arXiv:2303.00702v2 [math.PR] 2 Mar 2023

希尔伯特空间中随机流的 Karhunen–Loève 定理

Leonardo V. Santoro Kartik G.Waghmare Victor M.
Panaretosleonardo.santoro@epfl.ch kartik.waghmare@epfl.ch victor.panaretos@epfl.ch

洛桑联邦理工学院数学研究所

2023年3月3日

摘要

我们将 Mercer定理推广至无限维希尔伯特空间中的算子值核。随后，我们应用这一结果推导出一个适用于均方连续希尔伯特型函数数据（即希尔伯特空间中的流）的 Karhunen–Loève 定理。也就是说，我们证明了希尔伯特空间中平方可积随机流的级数展开，其系数彼此不相关，且该展开在时间上是一致成立的。

MSC2020类别： 60G12,62R10
关键词： Mercer定理、函数主成分、希尔伯特型函数数据、随机级数展开

1 引言

卡鲁宁–洛夫定理是关于随机过程的一项基本结果，在随机过程的概率构造、数值分析和统计推断中发挥着核心作用（Adler，2010；Le Maître 和 Knio，2010；Hsing 和 Eubank，2015年）。在最简单的形式下，它将一个以时间为索引的实值随机函数分解为可数个“互正交”傅里叶级数项，从而将其随机成分与函数成分分离：基函数是确定性的正交函数，其系数则是不相关的随机变量。当基函数按系数方差递减的顺序排列时，k 截断展开可在均方意义下为该过程提供最佳的 k维近似。重要的是，对于均方连续的随机函数，该级数可以作逐点解释。这些性质说明，该展开在函数数据分析领域发挥（并将继续发挥）着催化作用（Wang 等人，2016）；事实上，该领域的起源正是 Grenander 将这一展开用作对随机函数进行推断的坐标系（Grenander，1950）。

传统上，函数型数据分析主要关注定义在紧区间上的实值函数。然而，函数型数据的复杂性正日益超出这一范畴。现代函数型数据可以表现为函数流（从区间映射到函数空间的随机映射），或更一般地表现为函数型随机场（从欧几里得集合映射到函数空间的随机映射）。为了使这类数据能够应用函数型数据分析的工具，我们将卡鲁宁–洛维展开推广到取值于抽象可分希尔伯特空间（可能为无限维）的随机流（以及更一般的随机场）上（定理2）。为此，我们首先建立了取值于可分希尔伯特空间的非负定核的 Mercer定理版本（定理1）。

2 背景与问题陈述

设 $\mathbb{H}$ 为一个内积为 $\langle\cdot,\cdot\rangle_{\mathbb{H}} : \mathbb{H}\times\mathbb{H}\to\mathbb{R}$、诱导范数为 $\|\cdot\|_{\mathbb{H}} : \mathbb{H}\to\mathbb{R}_+$ 的可分希尔伯特空间，其中 $\dim(\mathbb{H}) \in \mathbb{N}\cup\{\infty\}$。我们分别用 $\mathscr{B}(\mathbb{H})$、$\mathscr{B}_1(\mathbb{H})$ 和 $\mathscr{B}_2(\mathbb{H})$ 表示 $\mathbb{H}$ 上有界、迹类（TC）和希尔伯特–施密特（HS）线性算子的空间，其相应范数为：

$$\|B\|_{\mathscr{B}(\mathbb{H})} := \sup_{h\in\mathbb{H},\|h\|=1} \|Bh\|, \quad \|B\|_{\mathscr{B}_1(\mathbb{H})} := \operatorname{tr}(\sqrt{B^*B}), \quad \|B\|_{\mathscr{B}_2(\mathbb{H})} := \sqrt{\operatorname{tr}(B^*B)},$$

其中，线性算子 $B\in\mathscr{B}(\mathbb{H})$ 的伴随算子 B^* 定义为：对所有 $u,v\in\mathbb{H}$，均有 $\langle Bu, v\rangle_{\mathbb{H}} = \langle u, B^*v\rangle_{\mathbb{H}}$。注意，$\|B\|_{\mathscr{B}(\mathbb{H})} \le \|B\|_{\mathscr{B}_2(\mathbb{H})} \le \|B\|_{\mathscr{B}_1(\mathbb{H})}$。对于 $u,v\in\mathbb{H}$，由 $(u\otimes_{\mathbb{H}} v)h = \langle u,h\rangle_{\mathbb{H}} v$ 定义的算子 $u\otimes_{\mathbb{H}} v : \mathbb{H}\to\mathbb{H}$ 是有界线性算子。

若有界算子 B 对任意有界序列 $\{h_n\}_{n\ge1}$（其中元素属于 $\mathbb{H}$），其像 $\{Bh_n\}_{n\ge1}$ 都含有收敛子序列，则称 B 为紧算子。设 $B\in\mathscr{B}(\mathbb{H})$ 是紧自伴随算子。记其特征向量为 $\{e_j\}_{j\ge1}$，对应的特征值为 $\{\lambda_j\}_{j\ge1}$，并按 $|\lambda_1|\ge|\lambda_2|\ge\dots$ 的顺序排列。则 $\{e_j\}_{j\ge1}$ 构成 $\operatorname{Im}(B)$ 的完备~~正交~~归一系统（CONS），因此我们可以写成 $B = \sum_{j\ge1}\lambda_j e_j\otimes_{\mathbb{H}} e_j$.

