(t,t'). \tag{2.8}$$

我们以关于近似映射 $A \mapsto A^{\mathsf{x}}$ 的一项一般性结果结束本节，该结果适用于每个 $\mathsf{x} \in \{\text{tr, pr, 主成分分析}\}$。我们证明这些映射均定义良好（即所得的核确实连续且正定）。此外，我们还证明近似映射是 Fréchet 可微的，这对于随后应用泛函 Delta 方法至关重要（例如，参见 van der Vaart 和 Wellner（1996）中的第 3.9 节）。

定理 2.2。设 $K_1 \subset \mathbb{R}^p$, $K_2 \subset \mathbb{R}^q$ 为紧致的非空集合，并设 $\tilde{A} \in \mathcal{C}((K_1 \times K_2)^2)$ 为一个˜协方差核，且满足 $A \neq 0$。此外，假设方程 (2.6) 和 (2.7) 均成立。那么映射

$$\begin{cases} \mathbf{F}_i^{\mathbf{x}} : \mathcal{C}((K_1 \times K_2)^2) \to \mathcal{C}(K_i^2) : A \mapsto A_i^{\mathbf{x}} \\ \mathbf{F}^{\mathbf{x}} : \mathcal{C}((K_1 \times K_2)^2) \to \mathcal{C}((K_1 \times K_2)^2) : A \mapsto A^{\mathbf{x}} \end{cases}$$

˜是针对 $i = 1,2$ 和 $\mathbf{x} \in \{\text{tr,pr,主成分分析}\}$ 在足够小的开邻域内定义良好的，且在 A 中弗雷歇可微。此外，$\mathbf{F}_i^{\mathbf{x}}[A]$ 和 $\mathbf{F}^{\mathbf{x}}[A]$ 仍然是协方差核。

该定理的证明见附录 A，其中我们还给出了导数的显式形式。在下一节中，我们将在时空数据统计推断的背景下运用这一结果。

3 连续协方差核的可分离性检验

本节将针对连续函数空间中的可分离协方差核假设，构建一种统计检验。首先，我们在第 3.1 节说明统计框架；其次，在第 3.2 节考察用于检验可分离性假设的检验统计量。引理 3.5 推出这些检验统计量（在原假设下）弱收敛至高斯过程的上确界，从而为可分离性检验提供理论工具。为近似检验统计量的渐近分位数，我们在第 3.3 节介绍一种适用于相关数据的乘子自助法。

3.1 记号和假设

设 $K_1 \subset \mathbb{R}^p$ 和 $K_2 \subset \mathbb{R}^q$ 为紧致的非空集合，且 $(X_n)_{n\in\mathbb{Z}}$ 是巴拿赫空间 $\mathcal{C}(K_1 \times K_2)$ 中的随机函数时间序列（定义见第2.1节）。下文假设 $(X_n)_{n\in\mathbb{Z}}$ 满足四阶平稳性，即对于任意指标 $(n_1, \ldots, n_4) \in \mathbb{Z}^4$ 和 $k \in \mathbb{Z}$，向量 $(X_{n_1}, \ldots, X_{n_4})$ 与 $(X_{n_1+k}, \ldots, X_{n_4+k})$ 具有相同的分布。特别地，假设 $\mathbb{E}\|X_1\|^2 < \infty$，则均值函数 $\mathbb{E}X_n(s,t)$ 和协方差

$$C(s,t,s',t') := \mathbb{E}\left[X_n(s,t)X_n(s',t')\right] - \mathbb{E}\left[X_n(s,t)\right]\mathbb{E}\left[X_n(s',t')\right]$$

不依赖于 n，并且分别在空间 $\mathcal{C}(K_1 \times K_2)$ 和 $\mathcal{C}((K_1 \times K_2)^2)$ 上定义良好。下文将构造一个用于检验协方差算子可分离这一假设的检验，即，

$$H_0 : C \text{ is separable} \qquad \text{vs.} \qquad H_1 : C \text{ is not separable,} \tag{3.1}$$

其中，可分离性在第2.1节中定义。为此，假设我们观测到一个由 N随机函数 $X_1, \ldots, X_N$（来自时间序列 $(X_n)_{n\in\mathbb{Z}}$）组成的样本。为估计 C，我们使用如下定义的（标准）经验协方差估计量：

$$\hat{C}_N(s,t,s',t') := \frac{1}{N}\sum_{n=1}^{N}\left(X_n(s,t) - \overline{X}_N(s,t)\right)\left(X_n(s',t') - \overline{X}_N(s',t')\right), \tag{3.2}$$

其中，$\overline{X}_N(s,t) := \frac{1}{N}\sum_{n=1}^{N} X_n(s,t)$、$(s,t) \in K_1 \times K_2$表示样本均值估计函数。在接下来的章节中，我们将通过将 C_N 与可分离

近似进行比较，为假设构造一个检验统计量。使用该统计量的动机在于，在关于依赖结构的适当假设下，根据（巴拿赫空间上的）大数定律，$\hat{}\, C_N$ 是 C 的相合估计量。为了量化依赖性，我们引入广为使用的 α-混合序列概念。

定义 3.1. 设$(X_n)_{n\in\mathbb{Z}}$ 是某个巴拿赫空间上的一列随机变量。对于指标集$\mathcal{I},\mathcal{J}\subset\mathbb{Z}$ 我们定义集合距离

$$\operatorname{dist}(\mathcal{I},\mathcal{J})=\min\{|i-j|:i\in\mathcal{I},j\in\mathcal{J}\}$$

并使用记号 $\mathcal{F}_{\mathcal{I}}:=\sigma(\{X_i:\ \ i\in\mathcal{I}\})$ 表示由随机变量族$\{X_i:\ \ i\in\mathcal{I}\}$生成的 σ–代数。于是， $r\in\mathbb{N}_0$ 第 r 个 α- 混合系数定义为

$$\alpha(r)=\sup\left\{|\mathbb{P}(A\cap B)-\mathbb{P}(A)\mathbb{P}(B)|\ :\ A\in\mathcal{F}_{\mathcal{I}},\ B\in\mathcal{F}_{\mathcal{J}},\ \operatorname{dist}(\mathcal{I},\mathcal{J})\geq r,\ \mathcal{I},\mathcal{J}\subset\mathbb{Z}\right\}.$$

若序列$(X_n)_{n\in\mathbb{Z}}$ 满足 $\alpha(r)\ \to 0$（如 $r\to\infty$），则称其为 α-混合序列。

We 现在可以陈述本节所建立的可分离性检验的理论假设 .

假设 3.2.

i) 序列$(X_n)_{n\in\mathbb{Z}}$ 由 $(K_1\times K_2)$ 中的中心化随机函数构成，且为四阶平稳。

ii) 存在一个非负随机变量 M、参数 $\beta\in(0,1]$ 和 $J\geq 0$，满足约束条件$J\beta>\lceil 2(\dim(K_1)+\dim(K_2))\rceil+1$，使得

$$\mathbb{E}(\|X_1\|^J M^J)<\infty,\quad \textit{and}\quad |X_n(s,t)-X_n(s',t')|\leq M\max\{\|s-s'\|^\beta,\|t-t'\|^\beta\},\tag{3.3}$$

对于所有 $(s,t),(s',t')\in K_1\times K_2$ 以及所有 $n\in\mathbb{Z}$，该条件几乎必然成立。

iii) 对于某个 $\gamma>\max\{J,8\}$，有 $\mathbb{E}\|X_1\|^\gamma<\infty$.

iv) 序列 $(X_n)_{n\in\mathbb{Z}}$ 按照定义 (3.1) 的意义是 α-混合的，其中 $\alpha(r)\ \leq\kappa/(1+r)^a$，其中$\kappa>0$ 是某个常数，且 $a>2\gamma/(\gamma-8)$。

我们简要讨论这些假设。

注 3.3。

i) 我们要求满足二阶平稳性，从而经验均值 $\overline{X}_N$ 和经验协方差算子 C_N 是相合估计量。更强的四阶平稳性假设保证 $\sqrt{N}(C_N-C)$ 的长期方差算子存在。检查我们的证明可以发现，这些假设还可进一步放宽为关于 X_n 的矩条件。不过，为了使表述简洁，我们不讨论这些在数学上较弱（但更难理解）的调整。我们的平稳性假设弱于相关文献中的假设（无论是在 L^2 空间还是 $\mathcal{C}$ 空间中）；相关文献考虑的要么是独立性，要么是严格平稳性（参见 Constantinou 等人（2017）；Aston 等人（2017）；Bagchi 和 Dette（2020）；Dette and Kokot（2022））。

ii) 为了构造可分离性渐近检验，我们在连续函数的巴拿赫空间上推导一个中心极限定理。此类结果需要验证紧性条件，而这些条件取决于底层空间的几何性质（由熵率体现）。对于独立同分布观测，Jain 和 Marcus（1975）给出了中心极限定理的充分条件；对于相关情形（在 α 混合和 ϕ 混合条件下），Dmitrovskii 等人（1984）给出了相应条件。具体而言，Jain 和 Marcus（1975）的定理 1 要求满足随机利普希茨条件，该条件与熵条件结合，可在 $\mathcal{C}(K)$ 上推出函数型中心极限定理。本文进一步放宽这一假设，只要求满足随机 β-Hölder 条件。该条件特别涵盖布朗运动的样本路径；对于 $\beta < 1/2$，这些路径几乎必然是 β-Hölder 连续的。不过，额外的光滑性有助于减少对数据施加的矩条件（见下一个假设）。

iii) – *iv*) 存在足够多的矩，是证明我们的中心极限定理（及其自助法变体）的关键。正如预期的那样，较强的矩条件可以与较弱的数据函数光滑性假设以及较温和的时间依赖条件相互权衡。我们通过强混合系数的衰减率来量化依赖性；衰减越慢，时间依赖性越强。与相关文献相比，我们的混合条件相当宽松，涵盖了大类相关时间序列。

3.2 在$\mathcal{C}((K_1 \times K_2)^2)$中的中心极限定理

我们首先通过证明标准化经验协方差算子 $\sqrt{N}(\hat{C}_N - C)$ 弱收敛到高斯过程 G，来展开理论推导。这一结果与相应的自助法收敛一同，对于在连续函数空间上对协方差进行统计研究具有独立意义（注意，定理 3.4 是定理 3.6 的一个特殊推论）。

定理 3.4. 假设假设 *3.2* 成立，则在空间 $\mathcal{C}((K_1 \times K_2)^2)$ 上存在一个中心高斯过程 G，使得

$$\sqrt{N}\big(\hat{C}_N - C\big) \xrightarrow{d} G.$$

现在，我们将经验协方差的弱收敛与可分近似映射的弗雷歇可微性（定理 2.2）结合起来。这使得对于第 A.2–A.3 节中讨论的任一种近似类型 $\mathbf{x} \in \{\text{tr, pr, 主成分分析}\}$，经验可分性度量 $\|\hat{C}_N - \hat{C}_N^{\mathbf{x}}\|$ 都弱收敛。

引理 3.5. 假设我们的假设 *3.2* 成立，$C \neq 0$ 是可分的，且满足(2.6)、(2.7) 。那么，在假设 H_0 （定义见(3.1)）下，有

$$\sqrt{N}\left(\hat{C}_N - \hat{C}_N^{\mathbf{x}}\right) \xrightarrow{d} (\mathrm{Id} - D_C\mathbf{F}^{\mathbf{x}})G \tag{3.4}$$

对所有$\mathbf{x} \in \{\text{tr}、\text{pr}、\text{主成分分析}\}$均成立。这里，$D_C\mathbf{F}^{\mathbf{x}}$ 是可分近似映射的导数（见第 *2* 节），Id 是恒等算子，而 G 是定理 *3.4* 中的高斯过程。特别地，在 H_0（可分性）条件下，有

$$\sqrt{N}\left\|\hat{C}_N - \hat{C}_N^{\mathbf{x}}\right\| \xrightarrow{d} \|(\mathrm{Id} - D_C\mathbf{F}^{\mathbf{x}})[G]\| \tag{3.5}$$

并且在 H_1 （非可分性）条件下

$$\sqrt{N}\left\|\hat{C}_N - \hat{C}_N^{\mathbf{x}}\right\| \xrightarrow{d} \infty.$$

式（3.4）中的弱收敛可通过将 Delta 方法的一个版本（例如参见 van der Vaart 和 Wellner（1996）中的第 3.9 节）应用于过程 $\sqrt{N}(\hat{}C_N - C)$ 得到。式（3.5）中的结论是这一结果的直接推论，使用了连续映射定理（参见 van der Vaart 和 Wellner ^ ^（1996）中的定理 1.3.6）。最后， H_1 下的发散性由 C_N 和 $C_N^{\mathbf{x}}$. 的一致性所导致

