arXiv:2302.02053v1 [stat.ME] 4 Feb 2023

基于模型的平滑与积分维纳过程及重叠样条

Ziang Zhang，统计科学系，多伦多大学；Alex Stringer，统计与精算科学系，滑铁卢大学；Patrick Brown，统计科学系，多伦多大学，圣迈克尔医院全球卫生研究中心；Jamie Stafford，统计科学系，多伦多大学

2023年2月7日

摘要

在许多涉及对未知光滑函数进行推断的应用中，对其导数进行推断往往与对函数本身进行推断同样重要。为了对函数及其导数进行联合推断，我们考虑一类称为 p^{th} 阶积分维纳过程（IWP）的高斯过程。已有构造 IWP 有限元法（FEM）近似的方法，但这些方法仅关注 $p = 2$ 阶情形，无法对导数进行恰当推断，而且其计算可行性依赖于对 FEM 本身的进一步近似。本文提出一种替代性的 FEM 近似，称为重叠样条（O-spline）。该方法通过选择检验函数直接实现计算可行性，并且其构造与 IWP 相呼应：O-spline 是由这些相同检验函数经过多重积分得到的。O-spline 近似适用于任意阶 $p \in \mathbb{Z}^+$，计算效率高，并可对直至 $p-1$阶的所有导数进行一致推断。理论分析和模拟实证结果均表明，随着节点数量增加，O-spline 近似将收敛到真实的 IWP。我们还基于预测标准差（PSD）的概念，提供了一种统一且易于解释的平滑参数先验定义方式；该概念具有以下性质：

它不受阶数p 和节点位置的影响。最后，我们通过模拟研究和 COVID死亡率分析展示了 O-spline 近似的实际应用；在该分析中，我们同时对函数及其导数进行推断，后者对于解释大流行的进程具有重要意义。

关键词：高斯过程、导数推断、平滑、近似贝叶斯推断、先验选择、层次模型

1. 引言

在许多涉及未知回归函数的统计应用中， g，对 g 的导数进行推断往往与对 g 本身进行推断同样重要（Li 和 Liu，2020；Swain 等，2016；De Brabanter 和 Liu，2015）。为了对 g 及其导数进行联合推断，我们考虑一种基于模型的平滑方法，为函数及其导数指定一个高斯过程（$\mathcal{GP}$）模型（Rasmussen，2003）。

由于与传统平滑样条（Wahba，1978）密切相关，p^{th} 阶积分维纳过程，记为 IWP$_p$（σ），是 $\mathcal{GP}$模型的一种常用选择（Lindgren 和 Rue，2008年）。标准差参数 $\sigma \geq 0$ 控制 IWP$_p$（σ）的协方差函数，也可解释为平滑参数：取值越大，推断得到的 g 变异性越大；取值接近零则会迫使推断得到的 g 保持在 p^{th} 阶多项式的张成空间内。

给定一个 $p-1$ 次连续可微函数 g，且满足 $p \in \mathbb{Z}^+$，若将 g指定为 IWP$_p$(σ)，则对于任意 $q < p$，其 q 阶导数 $g^{(q)}$ 也立即服从 IWP$_{p-q}$(σ)。得益于这种同时导数性质，采用 IWP$_p$(σ) 进行基于模型的平滑，可以对 g 及其导数进行可解释的联合推断。然而，在许多实际场景中，直接对 g 及其导数拟合 IWP 模型的计算开销很大，因为存储和分解大型稠密协方差矩阵的成本高昂。显著降低计算难度的一种方法，是通过有限元方法（FEM）将 IWP 模型近似为有限维近似；相关示例见 Lindgren 和 Rue（2008年）以及 Yue 等（2014）。不过，这些方法仅适用于阶数为 $p = 2$ 的情形，且无法对导数进行推断。此外，为了实现计算上的可行性，这些方法还需要对有限元方法本身再作一次近似。

本文提出一种替代性的有限元近似，称为重叠样条（O-spline）。该方法通过选择检验函数，直接实现计算上的可行性；同时，它也仿照 IWP 的构造方式，因为 O-spline 是由对这些相同检验函数进行多重积分得到的。本文的贡献如下：

(a) 我们通过有限元法提出了一种计算高效的有限维近似，适用于任意阶数 $p \geq 1$ 的积分维纳过程$_p(\sigma)$模型，称为重叠样条（O-样条）。我们将 g的 O 样条近似记为 $\tilde{g}_k$，其中k 是构造该近似所用的节点数。

(b) 我们通过理论分析和模拟表明，随着节点数 k 增加，$\tilde{g}_k$ 及其导数的联合分布将收敛至真实积分维纳过程$_p(\sigma)$下的分布。

(c) 我们提出了一种统一的方式，根据 h单位预测标准差（PSD）的概念，为参数 σ 定义先验；这一概念在不同阶数 p 下具有一致的解释。

本文其余部分安排如下。第 2 节介绍本文的建模背景，并提供积分维纳过程$_p(\sigma)$模型所需的一些背景知识。第 3 节介绍所提出的 O-spline 近似，讨论其统计和计算性质，以说明使用该近似的合理性，并介绍如何采用 Stringer 等人（2022）中的计算方法高效拟合该近似。我们还介绍了一种统一且易于解释的方式来为参数 σ定义先验。第 4 节和第 5 节通过模拟研究，以及对新冠肺炎死亡率及其随时间变化率的分析，展示所提出方法的实际用途。最后，我们在第 6 节进行讨论并总结全文。

复现本文所有结果和示例的代码可在相应的在线代码库 github.com/AgueroZZ/Smooth˙IWP˙code 中找到。

2. 使用积分维纳过程进行平滑

考虑以下层次模型:

$$
\begin{aligned}
&Y_i|\boldsymbol{\eta} \overset{ind}{\sim} \pi(Y_i|\boldsymbol{\eta},\kappa),\ i\in[n],\\
&\eta_i = \boldsymbol{v}_i^T\boldsymbol{\beta} + \sum_{l=1}^{L} g_l(x_{li}),\ x_{li}\in\Omega_l\subset\mathbb{R}, \boldsymbol{v}_i\in\mathbb{R}^r \qquad (1)\\
&g_l:\Omega\to\mathbb{R},\ g_l \overset{ind}{\sim} \mathrm{IWP}_{p_l}(\sigma_l), \forall l\in[L].
\end{aligned}
$$

这里 $\pi(Y_i|\boldsymbol{\eta},\ \kappa)$ 是一个二阶连续可导的密度，线性预测子 $\boldsymbol{\eta}=\{\eta_i, i\in[n]\}$、协变量 $\{\boldsymbol{v}_i, x_{li}:\ l\in[L],\ i\in[n]\}$ 以及超参数 κ。每个未知函数$g_l:\ \Omega_l\to\mathbb{R}$ 均被赋予一个相互独立的 IWP$_{p_l}(\sigma_l)$ 模型，其标准差参数为 σ_l ，阶数为 p_l。本文旨在建立一种对 g_l 及其导数进行同时推断的方法。

上述方程（1）中的模型类别属于 Stringer 等人（2022）的扩展潜在高斯模型，其中包括广泛使用的 Hastie 和 Tibshirani（1990）广义加性模型，以及其扩展模型，例如 Stringer 等人（2020）和 Zhang 等人（2022）的模型。为便于接下来几节的阐述，我们针对单个未知函数 g 考虑相应的推导。

2.1. 积分维纳过程

为了能够对 g 及其最高至 $p-1$ 阶的导数进行同时推断，我们通过以下构造，采用 p^{th} 阶积分维纳过程（IWP$_p(\sigma)$）作为 g 的 $\mathcal{GP}$ 模型:

$$
g(x) \overset{d}{=} \sum_{l=0}^{p-1}\gamma_l x^l + \sigma W_p(x), \qquad (2)
$$

我们将其记为 g ~积分维纳过程$_p(\sigma)$。令 $\boldsymbol{\gamma}=(\gamma_0,\ .....,\ \gamma_{p-1})^T \sim N(\mathbf{0},\ \Sigma_{\boldsymbol{\gamma}})$，其中 $\Sigma_{\boldsymbol{\gamma}}=\mathrm{diag}(\tau_0^2,\ .....,\ \tau_{p-1}^2)$。默认假设 $\tau_l^2=1000$ 对所有 $0\le l<p$ 均成立。过程 W_p 与 $\boldsymbol{\gamma}$ 相互独立，并通过以下方式定义（Shepp，1966年）:

$$
\frac{\partial^q}{\partial t^q}W_p(0)=0 \quad \forall\ q<p \qquad \text{and} \qquad \frac{\partial^p}{\partial t^p}W_p(x)\overset{d}{=}\xi(x) \qquad (3)
$$

其中 $\xi(x)$ 是广义高斯白噪声过程（Harvey，1990年；Lindgren 等人，2011年）。根据上述构造，给定系数为 γ_l，$q \le l < p$ 的 $g^{(q)} \sim$ 积分维纳过程$^{p-q}(\sigma)$，即可直接对 g 及其导数进行同时推断。我们称之为同时导数性质。全文中，阶数 p 均视为固定值并预先指定先验地。

在这一设定下进行推断的一项障碍涉及W_p(x)的协方差矩阵，其中 $\boldsymbol{x} = \{x_i, i \in [n]\}$。尽管 W_p 及其导数的协方差函数在（Robinson，2010）中有明确形式，但由此得到的协方差矩阵在每种情况下都是稠密矩阵，求逆的计算开销很大，需要 $O(n^3)$ 次浮点运算。一种可能高效的解决方案是利用积分维纳过程的马尔可夫性质（Rue 和 Held，2005；Robinson，2010），并采用稀疏矩阵存储、分解和求逆的高效算法（Rue，2001）。然而，协方差矩阵的规模仍会随着 n 增长，位置较多时计算依然可能变得困难。因此，我们提出使用有限元方法（FEM）对 W_p进行有限维近似，从而得到 g；该近似保留了同时导数性质，同时具备理想的计算和理论性质。