令 T 为欧几里得空间中的任意紧子集。我们将 $L^2(\mathrm{T};\mathbb{H})$ 记为平方可积的 $\mathbb{H}$值函数 $f:\mathrm{T}\to\mathbb{H}$ 所构成的希尔伯特空间，即：

$$L^2(\mathcal{T},\mathbb{H}) = \left\{ f : \int_{\mathcal{T}} \|f(t)\|_{\mathbb{H}}^2 dt < \infty \right\}.$$

相应的内积 $\langle\cdot,\cdot\rangle_{L^2(\mathrm{T};\mathbb{H})}$ 定义为 $\langle f,g\rangle_{L^2(\mathrm{T};\mathbb{H})} = \int_{\mathcal{T}}\langle f(t), g(t)\rangle_{\mathbb{H}} dt$，相应的范数 $\|\cdot\|_{L^2(\mathrm{T};\mathbb{H})}$ 定义为 $\|f\|^2_{L^2(\mathrm{T};\mathbb{H})} = \langle f,f\rangle_{L^2(\mathrm{T};\mathbb{H})}$。

令 $\chi\in L^2(\mathcal{T};\mathbb{H})$ 为 $\mathbb{H}$ 中具有有限二阶矩的（均值为零）随机流，$\mathbb{E}\|\chi\|^2_{L^2(\mathrm{T};\mathbb{H})} < \infty$。我们将 χ 称为 希尔伯特型流，如 Kim等人 （2020）所述，不过当 $\dim(\mathrm{T}) \ge 2$ 时，也可以称其为 希尔伯特型场 。记 $\mathscr{C}\in\mathscr{B}(L^2(\mathcal{T};\mathbb{H}))$ 为其协方差算子：

$$\mathscr{C} : L^2(\mathcal{T},\mathbb{H}) \to L^2(\mathcal{T},\mathbb{H}), \qquad \mathscr{C} = \mathbb{E}\left[\chi\otimes_{L^2(\mathcal{T},\mathbb{H})}\chi\right] \tag{1}$$

或者，等价地：

$$\langle \mathscr{C}U, V\rangle_{L^2(\mathcal{T},\mathbb{H})} := \mathbb{E}\left[\langle\chi,U\rangle_{L^2(\mathcal{T},\mathbb{H})}\langle\chi,V\rangle_{L^2(\mathcal{T},\mathbb{H})}\right], \quad \text{for } U,V\in L^2(\mathcal{T},\mathbb{H})$$

注意，$\mathscr{C}$ 是半正定的迹类算子。特别地，$\mathscr{C}$ 是紧的自伴算子；因此，它可按特征值—特征函数对进行如下谱分解（例如，Hsing 和 Eubank (2015年，定理 7.2.6）：

$$\mathscr{C} = \sum_{k=1}^{\infty} \lambda_k \Phi_k \otimes_{L^2(\mathcal{T},\mathbb{H})} \Phi_k,$$

其中，$\lambda_1 > \lambda_2 > \dots > 0$ 是特征值，$\Phi_k : L^2(\mathrm{T};\mathbb{H}) \to L^2(\mathrm{T};\mathbb{H})$是 $\mathscr{C}$对应的特征函数，它们构成一个完备正交归一系。
因此，χ具有如下分解

$$\chi = \sum_{k=1}^{\infty} \langle\chi,\Phi_k\rangle_{L^2(\mathcal{T},\mathbb{H})}\boldsymbol{\Phi}_k \tag{2}$$

其中，收敛是指在均值 $L^2(\mathcal{T}\ \mathbb{H})$ 范数意义下的收敛

$$\lim_{N\to\infty}\mathbb{E}\left\|\chi-\sum_{k=1}^{N}\langle\chi,\Phi_k\rangle_{L^2(\mathcal{T},\mathbb{H})}\Phi_k\right\|^2_{L^2(\mathcal{T},\mathbb{H})}=0 \tag{3}$$

并且还可以进一步证明 $\{\langle\log\Xi,\Phi_k\rangle_{L^2(\mathcal{T}\ \mathbb{H})}\}_{k\geq1}$ 是均值为零且不相关的随机变量。特别地，根据 Hsing 和 Eubank (2015年，定理 7.2.8)，上述分解是最优的：对于任意 $N<\infty$ 以及任意完备正交归一系 $\Psi_k\}_{k\geq1}$ 中的 $L^2(\mathcal{T}\ \mathbb{H})$，

$$\mathbb{E}\left\|\chi-\sum_{k=1}^{N}\langle\chi,\Phi_k\rangle_{L^2(\mathcal{T},\mathbb{H})}\Phi_k\right\|^2_{L^2(\mathcal{T},\mathbb{H})}\quad\leq\quad\mathbb{E}\left\|\chi-\sum_{k=1}^{N}\langle\chi,\Psi_k\rangle_{L^2(\mathcal{T},\mathbb{H})}\Psi_k\right\|^2_{L^2(\mathcal{T},\mathbb{H})}.$$

我们提出以下问题：

Is expansion (2) interpretable pointwise in $t\in\mathcal{T}$?
Is the convergence (3) valid uniformly over $t\in\mathcal{T}$?

当 $\mathbb{H}=\mathbb{R}$，并假设 χ具有均方连续性时，这两个问题长期以来已知可以通过著名的卡鲁宁–洛夫定理得到肯定回答（Karhunen (1946)，Loeve (1948)；另见 Kac 和 Siegert (1947)），其证明从根本上依赖于 Mercer 关于实值核分解的定理（Mercer，1909）。对取值于 d维欧几里得空间$\mathbb{H}=\mathbb{R}^d$ 的随机场也进行了扩展研究（Withers，1974），但一般的无限维情形尚未得到处理，也并不能 直接由 $\mathbb{H}=\mathbb{R}^d$情形推出。

3 算子值核的 Mercer 定理

考虑函数 $K:\mathrm{T}\ \times\mathrm{T}\ \to\mathscr{B}_1(\mathbb{H})$。我们将这样的函数 K 称为算子值核。如果 K 连续，则称其为连续核：

$$\|K(s,t)-K(s',t')\|_{\mathscr{B}_1(\mathbb{H})}\to0,\quad\text{as } t'\to t,s'\to s,$$