引理 3.5 给出了一个检验假设 H_0 的直接方法：如果已知高斯过程 G （或者等价地，其协方差），我们就可以轻松近似极限分布的上侧（$1-\alpha$）分位数 $q_{1-\alpha}$，并在以下情况下拒绝

$$\sqrt{N}\left\|\hat{C}_N - \hat{C}_N^{\mathbf{x}}\right\| > q_{1-\alpha}.$$

遗憾的是， G 的协方差 C_G 未知，而且极难近似。注意，它定义在乘积空间 $\mathcal{C}((K_1 \times K_2)^4)$ 上，因此几乎不可能对其进行计算或存储（请回想，我们的讨论源于 C 难以处理的问题，而 C_G 消耗的内存空间是其平方量级）。在有关 L^2 检验的文献中，Aston 等人（2017）；Constantinou 等（2018）曾尝试使用投影方法来规避这一问题。对于具有离散空间分量的函数型数据，这种方法是可行的；但在连续空间数据的情形下，效果似乎不那么理想（另见第 4.1 节中的模拟）。因此，作为一种更实用且完全函数型的备择方法，我们将在下一节讨论针对相关数据的乘子自助法。

3.3 乘子自助法

为了近似 $\sqrt{N}(\hat{}C_N - C)$ 的分布，我们针对相关数据提出一种乘子自助法，该方法的灵感来自 Bücher and Kojadinovic (2013) 的方法，且近期已由 Dette 等人 (2020) 推广至函数型数据。尽管函数型数据存在各种备择假设自助法（协方差方面的一个重要示例是 Paparoditis and Sapatinas (2016)），但其有效性通常是在 L^2 框架下证明的，因此不适用于我们的设定。

我们首先固定自助法重复样本的数量 $r \in \mathbb{N}$ 。对于 $1 \le k \le r$，我们考虑随机权重向量 $(w_1^{(k)}{}_{,N}, \ldots, w_N^{(k)}{}_{,N}) \in \mathbb{R}^N$，这些向量与数据 $X_1, \ldots, X_N$ 相互独立，且在 k 之间相互独立。我们假设每个权重向量都服从多元正态分布。更准确地说，变量 $w_i^{(k)}{}_{,N}$ 均假定为中心化且具有单位方差，并且为 $(l_N - 1)$-相关的

$$\mathbb{E}[w_{i,N}^{(k)} w_{j,N}^{(k)}] = (1 - |i-j|/l_N)$$

对任意 $|i-j| \le l_N$。这里 $l_N \in \mathbb{N}$ 是一个带宽参数，与分块自助法中的区块长度相当（例如，参见 Bücher and Kojadinovic (2013) 中的定理 2.1）。接下来，我们将自助过程 $B_N^{(k)}$ 逐点定义为

$$B_N^{(k)}(s,t,s',t') := \frac{1}{N}\sum_{n=1}^{N} w_{n,N}^{(k)} \left\{[X_n(s,t) - \overline{X}_N(s,t)][X_n(s',t') - \overline{X}_N(s',t')] - \hat{C}_N(s,t,s',t')\right\} \tag{3.6}$$

对 $1 \le k \le r$。

定理 3.6 (乘子自助法)。 假设假设 *3.2* 成立，且 l_N 满足限制条件 $l_N \to \infty$ 和 $l_N/\sqrt{N} \to 0$。那么，对任意 $r \in \mathbb{N}$ ，弱收敛

$$\sqrt{N}\left(\hat{C}_N - C, B_N^{(1)}, \ldots, B_N^{(r)}\right) \xrightarrow{d} \left(G, G^{(1)}, \ldots, G^{(r)}\right) \tag{3.7}$$

成立，其中， G 是定理 *3.4* 中的高斯过程， $G^{(1)}$、.....、 $G^{(r)}$ 是 G的独立同分布副本。

定理 3.6 的详细证明见附录，这里简要说明（3.7）成立的原因。首先注意，由于权重（数据的权重以及跨 k彼此独立）具有独立性，因此 $B_N^{(k)}$ 与ˆ $C_N - C$ 以及任何其他 $B^{(k')}{}_N$ 都不相关。因此，如果这些对象在极限下（联合）服从正态分布，那么它们也必然相互独立。由此留下两个问题：为什么 $B_N^{(k)}$ 服从正态分布，以及为什么它具有正确的协方差结构。从较高层面来看， $B_N^{(k)}$ 的渐近正态性显而易见，因为它是弱相关随机变量之和。仔细考察可知，（3.6）中的相关性由带宽 l_N决定；要使中心极限定理成立，该带宽不能增长过快（这将体现在假设 $l_N = o(\sqrt{N})$ 中）。另一方面， l_N 必须发散，以使 $N \to \infty$，从而 $B_N^{(k)}$ 具有与 G相同的渐近协方差。随着 l_N 增大， $B_N^{(k)}$ 中的相邻项具有（几乎）相同的ˆ权重。因此，它们的协方差与 C_N中各项的协方差（几乎）相同，只是没有乘数。

根据定理 3.6，自助法变量 $B_N^{(k)}$ 模拟了经验协方差 ˆ ˆ ˆC_N围绕 C的随机波动。因此，它可以帮助我们逼近可分离性度量 $\|C_N - C_N^{\mathbf{x}}\|$的分布；后者才是我们真正感兴趣的对象。为清楚起见，我们采用如下表示

$$\|\hat{C}_N - \hat{C}_N^{\mathbf{x}}\| = \|\{\hat{C}_N - C\} - ([C + \{\hat{C}_N - C\}]^{\mathbf{x}} - C)\|.$$

为了给出自助法近似，我们可以用 $B_N^{(k)}$ 替换 $\{\hat{C}_N - C\}$ 。此外，由于我们希望ˆ逼近原假设下的分布，因此必须在其他所有地方都用可分离估计量 $C_N^{\mathbf{x}}$替换 C 。由此得到自助法版本

$$B_N^{\infty,(k)} := \|B_N^{(k)} - ([\hat{C}_N^{\mathbf{x}} + B_N^{(k)}]^{\mathbf{x}} - \hat{C}_N^{\mathbf{x}})\|. \tag{3.8}$$

引理 3.7. 在定理 *3.6* 的假设下，有

$$\sqrt{N}(B_N^{\infty,(1)}, \cdots, B_N^{\infty,(r)}) \xrightarrow{d} \left(\|(\mathrm{Id} - D_C\mathbf{F}^{\mathbf{x}})[G_1]\|, \cdots, \|(\mathrm{Id} - D_C\mathbf{F}^{\mathbf{x}})[G_r]\| \right),$$

其中， G 是定理 *3.4* 中的高斯过程，而 $G^{(1)}, .., G^{(r)}$ 是 G的独立同分布副本。

引理 3.7 从理论上支撑了如下检验决策：令 ˆ$q_{1-\alpha}^{(r)}$ 为向量 $(B^{\infty,}{}_N{}^{(1)}, \cdots, B^{\infty,}{}_N{}^{(r)})$ 的经验 $(1-\alpha)$ 分位数。若满足以下条件，则拒绝假设 H_0（定义见 (3.1)）：

$$\|\hat{C}_N - \hat{C}_N^{\mathbf{x}}\| > \hat{q}_{1-\alpha}^{(r)}. \tag{3.9}$$

引理 3.7 表明，所得检验的渐近水平为 α，当 $r \to \infty$ 时，即

$$\lim_{r\to\infty} \lim_{N\to\infty} \mathbb{P}_{H_0} \left\{ \|\hat{C}_N - \hat{C}_N^{\mathbf{x}}\| > \hat{q}_{1-\alpha}^{(r)} \right\} = \alpha,$$

并且在备择假设 H_1 下，当 $N \to \infty$ 时是一致的，即

$$\lim_{N\to\infty} \mathbb{P}_{H_0} \left\{ \|\hat{C}_N - \hat{C}_N^{\mathbf{x}}\| > \hat{q}_{1-\alpha}^{(r)} \right\} = 1$$

对于任意 $r \in \mathbb{N}$。

注 3.8。

i) 数学上验证自助法程序有多种不同方式。一种方式是证明自助法测度（以数据为条件）收敛到正确的极限分布。这里，收敛是相对于概率测度上的某种距离来衡量的，例如 $\mathbb{R}^d$ 上概率测度的柯尔莫哥洛夫–斯米尔诺夫距离，或一般度量空间上的有界利普希茨距离（定义见 van der Vaart and Wellner（1996））。许多经典著作都采用了这种研究条件分布的方法，例如 Hall（1992）。另一种方法是推导无条件收敛结果（如定理 3.6 和引理 3.7），证明若干自助统计量收敛到同一极限分布的独立副本。这类结果反映了一个事实：对于大多数实际检验决策，都需要生成自助重复样本。与在测度空间上相对于一种难以处理的距离这一抽象收敛概念相比，这种方法的一个优点是更容易解释。然而，从理论角度来看，在许多情况下，这两种方法是等价的。特别是，引理 3.7 蕴含，对于条件自助法测度 $\mathbb{P}^{B_N^{(1)}|X_1,\ \ldots,\ X_N}$，依概率收敛

$$d\left(\mathbb{P}^{B_N^{(1)}|X_1,\cdots,X_N}, \mathbb{P}^{\|\hat{C}_N - \hat{C}_N^{\mathbf{x}}\|}\right) \xrightarrow{\mathbb{P}} 0,$$

其中，d是定义在 $\mathbb{R}$上的概率测度之间的 Kolmogorov–Smirnov 度量，$\mathbb{P}^{\|\hat{}C_N - \hat{}C_N^{\mathbf{x}}\|}$表示检验统计量 $\|C_N - C_N^{\mathbf{x}}\|$的概率测度。该结论的证明直接来自 Bücher and Kojadinovic（2019）中的引理 2.3；利用该文第 3 节中的技巧，也可以给出定理 3.6 的一个类似的条件版本。

ii) 定理 3.6 中的带宽参数 $l = l_N$ 与分块自助法中的块长度发挥着类似的作用。对于独立数据这一特殊情形，无需假设 $l_N \to \infty$，且选择 $l = 1$ 即可得到所需结果（关于独立情形下的乘子自助法，另见 van der Vaart and Wellner（1996）中的第 3.6 节）。我们还想强调，我们的假设$l_N/\sqrt{N} \to 0$ 比以往研究中的假设弱；以往研究通常假设 $l_N/\sqrt{N}$ 具有多项式衰减率（见 Dette and Kokot（2022））。

iii) 本文提出的自助法检验决策无需保存整个经验协方差算子或任何规模相当的其他对象即可实现。下面我们简要说明其原因。因此，我们重点讨论偏迹近似：要计算偏迹近似（$\hat{}C_N$）^{tr}，只需计算经验协方差的迹 Tr[$\hat{}C_N$]，以及偏迹（$\hat{}C_N$）$_1^{tr}$、（$\hat{}C_N$）$_2^{tr}$。所有这些对象都可以直接根据数据计算得到，例如，

$$\mathrm{Tr}[\hat{C}_N] = \frac{1}{N}\sum_{i=1}^{N}\int_{K_1}\int_{K_2} X_i^2(s,t)ds\,dt \tag{3.10}$$

or

$$(\hat{C}_N)_1^{tr}(s,s') = \frac{1}{N}\sum_{i=1}^{N}\int_{K_2} X_i(s,t)X_i(s',t)dt. \tag{3.11}$$

特别是，我们无需保存 $\hat{}\,C_N$ 即可计算 ($\hat{}C_N$)^{tr}。与 $\hat{}\,C_N$ 不同，三个对象 Tr[$\hat{}C_N$]、($\hat{}C_N$)$_1^{tr}$、($\hat{}\,C_N$)$_2^{tr}$ 都很小（离散化后所占内存与一个数据函数大致相同X_n），因此我们假设它们（进而包括 ($\hat{}\,C_N$)^{tr}）都易于处理。所以，从现在起我们

假设这些对象都已保存。现在，可以通过分块最大化距离 $|\hat{}C_N - \hat{}C_N^{tr}|$ 来计算检验统计量 $\|\hat{}C_N - \hat{}C_N^{tr}\|$（例如，只针对参数的某些子集（$s$、$t$、$s'$、$t'$）计算 $\hat{}C_N$，再对所有子集取最大值，很容易实现）。自助法统计量 $B^{\infty},_N{}^{(k)}$（定义见 (3.8)）的计算稍微复杂一些。这里，我们需要计算 $[\hat{}C_N^{tr} + B_N^{(k)}]^{tr}$（$B^{\infty},_N{}^{(k)}$ 中的其他对象仍然可以用直接的方法，针对索引子集计算）。计算 $[\hat{}C_N^{tr} + B_N^{(k)}]^{tr}$ 归结为分别计算 $\hat{}C_N^{tr}$ 和 $B_N^{(k)}$ 的（偏）迹（因为偏迹是线性的）。对于 $\hat{}C_N^{tr}$，它们等于 $(\hat{}C_N)_1^{tr}$、$(\hat{}C_N)_2^{tr}$，而我们此前已经计算并保存了这些量。对于 $B_N^{(k)}$，它们等于 $\frac{1}{N}\sum_{n=1}^{N} w_n^{(k)},_N \hat{}C_N$ 的偏迹（因此实质上就是 $\hat{}C_N$ 的偏迹）以及 $\frac{1}{N}\sum_{n=1}^{N} w_n^{(k)},_N [X_n - \overline{X}_N]\cdot[X_n - \overline{X}_N]$ 的偏迹。请注意，我们无法保存后一个对象（它与 C_N 大小相同）。因此，作为一种计算技巧，我们可以将其表示为复值数据 $\sqrt{\text{sign}(w_n^{(k)},_N)\cdot|w_n^{(k)},_N|}\cdot[X_n - \overline{X}_N]$ 的“协方差”。这些对象也足够小，可以保存；利用它们，我们可以直接计算 $\frac{1}{N}\sum_{n=1}^{N} \overline{w_n^{(k)},_N [X_n - \overline{X}_N]\cdot[X_n - \overline{X}_N]}$ 的偏迹（见 (3.10) 和 (3.11)）。请注意，实际计算中最后必须取实部，以消除由计算误差导致的复值余项。