3. 有限元方法与重叠样条

在本节中，我们使用有限元方法（FEM）对 W_p （进而对 g）进行基函数近似。不失一般性地，假设感兴趣区域$\widetilde{\Omega}$ 具有区间 $[0, a]$ 的形式，其中 $a \in \mathbb{R}^+$。在此，我们将 $W_p(x)$ 和 $\tilde{g}_k(x)$ 定义为

$$\widetilde{W}_p(x) = \sum_{j=1}^{k} w_j \varphi_j(x); \quad \tilde{g}_k(x) = \sum_{l=0}^{p-1} \gamma_l x^l + \sigma \widetilde{W}_p(x) \tag{4}$$

其中 $\mathbb{B}_k := \{\varphi_i : i \in [k]\}$ 是待选的一组基函数，而 $\boldsymbol{w} = (w_1, \ldots\ldots, w_k)^T$是有待推断的未知（随机）权重。通过进一步选择一组检验函数 $\mathbb{T}_k := \{\phi_i, i \in [k]\}$并强制满足以下分布近似，便可确定这些权重的分布：

$$\boldsymbol{w} \sim N(\mathbf{0}, \Sigma_{\boldsymbol{w}}) \tag{5}$$

其中 $\Sigma_{\boldsymbol{w}}^{-1} = \mathbf{B}^T\mathbf{T}^{-1}\mathbf{B}$ 和 $\mathbf{B}$，其中 $\mathbf{B}^{ij} = \int_\Omega \phi_i(x)\frac{\partial^p \varphi_j}{\partial t^p}(x)dx$，而 $\mathbf{T}^{ij} = \int_\Omega \phi_i(x)\phi_j(x)dx$~。因此，有限元近似 W_p 的分布，也就是 $\tilde{g}_k$，完全由检验函数 $\mathbb{T}_k$ 和基函数 $\mathbb{B}_k$的选择决定。这些选择也决定了该近似的其他性质。补充材料 B 对上述细节作了更形式化的阐述，其中涉及随机微分方程；相关内容也可见于文献（Shepp，1966年）。

Lindgren 和 Rue（2008年）给出了有限元法的一个示例，正式称为 IWP$_2$ 模型的伽辽金解。他们选择线性B样条作为 $\mathbb{B}_k = \mathbb{T}_k$。因此，$\mathbf{T}^{-1}$ 是稠密矩阵。为了高效计算精度矩阵，作者建议进一步近似：用$\mathbf{A}$替代$\mathbf{T}$，其中 A 是一个对角矩阵，每个对角元素等于$\mathbf{T}$中对应行的元素之和。所得近似称为连续二阶随机游走（RW2）。之后，Lindgren 等人（2011年）将同一策略推广用于连续空间变异建模，Yue 等（2014）则将其用于 IWP$_2$的自适应平滑。尽管效率较高，RW2 方法仍有以下缺点：

a. 该方法仅适用于阶数为 $p = 2$ 的情形；b. 为了使计算可行，该方法需要进行两次近似；c. IWP$_2$ 模型的样本路径是一次连续可微的，而近似模型的样本路径则不是。因此，同时导数性质不成立。

在下一节中，我们将提出一种新的有限元近似，直接兼顾同时推断与计算可行性，以解决上述问题。

3.1. 重叠样条（O-样条）

在本节中，我们利用有限元方法（FEM），为一般的 $p \in \mathbb{Z}^+$ 推导 IWP$_p(\sigma)$ 模型的有限维近似。我们直接追求同时求导性质和计算效率，构造一种形式上称为最小二乘解的解，随后证明该近似具有所需的性质。鉴于其构造方式，我们将所提出的方法称为重叠样条（O-spline）近似

其原因从构造过程显而易见。该方法的推导直观明了：为确保$\mathbf{T}$ 为对角矩阵，我们将检验函数 $\mathbb{T}_k$ 定义为分段常数

$$\phi_i(x) = \begin{cases} 0 & \text{if } x < s_{i-1}, \\ 1 & \text{if } x \in (s_{i-1}, s_i], \\ 0 & \text{if } x \geq s_i, \end{cases} \tag{6}$$

其中， $\boldsymbol{s} = \{s_i, i \in [k]\} \subset \Omega$ 是按递增顺序排列的 k 个不同位置的集合，且满足 $s_0 = 0$。最小二乘解使用满足以下条件的基函数（Lindgren 等人，2011年）：

$$\frac{\partial^p}{\partial x^p}\varphi_i(x) = \phi_i(x), \quad \frac{\partial^q}{\partial x^q}\varphi_i(0) = 0 \text{ for } \forall \ \ q < p. \tag{7}$$

这可通过反复积分检验函数 $\mathbb{T}_k$ 来定义（O-样条）基 $\mathbb{B}_k$，从而得到保证：

$$\varphi_i(x) = \begin{cases} 0 & \text{if } x < s_{i-1}, \\ \frac{(x-s_{i-1})^p}{p!} & \text{if } x \in (s_{i-1}, s_i], \\ \sum_{k=1}^{p} \frac{d_i^k (x-s_i)^{p-k}}{k!(p-k)!} & \text{if } x \geq s_i, \end{cases} \tag{8}$$

其中 $d_i = s_i - s_{i-1}$。上述基和检验函数的选择意味着，公式 (4) 中的两个矩阵$\mathbf{T}$ 和$\mathbf{B}$ 相同。因此，所提出的近似中基权重的精度矩阵是对角矩阵，即 $\Sigma_{\boldsymbol{w}}^{-1} = \mathbf{B}^T\mathbf{T}^{-1}\mathbf{B} = \mathbf{T}$，其第 i 个对角元素为 $\int_\Omega \phi_i^2(x)dx = d_i$。

从上述构造可以立即看出，O-spline 近似继承了同时导数性质

$$\frac{\partial^q}{\partial t^q}\widetilde{W}_p(x) = \frac{\partial^q}{\partial t^q}\sum_j^k w_j\varphi_j(x) = \sum_j^k w_j\varphi_j^{(q)}(x) = \widetilde{W}_{p-q}(x). \tag{9}$$

与 RW2 方法相比，它还有以下三个优点：

a. O样条近似可处理任意 $p \geq 1$ 下的 IWP$_p(\sigma)$，因此可以根据对 g 可微性的先验知识来选择 p； b. O样条近似的样本路径具有与原始 IWP$_p(\sigma)$ 模型相同的可微性，同时还能实现基于模型的

导数推断；

c. 由于 Φ 和 $\Sigma_{\boldsymbol{w}}$ 的上三角结构与对角结构已便于进行高效的矩阵计算，因此无需进一步进行矩阵近似，即可直接将 O-spline 近似推导为针对积分维纳过程模型的有限元近似。

O-spline 近似还具有以下部分所述的理想理论性质。

3.2. O-spline 近似的理论性质

除了所提出的 O-spline 近似具有直接优势外，它还具备理想的理论性质。这些结果的证明见附录。

首先，所提出的 O-spline 近似避免了节点选择和放置中的歧义；这一直是基于样条的近似方法需要解决的核心问题（Eilers and Marx，1996），相关讨论见 Eilers and Marx（2010）以及 Ruppert and Carroll（2000）。如引理1所述，随着更多节点在感兴趣区域内等距布置，所提出近似的协方差函数将以线性速率收敛至真实积分维纳过程$_p(\sigma)$的协方差，且适用于任意 $p \geq 1$。因此，就近似精度而言，在计算可行的范围内，使用尽可能多的等距节点效果更好。

引理1（O样条近似的收敛性）。设$\Omega = [a, b]$，其中 a、$b \in \mathbb{R}^+$，且令 $g \sim$积分维纳过程$_p(\sigma)$ 满足 $p \in \mathbb{Z}^+$。假设对于每个 $k \in \mathbb{N}$，节点 $\{s_1, ..., s_k\}$ 在Ω 上等距分布，且 $\tilde{g}_k$ 表示按式(4)定义的相应 p 阶（***O-spline***）近似，则：

$$||\mathcal{C} - \mathcal{C}_k||_\infty = O(1/k),$$

其中 $\mathcal{C}(s,t) =$ 协方差$[g(s)\ g(t)]$, $\mathcal{C}_k(s,t) =$ 协方差$[\tilde{g}_k(s)\ \tilde{g}_k(t)]$，且 $||\mathcal{C}-\mathcal{C}_k||_\infty \equiv$ 上确界$_{s,t\in\Omega}|\mathcal{C}(s,t)-\mathcal{C}_k(s,t)|$。

其次，O样条近似的导数同时性意味着，函数所用的节点序列和权重系数与导数所用的相同，唯一额外的步骤是在较低阶次下重新计算设计矩阵。

这一性质使我们提出的 O样条近似适用于函数及其导数的联合推断，且可由这一性质和引理1证明以下定理 1：

定理 1 (主定理)。 在与引理*1*相同的设定和记号下，对于任意非负整数 $q_1 \leq p-1$ 和 $q_2 \leq p-1$:

$$||\mathcal{C}^{(q_1,q_2)} - \mathcal{C}_k^{(q_1,q_2)}||_\infty = O(1/k),$$

其中 $\mathcal{C}^{(q_1,q_2)}(s,\ t) =$协方差$[g^{(q_1)}(s)\ g^{(q_2)}(t)]$ 且 $\mathcal{C}^{(q_1,q_2)}{}_k(s,\ t) =$协方差$[\tilde{g}_k^{(q_1)}(s)\ \tilde{g}_k^{(q_2)}(t)]$。

该定理给出了所提出的 O-spline 近似的一般收敛结果，这意味着近似与其导数之间的每个交叉协方差，进而其互相关，都将收敛到真实值。具体而言，这意味着在 IWP$_p$(σ) 模型下， $\tilde{g}_k$ 及其导数的任意有限维分布（f.d.d）都将收敛到真实的有限维分布：

推论 1 (有限维分布的收敛)。 考虑与引理*1*相同的设定和记号，令 $p, m \in \mathbb{Z}^+$，则：

$$[\tilde{g}_k^{(q_1)}(x_1), \tilde{g}_k^{(q_2)}(x_2), ..., \tilde{g}_k^{(q_m)}(x_m)]^T \xrightarrow{d} [g^{(q_1)}(x_1), g^{(q_2)}(x_2), ..., g^{(q_m)}(x_m)]^T,$$

如 $k \to \infty$，其中 $\{x_i \in \Omega:\ i \in [m]\}$ 和 $\{q_i \in \mathbb{Z}^+ :\ i \in [m]、\ q_i < p\}$ 均为任意值。