对任意 $s,t\in\mathrm{T}$ 均成立。如果 K满足 $K(s,t)=K(t,s)^*$，则称其为对称。最后，如果对任意 $n\geq1$以及序列 $\{v_j\}_{j=1,\dots,n}\subset\mathrm{T}$、$\{h_j\}_{j=1,\dots,n}\subset\mathbb{H}$，下式成立，则称 K为半正定：

$$\sum_{i,j=1}^{n}\langle K(v_i,v_j)h_i,h_j\rangle_{\mathbb{H}}\geq0. \tag{4}$$

给定一个算子值核，我们可以定义积分算子 $\mathscr{A}_K:L^2(\mathrm{T}\ \mathbb{H})\to L^2(\mathrm{T}\ \mathbb{H})$：

$$\mathscr{A}_K(f)(t)=\int_{\mathcal{T}}K(t,u)f(u)du \tag{5}$$

其中，式（5）中的积分应理解为 Bochner积分。

引理 1。设 K 为连续核。

(i) $\mathscr{A}_K$ 是紧的

(ii) 若 K 是对称的，则 $\mathscr{A}_K$ 是自伴的。

(iii) K 是半正定的，当且仅当 $\mathscr{A}_K$ 是半正定的。

证明。*(i)* 设 $\{e_i\}_{i\geq 1}$ 是 $\mathbb{H}$的一组完备正交归一系。那么，$\mathrm{P}_n := \sum_{i=1}^{n} e_i \otimes e_i$ 强收敛于恒等算子。对于 $n \geq 1$，令 $K_n(\cdot,\cdot) := \mathrm{P}_n K(\cdot,\cdot)\mathrm{P}_n$。注意到$\mathscr{A}_{K_n}$是紧的，因为它是有限秩的。因此，要证明 $\mathscr{A}_K$ 是紧的，只需证明当$n \to \infty$时，$\mathscr{A}_{K_n} \to \mathscr{A}_K$ 强收敛。现在：

$$\begin{aligned}\|(\mathscr{A}_{K_n} - \mathscr{A}_K)f\|_{L^2(\mathcal{T},\mathbb{H})} &= \left(\int_{\mathcal{T}} \|(K_n(t,u) - K(t,u))f(u)\|_{\mathbb{H}}^2 du\right)^2 \\ &\leq \sup_{t,s} \|(K_n(t,s) - K(t,s))\|_{\mathscr{B}(\mathbb{H})} \cdot \|f\|_{L^2(\mathcal{T},\mathbb{H})}.\end{aligned}$$

由连续性可知，要证明 $\mathscr{A}_K$ 的紧致性，只需证明:

$$\|(K_n(t,t') - K(t,s))\|_{\mathscr{B}(\mathbb{H})} \to 0, \quad \text{for all } t,s \in T.$$

注意， $P_n K(s,t) P_n \to K(s,t)$ 强收敛。由于 $K(s,t)$ 是希尔伯特–施密特算子，该收敛在算子范数下也成立。事实上，对于 $g,h \in \mathbb{H}$，我们可以写成

$$\begin{aligned}|\langle g, (K(s,t) - P_n K(s,t) P_n)h\rangle_{\mathbb{H}}| &\leq \sum_{i,j>n} |\langle K(s,t)e_i, e_j\rangle_{\mathbb{H}}| \cdot |\langle h, e_i\rangle_{\mathbb{H}}| \cdot |\langle g, e_j\rangle_{\mathbb{H}}| \\ &\leq \left[\sum_{i,j>n} |\langle K(s,t)e_i, e_j\rangle_{\mathbb{H}}|^2\right]^{1/2} \left[\sum_{i,j>n} |\langle h, e_i\rangle_{\mathbb{H}}|^2 |\langle g, e_j\rangle_{\mathbb{H}}|^2\right]^{1/2} \\ &\leq \left[\sum_{i,j>n} |\langle K(s,t)e_i, e_j\rangle_{\mathbb{H}}|^2\right]^{1/2} \|g\|_{\mathbb{H}} \|h\|_{\mathbb{H}}\end{aligned}$$

由此即可得出结论。

*(ii)*若 K 对称，则对任意 $f,g \in L^2(\mathrm{T}\ \ \mathbb{H})$:

$$\begin{aligned}\langle \mathscr{A}_K f, g\rangle_{L^2(\mathcal{T},\mathbb{H})} &= \int_{\mathcal{T}} \langle \int_{\mathcal{T}} K(t,u)f(u)du, g(t)\rangle_{\mathbb{H}} dt \\ &= \int_{\mathcal{T}} \int_{\mathcal{T}} \langle K(t,u)f(u), g(t)\rangle_{\mathbb{H}} dt du \\ &= \int_{\mathcal{T}} \int_{\mathcal{T}} \langle f(u), K(t,u)^* g(t)\rangle_{\mathbb{H}} dt du \\ &= \int_{\mathcal{T}} \int_{\mathcal{T}} \langle f(u), K(u,t) g(t) dt\rangle_{\mathbb{H}} du \\ &= \langle f, \mathscr{A}_K g\rangle_{L^2(\mathcal{T},\mathbb{H})}\end{aligned}$$

从而证明 $\mathscr{A}_K$ 是自伴的。

(iii) 我们遵循 Hsing 和 Eubank（2015 年定理 4.6.4）。给定 $n > 0$，选择 δ_n 使得当 $d((s_1, t_1), (s_2,t_2)) < \delta_n$ 时，$|K(s_2,t_2) - K(s_1,t_1)| < n^{-1}$。由于 T 是一个紧度量空间，存在 T 的有限划分 $\{E_{ni}\}$，使得每个 E_{ni} 的直径都小于 δ_n。令 v_i 为 E_{ni} 中的任意一点，并且对所有满足 $(s,t) \in E_{ni} \times E_{nj}$ 的情形，定义$K_n(s,t)$ 为 $K(v_i, v_j)$。K 的（均匀）连续性现在意味着