4 有限样本性质

本节中，我们研究新方法在有限样本下的表现。首先，我们进行一项模拟研究，将该检验的表现与 Constantinou 等（2018）提出的基准程序进行比较。随后，我们将该检验应用于一组罗曼语录音数据集，该数据集由 Aston 等人（2017）描述并使用。

4.1 模拟

遵循 Constantinou 等（2018）的方法，我们通过函数型 MA(1) 过程生成时空数据。具体而言，我们生成定义在单位正方形 $[0,1]^2$, 上的高斯过程 $e_0, \ldots, e_N$，其协方差核为 C，逐点定义如下：

$$C(s,t,s',t') := \frac{1}{(a|t-t'|+1)^{1/2}} \exp\left(-\frac{b^2|s-s'|^2}{(a|t-t'|+1)^c}\right), \qquad s,s',t,t' \in [0,1]\,. \tag{4.1}$$

与 Constantinou 等（2018）一样，我们设定 $a = 3$ 和 $b = 2$，并考虑以下观测值

$$X_n(s,t) := \sum_{s'=1}^{S} \exp\left\{-b^2(s-s')^2\right\}\left[e_n(t,s') + e_{n-1}(t,s')\right] \quad n = 1, \cdots, N.$$

注意，对于 $c = 0$，式（4.1）中的核 C 可分解为空间分量和时间分量。因此，X_n 的协方差对于 $c = 0$（该假设）是可分离的。另一方面，如果 $c > 0$，核 C 不可分离，这种不可分离性也会体现在 X_n的协方差中。因此，我们通过设置 $c = 0$，在 H_0 下生成数据；在备择假设下，则设置 $c = 1$。与 Constantinou 等（2018）一样，我们将时间分量 t离散化，即将单位区间划分为 $T = 50$ 个等距点（$1/T$、$2/T$、.....、$50/T$）。对于空间分量，Constantinou 等（2018）采用类似的离散化方式（$1/S$、$2/S$、.....、$(S-1)/S$），但网格点数量较少，$S = 4$、6、8、10、12、14。考虑到

更大的 S 会给 Constantinou 等人（2018）的方法带来问题，因为该方法依赖于估计 $\sqrt{N}$（$\hat{C}_N - C$）的渐近协方差算子——在高维情况下，这是一项代价高昂且困难的工作（我们在第 3.2 节末尾的讨论中提到了这一点）。为实施他们的方法，Constantinou 等人（2018）采用降维：对时间上的 $S \leq 8$ 进行降维，对空间和时间上的 $S > 8$ 进行降维。正如预期的那样，随着所解释的方差比例降低，这些投影会导致检验功效随 S 增大而下降。在本研究中，我们使用相同数量的网格点（以便进行有意义的比较），同时也考察更大的 $S = 20$ 和 $S = 30$。我们考虑的样本量为 $N = 50$、100、150、200，相应的区块长度为 $l_N = 2$、2、3、4。生成经验分位数的自助法重复次数固定为 400，名义水平为 $\alpha = 5\%$。所有报告的结果均基于 1000 次模拟运行。

表1报告了可分离性假设（$c = 0$）及备择假设（$c = 1$）下的经验拒绝概率。括号内为 Constantinou 等（2018）报告的拒绝概率（如有），其中我们每次均选择了最多的投影参数个数（对 $S \leq 8$，在时间维度上；对 $S > 8$，在空间和时间维度上），以获得检验功效最高的结果（有关投影参数变化下性能的完整情况，请参见他们的表 I–IV）。

	在 H_0下的拒绝概率				在 H_1下的拒绝概率			
S	$N = 50$	$N = 100$	$N = 150$	$N = 200$	$N = 50$	$N = 100$	$N = 150$	$N = 200$
4	3.3	5.8 (5.0)	5.2 (5.3)	5.4 (5.5)	36.0	91.0 (95.1)	99.2 (99.8)	100.0 (100.0)
6	2.5	6.0 (5.3)	5.2 (6.6)	4.5 (5.1)	40.5	92.8 (89.0)	99.1 (99.3)	100.0 (100.0)
8	3.75	6.5 (7.5)	4.2 (4.7)	5.1 (5.7)	34.8	92.3 (85.2)	98.0 (98.7)	99.9 (100.0)
10	2.75	7.6 (4.7)	3.4 (5.3)	3.7 (5.8)	41.8	90.6 (84.7)	98.6 (98.5)	99.9 (100.0)
12	1.5	7.2 (5.0)	3.1 (6.0)	5.2 (5.7)	37.8	90.4(82.9)	99.0 (97.8)	99.9 (100.0)
14	2.0	7.0 (4.6)	5.3 (5.7)	3.7 (5.6)	33.5	92.5 (78.9)	98.6 (93.7)	100.0 (95.7)
20	3.4	5.2	5.7	3.80	38.0	90.3	98.3	99.9
30	2.7	7.0	5.3	3.61	36.1	87.4	98.2	99.9

表1： 假设 *($c = 0$)* 和备择假设 *($c = 1$)* 下自助检验 (3.9) 的经验拒绝概率。若有可用数据，括号内给出了 *Constantinou* 等（*2018*）检验的基准值。

表1中的结果表明，自助检验表现良好。对于 $N \geq 100$，名义水平得到了较为合理的近似；而对于 $N = 50$，检验则略显保守。在大多数情形下，自助检验的检验功效都很高。虽然当 S 较小时，基准检验的表现略胜一筹，但自助检验的表现不会随着 S增大而变差，在这种情况下，它显然优于基准检验。即使将 S 提高到 20 或 30，检验表现也不会受到任何系统性影响（这正是连续过程理论所预示的）。在计算方面，自助检验尤其便于使用：生成自助样本时可以轻松并行化，因此易于实现。我们在一台标准台式计算机（3.2 GHz Apple

M1 Pro、八核、16 GB 内存）上运行了模拟，每次检验评估所需时间都不到一分钟（对于 $S \leq 10$，不到 2 秒），这证明了该方法的实用性。随后给出的数据示例是在同一台计算机上运行的，所用 T 和 S的值更大。

4.2 数据示例

我们将检验方法应用于 Aston 等人（2017）的声学语音 数据集，该数据集包含声学（对数）谱图。其论文第 4.2 节提供了简要讨论及更多参考文献。原始数据和预处理数据，以及详细说明，均包含在文件 Acoustic_Data_And_Code.zip 中；该文件可从第 67 卷的 C 系列数据集（ https://rss.onlinelibrary.wiley.com/hub/journal/ 14679876/series-c-datasets/67_5）获取。数据由五种不同罗曼语族语言的口语词语录音组成（本例中为数字一到十）。为便于统计使用，这些数据已转换为声学对数谱图的形式。我们仅使用预处理数据（即文件 WarpedPSD.RData 和 SVRF_WarpedPSD_SuppMat.RData）。总共有 219 个频率—时间域中的数据“函数”，具体描述见 Pigoli 等人（2018）第 2—4 节。

正如 Pigoli 等人（2018）的分析已经指出的，在这一实例中，可分离性的原假设（3.1）似乎不成立。在应用我们的检验统计量之前，我们先考察可分离性的相对度量。更具体地说，对于可分离迹近似，其相对度量为

$$\frac{\|\hat{C}_N - \mathbf{F}^{\mathbf{tr}}(\hat{C}_N)\|}{\|\hat{C}_N\|}.$$

事实上，作为初步分析，我们发现语言协方差算子的可分离性相对度量相当高，见表 2。

	法语	**意大利语**	**葡萄牙语**	**美洲西班牙语**	**伊比利亚西班牙语**
相对度量	0.631	0.895	0.868	0.947	0.882

表 2：该　　五种罗曼语的迹近似下的可分离性相对度量。

对于五种语言中的每一种，我们分别对各语言的残差（即中心化）曲面数据运行了 1,000 次重复的自助法统计量。因此，这些语言的可分离性假设均被拒绝，p–值高度显著，低于 0.1%。检验的 p值见表3第一行。为完整起见，第二行和第三行分别列出了 Aston 等人（2017）在频率维度和时间维度分别为 2 和 3、以及 8 和 10 时得到的 p值。我们的方法有多项优势：无需任何调节参数；相比之下，Aston 等人（2017）的经验自助法学生化版本需要额外选择特征方向的维数参数（请注意，他们的检验只能检测沿这些特征方向偏离可分离性的情况）。此外，当选取的特征方向过少时（见 Aston 等人（2017）第 4.2 节），可分离性假设无法在 5% 水平上被拒绝。其次，计算自助法时，我们只需保存数据，而无需保存整个协方差算子，因此可以避免存储问题。

	法语	意大利语	葡萄牙语	美洲西班牙语	伊比利亚西班牙语
测试（3.9）	<0.001	<0.001	0.001	<0.001	<0.001
经验学生t检验 (2, 3)	0.078	0.197	0.022	0.360	0.013
经验学生t检验 (8, 10)	0.001	0.002	0.001	0.001	<0.001

表3： p–值：五种罗曼语的三种不同自助法检验。第一行：本文提出的检验(3.9)。第二行和第三行：*Aston* 等人（*2017*）提出的经验自助法检验的学生化版本，频率维度和时间维度分别为$(2,\ 3)$和$(8,\ 10)$。

致谢。 本工作部分得到了德国研究基金会（DFG）研究小组5381——信息时代的数学统计学的支持。

参考文献

Aston, J. A. D., D. Pigoli, 和 S. Tavakoli（2017）。随机曲面的非参数协方差算子可分离性检验。《统计年刊》*45*(4), 1431–1461。Aue, A., G. Rice, 和 O. Sönmez（2018）。无需降维检测并确定函数型数据中结构突变的时间。《英国皇家统计学会杂志*B*辑：统计方法学》*80*(3), 509–529。Bagchi, P. 和 H. Dette（2020）。随机曲面协方差算子的可分离性检验。《统计年刊》*48*(4), 2303–2322。Bhatia, R.（2003）。偏迹与熵不等式。《线性代数及其应用》*370*(1), 125–132。Billingsley, P.（2012）。《概率与测度，周年纪念版》，第238卷，收录于*Wiley*概率与统计系列。Wiley & Sons。Bücher, A. 和 I. Kojadinovic（2013）。强混合条件下序贯经验Copula过程的相依乘子自助法。《伯努利》*22*(2), 927–968。Bücher, A. 和 I. Kojadinovic（2019）。关于自助法一致性结果中条件弱收敛与联合无条件弱收敛的简注。《理论概率杂志》*32*, 1145–1165。Cao, G., L. Yang, 和 D. Todem（2012）。基于密集函数型数据对均值函数进行同时推断。《非参数统计杂志》*24*(2), 359–377。Constantinou, P., P. Kokoszka, 和 M. Reimherr（2017）。时空函数型过程的可分离性检验。《生物统计学》*104*(2), 425–437。