证明。 利用定理 1 和 Cramér-Wold 装置，根据 Levy 连续性定理可直接得到推论 1。

□

请注意，由于 RW2 近似方法的样本路径不是 $p-1$ 次连续可微的，因此无法得到这一结果。

3.3. σ的先验与解释

在我们的语境中，一个关键考量是为参数 σ 选择适当的先验。鉴于对 σ 的解释取决于阶数 p，这一选择变得复杂。例如，具有相同 σ 值但阶数不同的 IWP 模型，其样本路径可能具有极为不同的变异性，这使得先验的选择对实际应用者而言颇为棘手。

为了解决这一问题，我们采用了类似于 Sorbye 和 Rue (2014) 的策略，并为一个仍涉及 σ 但在 p 取不同值时具有一致解释的量指定先验。为了便于解释，我们考虑函数 g 在未来位置 $x+h$ 处的标准差，条件是已知位置 x 处的 g 值及其导数，记作：

$$\sigma(h) = \mathrm{Var}\left[g(x+h)\middle| g(x), g^{(1)}(x), ..., g^{(p-1)}(x)\right]^{1/2} = \frac{\sqrt{h^{(2p-1)}}\sigma}{\sqrt{(2p-1)}(p-1)!}. \tag{10}$$

我们将该量 $\sigma(h)$ 称为以 h 个单位计的预测标准差（PSD），它量化了利用函数 g 截至当前位置的信息，对前方 h 个单位处的函数值进行预测时的不确定性。单位 h 的选择可依据其实际相关性确定。这不同于 Sorbye 和 Rue（2014）的方法，该方法使用边际标准差；但我们的做法有一个优势： $\sigma(h)$ 不随位置 x 变化，而边际标准差则会变化。同时，Sorbye 和 Rue（2014）将其先验 elicitation 方法应用于积分维纳过程 $_p(\sigma)$ 的有限维近似，而我们的方法完全基于原始积分维纳过程 $_p(\sigma)$，因此不受近似选择的影响，例如节点的数量和位置。

最后，受 Simpson 等人（2017）的启发，我们为 $\sigma(h)$ 考虑如下形式的指数先验：$P(\sigma(h) > u) = \alpha$，其中 $\{u, \alpha\}$ 将由用户选择。然后，通过将针对 $\sigma(h)$ 的指数先验缩放 $(p-1)!\sqrt{(2p-1)/h^{(2p-1)}}$，即可得到 σ 的先验。

3.4. 近似贝叶斯推断

在本文余下部分，我们将上述发展纳入 Stringer 等人（2022）的 ELGM 框架中。借助其中提出的自适应高斯-埃尔米特求积（`aghq`）算法，我们概述如何在公式 (1) 的 ELGM 中实现公式 (4) 所示的 O样条近似。

依照 Stringer 等人（2022）的详细说明，我们令 $\boldsymbol{\omega} = (\boldsymbol{w}, \boldsymbol{\gamma}, \boldsymbol{\beta})$ $\boldsymbol{\theta} = (\sigma, \kappa)$，并有 $\boldsymbol{\omega} \sim \mathrm{N}\ [\mathbf{0}, \Sigma(\boldsymbol{\theta})]$，其中 $\Sigma(\boldsymbol{\theta})$ 为 $\mathrm{diag}(\sigma^2\Sigma_{\boldsymbol{w}}, \Sigma_{\boldsymbol{\gamma}}, \Sigma_{\boldsymbol{\beta}})$，且 $\Sigma_{\boldsymbol{\beta}}$ 和 $\Sigma_{\boldsymbol{\gamma}}$ 为对角矩阵。目标是基于后验分布 $\pi(\boldsymbol{\omega}|\boldsymbol{y})$ 和 $\pi(\boldsymbol{\theta}|\boldsymbol{y})$，对 $\boldsymbol{\omega}, \boldsymbol{\theta}$ 进行完全贝叶斯推断。由于这两个后验分布都难以处理，我们使用 Stringer 等人的方法计算其对应的近似 $\widetilde{\pi}_{\text{LA}}(\boldsymbol{\theta}|\boldsymbol{y})$ 和 $\widetilde{\pi}(\boldsymbol{\omega}|\boldsymbol{y})$。

(2022）。首先，我们计算以下高斯近似和拉普拉斯近似 ：

$$\widetilde{\pi}_{\text{G}}(\boldsymbol{\omega}|\boldsymbol{\theta},\boldsymbol{y}) = N(\widehat{\boldsymbol{\omega}}_{\boldsymbol{\theta}}, \boldsymbol{H}(\boldsymbol{\theta})^{-1}),\ \widetilde{\pi}_{\text{LA}}(\boldsymbol{\theta},\boldsymbol{y}) = \frac{\pi(\widehat{\boldsymbol{\omega}}_{\boldsymbol{\theta}},\boldsymbol{\theta},\boldsymbol{y})}{\widetilde{\pi}_{\text{G}}(\widehat{\boldsymbol{\omega}}_{\boldsymbol{\theta}}|\boldsymbol{\theta},\boldsymbol{y})}, \tag{11}$$

其中， $\widehat{\boldsymbol{\omega}}_{\boldsymbol{\theta}} = \arg\max \log\ \pi(\boldsymbol{\omega},\boldsymbol{\theta},\boldsymbol{y})$为众数， $\boldsymbol{H}(\boldsymbol{\theta})$是在众数处计算的负海森矩阵。具体而言，我们将 $\boldsymbol{\eta}$写作 $\boldsymbol{\eta} = \boldsymbol{\Phi}\boldsymbol{w} + \boldsymbol{P}\boldsymbol{\gamma} + \boldsymbol{V}\boldsymbol{\beta}$，其中 $\boldsymbol{V}, \boldsymbol{P}$和 $\boldsymbol{\Phi}$ 为相应的设计矩阵，并计算 $\boldsymbol{H}(\boldsymbol{\theta})$如下：

$$\boldsymbol{H}(\boldsymbol{\theta}) = \begin{pmatrix} \sigma^2\Sigma_{\boldsymbol{w}} & 0 & 0 \\ 0 & \Sigma_{\boldsymbol{\gamma}} & 0 \\ 0 & 0 & \Sigma_{\boldsymbol{\beta}} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\Phi}^T \\ \boldsymbol{P}^T \\ \boldsymbol{V}^T \end{pmatrix} \ \partial^2_{\boldsymbol{\eta}}\ \log\pi(\boldsymbol{y}|\hat{\boldsymbol{\eta}}(\boldsymbol{\theta}),\kappa)\ \begin{pmatrix} \boldsymbol{\Phi} & \boldsymbol{P} & \boldsymbol{V} \end{pmatrix} \tag{12}$$

然后，我们使用 R 语言中的 aghq 程序包（Stringer, 2020；Bilodeau 等人，2021）计算归一化拉普拉斯近似 $\widetilde{\pi}_{\text{LA}}(\boldsymbol{\theta}|\boldsymbol{y})$，进而得到 $\widetilde{\pi}(\boldsymbol{\omega}|\boldsymbol{y})$。有了来自 $\widetilde{\pi}(\boldsymbol{\omega}|\boldsymbol{y})$的样本$\{\boldsymbol{\omega}_m\}_{m=1}^M$ 后，对于任意 $t \in \Omega$，都可以获得样本 $\{\widetilde{g}_m(t)\}_{m=1}^M$ 或其导数。

4. 模拟研究

4.1. 近似精度评估

本节将评估 O-spline 对积分维纳过程模型的近似质量。为比较不同 p取值下的结果，我们将分别基于真实 IWP 和其 O-spline 近似计算相关函数，而非协方差函数。不失一般性，我们假设公式 (2) 中每个 $0 \leq i < p$和 $\sigma = 1$ 均满足 $\gamma_i = 0$ 。

对于阶数为 $p = 1$，2，3 和 4 的情形，我们计算真实积分维纳过程的自相关函数 $\rho(5, x) = \text{Cor}[W_p(5)\ W_p(x)]$。随后，我们使用 O-spline 得到相应的近似，节点数分别为 $k = 5$，10，30 或 100，且节点均匀分布在区间 $[0, 15]$ 上。如图 1(a-d) 所示，当 $k \geq 30$ 时，O-spline 的近似在所有 p 上始终准确。随着 p 增大，在较大的 x 值处，较高 k 带来的近似质量提升会更加明显。

为了从联合分布的角度评估近似质量，我们随后比较积分维纳过程与其 q 阶导数之间的互相关函数，记作

$\rho^{(0,q)}(5, x) = \mathrm{Cor}[W_p(5)\ W_p^{(q)}(x)]$。我们考虑 $p = 2$，3，4 和 $q = 1$ 或 2。它们的 O-spline 近似采用与上述相同的设置得到。结果汇总于图 1(e-h)。与先前关于自相关的结果类似，当 $k \geq 30$ 时，对于所有 p，O-spline 近似始终能准确地近似互相关。

以上两个结果共同表明，即使 k 较小，我们提出的 O-spline 近似在 p 的不同阶数下仍具有实际精度，无论是就积分维纳过程内部的分布而言，还是就其与导数的联合分布而言。

4.2. 与精确积分维纳过程的计算比较

本节将通过比较所提出的 O 样条近似与精确积分维纳过程模型的实现，说明该近似在计算上的可行性。精确模型的实现采用 Rue 和 Held（2005）及 Robinson（2010）所述的增广空间方法。为确保结果具有可比性，两种方法均采用第3.4节所述的近似贝叶斯推断方法进行拟合。