$$\max_{(s,t)\in\mathcal{T}\times\mathcal{T}} |K(s,t) - K_n(s,t)| < n^{-1}.$$

现在令 $\mathscr{A}_{K_n}$ 为核为 K_n 的积分算子。根据这一选择，我们发现，对于任意 $f \in L^2(\mathrm{T}\ \ \mathbb{H})$,

$$\left|\langle \mathscr{K} f, f\rangle_{L^2(\mathcal{T},\mathbb{H})} - \langle \mathscr{K}_n f, f\rangle_{L^2(\mathcal{T},\mathbb{H})}\right| \leq n^{-1} \|f\|_{L^2(\mathcal{T},\mathbb{H})}^2$$

and

$$\langle \mathscr{K}_n f, f\rangle_{L^2(\mathcal{T},\mathbb{H})} = \sum_{i,j=1}^{n} \left\langle K(v_i, v_j) \int_{E_{ni}} f(t)dt, \int_{E_{nj}} f(t)dt \right\rangle_{\mathbb{H}}.$$

由（4）可知，这证明了 $\mathscr{A}_K$ 的正性。

反之，假设

$$\sum_{i,j=1}^{n} \langle K(v_i, v_j)h_i, h_j\rangle_{\mathbb{H}} < 0.$$

对于某些 $\{v_j\}_{j=1,\dots,n} \subset \mathrm{T}$、$\{f_j\}_{j=1,\dots,n} \subset \mathbb{H}$。由于 K 一致连续，存在可测不交集 E_1、…、$E_m \subset \mathrm{T}$，使得对于所有 i 都有 $|E_i| > 0$、$v_i \in E_i$：

$$\max_{\tilde{v}_i \in E_i} \sum_{i,j=1}^{n} \langle K(\tilde{v}_i, \tilde{v}_j)h_i, h_j\rangle_{\mathbb{H}} < 0.$$

这意味着：

$$\sum_{i,j=1}^{n} \int_{E_i}\int_{E_j} \langle K(u,v)h_i, h_j\rangle_{\mathbb{H}}\, du\, dv < 0$$

根据中值定理。注意到最后一个表达式对于 $f(\cdot) = \sum_{i=1}^{n}(|E_i|)^{-1}h_i \cdot \mathbb{1}_{E_i}(\cdot)$ 仅仅是 $\langle \mathscr{A}_K f, f\rangle_{\mathbb{H}}$，我们得出结论：$\mathscr{A}_K$ 也不是半正定的。

□

特别地，如果 K 是一个连续、对称且半正定的核，则 $\mathscr{A}_K$ 可根据其特征值-特征函数对进行谱分解。也就是说，对于 $\mathscr{A}_K$，$\{\Phi_j\}_{j\geq 1}$ 的特征函数构成 $L^2(\mathrm{T}\ \mathbb{H})$ 的完备正交归一系；如果 $\{\lambda_j\}_{j\geq 1}$ 表示相应的特征值，则可以写为：

$$\mathscr{A}_K = \sum_{j\geq 1} \lambda_j \Phi_j \otimes_{L^2(\mathcal{T},\mathbb{H})} \Phi_j. \tag{6}$$

此外，以下引理证明了 $\mathscr{A}_K$ 的特征函数是一致连续的。

引理 2. 设 K 是连续核。对于每个 $f \in L^2(\mathcal{T}\ \mathbb{H})$，$(\mathscr{A}_K f)(\cdot)$ 是一致连续的。

证明。 由 T 的紧致性以及 K 的一致连续性可知，对于给定的任意 $\epsilon > 0$，存在 $\delta > 0$，使得 $\|K(u,s) - K(u,t)\|_{\mathscr{B}_1(\mathbb{H})} < \epsilon$ 对所有满足 $|s-t| < \delta$ 的 u、s、$t \in \mathrm{T}$ 均成立。于是：

$$\|(\mathscr{A}_K f)(s) - (\mathscr{A}_K f)(t)\|_{\mathbb{H}} = \| \int_{\mathcal{T}} \big(K(s,u) - K(t,u)\big) f(u)du\|_{\mathbb{H}} \leq \epsilon \|f\|_{L^2(\mathcal{T},\mathbb{H})}.$$

□

以下引理将在证明我们对 Mercer定理的推广时发挥重要作用。

引理 3. *(a)* 对于任意 $t \in \mathcal{T}$，

$$\sum_{n=1}^{N} \lambda_j \Phi_j(t) \otimes_{\mathbb{H}} \Phi_j(t) \leq K(t,t) \tag{7}$$

(b) 设 $\{e_i\}_{i\geq 1}$ 是 $\mathbb{H}$ 的完备正交归一系。对于任意 $s,\ t\in\mathcal{T}$,

$$\left\|\sum_{n=1}^{N}\lambda_j\Phi_j(s)\otimes_{\mathbb{H}}\Phi_j(t)\right\|_{\mathscr{B}_1(\mathbb{H})}\leq\|K(s,s)\|^{1/2}_{\mathscr{B}_1(H)}\cdot\|K(t,t)\|^{1/2}_{\mathscr{B}_1(\mathbb{H})}$$

(c) 算子 $\sum_{n=1}^{\infty}\lambda_j\Phi_j(s)\otimes_{\mathbb{H}}\Phi_j(t)$ 是良定义的，并且相对于$\|\cdot\|_{\mathscr{B}_1(\mathbb{H})}$一致连续。此外，该级数一致收敛。

证明。*(a)*令

$$K_n(s,t)=K(s,t)-\sum_{j=1}^{n}\lambda_j\Phi_j(s)\otimes_{\mathbb{H}}\Phi_j(t)$$