Constantinou, P., P. Kokoszka, 和 M. Reimherr（2018）。时空函数型过程的可分离性检验。《时间序列分析杂志》*39*(5), 731–747。Crujeiras, R. M., Fernández-Casal, 和 W. González-Manteiga（2010）。时空函数型过程的可分离性检验。《环境计量学》*21*, 382–399。Degras, D.（2011）。函数型数据非参数回归的同时置信带。《统计学报》*21*, 1735–1765。Degras, D.（2017）。函数型数据均值的同时置信带。《*WIREs* 计算统计学》*9*(3), e1397。Dehling, H., T. Mikosch, 和 M. Sørensen（2002）。相关数据的经验过程技术。Birkhäuser。Dette, H., G. Dierickx, 和 T. Kutta（2022）。量化时空函数型过程中偏离可分离性的程度。《伯努利》*28*(4), 2909–2940。Dette, H. 和 K. Kokot（2022）。“检测函数型时间序列协方差算子中的相关差异——一种上确界范数方法”。《统计数学研究所年刊》*74*(2), 195–231。Dette, H., K. Kokot, 和 A. Aue（2020）。连续函数巴拿赫空间中的函数型数据分析。《统计年刊》*48*(2), 1168–1192。Dmitrovskii, V. A., S. V. Ermakov, 和 E. I. Ostrovskii（1984）。弱相依巴拿赫值变量的中心极限定理。《理论概率及其应用》*28*, 89–104。Filipiak, K., D. Klein, 和 E. Vojtkova（2018）。分块矩阵偏迹算子与分块迹算子的性质。《电子线性代数杂志》*33*, 3–15。Fuentes, M.（2006）。时空协方差函数的可分离性检验。《统计规划与推断杂志》*136*(2), 447–466。Garey, M. R. 和 D. S. Johnson（1979）。《计算机与难解性：*NP*完全性理论指南》（第一版）。数学科学丛书。W. H. Freeman。Genton, M. G.（2007, 11）。时空协方差矩阵的可分离近似。《环境计量学》*18*, 681–695。Gillis, N. 和 Y. Shitov（2017）。无穷范数下的低秩矩阵近似。《线性代数及其应用》*581*, 367–382。Giné, E. 和 R. Nickl（2016）。无限维统计模型的数学基础。剑桥统计与概率数学丛书。纽约：剑桥大学出版社。Gromenko, O., P. Kokoszka, 和 M. Reimherr（2016）。时空均值函数的变化检测。《英国皇家统计学会杂志*B*辑：统计方法学》*79*(1), 29–50。

Gromenko, O., P. Kokoszka, L. Zhu, 和 J. Sojka (2012). 空间索引曲线的估计与检验及其在电离层和磁场趋势中的应用。《应用统计年鉴》*6*(2), 669 – 696. Hall, P. (1992). 《自助法与 *Edgeworth* 展开》. 纽约：施普林格。Horváth, L. 和 P. Kokoszka (2012). 《函数型数据推断及其应用》. 纽约：施普林格统计学丛书。Hsing, T. 和 R. Eubank (2015). 《函数型数据分析的理论基础：兼论线性算子导论》. 纽约：威利。Jain, N. C. 和 M. B. Marcus (1975). 取值于C(S)的随机变量的中心极限定理。《泛函分析杂志》*19*, 216–231. Janson, S. 和 S. Kaijser (2015). 取值于巴拿赫空间的随机变量的高阶矩，载于第238卷美国数学会回忆录。美国数学会。King, M. C., A.-M. Staicu, J. M. Davis, B. J. Reich, 和 B. Eder (2018). 细颗粒物时空趋势与变化的函数型数据分析。《大气环境》*184*, 233–243. Kokoszka, P. 和 M. Reimherr (2013). 函数型时间序列主成分的渐近正态性。《随机过程及其应用》*123*(5), 1546–1562. König, H. (1986). 《紧算子的特征值分布》，载于第16卷算子理论：进展与应用。Birkhäuser。Martínez-Hernández, I. 和 M. Genton (2020). 复值函数型数据与空间相关函数型数据的最新进展。《巴西概率与统计杂志》*34*, 204–229. Masak, T., S. Sarkar, 和 V. M. Panaretos (2020). 通过部分内积进行主可分成分分析。https://arxiv.org/abs/2007.12175. Mateu, J. 和 R. Giraldo (2022). 《地统计函数型数据分析》。威利。Matsuda, Y. 和 Y. Yajima (2004). 关于多元时间序列可分离相关性的检验。《时间序列分析杂志》*24*(4), 501–528. Paparoditis, E. 和 T. Sapatinas (2016). 基于自助法的函数型数据均值函数相等或协方差算子相等检验。《生物统计学》*103*, 727–733. Pigoli, D., P. Z. Hadjipantelis, J. S. Coleman, 和 J. A. D. Aston (2018). 声学语音数据的统计分析：探究口语罗曼语族语言之间的差异。《英国皇家统计学会杂志：*C*辑（应用统计）》*67*, 1103–1145. Scaccia, L. 和 R. J. Martin (2005). 格点过程轴对称性与可分离性的检验。《统计规划与推断杂志》*131*(1), 19–39. Steinwart, I. 和 A. Christmann (2008). 《支持向量机。 纽约：施普林格科学与商业媒体。

van der Vaart, A. W. 和 J. A. Wellner (1996). 《弱收敛与经验过程：统计学中的应用》纽约：施普林格统计学丛书。Van Loan, C. F. 和 N. Pitsianis (1993). 利用克罗内克积进行逼近。《大规模与实时应用线性代数》*232*, 293–314. White, H. (2001). 《计量经济学家的渐近理论》。纽约：学术出版社。
Yoshihara, K.-i. (1978). 混合序列的矩不等式。《古代数学杂志》*1*, 316–328.

定理 2.2 的证明

本节中，我们研究近似映射 $(\cdot)^{\mathrm{tr}}$、$(\cdot)^{\mathrm{pr}}$、$(\cdot)^{\mathrm{PCA}}$ 的可微性。我们采用 Dette 等人（2022）（定理 3.4）的结果，其中证明了这些近似在 L^2 函数空间上的可微性。由于连续函数构成 L^2 的一个子空间，且赋予了更强的范数，因此（对于每个映射）弗雷歇微分在两个空间上必须一致（若其存在）。因此，我们只需证明“差商”在连续函数空间中收敛。请注意，我们并未证明任何近似具有半正定性；这在文献中已有充分讨论（关于偏迹，参见 Filipiak 等人（2018）的引理 2.4；关于其他近似类型，参见 Masak 等人（2020）中的讨论）。

A.1 偏迹的可微性

回顾式 (2.2) 中偏迹核的定义，以及式 (2.3) 中的偏乘积近似。

引理 A.1. 设 $K_1 \subset \mathbb{R}^p$, $K_2 \subset \mathbb{R}^q$ 为紧集。那么（偏）迹算子

$$\mathrm{tr} : \mathcal{C}((K_1 \times K_2)^2) \to \mathbb{R}$$
$$\mathrm{tr_i} : \mathcal{C}((K_1 \times K_2)^2) \to \mathcal{C}(K_i^2)$$

在第 *2.2.1* 节中定义的算子是连续的。

证明。 证明可由初等计算得出。对于 $A \in \mathcal{C}((K_1 \times K_2)^2)$，我们可以对积分给出上界

$$|\mathrm{tr}[A]| = \left| \int_{K_1 \times K_2} A(u, w, u, w) du\, dw \right| \leq |K_1| \cdot |K_2| \cdot \|A\|.$$

这里，$|K_i|$ 表示适当维数下 K_i 的勒贝格测度。利用第 2.2.1 节中的定义，可用类似论证证明 tr_1,tr_2 的连续性。

□

定理 A.2. 设 $K_1 \subset \mathbb{R}^p$, $K_2 \subset \mathbb{R}^q$ 为紧集。那么，在任意非零协方差核 C处，式(2.3) 中定义的映射$\mathbf{F}^{\mathrm{tr}}$均为弗雷歇可微。

证明。 由链式法则可知，$\mathrm{F}_i^{\mathrm{tr}} = \mathrm{H}^{\mathrm{tr}} \circ \mathrm{G}^{\mathrm{tr}}$的弗雷歇可微性来自这两个映射的可微性

$$\mathbf{G}^{\mathrm{tr}} : \mathcal{C}((K_1 \times K_2)^2) \to \mathcal{C}(K_1^2) \times \mathcal{C}(K_2^2) \quad \text{and} \quad \mathbf{H}^{\mathrm{tr}} : \mathcal{C}(K_1^2) \times \mathcal{C}(K_2^2) \to \mathcal{C}((K_1 \times K_2)^2)$$

逐点定义为

$$\begin{pmatrix} \mathbf{G}_1^{\mathrm{tr}}[T](s, s') \\ \mathbf{G}_2^{\mathrm{tr}}[T](t, t') \end{pmatrix} = \begin{pmatrix} \dfrac{\int_{K_2} T(s, w, s', w) dw}{\int_{K_1 \times K_2} T(u, w, u, w) du\, dw} \\ \int_{K_1} T(u, t, u, t') du \end{pmatrix}$$

$$\mathbf{H}^{\mathrm{tr}} \left[\begin{pmatrix} F \\ G \end{pmatrix} \right] (s, t, s', t') = F(s, s') G(t, t').$$

根据 Dette 等人（2022）中定理 3.4 的证明，我们声称 G^{tr} in$C$ 的弗雷歇微分，即 $D_C\mathrm{G}^{\mathrm{tr}}[T](s,t,s',t')$，可以逐点表示为

$$\begin{pmatrix} \dfrac{\int_{K_2} T(s,w,s',w)dw}{\int_{K_1\times K_2} C(u,w,u,w)dudw} - \dfrac{\int_{K_1\times K_2} T(u,w,u,w)dudw \int_{K_2} C(s,w,s',w)dw}{\left(\int_{K_1\times K_2} C(u,w,u,w)dudw\right)^2} \\ \int_{K_1} T(u,t,u,t')du \end{pmatrix}$$

映射 $D_C\mathrm{G}^{\mathrm{tr}}$ 的有界性直接由 tr、tr_1,tr_2 的连续性得出（见引理 A.1）。现在，我们验证 $D_C\mathrm{G}^{\mathrm{tr}}$ 确实是微分。由于第二个分量关于上确界范数是线性且连续的，因此它必然是自身的微分。因此，只需考虑第一个分量。为此，令 $H \in \mathcal{C}((K_1\times K_2)^2)$，使得 $\|H\| \to 0$。注意，当 H 足够小时，后续各对象中的所有（部分）迹都有定义。简单计算可得如下表示

$$\begin{aligned} &\mathbf{G}_1^{\mathrm{tr}}[C+H](s,t,s',t') - \mathbf{G}_1^{\mathrm{tr}}[C](s,t,s',t') - D_C\mathbf{G}_1^{\mathrm{tr}}[H](s,t,s',t') \\ =&\frac{-\left(\int_{K_2} H(s,w,s',w)dw\right)\left(\int_{K_1\times K_2} H(u,w,u,w)du\,dw\right)}{\left(\int_{K_1\times K_2}(C+H)(u,w,u,w)du\,dw\right)\left(\int_{K_1\times K_2} C(u,w,u,w)du\,dw\right)} \\ +&\frac{\left(\int_{K_1\times K_2} H(u,w,u,w)du\,dw\right)^2 \int_{K_2} C(s,w,s',w)dw}{\left(\int_{K_1\times K_2}(C+H)(u,w,u,w)du\,dw\right)\left(\int_{K_1\times K_2} C(u,w,u,w)du\,dw\right)^2}. \end{aligned}$$

仿照上文引理 A.1 的证明可以轻易看出，右侧两项均一致地为 $\mathcal{O}\left(\|H\|^2\right)$ 阶，这表明弗雷歇可微性所需的性质成立。接下来，我们讨论映射 H^{tr} 的可微性，而 Dette 等人（2022）中的定理 3.4（及其证明）再次给出了如下微分候选式

$$D_{(L_1,L_2)}\mathbf{H}^{\mathrm{tr}}\left[\begin{pmatrix} F \\ G \end{pmatrix}\right](s,t,s',t') = L_1(s,s')G(t,t') + F(s,s')L_2(t,t')$$

其中 $(L_1,L_2)^t := \mathrm{G}^{\mathrm{tr}}[C]$。证明 $D_{(L_1,L_2)}\mathrm{H}^{\mathrm{tr}}$ 确实是弗雷歇微分的过程，与定理 A.4（下文）中证明 H^{pr} 可微性时的计算类似，但更简单，因此略去。

□

A.2 偏乘积的可微性

回顾式 (2.4) 中偏乘积核的定义，以及式 (2.5) 中偏乘积近似的定义。

引理 A.3. 对于任意 $\psi \in \mathcal{C}(K_2^2)$，第 *2.2.2* 节中定义的偏乘积算子是连续的。

证明。 对于线性算子 $\mathrm{F}_1^{\mathrm{pr}}$ ，证明显然成立，因为它显然有界。为了证明

$\mathbf{F}_2^{\mathrm{pr}}$ 的连续性，设 $A, H \in \mathcal{C}(K_2^2)$ 满足 $\|H\| \to 0$。那么，根据其核的定义，我们有

$$\int_{K_1^2} H(u,t,u',t')(A+H)_1^{\mathrm{pr}}(u,u')du\,du' - \int_{K_1^2} A(u,t,u',t')\left((A+H)_1^{\mathrm{pr}}(u,u') - A_1^{\mathrm{pr}}(u,u')\right)du\,du' \tag{A.1}$$

$$= \int_{K_1^2} \left(H(u,t,u',t')H_1^{\mathrm{pr}}(s,s')(u,u') + H(u,t,u',t')A_1^{\mathrm{pr}}(u,u') + A(u,t,u',t')H_1^{\mathrm{pr}}(u,u')\right)du\,du'$$