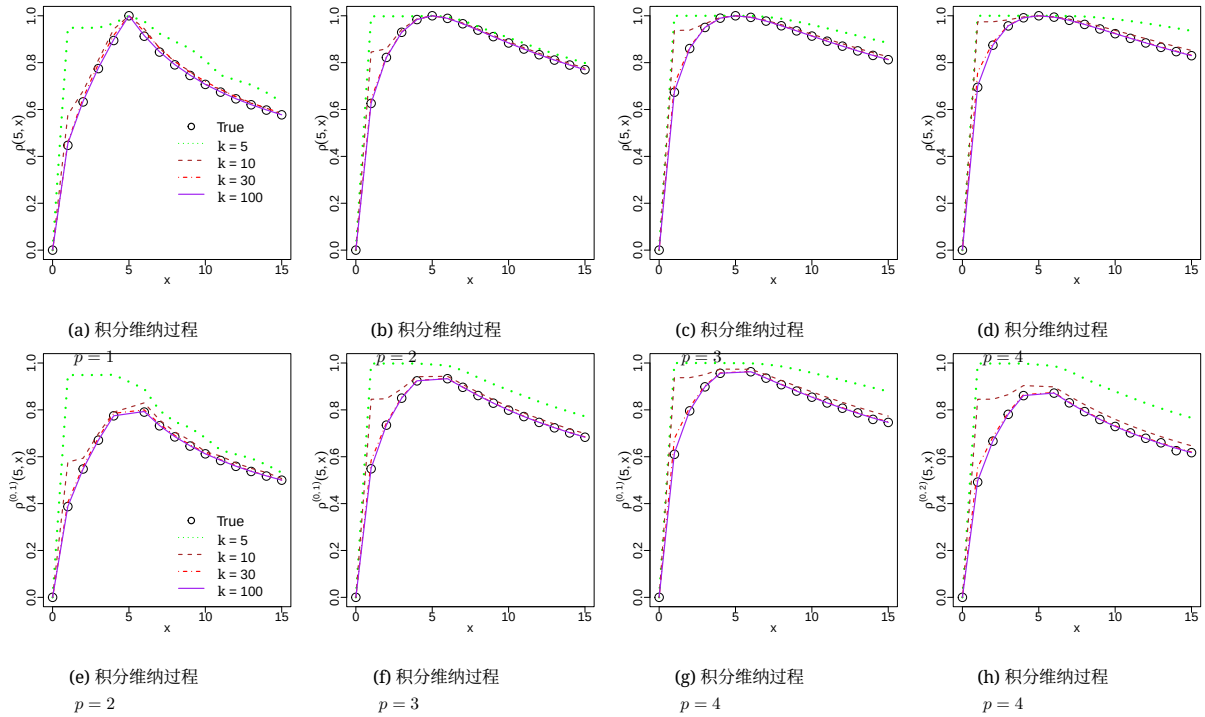
作为概念验证，我们假设 $\boldsymbol{x} = \{x_i\}_{i=1}^n$ 在 $\Omega = [0, 20]$ 和 $p = 3$ 上均匀分布。n 观测值由以下简单的单变量回归模型模拟得到：

$$\begin{aligned} y_i &= g(x_i) + \epsilon_i; \quad x_i \in \Omega; \\ \epsilon_i &\sim N(0, 1); \quad g(x) = \sqrt{3}\sin(x/2). \end{aligned} \tag{13}$$

对于 g 的推断，我们同时考虑精确积分维纳过程$_3(\sigma)$及其 O 样条近似。五单位功率谱密度 $\sigma(5)$采用指数先验，$\mathrm{P}(\sigma(5) > 3) = 0.01$。O 样条近似中使用的节点数 k 分别为 10，30，50 和 100，且均匀分布在 Ω 上。第 3.4 节中的所有计算均以单线程运行，自适应求积次数为 10，后验样本数为 3，000。

为比较各方法的数值稳定性，我们计算第3.4节中定义的 $\boldsymbol{H}(\boldsymbol{\theta})$ 的条件数，其定义为

$$\kappa[\boldsymbol{H}(\boldsymbol{\theta})] = \frac{\lambda_{\max}[\boldsymbol{H}(\boldsymbol{\theta})]}{\lambda_{\min}[\boldsymbol{H}(\boldsymbol{\theta})]}, \tag{14}$$



(a) 积分维纳过程 (b) 积分维纳过程 (c) 积分维纳过程 (d) 积分维纳过程

(e) 积分维纳过程 $p = 2$ (f) 积分维纳过程 $p = 3$ (g) 积分维纳过程 $p = 4$ (h) 积分维纳过程 $p = 4$

图 1：第4.1节中，积分维纳过程及其 O样条近似在 $p = 1, 2, 3$ 和 4 时的自相关 ρ (a-d) 与互相关 $\rho^{(0,q)}$ (e-h) 函数。图 (e-g) 计算了积分维纳过程与其一阶导数之间的相关性，最后一幅图 (h) 计算了积分维纳过程与其二阶导数之间的相关性。当 $k \geq 30$ 时，自相关和互相关均得到准确近似。

其中， $\lambda_{\max}$ 和 $\lambda_{\min}$ 分别表示最大奇异值和最小奇异值。随后，我们计算每种方法的最大条件数 κ_{max} ，其定义为

$$\kappa_{\max} = \max\{\kappa[\boldsymbol{H}(\boldsymbol{\theta}_j)] : j \in [10]\}, \tag{15}$$

其中， $\{\boldsymbol{\theta}_j : \ j \in [10]\}$ 表示由 10 个求积点组成的集合。表 1 展示了不同 n下每种方法的最大条件数。请注意，当 $n \geq 800$时，使用精确积分维纳过程方法的实现会因严重的数值奇异性问题而失效。然而，所提出的 O样条近似不存在数值问题。

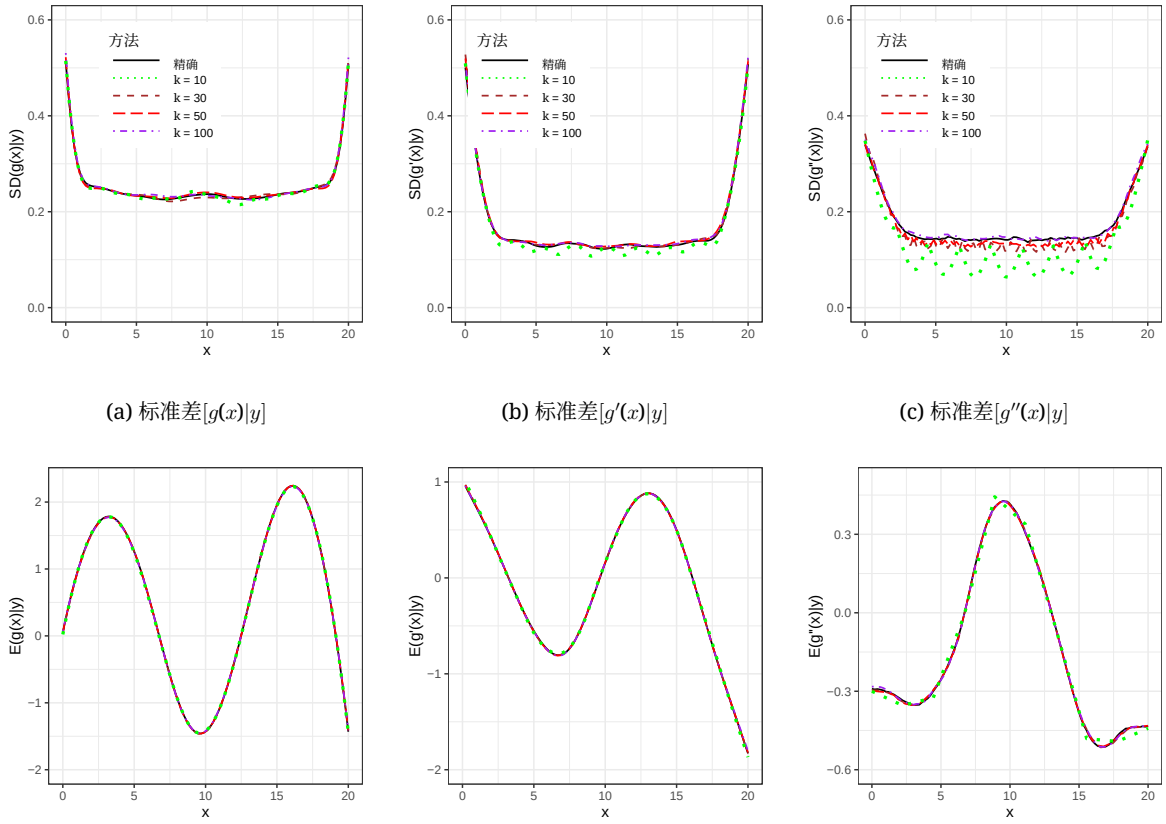
我们还比较了每种方法的平均运行时间。对于 n的每种取值，每个模型将独立拟合 10 次，并将相对运行时间与使用 $k = 10$ 个节点和 50 个观测值的 O样条近似的平均运行时间进行比较。表 1 展示了每个模型 10 次相对运行时间的均值和标准差。所提出的 O 样条近似的运行速度远快于使用精确积分维纳过程方法的实现，即使近似中使用了大量节点（$k \geq n$）也是如此。

最后，为了展示所提出的 O-spline 方法的近似质量，我们针对函数及其前两个导数，在 $n = 100$ 时比较了 O样条近似和精确 IWP 方法所得的推断结果。如图 2 所示，对于函数 g 及其导数，即使只有 $k = 10$，精确 IWP 方法与其 O样条近似所得推断结果的后验均值也非常相似。当 k 较小时，高阶导数的后验标准差存在一定程度的不一致，但随着 k 增加到约 30，这种不一致会大幅减小。

总之，结果表明，与精确 IWP 方法相比，所提出的方法得到的推断结果难以区分，但运行时间显著更短，且当感兴趣位置的数量较大时，数值稳定性更好。

4.3. 推断准确性的评估

本节将说明，与现有的基于二阶——的方法相比，高阶 O-spline 方法能够提供显著改进的推断结果。



(a) 标准差$[g(x)|y]$ (b) 标准差$[g'(x)|y]$ (c) 标准差$[g''(x)|y]$

(d) $\mathbb{E}[g(x)|y]$ (e) $\mathbb{E}[g'(x)|y]$ (f) $\mathbb{E}[g''(x)|y]$

图 2：比较在 $n = 100$ 时，精确积分维纳过程方法与其 O-spline 近似根据第4.2节所得的后验均值和标准差，涵盖函数及其导数。对于后验均值和标准差，O样条近似在 $k \geq 30$ 时都非常准确。