并取 $\mathscr{A}_{K_n}$ 为核为 K_n的积分算子。注意， K_n 是连续的，这是由 K 的连续性以及引理2所致。对任意 $f\in L^2(\mathcal{T}\ \mathbb{H})$:

$$\langle\mathscr{A}_{K_n}f,f\rangle_{L^2(\mathcal{T},\mathbb{H})}=\langle\mathscr{A}_{K}f,f\rangle_{L^2(\mathcal{T},\mathbb{H})}-\sum_{j=1}^{n}\lambda_j\langle f,\Phi_j\rangle^2_{L^2(\mathcal{T},\mathbb{H})}=\sum_{j=n+1}^{\infty}\lambda_j\langle f,\Phi_j\rangle^2_{L^2(\mathcal{T},\mathbb{H})}\geq 0$$

并且 $\mathscr{A}_{K_n}$ 必须是半正定的。由引理 1(iii)可知，这意味着 K_n 是半正定的，因此 $K_n(t,t)\geq 0$，从而证明了（7）。

(b) 设 $\{e_i\}_{i\geq 1}$ 是 $\mathbb{H}$的完备正交归一系。首先，注意:

$$\|K(t,t)\|_{\mathscr{B}_1(\mathbb{H})}\geq\|\sum_{j=1}^{N}\lambda_j\Phi_j(t)\otimes\Phi_j(t)\|_{\mathscr{B}_1(\mathbb{H})}=\sum_{i\geq 1}\sum_{j=1}^{N}\lambda_j\langle\Phi_j(t),e_i\rangle^2_{\mathbb{H}}.$$

然后，根据柯西–施瓦茨不等式，我们得到:

$$\begin{aligned}\left\|\sum_{n=1}^{N}\lambda_j\Phi_j(s)\otimes_{\mathbb{H}}\Phi_j(t)\right\|_{\mathscr{B}_1(\mathbb{H})}&=\sum_{i\geq 1}\sum_{j=1}^{N}\lambda_j\langle\Phi_j(t),e_i\rangle_{\mathbb{H}}\langle\Phi_j(s),e_i\rangle_{\mathbb{H}}\\&\leq\sum_{j=1}^{N}\left(\sum_{i\geq 1}\lambda_j\langle\Phi_j(t),e_i\rangle^2_{\mathbb{H}}\right)^{1/2}\cdot\left(\sum_{i\geq 1}\lambda_j\langle\Phi_j(s),e_i\rangle^2_{\mathbb{H}}\right)^{1/2}\\&\leq\|K(t,t)\|^{1/2}_{\mathscr{B}_1(\mathbb{H})}\|K(s,s)\|^{1/2}_{\mathscr{B}_1(\mathbb{H})}.\end{aligned}$$

(c) 固定 $\varepsilon>0$。由*(b)* 可得，存在 N_ε，使得:

$$\sup_{s,t\in\mathcal{T}}\left\|\sum_{n>N_\varepsilon}\lambda_j\Phi_j(s)\otimes_{\mathbb{H}}\Phi_j(t)\right\|_{\mathscr{B}_1(\mathbb{H})}\leq\varepsilon.$$

此外，对于任意 N， $\Phi_j(\cdot)$的一致连续性意味着存在 $\delta>0$，使得:

$$\left\|\sum_{n\leq N}\lambda_j\Phi_j(s)\otimes_{\mathbb{H}}\Phi_j(t)-\sum_{n\leq N}\lambda_j\Phi_j(s')\otimes_{\mathbb{H}}\Phi_j(t')\right\|_{\mathscr{B}_1(\mathbb{H})}\leq\varepsilon.$$

只要 $|s-s'|+|t-t'|<\delta$。由上面最后两个公式可知，存在 δ，使得:

$$\left\|\sum_{n\geq 1}\lambda_j\Phi_j(s)\otimes_{\mathbb{H}}\Phi_j(t)-\sum_{n\geq 1}\lambda_j\Phi_j(s')\otimes_{\mathbb{H}}\Phi_j(t')\right\|_{\mathscr{B}_1(\mathbb{H})}\leq\varepsilon.$$

只要 $|s-s'|+|t-t'|<\delta$。 □

定理 1（算子值核的Mercer定理）。设 K 是连续、对称且半正定的核，并记 $\mathscr{A}_K$ 为相应的积分算子。设 $\{\lambda_j,\ \Phi_j\}_{j\geq 1}$ 为 $\mathscr{A}_K$ 的特征值与特征函数对，其中 $\lambda_j \in \mathbb{R}$ 且 $\Phi_j \in L^2(\mathcal{T}\ \mathbb{H})$。则：

$$K(s,t) = \sum_{j=1}^{\infty} \lambda_j \Phi_j(s) \otimes_{\mathbb{H}} \Phi_j(t),$$

对所有 s，t，该级数均绝对且一致收敛。

证明。对于不同的连续核 $K_1(s,t)$ 和 $K_2(s,t)$，显然可以构造一个函数 f，使得 $\int_{\mathcal{T}} K_2(s,t)f(s)ds$ 与 $\int_{\mathcal{T}} K_1(s,t)f(s)ds$ 不同。因此，K 是定义 $\mathscr{A}_K$ 的唯一算子核。现在，具有连续核 $\sum_{j=1}^{\infty} \lambda_j \Phi_j(s) \otimes \Phi_j(t)$ 的积分算子与 $\mathscr{A}_K$ 具有相同的特征分解，因此是同一个算子。因此，对所有 $s,t \in \mathsf{T}$，均有 $K(s,t) = \sum_{j=1}^{\infty} \lambda_j \Phi_j(s) \otimes \Phi_j(t)$；根据引理 3，右侧绝对且一致收敛。

□

最后，我们现在可以看出，$\mathscr{A}_K$ 式（5）中的积分算子是迹类的。

命题 1. 设连续核 K 是对称且半正定的，$\mathscr{A}_K$ 是相应的积分算子。则：

$$\|\mathscr{A}_K\|_{\mathscr{B}_1(L^2(\mathcal{T},\mathbb{H}))} = \int_{\mathcal{T}} \|K(s,s)\|_{\mathscr{B}_1(\mathbb{H}))} ds.$$