其中，第二个等式用到了 $\mathrm{F}_1^{\mathrm{pr}}$ 的线性。第一个积分是 $\mathcal{O}(\|H\|^2)$ 阶的，这里用到了线性映射 $\mathrm{F}_1^{\mathrm{pr}}$ 的（利普希茨）连续性。显然，第二个和第三个积分是关于 H 的（有界）线性映射，对应于导数。

□

定理 A.4. 设 $K_1 \subset \mathbb{R}^p$、$K_2 \subset \mathbb{R}^q$ 为紧集，且选取 $\psi \in \mathcal{C}(K_2^2)$ 使得 C_1^{pr}（定义见 (2.4)）不恒为 0。那么，(2.5) 中定义的映射 $\mathbf{F}^{\mathrm{pr}}$ 在 C 上弗雷歇可微。

证明。与上文定理 A.2 的证明类似，我们也可以将 F^{pr} 分解为两个较简单的映射

$$\mathbf{G}^{\mathrm{pr}} : \mathcal{C}((K_1\times K_2)^2) \to \mathcal{C}(K_1^2)\times\mathcal{C}(K_2^2) \quad \text{and} \quad \mathbf{H}^{\mathrm{pr}} : \mathcal{C}(K_1^2)\times\mathcal{C}(K_2^2) \to \mathcal{C}((K_1\times K_2)^2)$$

逐点定义为

$$\mathbf{G}^{\mathrm{pr}}[T](s,t,s',t') = \begin{pmatrix} \int_{K_2^2} T(s,w,s',w')\psi(w,w')dw\,dw' \\ \int_{K_1^2} T(u,t,u',t')\left(\int_{K_2^2} T(u,w,u',w')\psi(w,w')dw\,dw'\right)du\,du' \end{pmatrix}$$

$$\mathbf{H}^{\mathrm{pr}}\left[\begin{pmatrix} F \\ G \end{pmatrix}\right](s,t,s',t') = \frac{F(s,s')G(t,t')}{\int_{K_1^2}(F(u,u'))^2 du\,du'}.$$

由线性和有界性（相对于上确界范数），$\mathbf{G}^{\mathrm{pr}}$ 的第一个分量关于 C 的导数就是该映射本身。第二个分量可微，这可由分解式 (A.1) 得出，其中最后两个积分是关于 H 的线性映射（即导数），第三个积分则是 $\mathcal{O}(\|H\|^2)$ 阶的。事实上，定义映射（它实际上仅依赖于变量 (t,t')）

$$D_C\mathbf{G}_2^{\mathrm{pr}}[H](s,t,s',t') = \int_{K_1^2} C(u,t,u',t')\left(\int_{K_2^2} H(u,w,u',w')\Delta(w,w')dw\,dw'\right)du\,du' + \int_{K_1^2} H(u,t,u',t')\left(\int_{K_2^2} C(u,w,u',w')\Delta(w,w')dw\,dw'\right)du\,du'.$$

并将其从 $(\mathbf{G}_2^{\mathrm{pr}}[C+H]-\mathbf{G}_2^{\mathrm{pr}}[C])(s,t,s',t')$ 中减去，我们根据 (A.1) 推断，该差值的阶为 $\mathcal{O}(\|H\|^2)$。接下来，我们计算 $\mathbf{H}^{\mathrm{pr}}$ 对一般数对 $(L_1,L_2)^t$, $L_i \in \mathcal{C}(K_i^2)$, $i=1,2$ 的导数

其中 $L_1 \neq 0$。其逐点形式为

$$D_{(L_1,L_2)^t}\mathbf{H}^{\mathrm{pr}}\left[\begin{pmatrix} F \\ G \end{pmatrix}\right](s,t,s',t') = \frac{F(s,s')L_2(t,t') + L_1(s,s')G(t,t')}{\int_{K_1^2}(L_1(u,u'))^2 du\,du'} - \frac{2\left(\int_{K_1^2} F(u,u')L_1(u,u')du\,du'\right)L_1(s,s')L_2(t,t')}{\left(\int_{K_1^2}(L_1(u,u'))^2 du\,du'\right)^2}$$

导数的有界性直接源于核 F, G 的有界性（以及分母远离 0 的有界性）。现在考虑 $H_i \in \mathcal{C}(K_i^2)$，其中 $\max(\|H_1\|, \|H_2\|) \to 0$，通过分解差值可得映射 $\mathbf{H}^{\mathrm{pr}}$ 的可微性

$$\begin{aligned}
&\left(\mathbf{H}^{\mathrm{pr}}\left[\begin{pmatrix} L_1+H_1 \\ L_2+H_2 \end{pmatrix}\right] - \mathbf{H}^{\mathrm{pr}}\left[\begin{pmatrix} L_1 \\ L_2 \end{pmatrix}\right] - D_{(L_1,L_2)^t}\mathbf{H}^{\mathrm{pr}}\left[\begin{pmatrix} H_1 \\ H_2 \end{pmatrix}\right]\right)(s,t,s',t') \\
&= \left(H_1(s,s')L_2(t,t') + L_1(s,s')H_2(t,t') - 2L_1(s,s')L_2(t,t')\frac{\int_{K_1^2} L_1(u,u')H_1(u,u')du\,du'}{\int_{K_1^2}(L_1(u,u'))^2\,du\,du'}\right) \\
&\quad \times \left(\frac{1}{\int_{K_1^2}((L_1+H_1)(u,u'))^2\,du\,du'} - \frac{1}{\int_{K_1^2}(L_1(u,u'))^2\,du\,du'}\right)
\end{aligned}$$

显然，第一个因子的阶为 $\mathcal{O}(\max\{\|H_1\|, \|H_2\|\})$。此外，经过简单计算可知，第二个因子也具有相同的速率。由此完成了映射 $\mathbf{G}^{\mathrm{pr}}$ 和 $\mathbf{H}^{\mathrm{pr}}$ 的可微性证明；根据链式法则以及恒等式 $\mathbf{F}^{\mathrm{pr}} = \mathbf{H}^{\mathrm{pr}} \circ \mathbf{G}^{\mathrm{pr}}$，偏乘积近似的可微性也随之成立。

□

A.3 SPCA 的可微性

为使以下推导更易于阅读，对于紧集 K，我们将连续对称核 $\mathcal{C}(K^2)^{Sym}$ 所构成的空间定义为

$$\mathcal{C}(K^2)^{Sym} := \{A \in \mathcal{C}(K^2) : A(x,y) = A(y,x)\ \forall x, y \in K\}.$$

任意核 A 都表示相应的希尔伯特–施密特（积分）算子。根据作用于 $L^2[0,1]$ 上的正规算子的谱定理，该算子可分解为

$$A(x,y) = \sum_{i\in\mathbb{N}} v_i^A(x)v_i^A(y)\lambda_i^A, \tag{A.2}$$

其中 $\{v_i^A\}_{i\in\mathbb{N}}$ 是特征函数，$\{\lambda_i^A\}_{i\in\mathbb{N}}$ 是相应的特征值。不失一般性，以下我们假设对于任意 $i \geq 3$，都有 $|\lambda_1^A| \geq |\lambda_2^A| \geq |\lambda_i^A|$。注意，先验地，恒等式 (A.2) 仅在 L^2意义下成立，但可以证明它在连续函数空间中也成立（相对于上确界范数），

参见 Mercer 定理，例如 König (1986) 中的定理 3.a.1。此外，可以证明特征函数可以选取一个连续代表元（参见引理 C.1）。因此，以下我们将假设 v_i^A是 $\mathcal{C}(K)$中的这个代表元。现在假设 $A_0 \in \mathcal{C}(K^2)^{Sym}$ 是一个满足$|\lambda_1^{A_0}| > |\lambda_2^{A_0}|$的核。这意味着 A_0 的第一个特征值是唯一的，且第一个特征函数（对应于 $\lambda_1^{A_0}$）也是唯一的。由于特征函数通常仅确定到一个符号，我们在此假设已为 $v_1^{A_0}$固定了某个符号选择。于是可得，存在一个充分小的 $\delta = \delta(A_0) > 0$，使得对于任意满足$\|A - A_0\| < \delta$ 的 $A \in \mathcal{C}(K^2)^{Sym}$，都有 $|\lambda_1^A| > |\lambda_2^A|$，从而使 A 的第一个特征值唯一。这也（差一个符号）确定了 A 的第一个特征函数 v_1^A，我们可以固定使距离$\int_K (v_1^{A_0}(t) - v_1^A(t))^2 dt$ 最小的符号选择（详见引理 C.3）。现在，“特征映射”

$$\Lambda : \begin{cases} U_\delta(A_0) \to \mathbb{R}, \\ A \mapsto \lambda_1^A \end{cases} \qquad V : \begin{cases} U_\delta(A_0) \to \mathcal{C}(K), \\ A \mapsto v_1^A \end{cases} \tag{A.3}$$

定义良好。由于可分离近似本质上是算子的秩一近似，证明 SPCA 近似可微性的关键步骤是证明特征函数映射和特征值映射的可微性。已知特征映射在 L^2 意义下可微（另见 Dette 等人（2022）中定理 3.4 的证明），因此我们只需验证微分仍是“差商”（如上述证明中所述）的极限。对于特征函数映射，这需要采用一种非标准的微分表示，以确保其映射结果仍属于连续函数空间。

引理 A.5。 假设 $A_0 \in \mathcal{C}(K^2)^{Sym}$ 满足$|\lambda_1^{A_0}| > |\lambda_2^{A_0}|$。那么当 $\delta = \delta(A_0)$ 充分小时，特征映射（定义见 (A.3)）是 *Fréchet* 可微的。

证明。 将 Dette 等人（2022）中的定理 3.4 的证明稍作调整，即可轻松得到特征值映射的情形，其微分为

$$D_{A_0}\Lambda[T] = \int_{K^2} T(x,y) v_1^{A_0}(x) v_1^{A_0}(y) dx\, dy.$$

相比之下，要为特征函数映射建立这一结果则更为复杂，因此值得详细讨论。这一次，我们不从微分入手，而是发现按步骤推导更为容易。为此，考虑 $H \in \mathcal{C}(K^2)^{Sym}$ with $\|H\| \to 0$ (尤其是，我们可以假设 $A_0 + H \in U_\delta(A_0)$)。现在我们来考察分解

$$\begin{aligned} v_1^{A_0+H} - v_1^{A_0} &= \frac{(A_0+H)}{\lambda_1^{A_0+H}} \left[v_1^{A_0+H}\right] - \frac{A_0}{\lambda_1^{A_0}} \left[v_1^{A_0}\right] \\ &= \left(\frac{1}{\lambda_1^{A_0+H}} - \frac{1}{\lambda_1^{A_0}}\right)(A_0+H)\left[v_1^{A_0+H}\right] + \frac{(A_0+H)}{\lambda_1^{A_0}}\left[v_1^{A_0+H} - v_1^{A_0}\right] + \frac{H}{\lambda_1^{A_0}}\left[v_1^{A_0}\right] \\ &=: T_1 + T_2 + T_3 \end{aligned}$$

接下来，我们将分别求出每一项的微分。先从第一项 T_1 开始，并注意到 $T_1 = -(\lambda_1^{A_0+H} - \lambda_1^{A_0})$ $v_1^{A_0+H}/\lambda_1^{A_0+H}$，因为 $(A_0+H)\left[v_1^{A_0+H}\right] = \lambda_1^{A_0+H} v_1^{A_0+H}$。我们声称

该微分为 $H \mapsto -v_1^{A_0}/\lambda_1^{A_0} \cdot \int_{K^2} H(x,y)v_1^{A_0}(x)v_1^{A_0}(y)dx\,dy$，显然关于上确界范数是线性且连续的。通过添加和减去交叉项，得到以下界

$$\left\| T_1 + \frac{v_1^{A_0}}{\lambda_1^{A_0}} \int_{K^2} H(x,y)v_1^{A_0}(x)v_1^{A_0}(y)dx\,dy \right\| \leq \int_{K^2} H(x,y)v_1^{A_0}(x)v_1^{A_0}(y)dx\,dy \frac{\|v_1^{A_0+H} - v_1^{A_0}\|}{\lambda_1^{A_0}}$$
$$+ \frac{\|v_1^{A_0}\|}{\lambda_1^{A_0}} \left| \int_{K^2} H(x,y)v_1^{A_0}(x)v_1^{A_0}(y)dx\,dy - (\lambda_1^{A_0+H} - \lambda_1^{A_0}) \right|$$