	精确		$\mathrm{OS}_{k=10}$		$\mathrm{OS}_{k=30}$		$\mathrm{OS}_{k=50}$		$\mathrm{OS}_{k=100}$	
n	相对运行时间	条件数（$\log_{10}$）	相对运行时间	条件数（$\log_{10}$）	相对运行时间	条件数 ($\log_{10}$)	相对运行时间	条件数 ($\log_{10}$)	相对运行时间	条件数 ($\log_{10}$)
50	3.29(0.06)	12.52	1.00(0.05)	6.35	1.10(0.06)	6.10	1.31(0.33)	6.03	1.65(0.30)	5.98
100	5.47(0.33)	14.13	1.11(0.03)	6.47	1.23(0.01)	6.20	1.37(0.05)	6.12	1.86(0.04)	6.10
200	10.40(0.62)	16.05	1.83(0.38)	6.98	1.83(0.05)	6.76	2.02(0.03)	6.96	2.82(0.06)	7.23
500	41.44(1.21)	17.34	5.80(0.34)	7.19	6.08(0.21)	7.08	6.69(0.35)	7.30	8.02(0.24)	7.56
800	–(–)	–	14.53(0.17)	7.29	15.07(0.28)	7.24	15.76(0.08)	7.46	18.15(0.24)	7.73
2000	–(–)	–	111.25(1.68)	7.49	112.00(1.16)	7.68	113.47(0.90)	7.86	120.68(1.38)	8.20
5000	–(–)	–	1000.72(25.03)	7.81	996.54(8.13)	7.90	999.43(5.04)	8.09	1010.83(6.77)	8.43

表 1：左列（相对运行时间）显示了按照第4.2节所述，精确积分维纳过程模型和各 O-spline 近似方法运行 10 次所得相对运行时间的均值（标准差）。相对运行时间通过将 O-spline 的平均运行时间除以 $k = 10,\ n = 50$ 计算得出。右列（条件数）显示了对数10 尺度下的 $\kappa_{\max}$。对于$n \geq 800$，精确方法因数值问题而失效，而 O-spline 方法不受影响。无论 k如何，O-spline 方法的运行时间都更短，数值稳定性也更好；并且随着 n 增加，差异会进一步扩大。

阶光滑性，尤其是导数的光滑性。

我们考虑真实函数 $g(x) = \sum_{i=1}^{3} \delta_i \phi(x - \mu_i)$，它是以混合高斯密度的形式模拟得到的，其中 ϕ 表示标准高斯密度。混合权重$\{\delta_1, \delta_2, \delta_3\}$ 设为 $\{0.6, 0.3, 0.1\}$，均值 $\{\mu_i$；每次重复实验中，$i \in [3]\}$ 均独立地从 $N(5,\ 4)$中模拟得到。观测位置 $\boldsymbol{x} = \{x_1, ..., x_n\}$在 $[0, 10]$ 上等间距分布，间距为 $n = 100$。随后对模拟得到的真实函数进行标准化，使每次重复实验中 $g(\boldsymbol{x})$ 的样本方差等于 1。每个观测值模拟为 $y_i = g(x_i) + e_i$，其中 $e_i \overset{iid}{\sim} N(0,\ 1/100)$，从而使信噪比为 100。

作为比较，我们拟合了 Lindgren 和 Rue（2008年）提出的连续二阶随机游走方法，该方法是对 $\mathrm{IWP}_2(\sigma)$ 的近似；还拟合了本文提出的 $\mathrm{IWP}_3(\sigma)$ 的 O样条近似。所有方法均使用 100 个等间距节点进行拟合，并采用第3.4节中的近似贝叶斯推断方法实现。我们采用第3.3节中的方法，对单位功率谱密度 $\sigma(1)$ 设定指数先验，使得 $P(\sigma(1) > 1) = 0.5$；随后将其缩放为相应的指数先验，分别用于 O样条实现中的 σ 和 RW2 实现中的 $p = 3$

O-spline 实现中为 $p = 2$，RW2 实现中为正态分布。

后验均值函数与真实函数之间的相对均方误差按 rMSE(g) $= \sqrt{\sum_{i=1}^{n}\{\mathbb{E}[g(x_i)|\boldsymbol{y}] - g(x_i)\}^2/n}$ 计算，对一阶和二阶导数也采用类似的计算方式。为了比较不同方法的推断准确性，我们再将每种方法的相对均方误差除以 O-spline 方法的相对均方误差中位数。我们对 300 次独立重复实验重复上述过程。

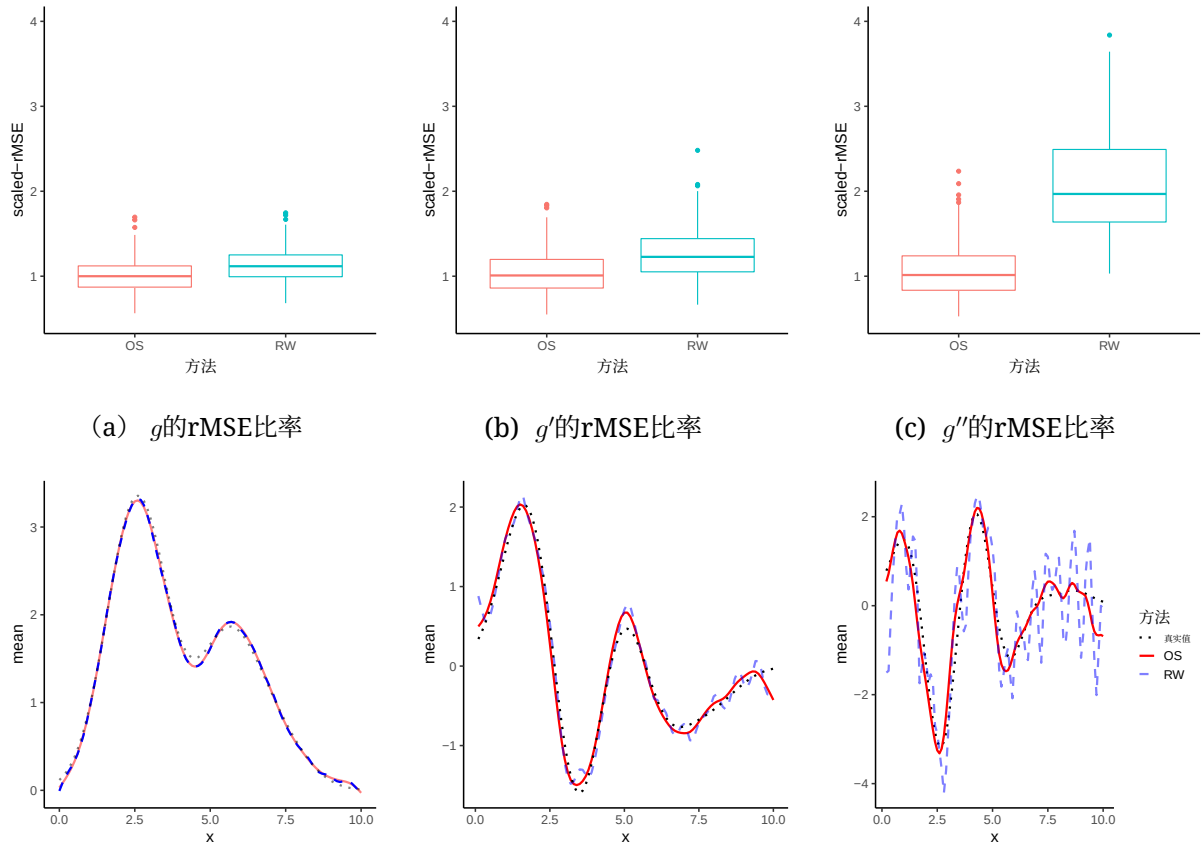
结果汇总于图3。图中显示，对于这类导数中存在少数峰谷的平滑真实函数，具有三阶光滑性的 O样条方法确实优于现有的基于 IWP-2 的方法。对于二阶导数的推断，这种差异最为明显 $g''(\boldsymbol{x})$：二阶方法得出的估计过于波动。其主要原因在于，所提出的具有三阶光滑性的 O样条方法假设目标过程的样本路径二阶连续可微，而二阶方法 RW2 则假设目标过程的样本路径仅一阶连续可微。

5. 示例

本节中，我们使用 O-spline 方法分析加拿大、丹麦、南非和韩国从 2020-01-23 至 2022-04-26 的新冠肺炎每日死亡率。在本例中，对死亡率导数进行基于模型的推断具有实际意义，可用于回答诸如新冠肺炎死亡率何时以最快速度增长，或新冠肺炎死亡率是在减缓还是加速等问题。数据来自约翰斯·霍普金斯大学系统科学与工程中心（CSSE）的 COVID-19 数据仓库（Dong等，2020）。各国的原始数据见在线补充材料。

令 y_t 表示时刻 t的每日新增死亡人数，其中 t 表示从 2020-03-01 至 2022-04-26 的天数。我们考虑一个泊松回归模型：

$$
\begin{aligned}
y_i &\sim \text{Poisson}(\exp(\eta_i)),\\
\eta_i &= \boldsymbol{v}_i^T\boldsymbol{\beta} + g(x_i) + \epsilon_i,\\
g(x) &\sim \text{IWP}_3(\sigma), \quad \epsilon_i \overset{iid}{\sim} N(0, \phi^2).
\end{aligned}
\tag{16}
$$