证明. 对于任意完备正交归一系 $\{e_i\}_{i\geq 1}$，其中 $\mathbb{H}$:

$$\begin{aligned}
\operatorname{tr}\{\mathscr{A}_K\} &= \sum_{j\geq 1} \lambda_j = \sum_{j\geq 1} \lambda_j \|\Phi_j\|_{L^2(\mathcal{T},\mathbb{H})} = \sum_{j\geq 1} \lambda_j \int_{\mathcal{T}} \|\Phi_j(u)\|_{\mathbb{H}}^2 du = \int_{\mathcal{T}} \sum_{j\geq 1} \lambda_j \|\Phi_j(u)\|_{\mathbb{H}}^2 du \\
&= \int_{\mathcal{T}} \sum_{i,j\geq 1} \lambda_j \langle \Phi_j(u), e_i\rangle_{\mathbb{H}} du = \int_{\mathcal{T}} \operatorname{tr}\left\{\sum_{j\geq 1} \lambda_j \Phi_j(u) \otimes_{\mathbb{H}} \Phi_j(u)\right\} du \\
&= \int_{\mathcal{T}} \operatorname{tr}\{K(u,u)\} du
\end{aligned}$$

这里我们用到了帕塞瓦尔等式，并根据富比尼定理交换了求和与积分的顺序。

□

4 希尔伯特流的 Karhunen–Loève 定理

设 $\{\chi(t)\ :\ t \in \mathsf{T}\}$ 是概率空间 $(\Omega\ \mathscr{F}, \mathbb{P})$ 上取值于 $\mathbb{H}$ 的随机过程。其均值 $m\ :\mathsf{T}$ Ð→ $\mathbb{H}$ 可定义为：

$$m(t) = \mathbb{E}[\chi(t)], \qquad t \in \mathcal{T} \tag{8}$$

及其协方差核 $K\ :\mathsf{T}\ \times\mathsf{T}\ \to \mathscr{B}(\mathbb{H})$ 为：

$$K(s,t) = \mathbb{E}[\chi(s) \otimes_{\mathbb{H}} \chi(t)], \qquad t,s \in \mathcal{T} \tag{9}$$

前提是上述期望均作为博赫纳积分有定义（参见 Hsing 和 Eubank（2015年，定义 7.2.1）。

若过程 $\{\chi(t) : t \in \mathcal{T}\}$ 满足以下条件，则称其为二阶过程：对于每个8和9均有定义，其中 $t \in \mathcal{T}$。注意，协方差核是对称算子值核，即：

$$K(s,t) = K(t,s)^*$$

我们称 $\chi(\cdot)$ 是均方连续的，如果：

$$\lim_{n\to\infty} \mathbb{E}\left[\|\chi(t_n) - \chi(t)\|_{\mathbb{H}}^2\right] \to 0$$

对于任意 $t \in \mathrm{T}$ 以及任何收敛到 t 的序列 $\{t_n\}_{n\geq 1}$。

引理4。 设$\{\chi(t) : t \in \mathcal{T}\}$ 是一个二阶过程。则 χ 均方连续，当且仅当其均值和协方差函数分别关于 $\|\cdot\|_{\mathbb{H}}$ 和 $\|\cdot\|_{\mathscr{B}_1(\mathbb{H})}$ 连续。

证明。 不失一般性，假设该过程是中心化的，即 $m = 0$.

$$\begin{aligned}
E[\|\chi(s) - \chi(t)\|_{\mathbb{H}}^2] &= \mathbb{E}\|(\chi(s) - \chi(t)) \otimes (\chi(s) - \chi(t))\|_{\mathscr{B}_1(\mathbb{H})} \\
&= \|\mathbb{E}[(\chi(s) - \chi(t)) \otimes (\chi(s) - \chi(t))]\|_{\mathscr{B}_1(\mathbb{H})} \\
&= \|\mathbb{E}[(\chi(s) - \chi(t)) \otimes \chi(s) - (\chi(s) - \chi(t)) \otimes \chi(t))]\|_{\mathscr{B}_1(\mathbb{H})} \\
&= \|\mathbb{E}[\chi(s) \otimes \chi(s) - \chi(t) \otimes \chi(s) - \chi(s) \otimes \chi(t) + \chi(t)) \otimes \chi(t)]\|_{\mathscr{B}_1(\mathbb{H})} \\
&= \|K(t,t) + K(s,s) - K(t,s) - K(t,s)\|_{\mathscr{B}_1(\mathbb{H})} \to 0, \quad \text{as } s \to t.
\end{aligned}$$

注意，核 K 是半正定的，含义见（4）。事实上，对于 $n \geq 1$以及任意序列 $\{v_j\}_{j=1,\ldots,n} \subset \mathrm{T}$、$\{f_j\}_{j=1,\ldots,n} \subset \mathbb{H}$:

$$\sum_{i,j=1}^{n} \langle K(v_i, v_j) f_i, f_j\rangle_{\mathbb{H}} = \mathbb{E}\left[\sum_{i,j=1}^{n} \langle \chi_{v_i}, f_i\rangle_{\mathbb{H}} \cdot \langle \chi_{v_j}, f_j\rangle_{\mathbb{H}}\right] = \left(\sum_{j=1}^{n} \mathbb{E}\langle \chi_{v_j}, f_j\rangle_{\mathbb{H}}\right)^2 \geq 0$$

□

最后，我们可以陈述：

定理 2 (希尔伯特流/场的卡鲁宁–洛维展开)。 设$\{\chi(t) : t \in \mathcal{T}\}$是均方连续的二阶过程。定义：

$$\chi_n(t) := \sum_{j=1}^{n} \langle \chi, \Phi_j\rangle_{L^2(\mathcal{T},\mathbb{H})} \Phi_j(t), \qquad t \in \mathcal{T}.$$

则：

$$\lim_{n\to\infty} \sup_{t\in\mathcal{T}} \mathbb{E}\left[\|\chi_n(t) - \chi(t)\|_{\mathbb{H}}^2\right] = 0.$$