这两项的阶均为 $\mathcal{O}\left(\|H\|^2\right)$。对于第一项，我们注意到积分显然为 $\mathcal{O}\left(\|H\|^2\right)$ 阶，而特征函数之差的界 $\mathcal{O}\left(\|H\|^2\right)$ 可由引理 C.2 的第 ii 部分推出）。右侧第二项为 $\mathcal{O}\left(\|H\|^2\right)$ 阶，这是由于特征值映射的可微性（见本证明开头），因此

$$D_{A_0}\Lambda[H] = \int_{K^2} H(x,y)v_1^{A_0}(x)v_1^{A_0}(y)dx\,dy$$

（详情请参阅 Dette 等人（2022）。）
我们继续分析 T_2。注意，根据 Kokoszka 和 Reimherr（2013）中引理2的（第二个）不等式，有

$$\|H[v_1^{A_0+H} - v_1^{A_0}]/\lambda_1^{A_0}\| \leq \frac{\|H\|}{|\lambda_1^{A_0}|} \left(\int_K \left(v_1^{A_0+H} - v_1^{A_0} \right)^2 (x)dx \right)^{1/2} < \kappa\|H\|^2.$$

这意味着 意味着 $T_2 = A_0[v_1^{A_0+H} - v_1^{A_0}]/\lambda_1^{A_0} + \mathcal{O}(\|H\|^2)$。我们可以重写不可忽略项 a s

$$\frac{A_0}{\lambda_1^{A_0}} \left[v_1^{A_0+H} - v_1^{A_0} \right] = \frac{A_0}{\lambda_1^{A_0}} \left[\sum_{i\geq 1} \left(\int_K \left(v_1^{A_0+H}(x) - v_1^{A_0}(x) \right) v_i^{A_0}(x)dx \right) v_i^{A_0} \right] \tag{A.4}$$

注意，右侧不过是 $v_1^{A_0+H} - v_1^{A_0}$ 关于标准正交基 $\{v_i^{A_0}\}_{i\in\mathbb{N}}$ 的 L^2-基展开。由于 A_0 是积分算子，这一关系在 L^2 范数下显然成立，但在上确界范数下未必成立。不过，我们认为它确实成立。首先，注意两边都是定义良好的连续函数。此外，注意

$$\sup_x \left| \int_K A_0(x,y) \left[v_1^{A_0+H}(y) - v_1^{A_0}(y) - \sum_{i\geq 1} \left(\int_K \left(v_1^{A_0+H}(u) - v_1^{A_0}(u) \right) v_i^{A_0}(u)du \right) v_i^{A_0}(y) \right] dy \right|$$
$$\leq \|A_0\| \int_K \left| v_1^{A_0+H}(y) - v_1^{A_0}(y) - \sum_{i\geq 1} \left(\int_K \left(v_1^{A_0+H}(u) - v_1^{A_0}(u) \right) v_i^{A_0}(u)du \right) v_i^{A_0}(y)dy \right| = 0$$

其中 a_0 是 A_0 的连续核。利用 Kokoszka 和 Reimherr（2013）中引理 1 的第一条和第二条等式，关于

$$\left(\int_K \left(v_1^{A_0+H}(x) - v_1^{A_0}(x) \right) v_1^{A_0}(x)dx \right) v_1^{A_0} \quad \text{and} \quad \sum_{i\geq 2} \left(\int_K \left(v_1^{A_0+H}(x) - v_1^{A_0}(x) \right) v_i^{A_0}(x)dx \right) v_i^{A_0}$$

分别可见，式（A.4）的右端等于

$$-\frac{A_0}{2\lambda_1^{A_0}} \left[v_1^{A_0} \right] \int_K (v_1^{A_0+H}(x) - v_1^{A_0}(x))^2 dx + \frac{A_0}{\lambda_1^{A_0}} \left[\sum_{i>1} \frac{v_i^{A_0}}{\lambda_1^{A_0+H} - \lambda_i^{A_0}} \int_{K^2} H(x,y)v_1^{A_0+H}(x)v_i^{A_0}(y)dx\,dy \right]$$

其中，按照上述相同的论证，两者在上确界范数下相等。由 Kokoszka 和 Reimherr（2013）中引理2的证明可轻易推出

$\int_K (v_1^{A_0+H}(x) - v_1^{A_0}(x))^2 dx = \mathcal{O}(\|H\|^2)$，因此第一项可忽略。采用与前面类似的推理（并结合引理 C.2 中的界），我们便可证明

$$\begin{aligned}
&\frac{A_0}{\lambda_1^{A_0}} \left[\sum_{i>1} \frac{v_i^{A_0}}{\lambda_1^{A_0+H} - \lambda_i^{A_0}} \int_{K^2} H(x,y) v_1^{A_0+H}(x) v_i^{A_0}(y) dx\, dy \right] \\
=&\frac{A_0}{\lambda_1^{A_0}} \left[\sum_{i>1} \frac{v_i^{A_0}}{\lambda_1^{A_0} - \lambda_i^{A_0}} \int_{K^2} H(x,y) v_1^{A_0}(x) v_i^{A_0}(y) dx\, dy \right] + \mathcal{O}(\|H\|^2),
\end{aligned}$$

这里为避免冗余，省略了具体计算。第二行中的其余级数关于 H 是线性的。简单计算可知，它也是有界的。最后，我们注意到，项 T_3 本身已经是关于 H 的线性有界映射。因此，我们可以将分解改写为

$$\begin{aligned}
v_1^{A_0+H} - v_1^{A_0} = T_1 + T_2 + T_3 = -v_1^{A_0}/\lambda_1^{A_0} \int_{K^2} H(x,y) v_1^{A_0}(x) v_1^{A_0}(y) dx\, dy \\
+ \frac{A_0}{\lambda_1^{A_0}} \left[\sum_{i>1} \frac{v_i^{A_0}}{\lambda_1^{A_0} - \lambda_i^{A_0}} \int_{K^2} H(x,y) v_1^{A_0}(x) v_i^{A_0}(y) dx\, dy \right] + \frac{H}{\lambda_1^{A_0}} \left[v_1^{A_0} \right] + \mathcal{O}(\|H\|^2).
\end{aligned}$$

非消失部分就是 Fréchet微分。注意，在 L^2 中，这个微分还可以进一步简化为更常见的表达式 $\sum_{i>1} \frac{v_i^{A_0}}{\lambda_1^{A_0} - \lambda_i^{A_0}} \int_{K^2} H(x,y) v_1^{A_0}(x) v_i^{A_0}(y) dx\, dy$（这里用到了 $A_0[v_i^{A_0}] = \lambda_i^{A_0} v_i^{A_0}$ 在 L^2 意义下成立这一点）；相关例子见 Kokoszka 和 Reimherr（2013）。然而，这个表达式不一定仍属于连续函数，因为特征函数的连续性对所有 $i \in \mathbb{N}$ 未必都成立。以上讨论完成了关于特征映射可微性的证明。

□

定理A.6. 设 $K_1 \subset \mathbb{R}^p$, $K_2 \subset \mathbb{R}^q$ 为紧集，$C \in \mathcal{C}((K_1 \times K_2)^2)$ 为可分离协方差算子。那么，映射 $\mathbf{F}^{\mathrm{PCA}}$（定义见 (2.8)）关于 C 的上确界范数弗雷歇可微。

证明。 既已证明对称算子的第一个特征值和特征函数具有可微性，接下来的证明便可逐步仿照 Dette 等人（2022）中定理 3.4 的证明。

□

B 定理 3.6 的证明

我们证明在连续函数空间中的弱收敛，并依赖 van der Vaart and Wellner（1996）所述的有界函数空间弱收敛理论。为此，需要验证两个条件：紧性和边际分布收敛，见他们的定理 1.5.4。首先，我们使用 Cramér–Wold 方法证明协方差算子的边际分布弱收敛（见 Billingsley（2012）中的定理 29.4）。我们采用经典的分块技术（用于处理自助法部分），并结合 Yoshihara（1978）针对强混合随机变量提出的矩不等式。其次，我们通过建立渐近等度连续性来证明紧性（例如，参见 van der Vaart and Wellner（1996）中的定理 1.5.7）。随后通过控制熵界来实现这一点。我们定义集合的覆盖数和熵。

为简洁起见，我们为所关注的对象引入以下逐点记号。因此，对于 (s,t)、$(s',t') \in K_1 \times K_2$，我们设

$$\tilde{X}_i(s,t) := X_i(s,t) - \mathbb{E}X_i(s,t)\,; \quad \tilde{C}_N(s,t,s',t') := \frac{1}{N}\sum_{i=1}^{N} \tilde{X}_i(s,t)\tilde{X}_i(s',t'),$$

and

$$\tilde{B}_N^{(k)}(s,t,s',t') := \frac{1}{N}\sum_{i=1}^{N}(\tilde{X}_i(s,t)\tilde{X}_i(s',t') - C(s,t,s',t'))\, w_{i,N}.$$

第二个等式由求和次序变换得到，其中对任意 $j \le 0$，我们令 $Z_j^{(k)} = 0$ （确定性的）。在最后一个等式中，我们以显然的方式定义了（随机）权重 $w_{i,\ N}$。
为使证明更加清晰，我们作两处简化：首先，我们不证明向量 $\sqrt{N}(\hat{}C_N - C, B_N^{(1)}, \cdots, B_N^{(r)})$ 的收敛，而只证明其中心化版本的收敛，即 $\sqrt{N}(\tilde{C}_N - C, \tilde{B}_N^{(1)}, \cdots, \tilde{B}_N^{(r)})$ 的收敛。根据下文证明中使用的类似但更简单的方法，可以证明这些向量之差为 $o_{\mathbb{P}}(1)$ 阶，因此略去。（这是 $X_i, i \ge 1$ 也满足中心极限定理这一事实的直接推论。）其次，不失一般性，我们令 $r = 1$，因为对 $r > 1$ 进行调整很简单，只会增加记号负担。

步骤 1：首先，我们通过 Cramér–Wold装置证明有限维分布的弱收敛（例如，参见 Billingsley（2012）中的定理 29.4）。为此，固定但任取 $p \in \mathbb{N}$，并考虑任意元组 $(s_m,t_m,s'_m,t'_m) \in (K_1 \times K_2)^2$ 以及数值 $a_m, a'_m \in \mathbb{R}$，其中 $1 \le j \le p$。请回想，要应用 Cramér–Wold装置，我们必须证明实值随机变量的弱收敛

$$CW_N := \sqrt{N}\sum_{m=1}^{p}\left\{a_m\left[\tilde{C}_N(s_m,t_m,s'_m,t'_m) - C(s_m,t_m,s'_m,t'_m)\right] + a'_m\tilde{B}_N^{(1)}(s_m,t_m,s'_m,t'_m)\right\}$$

to

$$\sum_{m=1}^{p}\left(a_m G(s_m,t_m,s'_m,t'_m) + a'_m G^{(1)}(s_m,t_m,s'_m,t'_m)\right),$$

其中 $G, G^{(1)}$ 是两个独立同分布的高斯过程。我们采用分块技术。根据定义，可将 CW_N 重写为

$$\begin{aligned} CW_N &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{m=1}^{p}\left(\tilde{X}_i(s_m,t_m)\tilde{X}_i(s'_m,t'_m) - C(s_m,t_m,s'_m,t'_m)\right)(a_m + a'_m w_{i,N}) \\ &= \sqrt{\frac{b_N}{N}}\sum_{j=1}^{\lfloor N/b_N\rfloor}\sum_{i=(j-1)b_N+1}^{jb_N}\sum_{m=1}^{p}\left(\tilde{X}_i(s_m,t_m)\tilde{X}_i(s'_m,t'_m) - C(s_m,t_m,s'_m,t'_m)\right)\frac{(a_m + a'_m w_{i,N})}{\sqrt{b_N}} + \mathrm{Rem}_1\,. \end{aligned}$$

这里，$(b_N)_{N\in\mathbb{N}}$是一个自然数序列，使得 $l_N/b_N \to 0$且$b_N/\sqrt{N} \to 0$，而 Rem_1是一个余项，用于涵盖索引介于 $b_N\lfloor N/b_N\rfloor$和 N之间的所有“悬余项”。利用三角不等式（并对项进行计数），不难看出 $\mathbb{E}|\mathrm{Rem}_1| = \mathcal{O}(b_N/\sqrt{N}) = o(1)$，因此

$\text{Rem}_1 = o_{\mathbb{P}}(1)$。为便于后文引用，现定义随机变量

$$Y_j := \sum_{i=(j-1)b_N+1}^{jb_N} \sum_{m=1}^{p} \left(\tilde{X}_i(s_m, t_m)\tilde{X}_i(s'_m, t'_m) - C(s_m, t_m, s'_m, t'_m) \right) \frac{(a_m + a'_m w_{i,N})}{\sqrt{b_N}} \tag{B.1}$$
$$= \sum_{i=(j-1)b_N+1}^{jb_N} \sum_{m=1}^{p} \breve{Y}_{i,m} \frac{(a_m + a'_m w_{i,N})}{\sqrt{b_N}},$$

其中，$\breve{Y}_{i,\ m}$ 以显而易见的方式定义。由此，我们可以写出 $CW_N = \sqrt{\frac{b_N}{N}} \sum_j Y_j + o_{\mathbb{P}}(1)$。下文将借助 Wooldridge–White 的中心极限定理（White（2001）中的定理5.20），证明该（不可忽略）总和的弱收敛。因此，我们需要检验三个条件：$(Y_j)_{j=1,\ \cdots,\ \lfloor N/b_N \rfloor}$的混合系数充分快速地衰减、四阶矩一致有界，以及渐近方差存在。