（a） g的rMSE比率　(b) g'的rMSE比率　(c) g''的rMSE比率

(d) g的后验均值　(e) g'的后验均值　(f) g''的后验均值

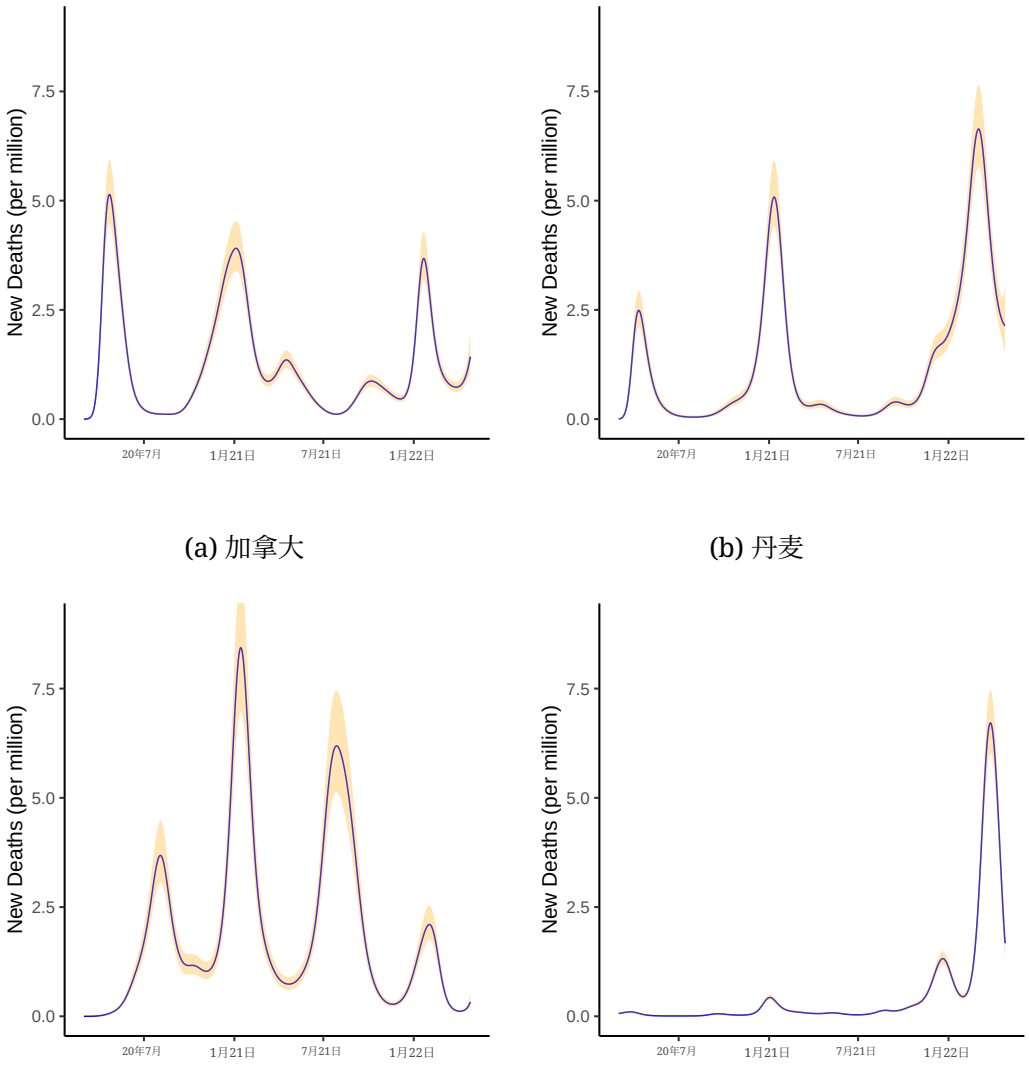
图3：第4.3节的模拟。图(a)–(c)显示了各方法的rMSE箱线图，并以O-spline方法的rMSE中位数进行缩放。图(d)–(f)显示了 g 及其导数的后验均值。O样条方法以红色表示；RW2 方法以蓝色表示。尽管 g 的推断质量相似，O样条方法对导数的推断更准确。

模型包含关于星期几变量的线性固定效应 $\boldsymbol{v}_i = (v_{i1}, ..., v_{i6})^T$，通过未知光滑函数 $g(x)$ x 对时间变量进行平滑，并包含观测层级随机效应 ϵ_i 以处理潜在的过度离散。星期几变量 $\boldsymbol{v}_i$ 的编码方式使得 $g(0)$ 表示平均工作日效应。因此，$\boldsymbol{\beta}$ 表示星期一至星期六相对于平均效应的额外工作日效应，而星期日的额外工作日效应则计算为 $-\sum_{i=1}^{6}\beta_i$。每个线性固定效应 β_i 都独立赋予 $N(0, 100)$先验。过度离散参数 ϕ 采用指数先验建模，中位数为 0.1。

未知函数 $g(x)$ 在公式 (2) 中采用积分维纳过程$_3(\sigma)$先验建模，其中 $\gamma_l \overset{iid}{\sim} N(0, 1/100)$ 用于 $0 \leq l < 3$。随后，我们使用所提出的 O- 样条方法近似积分维纳过程先验，并通过 $k = 100$ 平衡计算效率与近似精度。对于参数 σ，我们赋予 $\sigma(h)$指数先验，其中位数为 log(2)，这里 h 取 7 天。该先验可解释为：在一周内，死亡率有大约 50% 的概率会放大 4 倍或 $1/4$ 。

图4展示了函数 $\exp[g(x)]$ 的推断结果，该函数表示对工作日效应和人口规模进行调整后各国 COVID死亡率的变化。我们还对 COVID死亡率的导数 $g'(x)$ $\exp[g(x)]$ 进行了基于模型的推断，调整后的结果见图5。从图5可以看出，在2022年3月前后几轮疫情期间，韩国的 COVID死亡率增长最快；而在加拿大，增长最快的时期为2020年3月至2020年7月。对于南非，该方法表明 COVID死亡率在2020年底前后增长速度最快。在丹麦，三轮疫情达到峰值时的增长速度相近。

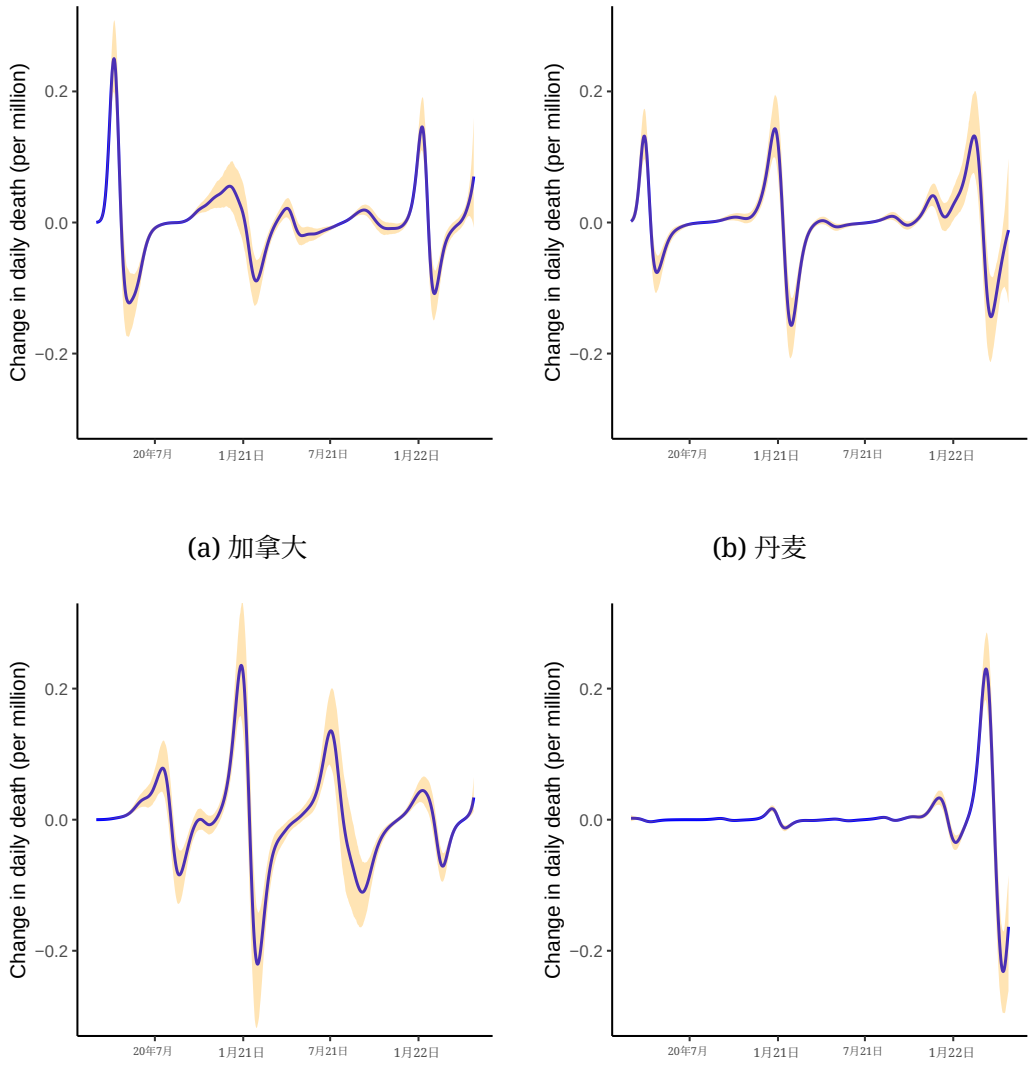
图6展示了各国工作日效应的后验分布，以及过度离散度和功率谱密度（PSD）的后验分布。对数相对风险 $g(x)$ 及其导数 $g'(x)$ 的推断结果也见在线补充材料。



(a) 加拿大　　(b) 丹麦

(c) 南非　　(d) 韩国

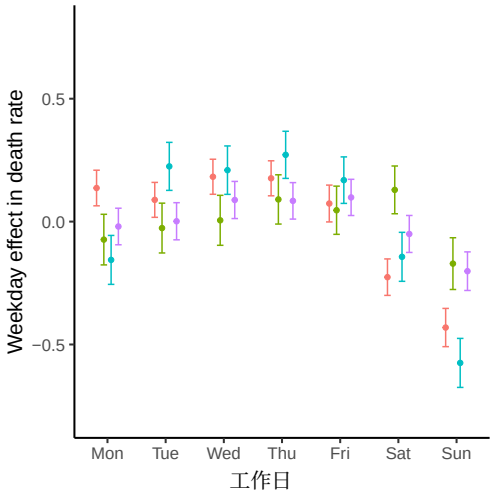
图4：第 5 节中 COVID数据分析的结果。在图 (a)–(d) 中，蓝线表示 O样条（OS）方法的 $\exp[g(x)]$ 后验均值；橙色范围表示 OS 方法的 95% 逐点后验可信区间。死亡人数根据各国 2020 年的人口规模（每百万人）进行了调整。