证明。 首先，请注意：

$$\begin{aligned}
\mathbb{E}\|\chi_n(t)\|_{\mathbb{H}}^2 &= \sum_{i,j=1}^{n} \langle \Phi_j(t), \Phi_i(t)\rangle_{\mathbb{H}}\, \mathbb{E}\left[\langle \chi, \Phi_j\rangle_{L^2(\mathcal{T},\mathbb{H})} \cdot \langle \chi, \Phi_i\rangle_{L^2(\mathcal{T},\mathbb{H})}\right] \\
&= \sum_{i,j=1}^{n} \langle \Phi_j(t), \Phi_i(t)\rangle_{\mathbb{H}}\, \langle \mathscr{A}_K \Phi_j, \Phi_i\rangle_{L^2(\mathcal{T},\mathbb{H})} \\
&= \sum_{i,j=1}^{n} \langle \Phi_j(t), \Phi_i(t)\rangle_{\mathbb{H}}\, \lambda_j \delta_{i,j} = \sum_{j=1}^{n} \lambda_j \|\Phi_j(t)\|_{\mathbb{H}}^2
\end{aligned}$$

此外：

$$\begin{aligned}
\mathbb{E}\left[\langle \chi_n(t), \chi(t)\rangle_{\mathbb{H}}\right] &= \sum_{j=1}^{n} \mathbb{E}\left[\langle \chi, \Phi_j\rangle_{L^2(\mathcal{T},\mathbb{H})} \cdot \langle \Phi_j(t), \chi(t)\rangle_{\mathbb{H}}\right] \\
&= \sum_{j=1}^{n} \int_{\mathcal{T}} \mathbb{E}\left[\langle \chi(u), \Phi_j(u)\rangle_{\mathbb{H}} \cdot \langle \Phi_j(t), \chi(t)\rangle_{\mathbb{H}}\right] du \\
&= \sum_{j=1}^{n} \int_{\mathcal{T}} \mathbb{E}\langle\, (\chi(u) \otimes_{\mathbb{H}} \chi(t))\Phi_j(u), \Phi_j(t)\,\rangle_{\mathbb{H}} \\
&= \sum_{j=1}^{n} \int_{\mathcal{T}} \langle\, K(u,t)\Phi_j(u), \Phi_j(t)\,\rangle_{\mathbb{H}} \\
&= \sum_{i=1}^{\infty}\sum_{j=1}^{n} \int_{\mathcal{T}} \langle\, \lambda_i \Phi_i(u) \otimes_{\mathbb{H}} \Phi_i(t)\Phi_j(u), \Phi_j(t)\,\rangle_{\mathbb{H}} du \\
&= \sum_{i=1}^{\infty}\sum_{j=1}^{n} \int_{\mathcal{T}} \langle\, \lambda_i \langle \Phi_i(u), \Phi_j(u)\rangle_{\mathbb{H}} \cdot \langle \Phi_i(t), \Phi_j(t)\rangle_{\mathbb{H}} du \\
&= \sum_{i=1}^{\infty}\sum_{j=1}^{n} \langle\, \lambda_i \langle \Phi_i, \Phi_j\rangle_{L^2(\mathcal{T},\mathbb{H})} \cdot \langle \Phi_i(t), \Phi_j(t)\rangle_{\mathbb{H}} \;=\; \sum_{j=1}^{n} \lambda_j \|\Phi_j(t)\|_{\mathbb{H}}^2
\end{aligned}$$

这里我们运用了富比尼定理来交换积分与期望，并将 Mercer 分解（定理 1）推广为 $K(u,t)$ 表示为 $\sum_{i=1}^{\infty} \lambda_i \langle \Phi_i(u) \otimes_{\mathbb{H}} \Phi_i(t)\rangle_{\mathbb{H}}$。

综合以上结果，我们得到：

$$\mathbb{E}\left[\|\chi_n(t) - \chi\|_{\mathbb{H}}^2\right] = \mathbb{E}\left[\|\chi_n(t)\|^2 + \|\chi\|_{\mathbb{H}}^2 - 2\langle \chi_n, \chi\rangle_{\mathbb{H}}\right] = \operatorname{tr}\{K(t,t)\} - \sum_{j=1}^{n} \lambda_j \|\Phi_j(t)\|_{\mathbb{H}}^2 \quad (10)$$

并注意到，根据 Parseval 恒等式：

$$\operatorname{tr}\left\{\sum_{j=1}^{n} \lambda_j \langle \Phi_j(t) \otimes_{\mathbb{H}} \Phi_j(t)\rangle_{\mathbb{H}}\right\} = \sum_{j=1}^{n}\sum_{i\ge 1} \lambda_j \langle \Phi_j(t), e_i\rangle_{\mathbb{H}}^2 = \sum_{j=1}^{n} \lambda_j \|\Phi_j(t)\|_{\mathbb{H}}^2$$

最后我们看到，(10) 由 Mercer 分解的推广（定理 1）一致收敛于零，从而证明了定理的结论。 □

5 计算方面的备注

现在，我们简要讨论如何计算一组实现流的展开分量。

设 $\chi^1, \dots, \chi^N$ 是 χ 的相互独立实现。设 $\mathbf{T}_n = \{t_k\}_{k=1}^n$ 是 $\mathbf{T}$ 的一种离散化。最后，设 $\mathbb{H}_m$ 是 $\mathbb{H}$ 的 m 维子空间，由 $\mathbb{H}$ 的某个完备正交归一系 $\{e_k\}_{k=1}^{\infty}$ 中的前 m 个向量张成。假设在每个节点 $\{t_k\}_{k=1}^n$，我们都可以观测到线性测量 $\chi^j(t_k)\ e_i$，其中每个流 $\{\chi_j\}_{j=1}^N$ 和每个基元素 $\{e_i\}_{i=1}^m$ 均如此。