混合性： 变量$(Y_j)_{j=1,\ \cdots,\ \lfloor N/b_N \rfloor}$构成一个由 α-混合随机变量组成的三角阵，其混合系数满足（对所有 $N \geq N_0$ 及某个充分大的 $N_0 \in \mathbb{N}$） $|j - j'| \geq 2$

$$\alpha(Y_j, Y_{j'}) \leq \kappa \left((|j - j'| - 1)(b_N - l_N + 2) \right)^{-a}. \tag{B.2}$$

这里我们结合假设3.2 iv)和随机变量 $Y_j, Y_{j'}$.~ ~的定义进行了推导。具体来说，我们利用了 Y_j 仅取决于函数 $X_{(j-1)b_N+1}, \cdots, X_{jb_N}$ 以及$w_{(j-1)b_N+1,\ N}, \cdots, w_{jb_N,\ N}$这一点。特别地， $|j - j'| \geq 2$的混合系数（一致地）收敛到0，因此 Wooldridge–White 定理中的衰减条件显然得到满足。

有界四阶矩： 现在我们希望证明， $\mathbb{E}[Y_j^4] \leq \kappa < \infty$ 在 j 中一致成立，并从以下计算开始

$$\begin{aligned}
\mathbb{E}[Y_j^4] \leq & \kappa \sum_{m=1}^{p} \mathbb{E}\left[\left| \sum_{i=(j-1)b_N+1}^{jb_N} \breve{Y}_{i,m} \frac{(a_m + a'_m w_{i,N})}{\sqrt{b_N}} \right|^4 \right] \\
= & \kappa \sum_{m=1}^{p} \mathbb{E}\left[\mathbb{E}\left[\left| \sum_{i=(j-1)b_N+1}^{jb_N} \breve{Y}_{i,m} \frac{(a_m + a'_m w_{i,N})}{\sqrt{b_N}} \right|^4 \middle| w_{1,N}, \ldots, w_{N,N} \right] \right] \\
\leq & \kappa \sum_{m=1}^{p} \mathbb{E}\left[\left(\sum_{i=(j-1)b_N+1}^{jb_N} \left(\frac{(a_m + a'_m w_{i,N})}{\sqrt{b_N}} \right)^2 \right)^4 \right] \leq \kappa.
\end{aligned} \tag{B.3}$$

在第一个不等式中，常数 κ 仅取决于 p。在第一个等式中，我们对 $w_{i,N}$,$1 \leq i \leq N$ 进行条件化（利用塔式性质）。因此，我们可以应用 Yoshihara（1978）的定理 3，得到第二个不等式。注意，我们假设 $\breve{Y}_{i,m}$ 具有有限的 $\gamma/2$ 阶矩；根据假设 3 的 iii)，这成立，因为 2 X_i 具有有限的 γ 阶矩。在上述定理中，我们选择参数 $m = 4$ 和 $\delta = \gamma/2 - 4$，其可和条件（ii）满足，因为根据我们的假设 3.2 的 iv)，$a > 2\gamma/(\gamma - 8)$。该不等式中的常数 κ 仅取决于混合系数的可和性（更多细节请参见 Yoshihara（1978）的定理 3 的证明）以及（ $\breve{Y}_i$的）矩，与其他因素无关。对于第三个不等式，我们注意到 $(a_m + a'_m w_{i,N})^2$ 具有任意阶的有界矩（$w_{i,N}$ 服从方差为 1 的正态分布），因此，通过对项计数即可得到最后一个不等式，其中

κ 现在还取决于 a_m, a'_m, $1 \le m \le p$。

方差的收敛：接下来，我们需要证明长期方差存在。为此，我们将其分解为

$$\mathbb{E}\left[\left(\sqrt{\frac{b_N}{N}}\sum_j Y_j\right)^2\right] = \frac{b_N}{N}\sum_{j=1}^{\lfloor N/b_N\rfloor}\mathbb{E}Y_j^2 + \frac{b_N}{N}\sum_{|j-j'|=1}|\mathbb{E}Y_jY_{j'}| + 2\sum_{h\ge 2}\sup_j|\mathbb{E}Y_jY_{j+h}|. \tag{B.4}$$

利用 混合条件 (B.2)，我们看到，对于充分大的 N，最后一个和可以被以下量所界定 y

$$\kappa b_N^{-a(\gamma-4)/\gamma}\sup_j(\mathbb{E}|Y_j|^{\gamma/2})^{4/\gamma}\sum_{h\ge 2}h^{-a(\gamma-4)/\gamma} = o(1).$$

这里我们使用了适用于 α-混合的标准协方差不等式（参见 Dehling 等人（2002）中的引理 3.11）来界定 $|\mathbb{E}Y_jY_{j+h}|$；对于较小的 α-速率，右侧的和在 $a > \gamma/(\gamma-4)$ 下收敛（参见假设 3.2 iv）。类似地，我们可以将另一个协方差和的上界估计为

$$\frac{b_N}{N}\sum_{|j-j'|=1}|\mathbb{E}Y_jY_{j'}| \le \kappa\sup_j|\mathbb{E}Y_jY_{j+1}|.$$

我们进一步对协方差 $|\mathbb{E}Y_jY_{j+1}|$ 进行界定（并证明它关于 j 一致收敛于 0）。为此，令 $(s_N)_{N\in\mathbb{N}}$表示一个递增的自然数序列，满足 $l_N/s_N \to 0$和 $s_N/b_N \to 0$。回顾 (B.1)，现在可以将其拆分为

$$Y_j = \sum_{i=jb_N+1}^{(j+1)b_N-s_N}\sum_{m=1}^{p}\breve{Y}_{i,m}\frac{(a_m+a'_mw_{i,N})}{\sqrt{b_N}} + \sum_{i=(j+1)b_N-s_N+1}^{(j+1)b_N}\sum_{m=1}^{p}\breve{Y}_{i,m}\frac{(a_m+a'_mw_{i,N})}{\sqrt{b_N}} =: Y_{j,1}+Y_{j,2},$$

$$Y_{j+1} = \sum_{i=(j+1)b_N+1}^{(j+2)b_N-s_N}\sum_{m=1}^{p}\breve{Y}_{i,m}\frac{(a_m+a'_mw_{i,N})}{\sqrt{b_N}} + \sum_{i=(j+2)b_N-s_N+1}^{(j+2)b_N}\sum_{m=1}^{p}\breve{Y}_{i,m}\frac{(a_m+a'_mw_{i,N})}{\sqrt{b_N}} =: Y_{j+1,1}+Y_{j+1,2}.$$

这里的变量 $Y_{j,1}, Y_{j,2}, Y_{j+1,1}, Y_{j+1,2}$ 按照显而易见的方式定义。采用与 (B.3) 中相同的技巧，可得 $\mathbb{E}[(Y_{j,1})^4], \mathbb{E}[(Y_{j+1,1})^4] < \infty$，并且（计数各项） $\mathbb{E}[(Y_{j,2})^4], \mathbb{E}[(Y_{j+1,2})^4] = o(1)$。因此，由柯西–施瓦茨不等式可得 $\mathbb{E}[Y_{j,1}Y_{j+1,2}], \mathbb{E}[Y_{j,2}Y_{j+1,2}], \mathbb{E}[Y_{j,2}Y_{j+1,1}] = o(1)$。此外，利用针对 α-混合的协方差不等式，可得 $\mathbb{E}[Y_{j,1}Y_{j+1,2}] \le \{\mathbb{E}[(Y_{j,1})^4]\mathbb{E}[(Y_{j+1,1})^4]\}^{1/2}\alpha(Y_{j,1}, Y_{j+1,1})^{1/2}$。现在，$\alpha(Y_{j,1}, Y_{j+1,1}) \le \kappa(s_N - l_N)^{-a} = o(1)$，这里我们使用了假设 3.2 iv)以及随机变量 $Y_{j,1}, Y_{j+1,1}$ 的定义。因此请注意，$Y_{j,1}$ 仅依赖于函数 ~ ~$X_{jb_N+1}, \dots, X_{(j+1)b_N-s_N}$以及权重 $w_{jb_N+1,N}, \dots, w_{(j+1)b_N-s_N,N}$。这些考虑表明，(B.4) 右侧的第二项和第三项，即混合项，其阶为 $o(1)$，因此方差等于 $\frac{b_N}{N}\sum_{j=1}^{\lfloor N/b_N\rfloor}\mathbb{E}Y_j^2$。最后，我们需要证明该方差收敛。因此，令 $1 < j \le \lfloor N/b_N\rfloor$, $\mathbb{E}Y_j^2$

$$= \mathbb{E}\left[\mathbb{E}\left[\left(\sum_{m=1}^{p}\sum_{i=jb_N+1}^{(j+1)b_N}\breve{Y}_{i,m}\frac{(a_m+a'_mw_{i,N})}{\sqrt{b_N}}\right)^2\Big|w_{1,N}\dots,w_{N,N}\right]\right] \tag{B.5}$$

$$= \sum_{m,l=1}^{p}\frac{1}{b_N}\sum_{i,k=jb_N+1}^{(j+1)b_N}\mathbb{E}[\breve{Y}_{i,m}\breve{Y}_{i,l}]\,\mathbb{E}[(a_m+a'_mw_{i,N})(a_l+a'_lw_{k,N})]$$

˘这里我们用到了 $Y_{i,m}$ 与 $w_{i,N}$, $1 \le i \le N$ 的独立性。根据我们的平稳性假设 3.2，有 $\mathbb{E}[\breve{Y}_{i,m}\breve{Y}_{k,l}] = \tau^{m,l}_Y(|i-k|)$，其中函数 $\tau^{m,l}_Y : \mathbb{N} \to \mathbb{R}$。此外，根据构造，$w_{i,N}$ 与 $w_{k,N}$ 的协方差也仅依赖于 $|i-k|$ 和 N。因此，我们可以一致地定义

$$\tau^{m,l}_{w,N}(|i-k|) := \mathbb{E}[(a_m + a'_m w_{i,N})(a_l + a'_l w_{k,N})] = a_m a_l + a'_m a'_l \mathbb{E}[w_{i,N} w_{k,N}]$$

当 $N \to \infty$ 和 $\mathbb{E}[w_{i,N}w_{k,N}] \to 1$ 时收敛（根据权重的定义即可得出；参见本证明开头）。因此，我们可以将 (B.5) 重写为

$$\sum_{m,l=1}^{p} \sum_{|h|<b_N} \tau^{m,l}_Y(|h|)\tau^{m,l}_{w,N}(|h|)(1-|h|/b_N).$$

注意，这个对象不依赖于 j，并且收敛到长期方差 $\sum_{m,l=1}^{p} \sum_{h\in\mathbb{Z}} \tau^{m,l}_Y(|h|)$。后一个收敛性可直接由控制收敛定理得出。事实上，首先注意到，对于 $(1-|h|/b_N) \le 1$，有 $|\tau^{m,l}_{w,N}(|h|)| \le \kappa$。其次，对于任意 m、l，项 $\tau^{m,l}_Y(|h|)$ 的级数均可求和，这同样可由关于 α-混合随机变量的协方差不等式得出。令 $|i-j| = |h| \ge 1$，则

$$\tau^{m,l}_Y(|h|) = \mathbb{E}[\breve{Y}_{i,m}\breve{Y}_{k,l}] \le \kappa\{\mathbb{E}[|\breve{Y}_{i,m}|^{\gamma/2}]\}^{2/\gamma}\{\mathbb{E}[|\breve{Y}_{k,l}|^{\gamma/2}]\}^{2/\gamma}(|h|+1)^{-a(\gamma-4)/\gamma},$$

由于 $a > \gamma/(\gamma-4)$ 且矩一致有界，该级数可求和，参见假设3.2的 iii)。通过简单修改可知，$\mathbb{E}Y_1^2$ 也收敛到相同的方差。根据上述讨论，(B.4) 中的方差收敛，因此我们可以应用 Wooldridge–White 的中心极限定理（White（2001）中的定理5.20），由此可得边际分布收敛。

剩下需要证明的是，$\sqrt{N}\sum_{p=1}^{N} a_m(\tilde{C}_N - C)(s_m,t_m,s'_m,t'_m)$ 和 $\sqrt{N}\sum_{p=1}^{N} a'_m \tilde{B}^{(1)}_N(s_m,t_m,s'_m,t'_m)$ 渐近独立，且具有相同的（渐近）方差。由序列 $(X_i)_{i\in\mathbb{Z}}$ 的不相关性、权重以及高斯极限，可直接得到渐近独立性。至于它们的方差相等，这可通过简单计算得出，所用技巧与证明第一部分时类似。