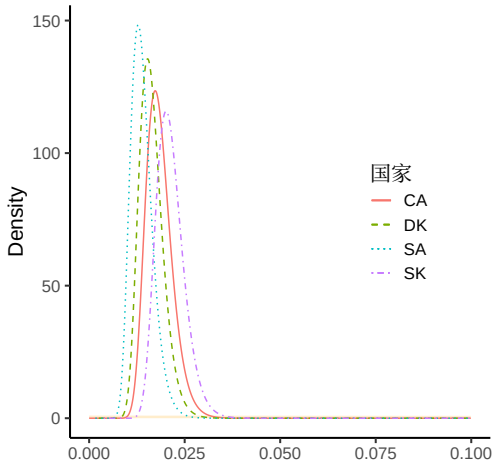
(a) 加拿大　　(b) 丹麦

(c) 南非　　(d) 韩国

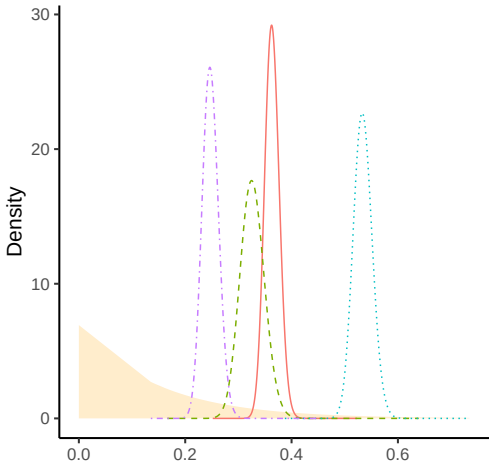
图5：第 5 节中 COVID 数据分析的结果。在图 (a)–(d) 中，蓝色显示的是使用 OS 方法得到的导数 $g'(x)$ $\exp[g(x)]$ 的后验均值；橙色范围表示 OS 方法的 95% 逐点后验可信区间。结果根据各国截至 2020 年的人口规模（每百万人）进行了调整。



（a）工作日效应



（b）7日功率谱密度



（c）过度离散 ϕ

图6：第 5 节 COVID数据分析的结果。图（a）展示了对各个国家工作日效应的推断。实心点表示每个 β_i 的后验均值，即相对于平均效应的工作日效应。线条表示 95 % 可信区间。图（b）展示了各个国家 7日功率谱密度 $\sigma(7)$ 的后验密度，（c）展示了各个国家过度离散参数ϕ 的后验密度。对于两幅图，橙色阴影区域表示先验分布。

6. 讨论

本文研究了基于积分维纳过程（IWP）的平滑方法，并提出了一种适用于任意阶 IWP 的新型有限维 O 样条近似。该近似能够对函数的所有导数进行一致推断；我们还通过理论分析和模拟研究，证明了它收敛于真实的 IWP。为选择和解释参数 σ 的先验，我们引入了一种基于 h-单位功率谱密度$\sigma(h)$ 概念的先验 elicitation 方法。与现有的基于近似的边际标准差的方法（Sorbye 和 Rue，2014）不同，该方法基于精确过程定义，因此不受节点选择的影响。我们通过模拟以及对四个不同国家 COVID 死亡率的数据分析示例，展示了该方法的实用性。

IWP$_p(\sigma)$ 模型的阶数隐含地体现了人们对函数平滑性的先验认识，而平滑性以未知函数能够连续求导的次数来衡量。如果先验知识无法明确确定阶数 p，可以考虑采用贝叶斯因子等方法进行选择，因为在所提出的近似框架下，边际似然是有明确定义的。

尽管所提出的方法在支持的积分维纳过程阶数以及导数推断方面已经具有灵活性，但仍可进一步改进和推广，以支持平滑参数随 σ 变化的自适应平滑；示例见 Yue 等（2014）。

为简化表述，我们默认假设积分维纳过程从感兴趣区域最左侧的点开始。不过，采用两个分别朝不同方向移动的独立积分维纳过程，同样可以将这一程序应用于其他起始值。在某些情形下，这还可以通过引入更稀疏的设计矩阵进一步简化计算。

另一个可能的改进方向是多元平滑，例如 Lindgren 等人（2011年）研究的空间设置。在这种情况下，将所提出的 O 样条基从单变量空间推广到多变量空间并非易事。这些问题留待未来工作解决。

利益披露声明

作者声明不存在需要披露的利益冲突。

补充材料

引理1和定理 1的证明见附录。有限元法程序的详细信息以及第 5 节的补充图表见在线补充材料。

用于复现论文正文中所有结果的代码可在在线代码仓库 github.com/AgueroZZ/Smooth˙IWP˙code 获取。

参考文献

Adler, R. J. (2010)。随机场几何。工业与应用数学学会。

Bilodeau, B.、A. Stringer 和 Y. Tang（2021）。贝叶斯推断中自适应求积的随机收敛速率及应用。*arXiv:2102.06801 [stat.ME]*。

De Brabanter, K. 和 Y. Liu（2015）。基于加权差分序列的平滑非参数导数估计。载于随机模型、统计学及其应用，第 31–38 页。施普林格。

Dong, E.、H. Du 和 L. Gardner（2020）。用于实时追踪新冠肺炎疫情的交互式网络仪表板。《柳叶刀·传染病》*20*(5)，533–534。

Eilers, P. H. 和 B. D. Marx（1996）。利用 B 样条和惩罚项进行灵活平滑。《统计科学》*11*(2)，89–121。

Eilers, P. H. 和 B. D. Marx（2010）。样条、节点与惩罚项。《威利跨学科评论：计算统计学》*2*(6)，637–653。

Harvey, A. C. (1990年)。预测、结构时间序列模型与卡尔曼滤波器。

Hastie, T. 和 R. Tibshirani（1990年）。广义加性模型。Chapman and Hall/CRC出版社。

Li, N. 和 X. Liu（2020）。基于置信分布的非参数曲线导数推断。《统计学通讯：理论与方法》*49*(11)，2607–2622。Lindgren, F. 和 H. Rue（2008）。关于不规则位置的二阶随机游走模型。《斯堪的纳维亚统计学杂志》*35*(4)，691–700。

Lindgren, F.、H. Rue 和 J. Lindstrom（2011）。高斯场与高斯马尔可夫随机场之间的显式联系：随机偏微分方程方法。《英国皇家统计学会杂志》*B*辑（统计方法学）*73*(4)，423–498。

Rasmussen, C. E.（2003）。机器学习中的高斯过程。载于机器学习暑期学校，第63–71页。施普林格。

Robinson, G.（2010）。用于序贯数据分析的连续时间布朗运动模型。《英国皇家统计学会杂志》*C*辑（应用统计）*59*(3)，477–494。

Rue, H.（2001）。高斯马尔可夫随机场的快速抽样。《皇家统计学会杂志：*B*辑（统计方法论）》*63*(2)，325–338。Rue, H. 和 L. Held（2005）。《高斯马尔可夫随机场：理论与应用》。Chapman and Hall/CRC出版社。Ruppert, D. 和 R. J. Carroll（2000）。理论与方法：用于样条拟合的空间自适应惩罚。《澳大利亚与新西兰统计学杂志》*42*(2)，205–223。

Shepp, L. A.（1966年）。高斯测度的拉东－尼科迪姆导数。《数理统计年刊》，321–354。

Simpson, D.、H. Rue、T. G. Martins、A. Riebler 和 S. H. Sorbye（2017）。惩罚模型组件复杂度：构建先验的一种有原则且实用的方法。《统计科学 *32*（1），1 – 28。

Sorbye, S. H. 和 H. Rue（2014）。“空间建模中内在高斯马尔可夫随机场先验的尺度调整”。空间统计 *8*，39–51。

Stringer, A.（2020）。使用自适应求积实现近似贝叶斯推断：aghq 包。*arXiv:2101.04468 [stat.CO]*。

Stringer, A.、P. Brown 和 J. Stafford（2020）。病例交叉模型的近似贝叶斯推断。生物统计学。

Stringer, A.、P. Brown 和 J. Stafford（2022）。扩展潜在高斯模型中后验分布的快速、可扩展近似。计算与图形统计学杂志，1–36。

Swain, P. S.、K. Stevenson、A. Leary、L. F. Montano-Gutierrez、I. B. Clark、J. Vogel 和 T. Pilizota（2016）。使用高斯过程推断时间导数，包括细胞生长率。自然通讯 *7*(1)，1–8。

Wahba, G.（1978）。回归中的不适定先验、平滑样条及防范模型误差问题。《英国皇家统计学会杂志》*B*辑（统计方法学）*40*（3），364–372。

Yue, Y. R.、D. Simpson、F. Lindgren 和 H. Rue（2014）。使用随机微分方程的贝叶斯自适应平滑样条。《贝叶斯分析》*9*（2），397–424。

Zhang, Z., A. Stringer, P. Brown 和 J. Stafford（2022）。具有部分似然、非线性协变量效应和相关观测值的 Cox 比例风险模型的贝叶斯推断。医学研究统计方法，即将出版。

附录：主要结果的证明

引理1 (O样条近似的收敛性)。 设$\Omega = [a, b]$ ， 其中 a、 $b \in \mathbb{R}^+$ ， 且令 $g \sim$积分维纳过程$_p(\sigma)$ ， 其中 $p \in \mathbb{Z}^+$。假设对于每个 $k \in \mathbb{N}$， 节点 $\{s_1, ..., s_k\}$ 在Ω 上等距分布， 且 $\tilde{g}_k$ 表示按式(4)定义的相应 p阶（*O*样条）近似，则：

$$||\mathcal{C} - \mathcal{C}_k||_\infty = O(1/k),$$

其中 $\mathcal{C}(s,\ t) =$协方差$[g(s)\ g(t)]$， $\mathcal{C}_k(s,\ t) =$协方差$[\tilde{g}_k(s)\ \tilde{g}_k(t)]$ 且 $||\mathcal{C}-\mathcal{C}_k||_\infty \equiv$上确界$_{s,\ t\in\Omega}|\mathcal{C}(s,\ t)-\mathcal{C}_k(s,\ t)|$。

证明。为了证明此定理适用于一般的整数阶 p，我们从 $p = 1$ 的情形开始。不失一般性地假设，对于公式 (2) 中的每个 i ， $\gamma_i = 0$ ，且 $\sigma = 1$。此外，不失一般性地假设 $\Omega = [0, 1]$ ，因此对于每个 i， $s_i = i/k$ 。此时 IWP 的真实协方差函数为

$$\mathcal{C}(s, t) = \min\{s, t\}.$$

为了计算近似的协方差，我们有：

$$\begin{aligned} \mathcal{C}_k(s,t) &:= \mathrm{Cov}\left(\tilde{g}_k(s), \tilde{g}_k(t)\right) \\ &= \mathrm{Cov}\left(\sum_{i=1}^{k} w_i\varphi_i(s), \sum_{i=1}^{k} w_i\varphi_i(t)\right) \\ &= k\Phi(s)^T\Phi(t), \end{aligned} \tag{17}$$