在这种测量方案下，实现 χ^j，$1 \le j \le N$ 被化简为列

$$\mathbf{X}_j = [\langle \chi^j(t_1), e_1\rangle, \langle \chi^j(t_1), e_2\rangle, \dots, \langle \chi^j(t_1), e_m\rangle, \langle \chi^j(t_2), e_1\rangle, \dots, \langle \chi^j(t_n), e_m\rangle]^{\top}$$

所有观测数据都可以简洁地概括为 $mn \times N$ 矩阵 $\mathbf{X} = [\mathbf{X}_1 \mathbf{X}_2 \cdots \mathbf{X}_N]$。一种朴素的计算特征分解 $\{(\lambda_j, \Phi_j)\}_{j=1}^{\infty}$的方法

与 χ 相关的方法是先计算经验协方差的离散化版本，再对所得矩阵进行特征分解。在我们的情形中，这相当于计算 $\mathbf{K} = \frac{1}{N}\mathbf{X}\mathbf{X}^\intercal$ 的特征分解。此操作的计算复杂度为 $O(m^3n^3 + m^2n^2N)$（其中 $O(m^3n^3)$ 用于特征分解，其余部分用于计算协方差），这会严重限制分辨率（n 和 m）。

幸运的是，通过一个经典技巧（Chambers，1977；Jolliffe，1986），我们可以通过直接计算$\mathbf{X}$的奇异值分解来避免这项计算开销。在实践中，$mn > N$ 这使我们能够将$\mathbf{X}$ = 的奇异值分解写作 UDV，其中$\mathbf{U} = [u_{ij}]$和$\mathbf{V}$是维数分别为 $mn \times N$和 $N \times mn$的正交矩阵，而$\mathbf{D} = [d_{ij}\delta_{ij}]$是一个$N \times N$对角矩阵。对于 $k = 1, \ldots, n$，我们有如下近似

$$\hat{\Phi}_j(t_k) = \sqrt{n}\sum_{l=k+1}^{k+m} u_{lj}e_l \quad \text{and} \quad \hat{\lambda}_j = d_{jj}^2/N$$

与朴素方法的复杂度 $O(m^3n^3 + m^2n^2N)$ 相比，这种近似方法的复杂度为 $O(\max(mn, N)\min(mn, N)) = O(mnN^2)$。

6 讨论

Kim等人(2020) 也研究了希尔伯特型函数数据，最后我们将讨论这两种情形之间的差异。Kim等人(2020通过对如下形式的表示进行优化，得到一个最优分解

$$\chi(t) = \sum_{j=1}^{\infty} \alpha_j(t)\Phi_j$$

其中 $\{\alpha_j\} \subset L^2(\mathbb{R})$ 是需要优化的完备正交归一系，而 $\{\Phi_j\}_{j=1}^{\infty} \subset \mathbb{H}$ 是环境希尔伯特空间中的随机“系数”。在这方面，它与原始的卡鲁宁–洛维展开截然不同：在原始展开中，系数与完备正交归一系的作用恰好“相反”——按照传统定义，完备正交归一系属于与过程相同的环境空间，而系数则是实值的。Kim等人（2020）的展开是通过对实值自协方差核 $C(s,t) = \mathbb{E}\chi(s)\ \chi(t)_{\mathbb{H}}$（其中 $s,t \in \mathbf{T}$）的特征分解得到的。因此，无论 $\mathbb{H}$ 的性质如何，他们的分析都会产生属于 $L^2(\mathbf{T} \to \mathbb{R})$ 的特征函数。

相比之下，我们的分解是通过对以下形式的表示进行优化得到的：

$$\chi(t) = \sum_{j=1}^{\infty} \alpha_j\Phi_j(t)$$

其中，$\{\Phi_j\}_{j=1}^{\infty} \subset L^2(\mathbf{T} \to \mathbb{H})$ 是待优化的完备正交归一系，$\{\alpha_j\} \subset \mathbb{R}$是随机实值系数。这源于对算子值核$K(s,t) = \mathbb{E}\chi(s) \otimes \chi(t)$的谱分析，因此产生了存在于环境空间 $L^2(\mathbf{T} \to \mathbb{H})$ 中的主成分。这些可以说更为自然，但无论如何都与传统的卡鲁宁–洛夫定理相容。

参考文献

Adler, R. J. (2010). 随机场的几何. SIAM. Chambers, J. M. (1977). 数据分析的计算方法. 应用统计学中的 *Wiley* 出版物 . Grenander, U. (1950). 随机过程与统计推断 . 数学档案 *1*(3), 195–277. Hsing, T. 和 R. Eubank (2015年). 函数型数据分析的理论基础：线性算子导论, 第997卷. John Wiley & Sons. Jolliffe, I. T. (1986). 主成分分析. Springer统计学系列. Springer- Verlag, 纽约. Kac, M. 和 A. Siegert (1947). 平稳高斯过程的显式表示. 《数学统计年刊》*18*(3), 438–442. Karhunen, K. (1946). 随机过程的谱理论. 芬兰科学院年报，*AI 34*. Kim, D., Y. K. Lee, 和 B. U. Park (2020). 希尔伯特型函数数据的主成分分析. 统计应用与方法通讯 *27*(1), 149– 161. Le Maître, O. 和 O. M. Knio (2010). 不确定性量化的谱方法：在计算流体动力学中的应用. Springer科学与商业媒体. Loeve, M. (1948). 二阶随机函数. 随机过程与布朗运动, 366–420. Mercer, J. (1909). XVI. 正定型与负定型函数及其与积分方程理论的联系. 伦敦皇家学会哲学汇刊，*A*辑：数学或物理性质论文 *209*(441-458), 415– 446. Wang, J.-L., J.-M. Chiou, 和 H.-G. Müller (2016). 函数型数据分析. 《统计学及其应用年度评论》*3*, 257–295. Withers, C. (1974). Mercer定理与Fredholm预解式. 《澳大利亚数学学会通报》*11*(3), 373–380.