在通过渐近等度连续性证明紧性之前，为完整起见，我们先给出打包数的定义。

定义 B.1. 设 $(\mathcal{X},d)$ 为一个度量空间，且 $B(x,r)$ 为一个半径为 $r>0$、以 $x\in\mathcal{X}$ 为中心的球。于是，对于 $\varepsilon>0$，我们将 ε-打包数 $D(\varepsilon,d)$ 定义为

$$\sup\left\{ n\in\mathbb{N} \,|\, \bigcup_{i=1}^{n} B(x_i,\varepsilon) \supset \mathcal{X} \text{ where } d(x_i,x_j)>\varepsilon,\ x_i\in\mathcal{X},\ 1\le i\le n \right\}.$$

显然，$\varepsilon>0$ 越小，打包数就越大；对于任何全有界集，打包数都有限。

在证明的后续部分，对于 $K\subset\mathbb{R}^p$，我们设 $\rho_K(x,y) := 1 \wedge \max_{i=1}^{p} |x_i - y_i|$。

步骤 2：我们证明过程 $\sqrt{N}(\tilde{C}_N - C, \tilde{B}^{(1)}_N)$ 依概率渐近一致地 $\tilde{\rho}$-等度连续，其中

$$\tilde{\rho}((s,t,u,v),(s',t',u',v')) := \max\{\rho_{K_1\times K_2}((s,t),(s',t')), \rho_{K_1\times K_2}((u,v),(u',v'))\}$$

是我们在 $(K_1 \times K_2)^2$ 上的度量。此外，回顾 van der Vaart 和 Wellner（1996）中的定理 1.5.7：过程的渐近等度连续性等价于紧性。对于 $\zeta > 0$，定义如下的点对集合

$$\mathcal{A}_\zeta := \{((s,t,u,v),(s',t',u',v')) \in (K_1 \times K_2)^4 \,|\, \tilde{\rho}((s,t,u,v),(s',t',u',v')) < \zeta\}.$$

现在我们将对 $\epsilon > 0$ 进行界定，

$$\begin{aligned}
&\limsup_{N\to\infty} \mathbb{P}\left\{ \sup_{(\mathbf{x},\mathbf{x}')\in\mathcal{A}_\zeta} \left| \sqrt{N}(\tilde{C}_N - C, \tilde{B}_N^{(1)})(\mathbf{x}) - \sqrt{N}(\tilde{C}_N - C, \tilde{B}_N^{(1)})(\mathbf{x}') \right| > \varepsilon \right\} \\
&\leq \limsup_{N\to\infty} \mathbb{E}\left(\sup_{(\mathbf{x},\mathbf{x}')\in\mathcal{A}_\zeta} \left| \sqrt{N}(\tilde{C}_N - C, \tilde{B}_N^{(1)})(\mathbf{x}) - \sqrt{N}(\tilde{C}_N - C, \tilde{B}_N^{(1)})(\mathbf{x}') \right| \right)^J / \varepsilon^J. \qquad \text{(B.6)}
\end{aligned}$$

根据 van der Vaart 和 Wellner（1996）中的定理 2.2.4，只需界定两个位置处差值的 J 阶矩。更确切地说，对于下文将指定的 J，我们对以下量给出上界

$$\mathbb{E}\left| \sqrt{N}(\tilde{C}_N - C)(s,t,u,v) - \sqrt{N}(\tilde{C}_N - C)(s',t',u',v') \right|^J \qquad \text{(B.7)}$$

and

$$\mathbb{E}\left| \sqrt{N}\tilde{B}_N^{(1)}(s,t,u,v) - \sqrt{N}\tilde{B}_N^{(1)}(s',t',u',v') \right|^J. \qquad \text{(B.8)}$$

对 α-混合随机变量应用 Yoshihara（1978）中的定理 3：

$$\tilde{X}_1(s,t)\tilde{X}_1(u,v) - \mathbb{E}\left[\tilde{X}_1(s,t)\tilde{X}_1(u,v)\right] - \tilde{X}_1(s',t')\tilde{X}_1(u',v') + \mathbb{E}\left[\tilde{X}_1(s',t')\tilde{X}_1(u',v')\right]$$

取权重 $a_i = 1/\sqrt{N}$ 和（任意但较小的） $\delta > 0$，我们可知 (B.7) 小于

$$\kappa\mathbb{E}\left(\left| \tilde{X}_1(s,t)\tilde{X}_1(u,v) - \mathbb{E}\left[\tilde{X}_1(s,t)\tilde{X}_1(u,v)\right] - \tilde{X}_1(s',t')\tilde{X}_1(u',v') + \mathbb{E}\left[\tilde{X}_1(s',t')\tilde{X}_1(u',v')\right] \right| \right)^J \qquad \text{(B.9)}$$

只要 $\sum_{i\geq 1}(i+1)^{J/2+1}\alpha(i)^{\delta/(J+\delta)} < \infty$.

后一项将利用我们在假设 3.2(*ii*) 中所假定的随机变量的连续性性质来界定。事实上，对关系式 3.3 取期望，可知

$$|\tilde{X}_i(s,t) - \tilde{X}_i(s',t')| \leq (M + \mathbb{E}M)\rho_{K_1\times K_2}((s,t),(s',t'))^\beta$$

˜因此，中心化后的随机变量是 Hölder 连续的，其新的随机常数为 $M := M + \mathbb{E}M$。简单计算可得

$$\left| \tilde{X}_1(s,t)\tilde{X}_1(u,v) - \tilde{X}_1(s',t')\tilde{X}_1(u',v') \right| \leq 2\|\tilde{X}_1\|\, \tilde{M}\tilde{\rho}^\beta((s,t,u,v),(s',t',u',v')),$$

与上文类似

$$\left| \mathbb{E}\left(\tilde{X}_1(s,t)\tilde{X}_1(u,v)\right) - \mathbb{E}\left(\tilde{X}_1(s,t)\tilde{X}_1(u,v)\right) \right| \leq 2\mathbb{E}|\tilde{M}|\tilde{\rho}^\beta((s,t,u,v),(s',t',u',v')).$$

利用上述两个界，我们可以将 (B.9) 的上界写为

$$\kappa\mathbb{E}\left[\tilde{M}^J \|X_1\|^J\right] \tilde{\rho}^{J\beta}((s,t,u,v),(s',t',u',v')),$$

其中 κ 可以依赖于 β、J，但不依赖于 N。

为界定 (B.8)，我们首先以权重 $w_{i\ N}^{(1)}$、$1 \leq i \leq N$为条件。接下来的论证与界定 (B.7) 时相同，即直接应用 Yoshihara（1978）中的定理 3。由此得到以下界

$$\kappa \mathbb{E}\left\{\left(\sum_{i=1}^{N} \frac{\left(w_{i,N}^{(1)}\right)^2}{N}\right)^J \mathbb{E}\left[\left((\tilde{X}_1 \cdot \tilde{X}_1 - C)(s,t,u,v) - (\tilde{X}_1 \cdot \tilde{X}_1 - C)(s',t',u',v')\right)^J \left| w_{i,N}^{(1)}, 1 \leq i \leq N \right.\right]\right\}$$

$$\leq \kappa \mathbb{E}\left(\sum_{i=1}^{N} \frac{\left(w_{i,N}^{(1)}\right)^2}{N}\right)^J \mathbb{E}\left(\tilde{M}^J \left\|X_1\right\|^J\right) \tilde{\rho}^{J\beta}((s,t,u,v),(s',t',u',v'))$$

其中我们也用到了 $\tilde{X}_i$与 $w_{i\ N}^{(1)}$ 的独立性。对l_N–相依序列$w_{i,N}$ 再次应用 Yoshihara 的定理 3，可得其 J 阶矩有界；由于 $l_N/N \to 0$，该矩仍然有限，如 $N \to \infty$ 所示。回顾一下，$\tilde{\rho} = \max\{\rho_{K_1}, \rho_{K_2}\}$，且有限维空间上的所有范数都等价。由于一般的有限维空间中下述界成立（例如参见 van der Vaart 和 Wellner（1996）第 2.1 节例 6）

$$D(\varepsilon, \tilde{\rho}^\beta) < \frac{\kappa}{\varepsilon^{2(d_1+d_2)/\beta}}$$

（这里 κ 取决于 $d_i := \dim(K_i)$、$i = 1$、2 以及 K_i的直径）。选择参数 $J = \lceil 2(d_1+d_2)/\beta \rceil + 1$，并结合马尔可夫不等式与 van der Vaart and Wellner（1996）中的定理 2.2.4，对于任意 $\nu > 0$，表达式 (B.6) 小于 $(\kappa/\varepsilon^J)(\eta^{-2(d_1+d_2)/(J\beta)+1} + \zeta\eta^{-4(d_1+d_2)^2/(J\beta^2)})$；通过选取足够小的 ζ，再选取足够小的 ν，该值可以任意小。因此，我们的过程是 $\tilde{\rho}$-依概率等度连续的。

C 补充结果

在本节中，我们始终假设算子 A 的特征值满足 $|\lambda_1^A| \geq |\lambda_2^A| \geq |\lambda_i^A|$，对任意 $i \geq 3$均成立。

引理 C.1. 设 $A \in \mathcal{C}(K^2)^{Sym}$ 为一个核，其特征值为 $|\lambda_1^A| > |\lambda_2^A|$，则我们可以找到首批特征函数 $v_1^A \in \mathcal{C}(K)$的一个连续代表元。

证明. 首先注意到，在 L^2意义下，等式 $v_1^A = A[v_1^A]/\lambda_1^A$成立。现在将 v_1^A逐点定义为表达式 $A[v_1^A]/\lambda_1^A$，可以看出它本身已经连续，因为

$$\begin{aligned} v_1^A(t+h) - v_1^A(t) &= (\lambda_1^A)^{-1} \int_K [A(s,t+h) - A(s,t)] v_1^A(s) ds \\ \leq &(\lambda_1^A)^{-1} \Big\{ \int_K [A(s,t+h) - A(s,t)]^2 ds \Big\}^{1/2} \leq \kappa \sup_t |A(s,t+h) - A(s,t)| = o(1). \end{aligned}$$

这里在第一个不等式中用到了柯西–施瓦茨不等式。小-o 表示当 $|h| \to 0$时的收敛；这是因为连续核 A 在紧集 K^2上一致连续。 □

引理 C.2. 设 $A_0 \in \mathcal{C}(K^2)^{Sym}$ 满足 A_0，且$|\lambda_1^{A_0}| > |\lambda_2^{A_0}|$成立。此外，对某个 $\delta > 0$，考虑算子 $A \in \mathcal{C}(K^2)^{Sym}$，其中 $\|A - A_0\| \leq \delta$，并选取一个特征函数，使得 $\int_K v_1^{A_0}(t) v_1^A(t) dt \geq 0$。那么存在常数 $\kappa := \kappa(A_0, \delta)$

i) 对于 $j = 1, 2$

$$\Big| |\lambda_j^A| - |\lambda_j^{A_0}| \Big| \leq \kappa \|A - A_0\|.$$

ii)

$$\Big\{ \int_K (v_1^A(t) - v_1^{A_0}(t))^2 dt \Big\}^{1/2} \leq \kappa \|A - A_0\|.$$

这些恒等式可直接由 Horváth 和 Kokoszka（2012）中的 *i)*（引理 2.2–3）应用于算子 A 和 A_0 得出。请注意，这里我们利用了上确界范数强于 L^2-范数这一事实。

引理 *C.3.* 设 $A_0 \in \mathcal{C}(K^2)^{Sym}$ 为一个核，其特征值为$|\lambda_1^{A_0}| > |\lambda_2^{A_0}|$，特征函数为 $v_1^{A_0}$（其中已固定特征函数符号的一种取法）。那么存在一个充分小的 $\delta = \delta(A_0) > 0$，使得对于任意满足 $\|A - A_0\| \leq \delta$ 的 $A \in \mathcal{C}(K^2)^{Sym}$，均有 $|\lambda_1^A| > |\lambda_2^A|$，且存在 v_1^A 的一种符号取法，使得 $\int_K (v_1^{A_0}(t) - v_1^A(t))^2 dt < \int_K (v_1^{A_0}(t) + v_1^A(t))^2 dt$。

证明。 该证明直接由前面的引理 C.2 得出。注意，

$$|\lambda_1^A| = (|\lambda_1^A| - |\lambda_1^{A_0}|) + (|\lambda_1^{A_0}| - |\lambda_2^{A_0}|) + (|\lambda_2^{A_0} - |\lambda_2^A|) + |\lambda_2^A| \geq |\lambda_2^A| - \kappa\delta,$$

其中 κ 来自引理 C.2。当 $\delta > 0$ 足够小时，不等式 $|\lambda_1^A| > |\lambda_2^A|$ 成立。最后，不等式 $\int_K (v_1^{A_0}(t) - v_1^A(t))^2 dt < \int_K (v_1^{A_0}(t) + v_1^A(t))^2 dt$ 直接由引理 C.2 的第 ii 部分得出，同样要求 δ 足够小。

□