其中 $\Phi(t) = (\varphi_1(t)$ ……， $\varphi_k(t))^T$ 且 $\Phi(s) = (\varphi_1(s)$ ……， $\varphi_k(s))^T$，这是因为按照我们的定义， $w_i \overset{iid}{\sim} N(0,\ k)$ 的分布如此。

现在不失一般性地假设 $s \le t$ 。由于 O样条基函数 φ_i 具有重叠性质，显然，只有位于区域 $[0, s]$ 内的节点才会

基函数 对上述内积产生非平凡贡献。因此 e,

$$\begin{aligned} k\Phi(s)^T\Phi(t) &= k\sum_{i=1}^{\lfloor ks\rfloor}\varphi_i(s)\varphi_i(t) \\ &= k\sum_{i=1}^{\lfloor ks\rfloor}\frac{1}{k^2}\ \text{since } p=1 \\ &= \lfloor ks\rfloor/k, \end{aligned} \tag{18}$$

因为 $\phi_i(s) = \phi_i(t) = \frac{1}{k}$ 对 $i \le \lfloor ks\rfloor$成立。

对于任意固定的 $t \in \Omega$，我们有

$$\begin{aligned} \sup_{s\in\Omega}|\mathcal{C}(s,t)-\mathcal{C}_k(s,t)| &\le \sup_{s\in[0,t]}|\mathcal{C}(s,t)-\mathcal{C}_k(s,t)| + \sup_{s\in[t,1]}|\mathcal{C}(s,t)-\mathcal{C}_k(s,t)| \\ &= \sup_{s\in[0,t]}|s-\lfloor ks\rfloor/k| + \sup_{s\in[t,1]}|t-\lfloor kt\rfloor/k| \\ &\le 2/k. \end{aligned} \tag{19}$$

由于 t 是任意的，这意味着 $\sup_{s,\ t\in\Omega}|\mathcal{C}(s,\ t) - \mathcal{C}_k(s,\ t)| \le 2/k$ ，因此我们证明了 $p = 1$的情形。

为了将结果推广到更高阶的 p，我们需要以下关于高斯过程的命题：

命题1 (高斯过程的积分**).** 假设 $f(t)$ 是一个样本路径连续的高斯过程，对 $t \in \Omega := [0, 1]$ 满足边界条件 $f(0) = 0$，那么$g(s) := \int_0^s f(t)dt$ 仍然是一个高斯过程，其在 $s^*, t^* \in \Omega$ 处的协方差函数可计算为

$$\mathcal{C}_g(s^*, t^*) = \int_0^{s^*}\int_0^{t^*}\mathcal{C}_f(s,t)dtds$$

其中 $\mathcal{C}_f(.)$ 是 f 的协方差函数。

该命题由 Adler（2010）第 2.4.3 节得出。

现在，考虑 $p = 2$ 的情况。由于重叠样条基具有导数一致性，因此显然， $\tilde{g}_k^{(1)}(x)$ 是一阶积分维纳过程 $g^{(1)}(x)$ 的一阶 O-spline 近似，其收敛性已在上文的证明中确立

中。此外，对于 $p = 1$，所提出的近似具有连续的样本路径，其协方差函数为 $\mathcal{C}_k(s,\ t) = \lfloor ks \rfloor / k \leq 1$，如上所示。

令 $\mathcal{C}^{[p]}(s,t)$ 和 $\mathcal{C}_k^{[p]}(s,t)$ 分别表示 p 阶积分维纳过程及其近似的协方差，则对于任意选择的 $s,t \in \Omega$，有：

$$
\begin{aligned}
&\sup_{s,t\in\Omega} |\mathcal{C}^{[2]}(s,t) - \mathcal{C}_k^{[2]}(s,t)| \\
&\quad = \sup_{s,t\in\Omega} \left| \int_0^s \int_0^t \mathcal{C}^{[1]}(s^*,t^*) - \mathcal{C}_k^{[1]}(s^*,t^*) dt^* ds^* \right| \\
&\quad \leq \sup_{s,t\in\Omega} \int_0^s \int_0^t \left| \mathcal{C}^{[1]}(s^*,t^*) - \mathcal{C}_k^{[1]}(s^*,t^*) \right| dt^* ds^* \\
&\quad \leq \sup_{s,t\in\Omega} \int_0^1 \int_0^1 \left| \mathcal{C}^{[1]}(s^*,t^*) - \mathcal{C}_k^{[1]}(s^*,t^*) \right| dt^* ds^* \\
&\quad \leq \int_0^1 \int_0^1 \frac{2}{k} dt^* ds^* \text{ by previous result on } p = 1 \\
&\quad = \frac{2}{k}.
\end{aligned}
\tag{20}
$$

由此，$p = 2$ 的协方差函数收敛性得以确立。此外，由于区域 $[0,s] \times [0,t] \subseteq \Omega \times \Omega$ 是紧致的，更高阶 p 的协方差函数仍将由 $2/k$ 界定。因此，通过归纳法，收敛结果可以推广到任意正整数 p。

□

定理 1（主定理）。 在与引理*1*相同的设定和记号下，对于任意非负整数 $q_1 \leq p-1$ 和 $q_2 \leq p-1$:

$$||\mathcal{C}^{(q_1,q_2)} - \mathcal{C}_k^{(q_1,q_2)}||_\infty = O(1/k),$$

其中 $\mathcal{C}^{(q_1,q_2)}(s,t) =$ 协方差$[g^{(q_1)}(s)\ g^{(q_2)}(t)]$ 且 $\mathcal{C}^{(q_1,q_2)}{}_k(s,t) =$ 协方差$[\tilde{g}_k^{(q_1)}(s)\ \tilde{g}_k^{(q_2)}(t)]$。

证明。 为了证明定理 1 中的一般收敛结果，我们先证明以下引理 2 给出的一个特殊情况：

引理 2 (交叉协方差的收敛)。 设 $\Omega = [a,b]$，其中 a、$b \in \mathbb{R}^+$，并且令 $p \geq 1$ 和 $q \leq p-1$ 为任意正整数。令 $g \sim$ 积分维纳过程$^p(\sigma)$。假设节点 $\{s_1, ..., s_k\}$

对于每个 $k \in \mathbb{N}$ 在 Ω 上等间距分布，且 $\tilde{g}_k(x)$ 表示式 (4)中定义的相应 p 阶（*O-spline*）近似，则：

$$||\mathcal{C}^{(0,q)} - \mathcal{C}_k^{(0,q)}||_\infty = O(1/k),$$

其中 $\mathcal{C}^{(0,q)}(s,\ t) =$ 协方差$[g(s)\ g^{(q)}(t)]$ 以及 $\mathcal{C}_k^{(0,q)}(s,\ t) =$ 协方差$[\tilde{g}_k(s)\ \tilde{g}_k^{(q)}(t)]$。

Fo为方便记号，定义微分算子和积分算子 a s

$$D_t^p(g(s,t)) := \frac{\partial^p g}{\partial t^p}(s,t),$$

and

$$I_t^p(g(s,t)) := \int_0^t \int_0^{t_1} ... \int_0^{t_{p-1}} g(s,t_p) dt_p dt_{p-1} ... dt_1.$$

两个算子都是线性的。同样为简便起见，我们考虑 $\Omega = [0,1]$，不失一般性。当 $p = 1$ 时，定理显然成立，因此我们考虑 $p > 1$ 的情形。

令 $s,\ t \in \Omega$ 固定，并令 g 服从 p 阶积分维纳过程，其 $\tilde{g}_k$ 表示其近似。令 $\mathcal{C}^{[p]}(s,\ t)$ 表示其自协方差，$\mathcal{C}_k^{[p]}(s,\ t)$ 表示该近似的自协方差。

根据引理1的证明，$\{\mathcal{C}_k^{[p]}(s,\ t)\}_k$ 一致有界。结合命题1的结果应用富比尼定理，可得：

$$\begin{aligned}
\mathcal{C}_k^{(0,q)}(s,t) &= \mathbb{E}[\tilde{g}_k(s)\frac{\partial^q}{\partial t^q}\tilde{g}_k(t)] \\
&= \frac{\partial^q}{\partial t^q}\mathbb{E}[\tilde{g}_k(s)\tilde{g}_k(t)] \\
&= D_t^q \mathcal{C}_k^{[p]}(s,t) \\
&= D_t^q I_t^{p-1} I_s^{p-1}[\mathcal{C}_k^{[1]}(s,t)] \\
&= I_t^{p-q-1} I_s^{p-1}[\mathcal{C}_k^{[1]}(s,t)].
\end{aligned} \tag{21}$$

类似地，$\mathcal{C}^{(0,q)}(s,t)$ 可写为：

$$\begin{aligned}
\mathcal{C}^{(0,q)}(s,t) &= D_t^q \mathcal{C}^{[q]}(s,t) \\
&= I_t^{p-q-1} I_s^{p-1}[\mathcal{C}^{[1]}(s,t)].
\end{aligned} \tag{22}$$

The利用引理1的结果，可以得到以下上界：

$$\begin{aligned}\sup_{s,t\in\Omega}|\mathcal{C}^{(0,q)}(s,t)-\mathcal{C}_k^{(0,q)}(s,t)| &= \sup_{s,t\in\Omega}|I_t^{p-q-1}I_s^{p-1}\mathcal{C}^{[1]}(s,t)-\mathcal{C}_k^{[1]}(s,t)| \\ &\le \sup_{s,t\in\Omega} I_t^{p-q-1}I_s^{p-1}|\mathcal{C}^{[1]}(s,t)-\mathcal{C}_k^{[1]}(s,t)| \\ &\le \sup_{s,t\in\Omega} I_t^{p-q-1}I_s^{p-1}\frac{2}{k} \\ &\le \frac{2}{k},\end{aligned} \tag{23}$$

其中，最后一个不等式由 Ω 的紧致性得出。引理 2 的证明至此结束。

结合引理 2 和式 (9) 中 O样条近似的性质，即可证明定理 1。

□